



Doctorado en Ingeniería Industrial, Informática y Medioambiental

Ingeniería de Sistemas y Automática

Tesis doctoral

Computación inteligente aplicada al mantenimiento del sector de la generación eléctrica insular

Daniel González Calvo

Directora: Dra. D^a Rosa María Aguilar China

Codirector: Dr. D. Jesús Miguel Torres Jorge

2021

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

1 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

2 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

D^a Rosa María Aguilar Chinaea, Doctora en Informática y Catedrática del Departamento de Ingeniería de Sistemas y Automática de la Universidad de La Laguna.

D. Jesús Miguel Torres Jorge, Doctor por la Universidad de la Laguna y Profesor del Departamento de Ingeniería Informática y de Sistemas de la Universidad de La Laguna.

CERTIFICAN:

que D. Daniel González Calvo, con la titulación de Master en Ingeniería Industrial, ha realizado bajo nuestra dirección la presente Tesis, titulada *Computación inteligente aplicada al mantenimiento del sector de la generación eléctrica insular*, para optar al grado de Doctor por la Universidad de La Laguna.

Con esta fecha, autorizamos la presentación de la misma.

En San Cristóbal de La Laguna, a 15/01/2021.

La Directora

Rosa María Aguilar Chinaea

El Codirector

Jesús Miguel Torres Jorge

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaea
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

3 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

4 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilan
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

A Marta.

*A mis padres y a mis hermanas,
Silvia, Lucía y Lola.*

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

5 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

6 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Agradecimientos

En primer lugar, quiero agradecer a la Dra. D.^a Rosa María Aguilar Chinaea, directora de este trabajo de investigación doctoral, sus consejos, su apoyo y por supuesto motivación continua, que han sido claves y necesarios para mantenerme firme en todo el desarrollo de esta investigación.

Agradecer también a mi codirector, el Dr. D. Jesús Miguel Torres Jorge, su ayuda y siempre buena disposición, además de los consejos en la escritura y programación LaTeX.

En segundo lugar agradecer a mi padre, el Dr. D. Luis Antonio González Mendoza, su asesoramiento continuo en el desarrollo de mis ideas y trabajos.

Agradecer también al Dr. D. Constantino Criado Hernández, por la dedicación prestada en este trabajo y su permanente respaldo en mis logros.

Quisiera agradecer a mis compañeros de Endesa, en particular a D. Saúl Barrio Ortega y al Dr. D. Tomás Álvarez Tejedor, por darme la oportunidad y el respaldo necesario para poder llevar a cabo este Doctorado Industrial.

Agradezco al equipo del Instituto de Investigación Tecnológica (IIT) de la Escuela Técnica de Ingeniería - ICAI de la Universidad Pontificia Comillas, el Dr. D. Antonio Muñoz San Roque, el Dr. D. José Portela González, el Dr. D. Miguel Ángel Sanz Bobi y el Dr. D. Eugenio Sánchez Ubeda, con los que he podido colaborar dentro del alcance de este trabajo, de manera extraordinaria.

A mis amigos y compañeros de trabajo, les agradezco su colaboración y apoyo para emprender y finalizar esta tesis doctoral.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaea
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

7 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Agradecimientos

Agradecerle a Marta, mi compañera de vida, su apoyo constante en todos los momentos de mi trayectoria.

A mi familia por supuesto, agradecerles su apoyo y cariño incondicional, incluyendo los periodos de aislamiento "justificados", necesarios para poder realizar una tesis doctoral. Mil gracias...

VIII

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

8 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Índice

Resumen	XXVII
Abstract	XXXI
Introducción	XXXIII
1. Motivación	XXXIII
2. Objetivos	XXXIV
3. Estructura de la tesis	XXXVIII
Capítulo 1. Estado del arte	1
1.1. Mantenimiento basado en datos	1
1.2. Métodos de computación inteligente	11
1.3. Entorno ambiental de la Macaronesia	23
1.4. Sistemas de generación eléctrica convencional	25
Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales	33
2.1. Estructura de los datos y arquitectura	33
2.2. Procedimientos	40
2.3. Resultados	47
2.4. Discusión	50

IX

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA

Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

9 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpFLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA

María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

Índice

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial	53
3.1. Estructura de los datos y arquitectura	53
3.2. Procedimientos	59
3.3. Resultados	67
3.4. Discusión	75
Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento	79
4.1. Estructura de los datos y arquitectura	80
4.2. Procedimientos	81
4.3. Resultados	90
4.4. Discusión	95
Conclusiones y trabajo futuro	99
Análisis multivariante mediante ensemble de redes neuronales artificiales	100
Influencia del polvo sahariano y el mantenimiento sobre un compresor axial	101
Metodología general y modelo para el aprendizaje de eventos de mantenimiento	102
Líneas futuras de estudio	103
Apéndice A - Código representativo del Capítulo 2	105
Apéndice B - Código representativo del Capítulo 3	111
Apéndice C - Código representativo del Capítulo 4	119
Bibliografía	125

x

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

10 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

10 / 181

Índice de tablas

2.1. Resultados y dispersión obtenidos para cada variable de entrada en el cálculo SIM y MGA, a partir del conjunto final promedio CVC-FFBPNN (DS1).	50
3.1. Resumen del estudio comparativo realizado con otros métodos de regresión para estimar el rendimiento isentrópico del compresor.	62
3.2. Resumen del estudio comparativo para observar el cálculo SIM utilizando diferentes configuraciones del ensemble red neuronal pre-alimentada y de propagación hacia atrás (FFBP-NN).	62
3.3. Registro histórico de eventos de mantenimiento (●) realizados en la turbina de gas durante todo el periodo de datos disponibles.	64
3.4. Resultados obtenidos para cada modelo de regresión entrenado.	69
3.5. Simulación simulaciones de mantenimiento (MBD.SIMU) 1. Horas de funcionamiento (MTWork1) sin realizar limpieza <i>offline</i> del compresor para condiciones de acumulación de arena y ensuciamiento de filtros.	74
4.1. Registro histórico de mantenimientos de revisión (●) realizados en el motor diesel de generación eléctrica.	86

xI

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

11 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpFLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

12 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Índice de figuras

1. Esquema de la línea de desarrollo para la metodología de MBD	xxxviii
1.1. Características principales del <i>Big Data</i> (Koseleva y Ropaite (2017)).	2
1.2. El <i>gemelo digital</i> - Evolución de la tecnología de simulación (Rosen <i>et al.</i> (2015)).	3
1.3. Evolución histórica del mantenimiento industrial	4
1.4. La arquitectura del sistema DT (Zheng <i>et al.</i> (2018)).	5
1.5. Modelo híbrido de un sistema para su análisis predictivo (Nikolaev <i>et al.</i> (2019)).	7
1.6. Modelo basado en la minería de datos (<i>data mining</i>) para identificar actividades de mantenimiento ineficientes (Mosaddar y Shojai (2013)).	10
1.7. Modelo no lineal de una neurona artificial. Adaptado de Haykin (1999).	12
1.8. Arquitectura clásica de un perceptrón multicapa (Haykin (1999)).	15
1.9. Esquema del algoritmo de <i>RF</i> (Ahmad <i>et al.</i> (2017)).	17
1.10. Ejemplo para la identificación de importancias SHAP, cuantificando el impacto de cada variable en la salida del modelo (Mckinley <i>et al.</i> (2020)).	22
1.11. Efecto de la concentración de arena sahariana en las Islas Canarias.	23
1.12. Frecuencia mensual de los días con predominio de aire sahariano en Canarias (1976-2003) (Dorta <i>et al.</i> (2005)).	24

xiii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

13 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

13 / 181

Índice de figuras

1.13. Variabilidad mensual de las concentraciones de aerosol mineral ($\mu\text{gr}/\text{m}^3$) en el aire medidas en Gran Canaria (Gelado <i>et al.</i> (2002)).	25
1.14. Diagrama general de una turbina de gas para generación de electricidad con un sistema filtrante multi etapa.	26
1.15. Diagrama estándar de la formación del fenómeno <i>fouling</i> en álabes de compresor (Aldi <i>et al.</i> (2015)).	28
1.16. Tendencia cualitativa de las variaciones de eficiencia en un compresor, con diferentes tipos de lavados (Dionysios <i>et al.</i> (2016)). .	28
1.17. Curvas de degradación por <i>fouling</i> en compresor de Turbina de Gas (Tarabrin <i>et al.</i> (2014)).	29
1.18. Adaptación del esquema general de un motor diesel para generación eléctrica.	31
2.1. Distribución del conjunto de datos para la fase de entrenamiento y test	34
2.2. Modelo de detección de arena BSC-DREAM8b para Canarias, operado por el Centro de Supercomputación de Barcelona. http://www.bsc.es/ESS/bsc-dust-daily-forecast	34
2.3. Datos originales vs 5-días SMA; Enero, Febrero y Marzo (Q1); Abril, Mayo y Junio (Q2); Julio, Agosto y Septiembre (Q3); Octubre, Noviembre y Diciembre (Q4)	35
2.4. Rosa de los vientos con DS1 en el sur de la isla de Tenerife	36
2.5. Topología de la estructura principal de RNA base, con cinco entradas, nueve nodos en una capa oculta, con función de activación tangente hiperbólica y un nodo en la capa de salida con una función de activación lineal.	37
2.6. Validación cruzada(10-fold) para DS1.	39
2.7. Esquema de aproximación MBD.0 en la línea de desarrollo para el mantenimiento basado en datos.	40
2.8. Flujograma del procedimiento computacional realizado en de este capítulo.	41

xiv

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

14 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

Índice de figuras

2.9. Diagrama principal del algoritmo de entrenamiento utilizado, ampliado en los bloques de aprendizaje ensemble mediante partición k -fold y validación final.	42
2.10. Diagrama principal del procedimiento SIM.	44
2.11. Diagrama ampliado del procedimiento SIM para eliminar una variable de entrada y sus pesos sinápticos hacia la capa oculta. . . .	45
2.12. Resultados de la predicción de la concentración de arena en la isla de Tenerife y comparación estadística del modelo final con DS1. Julio 2016 a Julio 2018.	47
2.13. Resultados de la predicción de la concentración de polvo sahariano en la isla de Tenerife y comparación estadística del modelo final con DS2. Julio 2018 a Diciembre 2018.	48
2.14. Histograma con la distribución del error del modelo final con DS2.	48
2.15. Grado de importancia relativa obtenida para los métodos SIM y MGA; Temperatura (Temp.); Presión Atmosférica (Atm.Pr); Velocidad del viento (Wind S.); Humedad relativa (RH.); Número del mes (Month N.).	49
2.16. Coeficientes de Variación (CV) obtenidos para cada variable de entrada del procedimiento SIM (A) y diagrama de cajas representando la dispersión del MAE para cada entrada, k veces con el método SIM (B).	49
3.1. Conjunto de datos DS1 y DS2, utilizados en la fase de entrenamiento y test de la RNA.	54
3.2. Topología de la estructura principal de la RNA. Diseñada con ocho nodos de entrada, ocho nodos en la capa oculta con función de activación tangente hiperbólica y un nodo en la capa de salida con función de activación lineal.	55
3.3. Muestra tomada de la zona de filtros de entrada de la turbina de gas (Wilcox <i>et al.</i> (2012)).	56

xv

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

15 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Índice de figuras

3.4. Difractograma de polvo, analizado por el laboratorio externo (CSIC-ICTJA) y la composición resultante de la muestra tomada en la zona de filtros de entrada de la turbina de gas.	56
3.5. Esquema de la línea de desarrollo para la metodología de MBD para el capítulo 3.	60
3.6. Flujograma del procedimiento principal utilizado en este trabajo. Sección 3.2.	61
3.7. Esquema del procedimiento MBD.1 para intervenciones de mantenimiento recurrentes.	63
3.8. Descripción del ciclo de funcionamiento del algoritmo MTWork1 para definir la variable calculada de mantenimiento del modelo.	63
3.9. Esquema del procedimiento de simulación utilizado en este trabajo para la línea de MBD.	66
3.10. Serie temporal (2013-2018) de la eficiencia isentrópica real medida del compresor, mostrando los períodos de limpieza <i>offline</i> del compresor de la turbina de gas.	67
3.11. Serie temporal (2013-2018) de la medida real de la eficiencia del compresor, suavizada con una SMA de tres días, con detalles de dos eventos (2015-2016) que implican una alta acumulación de polvo sahariano y una disminución de la eficiencia.	68
3.12. Diagrama de cajas que representa la cantidad de residuos para cada modelo de regresión entrenado con DS2 (Tabla 3.4).	70
3.13. Resultados y rendimiento del modelo final FFBP-NN CVC ($k=10$) basado en datos reales de DS2. Detalle ampliado del modelo de aprendizaje para un lavado de compresor <i>offline</i>	70

xvi

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

16 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

Índice de figuras

3.14. Niveles de importancia relativa aplicando el método SIM y su control de dispersión mediante el coeficiente de variación CV. Contador de polvo sahariano (C.S) ; Potencia eléctrica generada (Potencia El.); Temperatura ambiente (Temp.); Humedad relativa (HR.); Presión atmosférica (Pr.Atm); Presión diferencial - Filtros de entrada (PD Filtros); Número de mes (N.Mes); Horas MTWork1 (Horas Fcto.)	71
3.15. Simulación para la planificación del mantenimiento (MBD.SIMU 1); (1) Escenario de horas de funcionamiento; (2) Escenario de filtros de entrada PD.	73
3.16. Simulación de la planificación del mantenimiento (MBD.SIMU 2); Escenario para la limpieza del compresor <i>offline</i> y el mantenimiento de los filtros, por estación.	75
4.1. Conjunto de datos DS1 y DS2, utilizados en la fase de entrenamiento y test de la RNA.	80
4.2. Esquema de la línea de desarrollo para la metodología de MBD para el capítulo 4.	82
4.3. Esquema principal de la metodología MBD.2	83
4.4. Procedimiento computacional general de MBD, dividido en tres bloques principales (2A, 2B y 2C).	83
4.5. Diagrama MBD.2A, para el estudio de intervenciones de mantenimiento por cada nuevo evento temporal ME_n	84
4.6. Detalle ampliado para la determinación de la variable "ME" sobre MBD.2A.	84
4.7. Flujograma del método propio para el cálculo de IR en eventos de mantenimiento, MTImp.	87
4.8. Resultados de la serie temporal DS2 con valores reales de emisiones de NOx frente a la predicción realizada por el modelo <i>LGBM</i>	91
4.9. IR sobre un modelo de eventos de mantenimiento MBD.2 a partir del algoritmo MTImp.	92

xvii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

Índice de figuras

4.10. Niveles de IR para eventos de mantenimiento en emisiones de NOx mediante métodos SHAP y FeatImp.	93
4.11. Simulación MBD.SIMU 3 sobre el NOx sobre DS1, con los valores reales, predichos y si no se hubiera realizado un mantenimiento de revisión. Se analizan las intervenciones de mantenimiento MTO15000 2017, MTO10000 2016 y MTO10000 2018.	94

xviii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA

Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

18 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA

María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

Índice de algoritmos

2.1. Selective Importance Measure (SIM)	46
3.1. MTWork1	65
4.1. MTWork2	85
4.2. MTImp	89

xix

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA

Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

19 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA

María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

19 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

20 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Índice de acrónimos

MBD	mantenimiento basado en datos	xxxvi
MBD.SIMU	simulaciones de mantenimiento	xi
MBD.0	fase de aproximación al mantenimiento basado en datos	xxxviii
MBD.1	fase inicial del mantenimiento basado en datos	xxxix
MBD.2	última fase del mantenimiento basado en datos	xxxix
IR	importancias relativas	xxxv
CVC	comité de validación cruzada	xxxv
RNA	red neuronal artificial	xxxiv
FFBP-NN	red neuronal pre-alimentada y de propagación hacia atrás	xi
SMA	media móvil simple	35
MAE	error medio absoluto	38
RMSE	error medio cuadrático	38
VC	validación cruzada	6
SIM	algoritmo <i>Selective importance measure</i>	xxxv
DS	conjunto de datos	20
DS1	conjunto de datos para la fase de entrenamiento	33
DS2	conjunto de datos para la fase de test	34
R^2	coeficiente de determinación	38

xxi

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

21 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

21 / 181

Índice de algoritmos

ML	<i>machine learning</i>	xxxvi
CV	coeficiente de variación	45
MGA	algoritmo de Garson modificado	21
NOx	óxidos de nitrógeno	xxxvii
CSIC	Consejo Superior de Investigaciones Científicas	57
ICDD	<i>International Centre for Diffraction Data</i>	57
MTWork1	algoritmo para convertir en una variable computacional, un evento de mantenimiento de un equipo particular	xxxvi
MTWork2	algoritmo para convertir en una variable computacional, un evento general de mantenimiento	xxxvii
MTImp	algoritmo para el cálculo de importancias en eventos históricos de mantenimiento	xxxvii
GB	<i>Gradient boosting</i>	xxxiv
RF	<i>Random forest</i>	8
PD	<i>presión diferencial</i>	73
MWe	<i>megavatios eléctricos</i>	80
ME	eventos de mantenimiento	84
LGBM	Light Gradient Boosting Machine	81
SHAP	SHapley Additive exPlanations	21
FeatImp	Feature Importance	21
NLP	<i>Natural Language Processing</i>	104
MTW	variable contador de salida de los algoritmos MTWork	90
DT	<i>digital twin</i>	2
GOSS	<i>Gradient-based One-Side Sampling</i>	20
EFB	<i>Exclusive Feature Bundling</i>	20

xxii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

Símbolos y notación

$bara$	unidad de presión absoluta
$mmH2O$	unidad de presión - milímetro de columna de agua
\notin	no pertenece a
\in	pertenece a
[...]	vector o matriz
\approx	valor aproximado
M	modelo final validado de aprendizaje de máquina
μ	media del conjunto
S	observaciones escalares
A	vector variable
nn	conjunto de ensemble de redes neuronales artificiales
x	variable de entrada de un modelo de aprendizaje de máquina
y	variable de salida de un modelo de aprendizaje de máquina
w	peso sináptico de una red neuronal artificial
b	bías de una red neuronal artificial
nk	elemento n en un conjunto k de un ensemble computacional
k	índice de validación cruzada
w_{i_0xh}	peso sináptico de entrada inicial en un conjunto de ensemble, siendo h la neurona de la capa oculta e i la capa de entrada

XXIII

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

23 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

23 / 181

Símbolos y notación

b_{0h}	bías inicial de la capa de salida en un conjunto de ensemble, siendo h la neurona de la capa oculta
w_{0xh}	peso sináptico de salida inicial en un conjunto de ensemble, siendo h la neurona de la capa oculta y o la capa de salida
b_{i0xh}	bías inicial de la capa de entrada en un conjunto de ensemble, siendo h la neurona de la capa oculta
m	número de neuronas de entrada en una red neuronal artificial
N	número de neuronas de salida en una red neuronal artificial
$L1$	número de neuronas en una capa oculta de una red neuronal artificial
w_{ih}	pesos sinápticos entre la entrada i y la neurona h de la capa oculta
w_{ho}	pesos sinápticos entre la capa oculta h y la capa de salida o
σ	desviación estándar
ϕ	función de activación de una red neuronal artificial
\bar{y}	valor medio de y
\hat{y}	estimación de y
μ_{co}	rendimiento isentrópico del compresor (%)
γ	coeficiente adiabático
p_1	presión de entrada al compresor (<i>bara</i>)
p_2	presión de descarga del compresor (<i>bara</i>)
T_1	temperatura de entrada al compresor (°C)
T_2	temperatura de descarga del compresor (°C)
cp_h	calor específico del aire húmedo a presión constante ($kJ/kg^\circ K$)
cp_a	calor específico del aire seco a presión constante ($kJ/kg^\circ K$)
cv_h	calor específico del aire húmedo a volumen constante ($kJ/kg^\circ K$)
cv_a	calor específico del aire seco a volumen constante ($kJ/kg^\circ K$)
q_{air}	caudal másico de aire total de entrada al compresor (kg/s)
q_{airs}	caudal másico de aire seco de entrada al compresor (kg/s)

xxiv

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

24 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

24 / 181

wf caudal de humedad absoluta (kg/s)
 α fracción correspondiente al contenido de polvo de origen sahariano (%)
 ρ_{air} densidad del aire (kg/m^3)
 Q_{dust} totalizador de polvo de origen sahariano (gr)
 C_{dust} concentración de polvo medida en la superficie ($\mu gr/m^3$)
 x_t serie temporal con datos de fecha y hora
 x_{or} serie temporal con datos de régimen de operación
 x_a serie temporal con datos ambientales
 X_{train} conjunto general de datos de entrada para la fase de entrenamiento
 X_{test} conjunto general de datos de entrada para la fase de test
 Y_{train} conjunto general de datos de salida para la fase de entrenamiento
 Y_{test} conjunto general de datos de salida para la fase de test

xxv

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

25 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

25 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

26 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Resumen

La metodología de mantenimiento basado en datos se ha aplicado de forma progresiva en este trabajo, de forma que su estudio y desarrollo sirve para definir un procedimiento de aplicación computacional dentro de una línea de mantenimiento en el sector industrial.

Se parte de un enfoque dirigido hacia el análisis y la predicción de la concentración de polvo en el aire, vital para la actividad económica y la salud de la población. En este apartado del estudio, utilizamos un conjunto de redes neuronales artificiales que estructuramos mediante el aprendizaje de ensemble para obtener una variable compleja, como la concentración de polvo, basada en datos reales como la temperatura del aire, la humedad relativa, la presión atmosférica y la velocidad del viento. Los índices de rendimiento estadístico obtenidos muestran la eficacia del comité de validación cruzada propuesto. Por lo tanto, es fundamental disponer de un método de cálculo fiable para determinar las importancias relativas que pueden aplicarse a este tipo de arquitectura de ensemble mediante redes neuronales artificiales.

A diferencia de otros métodos de importancia relativa, en los que los cálculos se realizan basándose directamente en los pesos de la red neuronal artificial y cuyos resultados en configuración de ensemble muestran una alta dispersión, se propone un procedimiento original, que elige selectivamente la variación de las entradas para reajustar la arquitectura de la red neuronal. Esto nos permite medir aquellas variables con mayor efecto sobre la variable objetivo, obteniendo así la influencia multivariante sobre la concentración de polvo en la superficie a

XXVII

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

27 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

27 / 181

Resumen

través de un modelo computacional. Este método proporciona así una alternativa real para calcular y estimar la importancia relativa que puede generalizarse a cualquier tipo de problema para los sistemas multivariantes modelados mediante redes neuronales artificiales tanto para una configuración simple como para una arquitectura de ensemble.

Por otro lado, la planificación del mantenimiento industrial asociado a la producción de electricidad es vital, ya que permite obtener una foto actual y futura de un componente industrial para optimizar los recursos humanos, técnicos y económicos de la instalación. En un segundo bloque nos centramos en la degradación debida al ensuciamiento de una turbina de gas en las Islas Canarias, donde se analizan los niveles de ensuciamiento a lo largo del tiempo, en función del régimen de operación y de las variables meteorológicas locales. En particular, se estudia la relación entre la degradación de la máquina, las intervenciones históricas de mantenimiento realizadas y el polvo en suspensión de origen sahariano. Para ello, utilizamos de nuevo el procedimiento computacional inicial basado en el ensemble promedio de redes neuronales artificiales, con un enfoque que utiliza comités de validación cruzada, para estimar la eficiencia isentrópica del compresor.

El uso de estos modelos entrenados permite conocer de antemano cómo evolucionará el ensuciamiento local de un componente rotativo industrial, lo cual es útil para realizar simulaciones de predicción del estado de la máquina en el tiempo como ayuda a la planificación del mantenimiento y para calcular el impacto de las variables térmicas sobre la salida del sistema modelado.

El cierre de esta metodología basada en datos se realiza con un caso para el desarrollo de dos algoritmos que permiten transformar la información general de un conjunto de intervenciones históricas de mantenimiento, en diferentes variables de un modelo de aprendizaje de máquina, las cuales son utilizadas para el modelado de las emisiones ambientales de un motor diesel de generación eléctrica con métodos basados en árboles de decisión y manteniendo la misma

xxviii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

28 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

arquitectura de ensemble planteada inicialmente, cuyo resultado final pone de manifiesto la capacidad del procedimiento creado para adaptarse a cualquier caso de estudio y supervisión de mantenimiento industrial, por medio de diferentes algoritmos de computación inteligente.

xxix

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

29 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

29 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

30 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Abstract

The data-driven maintenance methodology has been progressively applied in this work, so that its study and development serves to define a computational application procedure within a maintenance line in the industrial sector.

It starts from an approach which is focus on the screening and predictive systems of the concentration of dust in the air, which is vital for economic activity and the health of the population. In this section of the study, we use a set of artificial neural networks that we structure through the learning in order to ensemble and to obtain a complex variable, such as the concentration of dust based on real data, air temperature, relative humidity, atmospheric pressure and wind speed among others. The statistical performance indices obtained show the effectiveness of the proposed cross-validation committee. Accordingly, it is essential to have a reliable calculation method to determine the relative importance that can be applied to this type of ensemble architecture by means of artificial neural networks.

In contrast to other methods of relative importance, in which the calculations are made directly based on the weights of the artificial neural network and whose results in ensemble configuration show a high dispersion, an original procedure is proposed. This selectively chooses the variation of the inputs to readjust the architecture of the neural network, allowing us to measure those variables with the greatest effect on the target variable, thus obtaining the multivariate influence on the surface dust concentration through a computational model. This method provides, at this way, a real alternative for calculating and estimating the relative

xxxI

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

31 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

31 / 181

Abstract

importance which can be generalized to any type of problem for multivariate systems modeled by artificial neural networks for both a simple configuration and an ensemble architecture.

On the other hand, the planning of industrial maintenance related to the electricity production is vital, since it allows us to obtain a current and future picture of an industrial component in order to optimize the human, technical and economic resources of the installation. In particular, we study the relationship between the degradation of the machine, the historical maintenance interventions carried out and the suspended dust of Saharan origin. For this purpose, we use again the initial computational procedure based on the average assembly of artificial neuronal networks, together with an approach that uses cross validation committees, to estimate the isentropic efficiency of the compressor.

The use of these trained models let us know in advance how the local fouling of an industrial rotary component will evolve, which is useful to perform predictive simulations of the machine's condition over time as an aid to maintenance planning and to calculate the impact of thermal variables on the output of the modeled system.

The closing of this data-driven methodology is done with a case for the development of two algorithms that allow transforming the general information of a set of historical maintenance interventions, into different variables of a machine learning model. These are used to model the environmental emissions of an electric generation diesel engine with methods based on decision trees and maintaining the same ensemble architecture as initially proposed. The final result shows the capacity of the procedure created to adapt to any case of study and supervision of industrial maintenance, by means of different intelligent computing algorithms.

xxxii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

32 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Introducción

1. Motivación

En la actualidad existe una tendencia al alza en la aplicación de métodos de computación inteligente y aprendizaje de máquina, aprovechando la ingente cantidad de datos que desde el sector industrial se disponen. Sin embargo, para aprovechar el potencial de estos modelos inteligentes, es necesario profundizar en su estudio para mejorar su entendimiento y procedimentar las diferentes etapas de su computación, de forma que puedan integrarse y así optimizar, los procesos de las principales actividades del sector industrial.

Por lo tanto, este proyecto de investigación dará respuesta, mediante el procedimiento de una metodología específica de mantenimiento basado en datos, a una serie de problemas de carácter industrial dentro del ámbito de la generación eléctrica, mediante la utilización de modelos de computación inteligente y también dentro de un escenario local con una climatología macaronésica.

Las dos primeras partes giran en torno a la problemática del ensuciamiento industrial. Para ello se plantea un caso particular asociado a la degradación por ensuciamiento de una máquina rotativa y un sistema filtrante, en particular un compresor de una turbina de gas para generación de energía eléctrica. Con el fin de proteger la actividad industrial de las islas, la información en tiempo real sobre la concentración de polvo en suspensión llega a ser una pieza clave para conocer el estado de la calidad del aire. Para ello, dentro de la búsqueda para mejorar el entendimiento y explicación de estos modelos de aprendizaje de máquina

xxxiii

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

33 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

33 / 181

Introducción

basados en redes neuronales artificiales, se ha propuesto un procedimiento de computación propio que permite determinar la importancia de cada variable en la formación y evolución de estos factores de ensuciamiento, cuyo conocimiento es vital para una planificación de mantenimiento en una ventana de medio/largo plazo. De esta manera permitirá disponer de una foto actual y futura de la degradación de la máquina, a la vez que se logra optimizar los recursos humanos, técnicos y económicos.

Dentro del tercer y último bloque se desarrollará y aplicará la metodología computacional para la integración de intervenciones de mantenimiento en los modelos inteligentes. Aplicando para esto un caso real de un motor diesel de generación eléctrica y el modelado de sus emisiones ambientales.

Para que se pueda disponer de una comparativa de modelos de estudio validados, asociados a esta problemática real, se utilizará un volumen y una variedad de datos que se sitúan dentro de la terminología computacional conocida como *Big Data*. Esto nos permite aplicar algoritmos de computación inteligente, particularmente una red neuronal artificial (RNA) en configuraciones de *ensemble*, además de otros modelos no lineales para regresión como *Gradient boosting* (GB).

De acuerdo a lo anterior, será fundamental obtener modelos sobre los diferentes bloques planteados, que permitan disponer de un sistema automático inteligente y autónomo, capaz de simular escenarios futuros, estudiar el grado de afectación de las intervenciones de mantenimiento y simularlas en función del entorno ambiental local y el régimen de funcionamiento de la planta.

2. Objetivos

La especificación de los propósitos de esta tesis doctoral se resumen a continuación y por capítulo de desarrollo, para situar al lector con el aporte fundamental de este trabajo:

xxxiv

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

34 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

2. Objetivos

2.1. Análisis multivariante mediante ensemble de redes neuronales artificiales

El desarrollo del bloque de aproximación al estudio de computación inteligente, posibilita la definición de los siguientes puntos:

- Estimar el valor de concentración de calima, que es el término local utilizado en las Islas Canarias, a partir de variables medidas básicas como temperatura ambiente, humedad relativa, velocidad de viento y presión atmosférica. Para ello, desarrollamos un modelo de RNA en el que aplicamos una arquitectura de *ensemble* o comité de validación cruzada (CVC).
- Abrir la *black box* de una RNA, donde para diseñar un sistema híbrido en el cálculo de importancias relativas (IR), se construye el algoritmo *Selective importance measure* (SIM), combinando un análisis de sensibilidad que elimina entradas de forma selectiva, con un nuevo ajuste de la arquitectura sobre las capas de la red y fijando los mismos pesos iniciales para cada nuevo entrenamiento.
- Validar la fortaleza del método *SIM*, estimando qué variables medidas dentro del entorno climático de las Islas Canarias tienen mayor impacto sobre la concentración de calima y justificar estos resultados con la literatura relacionada.
- Verificar la estabilidad del procedimiento *SIM* sobre un aprendizaje realizado mediante RNA en una arquitectura de ensemble, supervisando la dispersión de sus resultados y comparándolo con otro método clásico relacionado, donde se calculan las importancias directamente sobre los pesos de la RNA.

xxxv

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

35 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Introducción

2.2. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

Los desarrollos de este capítulo para iniciar el estudio de eventos de mantenimiento, conllevan la propuesta de las siguientes metas:

- Analizar la composición del ensuciamiento de un compresor axial de una turbina de gas de generación eléctrica en las Islas Canarias y medir el efecto que tiene la calima sobre la eficiencia térmica del compresor y los filtros de entrada. Para esto, se desarrollan modelos de RNA en arquitectura de ensemble CVC.
- Transferir a un modelo de computación inteligente o *machine learning* (ML), con el desarrollo de un *algoritmo para convertir en una variable computacional, un evento de mantenimiento de un equipo particular (MTWork1)*. Para aplicarlo sobre mantenimientos de limpieza *offline* del compresor de una turbina de gas, de forma que el modelo final sea capaz de entender y utilizar dicha información dentro de su aprendizaje.
- Evaluar la fortaleza de la arquitectura propuesta de ensemble CVC aplicada un caso de mantenimiento real, planteando una comparativa de diferentes modelos obtenidos con diferentes métodos de ML para problemas de regresión. Además se incluye una fase de evaluación para comparar diferentes combinaciones de *k-fold* y observar su evolución sobre los índices de rendimiento estadístico y cómo afectan al valor de importancias *SIM*.
- Aplicar los resultados del proceso computacional como parte del desarrollo de mantenimiento basado en datos (MBD), donde se propone construir matrices de simulación con el resultado de las actividades de limpieza de los compresores en cualquier estación del año dependiendo de las horas de operación, las condiciones climáticas locales y las condiciones actuales de funcionamiento de la máquina. Esto permite estudiar y simular cómo evoluciona la degradación debida al ensuciamiento a lo largo del tiempo.

xxxvi

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

36 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

2. Objetivos

2.3. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

La aplicación de esta metodología de mantenimiento basado en datos permitirá alcanzar los siguientes objetivos:

- Concluir y definir los pasos de una metodología de MBD y su procedimiento computacional de base.
- Desarrollo de un *algoritmo para convertir en una variable computacional, un evento general de mantenimiento (MTWork2)*, para trasladarlos al aprendizaje de un modelo de ML de un motor diesel de generación eléctrica y verificar la relación de estos mantenimientos de revisión con las emisiones de óxidos de nitrógeno (NOx).
- Crear un *algoritmo para el cálculo de importancias en eventos históricos de mantenimiento (MTImp)* para, a partir de cualquier modelo de ML entrenado, este pueda medir de forma autónoma, el grado de afección de cada uno de los eventos de mantenimiento históricos realizados sobre un elemento industrial, siendo capaz de utilizar su aprendizaje para tener en cuenta, tiempos de estabilización de la máquina y cualquier escenario ambiental y de funcionamiento sin necesidad de corrección de estos factores de forma teórica.
- Justificar los resultados de *MTImp* sobre el modelo realizado de GB, comparándolo con otros métodos de cálculo de IR relacionados con este modelo de ML.
- Evaluar el modelo entrenado realizando simulaciones en los diferentes eventos históricos de mantenimiento realizados en el motor diesel de generación eléctrica, de forma que se pueda comparar en cada intervención el resultado real, el predicho y simular qué hubiera pasado si no se hubieran realizado estas intervenciones de mantenimiento.

XXXVII

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

37 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Introducción

3. Estructura de la tesis

El documento de la tesis se estructura de la siguiente manera:

En el Capítulo 1 se describe el estado del arte, donde desarrollamos algunas de las investigaciones más relevantes de trabajos relacionados con el mantenimiento basado en datos, métodos de computación inteligente o aprendizaje de máquina y el entorno ambiental de la Macaronesia en las Islas Canarias, para finalmente concluir con los fundamentos de los sistemas de generación térmica estudiados en este trabajo.

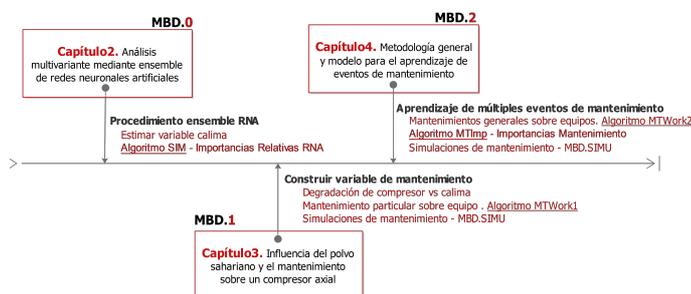


Figura 1. Esquema de la línea de desarrollo para la metodología de MBD

En el Capítulo 2 comienza la secuencia de desarrollos de la línea de MBD, siendo esta parte la correspondiente a la **fase de aproximación al mantenimiento basado en datos (MBD.0)**. En ella se realiza un estudio para la estimación de polvo sahariano o *calima* mediante el diseño y aplicación del procedimiento computacional de referencia en este trabajo de investigación. Además, se profundiza en la aplicación de redes neuronales artificiales y sobre éstas, se propone una nueva metodología para el estudio de la contribución de cada variable de entrada sobre la variable objetivo.

XXXVIII

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

3. Estructura de la tesis

En el Capítulo 3 se presenta el avance correspondiente a la **fase inicial del mantenimiento basado en datos (MBD.1)**. Aplicando los mismos procedimientos computacionales desarrollados en MBD.0, se estudia el grado de afección de la calima y las diferentes intervenciones históricas de mantenimiento de limpieza, sobre la degradación del compresor axial de una turbina de gas para generación eléctrica. Donde se establece el primer método de integración de una variable de mantenimiento en un modelo de aprendizaje de máquina.

En el Capítulo 4 se resume y define la metodología general de la **última fase del mantenimiento basado en datos (MBD.2)**, aplicando un caso de estudio sobre la relación de las diferentes intervenciones históricas de mantenimiento en un motor diesel de generación eléctrica, proponiendo un método general para el aprendizaje de importancias de cualquier intervención de mantenimiento, en función de las condiciones ambientales y de funcionamiento.

xxxix

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

39 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

39 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

40 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

40 / 181

Capítulo

1

Estado del arte

En este capítulo se presenta el estado del arte, donde de forma resumida, se tratan y se inician los conceptos y desarrollos que servirán de línea base para el correcto seguimiento por parte del lector, en los temas principales de este trabajo. Se desarrollará la evolución del *mantenimiento basado en datos*, el estudio de *técnicas de aprendizaje de máquina* y se concluye describiendo el escenario de trabajo, dentro del entorno de las Islas Canarias y el sector de la *generación eléctrica*.

1.1. Mantenimiento basado en datos

El proceso de mejora digital en las industrias se ha concebido como la cuarta revolución industrial o *Industria 4.0*, lo cual supone un cambio de paradigma impulsado por el auge y la madurez progresiva de las nuevas tecnologías de la información y las comunicaciones, aplicadas a los procesos y productos industriales. Desde la perspectiva de la ciencia de los datos, este cambio de paradigma permite extraer los conocimientos pertinentes de los activos vigilados mediante la adopción de estrategias inteligentes de vigilancia y fusión de datos, así como mediante la aplicación de métodos de aprendizaje y optimización de las máquinas (Diez-Olivana *et al.* (2019)).

Para poder desarrollar estos métodos de *ML*, es necesario estructurar el volumen de datos siguiendo las oportunidades que ofrece el *Big Data*. Estas son innegables, pero es necesario estudiarlas y darles un sentido de aplicación real.

1

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

41 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

41 / 181

Capítulo 1. Estado del arte

La tecnología industrial actualmente genera enormes cantidades de datos, por lo que debemos ser capaces de extraer la mejor información de ellos y usarla para tomar las mejores decisiones (Torrecilla y Romo (2018)). Utilizando tres direcciones diferentes: velocidad, complejidad (cantidad de variables y variedad de estructuras de datos) y tamaño de muestra de datos (Figura 1.1) (Koseleva y Ropaite (2017)).

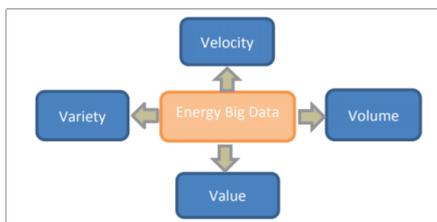


Figura 1.1. Características principales del Big Data (Koseleva y Ropaite (2017)).

En este contexto, el MBD ha ido adquiriendo una creciente relevancia en el campo de la investigación del mantenimiento en los últimos años. Junto al aprendizaje automático hace uso de modelos que pueden analizar esta cantidad significativa de datos y estimar un modelo de trabajo del sistema sólo en base a los datos recogidos por los sensores.

Además, los sistemas autónomos son máquinas inteligentes que ejecutan tareas de alto nivel sin una programación detallada y sin control humano. Son capaces de decidir entre un conjunto de acciones alternativas, orquestar y ejecutar habilidades (Rosen *et al.* (2015)). Para que esto suceda, los sistemas autónomos necesitarán acceder a modelos muy realistas del estado actual del proceso y de su propio comportamiento en interacción con su entorno en el mundo real, comúnmente llamado el *gemelo digital* o *digital twin* (DT). Elementos facilitadores para la transformación hacia la *Industria 4.0* y que se refieren a un duplicado virtual de un activo físico, producto o proceso, usado para entender, predecir y optimizar las características de rendimiento de la parte física (Herterich (2017)). Son nece-

2

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

42 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

42 / 181

1.1. Mantenimiento basado en datos

sarios e indispensables para el diseño virtual y la optimización de los sistemas industriales inteligentes (Murat *et al.* (2020)).

Una consecuencia de la autonomía es el fuerte aumento de la complejidad para asegurar el comportamiento adecuado del sistema durante el curso de la producción para alcanzar el objetivo de producción deseado. Esto sólo puede lograrse razonablemente mediante el uso extensivo de la simulación basada en modelos no sólo durante el diseño y la planificación, sino también durante las demás fases del ciclo de vida con fines tales como el diagnóstico y la optimización de las operaciones (Figura 1.2) (Rosen *et al.* (2015)).

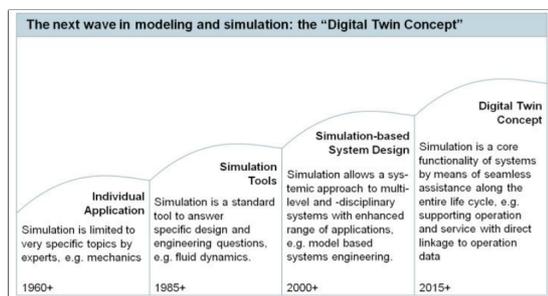


Figura 1.2. El gemelo digital - Evolución de la tecnología de simulación (Rosen *et al.* (2015)).

El gemelo digital tiene por objeto reducir el costo de las pruebas y la verificación del sistema. A diferencia del modelo virtual ordinario, que describe el producto sin ninguna imperfección, el gemelo digital del producto representa una instancia particular de éste en diferentes etapas del ciclo de vida (prueba, producción, mantenimiento, eliminación). Y debe utilizarse para el mantenimiento prescriptivo de un sistema o un producto en funcionamiento, existiendo diferentes formas de realizar el mantenimiento preventivo de un producto en funcionamiento, entre ellas el análisis descriptivo, diagnóstico, predictivo y prescriptivo (Nikolaev *et al.* (2019)).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 1. Estado del arte

El mantenimiento prescriptivo consiste en la evolución natural del mantenimiento predictivo, siendo hasta ahora la técnica más avanzada y compleja disponible (Figura 1.3). Aconseja cuándo intervenir en función de las condiciones y los recursos disponibles, de manera que afecte lo menos posible al ritmo normal de producción. Sin embargo, el mantenimiento tradicional tiende a ser reactivo, respondiendo a fallos en los equipos o dispositivos después del hecho. Este enfoque de describir los fallos después de que se hayan producido (mantenimiento correctivo) es el peor de los casos para el mantenimiento.

El mantenimiento prescriptivo va más allá del ámbito del mantenimiento correctivo, preventivo y predictivo. El descriptivo se centra en lo que ocurrió en el pasado. El análisis predictivo descubre opciones potenciales para el futuro. El mantenimiento prescriptivo aprovecha todos estos enfoques y capacidades. El ámbito de lo que debería suceder y la ejecución de un mantenimiento optimizado es precisamente el reino del mantenimiento prescriptivo. Con el mantenimiento prescriptivo, los dispositivos, en colaboración con los operadores, son participantes proactivos en su propio mantenimiento (Khoshafian y Rostetter (2015)).



Figura 1.3. Evolución histórica del mantenimiento industrial

La tecnología de gemelos digitales para sistemas en tiempo real ha impulsado la investigación en nuevas formas de convertir los datos en información útil para la toma de decisiones. Más específicamente, la aplicación de algoritmos de aprendizaje de máquina para el pronóstico y el mantenimiento predictivo ha ido aumentando.

Zheng y Ostrowski plantearon la sistemática en la tecnología de los gemelos

4

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

1.1. Mantenimiento basado en datos

digitales para el mantenimiento, la reparación y la revisión en el sector industrial (Figura 1.4). Esta cuarta revolución industrial está dotando a los fabricantes de soluciones para la reducción de costes, lo que se traduce en una ventaja competitiva. El sector de las operaciones de mantenimiento está a la cabeza de la innovación en ingeniería, desde las técnicas reactivas hasta las preventivas planificadas. El siguiente paso en la historia de los enfoques proactivos es el mantenimiento analítico predictivo (Zheng *et al.* (2018));(Ostrowski y Menyhárt (2019)).

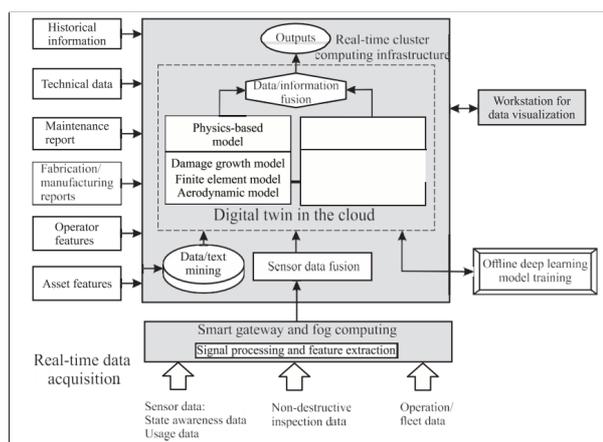


Figura 1.4. La arquitectura del sistema DT (Zheng *et al.* (2018)).

El estudio realizado por Edwin, presentó un modelo que utilizaba datos de inspección y modelos de fiabilidad para observar el envejecimiento de los componentes y los costes de mantenimiento asociados en un cuadro de riesgos y con los costes previstos en función de la estrategia de mantenimiento elegida por el usuario (Edwin *et al.* (2018)).

Con un enfoque similar, se presentó un novedoso enfoque para realizar la planificación prescriptiva del mantenimiento mediante la elaboración de conclusio-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 1. Estado del arte

nes basadas en la interacción entre los datos de los programas de producción, supervisión de las condiciones en tiempo real de la planificación a corto y medio plazo, así como los datos históricos de calidad y de fallos de las máquinas, obteniendo resultados excelentes. Sin embargo, es necesario trabajar en el futuro para adaptar automáticamente el conjunto de reglas propuesto a un entorno de producción dinámico, lo que actualmente se hace de forma periódica (Matyasa *et al.* (2017)).

Siguiendo este tipo de desarrollos y de forma procedimentada, Duer se centró en un conjunto de información determinado por una RNA para trabajar sobre un conjunto de información de mantenimiento que consiste en un grupo de aquellos elementos del equipo que necesitan ser intervenidos, un conjunto de actividades preventivas que regeneran elementos del equipo y un conjunto de normas de mantenimiento, con la idea de realizar un diagnóstico binario constante para determinar el momento en que debe comenzar la intervención del equipo (Duer *et al.* (2012)).

El sector industrial está llamando la atención ya que las técnicas de mantenimiento predictivo pueden aplicarse a diferentes campos. Específicamente, la aplicación del MBD al sector energético ha sido un campo de estudio reciente. Algunos estudios han aplicado modelos de RNA sobre la prognosis y diagnóstico de las turbinas de gas.

Kiakojoori *et al.* plantearon la utilización de redes neuronales dinámicas para medir la degradación por ensuciamiento del compresor. Sin embargo indicaron que su solución era mejorable si se aplicaban estrategias de aprendizaje de ensemble con RNA (Kiakojoori y Khorasani (2015)). Tahan *et al.*, avanzaron en el desarrollo de un modelo *multi-net*, donde entrenan un conjunto de RNA, consiguiendo su mejor resultado en una arquitectura monocapa, con una fase de entrenamiento que utiliza regularización Bayesiana, para finalmente combinarlas mediante la media resultante de una partición obtenida por el método de validación cruzada (VC) con un k -fold=4. Llegaron a aplicar este modelo de ensemble para la detección de ensuciamiento del compresor, aumentando la fiabilidad del diagnóstico y asegurando la detección de fallos en una etapa temprana (Mohamma-

6

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

46 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

46 / 181

1.1. Mantenimiento basado en datos

dreza *et al.* (2017)).

Amozegar, propone trabajar con ensemble dinámicos para monitorizar una turbina de gas. Específicamente hace distinción entre un tipo de ensemble homogéneo, que utiliza el mismo algoritmo de aprendizaje y uno heterogéneo donde cada miembro del ensemble utiliza un algoritmo de aprendizaje diferente. En ambos casos el rendimiento supera a los modelos que fueron entrenados de forma individual (Amozegar y Khorasani (2016)).

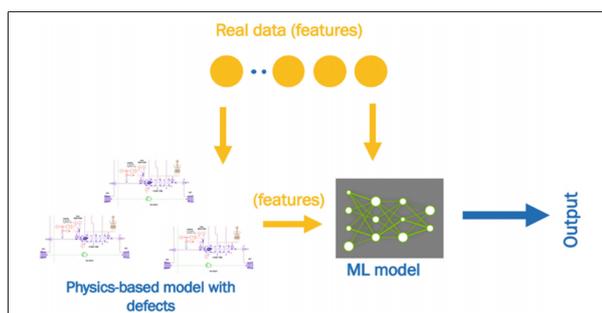


Figura 1.5. Modelo híbrido de un sistema para su análisis predictivo (Nikolaev *et al.* (2019)).

El enfoque de un modelado híbrido para el análisis prescriptivo fue probado también en el mantenimiento de una turbina de gas por Nikolaev (Figura 1.5), donde presentaron una aplicación de análisis de datos para el desarrollo de un modelo basado en la física de la turbina y su combinación para identificar las causas de las anomalías y defectos. Mediante el uso de la simulación y el modelado de aprendizaje automático de un inyector (Nikolaev *et al.* (2019)).

A la hora de establecer comparativas para verificar el rendimiento de los modelos computacionales aplicados, Baptista planteó contrastar diferentes métodos sobre un caso de uso basado en datos en la prognosis de una turbina de gas, en la que se probaron cinco métodos diferentes para los modelos de pronóstico, co-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 1. Estado del arte

mo la regresión lineal, los *K-Nearest Neighbour*, las RNA, Random forest (*RF*) y máquinas de soporte vectorial. Todos los métodos mostraron un rendimiento razonablemente bueno, pero con un gran potencial de mejora, especialmente en los campos de la selección de parámetros y la generalización a otras situaciones de pronóstico para estudiar metodologías formales en la validación de los enfoques basados en datos, e investigar las estimaciones (Baptista y Nascimento (2017)).

En línea con lo anterior, es fundamental plantear estrategias de mantenimiento dinámico para sistemas sujetos a envejecimiento y degradación, ya que durante el funcionamiento, el sistema pasa por un proceso de envejecimiento que afecta a la fiabilidad del sistema.

Liu lo analizó como primer paso, seguido de un modelo de mantenimiento en base a la tasa de coste esperada dentro de un ciclo de reparación. Proponía una estrategia de mantenimiento dinámico, en la que se actualizan los parámetros de degradación y se proporciona el tiempo óptimo de inspección posterior en cada acción de mantenimiento. Su modelo de mantenimiento propuesto tiene dos ventajas en comparación con los modelos existentes de mantenimiento preventivo basado en el tiempo. En primer lugar, incorpora los efectos tanto del envejecimiento como de la degradación. Y en segundo lugar, permite una toma de decisiones dinámica en el tiempo de inspección óptimo y las acciones de mantenimiento asociadas en función del nivel de degradación observado y el historial de funcionamiento, adaptándose así a condiciones de funcionamiento heterogéneas (Liu *et al.* (2019)).

Cabe destacar además, la aplicación de otros estudios sobre otra tecnología energética como es la de los motores diesel, donde por ejemplo Liu, aplicó diferentes métodos de *ML* para la predicción de NOx. Cuyo cálculo a priori, podría ser difícil de estimar debido a la complejidad del proceso de combustión.

Para la fase inicial de la modelización utilizaron el análisis de componentes principales para eliminar la redundancia y reducir la dimensión del conjunto de datos, mientras que el modelo final se entrenó con máquinas de soporte vectorial optimizando hiperparámetros mediante el algoritmo *GridSearch*. Este algoritmo

8

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

48 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

48 / 181

1.1. Mantenimiento basado en datos

de optimización presenta una estrategia de búsqueda en rejilla, no de forma aleatoria, sino en puntos de dicho espacio y repartidos de forma espaciada en una rejilla o rango determinado. Consiguiendo que el modelo de predicción para emisiones de NOx se estimara sobre la base de los datos de funcionamiento de un motor diesel obtenidos mediante pruebas, tanto en estados estables como transitorios. Por lo tanto el modelo de predicción propuesto puede utilizarse como un sensor sustituto, debido a su gran precisión y simplicidad en en la modelización (Liu *et al.* (2016)).

Sobre esta línea de investigación, Kowalski presenta un novedoso enfoque para el diagnóstico inteligente de fallos en motores marinos diesel de cuatro tiempos, donde la seguridad y la fiabilidad del barco dependen en gran medida del rendimiento de dicho motor; de lo que se deduce que la detección temprana de cualquier tipo de fallo que afecte al motor es crucial (Kowalski *et al.* (2017)).

Estos sistemas de diagnóstico automático son de especial importancia porque pueden operar continuamente en tiempo real, proporcionando así un control eficiente del rendimiento del motor. Y supervisando a la vez varias señales reales del mismo, utilizadas como entradas para un algoritmo de clasificación de patrones mediante un conjunto de modelos de *ML* que trabajan en un modo de descomposición y clasificación binaria (Kowalski *et al.* (2017)).

Otra aplicación importante del MBD es la evaluación de la eficacia de las acciones de mantenimiento y el poder cuantificar la importancia relativa de las diferentes acciones de mantenimiento. El análisis de estos datos puede proporcionar al sector industrial una enorme ventaja competitiva, identificando las actividades y procesos de mantenimiento improductivos. Mosaddar, centra su estudio en identificar y construir un modelo para identificar aquellas actividades de mantenimiento específicas que eran ineficientes, de forma que el modelo sirviera de ayuda a los encargados de tomar decisiones de mantenimiento para establecer objetivos en los sistemas de mantenimiento bajo su supervisión. En consecuencia, el sector industrial puede centrarse en la rectificación de estas actividades

9

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

49 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

49 / 181

Capítulo 1. Estado del arte

menos eficientes y por lo tanto, en la reducción de los costes asociados a su realización (Mosaddar y Shojaie (2013)).

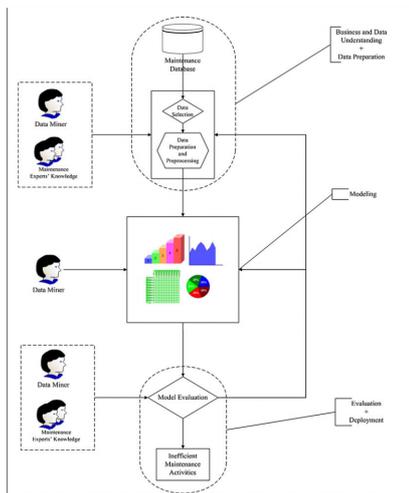


Figura 1.6. Modelo basado en la minería de datos (*data mining*) para identificar actividades de mantenimiento ineficientes (Mosaddar y Shojaie (2013)).

En su estudio, Mosaddar encontró que efectivamente, existe un vacío de investigación en la literatura disponible dentro del cálculo de importancias relativas de mantenimiento mediante técnicas de *ML* (Mosaddar y Shojaie (2013)). En su trabajo presentó un modelo basado en datos para identificar las actividades de mantenimiento ineficientes (Figura 1.6), donde agrupó los registros de mantenimiento de los diferentes equipos en *clusters* y utilizando la distancia euclidiana analizó las condiciones de los registros en los *clusters*. Finalmente mediante el uso de algoritmos de reglas de asociación, identificó las actividades que conducen a que los registros se coloquen en cada *cluster* y que mediante el uso técnicas de

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA Fecha: 20/01/2021 19:43:20
 Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA 20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

1.2. Métodos de computación inteligente

análisis exploratorio de datos, se podían extraer los indicadores de un mantenimiento ineficiente (Mosaddar y Shojaie (2013)).

1.2. Métodos de computación inteligente

Para poder conocer las características y principales métodos de aprendizaje de máquina en la línea base de los desarrollos de este trabajo de investigación, es necesario profundizar en el conocimiento de RNA y modelos basados en árboles de decisión como *RF* y *GB*, tanto en configuración simple como en una estructura de comité o *ensemble*.

Las RNA están inspiradas en las redes neuronales biológicas. Tratan de imitar el comportamiento del cerebro para desarrollar determinadas acciones tales como reconocimiento de patrones o clasificación de objetos (Jesan y Lauro (2003)).

El modelo de una neurona artificial se presenta en la Figura 1.7. En una neurona se obtiene una señal de salida sumando la multiplicación de las entradas por los pesos sinápticos de la neurona, el resultado de la sumatoria pasa por una función de activación (Haykin (1999)).

Haykin define las RNA de la siguiente manera: *“Una RNA es un procesador distribuido masivamente en paralelo, compuesto de unidades de procesamiento simples, que tiene una propensión natural a almacenar el conocimiento experimental y a ponerlo a disposición para su uso. Se asemeja al cerebro en dos aspectos.*

1. *El conocimiento es adquirido por la red a partir de su entorno mediante un proceso de aprendizaje.*
2. *Las fortalezas de la conexión interneuronal, conocidas como pesos sinápticos, se utilizan para almacenar el conocimiento adquirido”.*

Matemáticamente se puede expresar con las siguientes ecuaciones:

11

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

51 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

51 / 181

Capítulo 1. Estado del arte

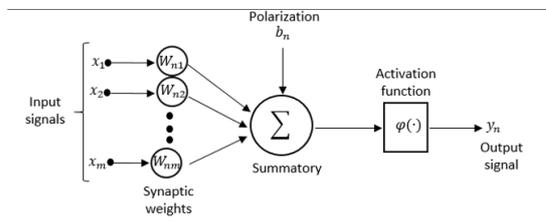


Figura 1.7. Modelo no lineal de una neurona artificial. Adaptado de Haykin (1999).

$$u_n = \sum_{j=1}^m w_{nj}x_j \quad (1.1)$$

$$y_n = \varphi(u_n + b_n) \quad (1.2)$$

Donde $x_1 \dots x_m$ son señales de entrada, $w_{n1} \dots w_{nm}$ son los pesos sinápticos de la neurona, b_n es la polarización o umbral, $\varphi(\cdot)$ es la función de activación e y_n es la señal de salida de la neurona. En la Ecuación (1.1), las señales de entrada se ponderan con las respectivas fortalezas sinápticas de la neurona; estas operaciones constituyen un tipo de combinador lineal (Haykin (1999)).

Podemos concentrar las características principales de una RNA en los siguientes aspectos (Haykin (1999)).

- **No linealidad:** Las neuronas son elementos de proceso generalmente no lineales. La interconexión de estos elementos genera estructuras de transformación de datos donde este carácter no lineal queda distribuido por toda la red.
- **Modelado de relaciones entrada-salida:** Capaces de aprender a partir de un conjunto de datos entrada-salida sin necesidad de conocer a priori las características del problema. Genera un mapa de entrada-salida para después

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

1.2. Métodos de computación inteligente

poder reproducir casos no representados en el conjunto de datos original.

- **Aprendizaje adaptativo:** Las redes neuronales artificiales tienen la cualidad de aprender a realizar tareas a partir de un entrenamiento, debido a la capacidad de "autoajuste" de los elementos de procesado (neuronas) que componen el sistema. En el proceso de entrenamiento, los pesos que conectan las neuronas se van ajustando de manera que la red reproduce los resultados buscados.
- **Tolerancia a fallos:** Las redes neuronales artificiales pueden aprender a reconocer patrones con ruido, distorsionados o incompletos siempre y cuando haya suficiente información correcta para ser capaz de distinguir errores.
- **Fácil inserción dentro de la tecnología existente:** El alto grado de conectividad de las redes neuronales artificiales y su rapidez de ejecución las hace ideales para su implementación en sistemas modulares. Una red individual puede ser entrenada para desarrollar una tarea y ser integrada rápidamente dentro de implementaciones de hardware para realizar este trabajo específico. Su forma de funcionamiento tipo caja negra (introducimos unas entradas y estas nos producen unas salidas), las hace recomendables para estos sistemas modulares.
- **Operación en tiempo real:** A pesar de que el entrenamiento de las redes neuronales artificiales pueda ser muy costoso en tiempo, una vez entrenadas, su respuesta ante las nuevas variables de entrada es muy rápida .

En los **tipos** de RNA, se encuentran las Regresivas y *Feedforward*, siendo la diferencia entre ambas RNA, que la primera tiene al menos un lazo de retroalimentación. Por ejemplo, una red *Feedforward* puede consistir en una capa de neuronas con cada neurona alimentando su señal de salida a las entradas de todas las otras neuronas sin bucles de retroalimentación hacia la propia entrada de la neurona (Haykin (1999)).

13

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

53 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

53 / 181

Capítulo 1. Estado del arte

Clasificándolas por el tipo de **arquitectura**, destaca la Monocapa *Feedforward*. En cuya RNA las neuronas están organizadas en forma de capas. En la forma más simple de una red de este tipo, se dispone de una capa de entrada de nodos fuente que se proyecta directamente en una capa de salida de neuronas (nodos de computación), pero no al revés.

Sin embargo la arquitectura Multicapa *Feedforward* se distingue por la presencia de una o más capas ocultas, cuyos nodos de cómputo se denominan correspondientemente neuronas o unidades ocultas; el término "oculto" se refiere al hecho de que esta parte de la red no se ve directamente desde la entrada o la salida de la misma (Haykin (1999)).

Otro concepto necesario y clave en la composición de las RNA son las **conexiones sinápticas** (Pesos y Bías), donde el conjunto de sinapsis, o enlaces de conexión, se caracterizan por un peso o fuerza propios. Específicamente, una señal x_j a la entrada de la sinapsis j conectada a la neurona n se multiplica por el peso sináptico w_{nj} . Es importante anotar la manera en que se escriben los subíndices del peso sináptico w_{nj} . El primer subíndice en w_{nj} se refiere a la neurona en cuestión, y el segundo subíndice se refiere al extremo de entrada de la sinapsis a la que el peso se refiere. A diferencia del peso de una sinapsis cerebral, el peso sináptico de una neurona artificial puede estar en un rango que incluye valores negativos y positivos. En otras palabras, estos pesos, permiten almacenar el conocimiento aprendido [35]. Los bias b_n , o umbral de polarización, tienen el efecto de aumentar o disminuir la entrada neta de la función de activación, dependiendo de si es positivo o negativo, respectivamente.

Para superar las limitaciones prácticas del perceptrón clásico de Rosenblatt (Figura 1.7), se desarrolla el **perceptrón multicapa** (Figura 1.8), donde cabe destacar algunas características básicas (Haykin (1999)) :

- El modelo de cada neurona en la red incluye una función de activación no lineal que es diferenciable.

14

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

54 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

54 / 181

1.2. Métodos de computación inteligente

- La red contiene una o más capas que están ocultas tanto de la entrada como nodos de salida.
- La red muestra un alto grado de conectividad, cuyo alcance se determina por pesos sinápticos de la red.

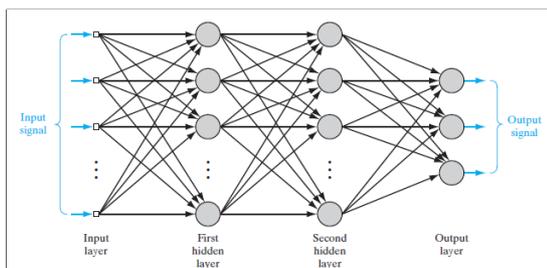


Figura 1.8. Arquitectura clásica de un perceptrón multicapa (Haykin (1999)).

Las **neuronas ocultas** actúan como captador de características y como tal, juegan un papel crítico en el funcionamiento de un perceptrón multicapa. A medida que el proceso de aprendizaje avanza a través del perceptrón multicapa, las neuronas ocultas comienzan a "descubrir" gradualmente las características sobresalientes que caracterizan los datos de entrenamiento. Lo hacen mediante la realización de una transformación no lineal en los datos de entrada en un nuevo espacio llamado espacio de características. En este nuevo espacio, las clases de interés en una tarea de clasificación de patrones, por ejemplo, pueden separarse más fácilmente las unas de las otras, de lo que podría ser el caso en el espacio de datos de la entrada original. De hecho, es la formación de este espacio funcional a través del aprendizaje supervisado lo que distingue al perceptrón multicapa del perceptrón de Rosenblatt (Haykin (1999)).

Para comprender las fases de aprendizaje de una RNA, dichas fases pueden clasificarse como supervisadas, según Haykin "*Learning with a teacher*" (Haykin

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 1. Estado del arte

(1999)). Como el caso del perceptrón multicapa, en el que los valores de salida deseados ya se conocen y se transmiten al sistema para que éste los produzca ajustando los pesos sinápticos para minimizar iterativamente el error hasta obtener el resultado deseado. Al contrario de un aprendizaje no supervisado, el cual no requiere un conjunto de datos preliminares, lo que significa que los resultados pueden no ser conocidos.

La idea de un **árbol de decisión** es dividir el espacio de las variables de entrada en áreas rectangulares homogéneas mediante un sistema de reglas basado en árboles. Cada división del árbol corresponde a una regla de *si o entonces* sobre alguna variable de entrada. Dividiendo el espacio de las características en varias regiones más pequeñas (no superpuestas) con valores de respuesta similares utilizando un conjunto de reglas de división. Las predicciones se obtienen ajustando un modelo más simple, por ejemplo, una constante como el valor de respuesta promedio, en cada región. Esos métodos de dividir y conquistar pueden producir reglas simples que son fáciles de interpretar y visualizar con diagramas de árbol (Breiman (1996)).

El enfoque más frecuente de la modelización basada en datos es construir sólo un único modelo predictivo sólido. Se puede considerar la posibilidad de construir un conjunto de modelo "fuerte", como las RNA, que pueden combinarse aún más para producir una mejor predicción. Sin embargo, en la práctica, el enfoque de conjunto se basa en la combinación de un gran número de modelos simples relativamente débiles, como los árboles de decisión, para obtener una predicción de conjunto más fuerte. Los ejemplos más destacados de esas técnicas de aprendizaje de máquina son los **RF** (Natekin y Knoll (2013)). Estos "bosques", se propusieron originalmente por Breiman y como su nombre indica, un **RF** crea un conjunto con múltiples árboles de decisión aleatorios y agrega sus predicciones por promedio simple (Breiman (1996)). El método **RF** es un algoritmo de aprendizaje supervisado tanto para tareas de clasificación como de regresión (Figura

16

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

56 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

56 / 181

1.2. Métodos de computación inteligente

1.9). Aunque un *RF* es una colección de árboles de decisión, hay diferencias que deben destacarse: mientras que los árboles de decisión generan reglas y nodos a partir del cálculo de la ganancia de información y del índice *Gini*, los *RF* generan árboles de decisión de forma aleatoria (Carvalho *et al.* (2019)).

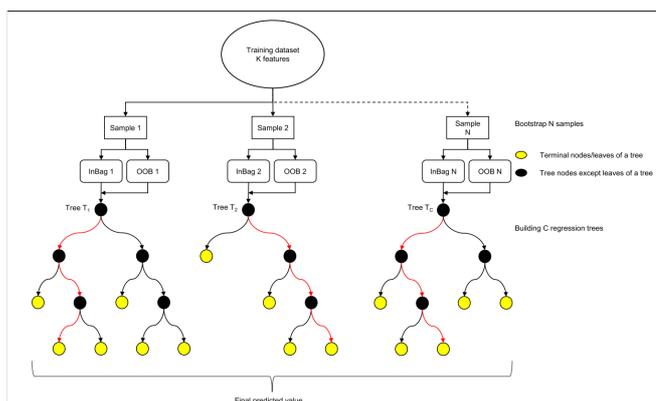


Figura 1.9. Esquema del algoritmo de *RF* (Ahmad *et al.* (2017)).

Además, si bien los árboles de decisión pueden sufrir de sobreentrenamiento, los *RF* pueden evitarlo en la mayoría de los casos, ya que trabajan con subconjuntos aleatorios de características y construyen árboles más pequeños a partir de esos subconjuntos (Breiman (1996)); (Prytz *et al.* (2015)); (Biau y Scornet (2016)).

También permiten evaluar la importancia relativa de las entradas, lo que es útil para que la reducción de la dimensionalidad mejore el rendimiento del modelo en conjuntos de datos de alta dimensión (Ahmad *et al.* (2017)). Para esto, *RF* cambia una de las variables de entrada mientras mantiene la constante restante, y mide la disminución media de la precisión de la predicción del modelo, que

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 1. Estado del arte

luego se utiliza para asignar una puntuación de importancia para cada variable de entrada (Breiman (1996)).

A la hora de comenzar a trabajar combinando modelos de *ML*, nos acercamos a una configuración colaborativa, donde varios modelos de computación inteligente son agrupados en diferentes **arquitecturas de comité o ensemble**, en la búsqueda de un mejor rendimiento estadístico del modelo final.

El concepto principal detrás de un modelo de aprendizaje de ensemble, representa la sencilla idea de un comité de expertos trabajando juntos para resolver un problema. Todos los miembros contribuyen con su propia experiencia e iniciativas, y el grupo en su conjunto puede elegir entre apoyar o rechazar una nueva idea por sus propios méritos (Zhou *et al.* (2002)). Otros trabajos similares, han estudiado la aplicación de arquitecturas basadas en ensemble, sobre turbinas de gas (Parmanto *et al.* (1996));(Gevrey *et al.* (2003)).

De acuerdo con Saviozzi, el método ensemble promedio se suele aplicar para lograr resultados más precisos que los de una única RNA (Saviozzi *et al.* (2019)). La idea principal de este método es entrenar diferentes redes y combinar sus resultados para tener una mejor predicción. Por ejemplo, Siddhartha observó que los algoritmos de predicción con esta arquitectura, superaban al modelo estándar (Mandal y Kishore (2020)). Weng *et al.* muestran en su experimento con métodos de ensemble, una mejora de un 30% en el rendimiento de la prueba en comparación con una RNA simple de dos capas ocultas con las mismas características y mismo número de neuronas (Weng *et al.* (2018)). Otros estudios recientes que utilizan una combinación de múltiples RNA mediante configuración de ensemble, obtienen un mejor rendimiento y han alcanzado mejores resultados. Esto es útil para hacer frente a problemáticas de aplicación real (Borra y Ciaccio (2010));(Parmanto *et al.* (1996));(Xia *et al.* (2011));(Dong y Han (2004)).

Otros estudios basados en una combinación de métodos de ensemble y aprendizaje profundo (*deep learning*), también llegaron a la conclusión de que la mejor opción es agrupar las RNA en una configuración de ensemble para resolver los

18

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

58 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

58 / 181

1.2. Métodos de computación inteligente

problemas de regresión y clasificación; en este caso, por ejemplo, mediante un diseño de ensemble por *bagging* se apoya en un nuevo aprendizaje semisupervisado, frente al aprendizaje clásico supervisado. Este nuevo método se basa en una fase de aprendizaje paralelo donde utiliza el método no supervisado para añadir nuevas características extraídas con un modelo aprendizaje de regresión *Kernel* (Yi *et al.* (2015)).

La labor emprendida por Liu se amplía con otro caso en el que se comparan modelos basados aprendizaje por método *Kernel*, cuyo mejor rendimiento y fiabilidad de predicción se derivan de nuevo de un modelo basado en ensemble (Yi *et al.* (2018)). Sin embargo, es importante señalar que en un conjunto ensemble, no todas las RNA entrenadas darán buenos resultados. Localizar las RNA que muestran un bajo rendimiento en el modelo final es una tarea compleja (Zhou y Liu (2016)). En este sentido, Parmanto y Murrugarra propusieron y comprobaron la eficiencia, en cuanto al rendimiento del modelo, diseñando una estructura de comité a partir de la distribución que da el índice de VC, ya que las divisiones que se hacen en este método pueden utilizarse para excluir algunos subconjuntos de entrenamiento con cada iteración (Parmanto *et al.* (1996)); (Llerena *et al.* (2012)).

Cabe destacar además, otro de los métodos de ensemble basado en árboles de decisión y diseñado por Friedman. El algoritmo de *GB* se basa en el algoritmo general de *boosting* para construir modelos de ensemble más sencillos como árboles de decisión (Friedman (2002)).

La idea principal del algoritmo de *boosting*, es añadir nuevos modelos al conjunto de forma secuencial. En esencia, este ataca el equilibrio entre sesgo y varianza empezando con un modelo débil, por ejemplo, un árbol de decisión con sólo unas pocas divisiones y aumenta secuencialmente su rendimiento continuando la construcción de nuevos árboles, en los que cada nuevo árbol de la secuencia trata de corregir los errores más grandes cometidos por el anterior. Es decir, cada nuevo árbol de la secuencia se centrará en las filas de entrenamiento en las que el

19

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

59 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

59 / 181

Capítulo 1. Estado del arte

árbol anterior tenía los mayores errores de predicción (Friedman (2001)). Según Natekin, donde los *RF* construyen un conjunto de árboles independientes profundos, los *GB* construyen un conjunto de árboles poco profundos en secuencia con cada árbol, aprendiendo y mejorando el anterior (Natekin y Knoll (2013)).

Aunque se han adoptado muchas optimizaciones en las implementaciones de *GB*, la eficiencia y la escalabilidad son todavía insatisfactorias cuando la característica y el tamaño de los datos es grande. En particular, Guolin *et al.*, desarrollaron un algoritmo optimizado de *GB*, *LightGBM*. Mediante el uso de dos técnicas novedosas, el *Gradient-based One-Side Sampling* (GOSS) y el *Exclusive Feature Bundling* (EFB), consiguieron excluir una proporción significativa de instancias de datos del conjunto de datos (DS) con pequeños gradientes, y utilizaron el resto para estimar la ganancia de información, agrupando además, características excluyentes para reducir el número de estas características. Resultando ser hasta veinte veces más rápido que un algoritmo de *GB* convencional (Guolin *et al.* (2017)).

Una de las principales características cuya aplicación se persigue, tras el desarrollo de estos métodos de *ML*, es el proceso de explicación de la relación existente entre las variables de entrada que participan en el modelado y la estimación de la variable objetivo de salida mediante el **cálculo de las IR**.

A la hora de analizar, por ejemplo, las IR de una RNA, algunos investigadores han utilizado métodos en un esfuerzo por obtener las contribuciones individuales de las variables de entrada de una RNA, como los métodos de análisis de sensibilidad que utilizan la perturbación de entrada o calculan las derivadas parciales (Gevrey *et al.* (2003));(Dimopoulos *et al.* (1995));(Zeng e Yeung (2003));(Wang *et al.* (2000));(Montaño y Palmer (2003)). Por ejemplo, Dimopoulos y Zeng llegaron a la conclusión de que el algoritmo de derivadas parciales y el algoritmo de perturbación de entrada funcionaban relativamente bien (Dimopoulos *et al.* (1995));(Zeng e Yeung (2003)). Maosen en su investigación, señala

20

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

60 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

60 / 181

1.2. Métodos de computación inteligente

el problema de inestabilidad de los cálculos basados en el análisis de sensibilidad, que no han prestado suficiente atención a las limitaciones de este tipo de método, que viene determinado por la estructura entrenada de la RNA (Maosen y Pizhong (2008)).

Otros algoritmos, como el propuesto por Garson, y modificado posteriormente por Goh, denominado algoritmo de Garson modificado (MGA), permiten obtener una aproximación de las importancias para la RNA mediante la realización de cálculos que implican diferentes pesos sinápticos para las capas de entrada, oculta y de salida (Goh (1995)). Sin embargo, los estudios de De Oña indicaron que este método de conexión de pesos para medir el IR con el MGA producía una alta variabilidad cuando se aplicaba a una arquitectura de conjunto de RNA que se había entrenado con diferentes pesos iniciales, como por ejemplo una configuración de ensemble (Oña y Garrido (2014)).

En el caso de aplicación de un análisis de IR para modelos basados en árboles de decisión, en particular, *RF* y *GB*, la literatura existente destacan varios métodos que proporcionan información para comprender qué características conducen hacia la predicción del modelo. El algoritmo clásico de *Feature Importance (FeatImp)* basado en la ganancia y el *split*, introducen un enfoque clásico de la importancia de las características (Breiman (1996)). Siendo la ganancia la reducción total de la pérdida o impureza aportada por todas las escisiones en una determinada característica.

El *split* simplemente tiene en cuenta cuántas veces se utiliza una entrada para realizar cada separación.

La aplicación del algoritmo *FeatImp* es bastante común y su utilización se ha utilizado ampliamente (Chebrolu *et al.* (2005));(Chen y Guestrin (2016));(Huynh-Thu *et al.* (2010));(Sandri y Zuccolotto (2010)). No obstante, el método más reciente y que aporta un enfoque novedoso sobre el proceso de explicación de un aprendizaje supervisado, es el algoritmo *SHapley Additive exPlanations (SHAP)*, ya que en su desarrollo recurrieron a recientes aplicaciones de la teoría de juegos cooperativos para distribuir equitativamente el reparto entre las características de

21

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

61 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

61 / 181

Capítulo 1. Estado del arte

entrada del modelo. Demostrando una mejora en el rendimiento computacional y una mejor coherencia con la intuición humana respecto a enfoques anteriores (Lundberg y Lee (2017));(Lundberg *et al.* (2018)).

Algunos de los resultados de la efectividad de SHAP aplicados a problemáticas reales, se pueden verificar en diferentes estudios recientes (Lubo-Robles *et al.* (2020));(Antwarg *et al.* (2020));(Mckinley *et al.* (2020)). Por ejemplo, Mckinley mediante la aplicación de modelos de GB, predecía las posibles averías en un sensor de emisiones de NOx en vehículos, mediante un análisis exploratorio de datos técnicos del mismo, horas de funcionamiento, datos meteorológicos e histórico de fallos anteriores (Figura 1.10).

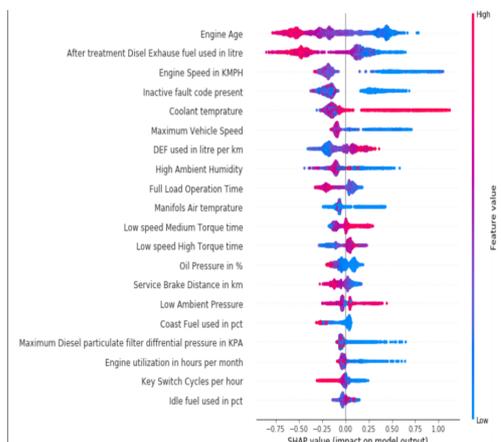


Figura 1.10. Ejemplo para la identificación de importancias SHAP, cuantificando el impacto de cada variable en la salida del modelo (Mckinley *et al.* (2020)).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

1.3. Entorno ambiental de la Macaronesia

1.3. Entorno ambiental de la Macaronesia

La calima es el término local con el que en las Islas Canarias se designa a la arena procedente del Sáhara. La llegada del polvo en suspensión se hace patente por un enturbiamiento del aire, disminuyendo la visibilidad y tomando el cielo un tono rojizo cuando la densidad de polvo es importante. El entrafiamiento del polvo en sus áreas de origen, localizadas en el interior del desierto del Sáhara, se puede producir tanto por el alisio como fenómenos convectivos. En ambos casos pueden arribar al archipiélago canario con vientos del 2º cuadrante (del este al sur). Debido a la cercanía al continente africano los episodios de calima son más frecuentes en Lanzarote, Fuerteventura y Gran Canaria.



Figura 1.11. Efecto de la concentración de arena sahariana en las Islas Canarias.

Parte del polvo sahariano que llega a Canarias se asienta por deposición seca, siendo especialmente llamativa la deposición húmeda o "lluvia de sangre", aunque no es un fenómeno dominante en Canarias (Torres-Padrón *et al.* (2002));(Friese *et al.* (2016));(Criado y Dorta (2003)). Las lluvias después de los eventos invernales pueden eliminar una cantidad significativa de polvo del aire. El aporte mineral sahariano en las islas reviste especial importancia, encontrándose minerales saharianos (tales como el cuarzo y la mica) en suelos y sedimentos formados a lo largo del Cuaternario; aportes de fosfatos y hierro parecen estar detrás de la fertilidad de los suelos y mares del archipiélago (Criado *et al.* (2012));(Dor-

23

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

63 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

63 / 181

Capítulo 1. Estado del arte

ta *et al.* (2002));(Pye (1987)); (Dorta *et al.* (2005)). Los eventos de calima causan graves daños a la economía, al afectar a la agricultura y a la salud de la población de las Islas Canarias (Dorta *et al.* (2002));(Torres-Padrón *et al.* (2002));(Middleton (2017)). Estos eventos casi siempre van acompañados de altas temperaturas y una importante disminución de la humedad relativa, lo que convierte a estos eventos, sobre todo en verano, en momentos de alto riesgo de incendios forestales [10]. Además, la presencia de polvo, que alcanza una media de 127,30 ($\mu\text{gr}/\text{m}^3$) frente a los 27,76 ($\mu\text{gr}/\text{m}^3$) con aire marino, reduce notablemente la visibilidad y provoca problemas respiratorios, especialmente en los grupos de riesgo como los niños pequeños, los ancianos y las personas con problemas respiratorios (Dorta *et al.* (2005));(García *et al.* (2001)).

La cercanía del desierto del Sáhara implica que Canarias se vea afectada, con frecuencia, por las masas de aire cálidas y muy secas que se sitúan sobre ese inmenso territorio. Siendo más frecuente en invierno, desde diciembre hasta febrero y menos desde mediados de primavera hasta principios del verano, especialmente entre abril y junio (Figura 1.12). Totalizando un 65 % del total de invasiones, frente a un 35 % en primavera y verano (Dorta *et al.* (2005)).

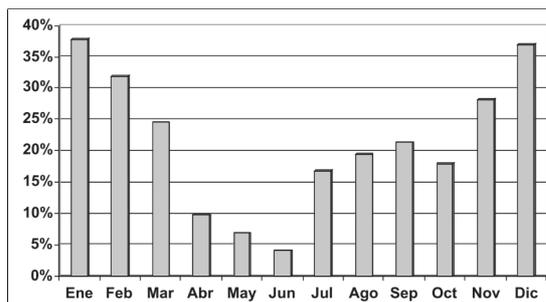


Figura 1.12. Frecuencia mensual de los días con predominio de aire sahariano en Canarias (1976-2003) (Dorta *et al.* (2005)).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA Fecha: 20/01/2021 19:43:20
 Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA 20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA Fecha: 26/01/2021 17:45:11
 María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA 27/05/2021 13:54:44

1.4. Sistemas de generación eléctrica convencional

Otros estudios indican que las mayores entradas de aerosol mineral se han registrado durante los meses de invierno y en particular durante enero-febrero (Figura 1.13). Este patrón diferencia los registros en Gran Canaria de otras series largas de aerosoles minerales en la región Atlántica Norte, donde los máximos en concentración de polvo sahariano se producen durante finales de la primavera y al principio y mitad del verano. El tamaño de partícula que domina el aerosol está comprendido entre 0.6 a 4.9 μm y su velocidad de deposición estimada a partir del flujo de deposición seca es de $1 \text{ cm} \cdot \text{s}^{-1}$ (Gelado *et al.* (2002)).

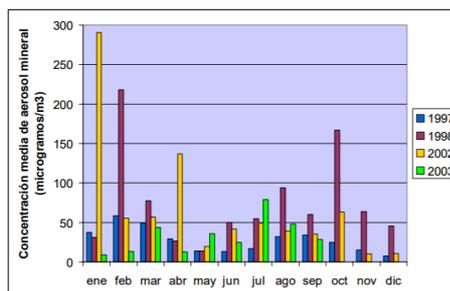


Figura 1.13. Variabilidad mensual de las concentraciones de aerosol mineral ($\mu\text{gr}/\text{m}^3$) en el aire medidas en Gran Canaria (Gelado *et al.* (2002)).

1.4. Sistemas de generación eléctrica convencional

Las centrales eléctricas convencionales y/o de ciclo combinado de un sistema aislado abarcan sistemas complejos donde el rendimiento térmico y el control de emisiones son dos aspectos fundamentales a la hora de supervisar la generación eléctrica insular.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 1. Estado del arte

Un conjunto típico es una **turbina de gas** para generar electricidad, que como se aprecia en la Figura 1.14, consiste básicamente en una sección generadora de gases y una sección de conversión de energía. La sección generadora de gases consiste en un compresor, una cámara de combustión y una turbina, que extrae suficiente energía para accionar el compresor. El resultado de esto es un gas a alta temperatura y presión a la salida de la turbina.

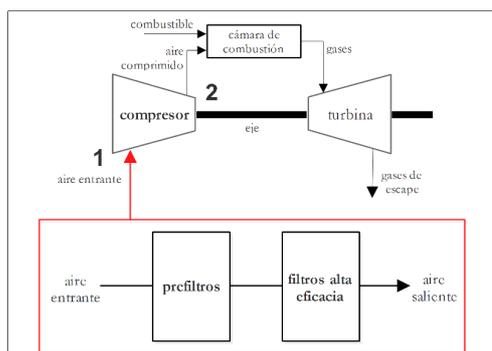


Figura 1.14. Diagrama general de una turbina de gas para generación de electricidad con un sistema filtrante multi etapa.

La compresión de grandes volúmenes de aire es esencial para el buen funcionamiento de una turbina de gas. La trayectoria del flujo en un compresor de flujo axial es paralelo al eje de giro. El compresor de flujo axial consume entre 55 % y 65 % de la potencia producida por la sección de turbina en la turbina de gas [1].

Para adaptar las turbinas de gas de una variedad de entornos y aprovechar al máximo su rendimiento y confiabilidad, es necesario tratar el aire que consumen (Figura 1). Incluso en entornos relativamente limpios, una turbina de gas puede ingerir gran cantidad de partículas cada año. Que esto cause o no un problema depende de la cantidad de estas partículas, sus propiedades mecánicas y su com-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

1.4. Sistemas de generación eléctrica convencional

posición química (Syverud *et al.* (2007)). El ensuciamiento de los compresores está provocado principalmente por sales y otros depósitos minerales contenidos en el aire entrante (Meher-Homji y Bromley (2004)).

Cualquier turbina de gas, debido a su diseño y la gran cantidad de aire consumido, es sensible a la calidad del aire. La filtración se aplica para proporcionar protección contra los efectos del aire contaminado que puede degradar el rendimiento y la vida de la turbina de gas: erosión, incrustaciones, corrosión y obstrucción del conducto de refrigeración. Por lo tanto, la necesidad de una filtración adecuada es vital para esta tecnología (Santini *et al.* (2013));(Kurz y Brun (2012)).

La estructura convencional de un bloque de filtración en la entrada del compresor de una turbina de gas se compone de etapa de un sistema de prefiltros y un bloque de filtración con un sistema de filtros de alta eficacia (Santini *et al.* (2013)). Dependiendo del tipo de sistema de filtración utilizado, pequeñas partículas en el rango de 0,1 - 2,0 μm podrían entrar en la turbina (Kurz y Brun (2012)). Estas partículas son demasiado pequeñas para causar problemas de erosión, pero son adecuadas para adherirse a la superficie de los álabes y causar incrustaciones (Figura 1.15). El estudio de compresores sucios ha revelado contaminación tanto en el lado de succión como en el lado de presión de los álabes del compresor. La detección de ensuciamiento en campo ha revelado que solo las primeras etapas se ven afectadas por los depósitos y que la ausencia de fuerza centrífuga en los álabes de la parte estática, conducen a depósitos más severos en comparación con el rotor de la turbina de gas (Syverud *et al.* (2007));(Tarabrin *et al.* (2014)). El estudio de Aker reveló que el ensuciamiento puede progresar alrededor del 40 y 50 % de las etapas del compresor (Aker y Saravanamuttoo (1989)).

El ensuciamiento o *fouling* que, de forma progresiva y exponencial, irá creciendo en función de las horas de funcionamiento de la turbina, se irá depositando en las primeras etapas de los álabes del compresor, reduciendo parámetros como el caudal de aire de entrada y su presión de descarga (Giampaolo (2003));(Tarabrin *et al.* (2015)). En contrapartida, observaremos un aumento en el consumo de

27

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

67 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

67 / 181

Capítulo 1. Estado del arte

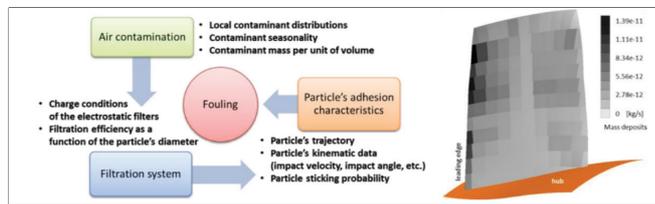


Figura 1.15. Diagrama estándar de la formación del fenómeno *fouling* en álabes de compresor (Aldi *et al.* (2015)).

combustible y en definitiva en el consumo específico de la turbina de gas (Meher-Homji y Bromley (2004)). Las pérdidas derivadas del *fouling* del compresor se catalogan como recuperables, ya que pueden ser erradicadas con la realización de mantenimientos de lavado (Figura 1.16). Existe un porcentaje de pérdidas que, en función de las horas de funcionamiento, se convierten en “no recuperables” y sólo pueden solventarse en una parada larga y/o sustitución de piezas por mantenimiento (Jiang *et al.* (2016));(Gayraud y Singh (1997)). Se estima y se achaca que, entre un 70-85 % de las pérdidas recuperables de una turbina de gas, están ligadas al ensuciamiento del compresor (Janawitz *et al.* (2010)).

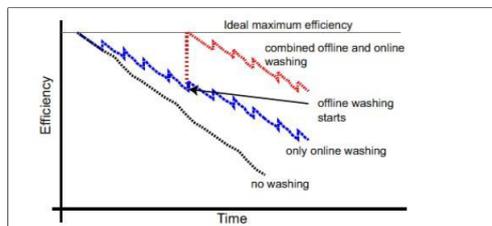


Figura 1.16. Tendencia cualitativa de las variaciones de eficiencia en un compresor, con diferentes tipos de lavados (Dionysios *et al.* (2016)).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

1.4. Sistemas de generación eléctrica convencional

La principal solución es la de realizar una serie de lavados *offline* y cambios en los elementos filtrantes de entrada que puedan normalizar los parámetros térmicos, a condiciones de inicio. Las horas de funcionamiento son proporcionales a este factor de ensuciamiento, y alcanzan su valor crítico, en torno a las 2000-2500 horas (Stalder (2001));(Dionysios *et al.* (2016)). Según Tarabrin *et al.* , el ensuciamiento tiende a ocurrir durante la operación inicial y sigue aproximadamente una fase exponencial, estabilizándose después de 1000 a 2000 horas de funcionamiento (Tarabrin *et al.* (2014)). Tarabrin también observó la siguiente relación por ensuciamiento del compresor. Un descenso de su presión de descarga, en torno a un 5.5 %, suponía una pérdida de producción en potencia eléctrica de salida de la turbina de gas en un 13 % y un aumento consumo específico del 6 % (Figura 1.17).

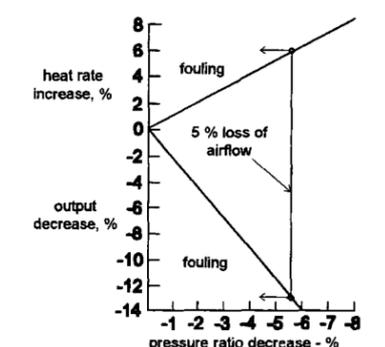


Figura 1.17. Curvas de degradación por *fouling* en compresor de Turbina de Gas (Tarabrin *et al.* (2014)).

En el estudio realizado por Inamul *et al.* , la caída en la temperatura y la eficiencia obtenidas, fueron del orden de 0.8 % y 0.5 % respectivamente sobre las horas de funcionamiento (1290h) (Inamul y Saravanamuttoo (2015)). En otro caso, el ensuciamiento que disminuye el flujo de aire de entrada en un 5 % cau-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 1. Estado del arte

sará aproximadamente un 2.5 % de disminución en la eficiencia del compresor. Esto dará como resultado de 8 a 10 % disminución de la potencia de salida y una disminución del 4 % en la eficiencia térmica (Diakunchak (2015)).

Otras investigaciones detallan y modelan este problema de ensuciamiento del compresor (Hanachi *et al.* (2017));(Cruz-Manzo *et al.* (2017));(Malekshan *et al.* (2017)), donde por ejemplo Mohammadi *et al.* , simularon la variación del ensuciamiento del compresor en función de diferentes escenarios de producción de potencia eléctrica, tanto a plena carga como parcial, llegando a la conclusión de que el ensuciamiento no era el mismo y por lo tanto variaba para cada situación de producción eléctrica de la turbina de gas (Mohammadi y Montazeri-Gh (2015)).

Otro conjunto típico de las centrales térmicas convencionales son los **motores diesel** de combustión interna con un consumo de combustible líquido similar al esquema planteado en la Figura 1.18. Un componente crucial en el funcionamiento de este tipo de plantas, donde la monitorización continua sobre los niveles de emisiones contaminantes de su producción es vital para el cumplimiento legal dentro de la regulación ambiental vigente.

El control de **emisiones ambientales** y la supervisión de parámetros térmicos en motores diesel fue abordado en varios trabajos de investigación. Raptotasio, aplicó con éxito un modelo de reducción de NOx en motores diésel de gran potencia mediante la recirculación de gases de escape, proporcionando resultados favorables. Con una evaluación de un modelo matemático basada en datos experimentales adquiridos en la literatura internacional, debido a la falta de datos experimentales para motores de dos tiempos (Raptotasio *et al.* (2015)).

El análisis de los resultados revela la capacidad de este modelo para predecir tanto el rendimiento del motor como las emisiones de NOx, pero lo más importante es la capacidad de predecir correctamente el efecto general de la recirculación de gases sobre las emisiones de NOx.

30

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

70 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

70 / 181

1.4. Sistemas de generación eléctrica convencional

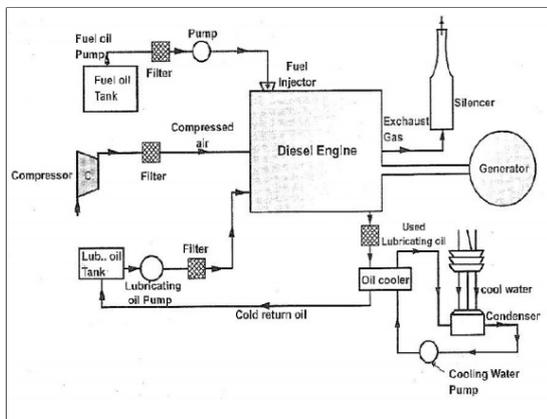


Figura 1.18. Adaptación del esquema general de un motor diesel para generación eléctrica.

Siguiendo el proceso de evolución digital, Vorkapic se acerca hacia un modelo de *ML* comparando varios algoritmos de computación como máquinas de soporte vectorial o *RF*, para monitorizar y supervisar un motor diesel marino de dos tiempos, estableciendo un modelo capaz de controlar parámetros operativos influyentes y la predicción del consumo de combustible (Vorkapic *et al.* (2020)).

Otra aproximación al modelado de este tipo de motores fue la realizada por Niu, añadiendo esta vez un modelo basado en RNA, para investigar la predicción de la respuesta de un motor diesel marino. Su estudio sugería que se deben realizar más investigaciones sobre la validación y aplicación de las máquinas de soporte vectorial sobre el rendimiento de los motores diesel y la predicción de los niveles de emisiones. Sin embargo, la estabilidad y la precisión de la RNA pueden mejorarse, optimizando los pesos iniciales y los bias (Niu *et al.* (2017)).

De la misma manera, Kowalski en su investigación, profundizó en la aplicación

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 1. Estado del arte

de RNA sobre la estimación de NOx en motores diesel marinos de dos tiempos. Presentando un procedimiento para estimar las emisiones ambientales del motor, sustituyendo la medición real de NOx por la obtenida en el modelo neuronal, que utilizaba los datos de entrada de quince parámetros de funcionamiento del motor, que influían en la emisión de NOx.

Tras probar varias configuraciones de arquitectura de la RNA y varias iteraciones en los pesos iniciales, consiguieron estimar con una precisión equiparable a la indicada por el marco regulatorio (Kowalski *et al.* (2017)).

32

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

72 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

72 / 181

Capítulo

2

Análisis multivariante mediante ensemble de redes neuronales artificiales

Como punto de partida para la etapa de desarrollo **MBD.0** y siguiendo la línea general de trabajo propuesta en la estructura de la Figura 1. Se realiza el análisis y la predicción de la concentración de polvo de origen sahariano en el aire, donde se define el procedimiento de aprendizaje para un conjunto de redes neuronales artificiales, estructuradas en configuración de unión o *ensemble*. De forma que una variable compleja como la calima se obtenga, a partir de datos reales medidos como la temperatura del aire, la humedad relativa, la presión atmosférica y la velocidad del viento.

Para poder profundizar en el conocimiento obtenido a través de una RNA y por ende abrir su *caja negra*, se diseña y analiza un nuevo sistema híbrido para el cálculo de IR en este tipo de métodos de computación inteligente, tanto para una arquitectura simple como para una de ensemble.

2.1. Estructura de los datos y arquitectura

El conjunto de datos se divide en dos bloques principales (Figura 2.1). El primero consiste en un conjunto de datos para la fase de entrenamiento (DS1) que contiene los datos de entrada y salida de julio de 2016 a julio de 2018, con una

33

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

73 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

73 / 181

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales



Figura 2.1. Distribución del conjunto de datos para la fase de entrenamiento y test

frecuencia **minutal** para los datos de entrada del entorno ambiental.

Los valores reales de la salida se corresponden a la concentración de polvo sahariano en la superficie de la isla de Tenerife, obtenida a partir del modelo BSC-DREAM8b v2.0 (Figura 2.2), utilizando interpolación lineal para adaptar sus intervalos de seis horas a la frecuencia de muestreo de un minuto utilizada en los datos de entrada.

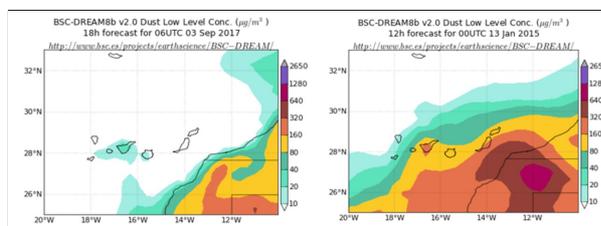


Figura 2.2. Modelo de detección de arena BSC-DREAM8b para Canarias, operado por el Centro de Supercomputación de Barcelona. <http://www.bsc.es/ESS/bsc-dust-daily-forecast>

El segundo conjunto de datos para la fase de test (DS2), independientes y que no se utilizarán para entrenar la RNA, pertenecen al periodo de julio de 2018 a diciembre de 2018 y también contiene datos de entrada y salida. Está destinado exclusivamente a la fase de test y verificación de resultados. Esto es, estimar la concentración de polvo utilizando sólo los datos de entrada ambientales de DS2

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

2.1. Estructura de los datos y arquitectura

y comparando los resultados estadísticamente, para ver si se ajustan a las salidas reales de este periodo.

Como parte de la **preparación de los datos** para la fase de entrenamiento RNA, y para facilitar el aprendizaje del algoritmo de entrenamiento, se suavizaron los datos de las series temporales (Figura 2.3), procesando una media móvil simple (SMA) con un periodo de *tres días*. La SMA es la media no ponderada de los S puntos de referencia anteriores (Chou (1975)).

$$SMA = \frac{1}{S} \sum_{i=0}^{S-1} A_{-i} \quad (2.1)$$

Donde A representa el vector variable compuesto de S observaciones escalares.

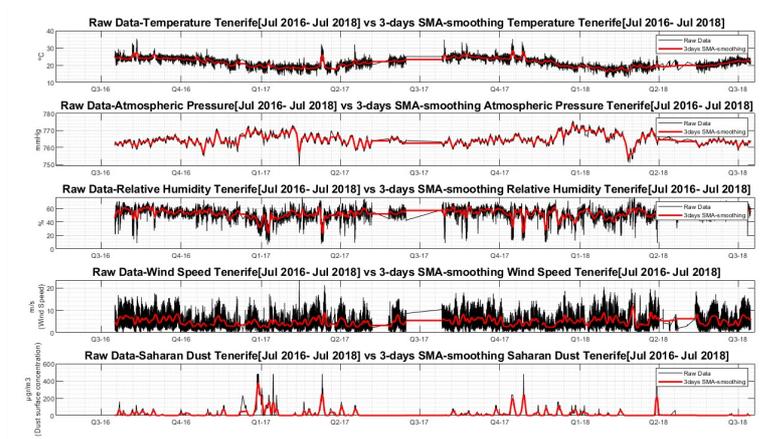


Figura 2.3. Datos originales vs 5-días SMA; Enero, Febrero y Marzo (Q1); Abril, Mayo y Junio (Q2); Julio, Agosto y Septiembre (Q3); Octubre, Noviembre y Diciembre (Q4)

35

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

75 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpFLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

Se han extraído además, datos de dirección y velocidad del viento, y se muestran gráficamente en una rosa de los vientos (Figura 2.4), donde se aprecian las direcciones predominantes del primer y segundo cuadrante características de las Islas Canarias (Criado *et al.* (2012)). Sin embargo, no se ha tenido en cuenta la variable dirección de viento para la construcción del modelo computacional de este capítulo, dada su variabilidad y complejidad de normalización.

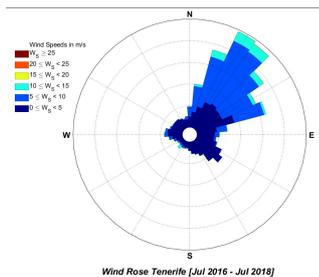


Figura 2.4. Rosa de los vientos con DS1 en el sur de la isla de Tenerife

La estructura principal de este desarrollo, se construye a partir de una RNA. Siendo el tipo de red neuronal que se presenta, una **FFBP-NN**. Con una **arquitectura** básica de cinco entradas (temperatura ambiente; presión atmosférica; velocidad del viento; humedad relativa para cada mes), una única capa oculta con nueve nodos con una función de activación sigmoide; que específicamente utiliza la tangente hiperbólica. Para la capa de salida se implementa una función de activación lineal sobre una variable (concentración de polvo sahariano) (Figura 2.5).

El número de **nodos de la capa oculta** se ha ajustado adaptando de forma iterativa el método propuesto por Huang, de forma que, en cada iteración al finalizar el entrenamiento del modelo, se verifica que su rendimiento estadístico se encuentre dentro de los límites definidos por la adaptación propia en el cálculo de límites por control estadístico (Huang (2003)).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

2.1. Estructura de los datos y arquitectura

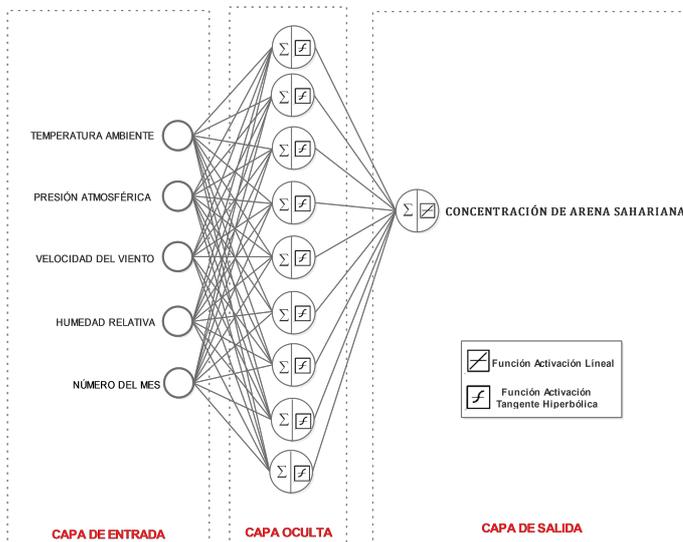


Figura 2.5. Topología de la estructura principal de RNA base, con cinco entradas, nueve nodos en una capa oculta, con función de activación tangente hiperbólica y un nodo en la capa de salida con una función de activación lineal.

El método de Huang para el cálculo de neuronas en la capa oculta, relaciona el número de entradas con las variables de salida utilizando modelos neurales artificiales en una capa oculta $L1$ (2.2) (Huang (2003)).

$$L1 = \sqrt{(N + 2) \cdot m} + 2 \cdot \sqrt{\frac{m}{(N + 2)}} \quad (2.2)$$

Siendo N el número de neuronas de salida y m el número de nodos de entrada de la RNA diseñada.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

Los índices de rendimiento estadístico utilizados a lo largo de este trabajo de investigación son, el **coeficiente de determinación** (R^2) y el **error medio absoluto (MAE)** (2.3). Según Willmott, el MAE ofrece una medida natural del error medio, a diferencia del error medio cuadrático (RMSE) (Willmott y Matsuura (2005)).

$$MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n} \quad (2.3)$$

Donde n es el número total de puntos de muestreo de datos, \hat{y}_i es la predicción e y_i el valor real.

Por otro lado, el coeficiente de determinación el R^2 se toma como medida de la validez de un modelo de regresión o de una estimación de regresión (2.4).

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.4)$$

Siendo n el número total de puntos de la muestra de datos, \hat{y}_i es el valor estimado de la variable dependiente para la observación i , tal y como se calcula con la ecuación de regresión, y_i es el valor observado de la variable dependiente para la observación i e \bar{y} es la media de todas las n observaciones de la variable dependiente.

Cuanto más se aproximen los valores estimados del modelo de regresión \hat{y}_i a los valores observados y_i , más se aproximará el coeficiente de determinación a 1 y más preciso será el modelo de regresión (Schneider *et al.* (2010)).

El **método de aprendizaje** aplicado es de tipo supervisado. Además, para mejorar la eficiencia del perceptrón multicapa, se elige un algoritmo de propagación del error hacia atrás FFBP-NN, con una inicialización adaptativa de los pesos, determinada mediante el criterio de *Nguyen-Widrow*, para reducir tiempos en la fase de entrenamiento (Nguyen y Widrow (1990)). Específicamente, se ha utilizado el método de optimización de *Levenberg-Marquardt*, modificado mediante la Regularización Bayesiana, diseñado para minimizar las funciones de error de la suma

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

2.1. Estructura de los datos y arquitectura

de cuadrados de sistemas no lineales (Sapna (2012));(Cervantes *et al.* (2017)).

La Regularización Bayesiana realiza una ligera modificación sobre el método de optimización, incorporando el teorema de Bayes en el algoritmo y compensando el posible sobre-entrenamiento de la RNA (Burden y Winkler (2009)).

La **división del conjunto de datos** se realiza a partir de la técnica de remuestreo o *resampling*, en la VC de k -iteraciones o k -fold, particionando el conjunto o *set* de datos de entrenamiento en k subconjuntos de la mismas dimensiones. De tal forma que la partición final de entrenamiento se construye utilizando $k-1$ subconjuntos y el restante queda definido como el set de validación o test. Donde el 90% de cada subconjunto se utilizará en la fase de entrenamiento y el 10% restante para la fase de test. De esta forma todo el conjunto de datos queda representado (Figura 2.6) (Berrar *et al.* (2006)).

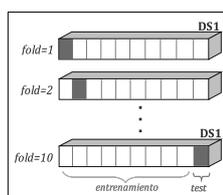


Figura 2.6. Validación cruzada(10-fold) para DS1.

En este trabajo de investigación, **se establece un valor estándar $k=10$** , ya que según Kohavi, es la partición óptima al aplicar técnicas de VC sobre un modelo (Kohavi (1995)). Otros estudios señalan además que las simulaciones realizadas con este valor de k , obtienen reiteradamente el mejor rendimiento estadístico (Borra y Ciaccio (2010));(Llerena *et al.* (2012)).

39

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

79 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

79 / 181

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

2.2. Procedimientos

Siguiendo la línea de trabajo para MBD, en este desarrollo se ha iniciado la primera aproximación MBD.0 en el planteamiento y propuesta de una solución computacional aplicada a un sistema de aprendizaje de máquina o *ML* para el modelado de una variable ambiental en la isla de Tenerife, tomando como referencia de entrada, otro conjunto de datos medidos en el mismo entorno (Figura 2.7).

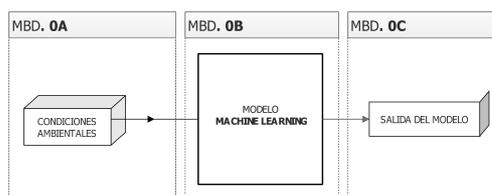


Figura 2.7. Esquema de aproximación MBD.0 en la línea de desarrollo para el mantenimiento basado en datos.

El **procedimiento general** de cálculo de este capítulo se muestra en la Figura 2.8, donde se indica la secuencia computacional utilizada para estimar la concentración de polvo sahariano dentro de una arquitectura de unión o *ensemble* para un conjunto de RNA. Siendo el método de *ensemble* utilizado el *CVC*, una estructura similar de *ensemble* a la utilizada por Parmanto, Xia, Dong y Llerena, que particiona los datos DS1 según la Sección 2.1, y resultando en 10x subconjuntos de las mismas dimensiones cada uno (Parmanto *et al.* (1996));(Xia *et al.* (2011));(Dong y Han (2004));(Llerena *et al.* (2012)).

Cada conjunto del *fold*, será utilizado para construir una FFBP-NN_x, donde además, los pesos sinápticos iniciales serán diferentes por cada partición de *k*. Obteniendo para nuestro sistema, un total de 10x FFBP-NN, de las cuales, el valor final del *ensemble* será la media de los resultados obtenidos con cada una (Figura 2.9) y donde obtendremos un MAE_x y un R²_x preliminar .

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

2.2. Procedimientos

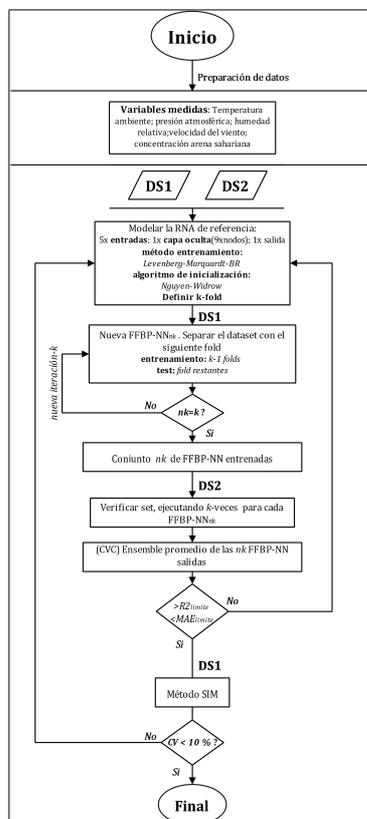


Figura 2.8. Flujograma del procedimiento computacional realizado en de este capítulo.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

Para la etapa **validación final** de los modelos, el ciclo de iteración de los entrenamientos finaliza cuando sobre el conjunto de datos DS2, se alcanza un MAE medio final inferior a un límite de control, resultante de una adaptación del método de control de calidad estadístico, *c-chart* y un valor de R^2 que presente un alto grado de correlación (Montgomery (2008)).

$$R^2_{límite} > 0.7 \tag{2.5}$$

$$MAE_{límite} = \bar{c} + 3\sqrt{\bar{c}} \tag{2.6}$$

Donde \bar{c} representa el valor medio de DS1, sobre la variable objetivo de salida.

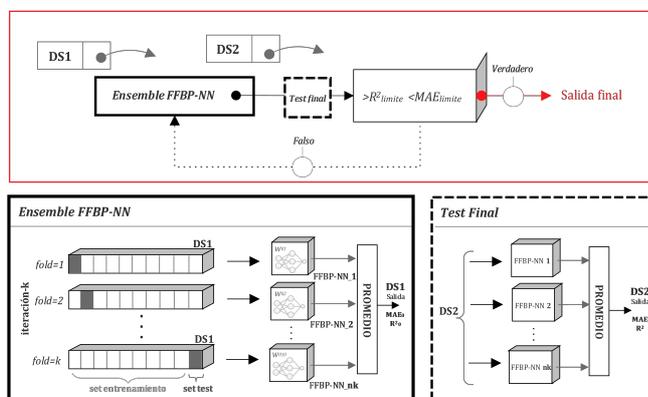


Figura 2.9. Diagrama principal del algoritmo de entrenamiento utilizado, ampliado en los bloques de aprendizaje ensemble mediante partición *k*-fold y validación final.

Estos resultados finales serán evaluados utilizando el valor medio de lo obtenido, combinando DS2 y cada FFBP-NN_x entrenada en el paso anterior con DS1, hasta conseguir que el rendimiento estadístico final supere lo establecido en (2.5)

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

2.2. Procedimientos

y (2.6), en términos de MAE y R^2 . Si no se supera el límite establecido, el modelo vuelve a reajustar los pesos iniciales y los nodos de la capa oculta de cada FFBP- NN_x (Figura 2.9). Si incluso después de este paso no se superan los límites, se debe considerar revisar la normalización del conjunto de datos o añadir una nueva variable que pueda ayudar a mejorar la fase de aprendizaje.

El último bloque de este procedimiento, incluye el estudio de métodos para medir el grado de IR de las diferentes de variables que conforman el modelo de nuestro sistema ambiental. Se ha diseñado un algoritmo específico en este trabajo, *SIM*, el cual representa un método de estudio y medida directa en el cálculo de la IR, tras el aprendizaje realizado por una sola RNA, o un ensemble de éstas (Algoritmo 2.1), que combina la alteración de las entradas y el nuevo entrenamiento de la para ajustar la nueva arquitectura, y un método clásico basado en los pesos de una RNA.

Tras haber superado previamente la fase de evaluación con nuevos datos DS2, sabemos que disponemos de un modelo que ha aprendido correctamente la dinámica del sistema. Por lo tanto, para esta nueva fase únicamente queremos comprobar qué variables aportan mayor o menos importancia al modelo. Y nos interesa verificarlo con el conjunto de datos que mayor rendimiento debe obtener, DS1.

$$nm(MAE_x) = \frac{\sum_{nk=1}^k (MAE_{(m-1)nk})}{k} \quad (2.7)$$

$$nm(MAE_0) = \frac{\sum_{nk=1}^k (MAE_{0nk})}{k} \quad (2.8)$$

Donde x indica la variable de entrada que se evalúa en base al promedio MAE_x , obtenido del conjunto de todos los k -fold, durante el nuevo entrenamiento del

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

FFBP-NN_{nk} con DS1. Las iteraciones (*k*) se llevan a cabo con el nuevo entrenamiento después de eliminar una variable de entrada con cada ciclo. Esto significa que para cada iteración realizada en cada FFBP-NN_{nk}, la nueva configuración de entrenamiento siempre tendrá *m*-1 entradas, siendo *m* el número total de entradas utilizadas en el sistema que se está modelando (2.7).

Finalmente, para obtener el valor IR final de cada variable de entrada (2.9), el nuevo entrenamiento $nm(MAE_x)$, se compara con el rendimiento promedio, resultante del modelo final del conjunto original con DS1 (2.8).

$$SIM = nm(MAE_x) - nm(MAE_0) \quad (2.9)$$

El proceso selectivo de eliminar variables de entrada implica, suprimirlas de forma individual junto con su conjunto de pesos, para calcular así el nuevo MAE obtenido en los nuevos entrenamientos, del ensemble $nm(MAE_x)$, de forma que se vaya analizando cuál es el resultado de alterar el modelo (Figura 2.10).

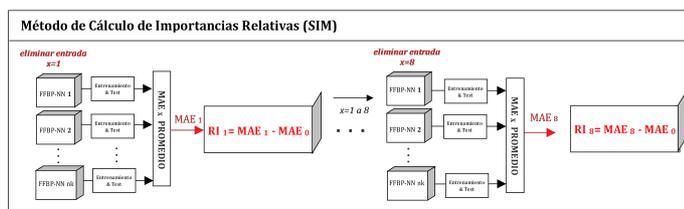


Figura 2.10. Diagrama principal del procedimiento SIM.

Para cada una de las 10x FFBP-NN_x finales entrenadas del conjunto de ensemble con particionamiento *k*-fold, se irán guardando todos sus pesos sinápticos de entrada, salida y bias iniciales, w_{i0xh} , w_{o0xh} , b_{i0xh} y b_{o0xh} , respectivamente. Siendo *h* la neurona de la capa oculta correspondiente, *i* la capa de entrada, *o* la capa de salida y el valor 0 indica su estado inicial. Estas variables de pesos y bias guardados, serán utilizados en la configuración inicial de los entrenamientos del nuevo con-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

2.2. Procedimientos

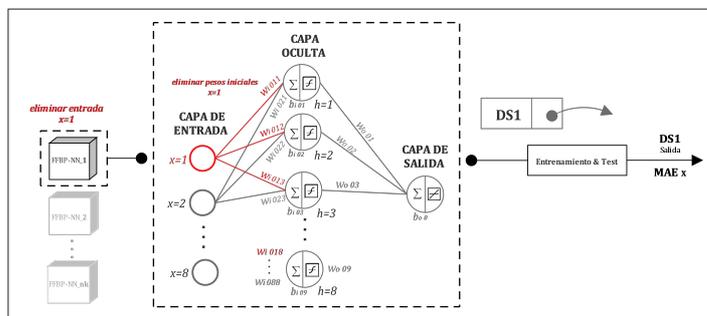


Figura 2.11. Diagrama ampliado del procedimiento SIM para eliminar una variable de entrada y sus pesos sinápticos hacia la capa oculta.

junto ensemble, $nn(MAE_x)$ y tras haber eliminado los pesos iniciales de entrada w_{i0x} que corresponden a la variable x que hemos quitado (Figura 2.11).

Para que este procedimiento sea válido, se establece una **condición final de control** que verifique de alguna forma el grado de dispersión de los resultados (2.10). Para ello se ha establecido un valor límite de coeficiente de variación (CV) del 10% sobre el conjunto final obtenido. Un valor por debajo del cual se considera que los resultados tendrán una baja variabilidad (Braz *et al.* (2017));(Gomes (1985)).

La expresión del CV se define como:

$$CV(\%) = \frac{\sigma_{nn(MAE_x)}}{\mu_{nn(MAE_x)}} \quad (2.10)$$

Donde μ indica la media del conjunto del ensemble $nn(MAE_x)$, obtenido por cada entrada y σ la desviación estándar media del mismo conjunto. Si este límite se ve superado en algunas de las variables de entrada y el grado de variación que se obtiene es alto, nos puede indicar que esa variable en particular no realiza

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

un aporte relevante al conjunto del modelo. Permittiéndonos incluso, descartarla, filtrarla del entrenamiento o revisar la normalización de su conjunto de datos inicial.

El pseudocódigo del procedimiento general *SIM* (2.1), desarrollado para el cálculo de IR en RNA, se muestra a continuación:

Algoritmo 2.1. Selective Importance Measure (SIM)

```

1: for FFBP-NNnk do
2:   if modelo final FFBP-NNnk (DS2) < MAElímite y R2límite then
3:     Seleccionar FFBP-NNnk (DS1).
4:     Almacenar variables iniciales  $w_{i_{0xh}}, w_{o_{0xh}}, b_{i_{0xh}}$  y  $b_{o_{0h}}$  de FFBP-NNnk
5:     for todas las  $m$  entradas de FFBP-NNnk do
6:       Eliminar la entrada  $x$  de FFBP-NNnk y los pesos asociados  $w_{i_{0xh}}$ 
7:       Preparar el entrenamiento de FFBP-NNnk con los nuevos  $w_{i_{0xh}}, w_{o_{0xh}}, b_{i_{0xh}}$  y  $b_{o_{0h}}$  de la etapa 6
8:       Iniciar el nuevo entrenamiento de FFBP-NNnk
9:       Almacenar el resultado en MAEx
10:      Cargar pesos iniciales  $w_{i_{0xh}}, w_{o_{0xh}}, b_{i_{0xh}}$  y  $b_{o_{0h}}$  de la etapa 4
11:     end for
12:      $n \leftarrow n + 1$ 
13:   end if
14: end for
15: if  $n > 0$  then
16:   Calcular el CV para las  $m$  entradas
17:   if CV < 10 then
18:     return SIM
19:   end if
20: end if
    
```

Por último, como medida de contraste de este método propio de cálculo de IR, específico para RNA en configuración simple o de ensemble, se aplica sobre los resultados finales, el MGA (2.11).

Este método está basado directamente sobre los pesos sinápticos de una RNA y deriva del primer algoritmo original propuesto por Garson y que más tar-

46

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

86 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

86 / 181

2.3. Resultados

de fue modificado por Goh, se representa como sigue a continuación (Garson (1991));(Goh (1995)):

$$MGA = \sum_{h=i}^j \frac{|w_{ih} \cdot w_{ho}|}{\sum_{i=1}^m |w_{ih} \cdot w_{ho}|} \quad (2.11)$$

Donde m y j indican el número de entradas y nodos de la capa oculta, respectivamente, w_{ih} son los pesos sinápticos entre la entrada i y la neurona de la capa oculta h y w_{ho} son los pesos entre la capa oculta h y la salida o .

2.3. Resultados

En total, para este caso se analizaron selectivamente un total de 50x FFBP-NN; es decir, 10x FFBP-NN por cada variable de entrada x . Esto significa que cada nodo de entrada genera un conjunto de 10 resultados en términos del MAE.

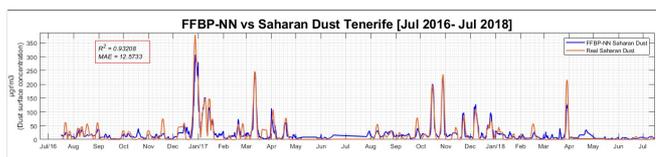


Figura 2.12. Resultados de la predicción de la concentración de arena en la isla de Tenerife y comparación estadística del modelo final con DS1. Julio 2016 a Julio 2018.

Para la primera predicción del valor de la concentración de arena con datos de DS1, donde su conjunto de datos sí participo en el entrenamiento de la FFBP-NN, los índices de rendimiento estadístico obtenidos fueron $R^2 \approx 0.93$ y un error absoluto medio de $MAE \approx 12.57 \mu gr / m^3$ (Figura 2.12).

En el caso de la estimación con datos de DS2, cuya salida se desconocía y cuyos datos de entrada no se utilizaron para entrenar al FFBP-NN, los índices de

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguilera 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

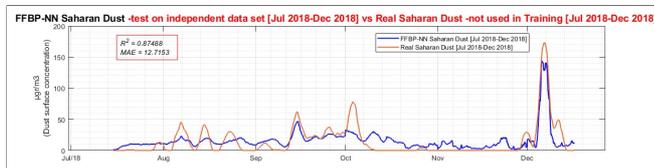


Figura 2.13. Resultados de la predicción de la concentración de polvo sahariano en la isla de Tenerife y comparación estadística del modelo final con DS2. Julio 2018 a Diciembre 2018.

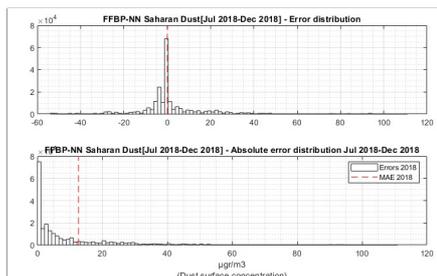


Figura 2.14. Histograma con la distribución del error del modelo final con DS2.

rendimiento estadístico obtenidos fueron $R^2 \approx 0.87$ y un error absoluto medio de $MAE \approx 12.71 \mu gr/m^3$ (Figura 2.13). La distribución del error se dispersa en torno a valores cercanos al cero, pero su error medio es bajo (Figura 2.14).

Habiendo completado con éxito la fase de evaluación con los nuevos datos de DS2, sabemos que tenemos un modelo que ha aprendido correctamente la dinámica del sistema. Por lo tanto, para esta nueva fase se comprueba la importancia de cada variable para el sistema, estudiando el conjunto de datos de mejor rendimiento, DS1.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar China UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

2.3. Resultados

A continuación presentamos los resultados de cada uno de los métodos descritos en la Sección 2.3, que miden el nivel de IR en términos de porcentaje; es decir, el valor de IR que cada variable de entrada del conjunto representa con respecto al conjunto total (Figura 2.15).

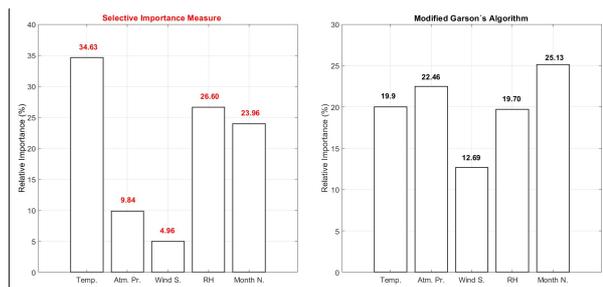


Figura 2.15. Grado de importancia relativa obtenida para los métodos SIM y MGA; Temperatura (Temp.); Presión Atmosférica (Atm.Pr); Velocidad del viento (Wind S.); Humedad relativa (RH); Número del mes (Month N.).

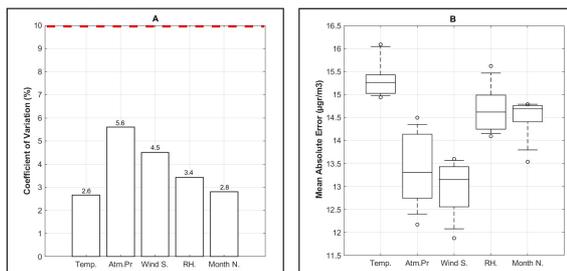


Figura 2.16. Coeficientes de Variación (CV) obtenidos para cada variable de entrada del procedimiento SIM (A) y diagrama de cajas representando la dispersión del MAE para cada entrada, k veces con el método SIM (B).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

Los principales resultados de los cálculos de la IR media, clasificados por relevancia, y las medidas de control del CV de nuestros modelos, se muestran en la Tabla 2.1, donde se comparan los resultados del método propio *SIM* con el MGA (Sección 2.4), basado directamente en los pesos de una RNA.

Tabla 2.1. Resultados y dispersión obtenidos para cada variable de entrada en el cálculo SIM y MGA, a partir del conjunto final promedio CVC-FFBPNN (DS1).

Variable	SIM Rank	SIM CV(%)	MGA Rank	MGA CV(%)
Temperatura	1	2.6	3	34
Presión atmosférica	4	5.6	2	31.7
Velocidad del viento	5	4.5	5	41.9
Humedad relativa	2	3.4	4	16.6
Número del mes	3	2.8	1	30.6

2.4. Discusión

Los resultados medios de nuestro método propuesto (Figura 2.15) *SIM* y medidos en la Tabla 2.1, en términos de las IR, concuerdan con las observaciones realizadas en las Islas Canarias durante la llegada de masas de aire tropical continental desde el desierto del Sahara, comúnmente conocido como tiempo sur, que presenta altas concentraciones de polvo en condiciones de alta temperatura y baja humedad relativa principalmente (Dorta (2001));(Cana (2002));(Avila *et al.* (1997)).

El modelo final entrenado en este capítulo, bajo la aplicación del algoritmo *SIM* (2.1), sitúa a la humedad relativa y a la temperatura ambiente como elementos claves para la formación de polvo (Dorta (2001));(Csavina *et al.* (2014));(Diaz-hernandez y Sanchez-Navas (2016)).

La variable de la velocidad del viento sin embargo, ocupa el último lugar y no tiene un alto grado de importancia en comparación con las demás variables. En su investigación, Tao y Jamalizadeh también analizaron y observaron la contri-

50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

90 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

90 / 181

2.4. Discusión

bución irregular de la variable de la velocidad del viento en diferentes escenarios, y encontraron algunos casos en los que no existía una correlación directa entre esta variable y la presencia de tormentas de arena (Gao *et al.* (2011));(Jamshid y Vahid (2008)).

Otro de los problemas encontrados en la bibliografía examinada, relativo al desarrollo de métodos para analizar importancias relativas en RNA, describen el problema de la aleatoriedad al reentrenar una RNA debido a la nueva inicialización de los pesos, que se produce en cada nuevo entrenamiento de la red. Sin embargo, la solución propuesta en este capítulo lo tiene en cuenta, por lo que el algoritmo, al eliminar las variables de entrada y en cada nuevo entrenamiento que debe realizar en cada caso, utiliza siempre los mismos pesos iniciales guardados y que se utilizaron en la primera RNA entrenada original. Por lo tanto, las condiciones iniciales del nuevo conjunto de entrenamiento FFBP-NN_{nk}, son las mismas en cada iteración para las diferentes variables de entrada x del modelo.

Las investigaciones de Maosen y otros, señalaron además de los problemas relacionados con los análisis de sensibilidad para la IR, la falta de determinación y claridad a la hora de definir el número de neuronas en la capa oculta de una RNA (Maosen y Pizhong (2008)). Para solucionarlo, en este procedimiento CVC presentado y como parte del diseño inicial de las RNA, se utiliza siempre el mismo método para seleccionar el número ideal de capas ocultas, que se establece en función del número de entradas y salidas (2.2) (Huang (2003)).

A la hora de establecer una comparativa del método propio que se ha desarrollado para el cálculo de IR, los hallazgos en el método MGA, reflejan un alto grado de variabilidad en las salidas, en términos del CV (Tabla 2.1) que se ha establecido como control de dispersión. Aumentando además su valor cuando se aplicó este método en una arquitectura de ensemble CVC y FFBP-NN. Al no tener este método en cuenta los diferentes pesos iniciales de cada RNA utilizada en este tipo de configuración de ensemble o agrupación. Por lo tanto, se confirma

51

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

91 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

91 / 181

Capítulo 2. Análisis multivariante mediante ensemble de redes neuronales artificiales

que este método, no garantiza la validez de los resultados de IR (Oña y Garrido (2014)).

Por último, cabe destacar que los resultados satisfacen el control de calidad establecido (2.10), lo que valida el método SIM. Basándonos en el grado de variabilidad (Figura 2.16), en el que los resultados se sitúan dentro del nivel definido como "bajo" (<10%) en el conjunto final resultante. Esto muestra que los 10x conjuntos finales de RNA FFBP-NN para cada una de las cinco variables de entrada utilizadas en este modelo, son relevantes e indican la correcta distribución y el buen funcionamiento de todo el conjunto de datos utilizados en DS1 para el cálculo de IR, en esta configuración de ensemble.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA
Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA
María de las Maravillas Aguiar Aguilar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

Capítulo 3

Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

En este tercer capítulo, se avanza en el estudio de la degradación debido al ensuciamiento del compresor axial de turbina de gas de generación eléctrica en las Islas Canarias y a través de un aprendizaje computacional inteligente, siendo éste vital para la planificación del mantenimiento, ya que permite obtener una foto actual y futura de un componente industrial a nivel individual, y capaz de analizar los niveles de ensuciamiento a lo largo del tiempo, en función del régimen de operación del equipo y de las variables meteorológicas locales.

En particular, se estudia la relación entre la degradación de la máquina, las intervenciones históricas de mantenimiento de limpieza realizadas y el polvo en suspensión de origen sahariano. Para ello, utilizamos un procedimiento computacional de CVC, descrito en el capítulo 2, para estimar la eficiencia isentrópica del compresor.

3.1. Estructura de los datos y arquitectura

Para este desarrollo computacional, el conjunto de datos se divide en dos bloques principales (Figura 3.1). El primer conjunto DS1 contiene los datos de entrada y salida desde enero 2013 hasta diciembre 2017, con una **frecuencia por**

53

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

93 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

minuto .

Este conjunto será utilizado exclusivamente para realizar todo el proceso de entrenamiento. El segundo conjunto de datos DS2, se sitúa entre enero 2018 has-



Figura 3.1. Conjunto de datos DS1 y DS2, utilizados en la fase de entrenamiento y test de la RNA.

ta diciembre 2018 y también incluye los datos de entrada y salida. Se utilizará exclusivamente para la fase de verificación y test de los resultados.

El tipo de RNA presentada es una **FFBP-NN**. Utilizando el mismo conjunto DS1 para su entrenamiento y aplicando una partición de datos para el subconjunto de entrenamiento y test por *k-fold* (Figura 2.6). Siendo el entrenamiento de la RNA, el método de optimización utilizado en el capítulo 2. El algoritmo de *Levenberg-Marquardt* modificado mediante la Regularización Bayesiana e inicializado mediante el criterio de *Nguyen-Widrow* (Sapna (2012));(Nguyen y Widrow (1990)).

En la **preparación de los datos** se eliminaron las filas que contenían valores constantes y se aplicó la normalización *min-max*. Una técnica que proporciona una transformación lineal en el rango original de los datos (Patro y Sahu (2015)).

La **arquitectura base** se establece de la siguiente forma. Ocho entradas (Temperatura ambiente; presión atmosférica; humedad relativa; número del mes; potencia eléctrica bruta generada; presión diferencial filtros de entrada; contador de mantenimiento del compresor realizado por *MTWork1* (3.1); **contador de arena sahariana**. Una única capa oculta de ocho nodos o neuronas, adaptando el número de nodos mediante el método propuesto en la ecuación (2.2) (Huang (2003)).

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

3.1. Estructura de los datos y arquitectura

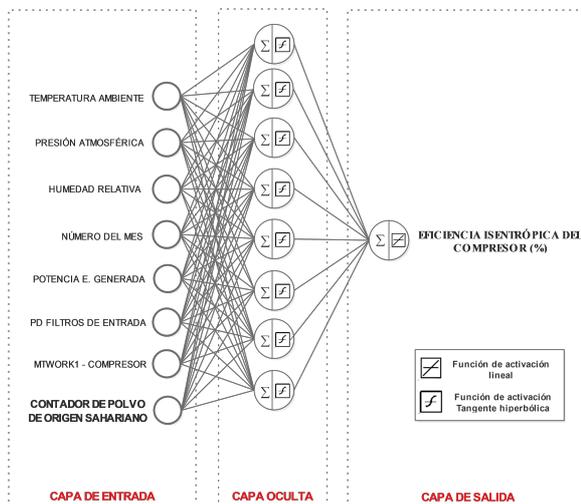


Figura 3.2. Topología de la estructura principal de la RNA. Diseñada con ocho nodos de entrada, ocho nodos en la capa oculta con función de activación tangente hiperbólica y un nodo en la capa de salida con función de activación lineal.

Se incluye además, una función de activación que utiliza la tangente hiperbólica y una capa de salida con función de activación lineal para la eficiencia o rendimiento isentrópico del compresor (Figura 3.2).

A diferencia del modelo anterior, donde las variables de entrada eran series temporales medidas, en este segundo desarrollo se combinan con tres **nuevas entradas calculadas** para la RNA. Como son el rendimiento isentrópico del compresor (1), el contador de polvo de origen sahariano (2) y el contador de eventos de limpieza del compresor (*MTWork1*) (3).

Los datos de la **concentración de polvo sahariano** ($\mu\text{gr}/\text{m}^3$), se sitúan en la

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

superficie de la isla de Tenerife. Combinando los datos del modelo del BSC-DREAM8b, junto con los datos de arena obtenidos por el modelo desarrollado en el capítulo 2.

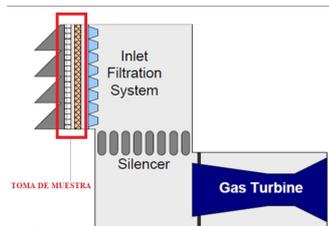


Figura 3.3. Muestra tomada de la zona de filtros de entrada de la turbina de gas (Wilcox et al. (2012)).

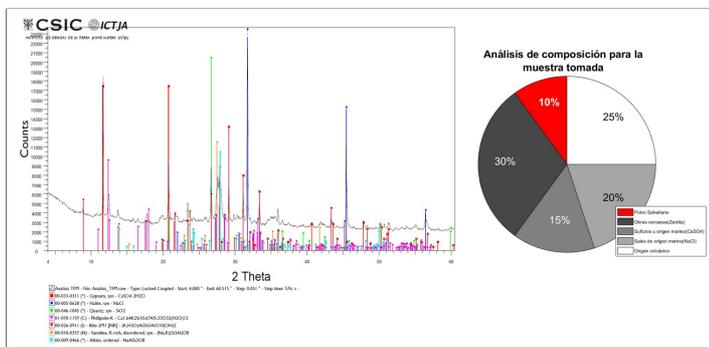


Figura 3.4. Difractograma de polvo, analizado por el laboratorio externo (CSIC-ICTJA) y la composición resultante de la muestra tomada en la zona de filtros de entrada de la turbina de gas.

Para poder construir la relación de polvo de origen sahariano en el modelo, se toma una muestra en la zona de admisión de la turbina de gas en los filtros de

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

3.1. Estructura de los datos y arquitectura

entrada (Figura 3.3) y se realiza análisis en laboratorio externo Consejo Superior de Investigaciones Científicas (CSIC), donde se ha adquirido un espectro de difracción mediante un difractor marca *Bruker*, modelo D8-A25 en geometría θ/θ , equipado con un tubo de rayos X de *Cu* (longitud de onda $\lambda=1,5405 \text{ \AA}$), rendijas *Soller* y detector ultra-rápido de tipo PSD (*Lynxeye*). Se ha realizado un barrido de 4° hasta 60° de 2θ , con pasos de $0,05^\circ$ y un tiempo de medida equivalente por paso de 576s. Para las medidas se han aplicado al tubo las siguientes condiciones de trabajo: voltaje de 40 kV , corriente de 40 mA . La interpretación de los espectros de difracción se ha realizado con el programa *DiffraSuiteTM de Bruker*, mientras que la identificación de las fases cristalinas ha sido realizada utilizando la base de datos *PDF-2* del *International Centre for Diffraction Data* (ICDD).

El difractograma de la muestra estudiada (Figura 3.4), permite identificar la presencia de las siguientes fases cristalinas: Yeso (15%), halita (20%), cuarzo (5%), feldespatos (sanidina (20%) y albita (5%)), la fase arcillosa moscovita/illita (5%) y zeolita phillipsita (30%).

Se han asociado los focos locales a las condiciones macaronésicas y se descarta la composición temporal, achacable a un aporte humano puntual. **La composición local de polvo sahariano comprende un 10 %** de la muestra analizada de ensuciamiento en la zona de filtrado de la turbina de gas, siendo la suma de los minerales cuarzo y moscovita, como parte de una composición normal de la arena de origen sahariano (Criado *et al.* (2012));(Dorta *et al.* (2005));(Dorta *et al.* (2002));(Pye (1987)).

Por otro lado y de cara a poder establecer un sistema de **control de polvo sahariano** que permita vigilar la acumulación de arena y discriminar todo aquel material que no provenga de la calima, se plantea la siguiente expresión para la acumulación de cada instante t y con una frecuencia de un minuto:

$$Q_{dust}(t) = Q_{dust}(t - 1) + \alpha \cdot \left(\left(\frac{q_{air} \cdot C_{dust}}{\rho_{air}} \right) \cdot 10^{-6} \right) \quad (3.1)$$

57

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

97 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

97 / 181

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

Donde el valor final de $Q_{dust}(gr)$, representa la cantidad de arena procedente de la calima que se va acumulando en el compresor de la turbina de gas, q_{air} $\left(\frac{kg}{s}\right)$ el caudal real de aire másico total que entra al compresor, C_{dust} $\left(\frac{\mu gr}{m^3}\right)$ es la concentración de arena medida en la superficie y recogida del modelo de polvo de origen sahariano desarrollado en el capítulo 2. La densidad del aire húmedo en función de la temperatura, humedad relativa y presión atmosférica calculada por el procedimiento CIPM-81/9 es ρ_{air} $\left(\frac{kg}{m^3}\right)$ (Davis (1991)).

Finalmente una constante α (%) representa el porcentaje correspondiente de polvo sahariano sobre el total de la muestra analizada en la zona local. Esto es, el resultado del análisis realizado que obtuvo un valor del 10 % (Figura 3.4).

El siguiente aspecto relevante para este desarrollo, es la **variable objetivo** y de salida. Partiendo de la base que el rendimiento de una máquina ya sea porque produzca o consuma trabajo, se expresa como una relación entre el trabajo que se ha transferiría de forma ideal y el que realmente se ha transferido. Dado que el funcionamiento de las turbos máquinas es prácticamente adiabático, el proceso que se toma como ideal es el isentrópico y el rendimiento que se obtiene se denomina rendimiento isentrópico.

Para disponer de una medida cercana a la realidad térmica de una turbina de gas, se define el **cálculo del rendimiento isentrópico del compresor**.

$$\mu_{co} = \frac{\left(\left(T_1 \cdot \left(\frac{p_2}{p_1} \right)^{\frac{\gamma-1}{\gamma}} \right) - T_1 \right)}{T_2 - T_1} \cdot 100 \quad (3.2)$$

Para este desarrollo, se dispone de las medidas reales de p_2 (bara) y p_1 (bara) como las presiones de entrada y descarga del compresor respectivamente. La presión de entrada tiene en cuenta la caída de presión por ensuciamiento de la

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

3.2. Procedimientos

zona filtrante. Además, T_1 y T_2 , son las temperaturas de entrada al compresor y a la descarga, respectivamente. Donde el coeficiente adiabático a calcular tiene en cuenta el calor específico, representado por la fracción de humedad en el ambiente y el rango de temperaturas medias reales medidas en el compresor (Zoran y Gvozdenac (2008));(Alhazmy y Najjarb (2004)).

$$\gamma = \frac{\left(cp_a \left(\frac{wf}{q_{airs} cp_a} \right) \cdot cv_h \right)}{\left(cv_a \left(\frac{wf}{q_{airs} cv_a} \right) \cdot cv_h \right)} \quad (3.3)$$

Siendo wf el caudal de humedad absoluta y q_{airs} el caudal másico de aire seco de entrada al compresor. Donde, para aproximarnos al valor real del aire húmedo entrante, definimos un conjunto de calores específicos (cp) como el cp_a a presión constante del aire seco, cp_h a presión constante de la fracción de aire húmedo, cv_a a volumen constante del aire seco y cv_h a volumen constante del aire húmedo. Todos ellos, obtenidos a partir del promedio real de temperaturas de entrada y salida del compresor (Santos y Andrade (2012)).

3.2. Procedimientos

Continuando con el esquema marcado de **MBD** (Figura 3.5), el procedimiento computacional de base, utilizado en esta sección dentro de **MBD.1**, viene dado por la Figura 3.6, donde en la salida final del modelo se ha estimado el rendimiento isentrópico del compresor, utilizando la humedad relativa, presión atmosférica, temperatura de entrada al compresor, potencia eléctrica bruta generada, contador *MTWork1* del compresor, la presión diferencial de entrada en los filtros de admisión y el totalizador de arena de origen sahariano.

Este procedimiento aplica la arquitectura de ensemble CVC (Figura 2.9) desarrollada en el capítulo 2. Sin embargo en este bloque evaluaremos además, la

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

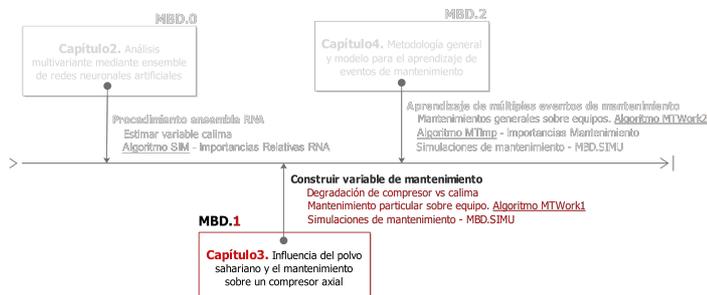


Figura 3.5. Esquema de la línea de desarrollo para la metodología de MBD para el capítulo 3.

robustez del CVC, comparando sus resultados junto con la **aplicación de otros métodos** clásicos de computación inteligente (Tabla 3.1).

En esta comparativa se han entrenado los modelos de *GB* y *RF* para un total de 500 y 340 estimadores respectivamente, siguiendo el procedimiento indicado en Tabla 3.1.

Los métodos *GB* y *RF* son modelos basados en árboles de decisión, utilizados para tratar problemas no lineales. Entre los diversos métodos de aprendizaje de máquina, el método *GB* propuesto por Friedman ha sido considerado como algoritmo de gran potencial, con muchas aplicaciones de éxito tanto para problemas de clasificación como de regresión en diferentes campos (Friedman (2002)).

El modelo *RF* es un procedimiento de aprendizaje de máquina muy popular que puede utilizarse para desarrollar modelos de predicción. Introducido por Breiman, los bosques aleatorios son una colección de árboles de clasificación y regresión, que utilizan divisiones binarias en variables predictoras para determinar las nuevas estimaciones (Breiman (2001)).

Además, para poder definir y verificar el número de *k*-fold definitivo en el CVC, se plantean diferentes combinaciones del valor *k*.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

3.2. Procedimientos

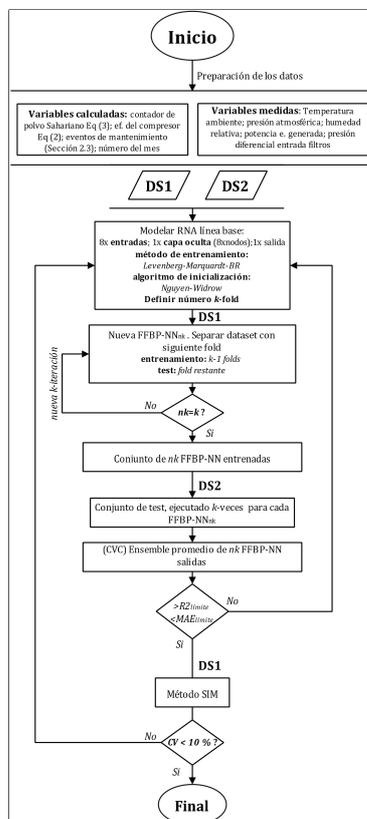


Figura 3.6. Flujograma del procedimiento principal utilizado en este trabajo. Sección 3.2.

61

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

101 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

Tabla 3.1. Resumen del estudio comparativo realizado con otros métodos de regresión para estimar el rendimiento isentrópico del compresor.

Junto con el procedimiento base, se aplican otros modelos para estimar el rendimiento del compresor utilizando algoritmos de GB, RF, una FFBP-NN simple y tres combinaciones diferentes del valor de k-fold, en el CVC.

Para cada modelo comparativo, a excepción del método principal de este capítulo (FFBP-NN CVC) se subdivide el conjunto de datos DS1 en un 70 % para entrenamiento y un 30 % para test.

La configuración de los hiperparámetros para los modelos GB y RF se realiza de forma iterativa mediante el algoritmo de optimización de rejilla, GridSearch optimizado para VC

De la misma forma que se ha buscado comparar el principal procedimiento de este trabajo CVC con otros métodos de ML (Tabla 3.1). Se realiza el mismo caso en la **aplicación del método SIM** para observar su respuesta y variación en la selección de diferentes configuraciones del ensemble CVC, utilizando para esto varias combinaciones de k-fold (Tabla 3.2).

Tabla 3.2. Resumen del estudio comparativo para observar el cálculo SIM utilizando diferentes configuraciones del ensemble FFBP-NN.

Se comparan sólo aquellos modelos de computación que están basados en RNA y se les aplica el método de SIM para el cálculo de IR.

El algoritmo SIM se aplica de forma individual para cuatro casos. El modelo simple de FFBP-NN y varias combinaciones del modelo principal de este trabajo (FFBP-NN CVC), con valores de $k=[5, 10 \text{ y } 15]$.

El resultado final de IR que se obtiene en cada modelo anterior, es el valor resultante, por cada entrada, de la etapa 18 indicada en el algoritmo SIM (2.1). Incluyendo el control de dispersión de la línea 16 de SIM, en términos del valor de CV.

En el caso de utilizar una FFBP-NN individual sin configuración de ensemble, esta supervisión de CV no es necesaria, ya que al ser un único conjunto neuronal, al aplicar el método SIM sólo tiene un valor de MAE_x para cada entrada (2.7) y por tanto no se verifica el grado de dispersión del MAE.

62

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

102 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

102 / 181

3.2. Procedimientos

Para poder integrar una **variable de mantenimiento** en modelos de computación inteligente, se desarrollan dos algoritmos fundamentales dentro de la metodología de MBD. El primero y el que será objeto de aplicación en este capítulo, el *MTWork1*, está enfocado a recoger todas aquellas intervenciones o eventos históricos de **mantenimiento específicos**, que sean **recurrentes o repetitivos** en el tiempo, sobre un equipo o elemento industrial particular. En este caso de aplicación se trasladarán a esta variable calculada, todas las intervenciones de limpieza fuera de línea del compresor de una turbina de gas entre los años 2013 y 2018 (Tabla 3.3).

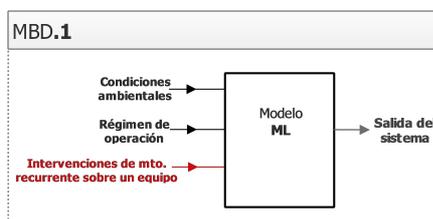


Figura 3.7. Esquema del procedimiento MBD.1 para intervenciones de mantenimiento recurrentes.

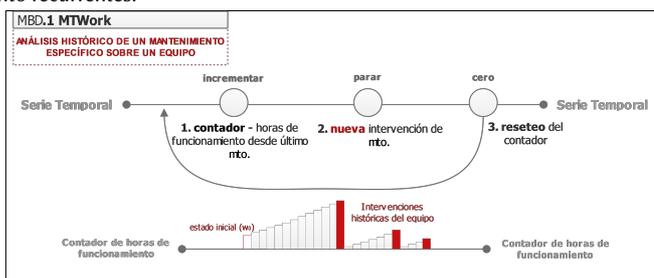


Figura 3.8. Descripción del ciclo de funcionamiento del algoritmo MTWork1 para definir la variable calculada de mantenimiento del modelo.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

El objetivo principal que tiene el algoritmo MTWork1, es determinar el tiempo que ha estado la turbina de gas arrancada y con una generación de potencia bruta de salida por encima de un megavatio eléctrico durante todo el periodo estudiado, estableciendo por tanto un contador acumulativo de tiempo de funcionamiento, que será *reseteado* bajo la condición de una nueva intervención de mantenimiento recurrente realizada. Esto significa que se pondrá de nuevo a cero en el instante que exista una limpieza *offline* del compresor, donde empezará de nuevo a acumular horas de funcionamiento hasta que se vuelva a producir otro nuevo ciclo de limpieza.

Además y en paralelo para este caso particular de aplicación, el **contador de arena de origen sahariano** y por lo tanto el Q_{dust} total (3.1), se *resetea* también a valor cero, cuando se realice cada vez un mantenimiento de limpieza del compresor. Junto a las intervenciones de mantenimiento reales asociadas a estas limpiezas *offline*, se añaden además como información de respaldo para el análisis, los trabajos de mantenimiento sobre los filtros de entrada del compresor. Resultando la siguiente secuencia histórica, asociada a DS1 y DS2.

Tabla 3.3. Registro histórico de eventos de mantenimiento (•) realizados en la turbina de gas durante todo el periodo de datos disponibles.

Tipo de Mantenimiento	DS1					DS2
	2013	2014	2015	2016	2017	2018
Limpieza Offline Compresor	•	•		•	•	•
Cambio de Prefiltros	•	•		•	•	•
Cambio de Filtros de Alta Eficacia		•				•

El primer algoritmo para crear variables de mantenimiento que puedan utilizarse dentro del proceso de MBD.1, se presenta a continuación:

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpFLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

3.2. Procedimientos

Algoritmo 3.1. MTWork1

```

1: Entradas:
2:  $x_{or}$ ,  $x_t$  ← series temporales del régimen de funcionamiento y los conjuntos de fecha y hora contenidos en
   DS[ $x_t, x_{or}, \dots$ ], de longitud  $P$  y expresado en la unidad de DS
3:  $w_0$  ← estado inicial del contador de horas
4:  $ME$  ← conjunto que contiene las fechas de los eventos de mantenimiento dentro de un rango dado de inicio y fin
5:  $t$  ← intervalo de tiempo de DS expresado en la unidad por defecto de DS
6: Salida:
7:  $MTW$  ← evolución del contador con todos los eventos históricos de mantenimiento, de longitud  $P$ 

8: for  $i=1$  to  $P$  by 1 do
9:    $MTW_{[i]} \leftarrow w_0$ 
10:  if  $x_{or[i]} > 1$  and  $x_{t[i]} \notin ME$  then
11:     $MTW_{[i]} \leftarrow MTW_{[i-1]} + t$ 
12:  else if  $x_{or[i]} \leq 0$  and  $x_{t[i]} \notin ME$  then
13:     $MTW_{[i]} \leftarrow MTW_{[i-1]}$ 
14:  else if  $x_{t[i]} \leq 0 \in ME_n$  then
15:     $MTW_{[i]} \leftarrow 0$ 
16:  end if
17: end for
    
```

Finalmente se plantea la sistemática definida para poner en servicio las **MBD.SIMU**.

De esta forma se podrá utilizar el modelo final entrenado mediante FFBP-NN CVC para poder evaluar diferentes escenarios como ayuda y respaldo en la toma de decisiones basadas en el aprendizaje de un modelo de *ML* (Figura 3.9).

Se han definido para ello dos tipos de simulaciones de cara a observar los resultados que se alcanzan en la variable objetivo de salida: el rendimiento isentrópico del compresor, bajo diferentes escenarios ambientales y de régimen de funcionamiento de la turbina de gas. En la primera simulación MBD.SIMU 1, se toman como datos de partida en las entradas, los correspondientes a la foto de la turbina de gas en el último mes de DS2 (diciembre-2018), seleccionando el promedio mensual para la temperatura ambiente, humedad relativa, presión atmosférica y el régimen de funcionamiento de la turbina de gas en ese periodo, en términos de la potencia eléctrica bruta generada. A partir de este instante, se irán simulando

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

horas de funcionamiento acumuladas sin realizar mantenimiento de limpieza del compresor, a medida que se incrementan los valores del contador total de arena sahariana Q_{dust} (3.1) y los valores de presión diferencial de la entrada de filtros, en intervalos de +20 gr y +5 mmH2O respectivamente, para cada escenario de horas de funcionamiento.

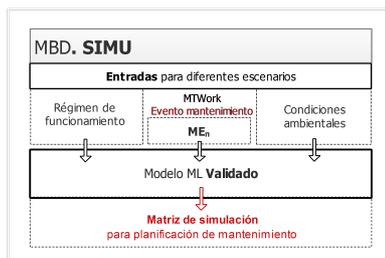


Figura 3.9. Esquema del procedimiento de simulación utilizado en este trabajo para la línea de MBD.

En paralelo además, se comparó con otra simulación donde se mantuvieron en un valor constante, las horas de funcionamiento de *MTWork1* acumuladas sin lavar el compresor. Fijándolo por tanto en 720 horas e incrementando sólo el valor total de Q_{dust} , para varios escenarios de presión diferencial de la zona filtrante de entrada.

Para la segunda simulación MBD.SIMU 2, se observa el resultado de la limpieza del compresor en función de la estación del año y en diferentes condiciones de funcionamiento, en términos de la potencia eléctrica bruta generada en la turbina de gas. Para esto, se eligieron cuatro escenarios donde se adapta únicamente la temperatura de entrada del aire al compresor en las condiciones medias de cada estación del año. El resto de variables de entrada ambientales serán el promedio de DS2.

Los valores Q_{dust} , las horas de *MTWork1* y la presión diferencial de los filtros serán reiniciados a valores iniciales, simulando así un mantenimiento de limpieza

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

3.3. Resultados

de lavado de compresor y un cambio de filtros de entrada en todos los casos de esta simulación.

3.3. Resultados

En primer lugar, se han representado todas las intervenciones de mantenimiento relacionadas con las limpiezas del compresor axial de la turbina de gas desde 2013 hasta 2018 utilizando las variables calculadas del rendimiento isentrópico del compresor (μ_{co}) (3.2) y el totalizador de arena de origen sahariano (Q_{dust}) (3.1). Tal y como se indica en la Sección 3.2, estas tareas de limpieza reinician el contador de horas de mantenimiento y reducen la cantidad del totalizador de arena a cero. (Figura 3.10).

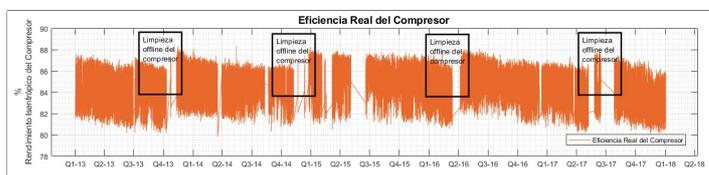


Figura 3.10. Serie temporal (2013-2018) de la eficiencia isentrópica real medida del compresor, mostrando los períodos de limpieza *offline* del compresor de la turbina de gas.

Para poder visualizar la evolución del rendimiento μ_{co} de forma adecuada, se ha utilizado una ventana temporal más amplia de varios años y aplicando además un suavizado SMA de la misma variable (Figura 3.11).

Los resultados obtenidos de **acumulación de polvo en el tiempo** en el interior de la turbina de gas, con el ajuste de composición de arena sahariana (α) frente al rendimiento μ_{co} de DS1 suavizado, se muestran asimismo en la Figura 3.11 y con una SMA de tres días para facilitar su interpretación gráfica y reflejan la relación existente entre ambas variables calculadas.

La evolución normal de ensuciamiento parece a simple vista, constante tras

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpFLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

cada fase de limpieza del compresor, coincidiendo con una reducción del rendimiento del orden de 1.5-2 %, transcurrido un periodo aproximado de nueve meses. Durante el primer trimestre la reducción del rendimiento del compresor se acelera y a partir del cuarto mes continua disminuyendo a menor velocidad.

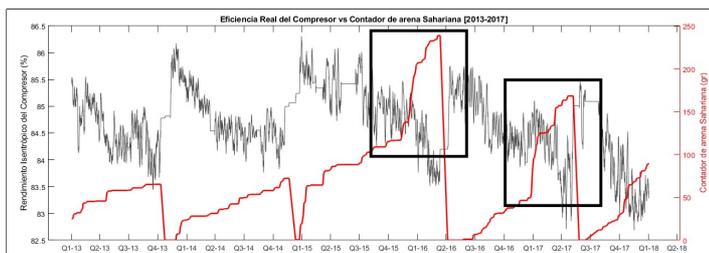


Figura 3.11. Serie temporal (2013-2018) de la medida real de la eficiencia del compresor, suavizada con una SMA de tres días, con detalles de dos eventos (2015-2016) que implican una alta acumulación de polvo sahariano y una disminución de la eficiencia.

En el intervalo 2015-2016, se produce un episodio donde no se realizó ninguna intervención de mantenimiento para limpieza, produciéndose una mayor acumulación de ensuciamiento y una fuerte subida en el contador de arena en el tramo final de este intervalo, debido a numerosos episodios de calima en la isla. Durante los periodos situados entre 2016 y 2017, bajo unas condiciones de episodios continuados de Q_{dust} por encima de los 150 gr, se aprecia una importante disminución del rendimiento del compresor en torno al 1 %, en una ventana de tiempo de sólo tres meses, a partir de valores $Q_{dust} > 100$ gr. Llegando a un mínimo de rendimiento del 83.5 %, esto es casi un 2.5 % menos con respecto al valor promedio de rendimiento obtenido en las limpiezas.

En estas condiciones, si lo comparamos con los episodios de 2013 y 2014, en el que se observan menores llegadas de arena sahariana y por lo tanto con menor afección sobre el rendimiento del compresor, podemos afirmar que existe una

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

3.3. Resultados

relación entre este fenómeno macaronésico y el rendimiento de un elemento industrial, como el compresor de una turbina de gas.

Analizando a continuación los **resultados del modelo neuronal** FFBP-NN CVC. Los valores predichos por el conjunto de ensemble ($k = 10$) y que han superado los límites estadísticos planteados en el procedimiento computacional, obtienen para DS1, un valor de $R^2 \approx 0.932$ y un $MAE \approx 0.36\%$ en términos de μ_{co} . Confirmando finalmente los buenos resultados de predicción del modelo, cuando en DS2, se alcanzó un $R^2 \approx 0.925$ y un $MAE \approx 0.53\%$, con una distribución del error agrupada mayoritariamente sobre el valor cero (Figura 3.12).

Cuando se comparan los rendimientos estadísticos obtenidos con otros modelos de ensemble basados en árboles de decisión y una RNA simple según lo indicado en la Tabla 3.1, es apreciable como la combinación del CVC con valores de $k = 5$, $k = 10$, y $k = 15$, pueden proporcionar el mejor escenario en términos del rendimiento estadístico (Tabla 3.4). Sin embargo, después de aplicarles el método *SIM* y observar su coeficiente de dispersión final, el modelo más estable se presenta para un valor de $k = 10$ o $k=5$.

Tabla 3.4. Resultados obtenidos para cada modelo de regresión entrenado.

Método de computación	DS1	DS1	DS2	DS2
	MAE	R^2	MAE	R^2
GB	0.34	0.89	0.49	0.80
RF	0.38	0.86	0.62	0.71
FFBP-NN	0.38	0.93	0.54	0.91
FFBP-NN CVC ($k=5$)	0.37	0.93	0.51	0.90
FFBP-NN CVC ($k=10$)	0.36	0.93	0.53	0.92
FFBP-NN CVC ($k=15$)	0.37	0.93	0.54	0.93

69

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

109 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

109 / 181

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

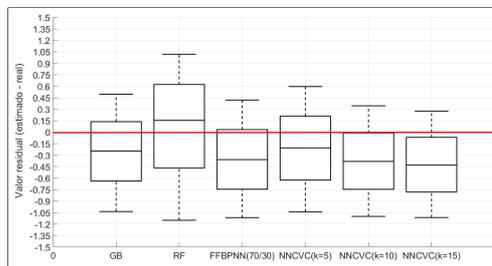


Figura 3.12. Diagrama de cajas que representa la cantidad de residuos para cada modelo de regresión entrenado con DS2 (Tabla 3.4).

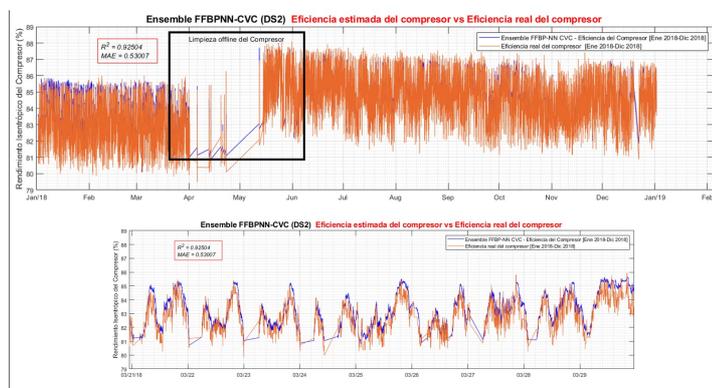


Figura 3.13. Resultados y rendimiento del modelo final FFBP-NN CVC (k=10) basado en datos reales de DS2. Detalle ampliado del modelo de aprendizaje para un lavado de compresor offline.

La respuesta aprendida sigue prácticamente la misma dinámica que la real (Figura 3.13). Además, el modelo no sólo ha aprendido a estimar satisfactoriamente la variable calculada de eficiencia del compresor, sino que cuando se le indica que

70

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

110 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

3.3. Resultados

en un determinado momento de la serie temporal se produce un evento de mantenimiento de limpieza sobre el compresor, este es capaz de reaccionar con un salto ascendente similar a la foto real obtenida tras un mantenimiento de la limpieza (Figura 3.10).

Tras confirmar los buenos resultados alcanzados del modelo de ensemble en términos del rendimiento estadístico sobre DS2 (Tabla 3.4) y siguiendo lo indicado en la Tabla 3.2, se registran los valores de SIM (Figura 3.14) para cada conjunto basado en RNA.

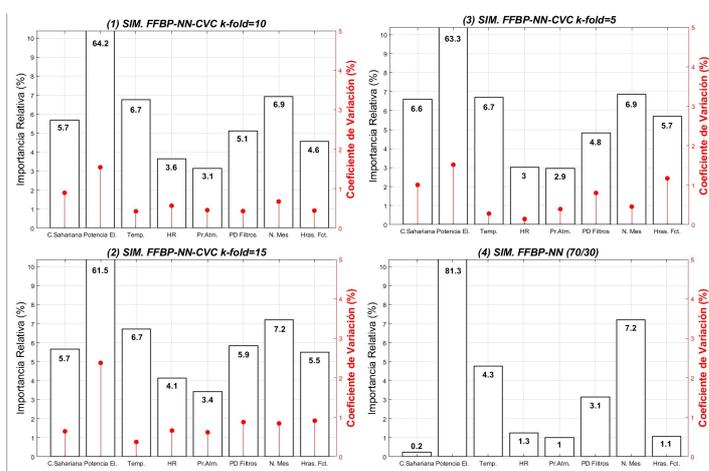


Figura 3.14. Niveles de importancia relativa aplicando el método SIM y su control de dispersión mediante el coeficiente de variación CV. Contador de polvo sahariano (C.S) ; Potencia eléctrica generada (Potencia El.); Temperatura ambiente (Temp.); Humedad relativa (HR.); Presión atmosférica (Pr.Atm); Presión diferencial - Filtros de entrada (PD Filtros); Número de mes (N.Mes); Horas MTWork1 (Horas Fcto.)

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

Esto ha permitido clasificar las variables más importantes utilizadas por nuestro modelo FFBP-NN CVC, y por lo tanto identificar aquellas que son claves para la eficiencia isentrópica del compresor o que, por el contrario, no aportan nada al sistema modelado (Figura 3.14).

Para reforzar la validez de los resultados en términos de la IR mediante el cálculo SIM, se estiman los valores para las diferentes combinaciones del modelo CVC, indicado en la Tabla 3.2.

Los resultados que se obtienen validan la fiabilidad del modelo SIM, ya que los modelos basados en RNA fueron entrenados de forma independiente, mostrando que para cada caso diferente del valor de k -fold, incluyendo la FFBP-NN individual donde la metodología de ensemble no se aplicaba. Unos niveles de IR que se mantienen en el mismo rango para cada uno de ellos.

Cabe destacar como variable relevante la potencia eléctrica de salida, por encima del resto, junto a la temperatura ambiente, el mes y el contador de arena. Sin embargo, se puede apreciar en la Figura 3.14 la RNA individual donde el peso del contador Q_{dust} ha sido penalizado frente al incremento de la potencia eléctrica bruta generada. Esta diferencia reside principalmente en la división de datos realizada para este modelo particular de FFBP-NN simple, donde la partición de su conjunto de entrenamiento era del 70 % frente al 90 % de los modelos de ensemble CVC, donde las iteraciones por cada $fold$ nos aseguran la participación al completo en la fase de entrenamiento y test, de todo el conjunto de datos (Sección 2.2).

La aplicación de este procedimiento computacional, implicó el entrenamiento de un total de 40x, 80x y 120x FFBP-NN respectivamente (Figura 3.14). Los resultados para los ensemble CVC se sometieron además, a la etapa de verificación (Figura 3.14) aplicando un coeficiente de variación CV para todas las variables involucradas.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

3.3. Resultados

Se confirma que el modelo final ha aprendido las diferentes intervenciones de mantenimiento en la limpieza *offline* del compresor y su correspondiente dinámica. Por lo tanto, estamos en disposición de realizar diversas **simulaciones** en varios escenarios de funcionamiento y entorno ambiental para presentarlas en términos del rendimiento del compresor. Buscando obtener una respuesta similar a la que tendrá realmente la máquina para disponer de un modelo inteligente que de soporte en la planificación del mantenimiento de la instalación (Figura 3.15).

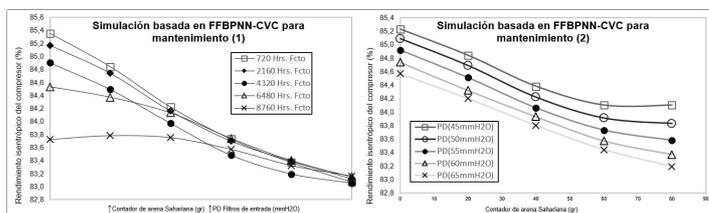


Figura 3.15. Simulación para la planificación del mantenimiento (MBD.SIMU 1); (1) Escenario de horas de funcionamiento; (2) Escenario de filtros de entrada PD.

En la primera simulación de **MBD.SIMU 1**, al trabajar en diferentes estados de horas de funcionamiento, observamos una degradación más rápida en el compresor con el incremento de arena sahariana y la *presión diferencial* (PD) en los filtros de entrada, cuando este lleva menos tiempo sin lavar (Figura 3.15). Sin embargo, los niveles de rendimiento isentrópico son mayores en los casos donde partimos de un número bajo de horas de funcionamiento sin realizar mantenimiento. Reflejando así, dos bloques principales, el primero situado entre las 720 y 4320 horas donde la tendencia es similar y la eficiencia del compresor pierde entre 0.4 y 0.5%. El segundo a partir de las 6480 horas de funcionamiento sin lavar, para el cual la degradación la encontramos en niveles de un 1 y 1.5% por debajo del valor que tendría en un escenario de funcionamiento próximo al mantenimiento realizado.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

Todos los supuestos, hasta las 6480 horas de funcionamiento, coinciden en una franja donde la acumulación de arena sahariana en el compresor y la PD de los filtros de entrada alcanzan prácticamente un estado estacionario. Esto es, a partir de valores de Q_{dust} igual a 60 gr y una PD de 55 mmH2O.

Tabla 3.5. Simulación MBD.SIMU 1. Horas de funcionamiento (MTWork1) sin realizar limpieza *offline* del compresor para condiciones de acumulación de arena y ensuciamiento de filtros.

$Q_{dust}(gr)$	PD(mmH2O)	720h	2160h	4320h	6480h	8760h
		$\mu_{co}(\%)$				
0	40	85.35	85.17	84.9	84.53	83.72
20	45	84.84	84.74	84.49	84.37	83.78
40	50	84.22	84.16	83.97	84.13	83.75
60	55	83.73	83.7	83.48	83.74	83.57
80	60	83.37	83.39	83.19	83.4	83.32
100	65	83.07	83.12	83.05	83.15	83.16

Por otro lado y en paralelo en la segunda simulación de MBD.SIMU 1, se detecta en la Figura 3.16 que para un mismo escenario de funcionamiento tras un lavado, los niveles de eficiencia del compresor se encuentran en torno a un 0.6 % por encima cuando disminuye la PD de los filtros de entrada en aproximadamente 20 mmH2O. Esto significa que la eficiencia del compresor será mayor siempre y cuando los filtros se mantengan adecuadamente. También en este caso se observa una zona de estabilización de la matriz de simulación para todos los niveles de PD en filtros de entrada calculados, a partir de una acumulación de arena aproximada de 60 gr. A partir de ese punto, se produce una reducción de la velocidad de degradación del compresor por ensuciamiento.

Aplicando la matriz de simulación MBD.SIMU 2, planteada en la Sección 3.3, se muestra la relación entre el comportamiento de los valores de eficiencia del compresor para cada estación del año y para cada escenario de funcionamiento de la turbina de gas.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

3.4. Discusión

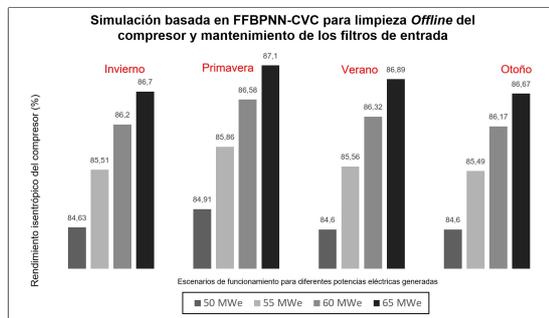


Figura 3.16. Simulación de la planificación del mantenimiento (MBD.SIMU 2); Escenario para la limpieza del compresor *offline* y el mantenimiento de los filtros, por estación.

A la vista de los resultados de la simulación para un escenario futuro, los meses de invierno y otoño son los periodos donde se sufre una mayor reducción sobre el resultado final de rendimiento, si realizáramos en esta franja un mantenimiento de limpieza *offline* del compresor. Con una diferencia entre invierno y primavera en torno a un 0.4 % de mejora en la eficiencia de del compresor.

3.4. Discusión

Los datos de rendimiento isentrópico del compresor han sido calculados para una turbina de gas de generación eléctrica dentro de un sistema eléctrico insular aislado. Su régimen de producción eléctrico bruto, entra dentro de la regulación secundaria de la frecuencia eléctrica de la isla (Guo e Yu (2004)). En otras palabras, no se mantiene esta turbina de gas en una misma carga base constantemente. Sino que se mantiene oscilando su producción con la frecuencia de la red en el

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 3. Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

balance diario de producción/demanda de la isla y normalmente con un margen de reserva rodante de su producción. Es por ello, que al observar la evolución del rendimiento (Figuras 3.10 y 3.11) en una ventana temporal reducida, veremos una oscilación continua de este valor de rendimiento de aproximadamente un $\pm 4\%$, derivado de este hecho.

Para resolver esta particularidad, se ha integrado como parámetro de entrada a la RNA, la potencia de salida producida en la arquitectura base, reflejando así estas oscilaciones de potencia en el modelo final entrenado.

El rendimiento promedio del conjunto en configuración de ensemble CVC, confirma lo que Atiq *et al.*, observaron. Si se combinan diferentes modelos promediando su salida conjunta, la precisión de la predicción mejora (Rehman y Bermak (2019)). En comparación con los resultados del procedimiento presentado en este trabajo, Pawelczyk *et al.* presentaron una comparativa de modelos de predicción mediante aprendizaje de la máquina, incluido una RNA individual de tres capas ocultas para la predicción de las características de funcionamiento de las turbinas de gas y sus posibles fallos. Sin embargo, se revisaron los modelos de RNA obtenidos y sus niveles de rendimiento estadístico fueron inferiores al conseguido por el modelo basado en el algoritmo de *RF* (Pawelczyk y Fulara (2020)).

Kohavi y Borra además, situaron en los resultados de su investigación, identificando un valor óptimo de k -fold = 10. Sin embargo, para este trabajo y después de estudiar el método de conjunto CVC aplicado, observamos que incluso con un valor de $k = 5$, el rendimiento obtenido es satisfactorio (Tabla 3.4) (Kohavi (1995));(Borra y Ciaccio (2010)). De hecho, tal y como se muestra en la Figura 3.12, el conjunto de residuos obtenidos para cada modelo confirma el valor k comentado.

Respecto a los resultados de los niveles de IR mediante *SIM*, la variable predominante por encima del resto de entradas, es la potencia eléctrica de salida. Esto es normal, ya que predomina la producción de energía eléctrica de la turbina de

76

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

116 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

116 / 181

3.4. Discusión

gas para mantener la regulación de la frecuencia de la red en el valor de referencia (Morren *et al.* (2005));(Mahat *et al.* (2009)). Por lo tanto, esto determinará las condiciones térmicas, sobre la presión de salida y la temperatura del compresor en todo momento (Kunitomi *et al.* (2001));(Thamir *et al.* (2011)).

Además, encontramos que la temperatura ambiente, el número de mes y el contador de arena sahariana son las variables con mayor impacto en la eficiencia del compresor. Esto confirma los resultados alcanzados en los estudios de De Sa, Maleki y Kurz, respecto a la relación de la temperatura ambiente y la época del año en la progresión del ensuciamiento en turbinas de gas (Sa y Zubaidy (2011));(Cruz-Manzo *et al.* (2017));(Wilcox *et al.* (2012)). Por lo tanto, estos hallazgos confirman y validan lo ya observado en la Figura 3.11 sobre la relación detectada entre la eficiencia isentrópica de un compresor de turbina de gas y el totalizador de polvo sahariano Q_{dust} (Inamul y Saravanamuttoo (2015));(Ujam *et al.* (2013));(Cuvelier y Belcher (2015));(Wilcox *et al.* (2012)).

En la simulación realizada dentro del escenario futuro situado por estaciones. Se sitúa a los meses de invierno y otoño como los periodos donde sufren una mayor reducción sobre el resultado final de rendimiento. Estos números coinciden con lo estudiado por Criado y Dorta, que también situaban en estos mismos periodos un mayor número de episodios de calima en las Islas Canarias (Criado *et al.* (2012));(Dorta *et al.* (2005)).

Por lo tanto, si realizáramos en esta franja una limpieza *offline* del compresor, teniendo en cuenta que la variable de eficiencia del compresor está directamente relacionada con el aumento en el consumo específico de la turbina de gas, supondría un mayor aumento del consumo de combustible se si planifican este tipo de mantenimiento en estas fechas.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

118 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo

4

Metodología general y modelo para el aprendizaje de eventos de mantenimiento

Finalizamos los desarrollos propuestos en la línea de trabajo que lleva hacia un método de análisis para el MBD, desde su diseño conceptual hasta los últimos algoritmos desarrollados para cerrar la metodología. De forma que un conjunto de intervenciones históricas de mantenimiento de revisión de un elemento industrial, se conviertan en un grupo de entrada de datos y puedan integrarse en un modelo de aprendizaje de máquina.

Esto combinado dentro de una arquitectura computacional definida, permite llevar desarrollar un nuevo algoritmo de análisis general de IR para observar la calidad de los mantenimientos, a partir de la información tomada en tiempo real. Para poder evaluar la metodología planteada, se desarrolla un modelo aplicado a un caso real de mantenimiento sobre emisiones ambientales de motores diesel de generación eléctrica, evaluando el sistema bajo sus condiciones operativas y el entorno atmosférico, permitiendo desarrollar nuevos modelos basados en datos para la implementación y mejora de la planificación del mantenimiento de un elemento industrial.

79

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

119 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

119 / 181

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

4.1. Estructura de los datos y arquitectura

El conjunto de datos del motor diesel de generación eléctrica consiste en un conjunto inicial DS1 y DS2 contienen los datos de entrada y salida $[X_{train}, Y_{train}]$ (Figura 4.4), desde Julio 2015 a Febrero 2019, con una **frecuencia horaria** para todo el conjunto (Figura 4.1). A diferencia de los capítulos 2 y 3, donde cada DS ocupaba diferentes periodos. En este desarrollo DS1 y DS2 tienen definido el mismo periodo, para de forma general y aleatoria asignarle a DS1 un 70% del total de la muestra y un 30% para DS2.

El único tratamiento inicial al que se ha sometido el conjunto de datos, es un **filtrado manual** definido por el rango de funcionamiento en términos de la potencia eléctrica bruta generada por el motor diesel. Estableciendo un rango de trabajo de entre 17.5 y 19.5 megavatios eléctricos (MWe).



Figura 4.1. Conjunto de datos DS1 y DS2, utilizados en la fase de entrenamiento y test de la RNA.

Siguiendo lo indicado en la metodología MBD, la **arquitectura del modelo** está formada por **tres variables medidas** ambientales; temperatura ambiente, humedad relativa y presión atmosférica, **dos variables medidas** relacionadas con el régimen de funcionamiento; potencia de salida generada y velocidad del motor y **ocho variables calculadas** que representan cada evento de mantenimiento (Tabla 4.1). Por último, en el único valor de salida, se encuentran los óxidos de nitrógeno emitidos por el motor diesel (NOx)¹.

¹Debido a la confidencialidad y sensibilidad de los datos reales de emisiones ambientales de NOx del motor diesel. El conjunto total de esta serie, se ha multiplicado por un factor de corrección.

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

4.2. Procedimientos

Dentro de la etapa de entrenamiento se aplica la misma **división de datos** que en los capítulos anteriores. Para cada DS, se creará un subconjunto de entrenamiento y test definido por k -fold con un valor estándar $k=10$ (Figura 2.6), que se unirán posteriormente en una configuración de ensemble promedio CVC.

Sin embargo en este caso, los 10x modelos serán entrenados mediante la técnica de *ML* basada en el ensemble de árboles de decisión para regresión con *GB*. Siendo el modelo base el algoritmo *Light Gradient Boosting Machine (LGBM)*.

Donde se aplica una selección de sus **hiperparámetros** no de forma aleatoria sino siguiendo el procedimiento de búsqueda de rejilla *GridSearch*, resultando finalmente en 520 estimadores o árboles, 50 número de hojas y un máximo de 8 en la profundidad del modelo.

4.2. Procedimientos

Para finalizar el esquema indicado de **MBD** (Figura 4.2), se define el procedimiento computacional de base utilizado en esta sección dentro de **MBD.2**, que viene dado por las Figuras 4.3 y 4.4. Este esquema final y de carácter general, recoge la sistemática que se ha seguido a la hora de crear desarrollos basados en *ML* y aplicarlos al mantenimiento industrial.

Tal y como se ha ido desarrollando en los anteriores capítulos, su aplicación permitirá el aprendizaje de un evento de mantenimiento en un modelo inteligente. Separando en tres etapas la metodología final (Figura 4.4). La selección de variables de entrada (2A); medidas o calculadas, la aplicación de un modelo de aprendizaje de máquina (2B) y la evaluación y simulación final (2C).

La aplicación en este caso se sitúa en un motor diesel de generación eléctrica. Por lo tanto dentro del grupo 2A necesitaremos añadir variables del entorno ambiental de la máquina y sus condiciones de explotación o funcionamiento. El cual se completa con el bloque de mantenimiento.

81

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

121 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

121 / 181

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

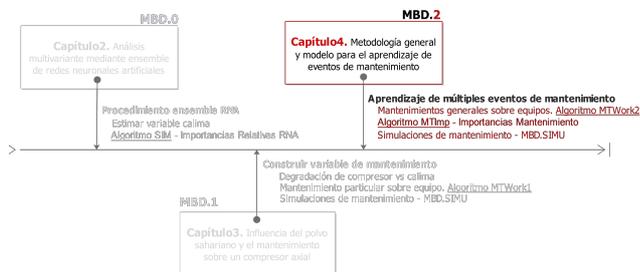


Figura 4.2. Esquema de la línea de desarrollo para la metodología de MBD para el capítulo 4.

En este estudio, el bloque recogerá la información relativa a los eventos de mantenimiento de revisión que se han realizado en el motor diesel de forma histórica. Desde revisiones menores, intermedias y mayores y por lo tanto con un punto de vista más general y no de cada equipo del motor en particular. El modelo aprenderá a estimar la afección del conjunto de todos los trabajos realizados en cada evento de cada una de estas revisiones del grupo sobre la variable de salida, los NOx.

Para resolver el problema, se ha planteado y adaptado un **procedimiento de base estándar** que pueda utilizarse en cualquier otro caso de computación inteligente aplicado en particular, al seguimiento y control de tareas o eventos de mantenimientos sobre un elemento industrial.

Siguiendo la misma línea de desarrollo planteada anteriormente, se divide en tres bloques fundamentales (Figura 4.4). Un primer bloque 2A, donde se toman los datos de proceso de la planta y se clasifican a su vez en tres grupos principales. Condiciones atmosféricas del entorno del elemento, su régimen de funcionamiento y su registro de mantenimiento. Siguiendo el proceso normal de un tratamiento de aprendizaje de máquina, los datos son filtrados y/o normalizados, donde posteriormente en el bloque 2B se realiza la separación de datos en con-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

4.2. Procedimientos

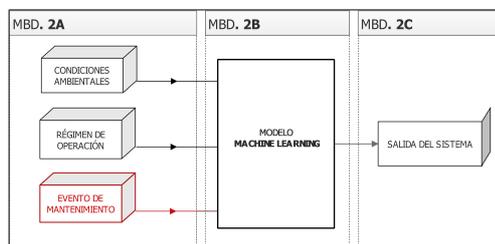


Figura 4.3. Esquema principal de la metodología MBD.2 .

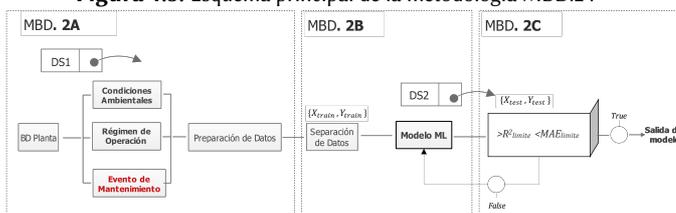


Figura 4.4. Procedimiento computacional general de MBD, dividido en tres bloques principales (2A, 2B y 2C).

juntos de entrenamiento y test $[X_{train}, Y_{train}]$ de un conjunto de datos conocido como DS1. En este punto el conjunto es alimentado a un modelo de aprendizaje de máquina, cuya salida será evaluada utilizando como contraste DS2, que no ha participado en el entrenamiento $[X_{test}, Y_{test}]$ y donde se establece un control de calidad final del rendimiento como el utilizado en la Sección 2.2, para asegurar la fiabilidad del modelo final entrenado.

A la hora de **construir la variable de mantenimiento** y a diferencia del primer algoritmo *MTWork1*, es necesario realizar una modificación de este para poder abordar una segunda variante que pueda reconocer, no sólo un único tipo de mantenimiento histórico de manera acumulada sino diferentes eventos de man-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

tenimiento generales realizados en un instante t , para así poder estudiarlos de forma desagregada en un modelo de ML .

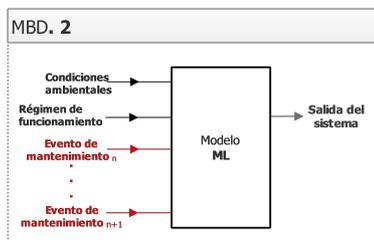


Figura 4.5. Diagrama MBD.2A, para el estudio de intervenciones de mantenimiento por cada nuevo evento temporal ME_n .

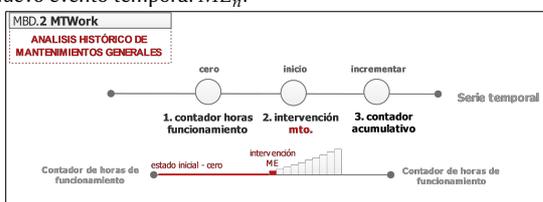


Figura 4.6. Detalle ampliado para la determinación de la variable “ME” sobre MBD.2A.

En la Sección 3.2 se relacionaba una acción de mantenimiento concreta sobre el rendimiento isentrópico de un compresor a lo largo del tiempo. El cual, a diferencia de este aprendizaje general de n eventos de mantenimiento, implicaba crear una única variable de entrada que representaba en el modelo de aprendizaje, todo el histórico de registro de mantenimiento del compresor.

Para conseguir adaptar este nuevo enfoque, se diseña el algoritmo *MTWork2*, donde se define un nuevo conjunto de n variables de entrada agrupadas dentro de *eventos de mantenimiento (ME)*. Creando $n+1$ registros, a medida que se va incre-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3170295	Código de verificación: a2/na6Ur
Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015. La autenticidad de este documento puede ser comprobada en la dirección: https://sede.ull.es/validacion/	
Identificador del documento: 3178786	Código de verificación: axpfLfm8
Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

4.2. Procedimientos

mentando por cada intervención real de mantenimiento de revisión que se ha realizado sobre el motor diesel (Figura 4.5).

Cada variable n por evento de mantenimiento, se construye a partir de un contador de horas de funcionamiento, el cual se iniciará en un valor cero constante, empezando a contar justo en el instante donde se produce la intervención de mantenimiento ME_n . A partir de ese momento, se irá incrementando en función del tiempo de funcionamiento del propio motor diesel (Figura 4.6). Este incremento se produce siempre que la variable de potencia eléctrica generada, sea superior a un MWe durante todo el periodo del estudio y además, a diferencia de MTWork1, en ningún caso volverá a reinicializarse a valor de cero.

Esta secuencia permite crear una nueva entrada ME_n , de un modelo de ML, por cada intervención histórica de mantenimiento y se recoge en el siguiente pseudocódigo:

Algoritmo 4.1. MTWork2

```

1: Entradas:
2:  $x_{or}$ ,  $x_t$  ← series temporales del régimen de funcionamiento y los conjuntos de fecha y hora contenidos en  $DS[x_t, x_{or}, \dots]$ , de longitud  $P$  y expresado en la unidad de DS
3:  $w_0$  ← estado inicial del contador de horas
4:  $ME$  ←  $n$  conjuntos de fechas con los eventos de mantenimiento dentro de un rango de inicio y fin
5:  $t$  ← intervalo de tiempo de DS expresado en la unidad por defecto de DS
6: Salida:
7:  $MTW$  ←  $n$  conjuntos con la evolución del contador de eventos históricos de mantenimiento de longitud  $P$ 
8: while  $n > 0$  do
9:   for  $i=1$  to  $P$  by 1 do
10:    if  $x_{t[i]} < ME_1$  or  $x_{t[i]} \in ME_n$  then
11:       $MTW_n[i] \leftarrow w_0$ 
12:    else if  $x_{or[i]} > 1$  and  $x_{t[i]} \notin ME_n$  then
13:       $MTW_n[i] \leftarrow MTW_n[i-1] + t$ 
14:    else if  $x_{or[i]} \leq 0$  and  $x_{t[i]} \notin ME_n$  then
15:       $MTW_n[i] \leftarrow MTW_n[i-1]$ 
16:    end if
17:  end for
18:   $n \leftarrow n-1$ 
19: end while
    
```

85

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

125 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

Los **casos históricos** de las diferentes intervenciones de mantenimiento de revisión que se presentan para el caso aplicado a motores diesel, tienen lugar y se codifican según la cronología que se indica en la Tabla 4.1. En esta secuencia de eventos de mantenimiento realizados desde 2016 hasta 2019, encontramos revisiones menores que se realizan cada 2500, 7500 y 12500 horas de funcionamiento respectivamente. Revisiones intermedias de 5000 y 10000 horas e intervenciones de mantenimiento mayor cada 15000 horas.

El volumen de trabajos que se realizan sobre el motor en una revisión de 15000h abarca el máximo número de tareas de mantenimiento que se realizan en este tipo de motores de generación eléctrica.

Tabla 4.1. Registro histórico de mantenimientos de revisión (•) realizados en el motor diesel de generación eléctrica.

Tipo de Mantenimiento de revisión	2016	2017	2018	2019
MTO10000	•		•	
MTO12500		•		•
MTO15000		•		
MTO2500		•		
MTO5000			•	
MTO7500			•	

De cara a poder medir el **grado de afición** que ha tenido cada intervención de mantenimiento sobre una variable de salida y esta relación de aprendizaje pueda vincularse de forma automática a diferentes escenarios ambientales y de régimen de funcionamiento, se presenta el algoritmo *MTImp* (4.2).

Este desarrollo es un segundo paso tras la aplicación del algoritmo *SIM* (2.1), utilizado en la aplicación específica de RNA para explicar los niveles de IR en una primera fase preliminar de los modelos de RNA y como inicio del procedimiento de MBD.

Sin embargo ahora el planteamiento se centra en un aprendizaje general y aplicable a cualquier método de computación para *ML*. Específicamente en analizar

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

4.2. Procedimientos

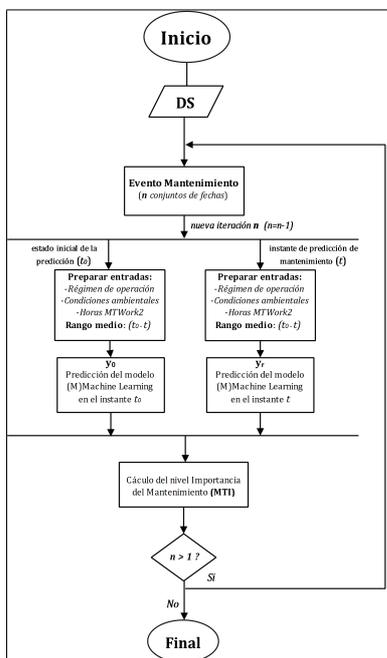


Figura 4.7. Flujograma del método propio para el cálculo de IR en eventos de mantenimiento, MTImp.

y comparar la contribución de diferentes eventos históricos de mantenimiento sobre variables del proceso de un elemento industrial.

El *MTImp* recoge la información de la serie temporal *DS1*, una vez que se ha verificado la validez y el rendimiento del modelo (Figura 4.7). Donde tendrá en cuenta, las variables medidas del ambiente y de funcionamiento $[x_a, x_{or}]$, la variable calculada que contiene los contadores *MTWork2* para cada evento de man-

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

tenimiento y las diferentes fechas de registro con el histórico de intervenciones ME.

Este algoritmo general que se presenta a continuación, permite estimar los niveles de importancia en diferentes escenarios de tiempo t , de forma que el modelo pueda tener en cuenta los tiempos de estabilización y ajuste que pueden producirse tras realizar una revisión de mantenimiento y ponerse de nuevo en servicio, en este caso, en un motor diesel de generación eléctrica.

Esta información se procesa y se simula de forma iterativa sobre el modelo de aprendizaje de máquina verificado, en dos escenarios de tiempo $[t_0, t]$, los cuales representan los instantes antes y después de cada intervención de mantenimiento respectivamente, para cada evento ME. En ambas simulaciones $[\hat{y}_0, \hat{y}_r]$, se utiliza el mismo valor promedio de entorno ambiental y régimen de funcionamiento. Esto es, el aprendizaje del modelo en toda la trayectoria histórica real del motor, permite poder hacer comparativas simulando las mismas condiciones de ambiente y de funcionamiento, sin necesidad de utilizar ningún otros factores de corrección empíricos dados por el fabricante. Finalmente se determina el nivel de importancia relativa en valor absoluto MTI , para cada intervención n .

En este capítulo además, se utiliza el modelo final para explicar y contrastar las IR sobre las variables calculadas de mantenimiento, los contadores de funcionamiento $MTWork2$. Aplicando para esto, los métodos *SHAP*, basado en un enfoque de la teoría de juegos para modelos de aprendizaje de máquinas y *FeatImp* específico para modelos de *GB*.

Sus resultados servirán para discutir y evaluar el desarrollo y aplicación del algoritmo propio *MTImp*.

Se concluye este último desarrollo aplicando el algoritmo de IR, el *MTImp* sobre el modelo *LGBM*, en diversos instantes $t=[48h, 168h]$ y manteniendo en cada caso, las mismas condiciones de funcionamiento y entorno ambiental para analizar, con el aprendizaje del modelo, cada revisión de mantenimiento ME.

88

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

128 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

128 / 181

4.2. Procedimientos

Algoritmo 4.2. MTImp

```

1: Entradas:
2:  $x_{or}, x_t, x_a \leftarrow$  series temporales del régimen de funcionamiento, los conjuntos de fecha y hora y condiciones ambientales contenidos en  $DS[x_t, x_{or}, \dots]$ , de longitud  $P$  y expresado en la unidad de DS
3:  $MTW \leftarrow n$  conjuntos con la evolución del contador de eventos históricos de mantenimiento de longitud  $P$  y expresado en la unidad de DS
4:  $ME \leftarrow n$  conjuntos con las fechas de los eventos de mantenimiento dentro de un rango de inicio y fin
5:  $t_0 \leftarrow$  estado inicial antes del mantenimiento, expresado en la unidad por defecto de DS (e.g.,  $t_0 = 0$ )
6:  $t \leftarrow$  estado después del mantenimiento, expresado en la unidad por defecto de DS (e.g.,  $t = 48h, 168h, \dots$ )
7:  $M \leftarrow$  modelo validado y entrenado de ML
8: Salida:
9:  $MTI \leftarrow n$  conjuntos con el cálculo de IR para cada evento de mantenimiento
10: while  $n > 1$  do
11:   for  $i = 1$  to  $P$  by 1 do
12:     if  $x_{t[i]} == ME_1$  then
13:        $k_0 \leftarrow i$ 
14:     else if  $x_{t[i]} == ME_{n+1}[1]$  then
15:        $k_f \leftarrow i$ 
16:     end if
17:   end for
18:    $x_{or}[k_0 + t] \leftarrow \bar{x}_{or}[k_0, k_0 + t]$ 
19:    $x_a[k_0 + t] \leftarrow \bar{x}_a[k_0, k_0 + t]$ 
20:    $\hat{y}_0 \leftarrow M(MTW[k_0 + t_0], x_{or}[k_0 + t], x_a[k_0 + t], \dots)$ 
21:    $\hat{y}_r \leftarrow M(MTW[k_0 + t], x_{or}[k_0 + t], x_a[k_0 + t], \dots)$ 
22:    $MTI_n \leftarrow |y_r - y_0|$ 
23:    $n \leftarrow n - 1$ 
24: end while
25: reset  $n$ 
26: while  $n > 1$  do
27:    $MTI_n \leftarrow \frac{MTI_n}{\sum_i^n MTI_n}$ 
28:    $n \leftarrow n - 1$ 
29: end while
    
```

Tras su aplicación sobre el caso de estudio, se plantea una **simulación en el modelo final** siguiendo el esquema MBD.SIMU (Figura 3.9) sobre DS1, de tal forma que se utilice el soporte del modelo para verificar en aquellas intervenciones con grados de importancia relevantes, cómo evolucionaron los NOx antes y después

89

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

129 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

129 / 181

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

del periodo de su intervención de mantenimiento, frente a un escenario que además simula qué hubiera pasado si no se hubiera realizado dicha intervención de mantenimiento. Fijando para esto en cero, el valor del contador sobre *variable contador de salida de los algoritmos MTWork (MTW)* del evento ME por cada intervención de mantenimiento simulada.

El aprendizaje del modelo validado y las variables contador de MTW, permiten forzar y crear una situación paralela donde no existiera dicho mantenimiento y observar así cómo hubiera afectado este nuevo caso a la variable de emisiones del motor diesel. De tal forma que sirva para contrastar la calidad y efectividad de cada uno de los tipos de mantenimiento de revisión.

4.3. Resultados

Los resultados obtenidos por el modelo, siguiendo la metodología MBD.2 aplicada a un caso real de un motor diesel, mediante el conjunto de datos DS1 obtuvieron un $R^2 \approx 0.88$ y un $MAE \approx 2.21 \text{ mg/m}^3$ en términos del valor de NOx.

Al evaluar el modelo final promedio resultante del CVC, con el conjunto DS2, se alcanza un $R^2 \approx 0.85$ y un $MAE \approx 2.48 \text{ mg/m}^3$. Estos resultados se encuentran por debajo del límite establecido en términos del control de calidad estadístico que se ha planteado a lo largo de este trabajo (2.5) y (2.6).

Con los buenos resultados obtenidos y evaluados en DS2, aplicamos MTImp sobre DS1, con un mayor rendimiento estadístico, observando que la revisión de mantenimiento que mayor efecto supuso para la variable NOx del motor, es la revisión de mantenimiento intermedia, MTO10000 2018. Tanto en un escenario de funcionamiento $t=48\text{h}$, como en $t=168\text{h}$, donde los tiempos de estabilización y ajustes finales del motor ya han podido realizarse (Figura 4.9).

En ambos escenarios también podemos encontrar las revisiones intermedia y menor MTO10000 de 2016 y MTO12500 de 2017 respectivamente, con un impacto

90

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

130 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

130 / 181

4.3. Resultados

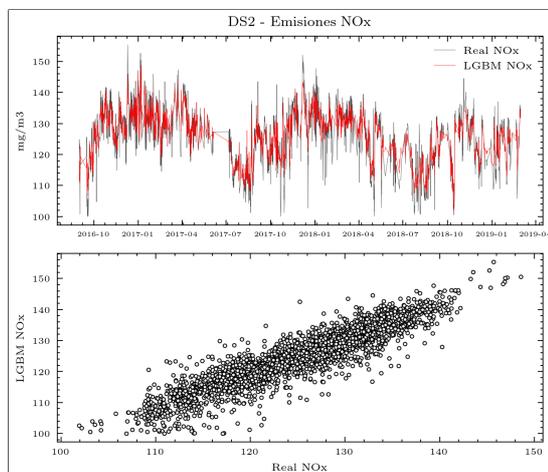


Figura 4.8. Resultados de la serie temporal DS2 con valores reales de emisiones de NOx frente a la predicción realizada por el modelo *LGBM*.

menor sobre la variable de salida. Aunque también encontramos diferencias entre ambos instantes de análisis t , confirmando una característica fundamental de este algoritmo, la posibilidad de realizar análisis de IR, para cualquier instante posterior a la finalización del mantenimiento.

La única revisión mayor que se registró en periodo de muestra MTO15000 2017, en los escenarios de funcionamiento de $t=48h$ y $t=168h$, no tuvieron una afección relevante en este periodo de estabilización, sin embargo este nivel de impacto se incrementa a partir de un $t=450h$ (Figura 4.11).

Si se hubieran analizado los niveles de IR con *MTImp* a partir de las 450 horas de funcionamiento, la revisión mayor de 15000h y la revisión intermedia de 10000h de 2018 hubieran aumentado su valor de relevancia con respecto a los niveles originales. Ya que en ambos casos la tendencia de emisiones evoluciona de

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

forma decreciente con respecto a t_0 . Este caso se justifica asumiendo que en una revisión mayor de 15000h, el motor se ha sometido a una serie de intervenciones y ajustes de carácter relevante durante más de 30 días de trabajos de mantenimiento. Por lo tanto, el periodo de estabilización y ajuste tras la puesta en marcha del mismo, es mayor que en otro tipo de revisión cuya duración y trabajos son de carácter intermedio o menor.

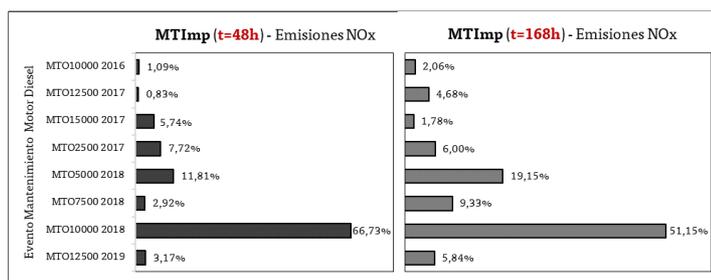


Figura 4.9. IR sobre un modelo de eventos de mantenimiento MBD.2 a partir del algoritmo MTImp.

Como medida de contraste y para observar la efectividad de otros algoritmos de IR para GB sobre los eventos de mantenimiento, se han obtenido las medidas de SHAP y FeatImp sobre DS1.

El resultado se observa en la Figura 4.10, donde la secuencia de importancias coinciden en ambos métodos, penalizando a los eventos de mantenimiento más recientes frente a los más antiguos y agrupando de forma dudosa para ello, los niveles de IR por orden cronológico.

Finalmente las simulaciones de MBD.SIMU 3, se construyen siguiendo el esquema propuesto y se aplican sobre tres mantenimientos de revisión. Dos mantenimientos intermedios de 10000h y una revisión mayor de 15000h. A partir del

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

4.3. Resultados

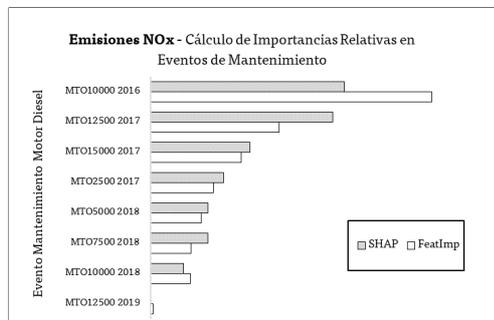


Figura 4.10. Niveles de IR para eventos de mantenimiento en emisiones de NOx mediante métodos *SHAP* y *FeatImp*.

grado de importancia obtenido por el aprendizaje del modelo en *MTImp*, estas revisiones de mantenimiento seleccionadas, representan los tres niveles de afectación más bajo, medio y más alto sobre los NOx del motor diesel para los instantes $t=48h$ y $t=168h$ (Figura 4.11).

Para el primer caso de los trabajos de revisión mayor de 15000h donde la predicción se comporta de forma similar a los valores reales de emisiones, se observa como el modelo marca en rojo una clara línea base a partir de las 300 horas *MTWork2* después de la puesta en marcha del motor, donde la simulación del modelo en tiempo real indicaba en cada nuevo instante que los niveles de NOx debían mantenerse en un nivel superior y sin embargo, después de un tiempo aproximado de 350 horas tras la intervención, se aprecia la efectividad de los trabajos de revisión y un descenso de unos $6 \text{ mg}/\text{m}^3$ frente al valor de referencia que indica la simulación del modelo.

Por otro lado, en la revisión de 2016, con una incidencia más baja sobre los NOx, se observa el caso opuesto al anterior. Esto es, los niveles de emisiones con

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguilar Aguilar 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

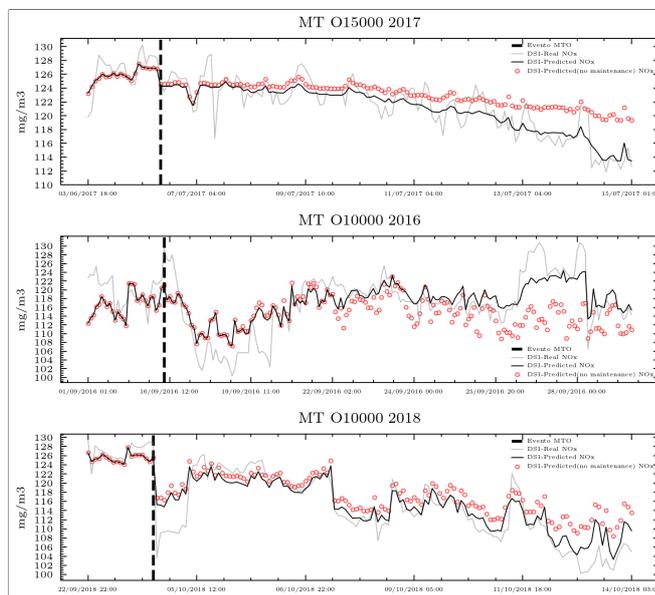


Figura 4.11. Simulación MBD.SIMU 3 sobre el NOx sobre DS1, con los valores reales, predichos y si no se hubiera realizado un mantenimiento de revisión. Se analizan las intervenciones de mantenimiento MTO15000 2017, MTO10000 2016 y MTO10000 2018.

respecto a la línea de referencia que el modelo simulaba para después de la intervención y a partir de las 200 horas, comienza a superarse progresivamente. Por lo que el resultado de los trabajos de mantenimiento sobre esta variable objetivo, no se señalaron como satisfactorios, siendo necesario una corrección y revisión posterior del motor para volver a reconducir los niveles de emisión.

El último caso, sin embargo, en el mismo tipo de revisión intermedia que la realizada en 2016, coincide con la referencia esperada por el modelo en cuanto a

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

4.4. Discusión

la trayectoria del resultado de la intervención sobre los valores de NOx. Aunque la reacción del modelo establece que, el valor de emisiones esperado sin mantenimiento se reduce y esto puede demostrarse por el tipo de mantenimiento realizado. En este caso, el evento de 2018 asocia un menor número de trabajos planificados que los ejecutados en una revisión mayor.

4.4. Discusión

De la misma forma que en el capítulo 3, los datos reales de emisiones de NOx se han tomado de un motor diesel de generación eléctrica dentro de un sistema eléctrico insular aislado y en una franja de 17.5 a 19.5 MWe. Por lo tanto su régimen de producción eléctrica, entra dentro de la regulación secundaria de la frecuencia eléctrica de la isla, resultando en una posible oscilación de este valor de emisiones (Guo e Yu (2004));(Jaime-Rojas y Latorre-Bayona (2004)). Para reflejar estas variaciones de potencia en el aprendizaje, se utiliza esta como una de las entradas de funcionamiento del modelo de GB.

La metodología general de este trabajo se desarrolla basándose en la necesidad de vigilar constantemente la gran cantidad de datos de la planta o la instalación, por lo tanto y de acuerdo con Coraddu *et al.*, estas estrategias de MBD requieren que se implementen modelos adecuados de ML para poder manejarlos (Coraddu *et al.* (2014)). Ya que el impacto del coste de mantenimiento es notable; como consecuencia, esto permite la necesidad de diseñar enfoques más eficaces basados en mantenimiento predictivo.

La aplicación y resultados de las simulaciones y estimaciones de este capítulo se ajustan a las conclusiones de sus estudios, donde tanto la aproximación con técnicas de ML como las simulaciones para estimación de degradación son claves para disponer de un sistema de soporte de respaldo y supervisión en el tiempo de un elemento industrial (Coraddu *et al.* (2014)).

95

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

135 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

135 / 181

Capítulo 4. Metodología general y modelo para el aprendizaje de eventos de mantenimiento

El procedimiento y arquitectura computacional planteada, coinciden con otros estudios en motores diesel y sistemas de propulsión naval, como los de Pontoppidan y Cipollini, que también aplicaban metodología de *ML* en casos donde los sistemas térmicos no pueden ser modelados fácilmente con una regla física a priori, debido a la gran cantidad de variables y datos que influyen en estos sistemas, similar al caso que se ha planteado en este trabajo sobre las emisiones de NOx (Pontoppidan *et al.* (2019));(Cipollini *et al.* (2018)). Por lo tanto, los resultados obtenidos del procedimiento computacional MBD.2 en este caso de aplicación permiten, la generación de modelos inteligentes basados en datos e información histórica.

En particular, el modelo desarrollado de *GB* para el motor diesel fue adaptado y entrenado mediante la configuración de sus hiperparámetros, según el algoritmo *Grid Search*, siguiendo un planteamiento similar al que realizaron Shin y Liu para modelar emisiones de NOx en motores diesel, donde además alcanzaron los mejores rendimientos de sus procedimientos, incluyendo tratamiento de VC por *k-fold* (Shin *et al.* (2020));(Liu *et al.* (2016));(Xiaojian *et al.* (2020)).

Coincidiendo con lo propuesto por Liu y a la vista de los buenos resultados de predicción del modelo de NOx desarrollado en este trabajo, este puede aplicarse como una medida de respaldo y de contraste, frente a la variable real medida por el equipo en campo.

Con relación a la metodología creada para medir el impacto en tareas de mantenimiento y revisiones, se dispone de un modelo que puede evaluar y entender de forma clara cómo se relacionan las diferentes intervenciones de mantenimiento sobre la variable del proceso. Por un lado utilizando los niveles de importancia y de afección de cada registro de mantenimiento, calculados por *MTImp* y por otro lado mediante el seguimiento y simulación de escenarios en tiempo real que marcan una línea base en el caso que no se hubiera realizado dicho mantenimiento para así observar su desviación positiva o negativa sobre el equipo.

Esto encaja con lo presentado por Rengasamy o Pelt, cuyos estudios planteaban

96

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

136 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

136 / 181

4.4. Discusión

que las predicciones basadas en datos pueden orientar las actividades de inspección y mejorar la eficiencia. Sin embargo y a diferencia de lo desarrollado anteriormente, citan que estos algoritmos de *ML* a menudo no son transparentes (Rengasamy *et al.* (2018));(Pelt *et al.* (2019)).

Por último, los métodos para medir el grado de importancia de un modelo de *GB*, como *SHAP* o *FeatImp*, utilizados en otros estudios específicos de emisiones ambientales y de verificación de entorno ambiental, dan respuesta en este tipo de métodos basados en arboles de decisión (Qadeer y Jeon (2019));(Lundberg y Lee (2017));(Park *et al.* (2020)).

Como se ha podido verificar en este Capítulo y con el desarrollo del algoritmo *SIM* para RNA. Estos métodos calculan los niveles de importancia para todo el DS, siendo capaces de evaluar y explicar correctamente de forma general, la relación de las diferentes variables de entrada, mayormente series cronológicas o temporales, con la variable objetivo.

Sin embargo, en este caso se observa que para la variable de conteo acumulativo no reseteable *MTWork2*, específica para el registro histórico de eventos de mantenimiento generales, los valores finales de importancias mediante *SHAP* y *FeatImp*, se clasifican totalmente de forma cronológica, penalizando a la variable más reciente, que en definitiva es la que su valor de conteo *MTWork2* tiene un valor menor, frente al verdadero nivel de importancia que pudiera tener este ME sobre la variable de salida. Además, para este caso específico de mantenimiento, se comprueba que estos métodos explicativos no permiten ajustar el rango de tiempos de estabilización que pueden producirse tras una intervención de mantenimiento y puesta en marcha del equipo en esta comparativa de importancias.

97

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

137 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

137 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

138 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

138 / 181

Conclusiones y trabajo futuro

Como parte de la necesaria y vital evolución digital del sector industrial hacia una línea basada en datos (*data-driven*), es clave un acercamiento al conocimiento y aplicación de métodos de aprendizaje de máquina e inteligencia artificial. De forma que se pueda transformar la gran cantidad de datos brutos disponibles, en una nueva materia prima de utilidad.

Dentro de la transición hacia una Industria 4.0, en esta investigación se ha dado respuesta a la aplicación de computación inteligente al sector de la generación eléctrica mediante el desarrollo de tres bloques principales que, de forma progresiva e interconectada, buscan definir una línea base de trabajo para una nueva metodología de mantenimiento basado en datos y dar respuesta a una serie de problemáticas reales de carácter industrial.

Se ha alcanzado el conocimiento para desarrollar un gemelo digital (DT) de un elemento industrial, modelando parámetros de una turbina de gas y un motor diesel de generación eléctrica junto con su entorno ambiental y su régimen de operación y mantenimiento insular. Además, los algoritmos que se han desarrollado en este trabajo, quedan agrupados dentro de una metodología computacional específica que conlleva procesos de aprendizaje de máquina configurados hacia el camino del mantenimiento prescriptivo. Haciendo posible enunciar una serie de conclusiones relevantes y exponer brevemente algunos detalles de la continuación de este trabajo hacia nuevas líneas futuras de estudio.

99

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

139 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Conclusiones y trabajo futuro

Análisis multivariante mediante ensemble de redes neuronales artificiales

- Se ha conseguido obtener un modelo que permite estimar una variable más compleja, como la concentración de calima en la superficie, a partir de series temporales básicas y medidas de forma online en la isla de Tenerife cuyos índices de rendimiento estadístico obtenidos muestran la eficacia del conjunto FFBP-NN propuesto, estructurado mediante modelos particionados con VC de $k=10$.

Los resultados se combinaron en un promedio final, después de haber sido validados utilizando un límite de control de rendimiento estadístico. Siendo esencial la normalización y el suavizado preliminar del conjunto de datos mediante una SMA.

- Los resultados del procedimiento CVC mejoran el rendimiento individual al combinar el promedio resultante de todas las RNA entrenadas, por lo tanto es vital disponer de un método de cálculo fiable para determinar sus IR y que pueda aplicarse a este tipo de arquitectura de ensemble con RNA.
- A diferencia de otros métodos de IR, en los que los cálculos se hacen en base a los pesos de la RNA y cuyos resultados en conjuntos de ensemble muestran una dispersión relevante y presentan gran aleatoriedad en cada nueva RNA entrenada, en este procedimiento se ha desarrollado un sistema híbrido *SIM*, que analiza el resultado de un nuevo método de entrenamiento basado en la eliminación de las variables de entrada de forma selectiva y que además reajusta la arquitectura de la RNA, fijando la inicialización de los pesos para evaluar su aprendizaje en términos del MAE. El grado de IR calculado con el método *SIM*, indicó correctamente dos de las variables clave que afectan a la concentración de arena: la humedad relativa (26,6%) y la temperatura del aire (34.65%). Frente a lo obtenido mediante el cálculo de *MGA*, donde aquí la temperatura ocupa la tercera posición (19.7%) y la

100

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

140 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

140 / 181

Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

humedad relativa la cuarta (19.9%). Sin embargo, el control de dispersión estadística realizado para el caso de *MGA*, basado en cálculo de los pesos, indica una alta variabilidad de sus resultados. Confirmando las conclusiones que se obtuvieron en otras investigaciones discutidas en este trabajo. Así pues, este método proporciona una alternativa real para el cálculo y la estimación de IR que puede generalizarse a cualquier tipo de problema en sistemas multivariantes modelados mediante RNA. Tanto para una configuración simple como para una arquitectura de ensemble.

Además, la información que se obtiene a través del cálculo de las IR utilizando *SIM*, puede utilizarse también para mejorar las predicciones o estimaciones de un modelo real, al eliminar variables irrelevantes.

Influencia del polvo sahariano y el mantenimiento sobre un compresor axial

- Los modelos obtenidos del sistema FFBP-NN CVC, tras haber sido sometidos a la recepción de nuevos datos que no conocía, han conseguido un buen rendimiento estadístico de salida en términos del rendimiento isentrópico del compresor de una turbina de gas. Asimismo, se han contrastado con otros modelos de regresión basados en árboles de decisión como *RF* o *GB*, obteniendo rendimientos en torno a un 10% y un 20% más bajos en comparación con los modelos basados en RNA.
- Se verifica que las estimaciones térmicas obtenidas de la turbina de gas, son fieles a las medidas reales, incluso a la hora de añadir una intervención real de mantenimiento sobre este elemento industrial, siendo capaces de predecir la evolución de su degradación por ensuciamiento en función del tiempo en servicio. Para esto ha sido vital, como parte del desarrollo de MBD, el desarrollo del algoritmo *MTWork1* que ha transformado esta información de mantenimiento en variables de entrada del modelo de *ML*.

101

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

141 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

141 / 181

Conclusiones y trabajo futuro

- El grado de importancia relativa tras aplicar el procedimiento computacional *SIM*, destaca las variables de potencia eléctrica de salida, mes del año, temperatura ambiente y el propio contador de arena sahariana. Esto ha permitido acotar las simulaciones a aquellas entradas con mayor grado de importancia y reducir así la dimensión del sistema de estudio.
- Del mismo modo, los estudios desarrollados en este trabajo han permitido verificar una relación directa entre la afección que tiene una variable climatología local, característica del clima de las Islas Canarias, y el ensuciamiento del compresor de una turbina de gas. Esta relación, definida en un contador de arena, ha permitido hacer simulaciones y predicciones que son fieles a la realidad y a la evolución de la degradación en el tiempo de un elemento rotativo industrial, en función del polvo sahariano y las horas de funcionamiento. Verificando su eficiencia en determinadas estaciones del año y bajo diferentes escenarios de funcionamiento. Lo que permite disponer de una foto de la turbina de gas en medio plazo, de cara a optimizar y seleccionar los intervalos ideales para realizar nuevas actividades de mantenimiento de limpieza.

Metodología general y modelo para el aprendizaje de eventos de mantenimiento

- A raíz de las diferentes aplicaciones que se han ido desarrollando a lo largo del trabajo, se establece un esquema general para una metodología de MBD. Este se ha revisado, utilizando el mismo planteamiento computacional de entorno ambiental, régimen de funcionamiento y la construcción de dos tipos de variables de mantenimiento. Ya que la metodología general de MBD necesitaba integrar una intervención recurrente sobre un equipo particular y una agrupación de diferentes trabajos que se realizan periódicamente.

Para asegurar la metodología, se ha optado por utilizar esta secuencia compu-

102

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

142 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

142 / 181

Líneas futuras de estudio

tacional pero introduciendo otro método de *ML* como es *GB*. Cuyos resultados son igualmente satisfactorios y nos habilitan para definir el procedimiento donde la combinación de este tipo de conjuntos de variables de entrada junto con la construcción de un ensemble promedio, aplicable a cualquier método de *ML* y supervisado por un control estadístico definido, permite dimensionar y modelar cualquier elemento industrial.

- En una primera fase de MBD, el primer algoritmo *SIM* debe utilizarse para explicar aquellas variables medidas de los modelos, pero no para evaluar los contadores calculados de mantenimiento *MTWork*. En este punto es donde debe aplicarse el método específico de importancias *MTImp*, el cual ha alcanzado el aprendizaje de un modelo de *ML* de forma autónoma, para estudiar y clasificar la importancia de las intervenciones de mantenimiento en instantes específicos, teniendo en cuenta por tanto, periodos de estabilidad y ajustes del equipo tras su puesta en servicio después de haber realizado diferentes intervenciones de mantenimiento.
- Finalizando la metodología MBD en las simulaciones planteadas, se realizó un nuevo tipo de simulación sobre tres casos reales de mantenimiento de revisión de un motor diesel para comprobar que el sistema es capaz de proyectar, además de estimaciones, un supuesto estado de mantenimiento no realizado en tiempo real sobre las emisiones del motor. Por lo que este escenario se utilizó para observar, respecto a lo esperado, el grado de mejora que existió en dos de los casos y una intervención donde se identificó un empeoramiento de la variable objetivo.

Líneas futuras de estudio

- Gran parte de los estudios que se investigaron en este trabajo con relación a las turbinas de gas, destacan que la reducción de la eficiencia del compresor lleva asociado un aumento en el consumo de combustible (Meher-Homji y Bromley (2004));(Janawitz *et al.* (2010));(Tarabrin *et al.* (2014)). Meher y

103

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

143 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

143 / 181

Conclusiones y trabajo futuro

Bromley añaden además otra pérdida asociada a la degradación del compresor y que afecta directamente a las emisiones ambientales de la turbina, en particular a los NOx. Donde estimaban que tan solo un 1% de disminución del rendimiento isentrópico del compresor (μ_{co}), supone un aumento de los NOx en unos 25 mg/m^3 (Meher-Homji y Bromley (2004)).

Por lo tanto, en la primera aproximación computacional para crear la simulación de escenarios futuros sobre una variable térmica de eficiencia del compresor de una turbina de gas, se debe asociar un factor de coste, enlazando una cadena de modelos de RNA para la estimación futura del consumo de combustible en función de la degradación. De forma que al modelar completamente la turbina de gas, el sistema pueda unir el conocimiento térmico total de la turbina junto a las condiciones de mantenimiento y la estimación en términos económicos para proponer un nuevo calendario optimizado de planificación del mantenimiento.

- Como parte del proceso de mejora continua de la metodología, se debe modificar el esquema computacional de MBD, añadiendo nuevas variables derivadas de fuentes de información que provengan del texto de informes detallados con las tareas de los mantenimientos de revisión realizadas y otras anomalías registradas en el sistema manual de averías, que se realiza durante todo el tiempo de funcionamiento del elemento industrial. Para esto, se aplicarán métodos basados en *Natural Language Processing* (NLP) dentro del procedimiento computacional propuesto y con el respaldo de un diccionario técnico creado por máquina, que sirva de línea de referencia para el aprendizaje y reconocimiento de este tipo de información.

104

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

144 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

144 / 181

Apéndice A - Código representativo del Capítulo 2

En este apéndice se muestra un extracto con las líneas de código de los procedimientos más representativos, desarrollados para la realización del Capítulo 2 y utilizando fragmentos de MATLAB:

FFBP-NN CVC

```
1 % Suavizado SMA y Filtrado
2 %SMA DS1
3 input=movavg(input,4320);
4 output=movavg(output,4320);
5 %SMA DS2
6 INPUT17=movavg(INPUT17,4320);
7 OUTPUT17=movavg(OUTPUT17,4320);
8 %Nan DS2
9 INPUT17(isnan(INPUT17))=0;
10 OUTPUT17(isnan(OUTPUT17))=0;
11 %Nan DS1
12 input(isnan(input))=0;
13 output(isnan(output))=0;
14 %Ensemble FFBPNN-CVC
15 % Definir valor del k-fold
16 k=10;
17 nodos1=size(input);
18 nodos1=(nodos1(:,2));
19 nodos2=size(output);
20 nodos2=(nodos2(:,2));
21 % Cálculo del número de L1 o capas ocultas
22 h0=sqrt((nodos2+2)*nodos1)+2*sqrt((nodos1/(nodos2+2)));
23 h0=round(h0);
24 h1=h0;
25 h2=h0+h0/2;
26 h2=round(h2);
27 % Transponer
28 input=input';
29 output=output';
30 INPUT17=INPUT17';
31 OUTPUT17=OUTPUT17';
```

105

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

145 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

145 / 181

Apéndice A - Código representativo del Capítulo 2

```
32 regresion17=0;
33 accuracy17=0;
34 c=mean0+3*sqrt(mean0);
35 ite=0;
36 tic
37 % Validación FFBPNN-CVC
38 while regresion17<0.7 && accuracy17<c
39 % Índices de VC para el valor dado de k
40 cvFolds = crossvalind('Kfold',output, k);
41 MAE=[1:k];
42 reg=[1:k];
43 for i = 1:k
44 % obtener división de conjunto de test utilizando índice VC
45 testIdx = (cvFolds == i);
46 % obtener división de conjunto de entrenamiento utilizando índice VC
47 trainIdx = ~testIdx ;
48 trInd=find(trainIdx);
49 tstInd=find(testIdx);
50 net=feedforwardnet(hl,'trainbr');
51 net.layers(1).transferFcn='tansig';
52 net.layers(2).transferFcn='purelin';
53 net.layers(1).initFcn='initnw';
54 net.layers(2).initFcn='initnw';
55 net.trainParam.epochs=300;
56 net.trainParam.goal=-2;
57 net.divideFcn = 'divideind';
58 net.divideParam.trainInd=trInd;
59 net.divideParam.testInd=tstInd;
60 net.inputs(1).processFcns = {'mapminmax'};
61 net.outputs(2).processFcns = {'mapminmax'};
62 net.performFcn='mse';
63 net.plotFcns={'plotperform', 'plottrainstate', 'ploterrhist', 'plotregression'};
64 % Entrenamiento RNA DSI
65 [net, tr]=train(net, input, output);
66 % Evaluar RNA DSI
67 test = net(input);
68 test(:, test<0)=0;
69 test=round(test, 0);
70 errors = gsubtract(output, test);
71 % Rendimiento RNA DSI
72 MAE(:, i) = mae(errors);
73 reg(:, i) = regression(test, output);
74 % Guardar la RNA DSI
75 eval(['NET' num2str(i) '=net;'])
76 eval(['test' num2str(i) '=test;'])
77 end
78 % Evaluar FFBPNN-CVC
79 test=0;
80 for i = 1:1:k
81 eval(['test' num2str(i) '=test;']);
82 end
83 test=test./k;
84 % Rendimiento DSI FFBPNN-CVC
85 accuracy=mean(MAE);
86 accuracy2=std(MAE);
87 regresion=regression(output, test);
88 MAE=mean(MAE);
89 accuracy=mean(MAE);
90 accuracy2=std(MAE);
```

106

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

146 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

146 / 181

```
91 regresion=regression(output, test);
92 MAE17=[1:k];
93 MAEM=mean(MAE);
94 MAE17=mean(MAE17);
95 reg17=[1:k];
96 for i = 1:k
97 % Guardar la RNA DS2
98 eval(['TEST17' num2str(i) '=NET' num2str(i) '(INPUT17);'])
99 eval(['TEST17' num2str(i) '(,TEST17' num2str(i) '<0)=0;'])
100 eval(['TEST17' num2str(i) '=round,(TEST17' num2str(i) ',0);'])
101 eval(['errores,=subtract(OUTPUT17,TEST17' num2str(i) ');'])
102 % Rendimiento DS2 FFBPNN-CVC
103 MAE17(:,i) = mac(errores);
104 eval(['reg17(:,i)=regression(TEST17' num2str(i) ',OUTPUT17);'])
105 end
106 TEST17=0;
107 for i= 1:1:k
108 eval(['TEST17,=,TEST17' num2str(i) '+,TEST17,;']);
109 end
110 TEST17=TEST17./k;
111 TEST17=round(TEST17, 0);
112 % Rendimiento DS2 FFBPNN-CVC
113 regresion17=regression(TEST17,OUTPUT17);
114 accuracy17=mean(MAE17);
115 accuracy172=std(MAE17);
116 ite=ite+1;
117
118 end
119 t = toc;
120 t=t/3600;
```

SIM

```
1 % Variables requeridas del modelo FFBP-NN CVC: cvFolds,h1,MAE0,input,output
2 % Numero de entradas que quieres analizar y número de RNA que tienes (CVC)
3 j=5;
4 y=10;
5 % Preparar variables
6 RIT=[1:j];
7 RIT100T=[1:j];
8 MAEM=[1:j];
9 SIDMAET=[1:j];
10 MAE0=MAE17;
11 for x= 1:1:j
12 % Se eliminan las filas de cada variable de entrada
13 input2=input;
14 input([x],:) = [];
15 MAE=[1:y];
16 % Comienza el proceso para definir nuevas RNA
17 for i= 1:1:y
18 eval(['BNET' num2str(i) '=,NET' num2str(i) ');'])
19 % Se obtienen los pesos iniciales de la RNA
20 eval(['BNET' num2str(i) '=revert(NET' num2str(i) ');'])
21 % Comienza el proceso para definir nuevas RNA
22 % Se guardan Pesos sinápticos iniciales hacia capa oculta
23 eval(['IW' num2str(i) '=BNET' num2str(i) '.IW;'])
24 % Se guardan Pesos sinápticos iniciales de capa oculta a salida
25 eval(['LW' num2str(i) '=BNET' num2str(i) '.LW;'])
```

107

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

147 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

147 / 181

Apéndice A - Código representativo del Capítulo 2

```
26 %Se guardan Bías iniciales
27 eval(['B' num2str(i) 'E=NET' num2str(i) '.b;'])
28 %Guardar pesos iniciales
29 eval(['BIW0' num2str(i) num2str(x) '=IW' num2str(i) ';'])
30 eval(['BLW0' num2str(i) num2str(x) '=LW' num2str(i) ';'])
31 eval(['BB0' num2str(i) num2str(x) '=B' num2str(i) ';'])
32 %Adaptar nuevas matrices de pesos sin la variable x eliminada
33 eval(['TW' num2str(i) '{1,~1}(:,x),~[,] ;'])
34 eval(['BW' num2str(i) num2str(x) '=IW' num2str(i) ';'])
35 %Se utiliza la misma partición de cvFolds de VC generada en la RNA original
36 testIdx = (cvFolds == i);
37 trainIdx = ~testIdx ;
38 trInd=find(trainIdx);
39 tstInd=find(testIdx);
40 trainmethod='trainbr';
41 netinicio=feedforwardnet(h1,trainmethod);
42 netinicio.layers(1).transferFcn='tansig';
43 netinicio.layers(2).transferFcn='purelin';
44 %Se realiza un primer entreno de 1 época sólo para guardar la configuración de las variables
45 netinicio.trainParam.epochs=1;
46 netinicio.divideFcn = 'divideind';
47 netinicio.divideParam.trainInd=trInd;
48 netinicio.divideParam.testInd=tstInd;
49 netinicio.inputs(1).processFcns = {'mapminmax'};
50 netinicio.outputs(2).processFcns = {'mapminmax'};
51 %Si hace falta transponer
52 %output=output';
53 %input=input';
54 %Se realiza el entrenamiento de cada nueva RNA
55 netinicio=train(netinicio,input,output);
56 netinicio.trainParam.epochs=100;
57 eval(['netinicio.IWc=IW' num2str(i) ';']) %input weights modificada la IW sin
58 eval(['netinicio.LWc=LW' num2str(i) ';']) %layer weights
59 eval(['netinicio.b_c=B' num2str(i) ';']) %bias
60 X=netinicio.IW;
61 X=X(1, 1);
62 %Se comprueban que son los mismos pesos
63 eval(['check=IW' num2str(i) '{1,~1} ;'])
64 if X==check
65 netinicio.divideFcn = 'divideind';
66 netinicio.divideParam.trainInd=trInd;
67 netinicio.divideParam.testInd=tstInd;
68 netinicio.inputs(1).processFcns = {'mapminmax'};
69 netinicio.outputs(2).processFcns = {'mapminmax'};
70 netinicio=train(netinicio,input,output);
71 testinicio = netinicio(input);
72 testinicio(:,testinicio<0)=0;
73 testinicio=round(testinicio,0);
74 erroresinicio=gsubtract(testinicio,output);
75 % Analisis MAE
76 MAE(:,i) = mae(erroresinicio);
77 else
78 f = msgbox('Error', 'Error','error');
79 end
80 end
81 input=input2;
82 %Dispersión MAE
83 accuracy=mean(MAE);
84 accuracy2=std(MAE);
```

108

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

148 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

```
85 % Resultados SIM
86 eval(['MAE' num2str(x) '=MAE;'])
87 eval(['MAEM' num2str(x) '=accuracy;'])
88 eval(['SIDMAE' num2str(x) '=accuracy2;'])
89 eval(['RI' num2str(x) '=abs(MAE-MAEM) num2str(x) ';'])
90 eval(['MAEM(:,x)=-MAEM' num2str(x) ';'])
91 eval(['RIT(:,x)=-RI' num2str(x) ';'])
92 eval(['SIDMAET(:,x)=-SIDMAE' num2str(x) ';'])
93 end
94 RITOT=sum(RIT);
95 for i= 1:1;j
96     eval(['RIT100T(:,i)=-RI' num2str(i) '/RITOT*100;']);
97 end
98 %Control de calidad SIM
99 CV=(SIDMAET./MAEM)*100;
```

109

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

149 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

149 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

150 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

150 / 181

Apéndice B - Código representativo del Capítulo 3

En este apéndice se muestra un extracto con las líneas de código de los procedimientos más representativos, desarrollados para la realización del Capítulo 3 y utilizando fragmentos de MATLAB y Python:

FFBP-NN CVC

```
1 %NaN DS2
2 INPUT17(isnan(INPUT17))=0;
3 OUTPUT17(isnan(OUTPUT17))=0;
4 %NaN DS1
5 input(isnan(input))=0;
6 output(isnan(output))=0;
7 %Ensemble FFBPNN-CVC
8 % Sustituir para el caso de la comparativa CVC k-fold
9 k=10; %k=5; k=15;
10 nodos1=size(input);
11 nodos1=(nodos1(:,2));
12 nodos2=size(output);
13 nodos2=(nodos2(:,2));
14 % Cálculo del número de L1 o capas ocultas
15 h0=sqrt((nodos2+2)*nodos1)+2*sqrt((nodos1/(nodos2+2)));
16 h0=round(h0);
17 h1=h0;
18 h2=h0+h0/2;
19 h2=round(h2);
20 % Transponer
21 input=input';
22 output=output';
23 INPUT17=INPUT17';
24 OUTPUT17=OUTPUT17';
25 regresion17=0;
26 accuracy17=0;
27 c=mean0+3*sqrt(mean0);
28 ite=0;
29 tic
30 % Validación FFBPNN-CVC
31 while regresion17 < 0.7 && accuracy17 <
```

111

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

151 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

151 / 181

Apéndice B - Código representativo del Capítulo 3

```
32 %Indices de VC para el valor dado de k
33 cvFolds = crossvalind('Kfold',output, k);
34 MAE=[1:k];
35 reg=[1:k];
36 for i = 1:k
37 %Obtener división de conjunto de test utilizando indice VC
38 testIdx = (cvFolds == i);
39 %Obtener división de conjunto de entrenamiento utilizando indice VC
40 trainIdx = ~testIdx ;
41 trInd=find(trainIdx);
42 tstInd=find(testIdx);
43 net=feedforwardnet(hl,'trainbr');
44 net.layers{1}.transferFcn='tansig';
45 net.layers{2}.transferFcn='purelin';
46 net.layers{1}.initFcn='initnw';
47 net.layers{2}.initFcn='initnw';
48 net.trainParam.epochs=300;
49 net.trainParam.goal=1e-2;
50 net.divideFcn = 'divideind';
51 net.divideParam.trainInd=trInd;
52 net.divideParam.testInd=tstInd;
53 net.inputs{1}.processFcn = {'mapminmax'};
54 net.outputs{2}.processFcn = {'mapminmax'};
55 %Sustituir para el caso de la comparativa con la RNA individual 90/10
56 %net.divideMode='sample';
57 %net.divideParam.trainRatio=0.70;
58 %net.divideParam.valRatio=0.15;
59 %net.divideParam.testRatio=0.15;
60 net.performFcn='mse';
61 net.plotFcn={'plotperform', 'plottrainstate', 'ploterrhist', 'plotregression'};
62 %Entrenamiento RNA DSI
63 [net, tr]=train(net, input, output);
64 %Evaluar RNA DSI
65 test = net(input);
66 test (:, test<0)=0;
67 test=round(test,0);
68 errors = gsubtract(output, test);
69 %Rendimiento RNA DSI
70 MAE(:, i) = mae(errors);
71 reg(:, i) = regression(test, output);
72 %Guardar la RNA DSI
73 eval(['NET' num2str(i) '=net;'])
74 eval(['test' num2str(i) '=test;'])
75 end
76 %Evaluar FFBPNN-CVC
77 test=0;
78 for i= 1:1:k
79 eval(['test' num2str(i) '=test;'])
80 end
81 test=test./k;
82 %Rendimiento DSI FFBPNN-CVC
83 accuracy=mean(MAE);
84 accuracy2=std(MAE);
85 regresion=regression(output, test);
86 MAEM=mean(MAE);
87 accuracy=mean(MAE);
88 accuracy2=std(MAE);
89 regresion=regression(output, test);
90 MAEL7=[1:k];
```

112

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

152 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

152 / 181

```

91 MAEM=mean(MAE);
92 MAEI7M=mean(MAEI7);
93 reg17=[1:k];
94 for i = 1:k
95 % Guardar la RNA DS2
96 eval(['TEST17' num2str(i) '=NET' num2str(i) '(INPUT17);'])
97 eval(['TEST17' num2str(i) '(:,TEST17' num2str(i) '<0)=0;'])
98 eval(['TEST17' num2str(i) '=round,(TEST17' num2str(i) ',0);'])
99 eval(['errores' _=subtract(OUTPUT17,TEST17' num2str(i) ');'])
100 % Rendimiento DS2 FFBFNN-CVC
101 MAEI7(:,i) = mae(errores);
102 eval(['reg17(:,i)=regression(TEST17' num2str(i) ',OUTPUT17);'])
103 end
104 TEST17=0;
105 for i= 1:1:k
106 eval(['TEST17' _=TEST17' num2str(i) '+_TEST17,;']);
107 end
108 TEST17=TEST17./k;
109 TEST17=round(TEST17,0);
110 % Rendimiento DS2 FFBFNN-CVC
111 regresion17=regression(TEST17,OUTPUT17);
112 accuracy17=mean(MAEI7);
113 accuracy172=std(MAEI7);
114 ite=ite+1;
115
116 end
117 t = toc;
118 t=t/3600;
    
```

SIM

```

1 % Variables requeridas del modelo FFBP-NN CVC: cvFolds,h1,MAE0,input,output
2 % Numero de entradas que quieres analizar y número de RNA que tienes en CVC
3 j=8;
4 y=10;
5 % Preparar variables
6 RIT=[1:j];
7 RIT100T=[1:j];
8 MAEMT=[1:j];
9 STDMAET=[1:j];
10 MAE0=MAEI71;
11 for x= 1:1:j
12 % Se eliminan las filas de cada variable de entrada
13 input2=input;
14 input([x],:) = [];
15 MAE=[1:y];
16 % Comienza el proceso para definir nuevas RNA
17 for i= 1:1:y
18 eval(['BNET' num2str(i) '=NET' num2str(i) ');'])
19 % Se obtienen los pesos iniciales de la RNA
20 eval(['BNET' num2str(i) '=revert(NET' num2str(i) ');'])
21 % Comienza el proceso para definir nuevas RNA
22 % Se guardan Pesos sinápticos iniciales hacia capa oculta
23 eval(['IW' num2str(i) '=BNET' num2str(i) '.IW;'])
24 % Se guardan Pesos sinápticos iniciales de capa oculta a salida
25 eval(['LW' num2str(i) '=BNET' num2str(i) '.LW;'])
26 % Se guardan Bías iniciales
27 eval(['B' num2str(i) '=BNET' num2str(i) '.b;'])
    
```

113

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
 UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
 UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

153 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
 UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
 UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Apéndice B - Código representativo del Capítulo 3

```
28 %Guardar pesos iniciales
29 eval(['BIW0' num2str(i) num2str(x) '=IW' num2str(i) ';' ])
30 eval(['BLW0' num2str(i) num2str(x) '=LW' num2str(i) ';' ])
31 eval(['BB0' num2str(i) num2str(x) '=B' num2str(i) ';' ])
32 %Adaptar nuevas matrices de pesos sin la variable x eliminada
33 eval(['TW' num2str(i) '{1,1}(:,x),=[,] ;' ])
34 eval(['BW' num2str(i) num2str(x) '=IW' num2str(i) ';' ])
35 %Se utiliza la misma particion de cvFolds de VC generada en la RNA original
36 testIdx = (cvFolds == i);
37 trainIdx = ~testIdx ;
38 trInd=find(trainIdx);
39 tstInd=find(testIdx);
40 trainmethod='trainbr';
41 netinicio=feedforwardnet(h1,trainmethod);
42 netinicio.layers(1).transferFcn='tansig';
43 netinicio.layers(2).transferFcn='purelin';
44 %Se realiza un primer entreno de 1 época sólo para guardar la configuración de las variables
45 netinicio.trainParam.epochs=1;
46 netinicio.divideFcn = 'divideind';
47 netinicio.divideParam.trainInd=trInd;
48 netinicio.divideParam.testInd=tstInd;
49 netinicio.inputs{1}.processFcns = {'mapminmax'};
50 netinicio.outputs{2}.processFcns = {'mapminmax'};
51 %Si hace falta transponer
52 %output=output';
53 %input=input';
54 %Se realiza el entrenamiento de cada nueva RNA
55 netinicio=train(netinicio ,input ,output);
56 netinicio.trainParam.epochs=100;
57 eval(['netinicio .IW_=_IW' num2str(i) ';' ]) %input weights modificada la IW sin
58 eval(['netinicio .LW_=_LW' num2str(i) ';' ]) %layer weights
59 eval(['netinicio .b_=_B' num2str(i) ';' ]) %bias
60 X=netinicio.IW;
61 X=X(1, 1);
62 %Se comprueban que son los mismos pesos
63 eval(['checkIW' num2str(i) '{1,1} ;' ])
64 if X==check
65     netinicio.divideFcn = 'divideind';
66     netinicio.divideParam.trainInd=trInd;
67     netinicio.divideParam.testInd=tstInd;
68     netinicio.inputs{1}.processFcns = {'mapminmax'};
69     netinicio.outputs{2}.processFcns = {'mapminmax'};
70     netinicio=train(netinicio ,input ,output);
71     testinicio = netinicio (input);
72     testinicio (:, testinicio < 0) = 0;
73     testinicio = round(testinicio , 0);
74     erroresinicio = gsubtract(testinicio ,output);
75 % Analisis MAE
76 MAE(:,i) = mae(erroresinicio);
77 else
78     f = msgbox('Error', 'Error', 'error');
79 end
80 end
81 input=input2;
82 %Dispersión MAE
83 accuracy=mean(MAE);
84 accuracy2=std(MAE);
85 %Resultados SIM
86 eval(['MAE' num2str(x) '=MAE;' ])
```

114

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA
Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

154 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA
María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

154 / 181

```
87     eval(['MAEM' num2str(x) 'accuracy;'])
88     eval(['SIDMAE' num2str(x) 'accuracy2;'])
89     eval(['RI' num2str(x) 'abs(MAE-MEM) num2str(x) ;'])
90     eval(['MAEM(:,x),=MAEM' num2str(x) ' ;'])
91     eval(['RIT(:,x),=RI' num2str(x) ' ;'])
92     eval(['SIDMAET(:,x),=SIDMAE' num2str(x) ' ;'])
93 end
94 RITOT=sum(RIT);
95 for i= 1:1:j
96     eval(['RIT100T(:,i),=RI' num2str(i) '/RITOT*100;']);
97 end
98 %Control de calidad SIM
99 CV=(SIDMAET./MAEM)*100;
```

RF

```
1 import pandas as pd
2 from sklearn.model_selection import train_test_split
3 from sklearn.datasets import make_regression
4 from sklearn.ensemble import RandomForestRegressor
5 from sklearn.metrics import mean_absolute_error
6 from sklearn.model_selection import GridSearchCV
7 from sklearn.metrics import r2_score
8 from sklearn.metrics import make_scorer
9 import numpy as np
10 import scipy.io
11 # Cargar DS original de Matlab ( 8 entradas y 1 salida)
12 mat = scipy.io.loadmat('RF.mat')
13 # DS1
14 df3_TR=mat["input"];
15 df32_TR=mat["output"];
16 # DS2
17 df3_TS=mat["INPUT17"];
18 df32_TS=mat["OUTPUT17"];
19 # Preparar DataFrame
20 df3_TS=np.transpose(df3_TS)
21 df32_TS=np.transpose(df32_TS)
22 df32_TR = pd.DataFrame(df32_TR)
23 df32_TS = pd.DataFrame(df32_TS)
24 df3_TS = pd.DataFrame(df3_TS)
25 df3_TR = pd.DataFrame(df3_TR)
26 # Comenzar separación
27 X_train={}
28 X_test={}
29 y_train={}
30 y_test={}
31 # Construir conjuntos de entrenamiento y test
32 X_train, X_test, y_train, y_test = train_test_split(df3_TR, df32_TR, test_size = 0.3)
33 # Crear la rejilla de parámetros en un rango determinado. Seleccionar manualmente el rango
34 param_grid = {
35     'bootstrap': [True],
36     'max_depth': range(80,120,10)
37     'n_estimators': range(200,600,80)
38 }
39 # Crear el modelo base de RF
40 rf = RandomForestRegressor()
41 # Iniciar la búsqueda en rejilla de los parámetros óptimos RF
42 grid_search = GridSearchCV(estimator = rf, param_grid = param_grid, cv = 10)
```

115

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA
Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

155 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA
María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

155 / 181

Apéndice B - Código representativo del Capítulo 3

```
43 grid_result=grid_search.fit(X_train, np.ravel(y_train, order='C'))
44 best_params = grid_result.best_params_
45 # Cargar los mejores parámetros encontrados
46 rfr = RandomForestRegressor(max_depth=best_params["max_depth"], n_estimators=best_params["n_estimators"])
47 # Comenzar predicciones de los resultados del modelo RF
48 RFy_pred_TS = rfr.predict(df3_TS)
49 RFy_pred_TR = rfr.predict(df3_TR)
50 y_test_TS=df32_TS
51 y_test_TR=df32_TR
52 # Rendimiento modelo final RF
53 r2=r2_score(y_test_TS, RFy_pred_TS)
54 r2TR=r2_score(y_test_TR, RFy_pred_TR)
55 mae=mean_absolute_error(y_test_TS, RFy_pred_TS)
56 maeTR=mean_absolute_error(y_test_TR, RFy_pred_TR)
```

GB

```
1 import pandas as pd
2 from sklearn.model_selection import train_test_split
3 from sklearn.datasets import make_regression
4 import lightgbm as lgb
5 from sklearn.metrics import mean_absolute_error
6 from sklearn.model_selection import GridSearchCV
7 from sklearn.metrics import r2_score
8 from sklearn.metrics import make_scorer
9 import numpy as np
10 import scipy.io
11 # Cargar DS original de Matlab ( 8 entradas y 1 salida)
12 mat = scipy.io.loadmat('GB.mat')
13 # DS1
14 df3_TR=mat["input"];
15 df32_TR=mat["output"];
16 # DS2
17 df3_TS=mat["INPUT17"];
18 df32_TS=mat["OUTPUT17"];
19 # Preparar DataFrame
20 df3_TS=np.transpose(df3_TS)
21 df32_TS=np.transpose(df32_TS)
22 df32_TR = pd.DataFrame(df32_TR)
23 df32_TS = pd.DataFrame(df32_TS)
24 df3_TS = pd.DataFrame(df3_TS)
25 df3_TR = pd.DataFrame(df3_TR)
26 # Comenzar separación
27 X_train={}
28 X_test={}
29 y_train={}
30 y_test={}
31 # Construir conjuntos de entrenamiento y test
32 X_train, X_test, y_train, y_test = train_test_split(df3_TR, df32_TR, test_size = 0.3)
33 # Crear la rejilla de parámetros en un rango determinado. Seleccionar manualmente el rango
34 param_grid = {
35     'bootstrap': [True],
36     'max_depth': range(80,120,10)
37     'n_estimators':range(200,600,80)
38 }
39 # Crear el modelo base de GB
40 for i in range(0, nsplit):
41     model = lgb.LGBMRegressor(
```

116

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

156 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

156 / 181

```
42     boosting_type="gbdt",
43     objective='regression',
44     random_state=10,
45     n_estimators=300,
46     num_leaves=50,
47     max_depth=8,
48     feature_fraction=0.5,
49     bagging_fraction=0.8,
50     bagging_freq=15,
51     learning_rate=0.01,
52     )
53     score_func = make_scorer(mean_absolute_error, greater_is_better=False)
54     params_opt = {'n_estimators':range(200, 600, 80), 'num_leaves':range(20,60,10)}
55     # Iniciar la búsqueda en rejilla de los parámetros óptimos GB
56     gridSearchCV = GridSearchCV(estimator = model, param_grid = params_opt, scoring=score_func, cv=10)
57     gridSearchCV.fit(X_train, y_train)
58     a=gridSearchCV.best_params_
59     b=(a.get('n_estimators'))
60     c=(a.get('num_leaves'))
61     # Entrenar modelo final ajustado de GB
62     gbm = lgb.LGBMRegressor(
63     boosting_type="gbdt",
64     objective='regression',
65     random_state=10,
66     n_estimators=b,
67     num_leaves=c,
68     feature_fraction=0.9,
69     bagging_fraction=0.8,
70     bagging_freq=5,
71     learning_rate=0.05,
72     )
73     # Comenzar predicciones de los resultados del modelo GB
74     LGBy_pred_TS = gbm.predict(df3_TS)
75     LGBy_pred_TR = gbm.predict(df3_TR)
76     y_test_TS=df32_TS
77     y_test_TR=df32_TR
78     # Rendimiento modelo final GB
79     r2=r2_score(y_test_TS, LGBy_pred_TS)
80     r2TR=r2_score(y_test_TR, LGBy_pred_TR)
81     mae=mean_absolute_error(y_test_TS, LGBy_pred_TS)
82     maeTR=mean_absolute_error(y_test_TR, LGBy_pred_TR)
```

117

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

157 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

157 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

158 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

158 / 181

Apéndice C - Código representativo del Capítulo 4

En este apéndice se muestra un extracto con las líneas de código de los procedimientos más representativos, desarrollados para la realización del Capítulo 4 y utilizando fragmentos de Python:

CVC GB

```
1 import numpy as np
2 import pandas as pd
3 from sklearn.model_selection import train_test_split
4 from sklearn.datasets import make_regression
5 from sklearn.metrics import mean_absolute_error
6 from sklearn.model_selection import GridSearchCV
7 from sklearn.model_selection import RandomizedSearchCV
8 from sklearn.metrics import r2_score
9 import lightgbm as lgb
10 from sklearn.model_selection import KFold
11 from sklearn.metrics import make_scorer
12 from sklearn.feature_selection import RFE
13 import pickle
14 import shap
15 # Lectura de datos hoja de entrada .xlsx
16 # DS1
17 df_TR = pd.read_excel('VARIABLES_MIO_TR_MOTOR_DIESEL_-_GD01.xlsx', sheet_name='Hoja1')
18 # DS2
19 df_TS = pd.read_excel('VARIABLES_MIO_TS_MOTOR_DIESEL_-_GD01.xlsx', sheet_name='Hoja1')
20 # Proceso de filtrado valores no numéricos
21 for i in [i for i in list(range(len(df_TR.columns))) if i not in range(0,281)]:
22     df_TR.iloc[:,i]=pd.to_numeric(df_TR.iloc[:,i], errors='coerce')
23     df_TR=df_TR.dropna()
24 for i in [i for i in list(range(len(df_TS.columns))) if i not in range(0,281)]:
25     df_TS.iloc[:,i]=pd.to_numeric(df_TS.iloc[:,i], errors='coerce')
26     df_TS=df_TS.dropna()
27 # Valores negativos a 0
28 for i in [i for i in list(range(len(df_TR.columns))) if i not in range(0,281)]:
29     df_TR.iloc[:,i][df_TR.iloc[:,i] < 0] = 0
30 for i in [i for i in list(range(len(df_TS.columns))) if i not in range(0,281)]:
31     df_TS.iloc[:,i][df_TS.iloc[:,i] < 0] = 0
```

119

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

159 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

159 / 181

Apéndice C - Código representativo del Capítulo 4

```
32 # Eliminar valores final Nan
33 df_TR=df_TR.dropna()
34 df_TS=df_TS.dropna()
35 df2_TR=df_TR
36 df2_TS=df_TS
37 # Filtrado manual por Potencia eléctrica generada (kWe)
38 df3_TR =df2_TR[(df2_TR.POWER_OUT>17500)]
39 df3_TR =df3_TR[df3_TR.POWER_OUT<=19500]
40 df3_TS =df2_TS[(df2_TS.POWER_OUT>17500)]
41 df3_TS =df3_TS[df3_TS.POWER_OUT<=19500]
42 df4_TR= df3_TR['FECHA']
43 df4_TS= df3_TS['FECHA']
44 # Agrupación de variables de entrada y salida del DS original
45 col_mtoNH=['MT_O10000_2016_OH', 'MT_O12500_2017_OH', 'MT_O15000_2017_OH', 'MT_O2500_2017_OH', and \
46 'MT_O5000_2018_OH', 'MT_O7500_2018_OH', 'MT_O10000_2018_OH', 'MT_O12500_2019_OH']
47 colbasic=[ 'POWER_OUT', 'REL_HUMIDITY', 'ATM_PRESSURE', 'AMB_TEMP', 'RFM']
48 col_targetNOX=['EMNOX']
49 # Definición de variables iniciales y otros ajustes
50 nsplit=10
51 X1_TR=df3_TR[colbasic+col_mtoNH+col_targetNOX]
52 X1_TS=df3_TS[colbasic+col_mtoNH+col_targetNOX]
53 X1_TR= X1_TR.reset_index(drop=True)
54 X1_TS= X1_TS.reset_index(drop=True)
55 X1_TR=X1_TR.dropna()
56 X1_TS=X1_TS.dropna()
57 kf = KFold(n_splits=nsplit, shuffle = True, random_state=None)
58 kf.get_n_splits(X1_TR[col_mtoNH+colbasic],X1_TR[col_targetNOX])
59 X_train={}
60 X_test={}
61 y_train={}
62 y_test={}
63 model={}
64 LGBY_pred_TS={}
65 LGBY_pred_TR={}
66 y_test_TS={}
67 y_test_TR={}
68 mae={}
69 r2={}
70 loaded_model={}
71 feat_imp={}
72 # Construccion de ensemble CVC para k-fold en [Xtrain,Xtest] & [Ytrain,Ytest]
73 for i in range(0,nsplit):
74     for train_index, test_index in kf.split(X1_TR[col_mtoNH+colbasic],X1_TR[col_targetNOX]):
75         X_train[i], X_test[i] = X1_TR[col_mtoNH+colbasic].loc[train_index], X1_TR[col_mtoNH+colbasic].loc[test_index]
76         y_train[i], y_test[i] = X1_TR[col_targetNOX].loc[train_index], X1_TR[col_targetNOX].loc[test_index]
77 # Definir modelo inicial de GB
78 for i in range(0,nsplit):
79     model = lgb.LGBMRegressor(
80         boosting_type="gbdt",
81         objective="regression",
82         random_state=10,
83         n_estimators=300,
84         num_leaves=50,
85         max_depth=8,
86         feature_fraction=0.5,
87         bagging_fraction=0.8,
88         bagging_freq=15,
89         learning_rate=0.01,
90     )
```

120

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

160 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

160 / 181

```

91     score_func = make_scorer(mean_absolute_error, greater_is_better=False)
92 # Crear la rejilla de parámetros en un rango determinado. Seleccionar manualmente el rango
93 params_opt = {'n_estimators':range(200, 600, 80), 'num_leaves':range(20,60,10)}
94 gridSearchCV = GridSearchCV(estimator = model, param_grid = params_opt, scoring=score_func, cv=10)
95 gridSearchCV.fit(X_train[i],y_train[i])
96 # Almacenar los mejores hiperparámetros para crear modelo definitivo
97 a=gridSearchCV.best_params_
98 b=(a.get('n_estimators'))
99 c=(a.get('num_leaves'))
100 gbm = lgb.LGBMRegressor(
101     boosting_type="gbdt",
102     objective='regression',
103     random_state=10,
104     n_estimators=b,
105     num_leaves=c,
106     feature_fraction=0.9,
107     bagging_fraction=0.8,
108     bagging_freq=5,
109     learning_rate=0.05,
110 )
111 # Entrenamiento de cada modelo del CVC con parámetros óptimos
112 gbm.fit(X_train[i], y_train[i],
113         eval_set=[(X_test[i], y_test[i])],
114         eval_metric='l2',
115         early_stopping_rounds=5)
116 # Guardar los k modelos en fichero externo
117 pickle.dump(gbm, open('{0}.p'.format(i), "wb" ))
118 # Cargar Modelos Entrenados
119 listfeat = []
120 listfeat2 = []
121 for i in range(0, nsplit):
122     loaded_model = pickle.load(open('{0}.p'.format(i), "rb"))
123     explainer = shap.TreeExplainer(loaded_model)
124 # Realizar las estimaciones con las entradas y obtener importancias SHAP y FeatImp
125 LGBy_pred_TS[i]=loaded_model.predict(X1_TS[col_mtoNH+colbasic], num_iteration=loaded_model.best_iteration_)
126 LGBy_pred_TR[i]=loaded_model.predict(X1_TR[col_mtoNH+colbasic], num_iteration=loaded_model.best_iteration_)
127 y_test_TS=X1_TS[col_targetNOX]
128 y_test_TR=X1_TR[col_targetNOX]
129 r2[i]=r2_score(y_test_TS, LGBy_pred_TS[i])
130 mae[i]=mean_absolute_error(y_test_TS, LGBy_pred_TS[i])
131 fea_imp_ = pd.DataFrame({'cols':X1_TR[col_mtoNH+colbasic].columns, 'fea_imp':loaded_model.feature_importances_})
132 shap_values = explainer.shap_values(X1_TR[col_mtoNH+colbasic])
133 fea_imp_.loc[fea_imp_['fea_imp'] > 0].sort_values(by=['cols'], ascending = True)
134 fea_imp_ = fea_imp_.sort_values(by=['cols', 'fea_imp'], ascending=[False, False])
135 listfeat.append(fea_imp_)
136 listfeat2.append(shap_values)
137 AV_LGBy_pred_TS=pd.DataFrame.from_dict(LGBy_pred_TS)
138 AV_LGBy_pred_TS['mean']= AV_LGBy_pred_TS.mean(axis=1)
139 AV_LGBy_pred_TR=pd.DataFrame.from_dict(LGBy_pred_TR)
140 AV_LGBy_pred_TR['mean']= AV_LGBy_pred_TR.mean(axis=1)
141 # Rendimiento estadístico modelo CVC sobre DSI & DSI
142 print('R^2,Validation_Score:_{:2f},\nMAENOC:_{:2f}'.format(r2_score(y_test_TS, AV_LGBy_pred_TS['mean']),mean_absolute_error(y_
143 print('R^2,Validation_Score:_{:2f},\nMAENOC:_{:2f}'.format(r2_score(y_test_TR, AV_LGBy_pred_TR['mean']),mean_absolute_error(y_
    
```

MTImp

```

1 def MTImp(dfdata,df2,lmodel,num=t):
2     dfdate=df2[['Fecha_inicio', 'Fecha_fin', 'CODLOC']]
    
```

121

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China Fecha: 20/01/2021 19:43:20
 UNIVERSIDAD DE LA LAGUNA
 Jesús Miguel Torres Jorge 20/01/2021 21:15:50
 UNIVERSIDAD DE LA LAGUNA

161 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
 UNIVERSIDAD DE LA LAGUNA
 María de las Maravillas Aguiar Aguilár 27/05/2021 13:54:44
 UNIVERSIDAD DE LA LAGUNA

161 / 181

Apéndice C - Código representativo del Capítulo 4

```
3 dfdate.index = dfdate['Fecha_inicio']
4 dfdate = dfdate[dfdate.index[-1]:dfdate.index[0]].sort_index()
5 larray=[]
6 for i in range(len(dfdate)):
7 # Predicción del estado de referencia del modelo t0 (por defecto instante de realización de revisión)
8 # Fecha inicial
9 a = (dfdate.iloc[i,1]).strftime("%m-%d_%H%M%S")
10 # Fecha final
11 b = dfdate[dfdate.iloc[:,7+i]>num].index[0].strftime("%m-%d_%H%M%S")
12 sdata = dfdate[a:b]
13 ref = sdata.iloc[:,7].mean(axis=0).to_frame().transpose()
14 ref2 = sdata.iloc[0:1,7:]
15 ref2.index = [0]
16 ref = pd.concat([ref,ref2],axis=1)
17 refx = ref.iloc[:,1:]
18 refy = ref.iloc[:,0:1]
19 # Simulación del modelo M en t0
20 rvalue = ALGB_prediction(lmodel,refx)
21 spred = refx.copy()
22 # Establecer mismas variables ambientales y funcionamiento en la comparativa
23 spred.iloc[:,6+i] = 0
24 # Simulación del modelo M en t
25 dfm = ALGB_prediction(lmodel,spred)
26 # Comparativa de dos instantes t y t0
27 res = ((rvalue-dfm)).iloc[0,0]
28 larray.append([res])
29 # Total de importancias en absoluto
30 total = 0
31 for k in larray :
32 total+=abs(k[0])
33 # Calculo relativo de las importancias %
34 larray = [(1/total)*100 for i in larray]
35 return larray
```

MBD.SIMU

```
1 # Simulación para la comparativa de NOx Real, predicho y si no se hubiera realizado mantenimiento
2 # Del DSI, selección de 3x revisiones en un intervalo aproximado de 500h de funcionamiento
3 XIN2=XI_TR.loc[2877:3027]
4 XIN3=XI_TR.loc[0:200]
5 XIN4=XI_TR.loc[7550:7700]
6 # Simular variable sinrealizar mantenimiento
7 XIN2 = XIN2.assign(MT_O15000_2017_OH=0)
8 XIN3 = XIN3.assign(MT_O10000_2016_OH=0)
9 XIN4 = XIN4.assign(MT_O10000_2018_OH=0)
10 LGBY_pred_XIN2={}
11 LGBY_pred_XIN3={}
12 LGBY_pred_XIN4={}
13 # Realizar nuevas predicciones con variable sin mantenimiento a partir del modelo entrenado originalmente.
14 for i in range(0,nsplit):
15 loaded_model = pickle.load(open('0).p'.format(i), "rb"))
16 explainer = shap.TreeExplainer(loaded_model)
17 LGBY_pred_XIN2[i]=loaded_model.predict(XIN2[col_mtoNH+colbasic], num_iteration=loaded_model.best_iteration_)
18 AV_LGBY_pred_XIN2=pd.DataFrame.from_dict(LGBY_pred_XIN2)
19 AV_LGBY_pred_XIN2['mean'] = AV_LGBY_pred_XIN2.mean(axis=1)
20 AV_LGBY_pred_XIN2 = pd.DataFrame(AV_LGBY_pred_XIN2['mean'], index=AV_LGBY_pred_XIN2['mean'].index)
21 # Realizar nuevas predicciones con variable sin mantenimiento a partir del modelo entrenado originalmente.
22 for i in range(0,nsplit):
```

122

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

162 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.

La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

162 / 181

```

23     loaded_model = pickle.load(open('{0}.p'.format(i), "rb"))
24     explainer = shap.TreeExplainer(loaded_model)
25     LGBBy_pred_XIN3[i]=loaded_model.predict(XIN3[col_mtoNH+colbasic], num_iteration=loaded_model.best_iteration_)
26     AV_LGBy_pred_XIN3=pd.DataFrame.from_dict(LGBBy_pred_XIN3)
27     AV_LGBy_pred_XIN3['mean']= AV_LGBy_pred_XIN3.mean(axis=1)
28     AV_LGBy_pred_XIN3 = pd.DataFrame(AV_LGBy_pred_XIN3['mean'], index=AV_LGBy_pred_XIN3['mean'].index )
29     # Realizar nuevas predicciones con variable sin mantenimiento a partir del modelo entrenado originalmente.
30     for i in range(0, nsplit):
31         loaded_model = pickle.load(open('{0}.p'.format(i), "rb"))
32         explainer = shap.TreeExplainer(loaded_model)
33         LGBBy_pred_XIN4[i]=loaded_model.predict(XIN4[col_mtoNH+colbasic], num_iteration=loaded_model.best_iteration_)
34         AV_LGBy_pred_XIN4=pd.DataFrame.from_dict(LGBBy_pred_XIN4)
35         AV_LGBy_pred_XIN4['mean']= AV_LGBy_pred_XIN4.mean(axis=1)
36         AV_LGBy_pred_XIN4 = pd.DataFrame(AV_LGBy_pred_XIN4['mean'], index=AV_LGBy_pred_XIN4['mean'].index )
37     # Preparar el resto de datos reales y estimaciones con mantenimiento
38     X1_TR2=X1_TR.loc[2877:3027]
39     X1_TR3=X1_TR.loc[0:266]
40     X1_TR4=X1_TR.loc[7550:7700]
41     X1_TR2REAL=X1_TR['EM_NOX']
42     X1_TR2REAL=X1_TR2REAL.loc[2877:3027]
43     X1_TR3REAL=X1_TR['EM_NOX']
44     X1_TR3REAL=X1_TR3REAL.loc[0:200]
45     X1_TR4REAL=X1_TR['EM_NOX']
46     X1_TR4REAL=X1_TR4REAL.loc[7550:7700]
47     X1_TR2FECHA=df4_TR[2877:3028].reset_index()
48     X1_TR2FECHA=X1_TR2FECHA['FECHA']
49     X1_TR2FECHA=X1_TR2FECHA.dt.strftime('%d/%m/%L, %H%M')
50     X1_TR3FECHA=df4_TR[0:201].reset_index()
51     X1_TR3FECHA=X1_TR3FECHA['FECHA']
52     X1_TR3FECHA=X1_TR3FECHA.dt.strftime('%d/%m/%L, %H%M')
53     X1_TR4FECHA=df4_TR[7550:7701].reset_index()
54     X1_TR4FECHA=X1_TR4FECHA['FECHA']
55     X1_TR4FECHA=X1_TR4FECHA.dt.strftime('%d/%m/%L, %H%M')
56     AVPRED_TR2=AV_LGBy_pred_TR.loc[2877:3027]
57     AVPRED_TR3=AV_LGBy_pred_TR.loc[0:200]
58     AVPRED_TR4=AV_LGBy_pred_TR.loc[7550:7700]
59     # F = factor de confidencialidad de los datos de NOx
60     REAL2N= X1_TR2REAL/F
61     PRED2N= AVPRED_TR2['mean']/F
62     SIN2N= AV_LGBy_pred_XIN2/F
63     REAL3N= X1_TR3REAL/F
64     PRED3N= AVPRED_TR3['mean']/F
65     SIN3N= AV_LGBy_pred_XIN3/F
66     REAL4N= X1_TR4REAL/F
67     PRED4N= AVPRED_TR4['mean']/F
68     SIN4N= AV_LGBy_pred_XIN4/F
    
```

123

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China	Fecha: 20/01/2021 19:43:20
<i>UNIVERSIDAD DE LA LAGUNA</i>	
Jesús Miguel Torres Jorge	20/01/2021 21:15:50
<i>UNIVERSIDAD DE LA LAGUNA</i>	

163 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO	Fecha: 26/01/2021 17:45:11
<i>UNIVERSIDAD DE LA LAGUNA</i>	
María de las Maravillas Aguiar Aguiar	27/05/2021 13:54:44
<i>UNIVERSIDAD DE LA LAGUNA</i>	

163 / 181



Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

164 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguilár
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

Bibliografía

Las siguientes referencias bibliográficas se presentan en orden alfabético por autor. Las referencias con más de un autor aparecen ordenadas en base al primero de los mismos.

M. W. Ahmad, M. Mourshed e Y. Rezgui. Trees vs neurons: Comparison between random forest and ann for high-resolution prediction of building energy consumption. *Energy and Buildings Volume*, 147:77–89, Jul 15, 2017. DOI: 10.1016/j.enbuild.2017.04.038.

G. F. Aker y H. I. H. Saravanamuttoo. Predicting gas turbine performance degradation due to compressor fouling using computer simulation techniques. *J. Eng. Gas Turbines Power*, 111(2):343–350, Apr 1989. DOI: 10.1115/1.3240259.

N. Aldi, M. Morini, M. Pinelli, P. R. Spina y A. Suman. An interdisciplinary approach to study the fouling phenomenon. págs. 280–285, 2015.

M. M. Alhazmy e Y. S. H. Najjarb. Augmentation of gas turbine performance using air coolers. *Applied Thermal Engineering*, 24(2-3):415–429, 2004. DOI: 10.1016/j.applthermaleng.2003.09.006.

M. Amozegar y K. Khorasani. An ensemble of dynamic neural network identifiers for fault detection and isolation of gas turbine engines. *Neural Networks*, 76:106–121, April 2016. DOI: 10.1016/j.neunet.2016.01.003.

125

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

165 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

165 / 181

Bibliografía

- L. Antwarg, R. M. Miller, B. Shapira y L. Rokach. Explaining anomalies detected by autoencoders using shap. Jun 30, 2020.
- A. Avila, I. Queralt y M. Alarcón. Mineralogical composition of african dust delivered by red rains over northeastern. *Geophysical Research Atmospheres*, 102 (D18), 1997. DOI: 10.1029/97JD00485.
- M. Baptista y C. Nascimento. A case for the use of data-driven methods in gas turbine prognostics. 2017.
- D. Berrar, I. Bradbury y W. Dubitzky. Avoiding model selection bias in small sample genomic datasets. *Bioinformatics*, 22(19):2453, 2006. DOI: 10.1093/bioinformatics/btl435.
- G. Biau y E. Scornet. A random forest guided tour. *TEST*, 25:197–227, 2016. DOI: 10.1007/s11749-016-0481-7.
- S. Borra y A. D. Ciaccio. Measuring the prediction error. a comparison of cross-validation, bootstrap and covariance penalty methods. *Computational statistics and data analysis*, 54(12):2976–2989, 2010. DOI: 10.1016/j.csda.2010.03.004.
- M. A. Braz, P. Santana, E. Júnior y A. Pellegrin. Classification of the coefficient of variation to variables in beef cattle experiments. *Ciência rural*, 47:11, 2017. DOI: 10.1590/0103-8478cr20160946.
- L. Breiman. Random forests. *Machine Language*, October 2001. DOI: 10.1023/A:1010933404324.
- L. Breiman. Bagging predictors. *Machine Learning*, 24:123–140, 1996. DOI: 10.1023/A:1018054314350.
- F. Burden y D. Winkler. Bayesian regularization of neural networks. *Literature review in methods in molecular biology*, 458:23–42, 2009. DOI: 10.1007/978-1-60327-101-1_3.

126

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

166 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

166 / 181

Bibliografía

- L. Cana. The saharan dust episode of 26 february 2000 over the canary archipelago: a synoptic overview. *Weather*, 57(10):385–389, 2002. DOI: 10.1256/wea.275.01.
- T. P. Carvalho, F. Soares, R. Vita, F. Francisco, J. P. Basto y S. Alcalá. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, Nov, 2019. DOI: 10.1016/j.cie.2019.106024.
- J. Cervantes, J. Taltempa, F. García-lamont, F. Castilla y A. Y. Rendon. Comparative analysis of the techniques used in a recognition system of plant leaves. *RIAI Revista Iberoamericana de automática e Informática Dndustrial*, 14(1):104–114, 2017. DOI: 10.1016/j.riai.2016.09.005.
- S. Chebrolu, A. Abraham y J. P. Thomasa. Feature deduction and ensemble design of intrusion detection systems. *Computers & Security*, 24(4):295–307, 2005. DOI: 10.1016/j.cose.2004.09.008.
- T. Chen y C. Guestrin. Xgboost: A scalable tree boosting system. págs. 785–794, 2016. DOI: 10.1145/2939672.2939785.
- Y. Chou. *No title*, cap. 17.9. Statistical analysis.: With business and economic applications. 1975. ISBN 0-03-089422-0.
- F. Cipollini, L. Oneto, A. Coraddu, A. J. Murphy y D. Anguita. Condition-based maintenance of naval propulsion systems with supervised data analysis. *Ocean Engineering*, 149:268–278, Feb 1, 2018. DOI: 10.1016/j.oceaneng.2017.12.002.
- A. Coraddu, L. Oneto, A. Ghio, S. Savio, D. Anguita y M. Figari. Machine learning approaches for improving condition-based maintenance of naval propulsion plants. 2014. DOI: 10.1177/1475090214540874.
- C. Criado y P. Dorta. An unusual blood rain over canary islands (spain). the storm of january 1999. *Journal of Arid Environments*, 55:765–783, 2003. DOI: 10.1016/S0140-1963(02)00320-8.

127

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA
Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

167 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA
María de las Maravillas Aguiar Aguiar 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

167 / 181

Bibliografía

- C. Criado, J. M. Torres, A. Hansen, P. Lillo y A. Naranjo. Intercalaciones de polvo sahariano en paleodunas bioclásticas de fuerteventura (islas canarias). *Cuaternario y geomorfología*, 26(1-2):73–88, 2012.
- S. Cruz-Manzo, S. Maleki, Y. Zhang, V. Panov y A. Latimer. Ieee international conference on prognostics and health management (icphm). En *Dallas, TX, USA*. IEEE, 19-21 June 2017 2017. DOI: 10.1109/ICPHM.2017.7998334.
- J. Csavina, J. Field, O. Félix y E. Betterton. Effect of wind speed and relative humidity on atmospheric dust concentrations in semi-arid climates. *Science of the Total Environment*, 487(1):82–90, 2014. DOI: 10.1016/j.scitotenv.2014.03.138.
- L. Cuvelier y M. D. Belcher. Asme 1990 international gas turbine and aeroengine congress and exposition. Belgium, 1990 2015. DOI: 10.1115/90-GT-177.
- R. S. Davis. Equation for the determination of the density of moist air (1981/91). *Metrologia*, 29(1), 1991.
- I. S. Diakunchak. Performance improvement in industrial gas turbines. (93-JPGC-GT-5), February 23, 2015. DOI: 10.1115/93-JPGC-GT-5.
- J. L. Diaz-hernandez y A. Sanchez-Navas. Saharan dust outbreak sand iberulite episodes. *JGR Atmospheres*, 121:7064–7078, 2016. DOI: 10.1002/2016JD024913.
- A. Diez-Olivana, J. D. Ser, D. Galar y B. Sierra. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0. *Information Fusion*, 50:92–111, October 2019. DOI: 10.1016/j.inffus.2018.10.005.
- Y. Dimopoulos, P. Bourret y S. Lek. Use of some sensitivity criteria for choosing networks with good generalization ability. *Neural Processing Letters*, 2:1–4, 1995. DOI: 10.1007/BF02309007.
- P. X. Dionysios, M. Georgios, M. Ciccotti y N. F. Thornhill. Operational optimization of networks of compressors considering condition-based maintenance. *Computers & Chemical Engineering*, 84:117–131, 2016. DOI: doi.o10.1016/j.compchemeng.2015.08.008.

128

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

168 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

168 / 181

Bibliografía

- Y. Dong y K. Han. A comparison of several ensemble methods for text categorization. En *IEEE international conference on services computing*, págs. 419–422, 2004. DOI: 10.1109/SCC.2004.1358033.
- P. Dorta. Aproximación a la influencia de las advecciones de aire sahariano en la propagación de los incendios forestales en la provincia de santa cruz de tenerife. En *XVII Congreso de Geógrafos Españoles*, págs. 158–162, Oviedo, 2001.
- P. Dorta, M. D. Gelado, J. J. Hernández, P. Cardona, C. Collado, S. Mendoza, M. J. Rodríguez, V. Siruela y M. E. Torres. Algunas consideraciones sobre la importancia del polvo de origen sahariano en el clima del archipiélago canario y su aporte a las aguas superficiales oceánicas: el episodio de abril de 2002, en el agua y el clima. En *III Congreso de la Asociación Española de Climatología*, págs. 13–24, Palma de Mallorca, 2002.
- P. Dorta, M. D. Gelado, J. J. Hernández, P. Cardona, C. Collado, S. Mendoza, M. J. Rodríguez, V. Siruela y M. E. Torres. Frecuencia, estacionalidad y tendencias de las advecciones de aire sahariano en canarias (1976-2003). *Investigaciones Geográficas*, (38), 2005.
- S. Duer, K. Zajkowski y R. Duer. Designing of an effective structure of system for the maintenance of a technical object with the using information from an artificial neural network. *Neural Computing and Applications*, Jul 2, 2012. DOI: 10.1007/s00521-012-1016-0.
- N. J. Edwin, H. Mjølnerød y B. A. Gran. *Data-driven and risk-based decision support for maintenance planning on electrical power grid systems*. Safety and Reliability – Safe Societies in a Changing World. 2018. ISBN 9781351174664.
- J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.
- J. H. Friedman. Stochastic gradient boosting. 38(4):367–378, February 2002. DOI: 10.1016/S0167-9473(01)00065-2.

129

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

169 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

169 / 181

Bibliografía

- C. A. Friese, M. V. der Does, U. Merkel, M. H. Iversen, G. Fischer, W. Gerhard y J. B. Stuut. Environmental factors controlling the seasonal variability in particle size distribution of modern saharan dust deposited off cape blanc. *Aeolian research*, 22:165–179, 2016. DOI: 10.1016/j.aeolia.2016.04.005.
- T. Gao, J. Han, Y. Wang, H. Pei y S. Lu. Impacts of climate abnormality on remarkable dust storm increase of the hunshdak sandy lands in northern china during 2001–2008. *Meteorology & Atmospheric Sciences*, 19(3):265–278, 2011. DOI: 10.1002/met.251.
- J. García, A. Hernández, A. Blasco, B. C. Rodríguez, E. Rancaño y S. Núñez. Invasión de viento sahariano y su impacto en la asistencia sanitaria urgente. *Emergencias*, 13:372–376, 2001.
- G. D. Garson. Interpreting neural network connection weights. *AI Expert*, 6:47–51, 1991.
- S. Gayraud y R. Singh. Asme 1997 turbo asia conference. vol. V001T08A006. Proc. ASME, 1997. DOI: 10.1115/97-AA-130.
- M. D. Gelado, P. J. Dorta, M. E. Torres, J. J. Hernandez, C. Collado, V. F. Siruela, P. Cardona y M. J. Rodríguez. Caracterización del aerosol sahariano en gran canaria. págs. 155–157, 2002.
- M. Gevrey, I. Dimopoulos y S. Leka. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*, 160(3):249–264, 2003. DOI: 10.1016/S0304-3800(02)00257-0.
- T. Giampaolo. *Compressor Fouling*, págs. 205. The Gas Turbine Handbook: Principles and Practices. Fairmont Press, 2003. ISBN 0881734144.
- A. T. C. Goh. Back-propagation neural networks for modeling complex systems. *Artif.Intell.Eng.*, 9:143–151, 1995. DOI: 10.1016/0954-1810(94)00011-S.
- F. P. Gomes. *No title*, págs. 467. Curso de estadística experimental. Nobel, São paulo, 1985.

130

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

170 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

170 / 181

Bibliografía

- Y. Guo y D. Yu. The influence of interconnection of electric power systems on load characteristic and frequency regulation. 70(1):23–29, June 2004. DOI: 10.1016/j.epr.2003.11.006.
- K. Guolin, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye y T. Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 2017.
- H. Hanachi, C. Mechefske, J. Liu, A. Banerjee e Y. Chen. Ieee international conference on prognostics and health management (icphm). IEEE, 19-21 June 2017 2017. DOI: 10.1109/ICPHM.2017.7998307.
- S. Haykin. *Neural networks and learning machines*. 1999. ISBN-10: 0-13-147139-2.
- M. Herterich. On the design of digitized industrial products as key resources of service platforms for industrial service innovation. págs. 364–380, 2017. DOI: 10.1007/978-3-319-59144-5_22.
- G. Huang. Learning capability and storage capacity of two-hidden-layer feed-forward networks. *IEEE Transactions on Neural Networks*, 14(2), 2003. DOI: 10.1109/TNN.2003.809401.
- V. A. Huynh-Thu, A. Irrthum, L. Wehenkel y P. Geurts. Inferring regulatory networks from expression data using tree-based methods. *Plos One*, Sept 28, 2010. DOI: 10.1371/journal.pone.0012776.
- H. Inamul y H. I. H. Saravanamuttoo. Asme 1991 international gas turbine and aeroengine congress and exposition. vol. V003T07A006, Orlando, Florida, USA, 2015. Proc. ASME. DOI: 10.1115/91-GT-067.
- F. Jaime-Rojas y G. Latorre-Bayona. establecimiento de la reserva operativa como servicio complementario. *REVISTA DE LA FACUL3.AD DE INGENIER(AS FI51COMECANICAS*, 3(2):117–126, Dec 2004.

131

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

171 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

171 / 181

Bibliografía

- P. Jamshid y A. Vahid. Dust storm prediction using anns technique: A case study-zabol city. *World Academy of Science, Engineering and Technology International Journal of Civil and Environmental Engineering*, 2(7), 2008. DOI: 10.5281/zenodo.1061318.
- J. Janawitz, J. Masso y C. Childs. Heavy-duty gas turbine operating and maintenance considerations. *Inf. Téc. GER-3620M (02/15)*, GE Power & Water, 2010. URL https://www.ge.com/content/dam/gepower-pgdp/global/en_US/documents/technical/unused%20assets/hdgt-operating-maintenance-considerations-report.pdf.
- J. P. Jesan y D. M. Lauro. Human brain and neural network behavior a comparison. *Ubiquity*, 4(37):12–18, 2003. DOI: 10.1145/962068.958078.
- X. Jiang, T. Lin y E. Mendoza. Annual conference of the prognostics and health management society. Phm Society, 2016.
- S. Khoshafian y C. Rostetter. Digital prescriptive maintenance. *Pega*, 2015. URL <https://www.pegasystems.com/system/files/resources/2019-01/Digital-Prescriptive-Maintenance.pdf>.
- S. Kiakojoori y K. Khorasani. Dynamic neural networks for gas turbine engine degradation prediction, health monitoring and prognosis. *Neural Computing and Applications*, 27, July 2015. DOI: 10.1007/s00521-015-1990-0.
- R. Kohavi. A study of cross validation and bootstrap for accuracy estimation and model selection. En *Proceedings of the 14th international joint conference on artificial intelligence*, vol. 2, págs. 1137–1143, 1995. ISBN:1-55860-363-8.
- N. Koseleva y G. Ropaite. Big data in building energy efficiency: understanding of big data and main challenges. *Procedia Engineering*, 172:544–549, 2017. DOI: 10.1016/j.proeng.2017.02.064.
- J. Kowalski, B. Krawczy y M. Woźniak. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. *Engineering Applications*

132

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

172 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

172 / 181

Bibliografía

- tions of Artificial Intelligence*, 57:134–141, Jan, 2017. DOI: 10.1016/j.engappai.2016.10.015.
- K. Kunitomi, A. Kurita, H. Okamoto, Y. Tada, S. Ihara, P. Pourbeik, W. W. Price, A. B. Leirbukt y J. J. Sanchez-Gasca. 2001 ieeepower engineering society winter meeting. 2001. DOI: 10.1109/PESW.2001.916935.
- R. Kurz y K. Brun. Fouling mechanism in axial compressors. *Journal of Engineering for Gas Turbines and Power*, 134(032401), 2012. DOI: 10.1115/GT2011-45012.
- B. Liu, J. Hu, F. Yan, R. Turkson y F. Ling. A novel optimal support vector machine ensemble model for nox emissions prediction of a diesel engine. *Measurement*, 92(183-192), Oct, 2016. DOI: 10.1016/j.measurement.2016.06.015.
- B. Liu, J. Ling, Z. Liangwei y K. Uday. A dynamic prescriptive maintenance model considering system aging and degradation. *IEEE Access*, 7, Jul 15, 2019. DOI: 10.1109/ACCESS.2019.2928587.
- N. E. M. Llerena, L. Berton y A. de Andrade Lopes. Graph-based cross-validated committees ensembles. En *International Conference on Computational Aspects of Social Networks*, 2012. DOI: 10.1109/CASoN.2012.6412381.
- D. Lubo-Robles, D. Devegowda, V. Jayaram y H. Bedle. Machine learning model interpretability using shap values: application to a seismic facies classification task. 2020. DOI: 10.1190/segam2020-3428275.1.
- S. Lundberg y S. I. Lee. A unified approach to interpreting model predictions. 2017.
- S. M. Lundberg, G. Erion y S. I. Lee. Consistent individualized feature attribution for tree ensembles. ArXiv, 2018.
- P. Mahat, Z. Chen y B. Bak-Jensen. 2009 ieeepower & energy society general meeting. IEEE, 2009. DOI: 10.1109/PES.2009.5275701.

133

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

173 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguiar Aguiar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

173 / 181

Bibliografía

- S. R. Malekshan, M. A. Shoorehdeli y M. Yari. Iranian conference on electrical engineering (icee). 2017. DOI: 10.1109/IranianCEE.2017.7985130.
- S. Mandal y K. M. Kishore. Ensemble averaging based assessment of spatio-temporal variations in ambient pm2.5 concentrations over delhi, india, during 2010–2016. *Atmospheric Environment*, 224(117309), 1 March 2020. DOI: 10.1016/j.atmosenv.2020.117309.
- C. Maosen y Q. Pizhong. Neural network committee-based sensitivity analysis strategy for geotechnical engineering problems. *Neural Computing and Applications*, 17(5-6):509–519, 2008. DOI: 10.1007/s00521-007-0143-5.
- K. Matyasa, T. Nemeth, K. Kovacs y R. Glawar. A procedural approach for realizing prescriptive maintenance planning in manufacturing industries. *CIRP Annals*, 66(1):461–464, 2017. DOI: 10.1016/j.cirp.2017.04.007.
- T. Mckinley, M. Somwanshi y D. B. S. Verma. Identifying nox sensor failure for predictive maintenance of diesel. 2020.
- C. B. Meher-Homji y A. Bromley. Gas turbine axial compressor fouling and washing. *Texas A&M University. Turbomachinery Laboratories*, 2004. DOI: 10.21423/R1S66R. URL <http://hdl.handle.net/1969.1/163249>.
- N. J. Middleton. Desert dust hazards: a global review. *Aeolian research*, 24:53–63, 2017. DOI: 10.1016/j.aeolia.2016.12.001.
- E. Mohammadi y M. Montazeri-Gh. A fuzzy-based gas turbine fault detection and identification system for full and part-load performance deterioration. 46: 82–93, October–November 2015. DOI: 10.1016/j.ast.2015.07.002.
- T. Mohammadreza, Z. A. A. Karim y M. Masdi. A multi-nets ann model for real-time performance-based automatic fault diagnosis of industrial gas turbine engines. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39, February 2017. DOI: 10.1007/s40430-017-0742-8.

134

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

174 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

174 / 181

Bibliografía

- J. Montaña y A. Palmer. Numeric sensitivity analysis applied to feedforward neural networks. *Neural Computing & Applications*, 12:119–125, 2003. DOI: 10.1007/s00521-003-0377-9.
- D. Montgomery. cap. 7, págs. 289. Introduction to statistical quality control. 2008. ISBN 10:0470233974.
- J. Morren, S. W. H. De Haan y J. A. Ferreira. 2005 international conference on future power systems. IEEE, 2005. DOI: 10.1109/FPS.2005.204253.
- D. Mosaddar y A. Shojaie. A data mining model to identify inefficient maintenance activities. *International Journal of Systems Assurance Engineering and Management*, 4(2), Jun, 2013. DOI: 10.1007/s13198-013-0148-7.
- Z. Murat, A. Abdussalam, Q. Zeeshan y O. Korhan. Digital twins for industry 4.0: A review. págs. 193–203, 2020. DOI: 10.1007/978-3-030-42416-9_18.
- A. Natekin y A. Knoll. Gradient boosting machines, a tutorial. *Frontiers in Neurobotics*, 7(21), 2013. DOI: 10.3389/fnbot.2013.00021.
- D. Nguyen y B. Widrow. Neural networks for self-learning control systems. *IEEE Control Systems Magazine*, 10(3):18–23, 1990. DOI: 10.1109/37.55119.
- S. Nikolaev, S. Belov, M. Gusev e I. Uzhinsky. Hybrid data-driven and physics-based modelling for prescriptive maintenance of gas-turbine power plant. 2019. DOI: 10.1007/978-3-030-42250-9_36.
- X. Niu, C. Yan, H. Wang e Y. Wang. Investigation of ann and svm based on limited samples for performance and emissions prediction of a crdi-assisted marine diesel engine. *Applied Thermal Engineering*, 111:1353–1364, Jan, 2017. DOI: 10.1016/j.applthermaleng.2016.10.042.
- J. Ostrowski y J. Menyhárt. Enhancing maintenance with a data-driven approach. *International Review of Applied Sciences and Engineering*, 10(2):135–140, Dec 1, 2019. DOI: 10.1556/1848.2019.0016.

135

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA 20/01/2021 21:15:50

175 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA 27/05/2021 13:54:44

175 / 181

Bibliografía

- J. Oña y C. Garrido. Extracting the contribution of independent variables in neural network models: a new approach to handle instability. *Neural computing and applications*, 25(3-4):859–869, 2014. DOI: 10.1007/s00521-014-1573-5.
- J. Park, J. Moon, S. Jung y E. Hwang. Multistep-ahead solar radiation forecasting scheme based on the light gradient boosting machine: A case study of jeju island. *Assessment of Renewable Energy Resources with Remote Sensing*, 12(14), 2020. DOI: 10.3390/rs12142271.
- B. Parmanto, P. Munro y H. Doyle. Improving committee diagnosis with resampling techniques. En *Proceedings of the 8th international conference on neural information processing systems*, págs. 882–888, 1996.
- S. G. K. Patro y K. K. Sahu. Normalization: A preprocessing stage. March 2015. DOI: 10.17148/IARJSET.2015.2305.
- M. Pawelczyk y S. Fulara. Industrial gas turbine operating parameters monitoring and data-driven prediction. *Eksploatacja i Niezawodność - Maintenance and Reliability*, 22(3):391–399, June 2020. DOI: 10.17531/ein.2020.3.2.
- M. Pelt, K. Stamoulis y A. Apostolidis. Data analytics case studies in the maintenance, repair and overhaul (mro) industry. 2019. DOI: 10.1051/mateconf/201930404005.
- N. H. Pontoppidan, T. Lehn-Schiøler y K. B. Petersen. Machine learning for condition monitoring and innovation. IEEE, 2019. DOI: 10.1109/ICASSP.2019.8683545.
- R. Prytz, S. Nowaczy, T. Rögnvaldsson y S. Byttner. Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. *Engineering Applications of Artificial Intelligence*, 41:139–150, May, 2015. DOI: 10.1016/j.engappai.2015.02.009.
- K. Pye. Aeolian dust and dust deposits, 1987.

136

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

176 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
 La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

176 / 181

Bibliografía

- K. Qadeer y M. Jeon. Prediction of pm10 concentration in south korea using gradient tree boosting models. 2019. DOI: 10.1145/3387168.3387234.
- S. Raptotasio, N. F. Sakellaridis, R. G. Papagiannakis y D. T. Hountalasa. Application of a multi-zone combustion model to investigate the nox reduction potential of two-stroke marine diesel engines using egr. *Applied Energy*, 157: 814–823, Nov, 2015. DOI: 10.1016/j.apenergy.2014.12.041.
- A. Rehman y A. Bermak. Averaging neural network ensembles model for quantification of volatile organic compound. *IEEE*, 24-28 June 2019. DOI: 10.1109/IWCMC.2019.8766776.
- D. Rengasamy, H. P. Morvan y G. P. Figueredo. Deep learning approaches to aircraft maintenance, repair and overhaul: A review. *IEEE*, 2018. DOI: 10.1109/ITSC.2018.8569502.
- R. Rosen, G. Wichert, L. George y D. B. Kurt. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, 48(3): 567–572, 2015. DOI: 10.1016/j.ifacol.2015.06.141.
- A. D. Sa y S. A. Zubaidy. Gas turbine performance at varying ambient temperature. *Applied Thermal Engineering*, 31(14-15):2735–2739, 2011. DOI: 10.1016/j.applthermaleng.2011.04.045.
- M. Sandri y P. Zuccolotto. Analysis and correction of bias in total decrease in node impurity measures for tree-based algorithms. *Statistics and Computing*, 20:393–407, 2010.
- M. Santini, G. Marchetti, F. Giuntini y C. Suarez. Gas turbine high efficiency filtration systems. *GE OIL & GAS*, 2013.
- A. P. Santos y C. R. Andrade. Analysis of gas turbine performance with inlet air cooling techniques applied to brazilian sites. *Aerospace Technology and Management*, 4(3), 2012. DOI: 10.5028/jatm.2012.04032012.

137

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

177 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

177 / 181

Bibliografía

- S. Sapna. Backpropagation learning algorithm based on levenberg marquardt algorithms. En *The fourth international workshop on computer networks and communications*, 2012. DOI: 10.5121/csit.2012.2438.
- M. Saviozzi, S. Massucco y F. Silvestro. Implementation of advanced functionalities for distribution management systems: load forecasting and modeling through artificial neural networks ensembles. *electric power systems research*. 167:230–239, 2019. DOI: 10.1016/j.epsr.2018.10.036.
- A. Schneider, G. Hommel y M. Blettner. Linear regression analysis. *Dtsch arztebl int*, 107(44):776–782, 2010. DOI: 10.3238/arztebl.2010.0776.
- S. Shin, Y. Lee, K. Minjae, J. Park, S. Lee y M. Kyoungdoug. Deep neural network model with bayesian hyperparameter optimization for prediction of no at transient conditions in a diesel engine. *Engineering Applications of Artificial Intelligence*, 94, Sept, 2020. DOI: 10.1016/j.engappai.2020.103761.
- J. P. Stalder. Gas turbine compressor washing state of the art: Field experiences. *Engineering for Gas Turbines and Power*, 123(2):363–370, 2001. DOI: 10.1115/1.1361108.
- E. Syverud, O. Brekke y E. B. L. Axial compressor deterioration caused by saltwater ingestion. *Journal of Turbomachinery*, 129(1):119–126, 2007. DOI: 10.1115/1.2219763.
- A. P. Tarabrin, V. A. Schurovsky, A. I. Bodrov y J. P. Stalder. Asme 1998 international gas turbine and aeroengine congress and exhibition. vol. V004T11A006, Stockholm, Sweden, 1998 2014. Proc. ASME. DOI: 10.1115/98-GT-416.
- A. P. Tarabrin, V. A. Schurovsky, A. I. Bodrov y J. P. Stalder. Asme 1996 international gas turbine and aeroengine congress and exhibition. vol. V001T01A093. Proc. ASME, 1996 2015. DOI: 10.1115/96-GT-363.

138

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

178 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguilar Aguilar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

178 / 181

Bibliografía

- K. I. Thamir, M. M. Rahmanab y N. A. Ahmed. Gas turbine configuration for improving the performance of combined cycle power plant. *Procedia Engineering*, 15:4216–4223, 2011. DOI: 10.1016/j.proeng.2011.08.791.
- J. L. Torrecilla y J. Romo. Data learning from big data. *Statistics and Probability Letters*, 136:15–19, May 2018. DOI: 10.1016/j.spl.2018.02.038.
- M. E. Torres-Padrón, M. D. Gelado-Caballero, C. Collado-Sánchez, V. Siruela-Matos, P. J. Cardona-Castellano y J. J. Hernández-Brito. Variability of dust inputs on the canigo zone. *Deep-sea research II*, 49:3455-3464, 2002. DOI: 10.1016/S0967-0645(02)00091-7.
- A. J. Ujam, P. O. Ekere y T. O. Chime. Performance evaluation of a gas turbine power plant by the application of compressor off-line and on-line water washing techniques. (a case study of 450mw sapele power station in delta state, nigeria). *IOSR Journal of Engineering*, 3(11):29–41, 2013. DOI: 10.9790/3021-031112941.
- A. Vorkapić, R. Radonja, K. Babić y S. Martinčić-Ipšić. Machine learning methods in monitoring operating behaviour of marine two-stroke diesel engine. *Transport*, 35(5):462–473, Dec, 2020. DOI: 10.3846/transport.2020.14038.
- W. Wang, P. Jones y D. Partridge. Assessing the impact of input features in a feedforward neural network. *Neural Computing & Applications*, 9:101–112, 2000. DOI: 10.1007/PL00009895.
- B. Weng, L. Lu, X. Wang, F. Megahed y W. Martinez. Predicting short-term stock prices using ensemble methods and online data sources. *Expert Systems with Applications*, 112:258–273, 2018. DOI: 10.1016/j.eswa.2018.06.016.
- M. Wilcox, R. Kurz y K. Brun. Technology review of modern gas turbine inlet filtration systems. *International Journal of Rotating Machinery*, 12, 2012. DOI: 10.1155/2012/128134.
- C. J. Willmott y K. Matsuura. Advantages of the mean absolute error (mae) over

139

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinaa Fecha: 20/01/2021 19:43:20
UNIVERSIDAD DE LA LAGUNA
Jesús Miguel Torres Jorge 20/01/2021 21:15:50
UNIVERSIDAD DE LA LAGUNA

179 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO Fecha: 26/01/2021 17:45:11
UNIVERSIDAD DE LA LAGUNA
María de las Maravillas Aguiar Aguilár 27/05/2021 13:54:44
UNIVERSIDAD DE LA LAGUNA

179 / 181

Bibliografía

- the root mean square error (rmse) in assessing average model performance. *Climate research*, 30:79–82, 2005. DOI: 10.3354/cr030079.
- R. Xia, C. Zong y S. Li. Ensemble of feature sets and classification algorithms for sentiment classification. *Information Sciences*, 181(6):1138–1152, 2011. DOI: 10.1016/j.ins.2010.11.023.
- X. Xiaojian, Z. Zhuangzhuang, X. Xiaobin, Y. Jianbo, C. Leilei y W. Guodong. Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models. *Knowledge-Based Systems*, 190, Feb 29, 2020. DOI: 10.1016/j.knosys.2019.105324.
- L. Yi, Z. Zhengjiang y C. Junghui. Ensemble local kernel learning for online prediction of distributed product outputs in chemical processes. *Chemical Engineering Science*, 137:140–151, 2015. DOI: 10.1016/j.ces.2015.06.005.
- L. Yi, Y. Chao, G. Zengliang e Y. Yuan. Ensemble deep kernel learning with application to quality prediction in industrial polymerization processes. *Chemometrics and Intelligent Laboratory Systems*, 174:15–21, 2018. DOI: 10.1016/j.chemolab.2018.01.008.
- X. Zeng y D. Yeung. A quantified sensitivity measure for multilayer perceptron to input perturbation. *MIT Press*, 15(1):183–212, 2003. DOI: 10.1162/089976603321043757.
- L. Zheng, N. Meyendorf y N. Mrad. The role of data fusion in predictive maintenance using digital twin. 2018. DOI: 10.1063/1.5031520.
- Z. Zhou y D. Liu. Prognostics and system health management conference. Chengdu, China, 2016 2016.
- Z.-H. Zhou, J. Wu y W. Tang. Ensembling neural networks: Many could be better than all. *Artificial Intelligence*, 137(1-2):239–263, 2002. DOI: 10.1016/S0004-3702(02)00190-X.

140

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar Chinae UNIVERSIDAD DE LA LAGUNA	Fecha: 20/01/2021 19:43:20
Jesús Miguel Torres Jorge UNIVERSIDAD DE LA LAGUNA	20/01/2021 21:15:50

180 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO UNIVERSIDAD DE LA LAGUNA	Fecha: 26/01/2021 17:45:11
María de las Maravillas Aguilar Aguilar UNIVERSIDAD DE LA LAGUNA	27/05/2021 13:54:44

180 / 181

Bibliografía

K. M. Zoran y D. D. Gvozdenac. *Toolbox 6: Thermodynamic and Transport Properties of Moist Air*. Applied Industrial Energy and Environmental Management. John Wiley & Sons, 2008. DOI: 10.1002/9780470714379.

141

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3170295 Código de verificación: a2/na6Ur

Firmado por: Rosa María Aguilar China
UNIVERSIDAD DE LA LAGUNA

Fecha: 20/01/2021 19:43:20

Jesús Miguel Torres Jorge
UNIVERSIDAD DE LA LAGUNA

20/01/2021 21:15:50

181 / 181

Este documento incorpora firma electrónica, y es copia auténtica de un documento electrónico archivado por la ULL según la Ley 39/2015.
La autenticidad de este documento puede ser comprobada en la dirección: <https://sede.ull.es/validacion/>

Identificador del documento: 3178786 Código de verificación: axpfLfm8

Firmado por: DANIEL GONZALEZ CALVO
UNIVERSIDAD DE LA LAGUNA

Fecha: 26/01/2021 17:45:11

María de las Maravillas Aguiar Aguiar
UNIVERSIDAD DE LA LAGUNA

27/05/2021 13:54:44

181 / 181