

Predicción Temprana del Fracaso en una Asignatura de Electrónica con Técnicas de Aprendizaje Automático

Raúl Alcaraz
Escuela Politécnica
Universidad de Castilla-La Mancha
Cuenca, Spain
Raul.Alcaraz@uclm.es

Arturo Martínez-Rodrigo
Facultad de Comunicación
Universidad de Castilla-La Mancha
Cuenca, Spain
Arturo.Martinez@uclm.es

Roberto Zangróniz
Escuela Politécnica
Universidad de Castilla-La Mancha
Cuenca, Spain
Roberto.Zangroniz@uclm.es

José J. Rieta
Escuela Politécnica Superior de Gandía
Universitat Politècnica de Valencia
Gandía (Valencia), Spain
jjrieta@upv.es

Resumen—La minería de datos en educación se ha convertido en una disciplina muy útil para la recogida de datos y evaluación del rendimiento académico de los estudiantes. Una de las aplicaciones directas, consiste en detectar aquellos estudiantes con bajo rendimiento que podrían suspender en una determinada materia, para llevar a cabo actuaciones de refuerzo docente que reduzcan este efecto. Estos sistemas de aviso temprano se han aplicado tanto a cursos online, como en la modalidad de clases presenciales. Sin embargo, para que estos sistemas sean eficientes, requieren una evaluación del rendimiento del alumno lo más temprana posible, cosa que no siempre puede conseguirse con sistemas de predicción generalistas. Por esta razón, estudios recientes han comprobado la mejora en el rendimiento de clasificación cuando se utilizan sistemas adaptados a cada asignatura concreta. Así, en este estudio, se han elaborado dos modelos predictivos para identificar a los estudiantes con bajo rendimiento que podrían suspender la asignatura de Sistemas Electrónicos de Potencia. Los resultados mostraron unas precisiones globales en la clasificación alrededor del 75 % para los dos modelos, utilizando las calificaciones obtenidas por los alumnos durante las cuatro primeras semanas de clase. Sin embargo, algunas de las características evaluadas mostraron diferencias significativas a partir de la segunda semana de clase. Este hecho abre la puerta a futuros estudios en los que se analice el momento temporal óptimo para la evaluación de los estudiantes, de tal forma que exista un balance entre la precisión de la clasificación obtenida y el margen de actuación para prevenir un bajo rendimiento académico.

Index Terms—Sistemas de aviso temprano, estudiantes en riesgo, rendimiento de las predicciones, minería de datos en educación, sistemas electrónicos de potencia.

I. INTRODUCCIÓN

Hoy en día, gracias a los constantes avances en las tecnologías de la información y las comunicaciones (TICs), es posible la monitorización continua del rendimiento de los

Este trabajo se ha financiado a través del Plan Estratégico 2017–2020 de la Escuela Politécnica de Cuenca (programa CD–AFP20), así como con fondos propios de la Facultad de Comunicación, Universidad de Castilla-La Mancha

alumnos en su entorno de aprendizaje. En este sentido, existen herramientas TIC que permiten un seguimiento cercano del progreso del alumnado, así como la recopilación efectiva de los resultados de aprendizaje [1]. Este hecho está fomentando la recolección masiva de datos, que permite generar conocimiento sobre cómo aprenden los estudiantes, con el objetivo de mejorar el sistema educativo [2]. De hecho, un campo emergente en los últimos años consiste en la minería de datos en educación (MDE), que aplica diferentes técnicas usadas en la minería de datos tradicional, al campo de la educación [1]. De forma más concreta, *The International Educational Data Mining Society* define a la MDE como una disciplina emergente, preocupada por desarrollar métodos que exploren los tipos de datos provenientes de entornos educativos, con el objetivo de comprender mejor a los estudiantes y su entorno de aprendizaje [3].

La recolección de los datos por medio de la MDE puede utilizarse en diferentes áreas de aplicación. No obstante, uno de los campos más importantes consiste en la predicción del rendimiento académico de los estudiantes [4]. Esta predicción puede darse a diferentes niveles de detalle, intentando predecir las calificaciones de los estudiantes en una determinada tarea, en una asignatura o simplemente estimar si un alumno suspenderá o aprobará el curso al final del año [5]. Cualquiera que sea el caso, resulta deseable la predicción temprana del resultado académico para permitir una enseñanza proactiva y enfocada al estudiante, de tal manera que se maximice la tasa de éxito [6]. En esta línea, algunos estudios recientes han demostrado que algunas herramientas, tales como los sistemas de tutorización inteligente, los sistemas de alarma temprana (SAT) y los sistemas de recomendación, pueden ser realmente útiles en la enseñanza universitaria [7].

Las herramientas SAT ya han sido testeadas con éxito en aprendizaje en formato online, donde los ratios de abandono

son notables [8]. Una de las principales razones de este éxito radica en que las enseñanzas online requieren que el alumnado se conecte y cumplimente diariamente las tareas en las plataformas de e-learning. Por lo tanto, resulta fácil la monitorización continua de esos alumnos, así como la identificación precisa del alumnado que se encuentra en riesgo de abandono [9] y la elaboración de sistemas de intervención para su retención [10]. Sin embargo, estos sistemas no son extrapolables cuando se trata de sistemas de enseñanza presencial, ya que la mayoría de las actividades se llevan a cabo fuera de una plataforma de e-learning [11]. Esto ha originado que en los últimos años, se hayan puesto muchos esfuerzos en la creación de diseños efectivos de SAT enfocados a funcionar en enseñanza tradicional y en asignaturas con altos ratios de suspensos, como pueden ser las que se imparten en estudios pertenecientes a la programación, la física, las matemáticas o la ingeniería.

En lo concerniente al área de ingeniería, los cursos que tratan sobre electrónica de potencia pueden llegar a ser muy desalentadores para los alumnos, ya que a menudo se requiere mucho esfuerzo para conseguir comprender los fenómenos físicos complejos que suceden en las operaciones de muchos circuitos [12]. Efectivamente, los estudiantes necesitan entender y aplicar los principios fundamentales de fenómenos electrostáticos, magnéticos, electromagnéticos o térmicos, además de manejar magnitudes no visibles, como el flujo magnético [12]. En este sentido, se requiere un esfuerzo considerable de abstracción, así como una elevada inversión de tiempo por parte del alumno, que a menudo desemboca en el abandono de la asignatura en una fase temprana o la no adquisición de los conceptos básicos de conocimiento. Por lo tanto, este tipo de cursos podría beneficiarse del uso de herramientas SAT para predecir qué alumnos con bajo rendimiento abandonarían la asignatura.

Desafortunadamente, los sistemas SAT ya diseñados para otros cursos de ingeniería podrían no ser adecuados para la asignatura de electrónica de potencia presentada en este estudio. Efectivamente, numerosos estudios recientes han demostrado que los modelos generalistas de predicción SAT no podrían identificar adecuadamente a los alumnos en riesgo, mientras que los modelos SAT específicos, entrenados para una asignatura en concreto, son mucho más eficientes en la prevención del bajo rendimiento [11]. En este sentido, las diferencias entre distintos cursos en términos de estructura, objetivos, actividades o esquemas de evaluación, han demostrado tener un gran impacto en la selección de los factores más predictivos para cada caso particular [13]. Por lo tanto, en este trabajo se presenta un estudio que explora la posibilidad de detectar estudiantes que podrían suspender una asignatura concreta de circuitos electrónicos de potencia, con el fin de implantar un esquema SAT específico para esta asignatura.

II. MATERIALES Y MÉTODO

II-A. Principales Características del Curso

Sistemas Electrónicos de Potencia es en una asignatura obligatoria de segundo curso impartida en el Grado de Inge-

nería de Tecnologías de Telecomunicación en la Universidad de Castilla-La Mancha (UCLM). Esta asignatura muestra a los estudiantes por primera vez en el grado, los conceptos fundamentales de voltaje y potencia trifásica, transformadores, rectificadores monofásicos y trifásicos, convertidores DC-DC aislados y no aislados, inversores monofásicos y trifásicos, e instalaciones fotovoltaicas.

El curso está dividido en cuatro créditos de teoría, según el Sistema Europeo de Transferencia de Créditos (ECTS) y dos créditos prácticos ECTS. La carga docente se reparte en 40 horas de clase y 20 horas de laboratorio, que se imparten a lo largo de 15 semanas. Por lo tanto, los alumnos han de atender tres sesiones de 90 minutos a la semana, donde dos sesiones están dedicadas a clases teóricas y una sesión está destinada a prácticas de laboratorio. En lo referente a la evaluación de la asignatura, la Tabla I resume las actividades docentes más importantes y su peso dentro de la asignatura. Como puede observarse, el 50 % de la calificación final de la asignatura está repartida a lo largo del curso en diferentes actividades pensadas para fomentar la participación activa del alumnado durante la asignatura [14]. Más concretamente, el esquema de calificaciones se reparte en un 5 % en actividades realizadas en clase (AC), un 5 % en tareas para casa (TC), un 20 % en las actividades de laboratorio (AL), un 20 % en un examen parcial a la mitad del cuatrimestre (EP) y, por último, un 50 % del examen final de la asignatura (EF). El EP se lleva a cabo en la semana 9, de tal forma que los alumnos se examinan de los conocimientos desde el principio del curso, hasta la semana 8, aproximadamente la mitad de la duración del curso. Por otro lado, el EF se lleva a cabo siempre dos semanas después del final de la asignatura.

TABLA I
REPARTO DE PONDERACIONES ENTRE LAS ACTIVIDADES DOCENTES
PLANTEADAS EN CADA CURSO

Metodología docente	Ponderación
Actividades en clase (AC)	5 %
Tareas para casa (TC)	5 %
Actividades en laboratorio (AL)	20 %
Examen parcial (EP)	20 %
Examen final (EF)	50 %
TOTAL	100 %

La evaluación continua del alumno está planteada en una estructura semanal. Así, cada semana comienza con dos clases teóricas donde el profesor enseña conceptos teóricos usando presentaciones en PowerPoint y resuelve en pizarra algunos ejemplos representativos relacionados con la teoría. En los últimos 45 minutos de la segunda sesión de clase, se les pide a los alumnos que resuelvan un ejercicio por ellos mismos. Aunque el trabajo es de carácter individual, los alumnos pueden compartir y discutir posibles soluciones con los compañeros. Esta actividad está enfocada a que el alumnado pueda aplicar de forma práctica lo aprendido en las sesiones teóricas, así como resolver sus dudas particulares. De hecho, mientras que los alumnos resuelven el problema propuesto, el profesor camina entre los pupitres para analizar el progreso general y

resolver dudas particulares de cada alumno. Por otro lado, para animar a los alumnos a continuar practicando los contenidos teóricos vistos en clase, se plantea otro problema práctico distinto para que los alumnos lo resuelvan en su tiempo fuera de clase, a partir de los dos días siguientes a la finalización de la segunda clase teórica. Tanto las AC como las TC son de un nivel similar de dificultad.

La tercera sesión semanal de la asignatura está destinada al laboratorio. En esta clase, los alumnos han de analizar y comprobar el funcionamiento de diferentes circuitos usando la herramienta Simulink de Matlab y montajes reales hechos sobre placas de prototipado. El profesor orienta a los alumnos mediante instrucciones directas para la realización de los ejercicios. Además, los alumnos trabajan en parejas permanentes y antes de que termine la sesión, suben a Moodle las simulaciones y medidas realizadas sobre los circuitos reales, junto con las conclusiones.

Por último, es importante resaltar que todas las actividades realizadas, AC, TC y AL, así como los exámenes parciales de mitad de cuatrimestre fueron corregidos por el profesor en la misma semana en la que el alumno los realizó, asignando puntuaciones en una escala entre 0 y 10.

II-B. Participantes y Variables Predictoras

En este estudio participaron un total de 78 estudiantes matriculados por primera vez en la asignatura durante dos años académicos consecutivos. Es importante destacar que en ambos años académicos, los días de clase a la semana fueron los mismos, repartidos de la misma manera, e impartidos por el mismo profesor. El contenido de la asignatura fue exactamente el mismo en los dos cursos y los temas fueron impartidos en el mismo orden. Además, el sistema de calificaciones descrito anteriormente se mantuvo inalterado de tal forma que las entregas de trabajos fuera de clase, prácticas y exámenes, se llevaron a cabo aproximadamente en las mismas fechas durante el semestre. Tanto el EP como el EF consistieron en cuatro preguntas cortas a contestar y dos problemas con el mismo nivel de dificultad. Por último, destacar que la misma plataforma (Moodle) se usó como herramienta online para intercambio de documentos, desde donde los alumnos se descargaron el material docente y subieron las tareas e informes de laboratorio.

Para aprobar la asignatura los estudiantes tuvieron que cumplir dos requerimientos. Por un lado, la nota global de la asignatura, teniendo en cuenta todos los ítems evaluables, debía ser igual o superior a 5 puntos. Por otro lado, la nota mínima del examen final debía ser igual o mayor que 4 puntos. Para el resto de ítems evaluables no se pusieron requisitos de nota mínima, ya que la regulación de la propia UCLM restringe la asistencia obligatoria para los cursos de grado. Además, tampoco se exigió nota mínima para el examen parcial, pues el objetivo docente de esta prueba era la de dar una posibilidad al estudiante de autoevaluar sus conocimientos de cara al examen final y recibir realimentación de su progreso en la asignatura.

De los 78 estudiantes que participaron en el experimento durante los dos cursos académicos, 48 aprobaron la asignatura (61,5%), mientras que 30 la suspendieron (38,5%), bien porque abandonaron la asignatura o bien porque no cumplieron los requisitos mínimos. Esta clasificación binaria (aprobado o suspenso) es la que se utilizó como variable de salida para el entrenamiento y la validación de los modelos de predicción obtenidos. Además, las calificaciones de AC, TC y AL conseguidas por los alumnos durante las cuatro primeras semanas de clase se utilizaron como variables de entrada para la construcción de los modelos predictivos. Este punto de corte se eligió teniendo en cuenta que en este tiempo se han impartido aproximadamente un tercio de los contenidos de la asignatura y se han recopilado suficientes datos objetivos de los alumnos para el entrenamiento de los modelos, a la vez que no es demasiado tarde para aplicar herramientas SAT en la prevención del fracaso del alumnado.

II-C. Modelos Predictivos

Durante los últimos años, se ha utilizado una amplia variedad de algoritmos de clasificación para predecir el rendimiento académico de los alumnos en educación superior [4]. No obstante, no existe un algoritmo universal que funcione de forma correcta en cada escenario educativo, de tal forma que los investigadores suelen utilizar las técnicas que mejor se adaptan a cada caso particular [5]. En este estudio, se analizaron dos técnicas de clasificación muy habituales y extendidas en la evaluación del rendimiento académico del alumnado: análisis lineal discriminante (AL) y árbol de decisión (AD).

Más concretamente, el AL es una técnica estadística que estima una variable de salida en función de la combinación lineal de varias características de entrada. Su objetivo es predecir la probabilidad de ocurrencia de una variable de salida mediante el ajuste de los datos a través de una ecuación lineal, de tal forma que se minimice el error producido en la clasificación. De esta forma, en el periodo de aprendizaje, cada variable de entrada aporta un peso o coeficiente a la ecuación de discriminación de los datos, dependiendo de la relevancia que tenga en la clasificación. Dada su relativa simplicidad, el método de AL ha sido muy utilizado en la predicción del rendimiento académico [15]. Por otro lado, la clasificación por AD consiste en un conjunto de normas en forma de árbol donde los datos se separan en función de una variable, hasta que cada nodo que forma el árbol contiene sólo muestras de una categoría o todas las variables de entrada han sido utilizadas. Este clasificador se caracteriza por su facilidad de implementación, pues consiste en simples reglas de decisión tipo *if-else*, así como en la facilidad para interpretar los datos por su representación visual [15]. Finalmente, cabe destacar que, a pesar de existir diferentes criterios de puntuación para elegir la variable de entrada que mejor separa los datos en cada nivel del árbol, en este estudio se utilizó el índice de Gini, el más común de todos.

II-D. Descripción del Experimento y Evaluación Estadística

Para llevar a cabo el entrenamiento y la validación de los dos clasificadores, se utilizaron como variables de entrada las puntuaciones en crudo de los alumnos, correspondientes a las actividades de AC, TC y AL a lo largo de las cuatro primeras semanas de cada curso. En este periodo de tiempo, ya se han abordado en clase los conceptos básicos sobre circuitos trifásicos y transformadores, que son esenciales para entender sistemas de potencia más avanzados, como por ejemplo, rectificadores e inversores. Por lo tanto, se usó un total de doce características o variables de entrada para la construcción de los modelos, tal que las cuatro primeras características se correspondieron con las calificaciones conseguidas en AC_w , otras cuatro características fueron extraídas de las calificaciones obtenidas en TC_w y, finalmente, las cuatro últimas características se obtuvieron mediante la calificación de AL_w , siendo $w = [1, 4]$ la semana correspondiente a cada actividad.

Por otro lado, la evaluación del rendimiento de clasificación de los dos modelos estudiados se llevó a cabo cuantificando los estudiantes con bajo rendimiento, o verdaderos positivos (VP), estudiantes con rendimiento normal correctamente identificados, o verdaderos negativos (VN), estudiantes con un rendimiento normal clasificados como alumnado de bajo rendimiento, o falsos positivos (FP) y, finalmente, estudiantes con bajo rendimiento clasificados como estudiantes con rendimiento normal, o falsos negativos (FN). Después, se obtuvieron los valores de sensibilidad (Se), especificidad (Sp) y precisión (Ac) de cada modelo, a través de las ecuaciones 1, 2, y 3, respectivamente.

$$Se = \frac{VP}{VP + FN} \quad (1)$$

$$Sp = \frac{VN}{VN + FP} \quad (2)$$

$$Ac = \frac{VP + VN}{VP + VN + FP + FN} \quad (3)$$

Así, la Se representa la habilidad del modelo de identificar correctamente los estudiantes que podrían suspender la asignatura, la Sp expresa la capacidad del modelo de rechazar correctamente los alumnos que posiblemente aprobarán y la Ac representa la habilidad del algoritmo de diferenciar entre los alumnos que probablemente suspenderán, de los que posiblemente aprobarán la asignatura.

Finalmente, aunque el método de resustitución ha sido ampliamente utilizado en técnicas de MDE, entrenar y validar los algoritmos de clasificación con las mismas muestras a menudo provoca un efecto de sobreentrenamiento de los modelos de clasificación. Así, el modelo resultante ajusta muy bien para las muestras con las que fue entrenado y validado, consiguiendo un alto rendimiento de clasificación, pero falla cuando se testea con nuevas muestras de población [16]. Por lo tanto, para evitar esta limitación, en este estudio se llevó a cabo un procedimiento estadístico de validación cruzada k -fold

[16], con $k = 2$. En este sentido, el conjunto de datos primero se divide de forma aleatoria en 2 subconjuntos del mismo tamaño, pero asegurándose que la pertenencia a las dos clases está repartida en los subconjuntos de igual manera. Después, se llevan a cabo 2 iteraciones de entrenamiento y validación, de tal manera que en la primera iteración el entrenamiento se realiza con un subconjunto de datos, mientras que la validación se lleva a cabo con el subconjunto restante. En la segunda iteración, se invierten los roles de entrenamiento y validación entre los subconjuntos. Finalmente, se calculan los resultados promedio de las métricas de rendimiento (Se , Sp y Ac) para cada iteración y cada modelo de clasificación.

III. RESULTADOS

Un análisis previo de la muestra recolectada reveló que los datos eran homocedásticos y tenían una distribución normal, de tal manera que en la Tabla II se puede observar la media y desviación estándar para cada una de las características estudiadas, en la segunda y tercera columna, respectivamente. De la misma forma, la cuarta columna de la Tabla II muestra la significación estadística de cada variable ($p < 0,05$), calculada a través de un test de *Student*.

Como se puede observar, existe una tendencia decreciente de las calificaciones para AC, TC y AL, conforme va transcurriendo el curso, con un incremento en su desviación estándar. Los resultados muestran que las calificaciones de los alumnos son cada vez más desiguales y con unas calificaciones promedio más bajas para todos los ítems evaluables. Este hecho se confirma observando el valor de significación estadística para cada variable, donde se puede comprobar que la diferencia entre las medias de ambos grupos es más significativa conforme transcurren las semanas de clase. Las actividades de clase requieren una mención especial, se puede observar que a partir de la segunda semana las diferencias entre grupos ya son significativas ($p = 0,02$), llegando a alcanzar una significación de $5,29 \times 10^{-8}$ en la cuarta semana de clase. Por el contrario, las calificaciones de los trabajos encargados para realizar en casa, son las que menores diferencias muestran entre los aprobados y suspensos, teniendo que esperar hasta a la tercera semana de clase para obtener resultados significativos.

Por otro lado, los resultados obtenidos tras la validación de los modelos de clasificación AL y AD, se muestran en la figura 1 y 2, respectivamente, donde se puede observar una matriz de confusión para cada modelo, junto con los ratios de verdaderos positivos y falsos negativos. En primer lugar, el modelo de AL consiguió una precisión global del 79,5% diferenciando los alumnos con bajo rendimiento de aquellos con un aprovechamiento normal de la asignatura. En este modelo, cabe destacar el balance entre la capacidad de rechazar los alumnos que tienen un rendimiento normal, con los estudiantes que posiblemente suspenderán la asignatura. Además, la capacidad de predicción del alumnado que suspenderá es del 84%, frente a la capacidad de predecir satisfactoriamente el aprobado, que es del 72%.

TABLA II
 MEDIA, DESVIACIÓN ESTÁNDAR Y SIGNIFICACIÓN ESTADÍSTICA PARA CADA CARACTERÍSTICA USADA EN EL APRENDIZAJE Y VALIDACIÓN DE LOS CLASIFICADORES

Características	Media calificaciones	Desviación estándar	Significación estadística
AC ₁	7,96	0,36	0,34
AC ₂	7,84	0,38	0,02
AC ₃	7,74	0,43	4,18×10 ⁻⁵
AC ₄	7,61	0,48	5,29×10 ⁻⁸
TC ₁	7,18	0,44	0,76
TC ₂	7,07	0,46	0,11
TC ₃	6,87	0,45	0,01
TC ₄	6,78	0,56	2,0×10 ⁻³
AL ₁	6,52	0,44	0,17
AL ₂	6,41	0,51	0,05
AL ₃	6,30	0,56	6,4×10 ⁻³
AL ₄	6,13	0,55	1,1×10 ⁻³

En cambio, el clasificador AD obtuvo una precisión global de un 69,2%, lo que supone menos de un 10% de rendimiento, con respecto al clasificador anterior. Además, existe un mayor desequilibrio entre la habilidad de identificar estudiantes que suspenderán la asignatura de aquellos que no lo harán. De hecho, la habilidad de clasificar correctamente los estudiantes que suspenderán es del 76%, lo que supone un decremento del 8% en relación al clasificador basado en AL.

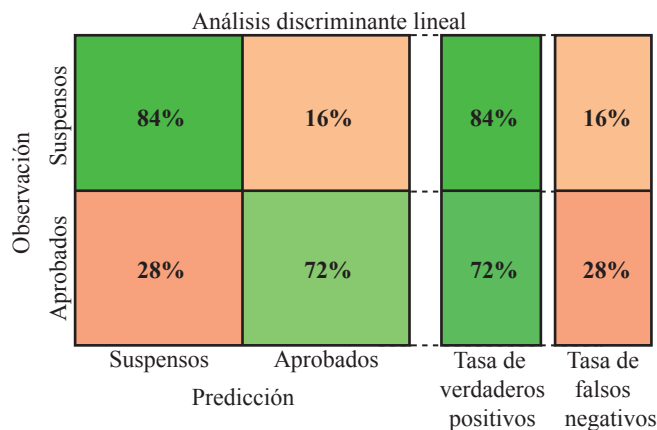


Figura 1. Matriz de confusión obtenida para el modelo de análisis discriminante, junto con sus ratios de aciertos y fallos para los dos grupos.

IV. DISCUSIÓN

Los dos modelos de clasificación presentados en este trabajo fueron entrenados y validados con el objetivo de identificar correctamente los alumnos con bajo rendimiento en el curso de Sistemas Electrónicos de Potencia. Ambos modelos demostraron su efectividad, clasificando correctamente estudiantes con posibilidad de suspender y obteniendo eficiencias de discriminación globales en torno al 75%. No obstante, los resultados de clasificación mostrados por el modelo de AL mejoraron en casi un 10% las del clasificador basado en AD. Además, el modelo de AL consiguió una habilidad de predecir

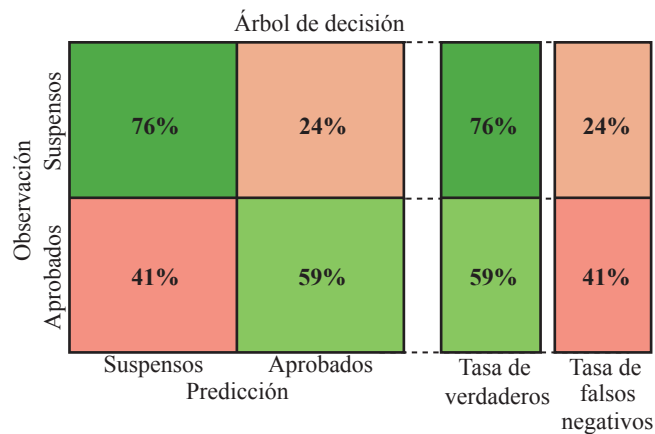


Figura 2. Matriz de confusión obtenida para el modelo de árbol de decisión, junto con sus ratios de aciertos y fallos para los dos grupos.

alumnos suspensos en la asignatura por encima del 84%, frente al 76% del modelo de AD. Esta característica es de especial relevancia en este estudio, porque en los esquemas basados en SAT se persigue sobre todo el incremento de la identificación de los alumnos que podrían suspender la asignatura (es decir, aumentar la *Se*), siempre que la precisión global (*Ac*) se mantenga lo más alta posible, para evitar así demasiadas falsas alarmas [8]. En este sentido, no es muy crítico si se falla en la clasificación de los estudiantes que superarán la asignatura, ya que éstos podrían beneficiarse de recibir enseñanza complementaria y mejorar su rendimiento. Sin embargo, si el error de precisión fuera muy alto, estaríamos invirtiendo tiempo y recursos en estudiantes que no necesitarían de este refuerzo docente extra [17].

Una de las posibles causas de esta diferencia de rendimiento entre clasificadores podría deberse a las limitaciones impuestas en la construcción de los modelos. Así, mientras que en el modelo de AL se incluyeron todas las características y no se impuso ninguna restricción, en el modelo de AD se limitó el crecimiento del árbol a un máximo de cuatro saltos de decisión, tal y como puede observarse en la figura 3. Además, como se puede apreciar, para formar el árbol se utilizaron sólo 3 características de las 12 posibles (AC₄, AL₃ y TC₄). En este sentido, resulta interesante que todas ellas correspondan a variables que cuantifican rendimientos del alumnado a partir de la tercera semana de clase. Efectivamente, cabe destacar que las características estudiadas se mostraron más significativas diferenciando entre grupos de estudio conforme transcurría cada semana. Concretamente, a partir de la segunda semana de clase, las AC y AL se mostraron estadísticamente significativas, mientras que hubo que esperar a la tercera semana de clase para que las TC mostraran el mismo efecto. Esto puede deberse a que las TC no reflejan fielmente el seguimiento de los alumnos en la asignatura, ya que los estudiantes pueden consultar material adicional en casa o pedir ayuda a otros compañeros fuera de clase.

En vista de los resultados, sería posible pensar que conforme

vayan transcurriendo las semanas, la precisión en la clasificación de alumnos con bajo rendimiento se incrementará, pero con el inconveniente de que quedará menos tiempo para implantar medidas correctoras que podrían minimizar la tasa de suspensos. En este sentido, existen trabajos en la literatura que han mostrado el momento ideal de predicción, teniendo en cuenta la precisión alcanzada y el tiempo transcurrido. Por ejemplo, Howard et al [14] concluyó que el momento óptimo para clasificar alumnos con probabilidad de suspender un curso de estadística era la quinta semana. Por otro lado, Lu et al. [18] informó que la sexta semana era el punto de corte ideal en un curso de cálculo avanzado para ingeniería. En el presente estudio, se alcanzaron ratios de alumnos con probabilidad de suspender superiores al 84 %, utilizando exclusivamente las actividades de las cuatro primeras semanas de clase, lo que dejaría margen de tiempo para aplicar acciones correctoras sobre los alumnos que lo necesiten. No obstante, hay que tomar estas comparaciones con mucho cuidado, pues los puntos temporales óptimos para la evaluación de estudiantes depende en gran medida de las características de cada curso y la duración de los mismos.

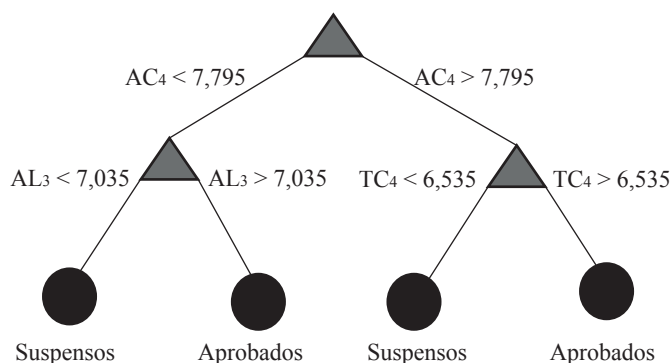


Figura 3. Modelo discriminante basado en un árbol de decisión obtenido considerando las actividades realizadas en clase y las tareas para casa durante la cuarta semana (AC_4 y TC_4 , respectivamente), así como las actividades de laboratorio durante la tercera semana (AL_3).

V. LIMITACIONES

A pesar de lo interesante del experimento, es necesario resaltar algunas limitaciones del estudio. Primero de todo, solamente se han utilizado dos clasificadores básicos, cuyos algoritmos son fáciles de implementar. Sin embargo, sería interesante llevar a cabo estudios adicionales utilizando clasificadores más avanzados, como redes neuronales o máquinas de vector soporte. De la misma forma, solamente se han utilizado algoritmos de clasificación, cuando también se podrían haber tenido en cuenta modelos de regresión, ya usados en otros estudios con la misma finalidad [19]. Además, solamente se han utilizado variables relativas al rendimiento académico, cuando algunos estudios recientes señalan algunas características demográficas o socio-económicas como factores influyentes

en el rendimiento académico [4]. Finalmente, destacar que la detección temprana de estudiantes con bajo rendimiento es sólo el primer paso para la construcción de un modelo SAT eficiente y que, por lo tanto, es necesario llevar a cabo futuras investigaciones sobre la clase de refuerzo académico que deberían recibir estos alumnos.

VI. CONCLUSIONES

En este trabajo se analiza la posibilidad de predecir estudiantes que podrían suspender una asignatura presencial de Sistemas Electrónicos de Potencia. Para llevar a cabo el estudio se programaron dos modelos de clasificación, que dieron unas tasas de precisión globales en la clasificación en torno al 75 %. El punto temporal elegido para hacer esta predicción fue hasta la cuarta semana de clase. Sin embargo, los resultados apuntan a que la precisión podría incrementarse conforme fueran avanzando las semanas del curso. Futuros estudios deberían evaluar el punto óptimo de corte de tal manera que se maximice la precisión y no sea demasiado tarde para actuar desde una perspectiva de refuerzo docente.

VII. AGRADECIMIENTOS

Los autores de este trabajo agradecen a todos los estudiantes que tomaron parte en este estudio.

REFERENCIAS

- [1] A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," *IEEE Access*, vol. 5, pp. 15991–16005, 2017.
- [2] B. Bakhshinategh, O. R. Zaiane, S. ElAtia, and D. Ipperciel, "Educational data mining applications and tasks: A survey of the last 10 years," *Education and Information Technologies*, vol. 23, no. 1, pp. 537–553, 2018.
- [3] [Online]. Available: <http://www.educationaldatamining.org>
- [4] A. Hellas, P. Ihantola, A. Petersen, V. V. Ajanovski, M. Gutica, T. Hyninen, A. Knutas, J. Leinonen, C. Messom, and S. N. Liao, "Predicting academic performance: A systematic literature review," in *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, 2018, pp. 175–199.
- [5] R. Asif, A. Mercer, S. A. Ali, and N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," *Computers & Education*, vol. 113, pp. 177–194, 2017.
- [6] P. A. Murtaugh, L. D. Burns, and J. Schuster, "Predicting the retention of university students," *Research in Higher Education*, vol. 40, no. 3, pp. 355–371, 1999.
- [7] A. Sandoval, C. Gonzalez, R. Alarcon, K. Pichara, and M. Montenegro, "Centralized student performance prediction in large courses based on low-cost variables in an institutional context," *The Internet and Higher Education*, vol. 37, pp. 76–89, 2018.
- [8] W. Chen, C. G. Brinton, D. Cao, A. Mason-singh, C. Lu, and M. Chiang, "Early detection prediction of learning outcomes in online short-courses via learning behaviors," *IEEE Transactions on Learning Technologies*, vol. 12, no. 1, pp. 44–58, 2019.
- [9] J. Kuzilek, M. Hlosta, D. Herrmannova, Z. Zdrahal, and A. Wolff, "Ou analyse: analysing at-risk students at the open university," *Learning Analytics Review*, pp. 1–16, 2015.
- [10] C. Burgos, M. L. Campanario, D. de la Pena, J. A. Lara, D. Lizcano, and M. A. Martínez, "Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout," *Computers & Electrical Engineering*, vol. 66, pp. 541–556, 2018.
- [11] F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, "Models for early prediction of at-risk students in a course using standards-based grading," *Computers & Education*, vol. 103, pp. 1–15, 2016.
- [12] R. Aldunate and M. Nussbaum, "Teacher adoption of technology," *Computers in Human Behavior*, vol. 29, no. 3, pp. 519–524, 2013.

- [13] D. Gašević, S. Dawson, T. Rogers, and D. Gasevic, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," *The Internet and Higher Education*, vol. 28, pp. 68–84, 2016.
- [14] E. Howard, M. Meehan, and A. Parnell, "Contrasting prediction methods for early warning systems at undergraduate level," *The Internet and Higher Education*, vol. 37, pp. 66–75, 2018.
- [15] G. Bonaccorso, *Machine learning algorithms*. Packt Publishing Ltd, 2017.
- [16] Y. Jung and J. Hu, "A K-fold averaging cross-validation procedure," *Journal of Nonparametric Statistics*, vol. 27, no. 2, pp. 167–179, 2015.
- [17] Y.-H. Hu, C.-L. Lo, and S.-P. Shih, "Developing early warning systems to predict students' online learning performance," *Computers in Human Behavior*, vol. 36, pp. 469–478, 2014.
- [18] O. H. Lu, A. Y. Huang, J. C. Huang, A. J. Lin, H. Ogata, and S. J. Yang, "Applying learning analytics for the early prediction of students' academic performance in blended learning," *Journal of Educational Technology & Society*, vol. 21, no. 2, pp. 220–232, 2018.
- [19] G. Kostopoulos, S. Kotsiantis, C. Pierrakeas, G. Koutsonikos, and G. A. Gravvanis, "Forecasting students' success in an open university," *International Journal of Learning Technology*, vol. 13, no. 1, pp. 26–43, 2018.