

Tipo de artículo: Artículo de revisión

Empleo de las redes bayesianas para apoyar la toma de decisiones sobre la propagación de la Covid-19

Use of Bayesian networks to support decision-making on the spread of Covid-19

Freddy E. Torres Cordero^{1*} , <https://orcid.org/0000-0002-5018-4632>
Neilys González Benítez² , <https://orcid.org/0000-0001-8691-445X>
Omar Mar Cornelio³ , <https://orcid.org/0000-0002-0689-6341>

¹ Especialista en Medicina Interna y Especialista en Cuidados Intensivos del Adulto Policlínica Docente Universitaria Pedro Borrás Astorga, Provincia, Pinar del Río, Cuba.

² Doctora en Ciencias, Centro Meteorológico de Pinar del Río, Cuba.

³ Centro de Estudio de la Matemática Computacional. Facultad de Ciencias y Tecnologías Computacionales. Universidad de las Ciencias Informáticas. Cuba.

* Autor para correspondencia: freddyelie@nauta.cu

Resumen

Las redes bayesianas proveen información sobre las relaciones de dependencia e independencia condicional existentes entre las variables. La inclusión de las relaciones de independencia en la propia estructura de la red, hace de las redes bayesianas una buena herramienta para representar conocimiento de forma compacta pues se reduce el número de parámetros necesarios. En general, su aplicación para apoyar la toma de decisiones sobre la propagación de la Covid-19 es escasa. La toma de decisiones para la Covid-19, ha estado centrada en el uso de modelos matemáticos que utilizan datos discretos en la búsqueda del conocimiento sobre la evolución de la epidemia, considerando a los sujetos susceptibles, expuestos, infectados y recuperados. En el presente trabajo se realiza un estudio de las redes bayesianas como una herramienta para resolver distintos problemas en los modelos matemáticos que utilizan datos discretos en la búsqueda del conocimiento sobre la evolución de la epidemia: considerando a los sujetos susceptibles, expuestos, infectados y recuperados. La posibilidad de trabajar con datos discretos y continuos simultáneamente, la variedad de problemas que pueden resolver, y la flexibilidad en la estructura del modelo, convierten a las redes bayesianas en una herramienta apropiada, en los modelos para apoyar la toma de decisiones sobre la propagación de la Covid-19.

Palabras clave: Modelos gráficos probabilísticos; toma de decisiones; Covid-19.

Abstract

They sum up the Bayesian's nets decide information on the relations of dependence and conditional existent independence between the variables. The inclusion of the relations of independence in the own structure of the net, does of the Bayesian's nets a good tool to represent cone, foundation in a compact way because reduces to him the number of necessities parameters. In general, your application to lean it takes of decisions on the propagation of the Covid -19 is scarce. It takes it of decisions for the Covid -19, has being centered in the use of mathematical models that use discrete data in the search of the knowledge on the evolution of the epidemic, by considering to the susceptible subjects, exposed, infected and recovered. At present work proposes insert the Bayesian's nets as a credit tool to resolve different problems in the mathematical models that use discrete data in the search of the knowledge on the evolution of the epidemic: by considering to the susceptible subjects, exposed, in-fects-two and recovered. The possibility to work with discrete and endless data simultaneously, the variety of problems that can resolve, and the



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

flexibility in the structure of the model, convert to the Bayesian's nets in an appropriate tool, in the models to lean it takes of decisions on the propagation of the Covid, 19.

Keywords: *Graphic models probabilistic; take of decisions; Covid-19.*

Recibido: 18/02/2021
Aceptado: 20/04/2021

Introducción

La Covid-19 es una enfermedad producida por el virus SARS-CoV-2. Este virus pertenece a la familia Coronaviridae de ácido ribonucleico (ARN) monocatenario. Desde una perspectiva virológica, ha presentado múltiples mutaciones entre las que se encuentran la nueva cepa de Covid-19 (Velavan & Meyer, 2020). La Covid-19, es un evento sanitario que se originó en Wuhan, China en diciembre de 2019 y se sospecha que su propagación se debió a una zoonosis entre murciélagos, un animal indeterminado como vector y humanos (Singhal, 2020).

En Cuba, este evento se comenzó a monitorizar desde el mes de marzo de 2020, cuando se notificaron las primeras personas contagiadas de Covid-19. Los síntomas de las personas que poseen la Covid-19 son fiebre, cefalea, rinitis, disnea, mialgia y tos (Singhal, 2020), (Mar Cornelio et al., 2020). El virus tiene un periodo de incubación de 5 a 7 días.

A nivel mundial, la Covid-19, ha infectado a más de un millón de personas, se han reportado elevados fallecimientos a causa de este virus y con una tasa de letalidad alta y variable, debido a la localización geográfica.

A pesar de tener una baja letalidad, la Covid-19, presenta un contagio de dos a 3 veces mayor en comparación a la influenza (Organization, 2020). Una de las medidas ampliamente utilizada ha sido el aislamiento social, generando una contracción económica mundial y una alta demanda de los servicios de salud para las poblaciones afectadas (Jackson et al., 2020).

Debido al comportamiento del virus, se hace necesario contar con proyecciones con el fin de estimar la magnitud de la propagación del virus. Reviste una significativa importancia la utilización de modelos matemáticos para apoyar la toma de decisiones, con el fin de evitar la propagación del agente viral.

Al respecto, se utilizan los modelos matemáticos para conocer pronósticos y se utilizan para comprender los procesos epidémicos en pos de generar estrategias para disminuir la diseminación del virus. Los modelos matemáticos en salud han sido utilizados desde las primeras aproximaciones de Daniel Bernoulli con la epidemia de la peste durante el siglo XVII (Casals et al., 2009), (Mar Cornelio, 2019).

Los modelos epidemiológicos predictivos deterministas, fueron utilizados según las características del agente etiológico (Casals et al., 2009). Entre los modelos utilizados desde el punto de vista epidemiológico se encuentran; el



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

modelo SIS(Susceptible, infectado, susceptible), útil para el análisis de enfermedades infecciosas que no generan inmunidad, el modelo SIR(Susceptible, infectado, recuperado), útil para el análisis de enfermedades infecciosas que generan inmunidad y las personas son declaradas como “recuperadas” y el modelo SEIR(Susceptible-Expuesto-Infectado-Recuperado), útil para el análisis de enfermedades con periodos de latencia, por lo que se reconoce el estado de “expuestos”.

Los modelos referidos presentan características variables y no son suficientes para apoyar la toma de decisiones sobre la propagación de la Covid-19. Hecho que muestra la necesidad de una herramienta de apoyo para el análisis certero de la propagación de la Covid-19, con el fin de obtener un mejor grado de acierto en las propagaciones de la enfermedad y que contribuya a apoyar la toma de decisiones, con el fin de disminuir las afectaciones que esta enfermedad provoca.

La utilización de redes bayesianas (RB) y el desarrollo acelerado de la tecnología ha permitido el almacenamiento y tratamiento de grandes volúmenes de información compuesta de diferentes tipos de datos, los cuales no siempre son tan precisos y completos como se hace necesario. Con frecuencia se dispone de datos imperfectos ya que habitualmente se emplean cuantificadores imprecisos (ejemplo: elevado, bastante joven, muy mayor, muy pequeño, demasiado grande, etc.), en ocasiones se cometen errores en los instrumentos de medida utilizados al obtener las informaciones, se presentan ausencia de datos, etc.

Las RB, son una herramienta estadística aplicada en entornos bajo condiciones de incertidumbre. En una revisión llevada a cabo por (Aguilera et al., 2011), se observa que las RB, han sido muy desarrolladas a nivel matemático, mientras que su aplicación en el campo de la modelización es escasa.

Más del 80 % de los artículos revisados emplean datos de tipo discreto, pese a la capacidad de las RB para tratar datos continuos e, incluso, híbridos (datos continuos y discretos simultáneamente). Puesto que fueron originariamente propuestas para el tratamiento de variables discretas, existe una amplia bibliografía, metodología y software al respecto (Nyberg et al., 2006).

Sin embargo, es muy común que los datos epidemiológicos presenten variables continuas y discretas simultáneamente. La solución más utilizada es la discretización de las variables continuas y su tratamiento como discretas. Este proceso supone una pérdida de información estadística, que reduce la exactitud y precisión del modelo final (Uusitalo, 2007).

A pesar de que la imperfección está presente de forma directa en los datos, todavía el número de técnicas dentro de la disciplina del Análisis Inteligente de Datos que permitan el tratamiento de forma explícita de la información es bastante escaso (González Benítez, 2017). De esta forma, si las técnicas no son capaces de tratar con datos



imperfectos, estos datos son transformados en precisos y durante este proceso de transformación es posible que se produzca una pérdida de información relevante que afecte de forma directa a la calidad de los resultados esperados (Martínez España, 2015).

Materiales y métodos

Para apoyar la toma de decisiones sobre la propagación de la Covid-19, se propone el uso de RB como una herramienta probabilística, capaz de aprender de datos imperfectos, tal y como lo propuso (González Benítez, 2017), con una estructura gráfica y fundamento probabilístico que contiene el proceso de gestión del conocimiento y Análisis Inteligente de los Datos (AID) para obtener información relevante que sea capaz de modelar el conocimiento



adquirido sobre la evolución de la epidemia de la Covid-19, (Fig. 1).

Figura 1. Marco de trabajo para el uso de una Red Bayesiana como herramienta probabilística en el apoyo a la toma de decisiones sobre la propagación de la Covid-19.

El marco de trabajo para el uso de una Red Bayesiana como herramienta probabilística en el apoyo a la toma de decisiones sobre la propagación de la Covid-19, está compuesto por tres componentes fundamentales. Los componentes son:

1. Gestión del conocimiento
2. Análisis Inteligente de los Datos
3. Red bayesiana

Cada componente se describe a continuación:

1. Componente de Gestión del Conocimiento



Constituye un pilar para el desarrollo de los procesos que lo integran pues ellos tienen como base el empleo del conocimiento para la toma de decisiones. Las herramientas utilizadas en el centro objeto de estudio hacen un uso limitado de la información y el conocimiento. Ante esta situación una acción indispensable fue potenciar la gestión de la información y el conocimiento (seleccionarla, organizarla, socializarla a todos los que la necesiten).

La información gestionada es amplia y lo que se trata es de lograr una mayor utilización de la misma para apoyar la toma de decisiones. El componente de gestión del conocimiento describe los elementos fundamentales a tener en cuenta para ser incluidos en los sistemas de información y vigilancia de la propagación de la Covid-19. Esto permite una gestión más adecuada de la información y el conocimiento, considerado como imprescindible para apoyar la toma de decisiones.

En el proceso de atención de los pacientes enfermos, se gestiona información relacionada con los síntomas que se observan en los pacientes. La información del laboratorio se enriquece a partir del proceso de evaluación de los pacientes enfermos, incrementándose el conocimiento de los especialistas de salud para ejecutar el diagnóstico de la enfermedad.

La información de frecuencia de aparición de la Covid-19, se gestiona a través de las tasas de incidencia (ecuación 1), que mide la probabilidad de que los pacientes sanos desarrollen enfermedades durante un período de tiempo específico. Representa el número de casos nuevos de la enfermedad en una población durante un período de tiempo definido.

$$\text{Incidencia} = \frac{\text{número de casos nuevos}}{\text{población en riesgo}} \text{ en un período de tiempo(1)}$$

En la Tabla 1, se muestra la información y conocimiento a gestionar como parte del componente de gestión del conocimiento.

Tabla 1. Información y conocimiento a gestionar para el uso de una Red Bayesiana en el apoyo a la toma de decisiones sobre la propagación de la Covid-19.

Categorías del conocimiento	Descripción	Ubicación	Información y conocimiento a gestionar
Explicito	Incrementa la eficiencia y eficacia organizacional gestionando el conocimiento, codificándolo a través del uso de las tecnologías de la información.	Información de pacientes enfermos y en la información de laboratorio.	Casos nuevos, recuperados y en situación, y los síntomas que se observan en los pacientes con probabilidad de padecimiento de la Covid-19. Procedimientos, para evolucionar los pacientes con probabilidad de padecimiento de la Covid-19.
Tácito	Toma un enfoque de	Información de los pacientes	Existencia de casos de nuevos



	personalización, es aquel que no puede ser fácilmente explicado, es casi propio, intrínseco.	susceptibles a la enfermedad y de posibles muertes por Covid-19.	pacientes enfermos, sobre la población en riesgo, en un período de tiempo. Información que se corresponde con la incidencia de la enfermedad. Número existente de casos sobre la población total en un punto de tiempo.
--	--	--	--

El conocimiento gestionado se socializa a través de la participación activa de los especialistas de salud en capacitaciones relacionadas con la gestión de la información para ejecutar diagnóstico de enfermedades, desde los síntomas que se observan en los pacientes enfermos, en aras de apoyar la toma de decisiones. Proporciona una representación simplificada de la realidad, para obtener mejoras en los diagnósticos de enfermedades.

2. Componente Análisis Inteligente de los Datos

Los datos de entrada seleccionados de las fuentes mencionadas se pre procesaron para eliminar los datos que causen contratiempo en los resultados esperados, en aras de garantizar su disponibilidad, completitud y fidelidad.

El pre procesamiento de datos, para apoyar la toma de decisiones sobre la propagación de la Covid-19, se ejecuta utilizando la tarea de limpieza de datos, para luego incorporar los datos limpios a la base de datos creada previamente.

Esta tarea se realiza de forma automática, utilizando el algoritmo K-Means.

El algoritmo K-Means tienen como base la optimización de una función criterio, para la toma de decisiones sobre la propagación de la Covid-19, la función de criterio se denomina F , el valor de esta función depende de las particiones del conjunto de datos $\{C_1, \dots, C_k\}$.

$$F: P_k(X) \rightarrow \mathbb{R} \quad (2)$$

Donde:

$P_k(X)$, son las particiones del conjunto de datos $X = \{x_1, \dots, x_n\}$ en K grupos no vacíos. x_i , es un vector n -dimensional (objeto) del conjunto de datos X .

El algoritmo K-Means converge a un mínimo local, utilizando la función criterio F , de la sumatoria de las distancias L_2 entre cada objeto y su centroide más cercano. A este criterio normalmente se le denomina error cuadrático y se obtiene a través de la expresión 3.

$$F(\{C_1, \dots, C_k\}) = \sum_{i=1}^k \sum_{j=1}^{p_i} \|x_{i_j} - \bar{c}_i\|^2 \quad (3)$$

Donde:



K es el número de grupos, p_i es el número de objetos del grupo i , x_{ij} es el j -ésimo objeto del i -ésimo grupo y \bar{C}_i es el centroide del i -ésimo grupo el cual es calculado a través de la expresión 4.

$$\bar{C}_i = \frac{1}{p_i} \sum_{j=1}^{p_i} X_{ij}, i = 1, \dots, K(4)$$

El conjunto de pasos lógicos del algoritmo K-Means es el que se presenta a continuación:

Paso 1. Selecciona los K centroides iniciales $\{C_1, \dots, C_k\}$.

Paso 2. Asigna los objetos x_i del conjunto de datos X , a su centroide más cercano.

Paso 3. Recalcula los nuevos centros, regresa al paso 2, hasta que el algoritmo converge.

El algoritmo se inicia seleccionando o calculando los centroides iniciales, dependiendo del criterio de selección de centroides, posteriormente asigna los objetos a su centroide más cercano, para después recalcular los nuevos centroides esto lo realiza hasta que el algoritmo converja (paso 3).

Pre procesados los datos y aplicada la técnica de limpieza de datos, los datos faltantes se rellenan, utilizando el método de imputación por media. Método que sustituye los valores faltantes de una variable mediante la media de las unidades observadas en esa variable.

Para apoyar la toma de decisiones sobre la propagación de la Covid-19, se realiza una imputación por media condicional, método que imputa medias condicionadas a valores observados y consiste en agrupar los valores observados y no observados en clases e imputar los valores faltantes por la media de los valores observados en la misma clase. El método de imputación de valores faltantes contribuye a reducir la pérdida de los datos faltantes en la base de datos (Castro et al., 2014), (Mar et al., 2020).

3. Componente Redes bayesianas

Las Redes bayesianas como herramienta probabilística para apoyar la toma de decisiones sobre la propagación de la Covid-19, son herramientas de modelado estadístico, destinadas a representar un conjunto de incertidumbre relacionada. Su estructura gráfica y su fundamento probabilístico las hace apropiadas para modelar sistemas multivariados orientados al diagnóstico médico para apoyar la toma de decisiones.

El proceso de construcción de redes bayesianas consta de tres fases:

1. Definición del grafo; se identifican las variables del problema y las relaciones de dependencia e independencia entre ellas.
2. Identificación de los modelos canónicos; para ajustar los grupos de variables que se representan en la red bayesiana y reducir el número de parámetros a obtener.



3. Obtención de datos cuantitativos; se obtienen las probabilidades a priori de las variables que no tienen padres y las probabilidades condicionadas para el resto de las variables.

En la construcción de una red bayesiana se crean relaciones de inferencia causal entre dos variables. La presencia o ausencia de que una enfermedad tiene inferencia en que el resultado de las pruebas de laboratorio sea positivo o negativo y definen reglas de diagnóstico del tipo “sí, entonces”.

Definido el grafo, se introduce la información numérica en la red bayesiana, proceso que constituye el modo de aprendizaje de la red bayesiana propuesta y que se corresponde con la tabla de probabilidad condicionada (TPC) de cada nodo. La TPC se obtiene en base a la experiencia de los médicos y en base a los datos utilizando el teorema de Bayes, donde se debe conocer:

- Prevalencia: cantidad expresada en % de personas que tuvieron la enfermedad en un período determinado.
- Sensibilidad: probabilidad de que el resultado sea positivo, si los pacientes tienen la enfermedad.
- Especificidad: probabilidad de que el resultado sea negativo, si los pacientes no tienen la enfermedad.

Obtenida las TPC, se realiza la inferencia bayesiana, la que se utiliza para la búsqueda de la hipótesis más factible, dentro de un conjunto de datos y conocimientos previos de la probabilidad de cada hipótesis. Como método bayesiano permite modificar valores cuando se dispone de nueva información.

Realizada la inferencia bayesiana, se comprueba que se tiene evidencia de Covid-19, útil para el cálculo de la probabilidad a priori de Covid-19. Obtenidos los resultados de las probabilidades a priori para la Covid-19 dado los síntomas que con frecuencia se presentan en las personas que padecen la enfermedad, se introducen hallazgos para obtener las probabilidades a posteriori (condicional).

Al contar con las probabilidades a priori y a posteriori, se calcula el valor predictivo positivo (VPP) que representa la certeza con la que se diagnostica la enfermedad, cuando los síntomas están presentes, si: $P(\text{Enfermedad}=\text{sí} | \text{síntoma}=\text{sí})$, obteniéndose así probabilidad de la presencia de la Covid-19. El valor predictivo, es útil para poder apoyar la toma de decisiones sobre la presencia de Covid-19, dado síntomas presentes y ausentes en personas enfermas.

Resultados y discusión

El uso de las técnicas de IA y en particular de las redes bayesianas, en los sistemas de ayuda a la toma de decisiones, para el diagnóstico médico, tiene como propósito mejorar la calidad de los procesos médicos – asistenciales en la predicción y clasificación de enfermedades, con el fin de enriquecer el cuidado de los enfermos (Monsalve Torra, 2017). Por tal motivo para apoyar la toma de decisiones sobre la propagación de la Covid-19, se describió la metodología a seguir en epígrafe anterior.



La red bayesiana propuesta, es un grafo que contiene información pre procesada con el fin de que la imperfección presente en los datos sea mínima, lo que contribuye al diagnóstico certero dados síntomas, signos y hallazgos de la presencia de la Covid-19. En efecto, de acuerdo con los datos entregados por el Ministerio de Salud Pública y la modelación con una generalización de Susceptible-Expuesto-Infectado-Recuperado en diferentes escenarios, se constató variabilidad en los resultados obtenidos, siendo alarmante la proyección de fallecidos y la cantidad de enfermos activos.

La red bayesiana para apoyar la toma de decisiones sobre la propagación de la Covid-19, que se propone se apoya en la guía confeccionada por los Centros para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés) de los EEUU, para clarificar los síntomas de las enfermedades que tienen a confundir en sus inicios sin realizar pruebas diagnósticas de laboratorio (Figura 2). La guía se confeccionó con el fin de que las personas pudieran diferenciar los síntomas de la Covid-19, los síntomas de un resfriado común, los síntomas de alergia o los síntomas de una gripe.

Figura 2. Síntomas comunes, útiles para conocer la probabilidad de presencia de la Covid-19, resfriado común o alergia. Fuente: CDC, NIH.

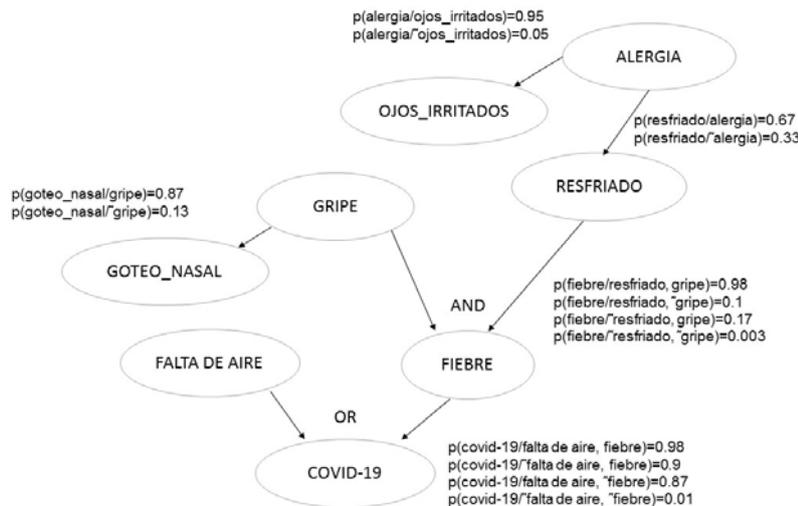


Identificadas las variables del problema y las relaciones de dependencia e independencia entre ellas, se construye la red bayesiana para apoyar la toma de decisiones sobre la propagación de la Covid-19. El grafo obtenido con sus respectivos resultados probabilísticos se muestra en la Figura 3.



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

Figura 3. Grafo para apoyar la toma de decisiones sobre la propagación de la Covid-19.



Se muestra (Figura 3), que la Covid-19, es la enfermedad donde los síntomas con mayor relevancia son: Fiebre y Falta de aire, por lo que los arcos representan una dirección de diagnóstico. El síntoma de Fiebre, es un síntoma común que identifica la presencia de gripe y resfriado, fiebre y resfriado son comunes para Alergia y Gripe, por lo que la dirección del arco es desde el síntoma hacia la enfermedad, representando una dirección de diagnóstico.

En el grafo mostrado en la figura (Figura 3), se introduce la información numérica en la red bayesiana, que se corresponde con la tabla de probabilidad condicionada (TPC) de cada nodo. La TPC se obtiene en base a la experiencia de los médicos y en base a los datos utilizando el teorema de Bayes.

La muestra seleccionada fue un grupo de pacientes con los síntomas referidos en el periodo de aparición de la Covid-19, en Cuba. El grupo de pacientes estuvo conformado por diferentes grupos etarios, para la modelación se utilizaron 520 pacientes acogidos en el hospital de aislamiento de la Universidad de las Ciencias Informáticas de la Habana.

Se obtiene que los síntomas de mayor peso sean fiebre y falta de aire. El resto de los síntomas como la tos, la fatiga y la debilidad son propias de las enfermedades respiratorias, la diferencia radica en que, si se experimentan los síntomas descritos y no poseen falta de aire, entonces los pacientes pueden estar padeciendo de una gripe.

Otros síntomas a tener en cuenta son los ojos irritados, si se está en presencia de dicho síntoma, pueden las personas estar padeciendo de alergia. Otro síntoma de alergia son los estornudos y el goteo nasal.



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

Si las personas con padecimientos de estos síntomas, si no tienen los ojos irritados, podría tratarse de un resfriado común. Otros síntomas del resfriado son los estornudos, el goteo nasal y una leve molestia en el pecho.

Discusiones

De acuerdo con (Casals et al., 2009), los modelos matemáticos/epidemiológicos son limitados en el entendimiento global de los eventos epidémicos, sus aproximaciones se realizan bajo un paradigma cuantitativo y las consecuencias de la Covid-19 son multidimensionales. En ese sentido, el empleo de las RB, para apoyar la toma de decisiones sobre la propagación de la Covid-19, constituye un aporte analítico provechoso toda vez que se incluyan en otras investigaciones, discusiones o aproximaciones, elementos socio-estructurales y de determinación social.

Resulta imperativo incluir estos elementos en el análisis global del Covid-19, debido al análisis cuantitativo hegemónico, observado en este evento. Las redes bayesianas poseen ventajas, frente a los modelos matemáticos/epidemiológicos, debido a la facilidad que poseen para plantear escenarios futuros sin tener que introducir información en todas las variables, e independientemente de que las variables sean discretas o continuas.

El proceso de aprendizaje en RB, puede llevarse a cabo de diversas formas, a partir de los datos, mediante expertos o una combinación de ambos, adoptando estructuras muy variadas en cuanto a complejidad, desde modelos fijos sencillos como el naïve Bayes, a estructuras adaptadas a los datos con un alto número de relaciones entre las variables. En la actualidad, el tratamiento de datos continuos a través de distintos modelos como el Gausiano, o los modelos “Mixtures of Truncated Exponentials”, “Mixtures of Polynomials” o “Mixtures of Truncated Basis Functions”, y su aplicación en el campo de la modelización de la Covid-19, constituye una nueva línea de investigación.

Las características referidas sobre las RB, las hacen útiles como modelos probabilísticos, no sólo aplicable en diagnósticos médicos, sino también en los estudios relacionados con en distintas áreas del conocimiento.

Conclusiones

Se estudian las redes bayesianas para insertarlas como herramienta para resolver distintos problemas en los modelos matemáticos que utilizan datos discretos en la búsqueda del conocimiento sobre la evolución de la Covid-19. La posibilidad de trabajar con datos discretos y continuos simultáneamente, la variedad de problemas que pueden



resolver, y la flexibilidad en la estructura del modelo, convierten a las redes bayesianas en una herramienta apropiada, en los modelos para apoyar la toma de decisiones sobre la propagación de la Covid-19.

Se modelizó una muestra significativa de pacientes con probabilidad de padecimiento de la Covid-19, y se obtuvo que los síntomas de mayor peso fueron la fiebre y la falta de aire, otros síntomas son comunes para enfermedades respiratorias, no obstante, se tuvieron en cuenta para la toma de decisiones con pacientes susceptibles a la enfermedad. La propuesta de red bayesiana con datos analizados previamente y gestionados tiene un impacto positivo en la certeza del diagnóstico de la Covid-19, lo que incide favorablemente en la toma de decisiones.

Conflictos de intereses

Los autores declaran que no poseen conflictos de intereses.

Contribución de los autores

Conceptualización: Freddy E. Torres Cordero, Neilys González Benítez.

Curación de datos: Freddy E. Torres Cordero.

Análisis formal: Neilys González Benítez.

Investigación: Freddy E. Torres Cordero, Neilys González Benítez.

Metodología: Neilys González Benítez.

Administración del proyecto: Neilys González Benítez.

Software: Freddy E. Torres Cordero.

Supervisión: Freddy E. Torres Cordero.

Validación: Neilys González Benítez, Omar Mar Cornelio.

Visualización: Neilys González Benítez, Omar Mar Cornelio.

Redacción – borrador original: Freddy E. Torres Cordero, Neilys González Benítez, Omar Mar Cornelio.

Redacción – revisión y edición: Freddy E. Torres Cordero, Neilys González Benítez, Omar Mar Cornelio.

Financiamiento

La investigación no requirió fuente de financiamiento externa.

Referencias



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)

- Aguilera, P. A., Fernández, A., Fernández, R., Rumí, R., & Salmerón, A. (2011). Bayesian networks in environmental modelling. *Environmental Modelling & Software*, 26(12), 1376-1388. <https://www.sciencedirect.com/science/article/pii/S1364815211001472>
- Casals, M., Guzmán, K., & Caylà, J. A. (2009). Mathematical models used in the study of infectious diseases. *Revista española de salud pública*, 83(5), 689-695. <https://europepmc.org/article/med/20111817>
- Castro, M., Gude, F., & Pérez, A. (2014). *Imputación de datos faltantes en un modelo de tiempo de fallo acelerado* Tesis de fin de Máster en Técnicas Estadísticas). 2014. Universidad de ...].
- González Benítez, N. (2017). Modelo basado en redes bayesianas para el diagnóstico de la Fasciolosis Bovina. <https://repositorio.uci.cu/handle/123456789/7936>
- Jackson, J. K., Weiss, M. A., Schwarzenberg, A. B., & Nelson, R. M. (2020). Global economic effects of COVID-19. https://case.house.gov/uploadedfiles/r46270_-_global_impact.pdf
- Mar Cornelio, O. (2019). Modelo para la toma de decisiones sobre el control de acceso a las prácticas de laboratorios de ingeniería de control II en un sistema de laboratorios remoto.
- Mar Cornelio, O., Gulín González, J., Bron Fonseca, B., & Garcés Espinosa, J. V. (2020). Sistema de apoyo al diagnóstico médico de COVID-19 mediante mapa cognitivo difuso. *Revista Cubana de Salud Pública*, 46(4).
- Mar, O., Ching, I., & González, J. (2020). Operador por selección para la agregación de información en Mapa Cognitivo Difuso. *Revista Cubana de Ciencias Informáticas*, 14(1), 20-39. <https://rcci.uci.cu/?journal=rcci&page=article&op=download&path%5B%5D=1876&path%5B%5D=794>
- Martínez España, R. (2015). Metodologías basadas en minería de datos para el diseño y optimización de técnicas de clasificación automática. *Proyecto de investigación*. <https://digitum.um.es/digitum/bitstream/10201/42697/1/TesisDoctoralRaquelMartinez.pdf>
- Monsalve Torra, A. (2017). Sistemas de ayuda a la decisión clínica en enfermedades de diagnóstico complejo. https://rua.ua.es/dspace/bitstream/10045/65334/1/tesis_monsalve_torra.pdf
- Nyberg, J. B., Marcot, B. G., & Sulyma, R. (2006). Using Bayesian belief networks in adaptive management. *Canadian Journal of Forest Research*, 36(12), 3104-3116. <https://www.fs.usda.gov/treearch/pubs/download/29727.pdf>
- [Record #3939 is using a reference type undefined in this output style.]
- Singhal, T. (2020). A review of coronavirus disease-2019 (COVID-19). *The indian journal of pediatrics*, 87(4), 281-286. <https://link.springer.com/article/10.1007/s12098-020-03263-6>



Uusitalo, L. (2007). Advantages and challenges of Bayesian networks in environmental modelling. *Ecological modelling*, 203(3-4), 312-318. <https://www.sciencedirect.com/science/article/pii/S0304380006006089>

Velavan, T. P., & Meyer, C. G. (2020). The COVID-19 epidemic. *Tropical medicine & international health*, 25(3), 278. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7169770/>



Esta obra está bajo una licencia *Creative Commons* de tipo **Atribución 4.0 Internacional** (CC BY 4.0)