



Estudio y selección de las técnicas de Inteligencia Artificial para el diagnóstico de enfermedades

Study and selection of artificial intelligence techniques for the diagnosis of diseases

Neily González Benítez,¹Vivian Estrada Sentí,²Aylin Febles Estrada ³

¹ Licenciada en Informática. Especialidad Informática y en Meteorología. Doctora en Ciencias Técnicas. Centro Meteorológico Provincial. Pinar del Río. Cuba. neilysgonzalezbenitez@gmail.com

² Ingeniera en Ciencias Informáticas. Doctora en Ciencias Informáticas. Profesora Titular. Universidad de las Ciencias Informáticas. La Habana. Cuba. vivian@uci.cu, vivianestrada4@yahoo.es

³ Licenciada en Informática. Especialidad Informática. Doctora en Ciencias Técnicas. Presidenta de la Unión de Informáticos de Cuba. afebles@gmail.com, ailyn.febles@uniondeinformaticos.cu

Recibido: 19 de enero de 2018

Aprobado: 12 de abril de 2018

RESUMEN

Introducción: el proceso de diagnóstico de enfermedades es complejo, ya que, a menudo, los datos médicos y la información pueden poseer incertidumbre, y requieren ser tratados con técnicas de Inteligencia Artificial en aras de asistir con mayor certeza al apoyo hacia la toma de decisiones.

Citar como: González Benítez N, Estrada Sentí V, Febles Estrada A. Estudio y selección de las técnicas de Inteligencia Artificial para el diagnóstico de enfermedades. Rev Ciencias Médicas. 2018; 22(3). Disponible en: www.revcmpinar.sld.cu/index.php/publicaciones/article/view/3465

Objetivo: comparar las técnicas de Inteligencia Artificial que con frecuencia se utilizan para el diagnóstico de enfermedades, cuando se tienen datos almacenados sobre el comportamiento de las enfermedades que con frecuencia afectan a una población en cuestión.

Método: es una investigación apoyada en la Teoría de la Decisión Multicriterio Discreta, útil para tomar decisiones, en particular, sobre la técnica de Inteligencia Artificial que mejor se ajusta para ejecutar diagnóstico de enfermedades con mayor certeza.

Resultados: se presenta la Teoría de la Decisión Multicriterio Discreta para la toma de decisiones sobre la técnica de Inteligencia Artificial que mejor se ajusta para el diagnóstico de enfermedades con mayor certeza.

Conclusiones: se alcanzó identificar la técnica de Inteligencia Artificial que mejor se ajusta para el diagnóstico de enfermedades con mayor certeza, utilizando la Teoría de la Decisión Multicriterio Discreta, la cual posibilita evaluar síntomas, signos y factores de riesgo presentados en los enfermos.

DeCS: DIAGNÓSTICO; INTELIGENCIA ARTIFICIAL; CERTEZA.

ABSTRACT

Introduction: the process of diseases diagnosis is complex, given that medical data and information can often have uncertainty and require being treated with artificial intelligence (AI) techniques in order to assist patients' diseases with greater assurance supporting the decision making.

Objective: to compare artificial intelligence techniques that are frequently used for the diagnosis of diseases, when the data stored are kept on the behavior of diseases that frequently affect a population in question.

Method: a research supported by the theory of discrete *multicriteria decision support* process, which is useful for the decision making, particularly in relation to the artificial intelligence techniques that best suits to perform the diagnosis of diseases with greater certainty.

Results: *Discrete Multicriteria Decision Making Theory* is presented for decision making in relation to the artificial intelligence techniques that best suit the diagnosis of diseases with greater certainty.

Conclusions: it was possible to identify the artificial intelligence techniques that well-suited the diagnosis of diseases with greater certainty, using Discrete Multicriteria Decision-making Theory, making possible the assessment of symptoms, signs and the risk factors presented in patients.

DeCS: DIAGNOSIS; ARTIFICIAL INTELLIGENCE; CERTAINTY

INTRODUCCIÓN

EL diagnóstico de enfermedades es un proceso cognitivo complejo que implica capacitación, experiencia, reconocimiento de patrones y cálculo de probabilidad condicional, entre otros componentes menos comprendidos. En las últimas décadas se han realizado varios esfuerzos por aplicar el análisis predictivo en los sistemas de salud, así como lanzar sistemas de aprendizaje automático para facilitar el diagnóstico.

En la actualidad la medicina utiliza considerables adelantos que involucran el uso intensivo de alta tecnología como el diagnóstico por imágenes, la robótica, entre otros, especialmente la aplicación de técnicas de Inteligencia Artificial (IA) las cuales son factibles cuando aprovechan los datos disponibles y la experiencia clínica.

Las técnicas de Inteligencia Artificial aplicadas al diagnóstico de enfermedades han sido utilizadas en estudios de problemas complejos, alcanzando un aceptado grado de certeza en los resultados obtenidos con respecto a la identificación de un tipo específico de enfermedad. Estas aplicaciones han resultado ventajosas debido a su facilidad en estudios de sistemas capaces de aprender a partir de diferentes conjuntos de datos y mejorar procesos de clasificación y predicción de enfermedades.

La Inteligencia Artificial es un área del conocimiento donde convergen diferentes técnicas, que tienen la finalidad de llevar a cabo tareas como el diagnóstico médico¹, estudios que han demostrado efectividad son:

- a) Las redes neuronales artificiales. ^(2, 3)
- b) El razonamiento basado en casos. ^(4, 5)
- c) Las redes bayesianas. ⁽⁶⁾

Las redes neuronales artificiales han demostrado que producen buenos resultados en bases de datos con muchas instancias. Las redes neuronales no consideran la incertidumbre, actúan como una caja negra, en el sentido de que no es posible saber cómo se ha llegado a los resultados obtenidos y los nodos intermedios no pueden ser interpretados. ⁽⁷⁾

El razonamiento basado en casos se ha utilizado con éxito en el diagnóstico médico asistido, esta técnica se basa en el paradigma de resolución de problemas en el que un problema nuevo se soluciona atendiendo casos similares del pasado. ⁽⁸⁾El ejemplo más conocido, es el vecino más cercano, que consiste en comparar cada caso problema con una serie de casos almacenados; el caso problema nuevo y los casos almacenados consisten en un número de atributos o variables y éstos se comparan para encontrar el caso almacenado que más se parece al caso problema, de esta manera, se puede seleccionar la enfermedad conocida que más se parezca a la descripción del paciente. Una vez resuelto, este nuevo caso pasa a formar parte de la biblioteca de casos disponibles para comparar, con lo que el sistema se vuelve cada vez más sólido.

Las redes bayesianas calculan la probabilidad de que un caso con cierto atributo pertenezca a una categoría. En medicina se han usado para predecir recaída de enfermedad o riesgo de cáncer. ⁽⁹⁾El filtro no sabe esto por adelantado y tiene que aprender de la mano del usuario. El programa calcula en lo sucesivo la probabilidad de que cada atributo corresponda a una categoría o a otra.

Basado en las características que se describen sobre las técnicas de IA que con frecuencia se utilizan para el diagnóstico de enfermedades, se persigue como objetivo comparar dichas técnicas para seleccionar la que mejor se ajusta al diagnóstico de enfermedades, cuando se tienen datos almacenados sobre el comportamiento de las enfermedades que con frecuencia afectan a una población en cuestión.

MÉTODO

Para el estudio comparado de las técnicas de IA que se utilizan con frecuencia en el diagnóstico de enfermedades, se utilizó la Teoría de Decisión Multicriterio Discreta (DMD), teoría que contribuye a seleccionar la técnica que mejor cumple con los criterios previstos para ejecutar el diagnóstico de enfermedades. La teoría de la DMD es considerada una metodología para el apoyo a la toma de decisiones, ha resultado idónea en varios campos de aplicación, específicamente donde hay que decidir entre varias alternativas, teniendo en cuenta diversos criterios o puntos de vista. ⁽¹⁰⁾

Estudios realizados han demostrado que los métodos de evaluación y decisión multicriterio contribuyen a encontrar soluciones posibles, pero no necesariamente óptimas, por lo que en este proceso juegan un importante papel las preferencias del decisor, una vez definidos los objetivos a evaluar. ⁽¹¹⁾

En base a lo anteriormente expuesto, es de destacar que para un adecuado trabajo con los métodos multicriterio deben tenerse en cuenta aspectos como:

1. Seleccionar la(s) mejor(es) alternativa(s).
2. Aceptar alternativas buenas y rechazar aquellas malas.
3. Generar una ordenación (*ranking*) de las alternativas consideradas (de la "mejor a la peor").

Para cumplir con los aspectos antes mencionados, existen diferentes enfoques, métodos y soluciones, a tener en cuenta, como son:

1. Ponderación Lineal (*Scoring*).
2. Utilidad Multiatributo (MUAT).
3. Relaciones de superación y procesos de Análisis Jerárquico (AHP – *The Analytic Process* – Proceso Analítico Jerárquico).

Para validar los resultados se utilizaron pruebas de aceptación con el objetivo de medir satisfacción a partir de los criterios evaluados por los usuarios. Las pruebas consistieron en realizar un cuestionario a 20 especialistas de salud pertenecientes al Centro Nacional de Genética Médica con más de 10 años de experiencia. Se utilizó la escala de **Osgood**. Los especialistas seleccionados aportaron una valoración sobre los resultados de la investigación realizada.

Las escalas utilizadas son instrumentos de recolección de información con base en una lista de ítems, reactivos, o frases cuidadosamente seleccionados, de forma que constituyen un criterio sistemático, confiable, válido y específico para medir cuantitativamente alguna forma de fenómeno. Se evalúa, en este caso, las opiniones de los especialistas de salud sobre la validez de los resultados de la técnica de IA que mejor se ajusta al diagnóstico de enfermedades.

Para el estudio comparado de las técnicas de IA que mejor se ajustan para el diagnóstico de enfermedades, se selecciona el método de ponderación lineal (*Scoring*), donde el decisor juega un papel fundamental de acuerdo a los objetivos definidos por cada técnica de IA, el mencionado método se utilizó con el fin de definir una estructura de preferencia entre las alternativas identificadas, tal como ocurre en la práctica y en particular en el diagnóstico de enfermedades. El esquema básico que sigue el método de ponderación lineal consiste en construir una función de valor S_j para cada alternativa, de acuerdo a como se plantea en la expresión 1. Resultados que se muestran en la tabla 1.

$$S_j = \sum_i w_i r_{ij}$$

Dónde:

w_i , es el peso del criterio j y r_{ij} la evaluación (*rating*) de la alternativa i respecto al criterio j .

En este caso se aplica para validar el resultado de una investigación ya concluida y que ha sido observada directamente por los evaluadores, con la escala que se explica a continuación:

Muy satisfactorio 5 4 3 2 1 Insatisfactorio

A los especialistas de salud se le pidió que evaluaran de la siguiente forma:

- 5 Muy satisfactorio
- 4 Satisfactorio
- 3 Neutro
- 2 Poco satisfactorio
- 1 Insatisfactorio

Luego de obtener los resultados a través de la escala de **Osgood** es necesario el cálculo del diferencial semántico con el objetivo de evaluar la aceptación de la técnica de IA que mayor se ajusta al diagnóstico de enfermedades, para ello se escogieron otros pares de adjetivos opuestos y 5 posibilidades de seleccionar teniendo en cuenta el siguiente ítem.

Ítem a evaluar: Las redes bayesianas como técnica de IA, para el diagnóstico de enfermedades con un mayor de certeza.

Espacio semántico: 1, 2, 3, 4, 5

Buena ----- Mala	➡	Los 20 especialistas en salud marcaron la primera opción.
Adecuada ----- Inadecuada	➡	Los 20 especialistas en salud marcaron la primera opción.
Aplicable ----- Inaplicable	➡	Los 20 especialistas en salud marcaron la primera opción.
Innovadora ----- Obsoleta	➡	Los 20 especialistas en salud marcaron la primera opción.
Completa ----- Incompleta	➡	19 la primera opción, un especialista en salud animal la segunda.

RESULTADOS

A modo de guía para la selección de la técnica más adecuada para ejecutar diagnóstico de enfermedades se utiliza la información extraída del análisis documental que se realiza en la presente investigación, la cual está en correspondencia con el tema que se presenta. En la tabla 1 se muestran los criterios y ponderaciones de las técnicas de IA que con frecuencia se han utilizado para el diagnóstico de enfermedades.

Tabla 1. Criterios y ponderación de las técnicas de Inteligencia Artificial a utilizar para el diagnóstico de enfermedades.

Criterios	Peso por criterio (W_i)
1. Utilización de mecanismo de aprendizaje	5
2. Trabajo con datos que presentan incertidumbre	1
3. Posibilidad de combinar conocimiento y datos	3
4. Obtención de conocimiento verificable	5
5. Obtención de resultados de forma rápida y precisa	5
6. Grado de organización de la información	4
7. Utilización de analogías	3
8. Información excesiva o redundante	1
9. Flexibilidad	4

Fuente: Elaboración propia.

El resultado de la aplicación de la teoría de DMD aplicada a las técnicas de IA más apropiada para el diagnóstico de enfermedades. (tabla 2)

Tabla 2. Resultados de la Ponderación Lineal de cada criterio para la selección de la técnica de IA para el diagnóstico de enfermedades.

Criterios	Ponderación W_i	Razonamiento basado en casos r_{i_1}	Redes neuronales artificiales r_{i_2}	Redes bayesianas r_{i_3}
Utilización de mecanismo de aprendizaje	5	7	9	9
Trabajo con los datos que presentan incertidumbre	1	3	7	9
Posibilidad de la combinación del conocimiento y los datos	3	8	6	9
Obtención de conocimiento verificable	5	8	1	9
Obtención de resultados de forma rápida y precisa	5	9	5	9
Grado de organización de la información	4	8	4	8
Utilización de analogías	3	9	1	2
Información excesiva o redundante	1	2	4	3
Flexibilidad	4	8	5	9
Scoring S_j		240	143	248

Fuente: Elaboración propia.

Basado en el análisis realizado, la técnica de IA que reflejó resultados más apropiados, para el diagnóstico de enfermedades en la ganadería se correspondió con el uso de las redes bayesianas. Resultado que está en correspondencia con el estado del arte relacionado con el uso de las técnicas de Inteligencia Artificial que con frecuencia se utilizan para el diagnóstico de enfermedades.

Con la finalidad de determinar una valoración sobre los resultados de la investigación realizada se realizó el análisis de las respuestas mediante los ítems a evaluar a través de la escala de **Osgood**(Tabla 3).

Tabla 3. Resumen de las valoraciones de escala realizada por especialistas de salud sobre los resultados de la investigación realizada.

Ítem a evaluar	1	2	3	4	5
La manera en que se tratan los diagnósticos de enfermedades en general y en particular, utilizando técnicas de IA.				6	14
Lo beneficioso de utilizar las TIC en el desarrollo de los diagnósticos de enfermedades.			2	6	12
Importancia de las acciones planificadas para utilizar técnicas de IA para el diagnóstico de enfermedades.			1	5	14
La concepción general de las técnicas de IA para el diagnóstico de enfermedades.				5	15
Impacto de las técnicas de IA en el diagnóstico de enfermedades, con un mayor grado de certeza.				2	18
Opinión sobre las indicaciones metodológicas para técnicas de IA en el diagnóstico de enfermedades.				6	14
Resultados visibles con la aplicación de las técnicas de IA en el diagnóstico de enfermedades en áreas de salud.				3	17
Suma total de la valoración			3	33	104

Fuente: Elaboración propia.

Los datos representados en la tabla 3 revelan una buena evaluación de los resultados de la investigación, lo cual se representa en la figura 1.

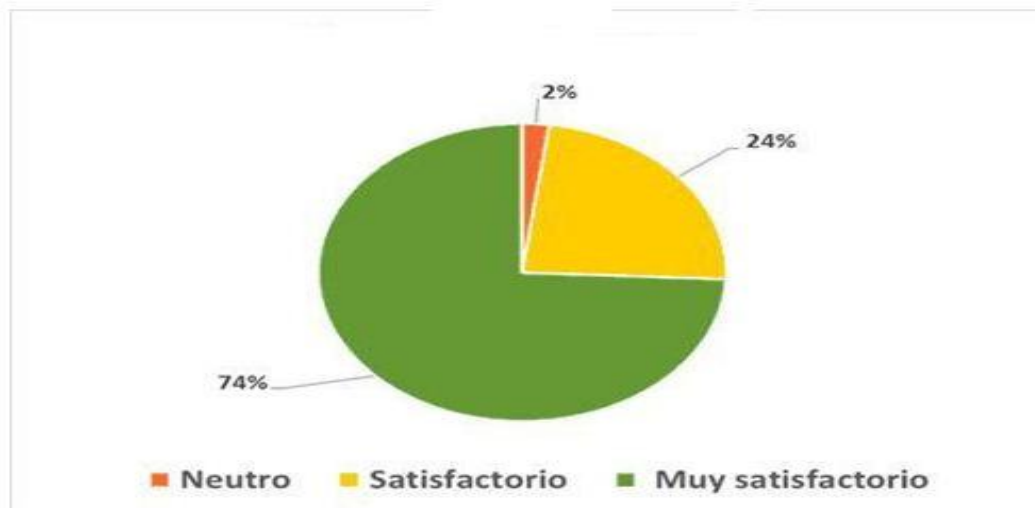


Fig. 1 Respuestas en la escala de *Osgood*.
Fuente: Elaboración propia.

Como se aprecia, los valores obtenidos son una evidencia de la evaluación favorable de las técnicas de IA para el diagnóstico de enfermedades y en particular las redes bayesianas.

DISCUSIÓN

Las redes bayesianas poseen ventajas frente a las redes neuronales, el razonamiento basado en casos que las distinguen para ejecutar diagnóstico de enfermedades, entre estas ventajas se encuentran:

Tienen una interpretación sencilla, apoyada en métodos probabilísticos

Son robustas, en el sentido que pequeñas alteraciones en el modelo no afecta la estructura de la red.

Se pueden incorporar los conocimientos o creencias previas de un experto a través de la distribución a priori.

Están especialmente diseñadas para descubrir las relaciones entre variables que son causales y no poder ser observadas de otro modo, permitiendo hacer predicciones más precisas.

No se ven afectadas por el sobreajuste de los datos.

Asimismo, poseen como desventaja; la escasa relación que se da en algunas situaciones entre el razonamiento lógico humano y los métodos probabilísticos ¹². Desventaja que en la práctica ha sido disminuida debido a que el trabajo con redes bayesianas para el diagnóstico de enfermedades se realiza comúnmente desde datos almacenados en bases de datos y desde el conocimiento de los expertos, lo que contribuye a obtener relaciones de razonamiento lógico humano útiles para el trabajo con métodos probabilísticos.

La utilización de las redes bayesianas facilitaría el diagnóstico de enfermedades desde los datos debido a la capacidad que poseen para modelar los procesos del razonamiento

médico combinando el conocimiento de los expertos con los datos médicos que ellos aportan y los datos clínicos existentes.

Se concluye que la aplicación de la escala de **Osgood** proporciona una valoración sobre los resultados de la investigación realizada, evidenciando la posibilidad de utilizar técnicas de IA para el diagnóstico de enfermedades desde los datos y en particular haciendo uso de las redes bayesianas.

REFERENCIAS BIBLIOGRÁFICAS

1. Lugo S, Maldonado G, Murata Ch. *Inteligencia artificial para asistir el diagnóstico clínico en medicina*. Revista Alergia México 2014 [Internet];61:110-120. Available from: <http://revistaalergia.mx/ojs/index.php/ram/article/download/33/46>.
2. Bernal EA. Sistema prototipo de entrenamiento pediatra para el proceso de adaptación neonatal. Tesis de Maestría en Ingeniería de Sistemas y Computación, Línea de investigación: Ingeniería del Software, Sistemas Inteligentes, ÁREA: Sistemas Inteligentes 2014 [Internet]. Disponible en: <http://slideplayer.es/slide/94359/>.
3. Atkov OY, Gorokhova SG, Sboev AG, Generozov EV, Muraseyeva EV. Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. J Cardiol 2012 [Internet]; 59(2):190- 194. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22218324>.
4. Monsalve Torra AE. Sistema de ayuda a la decisión clínica en enfermedades de diagnóstico complejo. Tesis en opción al grado científico de Doctor en Ciencias Informáticas. Departamento de Tecnología, Informática y Computación. Escuela Politécnica Superior de Alicante, España 2017 [Internet]. Available from: https://rua.ua.es/dspace/bitstream/10045/65334/1/tesis_monsalve_torra.pdf.
5. Ochoa CA, Mena LJ, Felix VG, García V. Diagnóstico para Síndrome de Asperger utilizando una aplicación inteligente a partir del razonamiento basado en casos; 2015. Revista Electrónica de Divulgación de la Investigación 2015 [Internet] Vol. 10. Available from: https://www.researchgate.net/publication/291698853_Diagnostico_para_Sindrome_de_Aasperger_utilizando_una_aplicacion_inteligente_a_partir_del_razonamiento_basado_en_casos.
6. Fernández Galán S. Redes Bayesianas temporales: aplicaciones médicas e industriales. Tesis doctoral. Universidad Nacional de Educación a Distancia [Internet]. Madrid 2002. Disponible en: <http://www.cisiad.uned.es/tesis/tesis-seve.pdf>.
7. Seixas JM, Faria J, Souza JBO, Vieira AFM, Kritski A et al. Artificial neural network models to support the diagnosis of pleural tuberculosis in adult patients. Int J Tuberc Lung Dis 2013[Internet] [cited 2013 Nov 18];17(5):682-686. Available from: <http://www.ingentaconnect.com/content/iuatld/ijtlid/2013/00000017/00000005/art00021?token=00581a1cffd6a264d37e41225f40384d576b4628486b253e2c49576b3427656c3c6a333f2566e4ed81d7b599>

8. Febles O. Modelo para el desarrollo de aplicaciones compuestas basadas en arquitecturas orientadas a servicios, Tesis en opción al grado científico de Doctor en Ciencias Técnicas, Especialidad – Informática. Universidad de las Ciencias Informáticas [Internet]. Ciudad de la Habana 2012. Disponible en: <http://tesis.cujae.edu.cu/bitstream/handle/123456789/891/Tesis283.pdf?sequence=1&isAllowed=y>.
9. Laws RJ, Kesler DC. A Bayesian network approach for selecting translocation sites for endangered island birds, 2012. Biological Conservation 2012 [Internet]; 155: 178-185. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0006320712002650?via%3Dihub>.
10. Barba-Romero S. Evaluación multicriterio de proyectos, en Martínez, 1994. E. (Ed.), Ciencia, Tecnología y Desarrollo: Interrelaciones Teóricas y Metodológicas, (Ed. Nueva Sociedad, Caracas, pp. 455-507).
11. Martínez E, Escudey M. Evaluación y Decisión Multicriterio. Reflexiones y Experiencias, 1998. Editorial Universidad de Santiago/UNESCO, Santiago de Chile.
12. Leyva M. Modelo de ayuda a la toma de decisiones basado en Mapas Cognitivos Difusos. Tesis presentada en opción al Grado Científico de Doctor en Ciencias Técnicas, 2013. Available from: https://www.researchgate.net/profile/Maikel_Leyva-Vazquez/publication/263221297_MODELO_DE_AYUDA_A_LA_TOMA_DE_DECISIONES_BASADO_EN_MAPAS_COGNITIVOS_DIFUSOS/links/0046353a2e6a8a1e8d000000/MODELO-DE-AYUDA-A-LA-TOMA-DE-DECISIONES-BASADO-EN-MAPAS-COGNITIVOS-DIFUSOS.pdf.

Neilyz González Benítez: Licenciada en Informática. Especialidad Informática y en Meteorología. Doctora en Ciencias Técnicas. Centro Meteorológico Provincial. Pinar del Río. Cuba. ***Si usted desea contactar con el autor de la investigación hágalo [aquí](#)***