

Inteligencia Artificial en Sistemas de recomendación Médica: Una revisión de la Literatura.

Artificial Intelligence in Medical Recommendation Systems: A Literature Review.

Javier A. Calderón-Blas¹
javierblas01721@gmail.com
María Angélica Cerdán^{1,2}
acerdan@uv.mx
Ángel J. Sánchez-García¹
angesanchez@uv.mx
Saúl Domínguez-Isidro¹
sauldominguez@uv.mx

¹Facultad de Estadística e Informática, Universidad Veracruzana, Xalapa, México

²Instituto Tecnológico Superior de Xalapa, Xalapa, México

Resumen: En México, aproximadamente una de cada tres personas busca información sobre atención médica a través de Internet, considerando factores como la especialidad, horarios, ubicación, costo, eficacia del tratamiento y calidad del trato médico. Para abordar estas búsquedas, se han desarrollado sistemas de recomendación que ofrecen sugerencias personalizadas basadas en información explícita e implícita del usuario. Estos sistemas utilizan diversas fuentes de datos, incluyendo el Internet de las cosas, para equilibrar la precisión, valor agregado, dispersión y estabilidad en las recomendaciones. A pesar de la amplia aplicación en eSalud, la investigación actual carece de un enfoque específico en los sistemas de recomendación de médicos. Este estudio tiene como objetivo presentar los trabajos más recientes en esta área, analizando enfoques, métodos, técnicas y tendencias a través de una revisión sistemática de la literatura siguiendo el proceso de la ingeniería de software basada en evidencia (EBSE). Los resultados muestran una mayor cantidad de técnicas de recomendación basadas en aprendizaje automático, lógica heurística y combinación de ambos enfoques (híbridos). Finalmente, se identifican áreas de investigación abiertas y tendencias en este campo, que serán de interés para futuras aplicaciones en el ámbito médico.

Palabras Clave: eSalud, sistemas de recomendación, Inteligencia Artificial, Software, revisión sistemática de la literatura.

Abstract: In Mexico, approximately one in three people search for information about medical care through the Internet, considering factors such as specialty, hours, location, cost, treatment effectiveness, and quality of medical treatment. In order to address these searches, recommender systems have been developed that offer personalized suggestions based on explicit and implicit user information. These systems use various data sources, including the Internet of Things, to balance accuracy, added value, dispersion, and stability in recommendations. Despite the wide application in eHealth, current research needs a specific focus on physician recommendation systems. This study aims to present the most recent works in this area, analyzing approaches, methods, techniques, and trends through a systematic literature review following the evidence-based software engineering (EBSE) process. The results show more recommendation techniques based on machine learning, heuristic logic, and a combination of both approaches (hybrids). Finally, open research areas and trends in this field are identified, which will be of interest for future applications in the medical field.

Keywords: eHealth, recommender system, Artificial Intelligence, Software, systematic literature review.

1. Introducción

Las Tecnologías de la Información y Comunicaciones, han tenido un gran impacto en todas las áreas del quehacer humano. Su uso para la atención de las necesidades en temas de salud se ha definido como eSalud (OMS, s.f.). Uno de los temas clave dentro de esta área, es el manejo de datos al servicio del consumidor de servicios sanitarios (Priego Álvarez, 2005).

La adquisición de servicios de salud como cualquier proceso de consumo, se ha visto influenciado por el uso masivo de Internet, incluyendo: la evaluación de alternativas, la decisión y el comportamiento post-consumo. En este sentido, más del 70% de las personas, recurren a la web en búsqueda de información sobre la salud y un aproximado del 30%, toma decisiones con base a la información encontrada (Bravo, 2021).

Esta tendencia se reforzó derivado de la pandemia de COVID-19, ya que, ante el confinamiento, las personas vieron en los medios digitales una alternativa para la búsqueda de medicamentos, insumos, servicios de asistencia y servicios médicos (Iglesias Ososres y Saavedra Camacho, 2020). Según (Villafranco, 2019), en México, una de cada tres personas, buscan información sobre atención médica, por Internet o por teléfono.

En la búsqueda de atención médica, no sólo se consideran factores como la especialidad, horarios de atención, ubicación y costo, sino que, a los pacientes o sus cuidadores, también les interesan otros aspectos, como los relativos a la eficacia de su atención y el trato personal que brinda el profesional, lo que ha derivado en la creación de los sitios web para calificación de médicos (Deng et al., 2019).

Para resolver este tipo de búsquedas, se han desarrollado los sistemas de recomendación, mecanismos inteligentes que brindan al usuario sugerencias personalizadas, a través del análisis de la información que se proporciona de manera explícita y de la derivada de la interacción que el usuario tiene con la plataforma de manera implícita, con la finalidad de ofrecer una mejor sugerencia (Ramos, 2019).

Estos sistemas utilizan diferentes fuentes de información, usando datos demográficos, contenido propio de la oferta, obtenido de la colaboración, de las redes sociales y más recientemente, de la adquisición de datos derivada del Internet de las cosas, para realizar las predicción y recomendación, equilibrando factores de precisión, valor agregado, dispersión y estabilidad en las recomendaciones, buscando cumplir las expectativas y requisitos de los usuarios. Generalmente, la mayor parte de los recomendadores aplican enfoques híbridos (Bobadilla et al., 2013).

Dada la amplia aplicación de los sistemas de recomendación en la eSalud en procesos de diagnóstico, pronóstico y tratamiento médico, se han desarrollado análisis de los temas más relevantes en este campo (Calero et al., 2016), sin embargo, no se particulariza en el proceso de recomendación de médicos o se incluye de manera parcial (Trang et al., 2020), sin profundizar en las más recientes aportaciones en particular. Por lo que el presente estudio, tiene por objetivo informar sobre los trabajos más recientes sobre los sistemas de recomendación de médicos, para analizar los enfoques con los que son abordados, identificando los métodos, técnicas, herramientas utilizadas y la evaluación de resultados, para identificar temas abiertos y tendencias, que pueden ser de útiles para los interesados en nuevas o futuras aplicaciones.

Este artículo está estructurado de la siguiente manera. La Sección 2 presenta trabajos relacionados a esta investigación. La Sección 3 describe el método de búsqueda para la selección de estudios primarios. La Sección 4 muestra los resultados y análisis de cada pregunta de investigación. La Sección 5 describe algunas amenazas a la validez de este trabajo. Finalmente, la sección 6 extrae las conclusiones del trabajo y presenta el trabajo futuro.

2. Trabajo Relacionado

Se realizaron búsquedas manuales en las que se identificaron artículos que proponen sistemas de recomendación en diversos campos, incluyendo recomendación laboral, elementos de comercio electrónico, rutas de aprendizaje, entre otros. De los estudios centrados en el análisis de sistemas de recomendación encontrados, únicamente dos son revisiones sistemáticas de la literatura (RSL).

En el estudio de Katarya (2017), se llevó a cabo una revisión sistemática de sistemas de recomendación de grupos (GRS por sus siglas en inglés), que ofrecen recomendaciones personalizadas a grupos de usuarios, considerando sus intereses, pasatiempos, entretenimiento, experiencias pasadas y planificación futura. Los autores clasificaron las técnicas de GRS en once categorías de filtrado, incluyendo enfoques basados en contenido, colaborativos, basados en contexto, semántica, comunitarios, etiquetado, análisis de sentimiento, híbridos, demografía, modelos de grupo y filtrado de grupos.

Por otro lado, en el artículo de Patel et al. (2017), se realizó una clasificación de diversas técnicas de recomendación en tres categorías: filtrado basado en contenido, filtrado colaborativo y filtrado híbrido. Se analizaron varios algoritmos en cada una de estas categorías, concluyendo que las técnicas de filtrado colaborativo presentan ventajas significativas en comparación con las otras técnicas. Aunque los enfoques híbridos buscan mejorar la calidad del sistema de recomendación, su implementación se considera compleja y costosa.

En el ámbito médico, se encontró el trabajo de Pincay et al. (2019), que realizó una revisión sobre los Sistemas de Recomendación de Salud (HRS por sus siglas en inglés). Estos sistemas están diseñados para respaldar a los profesionales médicos en la detección de enfermedades y brindar orientación a los pacientes en la promoción de su salud y el mantenimiento de su bienestar. El análisis incluyó artículos publicados entre 2006 y 2018 en cinco bases de datos científicas diferentes. Una conclusión interesante del estudio es que un "porcentaje significativo de las iniciativas empleó técnicas de inteligencia artificial y lógica difusa para lograr una mayor precisión y rendimiento de los sistemas". Por tanto, nuestro trabajo contribuirá a la clasificación de las técnicas de inteligencia artificial utilizadas en este tipo de sistemas.

En Thongchotchat et al. (2021), se llevó a cabo una RSL centrada en el estilo de aprendizaje, que es la preferencia de aprendizaje de un estudiante. El objetivo era identificar técnicas para desarrollar sistemas de aprendizaje con soporte informático que puedan crear un camino de aprendizaje personalizado para cada alumno. Este estudio involucró fuentes de datos como IEEE Xplore y ScienceDirect y encontró que la teoría de Felder & Silverman fue la más utilizada, con 72.5%, y que la aplicación adecuada es el algoritmo de recomendación más utilizado, con 42.5% del total.

Finalmente, en 2022, se publicó la revisión de Gupta y Bindal (2022), que evaluó los sistemas de recomendación multimodales. Los autores concluyeron que se han desarrollado diversas estrategias de recomendación para satisfacer las necesidades en diversas circunstancias, incluyendo sistemas basados en contenido, colaborativos, demográficos y enfoques híbridos. Este artículo también exploró los diferentes desafíos que pueden surgir en estos sistemas, como el arranque en frío, la escasez, la escalabilidad y la diversidad."

Sin embargo, no se encontró ninguna revisión sistemática en el área médica o afines. Este hecho es de gran relevancia, dado que existen diferencias entre recomendar productos y recomendar a un médico, ya que la veracidad de la sugerencia de un buen médico tiene un impacto directo en un diagnóstico preciso y, en consecuencia, en un tratamiento efectivo. Por lo tanto, el objetivo de esta revisión es identificar los componentes clave relacionados con la recomendación en este ámbito, centrándose específicamente en las técnicas de Inteligencia Artificial utilizadas para realizar sugerencias de calidad y cómo se evalúan estos sistemas.

3. Método de Investigación

El método utilizado para esta investigación es el propuesto por Kitchenham, et al. (2015), el cual se divide en las etapas de planeación y la conducción. A continuación, se describen los elementos para llevar a cabo esta revisión.

3.1 Etapa de Planeación

En esta etapa, se plantean las preguntas de investigación que guían esta investigación. En la Tabla 1 se presentan las preguntas de investigación junto con su motivación asociada.

Pregunta de investigación	Motivación
PI1. ¿Qué estrategias de Inteligencia Artificial utilizan los sistemas de recomendación médica?	Presentar un panorama general de las técnicas de Inteligencia Artificial que se han utilizado para hacer una recomendación precisa, y que pueden ser utilizadas o modificadas por los desarrolladores de software.
PI2. ¿Cómo se evalúan los sistemas de recomendación médica?	Conocer la manera en que evalúan los sistemas de recomendación, especialmente en el área médica, con el fin de generar propuestas que puedan compararse con los sistemas encontrados en la literatura.
PI3. ¿Cuáles fueron los resultados de la evaluación de los sistemas de recomendación médica?	Conocer los resultados de las evaluaciones de este tipo de sistemas permitirá generar propuestas que puedan mejorar los resultados encontrados en la literatura.
PI4. ¿Qué trabajo futuro se identificó en los estudios seleccionados?	Identificar las propuestas de mejora a los sistemas encontrados en los estudios primarios, con el fin de mejorar los resultados de la literatura.

Tabla 1. Preguntas de Investigación

De las fuentes sugeridas por las guías Kitchenham et al. (2015) y Petersen et al. (2015) se escogieron IEEEExplore, ACM Digital Library, Science Direct y Springer Link, debido a que en estas fuentes se publican artículos sometidos a rigurosos procesos de revisión por pares, lo que garantiza la calidad y validez de los resultados presentados. Además, considerando que el dominio de aplicación es relativa al área médica, fuentes como Science Direct y Springer Link brindan artículos de áreas multidisciplinarias. Las fuentes seleccionadas se pueden visualizar en la Tabla 2.

Fuente	Sitio web
IEEEExplorer	https://ieeexplore.ieee.org/Xplore/home.jsp
ACM Digital Library	https://dl.acm.org/
Science Direct	https://www.sciencedirect.com/
Springer Link	https://link.springer.com/

Tabla 2. Fuentes de información seleccionadas

Para la generación de la cadena de búsqueda, se identificaron los términos claves para la búsqueda de artículos. Los términos definidos para esta RSL se muestran en la Tabla 3.

Término clave	Término en inglés	Término relacionado
Sistema	System	-
Doctor	Doctor	-
Recomendador médico	Medical recommender	Medical assistance Medical services Healthcare
Paciente	Patient	-

Tabla 3. Términos clave de búsqueda

La cadena de búsqueda representa la entrada para recuperar estudios relevantes de las fuentes de datos mostradas en la Tabla 2. Se generaron diversas cadenas de búsqueda a partir de las palabras de la Tabla 3 enlazados con operadores booleanos (OR y AND), y se seleccionó aquella que obtuvo un mejor desempeño, en términos de *precision* y *recall*.

La métrica *precision* es la proporción de documentos recuperados que son relevantes para la consulta, véase la ecuación (1); por su parte, la métrica *recall* es la proporción de documentos relevantes que se recuperan con éxito, véase la ecuación (2).

$$Precision = \frac{\text{Documentos relevantes recuperados}}{\text{Total de documentos recuperados}} \quad (1)$$

$$Recall = \frac{\text{Documentos relevantes recuperados}}{\text{Total de documentos relevantes}} \quad (2)$$

Se seleccionó una cadena de búsqueda diferente para cada fuente de datos. La Tabla 4 muestra la cadena de búsqueda aplicada para cada fuente, con su respectiva medida de *recall* y *precision*. Como se puede observar en la Tabla 4, en IEEEExplore y Science Direct se alcanzó un *recall* de 0.8 (sugerido por Kitchenham et al. (2015)), mientras que para Springer y ACM Digital Library los valores más altos obtenidos fueron inferiores a 0.6, pero se tomó la cadena de búsqueda con mayor valor de *recall*, y como segundo criterio el valor de precisión más alto.

Fuente	Cadena de búsqueda	Valor de <i>recall</i>	Valor de <i>precision</i>
IEEEExplore	"Recommender system" AND (Doctor OR Patient OR Medical Assistance)	0.83	0.26
ACM Digital Library	(Doctor OR Medical Assistance) AND "Recommender system"	0.59	0.14
Science Direct	"Recommendation system" AND (Doctor OR "Medical services" OR "Medical Assistance" OR Healthcare)	0.8	0.22
Springer Link	"Recommender system" AND (Doctor OR Medical services OR Medical Assistance)	0.56	0.1

Tabla 4. Cadenas de búsqueda aplicadas por fuente de datos

3.2 Etapa de Conducción

En este apartado se describen los criterios de selección de los artículos encontrados con las cadenas de búsqueda y se explica el proceso de selección de los estudios primarios que serán analizados.

3.2.1 Criterios de selección

Se propusieron cinco criterios de inclusión (CI) y dos criterios de exclusión CE, para seleccionar los estudios primarios de los estudios recuperados en las fuentes seleccionadas, véase Tablas 5 y 6, respectivamente.

Clave	Descripción
CI1	Artículos publicados entre 2016 y 2023.
CI2	Publicaciones escritas en idioma inglés.
CI3	Artículos con acceso completo.
CI4	Al leer el título y el resumen, el artículo da indicios de estar relacionado con los sistemas de recomendación médica.
CI5	Al leer el contenido de la publicación, da indicios de responder al menos una pregunta de investigación.

Tabla 5. Criterios de inclusión

Clave	Descripción
CE1	El artículo es un resumen, póster o capítulo de libro.
CE2	Investigaciones duplicadas.

Tabla 6. Criterios de exclusión

3.2.2 Proceso de selección

Se definieron cuatro etapas para el proceso de selección de estudios. La Figura 1 muestra los criterios aplicados en cada etapa.

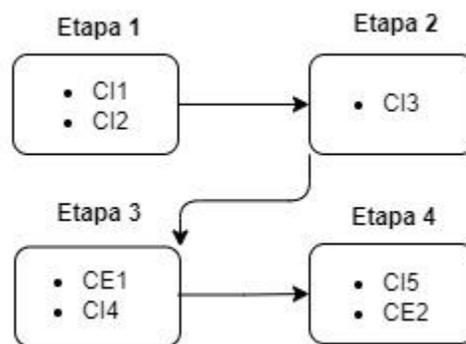


Figura 1. Proceso de selección de estudios primarios.

La Tabla 7 muestra el número de estudios primarios seleccionados por etapa, donde se puede observar que IEEEExplore seleccionó la mayor cantidad de estudios. Asimismo, la etapa 2 fue en la que se descartó la mayor cantidad de artículos. Finalmente, no se seleccionó ningún estudio de la biblioteca digital ACM Digital Library ni de Springer Link.

Fuente	Resultados	Etapa 1	Etapa 2	Etapa 3	Etapa 4
IEEEExplore	90	61	61	9	7
ACM Digital Library	2212	43	3	0	0
Science Direct	1558	882	32	2	2
Springer Link	1957	218	81	4	0
Total	5817	1204	177	15	9

Tabla 7. Número de estudios primarios seleccionados por etapa.

3.2.3 Proceso de Snowballing

Debido al escaso número de estudios seleccionados, se decidió aplicar una iteración del proceso de bola de nieve (snowballing) hacia adelante y una iteración hacia atrás como se muestra en la Figura 2.

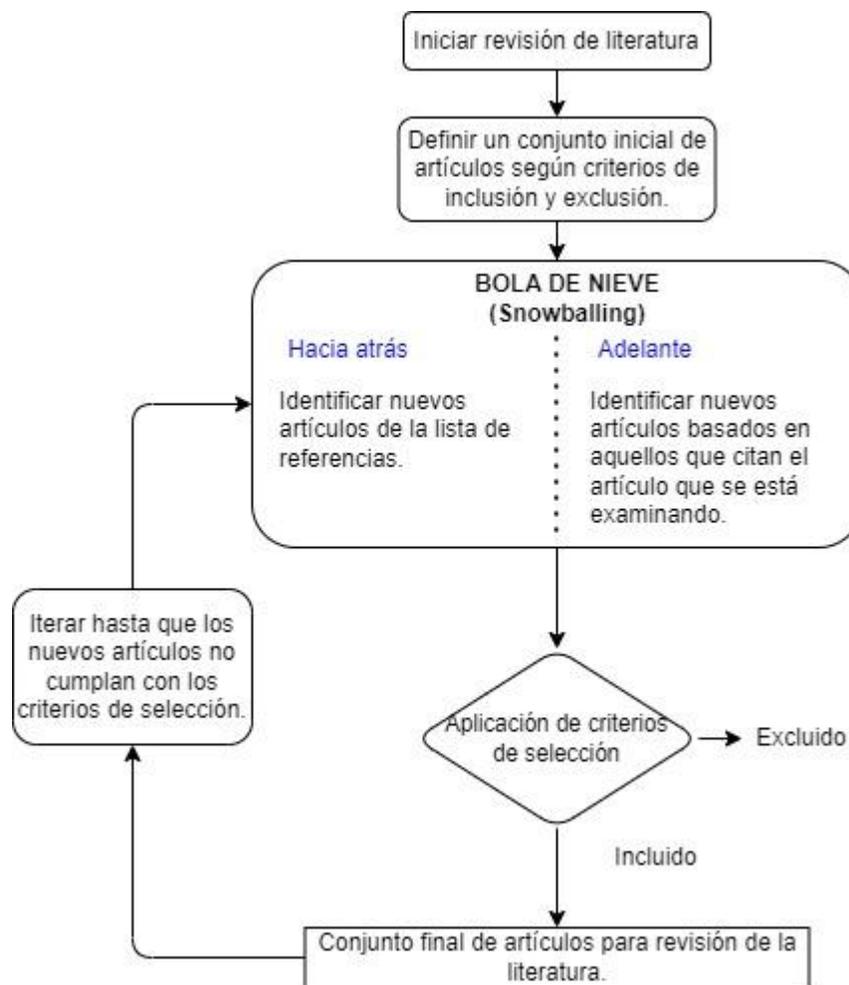


Figura 2. Proceso de bola de nieve (Snowballing).

Después de aplicar el proceso de bola de nieve, se agregaron tres estudios primarios más de dos fuentes diferentes que cumplieron con los criterios de selección, como se muestra en la Tabla 8.

Fuente	Sitio web	Número de estudios seleccionados
IOPscience	https://iopscience.iop.org/	2
JMIR Publications	https://www.jmir.org/	1

Tabla 8. Fuentes de datos agregadas después del proceso de snowballing.

4. Resultados

La distribución final de los estudios seleccionados por fuente se muestra en la Figura 3, donde IEEEExplore, con siete publicaciones de artículos, es la fuente seleccionada con la mayor cantidad.

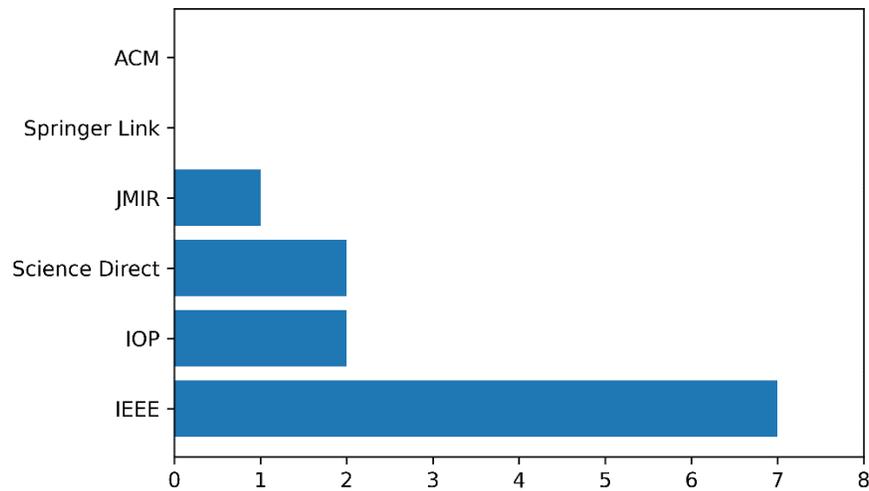


Figura 3. Número final de estudios primarios seleccionados por fuente.

La proporción de los estudios seleccionados por tipo de publicación se muestra en la Figura 4, siendo el mismo número de artículos de revistas y congresos. Por otra parte, en la Figura 5 se muestra la distribución de los estudios por año entre 2016 y 2023, se puede observar que 2018 y 2020 fueron los años con más artículos publicados en relación con los sistemas de recomendación médica. Los estudios primarios seleccionados se muestran en la Tabla 9 agrupados por fuente de información.

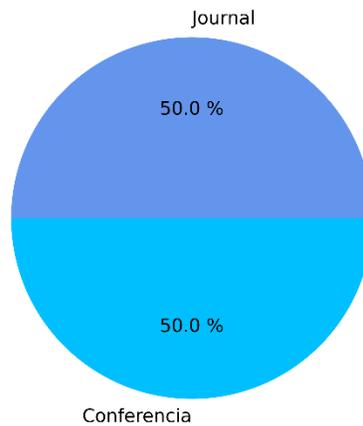


Figura 4. Proporción de estudios por tipo de publicación.

Fuente	Referencia
IEEEExplore	(Han et al.,2018), (Han et al.,2018b), (Bjansali & Nagwani, 2021) (Venkatesh et al., 2022), (Singh et al., 2016), (Swarnalatha et al., 2019), (Meng et al., 2019)
IOP	(Mawardi et al., 2020), (Kannan et al., 2020)
Science Direct	(Mondal et al., 2020), (Zhang et al., 2017)
JMIR	(Guo et al., 2016)

Tabla 9. Estudios primarios seleccionados

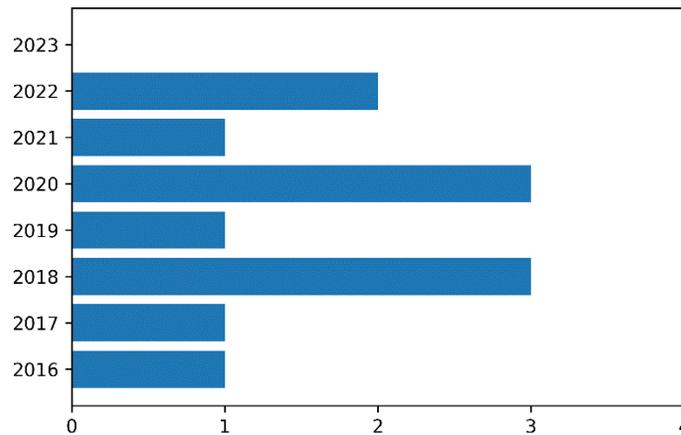


Figura 5. Distribución de año de publicación en el periodo 2016-2023.

Para el proceso de síntesis de la información se optó por la estrategia de síntesis narrativa. Este tipo de síntesis descrita por Popay et al. (2006) implica integrar los hallazgos de múltiples estudios en una narrativa coherente que resuma la evidencia disponible sobre nuestro tema de interés. A continuación, se presentan las respuestas a cada pregunta de investigación.

4.1 PI1. ¿Qué estrategias de Inteligencia Artificial utilizan los sistemas de recomendación médica?

Un total de 19 técnicas o algoritmos fueron identificados en los 12 estudios primarios. La Figura 6 presenta las diferentes técnicas encontradas para la mejora de la precisión en las recomendaciones de los sistemas de recomendación médica (SRM), siendo el Filtrado colaborativo la más recurrente, seguido de los algoritmos de clasificación C4.5 y Naive Bayes y un modelo híbrido. Cabe mencionar que algunos estudios primarios, al proponer experimentos para una comparación de entre técnicas para recomendación, abordaron más de una técnica.

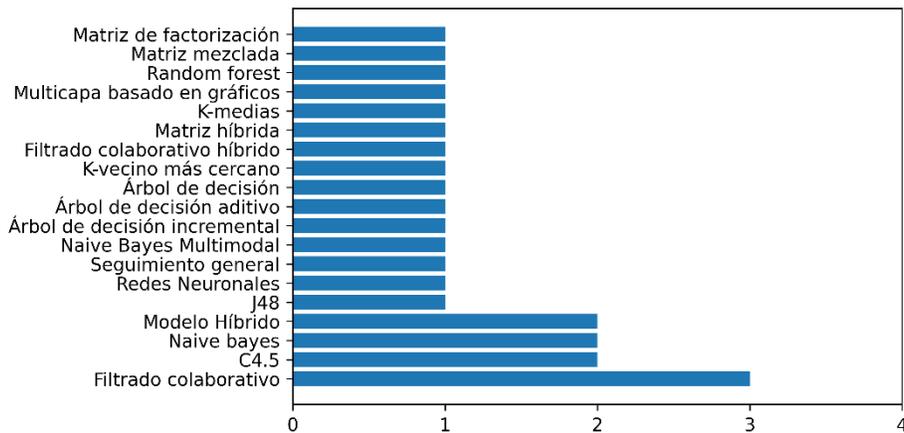


Figura 6. Distribución de técnicas encontradas para la recomendación médica.

Se definieron y agruparon las técnicas reportadas para hacer recomendaciones precisas mediante sistemas de recomendación médica en seis estrategias o enfoques. La Tabla 10 muestra la agrupación de las técnicas de la Figura 6 en seis enfoques.

Enfoque	Técnicas
Aprendizaje máquina	Aprendizaje supervisado: <ul style="list-style-type: none"> • Árboles de decisión (C4.5, incremental, aditivo, J48) • Modelos Bayesianos (Naive Bayes) • K-Vecino más cercano (K-Nearest Neighbor) • Clasificadores ensamblados (Random Forest) • Redes Neuronales
Lógica Heurística	Aprendizaje No supervisado <ul style="list-style-type: none"> • K-medias (K-means)
Híbrido	<ul style="list-style-type: none"> • Filtrado colaborativo • Filtrado colaborativo híbrido • Filtrado colaborativo basado en reglas
Matricial	<ul style="list-style-type: none"> • Basado en contenido y filtrado colaborativo • Filtrado colaborativo e ICD-9 • Filtrado colaborativo, K-Vecino más cercano y árboles de decisión
Basado en reglas	<ul style="list-style-type: none"> • Matriz de factorización (Factorization matrix) • Matriz mezclada (Mixed matrix) • Matriz Híbrida (Hybrid matrix)
Procesamiento de Lenguaje Natural	<ul style="list-style-type: none"> • Basado en reglas • Revisiones de usuarios

Tabla 10. Técnicas agrupadas por enfoque.

El aprendizaje automático, una rama de la Inteligencia Artificial, permite a las máquinas adquirir conocimiento sin necesidad de ser programadas explícitamente para ello. Esta disciplina se subdivide en tres categorías principales: aprendizaje supervisado, aprendizaje no supervisado y semi-supervisado. Sin embargo, en los estudios revisados, se han abordado únicamente las dos primeras categorías.

En el aprendizaje supervisado, los algoritmos aprenden a partir de un conjunto de datos de entrenamiento, lo que requiere la intervención humana para proporcionar un resultado esperado. Por otro lado, en el aprendizaje no supervisado, los algoritmos aprenden de los datos buscando patrones o relaciones entre ellos. En este tipo de enfoque, los algoritmos no dependen de la intervención humana y tienen la capacidad de agrupar registros con características similares. En la Tabla 10 los algoritmos de aprendizaje máquina fueron agrupados por tipo (supervisado y no supervisado).

La lógica heurística (Burgos et al., 2007) es un procedimiento para el cual existe un alto grado de confianza que se clasifica como método heurístico, encontrando soluciones de acuerdo con la información disponible. En el razonamiento basado en reglas, la base de conocimiento inicial se basa en variables y un conjunto de reglas compuestas por una premisa y una conclusión, definidas para el problema y cómo se puede resolver (Moya-Rodríguez et al., 2012). El procesamiento del lenguaje natural mencionado por (Russell & Norving., 2004), es otra disciplina de la Inteligencia Artificial que se ocupa de la formulación e investigación de mecanismos para la comunicación humano-máquina mediante el uso del lenguaje natural, ya sea escrito, hablado o interactuado físicamente.

Por otro lado, los enfoques matriciales se representan a través de gráficos, representación basada en una indexación bidimensional de los nodos (Ponz et al., 2011). Finalmente, los modelos híbridos están formados por la integración de varios enfoques, donde cada uno tiene una representación y sigue un proceso diferente, pero cuando el modelado finaliza, éste tiende a generar una solución casi perfecta para brindar sugerencias al usuario, abonando con las ventajas de cada enfoque que forma el modelo híbrido.

La Figura 7 presenta la distribución de estudios primarios que tratan distintos enfoques definidos. Es importante señalar que algunos de estos estudios primarios, al diseñar experimentos para comparar técnicas de recomendación, abordaron más de un enfoque en sus investigaciones.

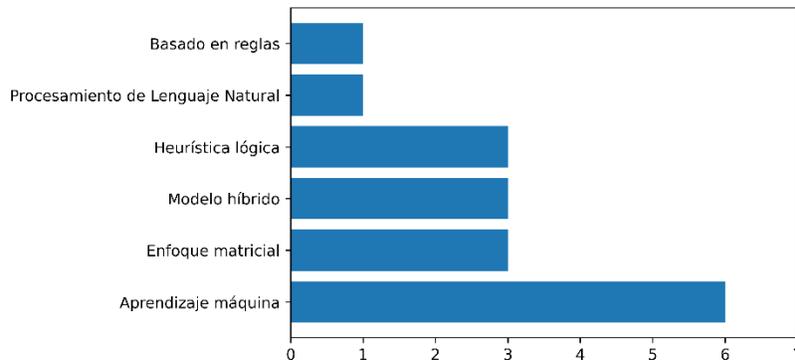


Figura 7. Distribución de estudios primarios que abordan los diferentes enfoques en la recomendación médica.

Como se observa en la Figura 7, el enfoque más frecuentemente mencionado en los estudios primarios es el uso del aprendizaje automático. Esto se debe a que se considera que no es muy complejo de implementar y que el entrenamiento proporcionado puede lograr una precisión adecuada en las recomendaciones, como se destacó en el estudio de Zhang et al. (2017). La implementación de técnicas de aprendizaje automático se ha centrado en la optimización del proceso de recomendación, particularmente en la reducción del consumo de recursos.

El segundo enfoque más utilizado es la lógica heurística. A pesar de no ser tan riguroso como otros enfoques, ha demostrado ofrecer un buen rendimiento en las sugerencias proporcionadas.

A continuación, se encuentran los modelos híbridos, que se distinguen por la integración de dos o más algoritmos que trabajan en conjunto para procesar la información del paciente y brindar recomendaciones altamente personalizadas.

Finalmente, se mencionan otros enfoques, como los enfoques matriciales, basados en reglas y de procesamiento del lenguaje natural. Estos enfoques son viables según la literatura, ya que ofrecen recomendaciones altamente personalizadas para los pacientes. Sin embargo, se ha observado que pueden tener un impacto negativo en el rendimiento del sistema y requerir un consumo significativo de recursos para su operación.

4.2 PI2. ¿Cómo se evalúan los sistemas de recomendación médica?

Tal como se estableció en la planificación de esta investigación, el objetivo de esta segunda pregunta es identificar la forma de evaluar el desempeño de los sistemas de recomendación médica. Para este tipo de sistemas, se han reportado diversas métricas de evaluación, siendo la métrica de precisión (*precision*) la más destacada. La precisión mide la proporción de verdaderos positivos en relación con el número total de valores positivos predichos, permitiendo así evaluar la calidad del modelo en tareas de clasificación.

Otra métrica importante es la recuperación (*recall*). Esta métrica proporciona información sobre la capacidad del modelo para identificar correctamente la cantidad de registros de interés. En otras palabras, evalúa el porcentaje de usuarios que están interesados en un servicio específico y cuántas recomendaciones acertadas se logran en este sentido.

Otra métrica de rendimiento destacada es la Raíz del Error Cuadrático Medio (*RMSE* por sus siglas en inglés), que calcula el valor medio de la diferencia. Se utiliza principalmente cuando se trata de aprendizaje automático supervisado. Cuanto menor sea el *RMSE*, más precisas serán las predicciones. Un *RMSE* igual a 1 es el valor óptimo en el que los pronósticos tienen una precisión excepcional. El *RMSE* tiene ventajas sobre la métrica de error absoluto medio, ya que destaca los errores grandes, entre los pequeños, destacando los valores atípicos que dan importancia a los datos de salida.

La Tabla 11 detalla cada métrica de evaluación encontrada en los estudios primarios seleccionados.

Métrica	Referencias
Precisión (Precision)	(Han et al.,2018)(Bjansali & Nagwani, 2021)(Swarnalatha et al., 2019)
Recuperación (Recall)	(Swarnalatha et al., 2019)(Meng et al., 2019)
Root Mean Square Error (RMSE)	(Meng et al., 2019)(Zhang et al., 2017)
F-Measure	(Bjansali & Nagwani, 2021)(Swarnalatha et al., 2019)
Métrica propia	(Mawardi et al., 2020)
Área bajo la curva (AUC)	(Meng et al., 2019)
Mean Absolute Error (MAE)	(Meng et al., 2019)
Proporción de Falsos positivos (FPR)	(Swarnalatha et al., 2019)
Accuracy	(Swarnalatha et al., 2019)
Hit Ratio@n	(Han et al.,2018)

Tabla 11. Métricas de evaluación reportadas por los estudios primarios seleccionados.

4.3 PI3. ¿Cuáles fueron los resultados de la evaluación de los sistemas de recomendación médica?

Con las técnicas de Aprendizaje Máquina encontradas y las métricas de evaluación reportadas, es importante conocer los resultados de sus evaluaciones, con el fin de facilitar la toma de decisiones en futuras implementaciones de sistemas de recomendación. Para ello aplica la síntesis narrativa para describir los diferentes resultados y conclusiones de los estudios.

Entre los hallazgos más importantes, podemos mencionar que el filtrado colaborativo y un modelo de confianza basado en el historial de consultas que tiene un paciente con el médico, obtienen márgenes considerablemente buenos a través de un modelo heurístico, en la medición de confianza y recomendando un 3% más de pacientes con médicos relevantes para ellos (Han et al.,2018).

El modelo de matriz de factorización mixta combinado con un proceso de selección basado en reglas funcionó mejor que simplemente usar el filtrado colaborativo. La aplicación de la codificación ICD-9 ayudó al rendimiento del modelo (Han et al., 2018b).

El clasificador Random Forest supera a los clasificadores C4.5 y Naïve Bayes en cuanto a su precisión. Los algoritmos de clasificación C4.5 Naïve Bayes superaron al algoritmo J48 (Bjansali & Nagwani, 2021). El factor de confianza radica en qué tan familiarizado está un paciente con un médico (Mondal et al., 2020).

El *RMSE* obtenido de las matrices de factorización híbridas es significativamente bajo, lo que demuestra que proporciona una precisión adecuada en los recomendadores médicos (Zhang et al., 2017). El mejor método entre C4.5 y K-Vecino más cercano es el clasificador K-Vecino más cercano, que reporta una precisión del 85% (Mawardi et al., 2020). Un enfoque bayesiano tiende a ser el clasificador más adecuado para un sistema de recomendación médico, ya que proporciona una precisión casi perfecta. Sin embargo, este método requiere que el usuario ingrese suficiente información para que el clasificador pueda ser entrenado y tener los al menos un caso para cada valor de clase (Kannan et al., 2020). Finalmente, las valoraciones de los pacientes pueden ser extremas, es decir, satisfechos o insatisfechos según la escala de valoración. Este tipo de evaluaciones afectan las recomendaciones que se brindan (Guo et al., 2016).

4.4 PI4. ¿Qué trabajo futuro se identificó en los estudios seleccionados?

En esta cuarta y última pregunta de investigación, el objetivo fue identificar las oportunidades reportadas por los estudios encontrados sobre Sistemas de recomendación médica. A continuación, se describen los principales trabajos o mejoras futuras que se proponen.

En el estudio de Han et al. (2018a), se propone la expansión del modelo de filtrado colaborativo mediante la incorporación de información auxiliar, como metadatos médicos y de pacientes, transformándolo en un enfoque híbrido. Además, se plantea la implementación del sistema de recomendación en un entorno de producción y prueba, con el objetivo de obtener comentarios explícitos de los pacientes como retroalimentación.

En Han et al. (2018b), se sugiere la implementación de una función de emparejamiento dentro del sistema de recomendación, diseñada según las necesidades y preferencias de los pacientes, considerándola una característica clave en los servicios de salud digitales. También se propone evaluar la utilidad de las recomendaciones mediante encuestas controladas aleatoriamente y su integración en el sistema para obtener retroalimentación implícita, lo que contribuiría a la personalización de las sugerencias dentro del modelo. En Bhansali et al. (2021), se propone comparar y evaluar varios algoritmos de clasificación.

En Venkatesh et al. (2022), se destaca la importancia de evaluar el uso de la tecnología big data para generar recomendaciones de alimentos y medicamentos para los usuarios. Asimismo, se considera mejorar la eficiencia de los registros médicos personales y la adición de más algoritmos de recomendación en este estudio.

El trabajo de Singh et al. (2016) propone la ampliación de la base de datos de médicos en su sistema, permitiendo a los usuarios programar citas con los médicos, no limitándose a proporcionar recomendaciones, sino también facilitando el contacto con los profesionales. Además, se plantea la introducción de una función que permita a los usuarios compartir sus experiencias con los médicos y proporcionar comentarios sobre la eficacia de sus consultas.

En Zhang et al. (2017), se describe una línea de trabajo abierta que permite que las preferencias de los usuarios cambien con el tiempo y que importe datos de las redes sociales, lo que podría mejorar aún más la precisión de las recomendaciones.

Swarnalatha et al. (2019) propone la integración de redes heterogéneas y redes neuronales para mejorar el rendimiento de las recomendaciones. En Kannan et al. (2020), se plantea el desarrollo de un conjunto de datos que incluya información sobre enfermedades para relacionar a los médicos especialistas con esas enfermedades.

Además, en el mismo trabajo de Kannan et al. (2020), los autores se enfocan en implementar un modelo de operación por voz para permitir que el sistema se adapte al usuario. También experimentan con la adaptación de sugerencias personalizadas según la ubicación del paciente. Por último, en Guo et al. (2016), se propone incorporar un enfoque de agregación de clasificación no supervisada.

5. Amenazas a la validez

Es importante mencionar que existen algunas amenazas a la validez de este trabajo. En primer lugar, es relevante mencionar que los autores no son expertos en el campo médico. No obstante, la investigación se enfocó en el análisis de técnicas de Inteligencia Artificial que permiten una recomendación más precisa en el ámbito médico, especialmente en lo que respecta a la selección de médicos, al diseñar e implementar un software con estas características.

En segundo lugar, es importante tener en cuenta el número de trabajos obtenidos con los criterios de selección inicialmente propuestos. No obstante, se realizó un esfuerzo para abordar esta limitación mediante un proceso de búsqueda de "bola de nieve" hacia adelante y hacia atrás, lo que permitió la inclusión de otros estudios primarios incluso desde fuentes de datos adicionales, con el fin de ampliar la base de evidencia y reducir esta preocupación.

6. Conclusiones y trabajo futuro

Este artículo presentó una Revisión Sistemática de la Literatura, basada en la guía propuesta por por Kitchenham, et al. (2015), con el fin de identificar técnicas de Inteligencia Artificial y su desempeño, que apoyan a los ingenieros de software en la toma de decisiones a implementar en sistemas de recomendación, específicamente por recomendación de los médicos. En este trabajo, se seleccionaron 12 estudios primarios de seis fuentes de datos.

Como resultado de esta investigación se identificaron técnicas, que se clasificaron en: aprendizaje automático (aprendizaje supervisado y aprendizaje no supervisado), lógica heurística, basado en reglas, procesamiento del lenguaje natural, enfoque matricial y modelos híbridos.

Se identificaron diferentes métricas de evaluación en MRS, que fueron: Precisión, Recuperación, Error cuadrático medio (RMSE), Índice de aciertos, Medida f1, Precisión, Puntuación F, Tasa de falsos positivos, Error absoluto medio (MAE), Área bajo la curva (AUC) y cálculos manuales (Síntomas).

Además, se han identificado varias áreas para futuras investigaciones, siendo una de las más destacadas la inclusión de un mayor volumen de datos en el sistema. Esto podría ampliar la base de información disponible y enriquecer las recomendaciones. También se plantea la adición de funcionalidades adicionales para mejorar la experiencia del usuario y la eficacia del sistema.

Asimismo, se sugiere explorar la integración de más algoritmos y observar su comportamiento en el sistema, con el propósito de determinar su factibilidad y evaluar si pueden brindar resultados más sólidos y precisos.

Finalmente, se propone dar continuidad a esta investigación, explorando las oportunidades que ofrece la literatura para mejorar las recomendaciones. Esto podría incluir el uso de algoritmos ensamblados, una estrategia que rara vez se ha abordado en la literatura, o la implementación de enfoques híbridos que aprovechen las ventajas de múltiples técnicas para lograr recomendaciones más efectivas.

Referencias

- [OMS] Organización Mundial de la Salud (s.f.). Observatorio Global para la eSalud. <https://www.who.int/observatories/global-observatory-for-ehealth>
- Bhansali, A., & Nagwani, N. K. (2021, May). A Prototype of Doctor Recommendation System Using Classification Algorithms. In 2021 Emerging Trends in Industry 4.0 (ETI 4.0) (pp. 1-4). IEEE.
- Bobadilla, J., Ortega, F. Hernando, A. y Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46. Pp. 109-132.
- Bravo, J. (2021, 7 mayo). Doctor Google y la salud digital. *El Economista*. <https://www.economista.com.mx/opinion/Doctor-Google-y-la-salud-digital20210507-0041.html>
- Burgos, D., Herder, E., & Olmedilla, D. (2007). Ten Competence: Building the European Network for the continuous development of competences. *Inteligencia Artif.*, 11(33), 79-84.
- Calero Valdez, A., Ziefle, M., Verbert, K., Felfernig, A. y Holzinger, A. (2016). Recommender Systems for Health Informatics: State-of-the-Art and Future Perspectives. *ML for Health Informatics, LNAI 9605*, pp. 391-414.
- Deng, Z., Hong, Z., Zhang, W., Evans, R. y Chen, Y. (2019). The Effect of Online Effort and Reputation of Physicians on Patients' Choice: 3-Wave Data Analysis of China's Good Doctor Website. *Journal of medical internet research*, 21(3), p. 1.
- Guo, L., Jin, B., Yao, C., Yang, H., Huang, D., & Wang, F. (2016). Which doctor to trust: a recommender system for identifying the right doctors. *Journal of medical Internet research*, 18(7), e6015.
- Gupta, S., & Bindal, A. K. (2022, November). Multi-Modality Recommender Systems: A Review. In *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 88-93). IEEE.
- Han, Q., de Troya, I. M. D. R., Ji, M., Gaur, M., & Zejnilovic, L. (2018, June). A collaborative filtering recommender system in primary care: Towards a trusting patient-doctor relationship. In 2018 IEEE International Conference on Healthcare Informatics (ICHI) (pp. 377-379). IEEE.
- Han, Q., Ji, M., Martinez De Rituerto De Troya, I., Gaur, M., & Zejnilovic, L. (2018). A hybrid recommender system for patient-doctor matchmaking in primary care. *Proceedings-2018 IEEE 5th International Conference on Data Science and Advanced Analytics, DSAA 2018*, 481-490.
- Iglesias Osores, S. y Saavedra Camacho, J. L. (2020). Tendencias de búsquedas en internet por la pandemia COVID-19 en Perú. *Revista Cubana de Higiene y Epidemiología*, 57.
- Kannan, R. J., Tamakuwala, H., Kale, S., & Bhowmick, H. R. (2020, September). Doctor Finder: Find doctors on the Go. In *IOP Conference Series: Materials Science and Engineering* (Vol. 925, No. 1, p. 012038). IOP Publishing.
- Katarya, R. (2017, December). A systematic review of group recommender systems techniques. In *2017 International conference on intelligent sustainable systems (ICISS)* (pp. 425-428). IEEE.
- Kitchenham, B. A., Budgen, D., & Brereton, P. (2015). *Evidence-based software engineering and systematic reviews* (Vol. 4). CRC press.
- Mawardi, V. C., & Naga, D. S. (2020, December). Application of Recommendation Medical Specialty Doctors Based on User Symptoms Using the C4. 5 Method and K-Nearest Neighbor. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1007, No. 1, p. 012152). IOP Publishing

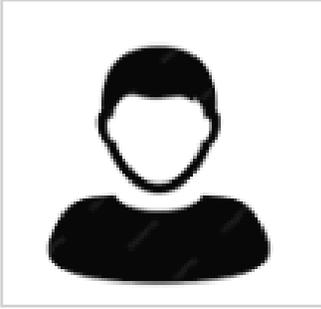
- Meng, S., Fan, S., Li, Q., Wang, X., Zhang, J., Xu, X., ... & Bhuiyan, M. Z. A. (2022). Privacy-aware factorization-based hybrid recommendation method for healthcare services. *IEEE Transactions on Industrial Informatics*, 18(8), 5637-5647.
- Mondal, S., Basu, A., & Mukherjee, N. (2020). Building a trust-based doctor recommendation system on top of multilayer graph database. *Journal of Biomedical Informatics*, 110, 103549.
- Moya-Rodríguez, J. L., Becerra-Ferreiro, A. M., & Chagoyén-Méndez, C. A. (2012). Utilización de Sistemas Basados en Reglas y en Casos para diseñar transmisiones por tornillo sinfín. *Ingeniería Mecánica*, 15(1), 01-09.
- Patel, B., Desai, P., & Panchal, U. (2017, March). Methods of recommender system: A review. In *2017 international conference on innovations in information, embedded and communication systems (ICIIECS)* (pp. 1-4). IEEE.
- Pincay, J., Terán, L., & Portmann, E. (2019, April). Health recommender systems: a state-of-the-art review. In *2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG)* (pp. 47-55). IEEE.
- Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and software technology*, 64, 1-18.
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., ... & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version, 1(1), b92.
- Ponz Tienda, J. L., Benloch Marco, J., Andrés Romano, C., & SENABRE, D. (2011). Un algoritmo matricial RUPSP/GRUPSP" sin interrupción" para la planificación de la producción bajo metodología Lean Construction basado en procesos productivos. *Revista de la Construcción*, 10(2), 90-103.
- Priego Álvarez, H. R. (2005). Consumo en salud. Análisis mercadológico del comportamiento del consumidor sanitario. [Tesis Doctoral]. Universidad Autónoma de Barcelona.
- Ramos, O. (2019, 2 diciembre). Sistemas de recomendación | Qué son, tipos y ejemplos. GraphEverywhere. <https://www.grapheverywhere.com/sistemas-de-recomendacion-que-son-tipos-yejemplos/>
- Russell, S. J., & Norvig, P. (2004). *Inteligencia Artificial: un enfoque moderno* (No. 04; Q335, R8y 2004.).
- Singh, A., Kaur, I. G. P., & Dabas, C. (2018, August). Get-a-Doc: A Doctor Recommender System. In *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 219-223). IEEE.
- Swarnalatha, S., Kesavarthini, I., Poornima, S., & Sripriya, N. (2019, February). Med-recommender system for predictive analysis of hospitals and doctors. In *2019 International Conference on Computational Intelligence in Data Science (ICCIDS)* (pp. 1-5). IEEE.
- Thongchotchat, V., Sato, K., & Suto, H. (2021, May). Recommender System Utilizing Learning Style: Systematic Literature Review. In *2021 6th International Conference on Business and Industrial Research (ICBIR)* (pp. 184-187). IEEE.
- Trang Tran, T. N., Felfernig, A., Trattner, Ch. y Holzinger, A. (2020). Recommender systems in the healthcare domain: state-of-the-art and research issues. *Journal of Intelligent Information Systems*. (2021)57, pp. 171-201.

Venkatesh, B. H., Sai, A. P., Reddy, M. R., & Fathimabi, S. K. (2022, June). Cloud Based Personal Health Record Management System and Medical Recommender System. In 2022 7th International Conference on Communication and Electronics Systems (ICCES) (pp. 1744-1749). IEEE.

Villafranco, G. (2019, 19 abril). Esta plataforma te ayuda a encontrar al médico especialista que necesitas. <https://www.forbes.com.mx/esta-plataforma-te-ayuda-a-encontrar-al-medico-especialista-que-necesitas/>

Zhang, Y., Chen, M., Huang, D., Wu, D., & Li, Y. (2017). iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization. *Future Generation Computer Systems*, 66, 30-35.

Notas Biográficas



Javier Alberto Calderón-Blas es estudiante de la Licenciatura en Ingeniería de Software de la Facultad de Estadística e Informática de la Universidad Veracruzana. Sus intereses se centran en la Inteligencia Artificial, Ciencia de Datos y UI/UX.



María Angélica Cerdán, Maestra en Ciencias de la Computación, docente de la Universidad Veracruzana Facultad de Estadística e Informática, Profesor con Perfil Deseable. Jefa de la División de la Ingeniería en Sistemas Computacionales, del Instituto Tecnológico Superior de Xalapa.



Ángel J. Sánchez-García es doctor en Inteligencia Artificial por la Universidad Veracruzana en 2018. Desde 2017 es profesor de tiempo completo de la Facultad de Estadística e Informática, Universidad Veracruzana. Sus intereses incluyen el aprendizaje automático y la Ingeniería de Software.



Saúl Domínguez-Isidro es doctor en Inteligencia Artificial, adscrito a la Facultad de Estadística e Informática de la Universidad Veracruzana. Sus intereses se centran en la solución de problemas mediante técnicas de inteligencia computacional y el desarrollo de software.



This work is under a Creative Commons
Attribution-NonCommercial-ShareAlike 2.5 Mexico license .