



**Aprendizaje automático para la predicción de la demanda de atención en  
urgencias. Revisión sistemática rápida.**

Juan Pablo Mogollón Avaunza

Universidad EAN

Facultad de Administración, Finanzas y Ciencias Económicas

Maestría en Administración

Bogotá, Colombia

15/06/2024

**Aprendizaje automático para la predicción de la demanda de atención en  
urgencias. Revisión sistemática rápida.**

**Juan Pablo Mogollón Avaunza**

Trabajo de grado presentado como requisito para optar al título de:

**Magíster en Administración de Empresas - MBA**

Director (a):

Fabian Gerardo Diaz Garzón

Modalidad:

**Monografía**

Universidad EAN

Facultad de Administración, Finanzas y Ciencias Económicas

Maestría en Administración

Bogotá, Colombia

15/06/2024

Nota de aceptación:

---

---

---

---

---

---

Firma del jurado

---

Firma del jurado

---

Firma del director del trabajo de grado

Bogotá, 15/06/2024

A Luisa,

Este trabajo es el resultado de  
innumerables horas de esfuerzo y  
dedicación, y no habría sido posible sin tu  
presencia constante y comprensión.

Gracias por ser mi compañera de vida, mi  
confidente y mi mayor admiradora. Este  
logro es tuyo tanto como mío.

JP

### **Agradecimientos**

Quiero expresar mi más profundo agradecimiento a Fabián, mi director de trabajo de grado, quien me animo y oriento en la ejecución y finalización de este trabajo. Gracias por tu orientación experta y tu confianza en mis habilidades, estoy profundamente agradecido por tu generosidad y compromiso.

JP

## Resumen

La demanda creciente de atenciones en los servicios de urgencias es un fenómeno permanente, que causa sobreocupación, incrementa los tiempos de atención y la insatisfacción de los pacientes y los equipos de salud. El aprendizaje automático ha tomado relevancia en el cuidado de la salud en un sin número de campos de actuación y podría ser útil para eliminar la incertidumbre del volumen de llegadas de pacientes a urgencias y realizar predicciones precisas de la demanda de atenciones en este servicio, lo cual podría contribuir a mejorar la planeación de recursos y en consecuencia a mejorar los tiempos de atención.

El propósito de este trabajo es analizar la literatura actual sobre el uso del aprendizaje automático en la predicción de la demanda de urgencias, identificando métodos efectivos, desafíos clave y posibles aplicaciones prácticas. Se realizó una revisión sistemática rápida usando como fuentes las bases de datos EBSCO Host, PubMed, SciELO, Scopus, Science Direct, y Ovid, limitada a las publicaciones realizadas entre 2014 y 2024. En total se incluyeron 21 publicaciones en la revisión.

Se utilizó una herramienta de inteligencia artificial para realizar la extracción de datos de los estudios seleccionados, identificar los algoritmos de pronóstico y los variables de predicción más frecuentemente utilizados, la precisión de cada uno de los modelos de pronóstico implementados, sus limitaciones y aplicaciones prácticas.

En conclusión, el aprendizaje automático es útil para prever la demanda de atenciones a urgencias con porcentajes absolutos de error bajos. Sin embargo, no es posible generalizar un modelo para su aplicación en múltiples instituciones, y las aplicaciones prácticas de estas predicciones deben ser exploradas con mayor profundidad.

**Palabras clave:** Aprendizaje automático, modelos de predicción, pronóstico, demanda de atenciones, urgencias.

### **Abstract**

The increasing demand for care in the emergency department is a permanent phenomenon, causing overcrowding, increasing care times and dissatisfaction of patients and healthcare teams. Machine learning has become relevant in healthcare in a number of fields and could be useful to eliminate the uncertainty of the volume of patient arrivals in the ED and to make accurate predictions of the demand for care in this service, which could contribute to improve resource planning and consequently to improve care times.

The purpose of this paper is to review the current literature on the use of machine learning in ED demand prediction, identifying effective methods, key challenges and possible practical applications. A rapid systematic review was conducted using EBSCO Host, PubMed, SciELO, Scopus, Science Direct, and Ovid databases as sources, limited to publications published between 2014 and 2024. A total of 21 publications were included in the review.

An artificial intelligence tool was used to extract data from the selected studies, identify the most frequently used forecasting algorithms and predictor variables, the accuracy of each of the forecasting models implemented, their limitations and practical applications.

In conclusion, machine learning is useful for forecasting ED demand with low absolute error rates. However, it is not possible to generalise a model for application in multiple institutions, and the practical applications of these predictions need to be further explored.

**Keywords:** Machine learning, prediction models, forecasting, demand for emergency care, emergency department.

## Contenido

	Pág.
Lista de Figuras.....	10
Lista de Tablas .....	11
Introducción .....	12
Objetivos.....	15
Objetivo general .....	15
Objetivos específicos.....	15
Justificación .....	16
Marco Teórico.....	18
Atención de urgencias .....	18
Generalidades del aprendizaje automático .....	19
Pronósticos de la demanda de un servicio .....	23
Revisión sistemática .....	24
Hipótesis .....	29
Variables .....	30
Metodología .....	32
Pregunta de investigación .....	32
Metodologías aplicables .....	32
Criterios de inclusión y exclusión.....	34
Estrategia de búsqueda .....	35
Tamizaje y selección de estudios a incluir en el informe final.....	37
Extracción de datos .....	38
Trabajo de Campo .....	39

<b>Identificación de publicaciones relevantes .....</b>	<b>39</b>
<b>Selección de estudios.....</b>	<b>39</b>
<b>Resultados.....</b>	<b>41</b>
<b>Descripción general de las publicaciones incluidas.....</b>	<b>41</b>
<b>Métodos de aprendizaje automático utilizados en la predicción de la demanda de     atenciones.....</b>	<b>44</b>
<i>Modelos de regresión. ....</i>	<i>47</i>
<i>Modelos de series temporales. ....</i>	<i>48</i>
<i>Modelos de redes neuronales. ....</i>	<i>51</i>
<i>Modelos de ensamble. ....</i>	<i>53</i>
<b>Predictores seleccionados.....</b>	<b>54</b>
<b>Desempeño predictivo .....</b>	<b>59</b>
<b>Limitaciones reconocidas en el desarrollo de los estudios .....</b>	<b>62</b>
<b>Aplicabilidad de los modelos predictivos de demanda de atenciones, en las     operaciones del servicio de urgencias .....</b>	<b>63</b>
<b>Propuesta de solución a la problemática .....</b>	<b>64</b>
<b>Discusión.....</b>	<b>66</b>
<b>Conclusiones y Trabajo Futuro.....</b>	<b>70</b>
<b>Conclusiones.....</b>	<b>70</b>
<b>Trabajo futuro.....</b>	<b>71</b>
<b>Referencias.....</b>	<b>72</b>

### Lista de Figuras

	Pág.
<b>Figura 1. Proceso de selección de publicaciones para incluir en la revisión .....</b>	<b>40</b>
<b>Figura 2. Número de publicaciones incluidas en la revisión por año de la publicación .....</b>	<b>43</b>
<b>Figura 3. Número de publicaciones incluidas en la revisión por País en la que se realizó el estudio .....</b>	<b>43</b>
<b>Figura 4. Número de publicaciones incluidas en la revisión de acuerdo con el tamaño estimado del servicio de urgencias.....</b>	<b>44</b>
<b>Figura 5. Error absoluto porcentual medio (MAPE) de los modelos evaluados con horizonte de tiempo de un día.....</b>	<b>61</b>

### Lista de Tablas

	Pág.
Tabla 1. Metodologías aplicables para el desarrollo del estudio .....	33
Tabla 2. Bases de datos y ecuaciones de búsqueda .....	35
Tabla 3. <i>Objetivo y lugar de realización de los estudios seleccionados para la revisión</i> .....	41
Tabla 4. <i>Modelos, métricas de desempeño y alcance del pronóstico en los estudios incluidos en la revisión</i> .....	45
Tabla 5. <i>Modelos de regresión empleados en los estudios incluidos en la revisión</i> .....	47
Tabla 6. <i>Modelos de series temporales empleados en los estudios incluidos en la revisión</i> .....	49
Tabla 7. <i>Modelos de redes neuronales empleados en los estudios incluidos en la revisión</i> .....	51
Tabla 8. Predictores utilizados en los estudios analizados en la revisión .....	55

## Introducción

La sobreocupación de los servicios de urgencias producto del incremento en la demanda de atenciones es un problema común que representa uno de los mayores desafíos de la atención sanitaria a nivel global. Colombia no es ajena a este fenómeno, no es infrecuente que los medios de comunicación denuncien que las urgencias hospitalarias se encuentran congestionadas y que los pacientes esperan atención durante horas (Asociación Colombiana de Empresas de Medicina Integral, ACEMI, 2018). Este fenómeno que afecta principalmente las instituciones de mediana y alta complejidad conlleva prolongación de los tiempos de atención, mayor riesgo de eventos adversos, aumento de las tasas de hospitalización y mortalidad, e insatisfacción de los usuarios (Estrada Atehortúa & Zuluaga Gómez, 2020; Sun et al., 2013).

Las causas de esta situación son múltiples, se han descrito factores asociados a la población y al servicio. A los pacientes se les atribuye: malos hábitos de uso, desconocimiento del concepto de urgencia, búsqueda de incapacidades, evasión del pago de cuotas moderadoras y búsqueda de atención médica en menor tiempo al ofertado en consulta ambulatoria (Asociación Colombiana de Empresas de Medicina Integral, ACEMI, 2012).

Entre las causas asociadas a la operación del servicio se cuentan: mala planeación en la ejecución de las actividades y desbalance entre oferta y la demanda en los servicios de urgencias. Esto genera cuellos de botella, prolonga los tiempos de atención y congestiona el servicio (Furterer, 2014).

A nivel institucional el desbalance entre la oferta y la demanda se presenta debido a la variabilidad de la demanda de pacientes, con comportamientos estacionales de acuerdo

con la época del año, la semana epidemiológica, el día de la semana, la hora del día, e incluso el clima, lo cual dificulta la asignación eficiente de recursos y la planificación.

De acuerdo con Estrada Atehortua y Zuluaga Gómez (2020) se han implementado diversas estrategias para gestionar la sobreocupación de los servicio de urgencias, ninguna de ellas relacionada con la predicción de la demanda de atenciones del servicio, los modelos de gestión implementados no han sido replicables y la situación ha empeorado.

Por otro lado, en los últimos años el aprendizaje automático ha tomado relevancia en el cuidado de la salud, teniendo aplicación en un sin número de campos de actuación (Javaid et al., 2022). Uno de esos campos es la mejora de la eficiencia operacional en la que el aprendizaje automático ha sido utilizado para realizar predicciones en diversas situaciones la mayoría de ellas relacionadas con la predicción de enfermedad, pero también ha sido útil para predecir la utilización de recursos y el costo médico, e incluso el riesgo de ocurrencia de eventos adversos (Barda et al., 2020; Berger et al., 2020).

El aprendizaje automático ofrece la posibilidad de analizar grandes cantidades de datos, identificar patrones y tendencias y podría constituirse como una herramienta prometedora para predecir la demanda, y a partir de estas predicciones optimizar la planificación de recursos.

El presente estudio busca examinar cómo se ha utilizado el aprendizaje automático en este contexto, analizando los métodos utilizados, las limitaciones existentes y las posibles aplicaciones futuras.

Este documento presenta una estructura organizada que aborda de manera integral la investigación propuesta. En primer lugar, se describen los objetivos que establecen el alcance del estudio. Posteriormente, en la justificación se expondrá el por qué del tema elegido, tratando de demostrar la relevancia y la importancia de este. A continuación, se

desarrolla un marco teórico que proporciona el contexto necesario del problema de investigación, seguido de la formulación de una hipótesis y la identificación de las variables relevantes. Se describe la metodología utilizada, y el trabajo de campo realizado. El documento cierra con un espacio para la discusión de los resultados obtenidos, y para enunciar conclusiones y posibles áreas de trabajo futuro.

## **Objetivos**

### **Objetivo general**

Analizar la literatura actual sobre el uso del aprendizaje automático en la predicción de la demanda de urgencias, identificando métodos efectivos, desafíos clave y posibles aplicaciones prácticas.

### **Objetivos específicos**

- Describir los métodos de aprendizaje automático más comúnmente utilizados en la predicción de la demanda diaria de atenciones del servicio de urgencias.
- Identificar el tipo de predictores que permiten mejorar el desempeño predictivo de un modelo de pronóstico de la demanda en urgencias.
- Establecer la aplicabilidad de un modelo de predicción de la demanda de atenciones en el contexto de un servicio de urgencias.

### **Justificación**

El volumen diario de personas que solicita atención en los servicios de urgencias es un factor crítico en la gestión de recursos de cualquier institución hospitalaria; un pronóstico preciso de esta demanda de atención puede mejorar la planeación y la asignación de recursos, y en consecuencia contribuir a mejorar la eficiencia del servicio.

El pronóstico de la demanda diaria y horaria de un servicio de urgencias involucra múltiples variables, las cuales exhiben patrones complejos, en general no lineales, y con un comportamiento que incluye estacionalidades por semana epidemiológica, día de la semana, hora del día, entre otros; en este contexto, la obtención de una predicción precisa requiere del enfoque avanzado que aportan los métodos de aprendizaje automático.

El aprendizaje automático ha demostrado ser útil en diversos campos de la atención sanitaria, y representa un avance tecnológico significativo que puede transformar positivamente la forma en que se administra y se entrega la atención en los servicios de urgencias. Su utilización en la predicción de las llegadas diarias a urgencias podría tener un gran impacto en la operación de este servicio, y permitiría a los hospitales anticipar y prepararse mejor para atender la demanda, asegurando oportunidad, seguridad y efectividad en la atención de los pacientes. El éxito en la predicción y el subsecuente ajuste de recursos podría contribuir a la reducción de los tiempos de espera al permitir una mejor gestión de la capacidad y la carga de trabajo.

Por otro lado, la aplicación práctica de este trabajo puede contribuir al desarrollo de algoritmos de predicción que podrían integrarse en sistemas de gestión hospitalaria para mejorar la eficiencia operativa. El desarrollo de la capacidad de predecir la demanda de

atenciones no solo beneficiaría a los servicios de salud en la actualidad, sino que podría contribuir en la mejora continua del sistema de salud en el largo plazo.

## Marco Teórico

### Atención de urgencias

Los servicios de urgencias tienen un papel fundamental en la sociedad para recibir a los pacientes que necesitan atención médica urgente, representando casi el 14% de las visitas hospitalarias en Estados Unidos y siendo las unidades más concurridas y congestionadas de los hospitales (Berikol & Berikol, 2020; K. J. W. Tang et al., 2021).

La atención en estos servicios suele involucrar un alto volumen de pacientes, recursos limitados y toma de decisiones de forma rápida (Piliuk & Tomforde, 2023, 2023); con implicaciones que potencialmente pueden cambiar la vida de las personas (Schneider et al., 1998). Muchos factores entre ellos el envejecimiento de la población, el aumento de la esperanza de vida y la mayor prevalencia de enfermedades crónicas han contribuido a incrementar su demanda y por ende su saturación, prolongando los tiempos de espera y la duración total de la atención (Bruballa et al., 2016).

En Colombia el acceso a urgencias está garantizado por el Sistema General de Seguridad Social en Salud (SGSSS), entendiendo urgencia como “la alteración de la integridad física y/o mental de una persona, causada por un trauma o por una enfermedad de cualquier etiología que genere una demanda de atención médica inmediata y efectiva tendiente a disminuir los riesgos de invalidez y muerte” (Decreto 412 de 1992, 1992, p. 412; Vanegas & Restrepo, 2022).

Si bien, los servicios de urgencias son responsables de organizar los recursos humanos, materiales, tecnológicos y financieros para asegurar el cuidado indispensable e inmediato de las personas que presentan una urgencia, esto no siempre se logra; haciendo que los pacientes deban esperar durante horas por la atención (Bobrovitz et al., 2017; Rodriguez et al., 2018). Esta situación frecuentemente conlleva que los pacientes

abandonen el servicio de urgencias sin recibir atención médica (Griner et al., 2020; Hobbs et al., 2000).

No solo por la necesidad de tomar decisiones de forma rápida para disminuir el riesgo de eventos adversos, de las tasas de hospitalización y mortalidad, sino para evitar la insatisfacción de los usuarios la gestión de los tiempos de urgencias resulta ser un factor crítico en la operación del servicio de urgencias (Asociación Colombiana de Empresas de Medicina Integral, ACEMI, 2018; Javaid et al., 2022; Vezyridis & Timmons, 2014).

En la búsqueda de soluciones, la aplicación de herramientas de inteligencia artificial y aprendizaje automático ha demostrado ser de utilidad en la gestión y organización de los recursos de los servicios de urgencias contribuyendo a mejorar la toma de decisiones en el servicio (Marco et al., 2021; Piliuk & Tomforde, 2023; K. J. W. Tang et al., 2021). Haciendo uso de estas herramientas es posible pronosticar el volumen de atenciones en un momento determinado, lo cual puede beneficiar la operación del servicio, entregando información que permita reorganizar los recursos de tal forma que se logren menores tiempos de atención y de esta forma mejores resultados en salud y una disminución de la insatisfacción de los usuarios del servicio.

Se ha demostrado que estas mejoras son plausibles y que pronosticando la cantidad de personas que solicitaran atención en un momento determinado, sus tiempos de estancia, y alineando estas predicciones con la asignación de recursos, en un trabajo mancomunado entre los equipos de salud y los científicos de datos se obtienen mejoras sostenibles en las operaciones de los servicios de salud (Griner et al., 2020; Gul & Celik, 2020; Sookmee et al., 2024).

### **Generalidades del aprendizaje automático**

El término "aprendizaje automático" (Machine Learning, ML) hace referencia a diversas técnicas estadísticas que permiten a los ordenadores aprender de la experiencia

sin ser programados explícitamente para ello (Javaid et al., 2022). Dicho de otra forma, el aprendizaje automático es una tecnología que utiliza programas informáticos para analizar grandes volúmenes de datos. Estos programas, que suelen llamarse algoritmos, se entrenan con información proporcionada por el usuario, por ejemplo, datos sobre pacientes y sus síntomas, o información histórica del volumen de atenciones; de esta forma el algoritmo puede identificar patrones y hacer predicciones.

En un sentido más romántico “el aprendizaje automático es un paradigma que puede referirse al aprendizaje a partir de la experiencia pasada (que en este caso son datos anteriores) para mejorar el rendimiento futuro” (K. P. Das & J, 2023).

Para que el aprendizaje automático pueda ser útil los datos estructurados son organizados en características únicas en un conjunto de datos que se utiliza para el entrenamiento del modelo, modelo que tras ser entrenado se utiliza para realizar inferencias sobre los nuevos datos (K. J. W. Tang et al., 2021). De esta forma la computadora puede encontrar puntos comunes, patrones en un conjunto de datos y llegar a una conclusión sobre la correlación de los datos sin estar específicamente programado para llegar a ese resultado (K. P. Das & J, 2023).

La capacidad de los algoritmos de aprendizaje automático para encontrar correlaciones entre diferentes atributos relacionados con ciertos eventos permite encontrar estrategias para optimizar los recursos sanitarios a través de modelos predictivos. Hay diferentes tipos de algoritmos de aprendizaje automático que adoptan diferentes enfoques para comprender los conjuntos de datos y construir modelos predictivos, ya sean de clasificación o de regresión. En los modelos de clasificación, el algoritmo clasifica el patrón de prueba en una clase específica. En regresión, el algoritmo intenta predecir un número específico basándose en el patrón de los datos de prueba.

Desde el punto de vista de las cuatro facetas funcionales de la analítica de datos: descriptiva, diagnóstica, predictiva y prescriptiva, el aprendizaje automático se constituye como un pilar de la analítica predictiva, a partir de la cual se pueden recomendar cursos de acción y conducir la toma de decisiones basada en datos (Gudivada, 2017).

El aprendizaje automático ya está presente en la atención sanitaria y ha beneficiado múltiples puntos de la cadena de valor de la atención sanitaria que incluyen mejoras en la experiencia de atención, diagnóstico, tratamiento, gestión del riesgo, entre otros aspectos; las posibilidades de aplicación futuras son ilimitadas, gracias a la informática, por supuesto, hasta las partes más ínfimas e insignificantes de cualquier operación pueden simplificarse hasta rozar la perfección (Abdelaziz et al., 2018; Ahmad et al., 2018; Char et al., 2020; Javaid et al., 2022).

Las herramientas de aprendizaje automático han demostrado ser capaces, a través de la automatización, de mejorar la capacidad de toma de decisiones inteligentes en la atención primaria/terciaria de una persona, lo que contribuye a mejorar la calidad de vida de miles de millones de personas en todo el mundo, la implementación de estas herramientas se ha visto favorecida por la implementación de registros médicos electrónicos y el impulso dado por los sistemas sanitarios que están utilizando herramientas de Big Data para el análisis de información (Gupta & Katarya, 2020; Sendak et al., 2020).

Como se mencionó previamente la literatura relevante ha demostrado que la sobreocupación de los servicios de urgencias y la subsecuente prolongación de los tiempos de atención tiene impactos en la calidad de atención del paciente incluso relacionados con mayores niveles de mortalidad. El aprendizaje automático y también el aprendizaje profundo ha demostrado que potencialmente es capaz de discernir patrones de datos recopilados a lo largo de los años y arrojar ideas para mejorar los procesos de

urgencias ayudando a los administradores el múltiples tareas de la operación de los servicios de urgencias dando como resultado inclusive una carga reducida para los profesionales médicos involucrados en la atención en este servicio (K. J. W. Tang et al., 2021).

La alineación de los procesos de ciencia de datos con la atención asistencial y los procesos administrativos en una institución de salud resulta ser útil para facilitar la planeación de servicios, la asignación de recursos y obtener mejores resultados operativos, de salud y en la satisfacción de los usuarios y del personal sanitario responsable de su atención (Griner et al., 2020).

No es posible implementar modelos de aprendizaje automático sin datos, por fortuna la evolución de los sistemas de información hospitalario (HIS), así como el de los sistemas de información de los servicios de apoyo diagnóstico, LIS en el laboratorio clínico y RIS en imágenes diagnósticas, entre otros, han facilitado que el cuidado de la salud se vea influenciado y beneficiado por la gestión de la información que proviene de las actividades clínicas, la generación de pronósticos a partir de estos grandes volúmenes de datos es solo una de las aplicaciones que se ha visto beneficiada con el almacenamiento y gestión que las historias clínicas electrónicas han propiciado (Azadi et al., 2023; Hong et al., 2018; L. Tang et al., 2022).

Sin embargo, no todo es color de rosa, si bien hoy en día la analítica de datos y el aprendizaje automático tienen un papel relevante en la organización de los sistemas sanitarios, y sus algoritmos ofrecen soluciones que puede ayudar a los administradores de servicio de salud a tomar la decisión correcta en el momento adecuado, aún se debe lograr que los usuarios finales de la información la empleen de forma coherente para lograr el propósito de mejorar la prestación de servicios de salud (Tawhid et al., 2021; Ward et al., 2014).

De regreso al contexto de urgencias el aprendizaje automático ha demostrado ser de utilidad en numerosas actividades del servicio de urgencias, siendo capaz de mejorar la exactitud en la asignación de la clasificación de triage, predecir mortalidad, e identificar riesgos críticos en la atención, así como la necesidad de recursos para atenderlos (Ivanov et al., 2021; Piliuk & Tomforde, 2023; Stover-Baker et al., 2012; Sulaiman et al., 2023); siendo capaces incluso de identificar la influencia del medio ambiente (temperatura, calidad del aire y polución) en el comportamiento de la demanda de urgencias (Jian et al., 2023; Kousha & Castner, 2016; Schaffer et al., 2012; Stieb et al., 2009); o de predecir la demanda de atenciones de poblaciones especiales, como los adultos mayores, niños, o pacientes con diagnóstico de asma u otras enfermedades específicas (Bunney et al., 2023; L. T. Das et al., 2017; Newcomb & Li, 2008; Wang et al., 2023). Pese a la gran cantidad de informes del uso los algoritmos analíticos predictivos en los servicios de urgencias, su utilización no ha podido ser escalada y generalizada.

### **Pronósticos de la demanda de un servicio**

La previsión es una herramienta fundamental en la toma de decisiones y la planificación. La fascinación por la incertidumbre del futuro y el desafío que esto significa, ha impulsado a individuos y organizaciones a tratar de predecirlo, con el ánimo de intentar reducir sus riesgos y aumentar sus beneficios. La amplia gama de aplicaciones de la previsión requiere una variedad diversa de métodos para enfrentar los desafíos del mundo real (Petropoulos et al., 2022).

Si simplificamos la cadena de valor del proceso típico de atención en urgencias, en pocos pasos: a) solicitud de atención; b) triage; c) registro y admisión; d) consulta; e) apoyo diagnóstico y terapéutico; f) egreso; fácilmente podemos concluir que se asemeja a la cadena de suministro de un servicio; en la cual la planificación y la gestión de la demanda del servicio juega un papel fundamental, asegurando que se cuente con la

capacidad adecuada para atender a los clientes sin demoras y de manera efectiva (Heizer et al., 2020; Slack & Lewis, 2019).

De tal manera que una buena predicción de la demanda contribuye significativamente en la mejora de la operación de la cadena de suministro de un servicio o producto, lo cual conlleva un impacto positivo en los niveles de servicio. Existen diversas técnicas para predecir la demanda: a) predicción clásicas basada en técnicas cualitativas, el juicio de expertos y la experiencia pasada; b) predicciones basadas en técnicas cuantitativas, modelos estadísticos y de aprendizaje automático; o c) una combinación de métodos cualitativos y cuantitativos (Petropoulos et al., 2022).

A pesar de la importancia que conlleva la predicción de la demanda en la cadena de operación de un servicio, tal como la atención de urgencias, no es común que en nuestro medio local se utilicen predicciones de la demanda para ajustar los recursos requeridos en el servicio de urgencias para atenderla. Esto puede deberse al desconocimiento de estas técnicas por los administradores del servicio; la no integración de personal capacitado en estas técnicas en los servicios de salud; la complejidad de los patrones de demanda del servicio; o la incapacidad del sistema de salud para direccionar la demanda de atención de urgencias a través de un modelo de asignación geográfica.

### **Revisión sistemática**

Las revisiones sistemáticas tienen su origen en las disciplinas de la salud y la medicina, sin embargo, también se utilizan en otras disciplinas como las ciencias sociales, las ciencias y la ingeniería (The University of Newcastle, 2024). Una revisión sistemática intenta identificar, evaluar y sintetizar toda la evidencia empírica que cumple con los criterios de elegibilidad preespecificados para responder una pregunta de investigación específica, utilizando métodos explícitos y sistemáticos con miras a

minimizar el sesgo, para producir hallazgos más confiables que sirvan de base para la toma de decisiones (Higgins & Green, 2011).

De acuerdo, con Higgins & Green (2011) las características clave de una revisión sistemática son:

- un conjunto de objetivos claramente establecidos con criterios de elegibilidad predefinidos para los estudios
- una metodología explícita y reproducible
- una búsqueda sistemática que intenta identificar todos los estudios que cumplirían los criterios de elegibilidad
- una evaluación de la validez de los hallazgos de los estudios incluidos,
- una presentación sistemática y una síntesis de las características y hallazgos de los estudios incluidos

Los tipos de revisión sistemática más comunes son:

- a) Revisión sistemática (Systematic Review): las revisiones sistemáticas son el tipo de revisión más conocido en determinados campos. Su objetivo es realizar una búsqueda exhaustiva y exhaustiva y recomendaciones para la práctica. Las características incluyen: Reunir y evaluar el conocimiento publicado dentro de un área o tema en particular; Adherirse a fijar pautas de conducta; Requiere revisión por pares antes de su publicación.
- b) Revisión de alcance (Scoping Review): Las revisiones de alcance proporcionan un “análisis ambiental” (evaluación preliminar) de la literatura actualmente disponible en un área de investigación. Las revisiones de alcance se diferencian de las revisiones de mapeo en que el resultado es sólo la revisión, no la realización de revisiones o investigaciones adicionales. Las

características incluyen: Ser una búsqueda rápida que pueda informar si se necesita una revisión completa. Acercamiento sistemático.

- c) Revisión de literatura (Literature Review). Las características incluyen: establecer un marco teórico para su tema/área temática; definir términos, definiciones y terminología clave; identificar estudios, modelos, estudios de casos, etc. que respalden su tema definir/establecer su área de estudio, es decir, su tema de investigación.
- d) Revisión Integrativa (Integrative Review). Las características incluyen: Generar o refinar una teoría o hipótesis; Combinar la investigación empírica y teórica; Examinar la investigación sobre un fenómeno de salud determinado; Informar las políticas y prácticas sanitarias.
- e) Metaanálisis (Meta-Analysis). El metaanálisis proporciona una combinación estadística de los resultados de estudios cuantitativos. Su objetivo es una búsqueda exhaustiva y exhaustiva. El metaanálisis se puede incluir en las revisiones sistemáticas, pero tenga en cuenta que no todas las revisiones sistemáticas contienen un metaanálisis. Las características incluyen: Capaz de reunir estudios más pequeños para contribuir a una imagen más amplia del impacto; Ahorro de tiempo para los tomadores de decisiones, en comparación con la revisión de estudios individuales.
- f) Revisión cualitativa (Qualitative Review). Las revisiones cualitativas son estudios interpretativos que pueden incorporar informes de usuarios y observaciones de profesionales para permitir una comprensión más amplia que la que permitirían los datos únicamente. Las características incluyen: Permite la síntesis de información no estadística. Puede resaltar temas en estudios individuales.

Específicamente una revisión sistemática rápida, sigue al menos los siguientes pasos:

- a) **Formulación de la Pregunta de Investigación:** Define una pregunta de investigación específica, clara y relevante para el contexto; y que puede ser respondida mediante este método de investigación.
- b) **Elaboración de un Protocolo Simplificado:** Diseña un protocolo de revisión conciso, con elementos clave: criterios de inclusión y exclusión de estudios, la estrategia de búsqueda preliminar y los métodos de síntesis de datos.
- c) **Búsqueda Rápida de Evidencia:** exploración rápida y eficiente en fuentes de literatura relevantes, como bases de datos académicas, buscadores especializados y literatura gris, a través del uso de términos de búsqueda.
- d) **Selección Rápida de Estudios:** Aplica criterios de inclusión y exclusión de manera rápida y eficiente para seleccionar los estudios relevantes. En algunas ocasiones, para acelerar el proceso, puede ser necesario limitar la cantidad de estudios incluidos.
- e) **Extracción Simplificada de Datos:** Extrae información clave de los estudios seleccionados de manera concisa, centrándote en los resultados pertinentes para responder la pregunta de investigación.
- f) **Síntesis Preliminar de la Evidencia:** Realiza una síntesis preliminar de los resultados de los estudios incluidos, identificando patrones o tendencias en la evidencia disponible.
- g) **Interpretación Preliminar de Resultados:** Interpreta los hallazgos de manera inicial, reconociendo las limitaciones inherentes a una revisión rápida y destacando áreas que requieren una exploración más detallada.

En general, se reconoce que una revisión sistemática rápida se enfoca en ser ágil y eficiente, priorizando la obtención de resultados en un tiempo reducido; estas limitaciones inherentes a la velocidad del proceso pueden conllevar que se pase por alto estudios relevantes o a que se realice una evaluación menos detallada de la calidad metodológica de los estudios incluidos (Featherstone et al., 2015; Tricco et al., 2017).

### **Hipótesis**

El uso de modelos de aprendizaje automático contribuye a identificar patrones y tendencias en los datos históricos de llegadas a urgencias, los cuales pueden incluir: tendencias estacionales, días festivos, condiciones meteorológicas y eventos especiales; siendo útiles para predecir de forma efectiva el número de pacientes que se espera que soliciten atención en un día y hora determinado en el servicio de urgencias.

La predicción de la demanda de los servicios de urgencias contribuye a reorganizar la planeación y disposición de los recursos disponibles, obteniendo mejoras en la operación del servicio que conlleven a una reducción de los tiempos de espera, incremento en la satisfacción de los usuarios y un mejor uso de los recursos disponibles para la atención médica.

## Variables

A continuación, se describen las principales variables o aspectos que se buscan identificar en el estudio:

- País en el que se realiza el estudio. Tipo de variable: Cualitativa nominal.
- Nombre de la institución en la que se realiza el estudio. Tipo de variable: Cualitativa nominal.
- Nombre del método o modelo aplicado. Nombre del algoritmo de aprendizaje computacional empleado en el estudio. Es una variable cualitativa nominal que identifica el modelo utilizado en cada estudio para obtener la predicción. Por ejemplo: ARIMA (AutoRegressive Integrated Moving Average): Un modelo que combina componentes autoregresivos, de media móvil e integrados para modelar series temporales; ETS (Exponential Smoothing methods): Métodos de suavizado exponencial que utilizan pesos exponencialmente decrecientes para observaciones pasadas; STLM (Seasonal decomposition of time series by LOESS): Un enfoque que descompone una serie temporal en componentes estacionales utilizando LOESS (Local Polynomial Regression Fitting); Modelos de Regresión Lineal Generalizada (GLM); Redes neuronales; Árboles de decisión, etc.
- Predictores. Variables o características utilizados para predecir la demanda. Variable cualitativa nominal, para identificar las covariables o variables de control utilizadas en la creación del modelo de pronóstico, por ejemplo: día de la semana, día festivo, atenciones del día anterior, temperatura, humedad, etc.
- Métrica de desempeño. Medida que se utilizó para evaluar la capacidad predictiva del modelo. Variable cualitativa nominal, utilizada para identificar el

nombre de la medida cuantitativa que se usó en el estudio para evaluar y comparar la eficacia y precisión de un modelo, ejemplo: Error Cuadrático Medio (Mean Squared Error, MSE), Error Absoluto Medio (Mean Absolute Error, MAE), Coeficiente de Determinación ( $R^2$ ), Error Porcentual Absoluto Medio (Mean Absolute Percentage Error, MAPE)

- Resultado de la métrica de desempeño. Variable cuantitativa continua, que identifica el valor obtenido en la métrica de desempeño del modelo de predicción.

## Metodología

Este estudio está enfocado en el análisis y evaluación de la utilización de técnicas de aprendizaje automático (machine learning) para predecir la demanda de atenciones en los servicios de urgencias.

### Pregunta de investigación

- P1. ¿Cuáles son los modelos o métodos de aprendizaje automático que han demostrado ser efectivos para predecir la demanda diaria de un servicio de urgencias en hospitales de cualquier nivel de complejidad, sean públicos o privados?
- P2. ¿Qué características o variables son relevantes en los modelos de aprendizaje automático para este propósito?
- P3. ¿Cómo son aplicadas las predicciones de demanda de atenciones en urgencias en la planeación y asignación de recursos a los servicios de urgencias?

### Metodologías aplicables

El objetivo del estudio y la respuesta a las preguntas de investigación puede ser logrado utilizando diversos enfoques metodológicos; entre ellos: revisión narrativa, integrativa, meta analítica, de *scoping* entre otros. En la tabla 1 se muestra un resumen de la aplicabilidad ventajas y desventajas de cada uno de estos métodos en el contexto de este estudio.

Teniendo en cuenta el equilibrio entre los recursos disponibles para la ejecución de este estudio, la aplicabilidad de los métodos, sus ventajas y desventajas. La revisión sistemática rápida es un método más estructurado y exhaustivo que una revisión narrativa, requiere menos recursos tanto de tiempo como de expertos disponibles que una revisión Delphi. Por otro lado, el desarrollo y alcance de los objetivos de este estudio

no requiere la exhaustividad ni integración de los resultados que aporta un metaanálisis; y dato que los objetivos propuestos requieren de una respuesta cuantitativa y no cualitativa, no se requiere de una revisión integrativa. Como objetivo no esta determinar si existen vacíos de conocimiento que puedan ser cubiertos en una investigación posterior por lo que una revisión tipo scoping sobre pasa el alcance de este estudio.

**Tabla 1**

*Metodologías aplicables para el desarrollo del estudio*

Metodología	Descripción	Aplicabilidad	Ventajas	Desventajas
Revisión Narrativa	Permite sintetizar y resumir la literatura sobre un tema sin seguir un método sistemático riguroso. Basada en la interpretación del investigador incluye una discusión amplia sobre conceptos, teorías y contextos (Baumeister & Leary, 1997).	Es útil para proporcionar una visión general sobre los modelos de aprendizaje automático, las variables relevantes y su aplicación en la planeación de recursos.	- Muy fácil de realizar. - Proporciona una visión amplia y contextual.	- Falta de rigurosidad. - Riesgo muy alto de sesgo. - Poco exhaustiva.
Revisión Integrativa	Combina datos de estudios teóricos y empíricos para proporcionar una comprensión más completa del fenómeno. Permite la inclusión de estudios cuantitativos y cualitativos (Whittemore & Knaf, 2005).	Útil para explorar de manera más amplia los métodos utilizados, las variables clave y las aplicaciones prácticas, integrando diferentes enfoques y perspectivas.	- Incluye una amplia variedad de estudios. - Proporciona una comprensión completa del tema. - Riesgo de sesgo: medio	- Complejidad en la síntesis de datos heterogéneos. - Toma más tiempo y esfuerzo.
Revisión Meta-analítica	Metodología estadística que combina los resultados de múltiples estudios cuantitativos para obtener una estimación más precisa de los efectos y relaciones investigadas (Borenstein et al., 2009).	Ideal para identificar y evaluar cuantitativamente los métodos y variables más efectivos en la predicción de la demanda de urgencias.	- Alta precisión en los resultados. - Método cuantitativo y robusto. - Riesgo de sesgo muy bajo	- Requiere estudios homogéneos. - Limitada a estudios cuantitativos. - Metodológicamente compleja.
Revisión de Alcance (Scoping Review)	Explora la extensión, alcance y naturaleza de la investigación sobre un tema, identificando lagunas en el conocimiento y áreas para futuras investigaciones (Arksey & O'Malley, 2005).	Adecuada para mapear los métodos utilizados, las variables relevantes y las aplicaciones prácticas, proporcionando una visión amplia de la literatura disponible.	- Identifica lagunas en la investigación. - Amplia cobertura del tema.	- No busca sintetizar resultados exhaustivamente. - Puede ser menos profunda en algunos aspectos.
Revisión Rápida	Utiliza métodos sistemáticos simplificados para evaluar y sintetizar la evidencia en un período de tiempo más corto que una revisión sistemática tradicional (Tricco et al., 2017).	Adecuada para una evaluación rápida y preliminar de la literatura sobre el uso del aprendizaje automático en la predicción de la demanda de urgencias.	- Tiempo de ejecución más corto. - Métodos sistemáticos.	- Sacrifica profundidad y exhaustividad, para conseguir un resultado en menor tiempo.
Estudio Delphi	Método estructurado de recolección de información basado en la consulta a un panel de expertos a través de múltiples rondas de cuestionarios, con el objetivo de llegar a un consenso (Hsu & Sandford, 2007).	Útil para identificar los métodos y predictores más relevantes según la opinión de expertos en el campo y evaluar su aplicabilidad en la planeación de recursos.	- Obtención de consenso entre expertos. - Identificación de prácticas y métodos relevantes.	- Proceso largo debido a múltiples rondas. - Dependencia de la disponibilidad y experiencia de los expertos.

Fuente: elaboración propia

Con base en lo expuesto se selecciona como metodología de desarrollo la revisión sistemática rápida, la cual permite seguir un método estructurado, aplicable en el desarrollo del estudio, y con la ventaja de poder obtener resultados en un tiempo más corto, sin comprometer el logro de los objetivos, o el riesgo de un sesgo medio, lo cual desde el punto de vista del autor es aceptable, para responder a las preguntas de investigación, y su aplicación en el contexto de la operación de un servicio de urgencias.

### **Criterios de inclusión y exclusión**

Se incluirán todas las publicaciones, realizadas en revistas científicas, libros y tesis, que apliquen métodos de aprendizaje automático para predecir la demanda de atención en los servicios de urgencias, publicados en los últimos 10 años (desde 2014 hasta la fecha actual). La publicación debe describir el desarrollo y evaluación de los modelos implementados; describir las variables utilizadas para la predicción (como datos demográficos, información clínica, condiciones meteorológicas, calendario, entre otros); y deben presentar las métricas de desempeño con las que evaluaron los modelos utilizados.

Se excluirán las publicaciones que cumplan los siguientes criterios:

1. Correspondan a cartas, resúmenes de conferencias o poster, o revisiones sistemáticas.
2. No relacionadas con la predicción de la demanda total de atención en los servicios de urgencias. Esto incluye las predicciones relacionadas con la demanda de poblaciones o enfermedades específicas, por ejemplo: demanda de atención de adultos mayores, COVID, infarto agudo de miocardio, etc); y las que realizan predicciones de demanda en otros servicios de atención diferentes a las urgencias, o predicen reingresos, mortalidad o la clasificación de triage. Aquellas que se relacionen con la predicción de admisiones a hospitalización, reingresos,

clasificación de triage, tiempos de espera o de estancia, también las que busquen cuantificar los niveles de congestión del servicio.

3. Texto completo no disponible.
4. Aquellas que no describan de forma detallada el modelo de pronóstico utilizado.
5. Las que no informen los predictores utilizados o no presenten los resultados de las métricas de desempeño del modelo.
6. En las que el alcance del pronóstico de la demanda no sea horario o diario.
7. Publicaciones en idioma diferente a inglés o español.

### Estrategia de búsqueda

Los términos de búsqueda que se utilizarán para identificar publicaciones de interés en las diferentes bases de datos son: “Machine learning”; “Predictive modeling”, Forecasting, Forecast, Prediction, “Emergency room”, “Emergency department”, “hospital emergencies”, Demand, Visits, Arrivals, Admission, “Time series análisis”. Estos términos deben estar contenidos en el título, resumen, o en las palabras claves o términos temáticos de la publicación.

Para realizar la identificación de las publicaciones de interés se utilizarán los recursos electrónicos de búsqueda disponibles, en la tabla 2 se listan las bases de datos que se consultaran con los correspondientes arreglos de los términos de búsqueda en cada una de ellas. En el caso de Science Direct no se permite realizar la búsqueda a través de una ecuación, la búsqueda se realiza usando términos de lenguaje natural.

**Tabla 2**

*Bases de datos y ecuaciones de búsqueda*

Aplicación de búsqueda y bases de datos	Términos de búsqueda
EBSCO Host <ul style="list-style-type: none"> <li>• eBook Academic Collection</li> </ul>	SU ( “Machine learning” or “Predictive modelling” or “Forecasting” or “Forecast” or “Prediction” or “Time series analysis”) AND (“Emergency room” or “Emergency department” or “Hospital emergencies”) AND (“Demand” or “Visits”

**Tabla 2**

*Bases de datos y ecuaciones de búsqueda*

Aplicación de búsqueda y bases de datos	Términos de búsqueda
<ul style="list-style-type: none"> <li>• Academic Search Premier</li> <li>• Business Source Premier</li> <li>• eBook Collection</li> <li>• EconLit with Full Text</li> <li>• Fuente Académica Premier</li> <li>• MathSciNe</li> </ul>	or "Arrivals" or "Admission" ) OR AB ( ("Machine learning" or "Predictive modelling" or "Forecasting" or "Forecast" or "Prediction" or "Time series analysis") AND ("Emergency room" or "Emergency department" or "Hospital emergencies") AND ("Demand" or "Visits" or "Arrivals" or "Admission" ) ) OR TI ( ("Machine learning" or "Predictive modelling" or "Forecasting" or "Forecast" or "Prediction" or "Time series analysis") AND ("Emergency room" or "Emergency department" or "Hospital emergencies") AND ("Demand" or "Visits" or "Arrivals" or "Admission" ) )
PubMed	(("Machine learning"[Title/Abstract] OR "Predictive modelling"[Title/Abstract] OR "Forecasting"[Title/Abstract] OR "Forecast"[Title/Abstract] OR "Prediction"[Title/Abstract] OR "Time series analysis"[Title/Abstract]) AND ("Emergency room"[Title/Abstract] OR "Emergency department"[Title/Abstract] OR "Hospital emergencies"[Title/Abstract]) AND ("Demand"[Title/Abstract] OR "Visits"[Title/Abstract] OR "Arrivals"[Title/Abstract] OR "Admission"[Title/Abstract])) OR (("Machine learning" or "Predictive modelling" or "Forecasting" or "Forecast" or "Prediction" or "Time series analysis") AND ("Emergency room" or "Emergency department" or "Hospital emergencies") AND ("Demand" or "Visits" or "Arrivals" or "Admission"))[MeSH Terms]
SciELO	(((ti:(Machine learning))) OR (ti:(Predictive modelling)) OR (ti:(Forecasting)) OR (ti:(Forecast)) OR (ti:(Prediction)) OR (ti:(Time series analysis))) AND ((ti:(Emergency room)) OR (ti:(Emergency department)) OR (ti:(Hospital emergencies))) AND ((ti:(Demand)) OR (ti:(Visits)) OR (ti:(Arrivals)) OR (ti:(Admission)))) OR (((ab:(Machine learning)) OR (ab:(Predictive modelling)) OR (ab:(Forecasting)) OR (ab:(Forecast)) OR (ab:(Prediction)) OR (ab:(Time series analysis))) AND ((ab:(Emergency room)) OR (ab:(Emergency department)) OR (ab:(Hospital emergencies))) AND ((ab:(Demand)) OR (ab:(Visits)) OR (ab:(Arrivals)) OR (ab:(Admission)))) (((ti:(aprendizaje automático)) OR (ti:(Modelo predictivo)) OR (ti:(Pronóstico)) OR (ti:(Predicción)) OR (ti:(Predecir)) OR (ti:(Series de tiempo))) AND ((ti:(emergencias)) OR (ti:(urgencias)) OR (ti:(hospital))) AND (ti:(Demand)) OR (ti:(visitas)) OR (ti:(llegadas)) OR (ti:(admisiones)))) OR (((ab:(aprendizaje automático)) OR (ab:(Modelo predictivo)) OR (ab:(Pronóstico)) OR (ab:(Predicción)) OR (ab:(Predecir)) OR (ab:(Series de tiempo))) AND ((ab:(emergencias)) OR (ab:(urgencias)) OR (ab:(hospital))) AND ((ab:(Demand)) OR (ab:(visitas)) OR (ab:(llegadas)) OR (ab:(admisiones))))
SCopus	TITLE ( ( "Machine learning" OR "Predictive modelling" OR "Forecasting" OR "Forecast" OR "Prediction" OR "Time series analysis" ) AND ( "Emergency room" OR "Emergency department" OR "Hospital emergencies" ) AND ( "Demand" OR "Visits" OR "Arrivals" OR "Admission" ) ) OR ABS ( ( "Machine learning" OR "Predictive modelling" OR "Forecasting" OR "Forecast" OR "Prediction" OR "Time series analysis" ) AND ( "Emergency room" OR "Emergency department" OR "Hospital emergencies" ) AND ( "Demand" OR "Visits" OR "Arrivals" OR "Admission" ) ) OR KEY ( ( "Machine learning" OR "Predictive modelling" OR "Forecasting" OR "Forecast" OR "Prediction" OR "Time series analysis" ) AND ( "Emergency room" OR "Emergency department" OR "Hospital emergencies" ) AND ( "Demand" OR "Visits" OR "Arrivals" OR "Admission" ) )
Science Direct	<ul style="list-style-type: none"> <li>• machine learning forecast Prediction emergency Arrivals visits Admission Demand</li> <li>• machine learning forecast emergency visits</li> </ul>
Ovid	(("Machine learning" or "Predictive modelling" or "Forecasting" or "Forecast" or "Prediction" or "Time series analysis") and ("Emergency room" or "Emergency department" or "Hospital emergencies") and ("Demand" or "Visits" or "Arrivals" or "Admission")).t Or (("Machine learning" or "Predictive modelling" or "Forecasting" or "Forecast" or "Prediction" or "Time series analysis") and ("Emergency room" or "Emergency department" or "Hospital emergencies") and ("Demand" or "Visits" or "Arrivals" or "Admission")).ab Or
<ul style="list-style-type: none"> <li>• Medline                         <ul style="list-style-type: none"> <li>○ Ovid MEDLINE(R)</li> </ul> </li> <li>• Cochrane Library                         <ul style="list-style-type: none"> <li>○ EBM Reviews - Cochrane Database of Systematic Reviews</li> <li>○ EBM Reviews - ACP Journal Club</li> </ul> </li> </ul>	

**Tabla 2**

*Bases de datos y ecuaciones de búsqueda*

Aplicación de búsqueda y bases de datos	Términos de búsqueda
<ul style="list-style-type: none"> <li>○ EBM Reviews - Database of Abstracts of Reviews of Effects</li> <li>○ EBM Reviews - Cochrane Clinical Answers</li> <li>○ EBM Reviews - Cochrane Central Register of Controlled Trials</li> <li>○ EBM Reviews - Cochrane Methodology Register</li> <li>● EBM Reviews - Health Technology Assessment</li> </ul>	<p>((("Machine learning" or "Predictive modelling" or "Forecasting" or "Forecast" or "Prediction" or "Time series analysis") and ("Emergency room" or "Emergency department" or "Hospital emergencies") and ("Demand" or "Visits" or "Arrivals" or "Admission"))).kw</p>

Fuente: Elaboración propia

**Tamizaje y selección de estudios a incluir en el informe final**

Se consultarán las bases de datos indicadas de acuerdo con las ecuaciones de búsqueda descritas, los resultados serán exportados de cada buscador en formato RIS y cargados en el gestor bibliográfico Zotero®, para iniciar el proceso de tamizaje y selección de las publicaciones. Utilizando Zotero® se identificarán y descartarán los registros duplicados.

Teniendo en cuenta que los resultados de la búsqueda se suelen encontrar más artículos de los que realmente se utilizarán en la revisión final. El proceso de selección permite filtrar los que no cumplen los criterios de inclusión. Durante este proceso se examinarán cada una de las publicaciones identificadas siguiendo los siguientes pasos:

1. Eliminar los registros duplicados de los resultados
2. Examinar los títulos y resúmenes para excluir los artículos que no cumplen los criterios de inclusión
3. Buscar el texto completo de los artículos pertinentes
4. Mantener juntos todos los informes que hagan referencia al mismo estudio
5. Examinar los artículos de texto completo según los criterios de inclusión
6. Tomar la decisión final sobre la inclusión de estudios

### **Extracción de datos**

Una vez seleccionados los estudios que se incluirán en la revisión se extraerán los datos relevantes de los estudios incluidos, con el fin de facilitar el análisis, la síntesis y la interpretación.

La extracción de datos se realizará apoyándose en el gestor bibliográfico Zotero® para los datos básicos de la publicación: Título, Resumen, e información que identifica la publicación; y en herramientas de inteligencia artificial como ChatPdf® para la extracción de datos específicos como las variables de interés descritas previamente.

Una vez eliminados los registros duplicados, se exportarán a un formato que pueda administrarse en una hoja de cálculo la cual contendrá las siguientes columnas: título de la publicación; autor(es); año; nombre de la publicación; volumen; número; páginas; resumen. También se agregarán las columnas objetivo; país; institución; método o algoritmo implementado; predictores; métrica de desempeño; resultados de la métrica de desempeño; limitaciones identificadas; aplicación práctica descrita; beneficios observados, las cuales serán diligenciadas para las publicaciones que se decida incluir en la revisión final.

## Trabajo de Campo

### Identificación de publicaciones relevantes

Durante el mes de abril de 2024 se realizó la búsqueda en cada una de las bases de datos de información científica utilizando las ecuaciones de búsqueda detalladas en la tabla 1. Para cada una de las bases de datos los resultados iniciales se filtraron por el año de publicación para obtener los estudios publicados entre 2014 y 2024. Como resultado de esta búsqueda se identificaron 7,083 posibles artículos relevantes en total, los cuales después del proceso de tamizaje y selección fueron reducidos a 23. El total de 7,083 artículos fue obtenido de EBSCO Host, (n = 578); PubMed, (n = 1,078); SciELO, (n = 83); Scopus, (n = 2,330); Science Direct, (n = 1,461); Ovid, (n = 1,553).

Los registros de cada base de datos fueron exportados de la plataforma online en formato RIS e importados al Zotero®, utilizando esta herramienta en una maquina local se identificaron y descartaron 2,565 registros duplicados. Los registros no duplicados fueron exportados en formato csv para posteriormente ser importados y tratados en una hoja de cálculo de MS Excel®, los campos de interés de la publicación fueron: título de la publicación; autor(es); año; nombre de la publicación; volumen; número; páginas; resumen. Esta hoja de cálculo se utilizó para realizar la lectura y análisis de los títulos y resúmenes de la publicación.

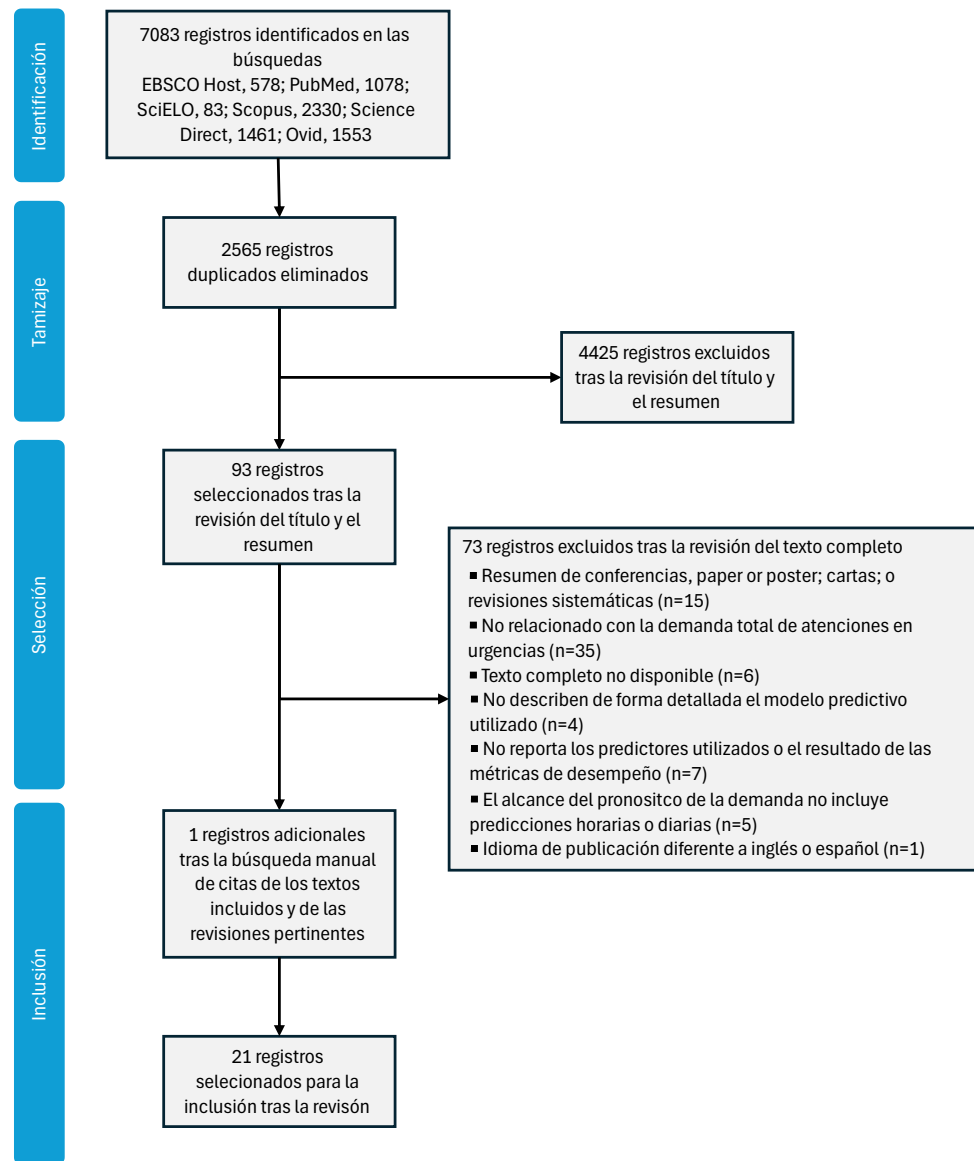
### Selección de estudios

Tras analizar los títulos y resúmenes de las publicaciones no duplicadas para verificar el cumplimiento de los criterios de inclusión fueron excluidos 4,425 registros. Se conservaron 93 registros para el análisis de texto completo. De estos, se descartaron 73 registros; la mayor parte de ellos (35) debido a que las predicciones no se enfocaban en la demanda total de visitas o atenciones en el servicio de urgencias, sino en la predicción

de reingresos, admisiones, tiempos de espera, tiempo total de atención, saturación del servicio, altas, flujo de los pacientes al interior del servicio, o enfermedades específicas. Posteriormente se incluyó un estudio realizado en Brasil en el año 2013, para un total de 21 registros incluidos en la revisión final (ver figura 1).

**Figura 1**

*Proceso de selección de publicaciones para incluir en la revisión*



Fuente: elaboración propia a partir de Prisma Statement (Page et al., 2021)

## Resultados

### Descripción general de las publicaciones incluidas

En la tabla 3 se presenta el año, objetivo y lugar de realización de los estudios incluidos en esta revisión. El 72% de los estudios incluidos corresponden a publicaciones del año 2020 o posteriores. Brasil, Portugal y Estados Unidos son los países en los que se encontró el mayor número de publicaciones relevantes (ver figuras 2 y 3).

**Tabla 3.**

Objetivo y lugar de realización de los estudios seleccionados para la revisión.

Publicación	Objetivo	País	Institución	Demanda máxima diaria (1)
Marcilio et al. (2013)	Predecir el volumen de pacientes en el departamento de emergencias que los gerentes de atención médica pueden emplear para una mejor planificación y asignación de recursos	Brasil	No informado	401 - 500
Calegari et al. (2016)	Desarrollar y comparar modelos de pronóstico para predecir la demanda futura de atención médica en un departamento de emergencias	Brasil	Hospital de Clínicas de Porto Alegre (HCPA)	50 - 100
Carvalho-Silva (2018)	Evaluar diferentes modelos de pronóstico para predecir la demanda de pacientes en el Departamento de Emergencias	Portugal	Hospital de Braga	501 - 600
Nas & Koyuncu (2019)	Utilizar algoritmos de Machine Learning para predecir la tasa de llegada de pacientes a un Departamento de Emergencias	Turquía	Hospital privado Ceyhan Çinar	No estimado
Whitt & Zhang (2019)	Investigar métodos de pronóstico para los totales diarios de llegadas y su aplicación para predecir el nivel de ocupación por hora en un departamento de emergencias	Israel	Rambam Hospital	101 - 200
Harrou et al. (2020)	Proponer un método efectivo para pronosticar las visitas diarias y por hora en un departamento de urgencias utilizando el algoritmo Variational autoencoder (VAE)	Francia	Centro Hospitalario Regional de Lille	50 - 100
Yousefi et al. (2020)	Investigar los factores que afectan la demanda diaria en un departamento de emergencias (ED) y proporcionar una herramienta de pronóstico en un hospital público para horizontes de hasta siete días	Brasil	Hospital Risoleta Tolentino Neves (RTNH), Belo Horizonte	501 - 600
Choudhury & Urena (2020)	Desarrollar un modelo de pronóstico preciso de las llegadas a la sala de emergencias por hora y realizar pruebas de ajuste del modelo	Estados Unidos	UnityPoint Health, Des Moines, Iowa,	No estimado
Erkamp et al. (2021)	Evaluar modelos de predicción para el número de visitas a urgencias basados en variables de calendario y clima	Países Bajos	Hospital Jeroen Bosch	50 - 100
Vollmer et al. (2021)	Desarrollar un marco predictivo novedoso para comprender la dinámica temporal de la demanda hospitalaria en los servicios de urgencias	Inglaterra	Hospitales St Mary's y Charing Cross, Londres	201 - 300
Rocha & Rodrigues (2021)	Desarrollar un modelo de pronóstico preciso para predecir el número de admisiones en el departamento de emergencias de un hospital	Portugal	Hospital de Braga	501 - 600
Zhang et al. (2022)	Mejorar la precisión en el pronóstico de llegadas de pacientes a departamentos de emergencia mediante la combinación de variables de calendario y meteorológicas	China	Primer Hospital Afiliado de la Universidad de	301 - 400

**Tabla 3.**

Objetivo y lugar de realización de los estudios seleccionados para la revisión.

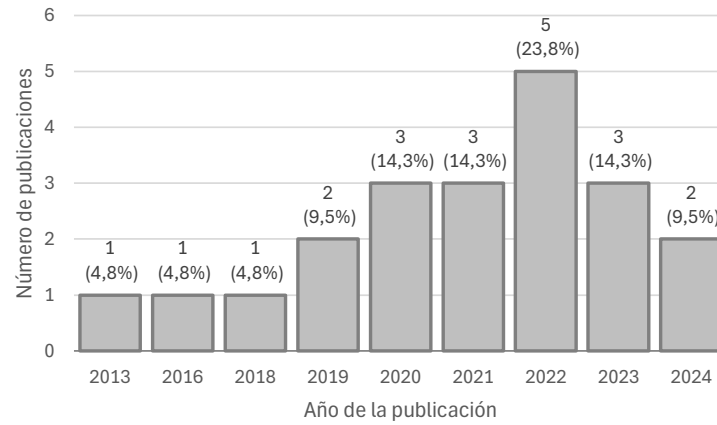
Publicación	Objetivo	País	Institución	Demanda máxima diaria (1)
Tuominen et al. (2022)	Investigar si la precisión predictiva de las llegadas del día siguiente podría mejorarse utilizando un alto número de variables explicativas potencialmente relevantes y documentar dos procesos de selección de características que buscan identificar qué subconjunto de variables está asociado con el número de llegadas del día siguiente	Finlandia	Medicina de Anhui, Hefei Hospital Universitario de Tampere	201 - 300
Murtas et al. (2022)	Predecir la demanda de visitas a los servicios de urgencias en Milán y desarrollar un sistema de alerta de 2 días para pronosticar alta demanda	Italia	No informado	201 - 300
Zhao et al. (2022)	Predecir las visitas diarias al departamento de emergencias de un hospital utilizando enfoques de aprendizaje profundo	Singapur	Singapore General Hospital (SGH)	401 - 500
Petsis et al. (2022)	Utilizar un algoritmo avanzado de ensamble basado en árboles y explicabilidad en Inteligencia Artificial (XAI) para pronosticar y explicar el número de visitas de pacientes al departamento de emergencias de un hospital público	Grecia	University Hospital of Ioannina	301 - 400
Hu et al. (2023)	Mejorar la precisión en la predicción de la demanda de atención en el departamento de emergencias (ED) mediante el uso de información en tiempo real	Estados Unidos	Hospital de la Universidad de Columbia, New York	No estimado
Gafni-Pappas & Khan (2023)	Examinar combinaciones de variables a través de aprendizaje automático para aumentar la precisión de la predicción y determinar los factores más predictivos de las visitas totales al departamento de emergencias	Estados Unidos	St. Joseph Mercy Hospital, Ann Arbor, MI, y NYU Langone Medical Center, Nueva York, NY.	201 - 300
Vieira et al. (2023)	Explorar las diferencias entre varias metodologías para pronosticar las admisiones diarias en un departamento de emergencias, considerando patrones estacionales únicos y múltiples	Portugal	Hospital de Braga	401 - 500
Peláez-Rodríguez et al. (2024)	Proponer dos enfoques interpretables basados en algoritmos de Machine Learning para predecir con precisión las visitas a los servicios de urgencias hospitalarias	España	Clínica Universidad de Navarra; Sedes Pamplona y Madrid	50 - 100
Tuominen et al. (2024)	Predecir la ocupación de urgencias en un departamento de emergencias utilizando modelos avanzados de aprendizaje automático y datos multivariantes	Finlandia	Centro de Cirugía Vasculare Radiología Intervencionista y el Hospital Universitario de Tampere	401 - 500

1. Estimación del volumen máximo de atenciones a partir de información gráfica contenida en la publicación. Este dato no es reportado explícitamente en ninguno de los estudios.

Fuente: Elaboración propia.

**Figura 2**

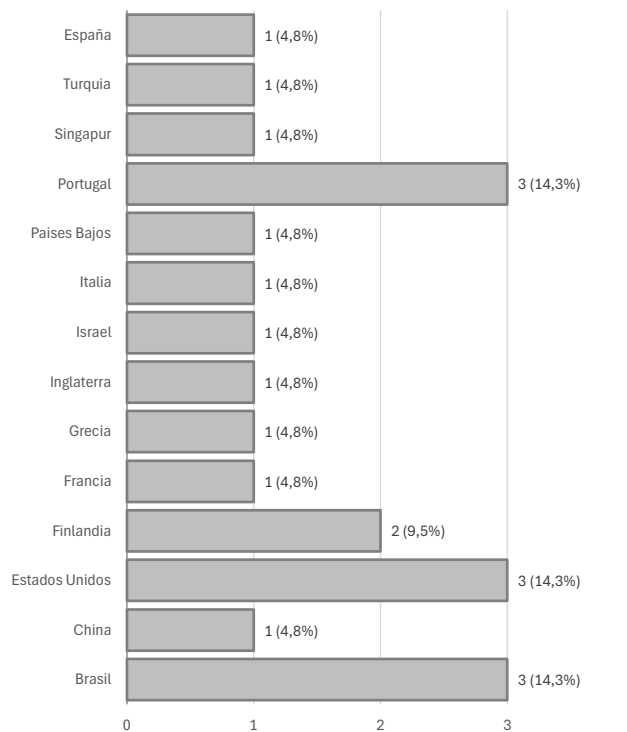
*Número de publicaciones incluidas en la revisión por año de la publicación*



Fuente: elaboración propia

**Figura 3**

*Número de publicaciones incluidas en la revisión por País en la que se realizó el estudio*

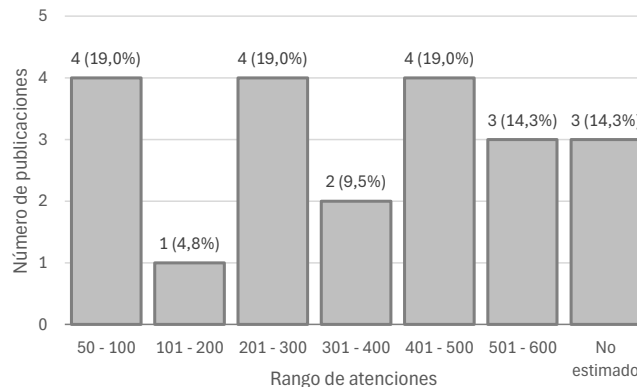


Fuente: elaboración propia

Si bien, en ninguno de los estudios se presenta explícitamente el tamaño de los servicios de urgencias en los que se realizaron las predicciones, se realizó una estimación del tamaño estimado de cada servicio con base en la información presentada en cada estudio, encontrando que el 23,8% de las instituciones tienen un flujo entre 50 y 200 atenciones día; en 28,5% de ellas el volumen máximo está entre 201 y 400 atenciones diarias; el 33,3% de las instituciones, en las que se puede estimar, el volumen máximo esta entre 401 y 600 pacientes por día; en tres estudios no fue posible estimar el volumen de atenciones diario (ver figura 4).

**Figura 4**

*Número de publicaciones incluidas en la revisión de acuerdo con el tamaño estimado del servicio de urgencias*



Fuente: elaboración propia

### **Métodos de aprendizaje automático utilizados en la predicción de la demanda de atenciones**

Los métodos empleados para pronosticar la demanda del servicio de urgencias son diversos, en la tabla 4 se muestra una visión general de los métodos utilizados, dividiéndolos en método principal y métodos de comparación; las métricas de desempeño empleadas y el modelo que obtuvo el mejor desempeño y el valor de la métrica de desempeño del mejor modelo.

**Tabla 4**

*Modelos, métricas de desempeño y alcance del pronóstico en los estudios incluidos en la revisión.*

Publicación	Modelo principal	Modelos de comparación	Métricas de desempeño	Mejor modelo	Resultado de la métrica de desempeño	Alcance del pronóstico
Marcilio et al. (2013)	GLM	GEE, SARIMA	MAPE	GLM	MAPE: 11.5%	Diario
Calegari et al. (2016)	SARIMA	SS, SMHW, MSARIMA	MAPE	SARIMA	MAPE: 6.23%	Diario
Carvalho-Silva (2018)	ARIMA	SARIMA, SS, MLR, HW	MAPE	ARIMAX	MAPE: 5.92%	Diario
Nas & Koyuncu (2019)	LSTM	DT, SVM, GBM, AB, SG, KNN, RF, MLP, LR	MAPE	LSTM (diaria) RF (horaria)	MAPE: 16.79% (diaria) 49.67% (horaria)	Diario
Whitt & Zhang (2019)	SARIMAX	SARIMA, ARIMA	MSE, MAPE	SARIMAX	MAPE: 8.4%	Diario
Harrou et al. (2020)	VAE	RNN, LSTM, BiLSTM, ConvLSTM, GRUs, RBM, CNN	R2, RMSE, MAE	VAE	MAE 0.295	Diario
Yousefi et al. (2020)	LSTM	MLR, ARIMA, SVR, GLM, GEE, SARIMA	MAPE, R2	LSTM	MAPE: 4.89%	Diario
Choudhury & Urena (2020)	ARIMA	RNN, TBATS, HW	RMSE	ARIMA	RMSE: 1.55	Horario
Erkamp et al. (2021)	MLR		MAPE	MLR + Calendar variables	MAPE: 8.6%	Diario
Vollmer et al. (2021)	ARIMA	SS, STLM, GBM, KNN, RF, GLM	MAE, MAPE	GLM	MAPE: 6.8%	Diario
Rocha & Rodrigues (2021)	Ensamble: SARIMA, RNN-1L, RNN-3L y XGBoost	ARIMA, SS, GLM	MAE, RMSE, sMAPE	Ensamble: SARIMA, RNN-1L, RNN-3L y XGBoost	sMAPE: 23.2%	Diario
Zhang et al. (2022)	LSTM	GLM, KNN, SVR, RIDGE, XGB, AB, GBM, RF	RMSE, MAE, MAPE	LSTM (horaria) SVR (diaria)	MAPE: 50% (horaria) MAPE: 8.81% (diaria)	Diario y horario
Tuominen et al. (2022)	ARIMA	ARIMAX, RLS, RF, LMS	MAPE	ARIMAX	MAPE: 6.6%	Diario
Murtas et al. (2022)	ARIMA		MAPE	ARIMA	MAPE: 8.1%	Diario
Zhao et al. (2022)	Deep Stacked Architecture	ARIMA, Prophet,	RMSE, MAE, MAPE	DSA	MAPE: 5.41%	Diario

**Tabla 4**

*Modelos, métricas de desempeño y alcance del pronóstico en los estudios incluidos en la revisión.*

Publicación	Modelo principal	Modelos de comparación	Métricas de desempeño	Mejor modelo	Resultado de la métrica de desempeño	Alcance del pronóstico
	(LSTM, GRU y RNN)	ConvLSTM, BiLSTM				
Petsis et al. (2022)	XGB		RMSE, MAE, MAPE	XGB	MAPE: 6.5%	Diario
Hu et al. (2023)	SARIMAX	SARIMA, XGB, DT, GLM	RMSE, MAPE	SARIMAX	MAPE: 8.7%	Diario
Gafni-Pappas & Khan (2023)	RF, GBM	ARIMA, SS, Prophet	RMSE	RF	RMSE: 18.94	Diario
Vieira et al. (2023)	ARIMA	ARIMAX, SARIMA, TBATS	RMSE, MAE, MAPE	ARIMA	MAPE: 6.7%	Diario
Peláez-Rodríguez et al. (2024)	LR	RT, RF, SVR, ELM, FCDNN	MAE	FCDNN: Pamplona LR: Madrid	MAE: 8.7 MAE: 8.6	Diario
Tuominen et al. (2024)	LightGBM	ARIMA, TFT, N-BEATS, RNN	MAE, RMSE	LightGBM,	MAE: 6.63	Diario

AB: AdaBoost, ARIMA: media móvil autorregresiva integrada, ARIMAX: regresión con errores ARIMA, BiLSTM: memoria bidireccional a corto plazo, CNN: Red neuronal convolucional, ConvLSTM: memoria convolucional a largo plazo, DeepAR: red neuronal recurrente autorregresiva, DSA: Arquitectura apilada profunda, DT: Árbol de decisión, ELM: Máquinas de Aprendizaje Extremo (Extreme Learning Machines), FCDNN: Redes Neuronales Profundas Completamente Conectadas (Fully Connected Deep Neural Networks), GBM: máquinas de gradiente incrementado, GEE: modelos de ecuaciones de estimación generalizada, GLM: Regresión lineal generalizada, GRU: Unidad Recurrente Controlada (Gated Recurrent Unit), HW: Holt-Winters, KNN: Vecinos más cercanos K, LMS: mínimos cuadrados medios, LR: regresión logística, LSTM: memoria a corto plazo larga, MLP: perceptrón multicapa, MLR: regresión lineal múltiple, MSARIMA: autorregresivo integrado de media móvil multivariado, N-BEATS: Análisis de Expansión de Bases Neuronales, RBM: máquina de Boltzmann restringida, RF: Bosques Aleatorios, RLS: mínimos cuadrados recursivos, RNN: Red neuronal recurrente, RT: Árboles de Regresión, SARIMA: modelo estacional autorregresivo integrado de media móvil, SARIMAX: Media móvil estacional autorregresiva integrada con regresores exógenos, SG: regresión por gradiente estocástico, SMHW: Holt-Winters multiplicativo estacional, SS: suavizado exponencial simple, STL: descomposición estacional de series temporales mediante LOESS, SVM: Máquina de vectores soporte, SVR: Máquinas de Vectores de Soporte para Regresión (Support Vector Machines for Regression), TFT: Transformador de Fusión Temporal, XGB: XGBoost

Fuente: Elaboración propia

Siguiendo la clasificación propuesta por Jiang et al. (2023) la cual está basada en los métodos de modelado de las llegadas de pacientes a los servicios de urgencias, se identifican tres clases principales de modelos: regresión, series temporales, inteligencia artificial basada en redes neuronales; y adicionalmente incluimos ensambles.

**Modelos de regresión.** Los modelos basados en la regresión suponen que los resultados previstos pueden explicarse por la interacción de las variables predictoras. La regresión es un método estadístico para estimar las relaciones causa-efecto entre una variable dependiente y una o más variables independientes. Los modelos de regresión que utilizaron en los estudios se describen brevemente en la tabla 5.

**Tabla 5**

*Modelos de regresión empleados en los estudios incluidos en la revisión.*

Modelo	Descripción breve
ARIMAX (Media Móvil Autorregresiva Integrada con Variables Exógenas)	se trata de una extensión del modelo ARIMA al que se le incorporan variables exógenas para mejorar la precisión de la predicción en series temporales (Huang & Petukhina, 2022).
DT (Árbol de Decisión)	es un modelo predictivo que utiliza una estructura de árbol para representar y clasificar las posibles decisiones y sus consecuencias (C. James et al., 2023).
GEE (Modelos de Ecuaciones de Estimación Generalizada)	son una extensión de los modelos lineales generalizados que permiten modelar datos correlacionados y heterocedásticos (Hardin & Hilbe, 2013).
GLM (Regresión Lineal Generalizada)	es un marco estadístico que generaliza la regresión lineal clásica para acomodar diferentes distribuciones de error y funciones de enlace (James et al., 2021; Kuhn & Johnson, 2013).
KNN (Vecinos más Cercanos K)	es un algoritmo de aprendizaje supervisado que clasifica los datos basándose en la similitud con los puntos de datos cercanos en el espacio de características (Kuhn & Johnson, 2013).
LMS (Mínimos Cuadrados Medios)	es un método de estimación de parámetros que minimiza la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos (Montgomery et al., 2012)
MLR (Regresión Lineal Múltiple)	es una extensión de la regresión lineal simple que modela la relación entre una variable dependiente y dos o más variables independientes (James et al., 2021; Kuhn & Johnson, 2013).
RIDGE (Regresión de Ridge)	es un método de regularización que agrega una penalización a los coeficientes para reducir la varianza del modelo y evitar el sobreajuste (Kuhn & Johnson, 2013).
RLS (Mínimos Cuadrados Recursivos)	es un algoritmo de estimación de parámetros que utiliza una técnica recursiva para actualizar continuamente los coeficientes del modelo (Kuhn & Johnson, 2013).
RT (Árboles de Regresión)	son una variante de los árboles de decisión que se utilizan para modelar relaciones no lineales entre las variables de entrada y de salida (James et al., 2021; Kuhn & Johnson, 2013).
SG (Regresión por Gradiente Estocástico)	es un algoritmo de optimización que utiliza el gradiente descendente para ajustar los parámetros del modelo de manera iterativa, utilizando solo una muestra de los datos en cada paso (Bottou, 2010).
SVM (Máquina de Vectores de Soporte)	es un método de aprendizaje supervisado utilizado para la clasificación y regresión que encuentra el hiperplano óptimo que maximiza el margen entre las clases (James et al., 2021; Kuhn & Johnson, 2013).
SVR (Máquinas de Vectores de Soporte para Regresión)	es una variante de las SVM que se utiliza para problemas de regresión en los que se busca encontrar una función que se ajuste a los datos con la mayor distancia posible a los puntos de entrenamiento (James et al., 2021; Kuhn & Johnson, 2013).

Fuente: Elaboración propia

Dos de los artículos revisados utilizaron métodos de regresión como el modelo principal para predecir el volumen de atenciones en urgencias; otros autores usaron modelos de regresión para comparar su resultado con el del modelo principal del estudio. Marcilio et al. (2013) usaron un modelo de regresión lineal generalizada como una clase de método de regresión de series temporales y lo compararon con los resultados de otros modelos de regresión de series temporales, el mejor desempeño predictivo lo obtuvieron con el modelo de regresión, el cual logro un MAPE de 11.5% en la predicción con alcance diario. Erkamp et al. (2021) utilizó un modelo de predicción lineal basado en variables de calendario y clima para realizar las predicciones, el cual calibro con el uso de diferentes conjuntos de predictores; en este caso el modelo no se comparó con otras técnicas de predicción, sino que se implementaron diversos set de datos para comparar con cuales variables obtendría el mejor resultado, este grupo logro obtener un poder predictivo con un MAPE de 8.6% utilizando solo variables asociadas al calendario. Peláez-Rodríguez et al. (2024) compararon múltiples modelos de regresión con una red neuronal completamente conectada (FCDNN) en dos servicios de urgencias de España (Madrid y Pamplona), en el servicio de urgencias de Madrid el mejor desempeño fue logrado con el modelo de regresión general (MAE = 8.6), mientras que en Pamplona el mejor resultado lo obtuvo la FCDNN (MAE=8.7).

**Modelos de series temporales.** El análisis de series temporales constituye un método muy importante en la predicción de las visitas a urgencias. Una serie temporal es una colección de medidas obtenidas mediante observaciones repetidas a lo largo del tiempo. El análisis de series temporales suele tiene en cuenta que los datos a lo largo del tiempo pueden mostrar ciertas tendencias o patrones internos, por lo que las predicciones futuras están basadas en los valores históricos y en la incorporación de estas tendencias en el modelo (Jiang et al., 2023). En la presente revisión se identifico el

uso de ARIMA, HW, MSARIMA, Prophet, SARIMA, SARIMAX, SMHW, SS, STLM, TBATS, TFT; como modelos de series temporales para realizar las predicciones, estos modelos se describen en la tabla 6.

**Tabla 6**

*Modelos de series temporales empleados en los estudios incluidos en la revisión.*

Modelo	Descripción breve
ARIMA (Media Móvil Autorregresiva Integrada)	es un modelo estadístico ampliamente utilizado en el análisis y predicción de series temporales. Combina componentes de media móvil, autorregresión e integración para modelar la estructura de autocorrelación y estacionalidad en los datos temporales (Box, 2016).
HW (Holt-Winters)	es un método de suavizado exponencial triple utilizado para pronosticar datos de series temporales con tendencia y estacionalidad. Incorpora componentes para modelar la tendencia, la estacionalidad y el nivel base de la serie temporal (Hyndman & Athanasopoulos, 2021).
MSARIMA (Autorregresivo Integrado de Media Móvil Multivariado)	es una extensión del modelo ARIMA para series temporales multivariadas. Permite modelar la autocorrelación y la estacionalidad en múltiples variables simultáneamente (Brockwell & Davis, 2016).
Prophet	es una herramienta de pronóstico de código abierto desarrollada por Facebook que se basa en un modelo aditivo para pronosticar datos de series temporales con tendencia, estacionalidad y días festivos (Taylor & Letham, 2018).
SARIMA (Modelo Estacional Autorregresivo Integrado de Media Móvil)	es una extensión del modelo ARIMA que incorpora términos estacionales para modelar series temporales con estacionalidad. Es útil para datos con patrones estacionales bien definidos (Healy, 1976).
SARIMAX (Media Móvil Estacional Autorregresiva Integrada con Regresores Exógenos)	es una extensión de SARIMA que permite incorporar variables exógenas en el modelo para mejorar la precisión de la predicción. Es útil cuando hay factores externos que afectan la serie temporal (Hyndman & Athanasopoulos, 2021).
SMHW (Holt-Winters Multiplicativo Estacional)	es una variante del método de suavizado exponencial triple que asume que la estacionalidad en los datos varía de forma multiplicativa en lugar de aditiva. Es útil cuando la magnitud de la estacionalidad varía con el nivel de la serie temporal (Hyndman & Athanasopoulos, 2021).
SS (Suavizado Exponencial Simple)	es un método de pronóstico simple que asigna pesos exponenciales decrecientes a las observaciones pasadas para pronosticar el siguiente punto en la serie temporal (Hyndman & Athanasopoulos, 2021).
STLM (Descomposición Estacional de Series Temporales mediante LOESS)	es un enfoque que utiliza la técnica de suavizado LOESS para descomponer una serie temporal en sus componentes de tendencia, estacionalidad y residuos (Cleveland et al., 1990).
TBATS (Especificación del Modelo de Series Temporales Box-Tiao)	es un método de pronóstico automático para datos de series temporales con múltiples estacionalidades y patrones de tendencia complejos. Es capaz de manejar series temporales irregulares y datos faltantes (De Livera et al., 2010).
TFT (Transformador de Fusión Temporal)	es un enfoque de pronóstico de series temporales que utiliza un modelo de transformador para capturar patrones temporales en los datos. Es capaz de manejar series temporales con múltiples frecuencias y patrones irregulares (Lim et al., 2020).
HW (Holt-Winters)	es un método de suavizado exponencial triple utilizado para pronosticar datos de series temporales con tendencia y estacionalidad. Incorpora componentes para modelar la tendencia, la estacionalidad y el nivel base de la serie temporal (Hyndman & Athanasopoulos, 2021).

Fuente: elaboración propia

Los modelos basados en series de tiempo constituyen uno de los métodos más frecuentemente utilizados para predecir la demanda de urgencias, en total nueve estudios utilizaron estos métodos como modelo principal; ARIMA fue el más frecuente, seis estudios lo usaron como método principal; SARIMAX se usó en dos publicaciones y SARIMA en una de ellas. En general todos los grupos compararon el desempeño del modelo principal con otros métodos de serie temporal o regresión lineal, excepto Choudhury & Urena (2020) quienes adicionalmente lo compararon con un modelo de redes neuronales.

En la comparación realizada por Vollmer et al. (2021) de modelos de serie temporal con modelos de regresión y ensambles basados en árboles de decisión el modelo que obtuvo el mejor resultado predictivo fue un modelo de regresión lineal que uso como predictores la demanda de días anteriores, indicadores para días festivos, datos meteorológicos e información de búsquedas específicas en internet; este modelo logró un MAPE de 6.8%, y fue el único estudio en el que un modelo de regresión fue superior a los modelos de series de tiempo.

Los modelos ARIMA obtuvieron el mejor desempeño en tres de los estudios, Murtas et al. (2022) reportó una MAPE de 8.1%; en el caso del estudio realizado Vieira et al. (2023) el MAPE del modelo fue 6.7% y fue superior a ARIMAX, SARIMA y TBATS; por último, Choudhury & Urena (2020) reportaron un RMSE de 1.55 para el modelo de ARIMA implementado, siendo este el mejor desempeño en comparación con los modelos de RNN, TBATS y HW con los que se comparó.

Los modelos de series de tiempo que permiten integrar variables exógenas a la serie de tiempo como SARIMAX y ARIMAX fueron destacados en cuatro estudios como los de mejor desempeño predictivo al integrar variables relacionadas con el clima, festividades y el calendario, obteniendo desempeños entre 5.9% y 8.7% en la métrica de desempeño

MAPE (Carvalho-Silva et al., 2018; Hu et al., 2023; Tuominen et al., 2022; Whitt & Zhang, 2019). Por otro lado, Calegari et al. (2016) reportaron un mejor rendimiento predictivo del modelo SARIMA con un MAPE de 6.23%

**Modelos de redes neuronales.** La tabla 7 presenta una descripción de los modelos de redes neuronales encontrados. En cinco de las publicaciones incluidas se usaron modelos de redes neuronales como algoritmo principal de predicción, comparando su capacidad predictiva con el resultado obtenido en otras redes neuronales (Harrou et al., 2020; Zhao et al., 2022), modelos de regresión o ensambles (Nas & Koyuncu, 2019; Yousefi et al., 2020; Zhang et al., 2022), y series de tiempo (Yousefi et al., 2020; Zhao et al., 2022).

**Tabla 7**

*Modelos de redes neuronales empleados en los estudios incluidos en la revisión.*

Modelo	Descripción breve
RNN (Red Neuronal Recurrente)	es un tipo de red neuronal diseñada para procesar datos secuenciales, donde la salida de cada capa se utiliza como entrada para la siguiente iteración. Son útiles para modelar dependencias temporales en datos de series temporales, texto, voz, entre otros (Xiao & Sun, 2021).
CNN (Red Neuronal Convolutacional)	es un tipo de red neuronal diseñada especialmente para procesar datos con estructura de cuadrícula, como imágenes. Utiliza capas de convolución para extraer características y capas de pooling para reducir la dimensionalidad (Fergus & Chalmers, 2022).
LSTM (Memoria a Corto Plazo Larga)	es un tipo de red neuronal recurrente (RNN) diseñada para modelar dependencias a largo plazo en secuencias. Utiliza unidades de memoria especiales para retener y actualizar información relevante a lo largo del tiempo (Fergus & Chalmers, 2022).
BiLSTM (Memoria a Corto Plazo Bidireccional)	es una variante de las redes neuronales recurrentes (RNN) que permite capturar la información de contexto tanto hacia adelante como hacia atrás en una secuencia (Laghari et al., 2023).
ConvLSTM (Memoria Convolutacional a Largo Plazo)	es una variante de las LSTM que integra operaciones de convolución en la estructura de la red, permitiendo aprender patrones espaciales y temporales en secuencias de datos (Shi et al., 2022).
DeepAR	es un modelo de pronóstico de series temporales basado en redes neuronales recurrentes (RNN) que puede manejar múltiples series de tiempo y capturar relaciones temporales complejas (Salinas et al., 2019).
ELM (Máquinas de Aprendizaje Extremo)	es un método de aprendizaje automático que utiliza una sola capa oculta de neuronas para aproximar una función de entrada-salida. Se caracteriza por su rapidez de entrenamiento y buen rendimiento (Castro & Jojoa, 2019).
CDNN (Redes Neuronales Profundas Completamente Conectadas)	se refiere a redes neuronales artificiales con múltiples capas ocultas donde todas las neuronas en una capa están conectadas a todas las neuronas en la capa siguiente. Son ampliamente utilizadas en tareas de clasificación y regresión (Goodfellow et al., 2016).

**Tabla 7**

*Modelos de redes neuronales empleados en los estudios incluidos en la revisión.*

Modelo	Descripción breve
GRU (Unidad Recurrente Controlada)	es una variante de las redes neuronales recurrentes (RNN) diseñada para abordar el problema de la desaparición y explosión del gradiente. Tiene una arquitectura más simple que LSTM pero aún es efectiva para modelar dependencias a largo plazo en secuencias (Fergus & Chalmers, 2022).
MLP (Perceptrón Multicapa)	es una arquitectura de red neuronal artificial donde múltiples capas de neuronas están conectadas entre sí en una estructura de alimentación hacia adelante. Es capaz de aprender representaciones no lineales de datos (Fergus & Chalmers, 2022).
N-BEATS (Análisis de Expansión de Bases Neuronales)	es un modelo de redes neuronales diseñado específicamente para pronósticos de series temporales. Utiliza bloques de redes neuronales para capturar patrones temporales en diferentes escalas de tiempo (Oreshkin et al., 2020).
RBM (Máquina de Boltzmann Restringida)	es un tipo de red neuronal no supervisada que utiliza una arquitectura probabilística para aprender representaciones eficientes de datos. Se compone de una capa visible y una capa oculta conectadas de forma restringida (Guido Montúfar, 2018).
VAE (Codificador Variacional Autoestable)	es un tipo de red neuronal generativa que utiliza técnicas de aprendizaje profundo para aprender y generar datos de alta dimensionalidad. Se basa en la teoría de inferencia variacional para aprender la distribución de probabilidad latente de los datos de entrada (Fergus & Chalmers, 2022).
DSA (Deep Stacked Architecture)	es un modelo de aprendizaje profundo que se basa en una estructura apilada de múltiples capas neuronales. Este enfoque permite aprender representaciones de datos cada vez más abstractas y complejas a medida que se profundiza en la red (LeCun et al., 2015).

Fuente: elaboración propia

LSTM resultó ser el diseño de red neuronal que mejor predijo la demanda en urgencias. En el estudio realizado en Brasil, en el Hospital de Risoleta Tolento de Belo Horizonte, Neves Yousefi et al. (2020) reporta que el algoritmo implementado de LSTM logro un MAPE de 4.89% en la predicción de la demanda diaria del servicio de urgencias. Para Nas & Koyuncu (2019) LSTM también fue el mejor modelo para predecir la demanda diaria obteniendo el mejor MAPE en comparación con los modelos de DT, SVM, GBM, AB, SG, KNN, RF, MLP, LR; no obstante RF logro el mejor desempeño en la predicción de la demanda horaria. Estos datos difieren de los publicados por Zhang et al. (2022) que reporta que LSTM fue el mejor modelo para predecir la demanda horaria, mientras que SVR logro el mejor desempeño en la predicción de la demanda diaria; en este estudio la capacidad predictiva de LSTM se comparó con GLM, KNN, SVR, RIDGE, XGB, AB, GBM, RF.

El modelo de aprendizaje profundo basado en una estructura apilada de múltiples capas neuronales (LSTM, GRU y RNN) utilizado por Zhao et al. (2022) para prever la demanda de atenciones en urgencias en el Hospital General de Singapore obtuvo el mejor desempeño predictivo en comparación con los algoritmos de ARIMA, Prophet, ConvLSTM, BiLSTM. La red neuronal generativa (VAE) desarrollada por Harrou et al. (2020) en el Centro Hospitalario Regional de Lille fue el algoritmo de mejor desempeño predictivo en comparación con RNN, LSTM, BiLSTM, ConvLSTM, GRUs, RBM, CNN.

**Modelos de ensamble.** Un ensamble es una técnica que combina múltiples modelos de predicción para mejorar la precisión y generalización de los modelos de aprendizaje automático, y son críticos en la construcción de sistemas de alta calidad en el campo del aprendizaje automático. Un ensamble puede estar hecho de diferentes tipos de modelos como los árboles de decisión, modelos lineales, redes neuronales, etc. Algunos de los ensambles populares incluyen Random Forests, y aquellos basados en bagging, y boosting (Kuhn & Johnson, 2013).

Petsis et al. (2022) utilizan como modelo principal un algoritmo de XGBoost con el que logran una predicción con un MAPE de 6.5%. Tuominen et al. (2024) y Gafni-Pappas & Khan (2023) utilizaron modelos basados en gradient boosting machine. Mientras que Rocha & Rodrigues (2021) realizaron un ensamble que combinó las mejores predicciones obtenidas con los modelos de SARIMA, RNN-1L, RNN-3L y XGBoost. Este último caso consigue el mejor desempeño de las predicciones horarias en comparación con los otros estudios que lo intentaron, alcanzando un MAPE de 23.3% en comparación con el 50% reportado por Zhang et al. (2022) y Nas & Koyuncu (2019) quienes usaron técnicas de redes neuronales.

### **Predictores seleccionados**

Para facilitar su presentación las variables utilizadas como predictores en los estudios seleccionados se clasificaron en tres grupos: calendario, meteorológicas y otras. En el grupo de calendario se incluyen aquellas asociadas con la identificación del día de la semana, mes del año, festivos, así como los rezagos de volumen. En meteorológicas se agrupan aquellas relacionadas con el clima, la temperatura, humedad, entre otras. Y en el tercer grupo otras: variables misceláneas como clasificación de triage, búsquedas en internet, vacaciones, eventos masivos, entre otros. En la tabla 8 se muestran los covariables o predictores utilizadas en cada uno de los estudios.

Todos los estudios utilizaron variables asociadas al calendario, siendo la variable más frecuente el día de la semana; solo en el estudio realizado por Yousefi et al. (2020) no fue utilizada, en este análisis se usó únicamente una variable para identificar si el día era un fin de semana o no.

Todos los autores coincidieron en que las variables que mayor poder predictivo agregan a sus modelos fueron las variables relacionadas con el calendario. En el análisis de Gafni-Pappas & Khan (223) quienes evaluaron la importancia de las variables el día de la semana fue la variable de mayor importancia en el modelo predictivo, coincidiendo con lo reportado por Hu et al. (223), y con la evaluación de impacto realizado por Petsis et al. (222); en estos estudios el día de la semana tuvo una mayor relevancia que el mes del año en la predicción; no obstante, para Tuominen et al. (222) el día de la semana y el mes del año fueron igualmente significativos en el resultado de la predicción.

**Tabla 8 Predictores utilizados en los estudios analizados en la revisión**

*Predictores utilizados en los estudios analizados en la revisión.*

Publicación	Calendario			Meteorológicos							Otros							
	Día de la semana	Mes del año	Estación del año	Festivo	Antes o después de un festivo	Fin de semana	Rezagos	Temperatura	Humedad	Precipitaciones	Condición del aire	Clima	Otros	Clasificación de triage	Búsquedas específicas Google	Temporada de gripas	Eventos masivos (Conciertos, partidos)	Vacaciones de verano
Marcilio et al. (213)	✓			✓	✓			✓								✓		
Calegari et al. (216)	✓	✓						✓	✓	✓	✓							
Carvalho-Silva (218)	✓	✓						✓		✓								
Nas & Koyuncu (219)	✓		✓	✓														
Whitt & Zhang (219)	✓	✓		✓	✓			✓		✓								
Harrou et al. (22)	✓	✓																
Yousefi et al. (22)				✓	✓	✓	✓											✓
Choudhury & Urena (2020)	✓													✓				
Erkamp et al. (2021)	✓	✓	✓	✓				✓		✓							✓	✓
Vollmer et al. (2021)	✓	✓		✓			✓	✓		✓					✓		✓	
Rocha & Rodrigues (2021)	✓	✓				✓		✓	✓									
Zhang et al. (2022)	✓	✓	✓	✓				✓			✓	✓						
Tuominen et al. (2022)	✓	✓						✓	✓	✓			✓		✓		✓	
Murtas et al. (2022)	✓		✓	✓	✓			✓	✓	✓			✓					
Zhao et al. (2022)	✓			✓			✓	✓		✓		✓						
Petsis et al. (2022)	✓	✓		✓	✓	✓	✓	✓		✓	✓							✓
Hu et al. (2023)	✓	✓	✓	✓	✓			✓		✓			✓	✓	✓			

**Tabla 8 Predictores utilizados en los estudios analizados en la revisión**

*Predictores utilizados en los estudios analizados en la revisión.*

Publicación	Calendario			Meteorológicos							Otros							
	Día de la semana	Mes del año	Estación del año	Festivo	Antes o después de un festivo	Fin de semana	Rezagos	Temperatura	Humedad	Precipitaciones	Condición del aire	Clima	Otros	Clasificación de triage	Búsquedas específicas Google	Temporada de gripas	Eventos masivos (Conciertos, partidos)	Vacaciones de verano
Gafni-Pappas & Khan (2023)	✓	✓					✓	✓	✓			✓						
Vieira et al. (2023)	✓	✓		✓				✓								✓		
Peláez-Rodríguez et al. (2024)	✓			✓			✓	✓		✓			✓					
Tuominen et al. (2024)	✓	✓		✓	✓			✓		✓	✓				✓			

Fuente: elaboración propia

Los rezagos, que representan la historia de las visitas a urgencias, se consideraron como variables importantes para capturar patrones históricos y tendencias en la demanda de servicios de urgencias en cinco de los estudios analizados. Estos rezagos le permitieron a Peláez-Rodríguez et al. (224) y Petsis et al. (222) ajustar los modelos de regresión desarrollados mejorando su capacidad predictiva, capturar patrones temporales y tendencias en la demanda de atención de urgencias. Zhao et al. (222) pudo identificar que los rezagos más importantes fueron los relacionados con la demanda de atenciones de la semana anterior.

Al menos una de las variables meteorológicas fue incluida en 17 de las 21 publicaciones analizadas en esta revisión. De acuerdo, con lo reportado por los diferentes autores, la forma en que estas variables pueden contribuir a mejorar el desempeño del modelo podría considerarse controversial.

Para Marcilio et al. (2013) los modelos que incluían información sobre el clima no mejoraron la precisión de los pronósticos y consideran que estos modelos podían

introducir inexactitudes adicionales debido a la dificultad de predecir con precisión las condiciones meteorológicas a largo plazo. Esto coincide el análisis realizado por Calegari et al. (2016) quienes también reportan que la inclusión de variables meteorológicas no aportó beneficios significativos a la capacidad predictiva del modelo SARIMA en el contexto específico de pronóstico de demanda en urgencias.

De acuerdo, con las pruebas de correlación de Pearson realizados por Carvalho-Silva (2018) para evaluar la fuerza y dirección de la relación entre las llegadas de pacientes y las variables ambientales, no se encontró correlación significativa de la temperatura con las llegadas de los pacientes, y en el caso de la lluvia se demostró una correlación débil y negativa con las llegadas de pacientes, lo que indicaba que las llegadas disminuían con un aumento en los niveles de precipitación.

El efecto de estas variables en la capacidad predictiva del modelo implementado por Erkamp et al. (2021) fue limitado; de igual forma, aunque se esperaba que las variables meteorológicas tuvieran un impacto en las visitas diarias, los resultados obtenidos por Zhao et al. (2022) mostraron que su inclusión no condujo a una mejora significativa en la precisión de las predicciones, coincidiendo esta conclusión con lo reportado por Vieira et al. (2023).

De otro lado en el estudio realizado por Whitt & Zhang (2019) el incremento de la temperatura tuvo una relación directa y positiva con el incremento de las llegadas diarias a urgencias, muy similar al efecto reportado por Vollmer et al. (2021) y Tuominen et al. (2022); y a la mención que hace Zhang et al. (2022) en relación a que las variables meteorológicas, como la temperatura, la velocidad del viento, la lluvia, la calidad del aire, entre otras, pueden afectar de manera importante la cantidad de pacientes que consultan urgencias, y que la incorporación de información meteorológica en el modelo de pronóstico puede mejorar su capacidad predictiva.

Hu et al. (2023) observaron que ciertas condiciones climáticas como la presencia de nieve, precipitación y temperaturas extremadamente frías o calientes tendían a reducir la demanda de atenciones, mientras que días con vientos fuertes estaban asociados con un incremento en la solicitud de atenciones, por lo cual su inclusión tenía un efecto significativo en la capacidad predictiva del modelo. Los estudios de Gafni-Pappas & Khan (2023), Vieira et al. (2023), Peláez-Rodríguez et al. (2024) y Tuominen et al. (2024) también reportan un efecto favorable en la capacidad predictiva de sus modelos al usar predictores relacionados con las condiciones meteorológicas.

En el grupo de predictores “otros” se incluyeron las variables que fueron usadas de forma menos frecuente en los estudios, en este grupo se destaca la inclusión de predictores como: el número de búsquedas de términos específicos en internet, la presencia de eventos masivos en el área de influencia de la institución y la temporada de gripa.

La inclusión de una variable relacionada con el número de búsqueda de términos relacionados con síntomas de enfermedades o enfermedades estacionales en Google ("flu", gripe; "emergency room", sala de emergencias; "abdominal pain", dolor abdominal; "respiratory infection", infección respiratoria; "chest pain", dolor de pecho; "depression", depresión; "heart attack", ataque al corazón; "abuse", abuso; "disorder", trastorno; "weather", clima; "hospital", hospital) en los modelos de pronóstico demostró que estas búsquedas están positivamente relacionadas con la demanda de atenciones en urgencias, y tuvo un impacto positivo en su capacidad predictiva, particularmente para las predicciones a corto plazo, horizonte t+1 (Hu et al., 2023; Tuominen et al., 2024; Vollmer et al., 2021).

La inclusión de la celebración de eventos masivos en el área de influencia del hospital como variable predictora demostró que esta variable puede contribuir a ajustar mejor la

precisión del pronóstico, dado que este tipo de eventos pueden tener un impacto positivo o negativo en la llegada de pacientes a urgencias. El estudio de Yousefi et al. (2020), realizado en Brasil, evidenció que la presencia de partidos de fútbol en la ciudad afectaba la afluencia de pacientes disminuyendo notablemente el número de visitas a urgencias durante estos eventos. Mientras que los carnavales en el área de influencia del Hospital Jeroen Bosch en Países Bajos, y del Hospitales St Mary's y Charing Crossen Londres incrementaron la llegada de personas a urgencias (Erkamp et al., 2021; Vollmer et al., 2021). No se descarta que la importancia del efecto de estos eventos en las predicciones tenga relación con su interacción con la estacionalidad del modelo predictivo (Tuominen et al., 2022).

Por último, la inclusión de un predictor que da cuenta de los eventos endémicos o epidémicos relacionados con gripe y en general enfermedades respiratorias permitió que los modelos capturaran y ajustaran la variabilidad adicional en las visitas diarias a urgencias al tener en cuenta el pico anormal de visitas provocado por esta situación (Marcilio et al., 2013; Vieira et al., 2023).

### **Desempeño predictivo**

Las métricas de desempeño más comúnmente utilizadas en los modelos de pronóstico para la demanda de atenciones en urgencias fueron: RMSE (Root Mean Squared Error) – Raíz del error cuadrático medio, MSE (Mean Squared Error) - Error cuadrático medio, MAE (Mean Absolute Error) - Error absoluto medio, MAPE (Mean Absolute Percentage Error) - Error absoluto porcentual medio.

MSE y RMSE tienen una mayor sensibilidad a los errores grandes debido al cuadrado de las diferencias, por lo que penalizan más los errores grandes en comparación con los pequeños, a diferencia de MSE, RMSE utiliza la misma unidad de medida de la variable que se está prediciendo. En comparación con MSE y RMSE, MAE proporciona una idea

clara del error promedio que se puede esperar y es menos sensible a los valores atípicos. Por otro lado, MAPE entrega una perspectiva clara sobre la magnitud del error relativo al valor real lo que hace que esta métrica sea muy útil para interpretar los errores en términos relativos, permitiendo comparar la precisión de los modelos en diferentes escalas (Gareth M. James et al., 2021; Kuhn & Johnson, 2013).

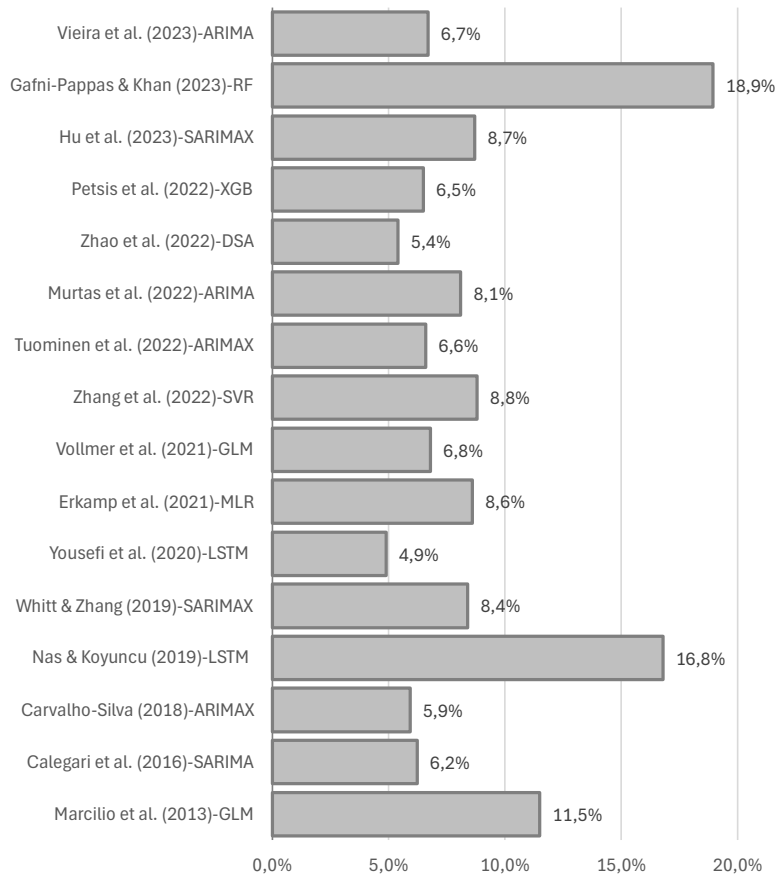
La métrica de desempeño que se utilizó con mayor frecuencia fue el error absoluto porcentual medio (MAPE). En las publicaciones realizados por Harrou et al. (2020), Choudhury & Urena (2020), Rocha & Rodrigues (2021), Peláez-Rodríguez et al. (2024), Tuominen et al. (2024) se prefirió el uso del error absoluto medio y la raíz del error cuadrático medio RMSE.

El resultado del MAPE de los 16 modelos en los que el alcance de la predicción fue diario y que reportaron esta métrica de desempeño se presenta en la figura 5. Trece de ellos lograron un error porcentual medio menor al 10% lo que los hace muy útiles para la toma de decisiones.

Se pudo observar que la precisión de los pronósticos se deteriora a la medida que se incrementa el horizonte de tiempo de la predicción, es decir que la mayor precisión se consigue en el día 1, y disminuye en la medida que el pronóstico se extiende a 7, 15 o 30 días. Este fenómeno fue reportado por Marcilio et al. (2013), Calegari et al. (2016), Nas & Koyuncu (2019), Yousefi et al. (2020) y Erkamp et al. (2021) quienes probaron diferentes horizontes observando que el MAPE del modelo se incrementó conforme se cambió el alcance del horizonte de tiempo de la predicción. No obstante, los modelos continuaron obteniendo desempeños destacados, con un horizonte de 30 días el modelo de Calegari et al. (2016) obtuvo un MAPE del 16%, mientras que el del modelo de Nas & Koyuncu (2019) fue del 12%; mientras que el mejor modelo de Yousefi et al. (2020) obtuvo un MAPE de 5.5% al cambiar el horizonte de tiempo de un día a una semana.

**Figura 5**

*Error absoluto porcentual medio (MAPE) de los modelos evaluados con horizonte de tiempo de un día.*



Fuente: elaboración propia

En contraste a los pronósticos diarios, los modelos pronósticos horarios tuvieron errores porcentuales altos. El mejor modelo logrado por Nas & Koyuncu (2019) para el pronóstico horario fue un RF que tuvo un MAPE de 49.67%, similar al LSTM de Zhang et al. (2022), ampliamente superados por el ensamble realizado por Rocha & Rodrigues (2021) que reporto un MAPE de 23.3%. Estos hallazgos contrastan con los de Harrou et al. (2020), en este estudio los modelos de redes neuronales obtuvieron un mejor desempeño en las predicciones por hora vs las predicciones diarias; el modelo VAE ofreció una mayor precisión predictiva horaria, en comparación con los otros modelos

probados en el estudio, obteniendo mejores resultados en las métricas empleadas: R2, RMSE y MAE. De acuerdo con los autores este resultado puede deberse a que los modelos de aprendizaje profundo muestran una mayor efectividad en la identificación y captura de patrones a corto plazo en los datos, tal como ocurre en los servicios de urgencias con las fluctuaciones horarias de la demanda, y por otro lado estos modelos pueden tener dificultades para capturar la influencia de los eventos estacionales, de las variables de calendario y meteorológicas en las predicciones diarias.

Un factor no menos importante en la precisión del modelo es el rango de los datos históricos que se utilizan en la predicción. Petsis et al. (2022) probaron alcances históricos diferentes del set de datos de entrenamiento: 6 meses, 1, 3 y 5 años; el dataset de seis meses permitió a los modelos de apilamiento profundo alcanzar una alta precisión en las predicciones, con una precisión media del 94.28% al 94.59%.

#### **Limitaciones reconocidas en el desarrollo de los estudios**

En los modelos que incluyeron covariables meteorológicas se reconoce que estos predictores conllevan un mayor grado de incertidumbre lo que puede afectar su tasa de error (Carvalho-Silva et al., 2018; Erkamp et al., 2021). Algunos autores coinciden en que la aplicabilidad de sus modelos está limitada a la institución en la que se aplicaron y que para poder generalizarlos se requiere aplicar los modelos en diferentes servicios de urgencias, regiones y entornos con contextos y características climáticas diferentes (Calegari et al., 2016; Hu et al., 2023; Petsis et al., 2022; Vollmer et al., 2021; Yousefi et al., 2020; Zhang et al., 2022).

Otras limitaciones mencionadas corresponden a las dificultades relacionadas con la disponibilidad de información, la capacidad para establecer relaciones entre las covariables utilizadas y el volumen de atenciones, y la selección correcta el horizonte de tiempo histórico que resulta más conveniente para entrenar los modelos.

### **Aplicabilidad de los modelos predictivos de demanda de atenciones, en las operaciones del servicio de urgencias**

Los estudios incluidos en la revisión reconocen dos aplicaciones prácticas del desarrollo de los pronósticos de las llegadas diarias a urgencias. La primera de ellas hace referencia a la planificación y asignación de recursos particularmente del recurso humano, permitiendo asignar más o menos personal y ajustar sus horarios, planificar mejor los inventarios de suministros médicos y medicamentos, y prever la asignación de otros recursos significativos de forma más efectiva. La segunda, es subsecuente a esta, y está relacionada con la gestión de calidad del servicio permitiendo reducir los tiempos de espera e incrementar la satisfacción de los pacientes. Aunque estas aplicaciones son ampliamente reconocidas por los autores, solo tres de ellos mencionan como el hospital uso los pronósticos para desarrollar una aplicación práctica de los mismos.

Utilizando como variables de entrada el resultado de las predicciones de las llegadas a urgencias Nas & Koyuncu (2019) mencionan haber desarrollado un modelo de simulación con el objetivo de minimizar el tiempo de espera de los pacientes y maximizar su satisfacción a partir de la asignación eficiente de recurso humano y camas. Por su parte el equipo de Murtas et al. (2022) desarrollaron un sistema de alerta temprana que permitió a los hospitales ajustar la planificación de sus recursos y contribuyó a una atención más eficiente y efectiva de los pacientes.

Petsis et al. (2022) reconocen que el staff administrativo y médico utilizó las predicciones para realizar una planificación de corto y mediano plazo del personal del servicio de urgencias del hospital, lo que les ayudó a ajustar los horarios de los turnos del próximo día en función del comportamiento esperado de las llegadas de pacientes a este servicio, así como a gestionar otros recursos como camas adicionales y la liberación de espacio en otros servicios que son requeridos por los pacientes en urgencias.

### **Propuesta de solución a la problemática**

De acuerdo con los hallazgos de esta revisión es posible utilizar diversas técnicas de aprendizaje automático para predecir con gran precisión la demanda diaria de atenciones en urgencias. Sin embargo, debido a la gran influencia del comportamiento histórico de la demanda, así como de otras variables que afectan el número de llegadas a urgencias, no es posible generalizar ninguno de los modelos de pronóstico identificados en una institución local.

En consecuencia, si se quiere usar el pronóstico de la demanda como una herramienta para reorganizar los recursos y de esta forma resolver la sobreocupación constante, reducir los tiempos de atención e incrementar la satisfacción de los usuarios de los servicios de urgencias, se debe realizar un estudio local preferentemente multicéntrico que permita identificar cual es el mejor modelo de pronóstico en nuestro medio.

Esta revisión entrega algunos puntos de partida interesantes. Los modelos que ofrecen una gran precisión y menor complejidad en su construcción son aquellos basados en series de tiempo que admiten variables exógenas, como ARIMAX y SARIMAX, mientras que las redes neuronales pueden ofrecer un mejor desempeño predictivo, pero su construcción y calibración puede ser más compleja; los modelos generales de regresión que integran rezagos en su construcción también mostraron un buen desempeño; por lo tanto, el uso de las series de tiempo o modelos generales de regresión puede ser el mejor punto de inicio de una investigación.

También se determinó, en diversas instituciones y localizaciones geográficas que las características que más aportan a la predicción son las relacionadas con el calendario, porque se podría minimizar el esfuerzo de la construcción del modelo evitando el uso de otras covariables, entre ellas las meteorológicas.

Con el fin de resolver el problema de congestión, demoras e insatisfacción planteado, es importante que la investigación tenga un alcance superior al pronóstico y evalúe como este puede ser usado en la gestión de los recursos del servicio de urgencias, y el impacto que esta gestión tiene en los indicadores de operación del servicio.

## Discusión

Los modelos de pronósticos basados en técnicas de aprendizaje automático que incluyen algoritmos de regresión, series de tiempo con o sin variables exógenas, diversos tipos de redes neuronales y ensambles ha sido utilizado en diversos escenarios para predecir la demanda de atenciones de urgencias, permitiendo que los equipos administrativos de los hospitales puedan abordar la gestión de sus servicios sin la incertidumbre del volumen de llegadas, lo que ha resultado útil en la planeación del recurso humano, sus horarios y otros recursos clave del servicio.

Si bien, como se observa en la figura 5, algunos algoritmos pueden tener un mejor desempeño que otros, ninguno de ellos puede ser generalizado más allá de la institución en la que fue realizado, lo que hace que en la práctica ningún tenga un desempeño preponderantemente mejor que el otro.

Las dificultades para generalizar un modelo en particular comienzan con la dificultad para establecer e interpretar los efectos que tienen las variables exógenas sobre la llegada de pacientes a urgencias, y su posible interacción con el contexto situacional de las instituciones. Es así como las variables meteorológicas tuvieron un efecto limitado o nulo en las predicciones en los estudios de Carvalho-Silva (2018), Marcilio et al. (2013) y Erkamp et al. (2021), mientras que Hu et al. (2023), Tuominen et al. (2022), Vollmer et al. (2021), Whitt & Zhang (2019), Zhang et al. (2022) observaron que las variables meteorológicas están relacionadas con el incremento o decremento de las solicitudes de atención, y por tanto tuvieron un impacto favorable en la capacidad predictiva del modelo.

Otra de las dificultades está relacionada con la selección del conjunto de datos adecuado para entrenar los modelos. Desde el horizonte de tiempo histórico como de la combinación de covariables con las que se puede obtener la mejor predicción. Zhao et al.

(2022) entrenaron sus modelos con cuatro horizontes de tiempo: 5 años, 3 años, 1 año y 6 meses; el mejor desempeño a corto plazo lo obtuvieron usando un conjunto de datos con un horizonte histórico de 6 meses y utilizando una arquitectura compleja de redes neuronales. Utilizando algoritmos de series de tiempo y horizontes de tiempo histórico del conjunto de datos otros autores obtuvieron errores porcentuales absolutos similares a los obtenidos por Zhao et al. (2022), Tuominen et al. (2022) utilizó un conjunto de datos de 4,5 años, Carvalho-Silva (2018) de dos años, y Calegari et al. (2016) usaron un horizonte histórico de 2.5 años; de tal manera que si bien este horizonte es relevante en la precisión del modelo no es posible recomendar un horizonte mínimo o máximo para crear el modelo de predicción.

La diversidad de covariables que se usó en los estudios se aprecia en la tabla 7 y el porcentaje de error absoluto obtenido por los mejores modelos en la figura 5. Diferentes combinaciones de predictores y modelos de predicción lograron porcentajes de error absoluto útiles para gestionar el servicio de urgencias. Por lo anterior no es posible generalizar el uso de alguna característica en particular y cada institución deberá evaluar cuales son las características que mejor se ajustan a sus dinámicas de demanda de atenciones.

Otro aspecto importante para tener en cuenta es el deterioro de la capacidad predictiva de los modelos que predicen la demanda diaria, al incrementar el horizonte de tiempo, tal como lo demuestran los hallazgos de Marcilio et al. (2013), Calegari et al. (2016), Nas & Koyuncu (2019), Yousefi et al. (2020) y Erkamp et al. (2021); este deterioro deberá ser tenido en cuenta en la toma de decisiones y en los cambios en la asignación de recursos, así como en la definición de los intervalos de reentrenamiento de los modelos, lo cual debe ser una actividad permanente que dificulta tanto su implementación, como la aplicación práctica de los pronósticos.

Si bien los pronósticos diarios ofrecen una alta precisión, no ocurre lo mismo con los modelos que intentaron realizar una predicción de la demanda horaria, con porcentajes de error cercanos al 50% (Nas & Koyuncu, 2019; Zhang et al., 2022), mejorados solo por el modelo de ensamble de Rocha & Rodrigues (2021) que reportó un MAPE de 23.3%. Por lo que los ajustes de los horarios basados en estos pronósticos podrían no ser útiles, y más bien deban ajustarse al pronóstico diario combinado con el entendimiento detallado de los patrones estacionales de la demanda horaria de atenciones.

La aplicación práctica de las predicciones se enfocó en la programación de recursos del servicio de urgencias y su efecto en los tiempos de espera y la satisfacción de los pacientes. El desarrollo del sistema de alerta temprano logrado por Murtas et al. (2022) y la planificación del personal realizada a partir de las predicciones de Petsis et al. (2022) son dos ejemplos de ello. El desarrollo de aplicaciones prácticas es un tema comúnmente tratado en la introducción o en la discusión de los hallazgos de otros estudios, sin embargo, esta limitado a mencionar el potencial de los pronósticos en la gestión de la atención de urgencias; sin que las mismas sean llevadas a un escenario real.

La limitación para que las aplicaciones prácticas se prueben en escenarios reales puede estar relacionada con la poca participación de los equipos administrativos y operativos de los servicios de urgencias o al menos de los hospitales en los estudios. Solo en una tercera parte de las publicaciones incluidas, los autores están vinculados al servicio de urgencias o al hospital donde se realizó el estudio: Calegari et al. (2016), Carvalho-Silva (2018), Tuominen et al. (2022), Gafni-Pappas & Khan (2023), Vieira et al. (2023), Peláez-Rodríguez et al. (2024), Tuominen et al. (2024); no obstante, en ninguna de ellas se menciona que la aplicación práctica de los pronósticos haya sido probada en un escenario real.

En nuestro medio la obtención y posterior aplicación práctica de pronósticos deberá superar limitaciones relacionadas con la disponibilidad de información, la posible inmadurez de las capacidades analíticas y de ciencia de datos de las instituciones sanitarias, y la poca flexibilidad de los esquemas de contratación del recurso humano que dificultarían la adaptación de los horarios del personal a los requerimientos que impone la dinámica estacional de la demanda de atención en urgencias.

## Conclusiones y Trabajo Futuro

### Conclusiones

El aprendizaje automático ha demostrado ser eficaz para prever tanto la demanda diaria como horaria de atenciones en el servicio de urgencias, teniendo un mayor poder predictivo y un menor error en la predicción diaria que horaria, y particularmente una mayor precisión en las predicciones de muy corto plazo, entre  $t+1$  y  $t+7$ .

Se identificaron 45 algoritmos diferentes para prever la demanda de urgencias, incluidos modelos de regresión, series de tiempo, redes neuronales y ensambles. Los algoritmos con el mayor poder predictivo fueron 17, de ellos tres son algoritmos de regresión, cuatro series de tiempo con o sin variables exógenas, seis son redes neuronales y cuatro corresponden a ensambles.

Los predictores o covariables que con mayor frecuencia se asociaron a una mayor precisión fueron los relacionados con el calendario, incluyendo los rezagos de la demanda.

La utilidad del uso de predictores meteorológicos es controversial, en algunos casos estuvo asociada al incremento o decremento de la demanda de atenciones y mejoro la precisión de los modelos, mientras que en otro su capacidad para mejorar la precisión fue limitada.

El volumen de búsqueda de términos específicos en Google, así como la realización de eventos masivos en el área de influencia de los hospitales podría estar relacionado con una mayor o menor demanda de atenciones, lo cual depende de la naturaleza del evento y del contexto geográfico de la institución.

Con base en los hallazgos de este estudio no es posible generalizar un algoritmo y cada institución debe determinar el modelo de pronóstico que mejor se adapta a su situación particular.

La aplicación práctica de estos modelos puede contribuir a mejorar la planeación de corto plazo de los recursos, disminuir los tiempos de atención e incrementar la satisfacción de los usuarios; esta utilidad práctica de las previsiones de demanda debe ser probada en escenarios reales.

### **Trabajo futuro**

Un estudio multicéntrico con la participación de científicos de datos y de los equipos operativos y administrativos del servicio de urgencias, y cuyo alcance se extienda del desarrollo de un modelo predictivo de la demanda a la utilización de estas predicciones para modelar la asignación de recursos a urgencias incluido el recurso humano y sus horarios, camas del servicio, disponibilidad de camas en hospitalización y cuidado crítico, salas de cirugía y apoyos diagnósticos, podría contribuir a establecer si es posible generalizar un modelo de pronóstico para varias instituciones y si las previsiones contribuyen con la mejora operativa del servicio.

## Referencias

- Abdelaziz, A., Elhoseny, M., Salama, A. S., & Riad, A. M. (2018). A machine learning model for improving healthcare services on cloud computing environment. *Measurement*, 119, 117–128.
- Ahmad, M. A., Eckert, C., & Teredesai, A. (2018). *Interpretable machine learning in healthcare. In Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics.*
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 81(1), 19–32.
- Asociación Colombiana de Empresas de Medicina Integral, ACEMI. (2012). *Cifras e indicadores del sistema de salud. Cifras 2008-2011.*
- Asociación Colombiana de Empresas de Medicina Integral, ACEMI. (2018). *Gestión de desperdicios e ineficiencias en el sistema de salud colombiano.*
- Azadi, M., Yousefi, S., Farzipoor Saen, R., Shabanpour, H., & Jabeen, F. (2023). Forecasting sustainability of healthcare supply chains using deep learning and network data envelopment analysis. *Journal of Business Research*, 154, 113357. <https://doi.org/10.1016/j.jbusres.2022.113357>
- Barda, A., Horvat, C., & Hochheiser, H. (2020). A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Medical Informatics and Decision Making*, 20, 257. <https://doi.org/10.1186/s12911-020-01276-x>
- Baumeister, R., & Leary, M. (1997). Writing Narrative Literature Reviews. *Review of General Psychology*, 1(3), 311–320.

- Berger, J. S., Haskell, L., Ting, W., Lurie, F., Chang, S. C., Mueller, L. A., & Alas, V. (2020). Evaluation of machine learning methodology for the prediction of healthcare resource utilization and healthcare costs in patients with critical limb ischemia—Is preventive and personalized approach on the horizon? *EPMA Journal*, 11(1), 53–64.
- Berikol, G. B., & Berikol, G. (2020). Use of artificial intelligence in emergency medicine. En *Artificial Intelligence in Precis. Health: From Concept to Applications* (pp. 405–413). Elsevier; Scopus. <https://doi.org/10.1016/B978-0-12-817133-2.00017-3>
- Bobrovitz, N., Lasserson, D. S., & Briggs, A. D. (2017). Who breaches the four-hour emergency department wait time target? A retrospective analysis of 374000 emergency department attendances between 2008 and 2013 at a type 1 emergency department in England. *BMC Emergency Medicine*, 17, 32. <https://doi.org/10.1186/s12873-017-0145-2>.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis* (1a ed.). John Wiley & Sons.
- Bottou, L. (2010). Large-Scale Machine Learning with Stochastic Gradient Descent. En Y. Lechevallier & G. Saporta (Eds.), *Proceedings of COMPSTAT'2010* (pp. 177–186). Physica-Verlag HD.
- Box, G. E. P. (2016). *Time series analysis: Forecasting and control* (Fifth edition.). Wiley.
- Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting* (3rd ed. 2016.). Springer International Publishing. <https://doi.org/10.1007/978-3-319-29854-2>
- Bruballa, E., Wong, A., Epelde, F., Rexachs, D., & Luque, E. (2016). A model to predict length of stay in a hospital emergency department and enable planning for non-

critical patients admission. *International Journal of Integrated Care (IJIC)*, 16(6), 1–2. Academic Search Ultimate.

Bunney, G., Tran, S., Han, S., Gu, C., Wang, H., Luo, Y., & Dresden, S. (2023). Using Machine Learning to Predict Hospital Disposition With Geriatric Emergency Department Innovation Intervention. *Annals of emergency medicine*, 81(3), 353–363. Ovid MEDLINE(R) <2023>.

<https://doi.org/10.1016/j.annemergmed.2022.07.026>

Calegari, R., Fogliatto, F. S., Lucini, F. R., Neyeloff, J., Kuchenbecker, R. S., & Schaan, B. D. (2016). Forecasting Daily Volume and Acuity of Patients in the Emergency Department. *Computational and Mathematical Methods in Medicine*, 2016, 3863268. <https://doi.org/10.1155/2016/3863268>

Carvalho-Silva, M., Monteiro, M. T. T., Sá-Soares, F. D., & Dória-Nóbrega, S. (2018). Assessment of forecasting models for patients arrival at Emergency Department. *Operations Research for Health Care*, 18, 112–118. Scopus.

<https://doi.org/10.1016/j.orhc.2017.05.001>

Castro, F. M., & Jojoa, P. E. (2019). Entrenamiento Comprimido Basado en Máquinas de Aprendizaje Extremo. *Informacion Tecnologica*, 30(4), 227–236.

<https://doi.org/10.4067/S0718-07642019000400227>

Char, D. S., Abramoff, M. D., & Feudtner, C. (2020). Identifying ethical considerations for machine learning healthcare applications. *The American Journal of Bioethics*, 20(11), 7–17.

Choudhury, A., & Urena, E. (2020). Forecasting hourly emergency department arrival using time series analysis. *British Journal of Health Care Management*, 26(1), 34–43. Scopus. <https://doi.org/10.12968/bjhc.2019.0067>

- Cleveland, R. B., Cleveland, W. S., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of official statistics*, 6(1), 3.
- Das, K. P., & J, C. (2023). A survey on artificial intelligence for reducing the climate footprint in healthcare. *Energy Nexus*, 9, 100167.  
<https://doi.org/10.1016/j.nexus.2022.100167>
- Das, L. T., Abramson, E. L., Stone, A. E., Kondrich, J. E., Kern, L. M., & Grinspan, Z. M. (2017). Predicting frequent emergency department visits among children with asthma using EHR data. *Pediatric pulmonology*, 52(7), 880–890. Ovid MEDLINE(R) <2017>. <https://doi.org/10.1002/ppul.23735>
- De Livera, A., Hyndman, R., & Snyder, R. (2010). Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing. *Journal of the American Statistical Association*, 106, 1513–1527.  
<https://doi.org/10.1198/jasa.2011.tm09771>
- Decreto 412 de 1992, 412 Decreto (1992).
- Erkamp, N. S., van Dalen, D. H., & de Vries, E. (2021). Predicting emergency department visits in a large teaching hospital. *International Journal of Emergency Medicine*, 14(1), 34. <https://doi.org/10.1186/s12245-021-00357-6>
- Estrada Atehortúa, A. F., & Zuluaga Gómez, M. (2020). Estrategias para la medición y el manejo de la sobreocupación de los servicios de urgencias de adultos en instituciones de alta complejidad con altos volúmenes de consulta. Revisión de la literatura. *Iatreia*, 33(1), 68–77.
- Featherstone, R. M., Dryden, D. M., Foisy, M., Guise, J.-M., Mitchell, M. D., Paynter, R. A., Robinson, K. A., Umscheid, C. A., & Hartling, L. (2015). Advancing knowledge of rapid reviews: An analysis of results, conclusions and recommendations from

published review articles examining rapid reviews. *Systematic Reviews*, 4(1), 50.

<https://doi.org/10.1186/s13643-015-0040-4>

Fergus, P., & Chalmers, C. (2022). *Applied deep learning: Tools, techniques, and implementation*. Springer.

Furterer, S. (2014). Hospital and emergency department throughput improvement. En *Lean Six Sigma Case Studies in the healthcare enterprise*. Springer-Verlag.

Gafni-Pappas, G., & Khan, M. (2023). Predicting daily emergency department visits using machine learning could increase accuracy. *The American journal of emergency medicine*, 65(aa2, 8309942), 5–11. Ovid MEDLINE(R) <2023>.

<https://doi.org/10.1016/j.ajem.2022.12.019>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning* (MIT Press).

<https://www.deeplearningbook.org/>

Griner, T. E., Thompson, M., High, H., & Buckles, J. (2020). Artificial Intelligence Forecasting Census and Supporting Early Decisions. *Nurs Adm*, 44(4), 316–328.

<https://doi.org/10.1097/NAQ.0000000000000436>

Gudivada, V. N. (2017). Chapter 2—Data Analytics: Fundamentals. En M. Chowdhury, A. Apon, & K. Dey (Eds.), *Data Analytics for Intelligent Transportation Systems* (pp. 31–67). Elsevier. <https://doi.org/10.1016/B978-0-12-809715-1.00002-X>

Guido Montúfar. (2018). *Restricted Boltzmann Machines: Introduction and Review* (pp. 75–115). Springer International Publishing. [https://doi.org/10.1007/978-3-319-97798-0\\_4](https://doi.org/10.1007/978-3-319-97798-0_4)

Gul, M., & Celik, E. (2020). An exhaustive review and analysis on applications of statistical forecasting in hospital emergency departments. *Health Systems*, 9(4), 263–284. Scopus. <https://doi.org/10.1080/20476965.2018.1547348>

- Gupta, A., & Katarya, R. (2020). Social media based surveillance systems for healthcare using machine learning: A systematic review. *Journal of biomedical informatics*, 108, 103500. <https://doi.org/10.1016/j.jbi.2020.103500>
- Hardin, J. W. (James W., & Hilbe, Joseph. (2013). *Generalized estimating equations* (2nd ed.). CRC Press. <https://doi.org/10.1201/b13880>
- Harrou, F., Dairi, A., Kadri, F., & Sun, Y. (2020). Forecasting emergency department overcrowding: A deep learning framework. *Chaos, Solitons, and Fractals*, 139, 110247. <https://doi.org/10.1016/j.chaos.2020.110247>
- Healy, M. J. R. (1976). The Analysis of Time Series: Theory and Practice. *Applied Statistics*, 25(3), 301–302. <https://doi.org/10.2307/2347243>
- Heizer, J., Render, B., & Munson, C. (2020). *Operations Management: Sustainability and Supply Chain Management* (13a ed.). Pearson.
- Higgins, J., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. The Cochrane Collaboration. [handbook.cochrane.org](http://handbook.cochrane.org)
- Hobbs, D., Kunzman, S. C., Tandberg, D., & Sklar, D. (2000). Hospital factors associated with emergency center patients leaving without being seen. *The American Journal of Emergency Medicine*, 18(7), 767–772. <https://doi.org/10.1053/ajem.2000.18075>
- Hong, L., Luo, M., Wang, R., Lu, P., Lu, W., & Lu, L. (2018). Big Data in Health Care: Applications and Challenges. *Data and Information Management*, 2(3), 175–197. <https://doi.org/10.2478/dim-2018-0014>
- Hsu, C.-C., & Sandford, B. (2007). The Delphi Technique: Making Sense Of Consensus. *Practical Assessment, Research and Evaluation*, 12.
- Hu, Y., Cato, K. D., Chan, C. W., Dong, J., Gavin, N., Rossetti, S. C., & Chang, B. P. (2023). Use of Real-Time Information to Predict Future Arrivals in the Emergency

- Department. *Annals of emergency medicine*, 81(6), 728–737. Ovid MEDLINE(R) <2023>. <https://doi.org/10.1016/j.annemergmed.2022.11.005>
- Huang, C., & Petukhina, A. (2022). *Applied Time Series Analysis and Forecasting with Python* (1st ed. 2022.). Springer International Publishing. <https://doi.org/10.1007/978-3-031-13584-2>
- Hyndman, R., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3a ed.). OTexts. <https://otexts.com/fpp3/>
- Ivanov, O., Wolf, L., Brecher, D., Lewis, E., Masek, K., Montgomery, K., Andrieiev, Y., McLaughlin, M., Liu, S., Dunne, R., Klauer, K., & Reilly, C. (2021). Improving ED Emergency Severity Index Acuity Assignment Using Machine Learning and Clinical Natural Language Processing. *Journal of Emergency Nursing*, 47(2), 265-278.e7. <https://doi.org/10.1016/j.jen.2020.11.001>
- James, C., Wood, R., & Denholm, R. (2023). A multi-granular stacked regression for forecasting long-term demand in Emergency Departments. *BMC Medical Informatics and Decision Making*, 23(1), 29. <https://doi.org/10.1186/s12911-023-02109-3>
- James, G. (Gareth M., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R* (Second Edition.). Springer US. <https://doi.org/10.1007/978-1-0716-1418-1>
- Javaid, M., Haleem, A., Pratap Singh, R., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. *International Journal of Intelligent Networks*, 3, 58–73.
- Jian, L., Patel, D., Xiao, J., Jansz, J., Yun, G., Lin, T., & Robertson, A. (2023). Can we use a machine learning approach to predict the impact of heatwaves on

- emergency department attendance? *Environmental Research Communications*, 5(4). Scopus. <https://doi.org/10.1088/2515-7620/acca6e>
- Jiang, S., Liu, Q., & Ding, B. (2023). A systematic review of the modelling of patient arrivals in emergency departments. *Quantitative Imaging in Medicine and Surgery*, 13(3), 1957–1971. <https://doi.org/10.21037/qims-22-268>
- Kousha, T., & Castner, J. (2016). The Air Quality Health Index and Emergency Department Visits for Otitis Media. *Journal of Nursing Scholarship*, 48(2), 163–171. Academic Search Premier.
- Kuhn, Max., & Johnson, Kjell. (2013). *Applied Predictive Modeling* (1st ed. 2013.). Springer New York. <https://doi.org/10.1007/978-1-4614-6849-3>
- Laghari, A. A., Sun, Y., Alhussein, M., Aurangzeb, K., Anwar, M. S., & Rashid, M. (2023). Deep residual-dense network based on bidirectional recurrent neural network for atrial fibrillation detection. *Scientific reports*, 13(1), 15109–15109. <https://doi.org/10.1038/s41598-023-40343-x>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature (London)*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lim, J. T., Dickens, B. S. L., & Cook, A. R. (2020). Modelling the epidemic extremities of dengue transmissions in Thailand. *Epidemics*, 33, 100402. <https://doi.org/10.1016/j.epidem.2020.100402>
- Marcilio, I., Hajat, S., & Gouveia, N. (2013). Forecasting daily emergency department visits using calendar variables and ambient temperature readings. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*, 20(8), 769–777. Ovid MEDLINE(R) <2013>. <https://doi.org/10.1111/acem.12182>

- Marco, C. A., Courtney, D. M., Ling, L. J., Salsberg, E., Reisdorff, E. J., Gallahue, F. E., Suter, R. E., Muelleman, R., Chappell, B., Evans, D. D., Vafaie, N., & Richwine, C. (2021). The Emergency Medicine Physician Workforce: Projections for 2030. *Annals of Emergency Medicine*, 78(6), 726–737.  
<https://doi.org/10.1016/j.annemergmed.2021.05.029>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (Fifth edition.). John Wiley & Sons Ltd.
- Murtas, R., Tunesi, S., Andreano, A., & Russo, A. G. (2022). Time-series cohort study to forecast emergency department visits in the city of Milan and predict high demand: A 2-day warning system. *BMJ Open*, 12(4), e056017.  
<https://doi.org/10.1136/bmjopen-2021-056017>
- Nas, S., & Koyuncu, M. (2019). Emergency Department Capacity Planning: A Recurrent Neural Network and Simulation Approach. *Computational and Mathematical Methods in Medicine*, 2019, 4359719. <https://doi.org/10.1155/2019/4359719>
- Newcomb, P., & Li, J. (2008). Predicting admissions for childhood asthma based on proximity to major roadways. *Journal of nursing scholarship : an official publication of Sigma Theta Tau International Honor Society of Nursing*, 40(4), 319–325. Ovid MEDLINE(R) <2008 to 2009>. <https://doi.org/10.1111/j.1547-5069.2008.00245.x>
- Oreshkin, B. N., Carpov, D., Chapados, N., & Bengio, Y. (2020). N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *arXiv.org*.  
<https://doi.org/10.48550/arxiv.1905.10437>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated

guideline for reporting systematic reviews. *BMJ*, 372, n71.

<https://doi.org/10.1136/bmj.n71>

Peláez-Rodríguez, C., Torres-López, R., Pérez-Aracil, J., López-Laguna, N., Sánchez-Rodríguez, S., & Salcedo-Sanz, S. (2024). An explainable machine learning approach for hospital emergency department visits forecasting using continuous training and multi-model regression. *Computer Methods and Programs in Biomedicine*, 245, 108033. <https://doi.org/10.1016/j.cmpb.2024.108033>

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., ... Ziel, F. (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, 38(3), 705–871.

<https://doi.org/10.1016/j.ijforecast.2021.11.001>

Petsis, S., Karamanou, A., Kalampokis, E., & Tarabanis, K. (2022). Forecasting and explaining emergency department visits in a public hospital. *Journal of Intelligent Information Systems*, 59(2), 479–500. Scopus. <https://doi.org/10.1007/s10844-022-00716-6>

Piliuk, K., & Tomforde, S. (2023). Artificial intelligence in emergency medicine. A systematic literature review. *International Journal of Medical Informatics*, 180, 105274. <https://doi.org/10.1016/j.ijmedinf.2023.105274>

Rocha, C. N., & Rodrigues, F. (2021). Forecasting emergency department admissions. *Journal of Intelligent Information Systems*, 56(3), 509–528. Scopus. <https://doi.org/10.1007/s10844-021-00638-9>

- Rodriguez, F., Jimenez, W., & Palencia, F. (2018). Uso de los servicios de urgencias en Bogotá, Colombia: Un análisis desde el Triage. *Universidad y Salud*, 20(3), 215–226.
- Salinas, D., Flunkert, V., & Gasthaus, J. (2019). DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks. *arXiv.org*.  
<https://doi.org/10.48550/arxiv.1704.04110>
- Schaffer, A., Muscatello, D., Broome, R., Corbett, S., & Smith, W. (2012). Emergency department visits, ambulance calls, and mortality associated with an exceptional heat wave in Sydney, Australia, 2011: A time-series analysis. *Environmental health : a global access science source*, 11(1), 3. Ovid MEDLINE(R) <2012>. <https://doi.org/10.1186/1476-069X-11-3>
- Schneider, S. M., Hamilton, G. C., Moyer, P., & Stapczynski, J. S. (1998). Definition of emergency medicine. *cad Emerg Med*, 5, 348–351.
- Sendak, M. A., D'Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., & Balu, S. (2020). A path for translation of machine learning products into healthcare delivery. *EMJ Innov*, 10, 19–172.
- Shi, Y., Wang, H., Zhang, L., Zhang, M., Shi, X., Pei, H., & Bai, Z. (2022). Nomogram Models for Predicting Delirium of Patients in Emergency Intensive Care Unit: A Retrospective Cohort Study. *International Journal of General Medicine*, 15, 4259–4272. <https://doi.org/10.2147/IJGM.S353318>
- Slack, N., & Lewis, M. (2019). *Operations Strategy* (5a ed.). Pearson.
- Sookmee, W., Liabsuetrakul, T., Tantarattanapong, S., & Wuthisuthimethawee, P. (2024). Emergency Department Length of Stay and in-Hospital Mortality of Non-Traumatic Patients in a University Hospital. *Journal of Health Science and Medical Research*, 42(3). Scopus. <https://doi.org/10.31584/jhsmr.20231018>

- Stieb, D. M., Szyszkowicz, M., Rowe, B. H., & Leech, J. A. (2009). Air pollution and emergency department visits for cardiac and respiratory conditions: A multi-city time-series analysis. *Environmental health : a global access science source*, 8(101147645), 25. Ovid MEDLINE(R) <2008 to 2009>. <https://doi.org/10.1186/1476-069X-8-25>
- Stover-Baker, B., Stahlman, B., & Pollack, M. (2012). Triage nurse prediction of hospital admission. *Journal of emergency nursing*, 38(3), 306–310. Ovid MEDLINE(R) <2012>. <https://doi.org/10.1016/j.jen.2011.10.003>
- Sulaiman, W. A., Nicolaou, A., Prentza, N., Stylianides, C., Panayides, A., Constantinou, I., Antoniou, Z., Kakas, A., Kyriacou, E., Palazis, L., Kyprianouv, T., & Pattichis, C. S. (2023). Emergency Department Triage Hospitalization Prediction Based on Machine Learning and Rule Extraction. *IEEE EMBS Spec. Topic Conf. Data Sci. Eng. Healthc., Med. Biol., IEEECONF*, 147–148. Scopus. <https://doi.org/10.1109/IEEECONF58974.2023.10405176>
- Sun, B., HSia, R., & Weis, R. (2013). Effect of emergency department crowding on outcomes of admitted patients. *Annals of Emergency Medicine*, 61(6), 605–611.
- Tang, K. J. W., Ang, C. K. E., Constantinides, T., Rajinikanth, V., Acharya, U. R., & Cheong, K. H. (2021). Artificial Intelligence and Machine Learning in Emergency Medicine. *Biocybernetics and Biomedical Engineering*, 41(1), 156–172. <https://doi.org/10.1016/j.bbe.2020.12.002>
- Tang, L., Li, J., Du, H., Li, L., Wu, J., & Wang, S. (2022). Big Data in Forecasting Research: A Literature Review. *Big Data Research*, 27, 100289. <https://doi.org/10.1016/j.bdr.2021.100289>
- Tawhid, A., Teotia, T., & Elmiligi, H. (2021). Chapter 13—Machine learning for optimizing healthcare resources. En P. Kumar, Y. Kumar, & M. A. Tawhid (Eds.), *Machine*

- Learning, Big Data, and IoT for Medical Informatics* (pp. 215–239). Academic Press. <https://doi.org/10.1016/B978-0-12-821777-1.00020-3>
- Taylor, S. J., & Letham, B. (2018). Forecasting at Scale. *The American statistician*, 72(1), 37–45. <https://doi.org/10.1080/00031305.2017.1380080>
- The University of Newcastle. (2024). *Systematic review guide: Overview*. Library Guides. <https://libguides.newcastle.edu.au/sysreviews>
- Tricco, A., Laglois, E., & Straus, S. (2017). *Rapid reviews to strengthen health policy and systems: A practical guide*. World Health Organization.
- Tuominen, J., Lomio, F., Oksala, N., Palomäki, A., Peltonen, J., Huttunen, H., & Roine, A. (2022). Forecasting daily emergency department arrivals using high-dimensional multivariate data: A feature selection approach. *BMC Medical Informatics and Decision Making*, 22(1), 134. <https://doi.org/10.1186/s12911-022-01878-7>
- Tuominen, J., Pulkkinen, E., Peltonen, J., Kanninen, J., Oksala, N., Palomäki, A., & Roine, A. (2024). Forecasting emergency department occupancy with advanced machine learning models and multivariable input. *International Journal of Forecasting*. Scopus. <https://doi.org/10.1016/j.ijforecast.2023.12.002>
- Vanegas, C., & Restrepo, S. (2022). *Sabes qué es una urgencia?* Universidad del Rosario. <https://urosario.edu.co/en/node/7482>
- Vezyridis, P., & Timmons, S. (2014). National targets, proces transformation and local consequences in an NHS emergency department (ED): A qualitative study. *BMC Emergency Medicine*, 14, 12.
- Vieira, A., Sousa, I., & Dória-Nóbrega, S. (2023). Forecasting daily admissions to an emergency department considering single and multiple seasonal patterns. *Healthcare Analytics*, 3. Scopus. <https://doi.org/10.1016/j.health.2023.100146>

- Vollmer, M. A. C., Glampson, B., Mellan, T., Mishra, S., Mercuri, L., Costello, C., Klaber, R., Cooke, G., Flaxman, S., & Bhatt, S. (2021). A unified machine learning approach to time series forecasting applied to demand at emergency departments. *BMC Emergency Medicine*, 21(1), 9. <https://doi.org/10.1186/s12873-020-00395-y>
- Wang, G.-Q., Gao, Y.-L., & Deng, P. (2023). Validation of a seven-question tool (PRISMA-7) in predicting prognosis of older adults in the emergency department: A prospective study. *The American journal of emergency medicine*, 73(aa2, 8309942), 131–136. Ovid MEDLINE(R) Epub Ahead of Print. <https://doi.org/10.1016/j.ajem.2023.08.030>
- Ward, M. J., Marsolo, K. A., & Froehle, C. M. (2014). Applications of business analytics in healthcare. *Business Horizons*, 57(5), 571–582. <https://doi.org/10.1016/j.bushor.2014.06.003>
- Whitt, W., & Zhang, X. (2019). Forecasting arrivals and occupancy levels in an emergency department. *Operations Research for Health Care*, 21, 1–18. Scopus. <https://doi.org/10.1016/j.orhc.2019.01.002>
- Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553.
- Xiao, C., & Sun, J. (2021). *Introduction to Deep Learning for Healthcare* (1st ed. 2021.). Springer International Publishing AG. <https://doi.org/10.1007/978-3-030-82184-5>
- Yousefi, M., Yousefi, M., Fathi, M., & Fogliatto, F. S. (2020). Patient visit forecasting in an emergency department using a deep neural network approach. *Kybernetes*, 49(9), 2335–2348. Academic Search Premier.
- Zhang, Y., Zhang, J., Tao, M., Shu, J., & Zhu, D. (2022). Forecasting patient arrivals at emergency department using calendar and meteorological information. *Applied*

*Intelligence (Dordrecht, Netherlands)*, 52(10), 11232–11243.

<https://doi.org/10.1007/s10489-021-03085-9>

Zhao, X., Lai, J. W., Ho, A. F. W., Liu, N., Ong, M. E. H., & Cheong, K. H. (2022).

Predicting hospital emergency department visits with deep learning approaches.

*Biocybernetics and Biomedical Engineering*, 42(3), 1051–1065. Scopus.

<https://doi.org/10.1016/j.bbe.2022.07.008>