Universidad Europea VALENCIA

GRADO EN CIENCIA DE DATOS

Modelo predictivo para la anticipación horaria de la demanda en urgencias médicas del Hospital Virgen de la Luz

Presentado por:

LEYRE VERDÚ FUENTE

Dirigido por:

MIGUEL ANGEL TORRES FONT

CURSO ACADÉMICO 2025-2026

RESUMEN

En los últimos años, los servicios de urgencias hospitalarias han experimentado un incremento en la afluencia de pacientes, lo que ha derivado en episodios recurrentes de saturación y dificultades para la planificación de los recursos sanitarios. Dado este contexto, el presente Trabajo Fin de Grado desarrolla un sistema predictivo orientado a estimar la llegada horaria de pacientes al servicio de urgencias del Hospital Virgen de la Luz de Cuenca.

El estudio se basa en el análisis de datos reales y anonimizados cedidos por el propio hospital, durante el periodo comprendido entre diciembre del 2021 y noviembre del 2025. A partir de estos datos, se construye una serie temporal horaria, enriquecida con variables de calendario, festivos y meteorología.

Se comparan diferentes modelos para estimar las llegadas en la hora siguiente y realizar consultas sobre fechas futuras. Los resultados muestran que es posible obtener estimaciones útiles de la demanda horaria, con errores medios próximos a dos pacientes por hora.

Como complemento, se desarrolla un panel de visualización interactivo diseñado para facilitar la interpretación e identificar patrones temporales y jornadas de mayor demanda.

Con este proyecto se busca ofrecer una herramienta útil e intuitiva que ayude a facilitar la planificación de las urgencias en el Hospital de Cuenca y que permita una mejor asignación de recursos, priorizando la calidad de la atención y la anticipación de necesidades, reduciendo las esperas de los pacientes.

Palabras clave: urgencias, predicción, demanda, afluencia, series temporales, modelado, visualización.

ABSTRACT

In recent years, hospital emergency departments have experienced an increase in the influx of patients, which has led to recurrent episodes of overcrowding and difficulties in planning healthcare resources. In this context, this Final Degree Project develops a predictive system aimed at estimating the hourly arrival of patients at the emergency department of Hospital Virgen de la Luz in Cuenca.

The study is based on the analysis of real, anonymised data provided by the hospital itself, covering the period between December 2021 and November 2025. Based on these data, an hourly time series is constructed, enriched with calendar, public holiday and meteorological variables.

Different models are compared to estimate arrivals in the following hour and to make queries regarding future dates. The results show that it is possible to obtain useful estimates of hourly demand, with average errors close to two patients per hour.

As a complement, an interactive visualisation dashboard is developed, designed to facilitate interpretation and identify temporal patterns and days of higher demand.

This project seeks to offer a useful and intuitive tool to help facilitate emergency department planning at Cuenca Hospital and enable a better allocation of resources, prioritising quality of care and the anticipation of needs, while reducing patient waiting times.

Key words: emergency, prediction, demand, attendance, time series, modelling, visualization.

ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN 14

2. OBJETIVOS 16

2.1 Objetivo General 17

2.2 Objetivos Específicos 18

3. MARCO TEÓRICO 19

3.1 Contexto asistencial del Hospital 19

3.2 Predicción de la demanda en urgencias hospitalarias 20

3.3 Patrones temporales de la demanda 21

3.4 Variables predictoras en urgencias 22

3.4.1 Variables temporales y de calendario 23

3.4.2 Variables clínicas y operativas 24

3.4.3 Variables exógenas: meteorología y contexto 24

3.4.4 Variables históricas, retardos y medias móviles 25

3.4.5 Calidad del dato: ruido, outliers y cambios de régimen 25

3.5 Series temporales y predicción 26

3.6 Familias de modelos predictivos 27

3.6.1 Modelos base o baselines 27

3.6.2 Modelos estadísticos clásicos 29

3.6.3 Modelos de regresión y conteo 31

3.6.4 Modelos de machine learning tabular 32

3.6.5 Modelos de deep learning 36

3.6.6 Comparación y criterios de selección 38

3.7 Validación temporal y métricas de evaluación 40

4. METODOLOGÍA 41

4.1 Enfoque metodológico del trabajo 41

4.2 Estrategias metodológicas consideradas 42
4.3 Diseño general del flujo de trabajo 43
4.4 Fuentes de datos y periodo de análisis 44
4.5 Estrategia de modelado predictivo 46
4.6 Diseño experimental y validación temporal 46
4.7 Métricas y criterios de evaluación 47
4.8 Recursos utilizados y reproducibilidad 48
4.9 Salida operativa: predicción futura y dashboard 48
5. DESARROLLO CIENTÍFICO-TÉCNICO 49
5.1 Visión general del desarrollo 49
5.2 Diseño del estudio 50
5.2.1 Tipo de estudio y formulación del problema 50
5.2.2 Población, periodo y unidad de análisis 51
5.3 Instrumentos y entorno tecnológico 52
5.4 Fuentes de datos 53
5.4.1 Datos hospitalarios 53
5.4.2 Datos capacidad 55
5.4.3 Datos meteorológicos (AEMET) 56
5.4.4 Calendario de festivos de Cuenca 57
5.4.5 Variables descartadas por no disponibilidad 57
5.5 Arquitectura del procedimiento 58
5.6 Limpieza del Excel original 58
5.6.1 Tratamiento de fechas y duraciones 58
5.6.2 Normalización de variables categóricas 59
5.6.3 Tratamiento de outliers 60
5.7 Construcción de la serie horaria 62
5.7.1 Variable objetivo: LLEGADAS_HORA 62

5.7.2 Ocupación simultánea 63
5.7.3 Indicadores de presión asistencial 63
5.7.4 Métricas operativas adicionales 64
5.8 Calendario y festivos 65
5.9 Incorporación de meteorología 66
5.10 Retardos y medias móviles 66
5.10.1 Imputación conservadora de meteorología restante 67
5.10.2 Codificación cíclica de variables temporales 67
5.10.3 Variables creadas a partir de valores anteriores (Lags del target) 67
5.10.4 Media móvil y desviación típica móvil 68
5.10.5 Lags de variables operativas observadas 68
5.10.6 Eliminación del periodo de calentamiento 69
5.11 Construcción del dataset predictivo final 69
5.11.1. Feature autoregresiva en t (LLEGADAS_HORA_T) 69
5.11.2 Variables excluidas y justificación 70
5.11.3 Verificaciones automáticas y salida final 71
5.12 Caracterización del dataset y análisis exploratorio 71
5.12.1 Distribución temporal y partición train/test 71
5.12.2 Distribución del target 72
5.12.3 Patrones temporales observados 73
5.13 Evolución del dataset predictivo 73
5.13.1 Metodología de búsqueda de features 73
5.13.2 Features descartadas tras el análisis 74
5.13.3 Dataset v2 variables aceptadas. 75
5.13.4 Dataset v3 variables aceptadas. 76
5.13.5 Ranking final de variables por correlación con el target 78
5.13.6 Estado final del dataset por versiones 78

5.14 Esquema de validación temporal 79
5.14.1 Hold-out final (test 2025) 79
5.14.2 TimeSeriesSplit con 3 folds sobre el train 79
5.15 Modelos de referencia (baselines) 80
5.15.1 Baseline 1: Naive estacional 168 h 81
5.15.2 Baseline 2: media histórica por (hora, día-semana, mes) 81
5.15.3 Baseline 3: MEDIA_HORA_DOW directa 81
5.15.4 Análisis de los baselines 82
5.16 Modelos de aprendizaje entrenados 83
5.16.1 Familias de modelos y justificación de la selección 83
5.16.2 SARIMAX 85
5.16.3 Prophet 85
5.16.4 LightGBM con optimización bayesiana 85
5.16.5 LSTM con MC-Dropout 86
5.16.6 Modelos adicionales sobre v2 87
5.17 Modelo futuro para predicción manual 88
5.17.1 Construcción del dataset futuro 88
5.17.2 Prevención de fuga de información temporal 89
5.17.3 Modelos entrenados en la versión futura 90
5.17.4 Modelos adicionales y ajustes específicos 91
5.17.5 Selección del modelo final futuro 92
5.17.6 Predicción manual de una fecha futura 92
5.17.7 Utilidad operativa del modelo futuro 94
5.18 Desarrollo panel interactivo 94
5.18.1 Estructura del panel 95
5.18.2 Utilidad dentro del proyecto 96
5.19 Código GitHub 96

6. RESULTADOS 97
6.1 Distribución horaria y semanal 97
6.2 Resultados evolución: v1, v2 y v3. 98
6.3 Comparación real-predicción v2 101
6.4 Utilidad del modelo v2 para el servicio de urgencias 102
6.5 Modelo v4 para predicción manual futura 103
6.5.1 Ejemplo de uso rango de estimación anticipada 105
6.6 Resultados del dashboard 106
6.6.1 Vista 1. Resumen ejecutivo de la actividad 106
6.6.2 Vista 2. Patrón horario y semanal 107
6.6.3 Vista 3. Evolución mensual y estacionalidad 108
6.6.4 Vista 4. Días críticos y picos extraordinarios 110
6.6.5 Vista 5. Triaje y tiempos de espera 111
6.6.6 Vista 6. Procedencia y destino asistencial 112
7. DISCUSIÓN 113
7.1 Interpretación de los principales hallazgos 114
7.2 Relación con investigaciones anteriores 114
7.3 Resultados discrepantes con la literatura 115
8. OBJETIVOS DE DESARROLLO SOSTENIBLE (ODS) 116
8.1 ODS 3. Salud y bienestar 116
8.2 ODS 9. Industria, innovación e infraestructura 116
8.3 ODS 10. Reducción de las desigualdades 117
8.4 ODS 11. Ciudades y comunidades sostenibles 117
9. CONCLUSIONES 117
9.1 Conclusiones respecto al objetivo general 118
9.2 Conclusiones respecto a los objetivos específicos 118
9.3 Limitaciones del estudio 120

9.4 Implicaciones prácticas y líneas futuras 120

BIBLIOGRAFÍA 122

ÍNDICE DE ECUACIONES

Ecuación 1. Modelo naive simple..... 28
Ecuación 2. Modelo naive estacional semanal aplicado a datos horarios. 28
Ecuación 3. Media histórica de llegadas por hora y día de la semana. 29
Ecuación 4. Notación general del modelo ARIMA..... 30
Ecuación 5. Especificación general del modelo SARIMA..... 30
Ecuación 6. Extensión del modelo SARIMA mediante variables exógenas. 30
Ecuación 7. Formulación aditiva del modelo Prophet para series temporales. 31
Ecuación 8. Formulación general de la regresión lineal múltiple 31
Ecuación 9. Modelo de regresión de Poisson con función de enlace logarítmica..... 32
Ecuación 10. Predicción media en un nodo hoja del árbol de decisión..... 33
Ecuación 11. Predicción agregada de un modelo Random Forest. 34
Ecuación 12. Predicción acumulativa en Gradient Boosting. 35
Ecuación 13. Predicción promedio de un modelo ensemble. 36
Ecuación 14. Formulación general de una red LSTM para predicción de series temporales. 37
Ecuación 15. Error Absoluto Medio (MAE) 38
Ecuación 16. Raíz del Error Cuadrático Medio (RMSE). 39
Ecuación 17. Sesgo medio de predicción (BIAS). 39

ÍNDICE DE FIGURAS

Figura 1. Hospital Virgen de la Luz Cuenca 15
Figura 2. Flujo de las fases metodológicas seguidas en el proyecto. 43
Figura 3. Distribución completa de LOS_MIN en escala logarítmica y boxplot con outliers extremos 61
Figura 4. Zoom de la distribución de LOS_MIN en el rango 0-2.000 minutos en escala lineal. 62
Figura 5. Estadísticos descriptivos del target. 72
Figura 6. Ranking de features por correlación absoluta con el target 78
Figura 7. Gráfico de los folds de validación temporal (TimeSeriesSplit, 3 folds). 80
Figura 8. Ejemplo de salida por consola de las variables añadidas manualmente. 93
Figura 9. Resultado mostrado por pantalla sobre la hora prevista. 93
Figura 10. Captura de pantalla del repositorio de GitHub 97
Figura 11. Demanda media de pacientes según el día de la semana y la hora del día. .. 97
Figura 12. Resultados dataset v1. 98
Figura 13. Resultados dataset v2 99
Figura 14. Resultados dataset v3 99
Figura 15. Resultados dataset v2 nuevos modelos añadidos. 100
Figura 16. Comparación entre llegadas reales y predicciones durante el mes de enero 2025. 101
Figura 17. Error medio por franja horaria: estimación basada en la semana anterior frente al modelo desarrollado 102
Figura 18. Resultados modelos v4 104
Figura 19. Ejemplo de estimación anticipada de la demanda por horas para un día completo. 105
Figura 20. Vista de resumen ejecutivo de la actividad registrada en el servicio de urgencias durante el periodo analizado. 107
Figura 21. Distribución horaria y semanal de la demanda asistencial mediante mapa de calor. 108
Figura 22. Evolución mensual de las atenciones y comparación de la demanda entre años. 109
Figura 23. Identificación de días críticos mediante el umbral del percentil 95 de demanda diaria. 111

Figura 24. Tiempo hasta el inicio del triaje según el nivel de prioridad asignado al paciente. 112
Figura 25. Distribución de la procedencia de los pacientes y del destino final tras la atención en urgencias..... 113

ÍNDICE DE TABLAS

Tabla 1. Fuentes de datos originales utilizadas en el proyecto. 45
Tabla 2. Resumen total de datasets utilizados en el proyecto. 45
Tabla 3. Librerías empleadas en el proyecto 53
Tabla 4. Variables originales del dataset hospitalario. 55
Tabla 5. Capacidad del Servicio de Urgencias. 56
Tabla 6. Variables documentadas, pero no disponibles en los datos cedidos. 58
Tabla 7. Definición de presión asistencial 64
Tabla 8. Variables excluidas del dataset predictivo y motivo de exclusión. 70
Tabla 9. Distribución temporal del dataset y partición train/test. 72
Tabla 10. Variables candidatas descartadas y motivo del descarte 75
Tabla 11. Variables grupo C 77
Tabla 12. Versiones finales de los datasets. 79
Tabla 13. Suelos de comparación para los modelos de ML. 82
Tabla 14. Modelos entrenados y su justificación. 84
Tabla 15. Modelos utilizados dataset futuro. 91
Tabla 16. Resumen estructura panel. 95
Tabla 17. Mejor resultado obtenido en cada dataset. 101
Tabla 18. Resumen comparación de los dos modelos principales para la selección final de v4. 105

1. INTRODUCCIÓN

En España, los servicios de urgencias hospitalarias constituyen uno de los ámbitos de mayor exigencia dentro del sistema sanitario. En los últimos años, han experimentado un aumento significativo en la afluencia de pacientes[1], lo que ha provocado episodios recurrentes de saturación que afectan tanto a la calidad de la atención asistencial como a la gestión diaria de los recursos sanitarios. Esta situación se ve agravada por la gran presión sobre el personal sanitario, que viene marcada por largas jornadas laborales, falta de profesionales y conflictos laborales derivados, en muchos casos, de condiciones salariales poco sostenibles. Esta situación ha dado lugar a movilizaciones y huelgas[2] en distintos centros hospitalarios, incrementando aún más la complejidad en la gestión de los servicios de urgencias.

La combinación de una alta demanda asistencial y una disponibilidad limitada de recursos humanos y materiales dificulta la planificación eficiente de los servicios, generando colapsos asistenciales, prolongados tiempos de espera y una elevada incertidumbre en la toma de decisiones. Además, en muchos hospitales, la ausencia de herramientas tecnológicas avanzadas para el análisis de datos históricos y la predicción de la demanda futura obliga a basar la planificación en estimaciones poco precisas, lo que limita la capacidad de anticipación ante situaciones críticas.

En este contexto, el uso de modelos predictivos basados en datos históricos se presenta como una alternativa eficaz para reducir dicha incertidumbre. La aplicación de técnicas de series temporales y aprendizaje automático puede ayudar a anticipar franjas de mayor afluencia y aportar información adicional para la planificación del servicio. Además, en estudios recientes se ha demostrado que la incorporación de variables de calendario y meteorología pueden mejorar la predicción de llegadas en servicios de urgencias y facilitar la planificación de recursos disponibles. *(Porto y Fogliatto, 2024).*

[1] Ministerio de Sanidad. (2025). *Informe Anual del Sistema Nacional de Salud 2024*. Apartado 5.1.4.2, "Urgencias hospitalarias", p. 74. El informe recoge 24.895.110 urgencias hospitalarias atendidas en 2023 y un aumento del 17,0 % en la frecuentación entre 2013 y 2023.

[2] Ministerio de Sanidad. (2025, 11 de diciembre). *Sanidad atiende todas las demandas del comité de huelga dentro de sus competencias y garantizando la unidad del Sistema Nacional de Salud*. La nota informa de las reuniones con las organizaciones sindicales convocantes de la huelga médica y de cuestiones relacionadas con turnos, guardias, carga de trabajo y condiciones laborales.

En este contexto surge el presente Trabajo Fin de Grado, (TFG, en adelante), cuyo objetivo no pretende sustituir la toma de decisiones clínica ni resolver por sí solo la saturación asistencial, sino proporcionar una herramienta analítica que permita anticipar la demanda horaria esperada y apoyar la planificación de recursos.

Para ello, se desarrolla un sistema predictivo basado en datos reales del Hospital Virgen de la Luz de Cuenca (Hospital, en adelante), estructurados en una serie temporal horaria y enriquecidos con variables temporales, de calendario, meteorológicas y de comportamiento histórico de la demanda. A partir de este conjunto de datos, se prueban distintos tipos de modelos predictivos, desde los más simples y tradicionales hasta técnicas más avanzadas de machine learning e inteligencia artificial, para determinar cuál se adapta mejor a la realidad del hospital en términos de precisión, fiabilidad y facilidad de uso.

Además, el trabajo incorpora una versión específica orientada a estimar la demanda en fechas futuras introducidas manualmente, utilizando únicamente información disponible o estimable antes del momento de la predicción.

Finalmente, se desarrolla una herramienta visual clara e intuitiva, destinada a visualizar la actividad histórica del servicio, identificar patrones temporales y detectar jornadas de mayor demanda. De este modo, el proyecto combina el desarrollo de modelos predictivos con una herramienta visual que facilita la interpretación de la información obtenida.



Figura 1. Hospital Virgen de la Luz Cuenca

2. OBJETIVOS

Antes de presentar los objetivos del trabajo, es importante concretar el alcance real del proyecto dentro del contexto del servicio de urgencias analizado. El TFG no se limita a describir cómo ha evolucionado la demanda de urgencias en el pasado, sino que desarrolla un sistema capaz de predecir cuántos pacientes se esperan en el futuro.

Para ello, se parte de los registros históricos de urgencias, que se organizan en una serie temporal horaria y se incorporan variables de calendario, festivos, meteorología y comportamiento histórico de la demanda. A partir de este conjunto de datos, se entrenan y comparan distintas familias de modelos predictivos con el fin de identificar qué enfoque funciona mejor en este contexto concreto.

Además, el proyecto no contempla una única versión del conjunto de datos, sino que por fases va evolucionando, buscando una mejora de los resultados. En primer lugar, se parte de una versión base con variables temporales, meteorológicas, lags y medias móviles. Posteriormente, se incorporan variables históricas más explicativas y, finalmente, se desarrolla una versión específica orientada a la predicción manual de fechas futuras.

Los datos se presentan a través de una herramienta visual interactiva, diseñada para que cualquier responsable del servicio pueda interpretar los resultados y usarlas como apoyo en la planificación diaria.

2.1 Objetivo General

El objetivo principal de este TFG es desarrollar un sistema capaz de predecir cuántos pacientes llegarán al servicio de urgencias en una hora concreta, utilizando datos históricos reales del hospital combinados con información de calendario, meteorología y patrones de demanda anteriores. Para alcanzar este objetivo, el TFG compara dos enfoques de modelado diferentes.

En primer lugar, se busca predecir las llegadas esperadas en la hora siguiente a partir de la información disponible en el momento actual. Donde se utilizan modelos de referencia, modelos estadísticos clásicos de series temporales, modelos de aprendizaje automático basados en árboles y se incorpora un modelo de aprendizaje profundo secuencial.

Por otro lado, a través de modelos como en el enfoque anterior y modelos más complejos, se desarrolla una versión del modelo diseñado para consultas manuales: el usuario introduce una fecha futura y el sistema devuelve una estimación de la demanda esperada para la franja horaria de ese día.

Este enfoque requiere un tratamiento distinto, ya que en el momento de la consulta no se dispone de información en tiempo real, como el número de pacientes llegados en las horas previas. Es por eso, que el modelo se construye exclusivamente usando únicamente información disponible de antemano, como el día, si es festivo, la previsión meteorológica y patrones históricos de demanda en casos donde esa información no se conozca.

La finalidad de esta comparación no es solo encontrar el error más bajo, sino el mejor equilibrio entre precisión y utilidad práctica para el hospital.

2.2 Objetivos Específicos

Para dar respuesta a esta necesidad, el trabajo se desarrolla a través de los siguientes objetivos concretos:

- Analizar y depurar los datos históricos del servicio de urgencias, garantizando la coherencia de fechas, registros horarios, tiempos de estancia y posibles valores anómalos o faltantes.
- Construir una serie temporal que organice los datos hora a hora y refleje cuántos pacientes llegan al servicio en cada franja horaria.
- Incorporar información externa relevante, como el día de la semana, festivos y condiciones meteorológicas, para enriquecer el contexto de cada predicción.
- Añadir variables que capturen el comportamiento reciente de la demanda, como los valores de horas anteriores, medias móviles y patrones históricos habituales.
- Diseñar y comparar distintas versiones del conjunto de datos predictivo, justificando la evolución desde una versión inicial hasta versiones más completas con nuevas variables.
- Desarrollar una versión del modelo diseñado para consultas manuales, en la que el usuario introduce una fecha y sus distintas franjas horarias futuras y el sistema estima la demanda esperada usando solo información disponible de antemano.
- Entrenar y comparar distintos tipos de modelos predictivos, desde los más simples hasta técnicas avanzadas de machine learning e inteligencia artificial.
- Comparar el rendimiento de los modelos mediante métricas como MAE, RMSE y sesgo, analizando también su comportamiento por franjas horarias y tipo de día.
- Seleccionar el modelo más adecuado para el hospital, valorando no solo su precisión, sino también su estabilidad, su facilidad de interpretación y su utilidad práctica.
- Diseñar un panel interactivo de visualización que permita visualizar la actividad histórica del servicio.

3. MARCO TEÓRICO

El presente marco teórico tiene como finalidad entender las decisiones metodológicas que se toman a lo largo del trabajo, para ello, es necesario establecer primero un marco de referencia conceptual y técnico. Es por eso que se empezará, describiendo la problemática asociada a la demanda asistencial en los servicios de urgencias y la necesidad de anticipar situaciones de elevada presión asistencial.

A continuación, se analiza la estructura temporal que caracteriza la afluencia de pacientes, con sus patrones horarios, semanales, estacionales y de calendario. Posteriormente, se revisan los principales tipos de variables utilizadas en este tipo de problemas predictivos, incluyendo variables temporales, meteorológicas, operativas e históricas.

Finalmente, se presentan las familias de modelos predictivos más relevantes para este tipo de problema, así como sus criterios de validación y evaluación necesarios para comparar su rendimiento de forma rigurosa. Este marco permite justificar las decisiones adoptadas en los apartados posteriores del trabajo y la selección final de la alternativa más adecuada.

3.1 Contexto asistencial del Hospital

El trabajo se desarrolla a partir de la actividad registrada en el Servicio de Urgencias del Hospital Virgen de la Luz de Cuenca. En el periodo analizado, este centro pertenecía al Servicio de Salud de Castilla-La Mancha (SESCAM) y era el hospital público de referencia de la provincia. Se trataba de un hospital general, de dependencia autonómica y con acreditación docente *(Ministerio de Sanidad, 2025)*.

El Servicio de Urgencias constituye un área relevante dentro de la actividad hospitalaria, ya que debe atender una demanda no programada y variable a lo largo del tiempo. Esta variabilidad hace que la organización de los recursos resulte compleja, especialmente cuando se producen franjas de mayor afluencia o situaciones de elevada presión asistencial.

Los datos utilizados corresponden a la actividad del Hospital hasta noviembre de 2025. A partir de esa fecha, el Hospital Universitario de Cuenca[3] pasó a asumir la atención sanitaria del centro anterior, incluido el Servicio de Urgencias. Este cambio no afecta al estudio, ya que el objetivo del trabajo es analizar la demanda observada durante el periodo del que se dispone de datos.

3.2 Predicción de la demanda en urgencias hospitalarias

Los servicios de urgencias hospitalarias presentan uno de los entornos de complejidad organizativa más elevados. Su naturaleza no programada, la diversidad en la gravedad de los pacientes atendidos y la necesidad de responder en tiempos reducidos los diferencian claramente de otros servicios con actividad planificada. A esto se le suma la necesidad de adaptarse continuamente a variaciones en la afluencia de pacientes, la disponibilidad de profesionales, la ocupación de espacios asistenciales y una correcta coordinación con otras áreas del hospital.

En este contexto, la saturación puede definirse como la situación en la que la demanda supera la capacidad operativa del servicio (recursos humanos, espacios asistenciales, circuitos y tiempos de atención), generando consecuencias como el aumento de los tiempos de espera, el incremento de pacientes pendientes de cama, la sobrecarga de los profesionales y dificultades para garantizar que todos los pacientes reciban una atención de calidad similar independientemente del momento en que llegan. *(Revista Emergencias, 2015)*

La predicción de la demanda asistencial permite abordar este problema desde una perspectiva preventiva. En lugar de limitarse a describir la actividad una vez que ya se ha producido, los modelos predictivos permiten estimar la afluencia esperada en un horizonte temporal futuro. Esta información puede ser útil para apoyar decisiones relacionadas con la planificación de turnos, preparar material o adelantarse a un momento de alta demanda antes de que ocurra.

[3] Servicio de Salud de Castilla-La Mancha. (2025, 19 de diciembre). *El Hospital Universitario de Cuenca toma hoy el relevo del Hospital Virgen de la Luz para seguir atendiendo la salud de los conquenses*.

Desde un punto de vista técnico, el problema consiste en estimar cuántos pacientes llegarán a urgencias en un determinado intervalo temporal. Para ello, pueden utilizarse diferentes fuentes de información, como el comportamiento histórico de la demanda, la hora del día, el día de la semana, los festivos, las condiciones meteorológicas o indicadores operativos del propio servicio. La combinación de estas variables permite construir modelos capaces de aprender patrones pasados y utilizarlos para anticipar la demanda futura. *(Zhang et al., 2022)*

3.3 Patrones temporales de la demanda

La demanda en urgencias no se comporta como un proceso aleatorio independiente, sino como una señal temporal con estructura. El número de llegadas puede representarse como una serie temporal de conteo, en la que cada observación representa el número de pacientes que llegan durante un intervalo concreto *(Evaluación de Modelos de Series Temporales Para la Previsión de la Demanda de Emergencias Sanitarias – Revista Emergencias, s. f.)*

En este tipo de series es habitual encontrar componentes sistemáticas:

- **Patrones intradiarios (por hora):** asociados a la distribución de llegadas a lo largo de las distintas horas del día. No todos los pacientes llegan a urgencias de forma uniforme a lo largo del día. Hay franjas horarias con mayor afluencia, influidas por los hábitos de la población, los horarios laborales, la disponibilidad de atención primaria. En cambio, luego también puede haber horas como la madrugada que suelen registrar una actividad notablemente menor.
- **Patrones semanales:** la demanda presenta diferencias significativas según el tipo de día. Los días laborables, los fines de semana y los festivos muestran comportamientos distintos, y estas variaciones se intensifican en puentes, días previos o posteriores a festivos y periodos vacacionales.
- **Patrones estacionales:** cambios por meses o estaciones (por ejemplo, periodos invernales con mayor carga respiratoria o periodos vacacionales con alteraciones en la movilidad).
- **Efectos de calendario:** festivos nacionales/regionales, puentes y periodos vacacionales.

- **Eventos exógenos:** existen picos de demanda asociados a eventos externos difíciles de anticipar, como episodios meteorológicos extremos, brotes epidémicos, accidentes o eventos locales. Estos picos representan el mayor reto para cualquier modelo predictivo, ya que no responden a patrones históricos estables y, en muchos casos, no dejan rastro en los datos disponibles

Este comportamiento implica que la demanda presenta autocorrelación, es decir, que lo que ocurre en una franja horaria está relacionado con lo que ha ocurrido en intervalos anteriores. Por ello, la predicción de la demanda en urgencias se aborda como un problema de forecasting, o predicción temporal, disciplina que tiene como objetivo estimar valores futuros de una variable a partir de su comportamiento pasado y variables contextuales relevantes. *(Serrano, 2022b)*

En este contexto, el intervalo temporal de la predicción marca una diferencia importante. Saber cuántos pacientes se esperan hora a hora resulta clave para gestionar el día a día del servicio: organizar los turnos, distribuir el personal y activar refuerzos cuando se anticipan picos de demanda. Disponer de predicciones para varios días, permite planificar con más antelación y tomar decisiones con mayor margen de tiempo.

3.4 Variables predictoras en urgencias

La predicción de la demanda asistencial puede abordarse desde dos enfoques principales. Por un lado, los modelos univariantes utilizan únicamente el historial de la propia serie, es decir, las llegadas pasadas de pacientes, para estimar valores futuros. Por otro lado, los modelos multivariantes incorporan variables adicionales que pueden ayudar a explicar la demanda, como el calendario, la meteorología, la situación operativa del servicio o determinados patrones históricos. *(Serrano, 2022ª; Rivas-Ruiz et al., 2021)*

La elección de las variables depende de tres factores: (i) con cuánta antelación se quiere predecir, (ii) qué información está disponible en ese momento y (iii) qué equilibrio se busca entre sencillez y precisión. Cuando se predice a corto plazo, es posible usar información reciente del propio servicio, como el número de pacientes llegados en las últimas horas. Sin embargo, cuando el usuario consulta una fecha futura concreta, esa información no existe todavía, por lo que el modelo debe apoyarse únicamente en

variables conocidas de antemano: el día de la semana, la hora, si es festivo, la previsión meteorológica y los patrones históricos de demanda.

3.4.1 Variables temporales y de calendario

Las variables temporales y de calendario son de las más útiles para predecir la afluencia en urgencias, ya que capturan los patrones que se repiten a lo largo del día, la semana y el año. Además, tienen la ventaja de que no requieren ninguna fuente externa: se obtienen directamente a partir de la fecha y hora de cada registro. *(Crisafulli et al., 2026)*

- Hora del día
- Día de la semana
- Mes
- La estación del año
- El trimestre
- La semana del año
- Los fines de semana
- Festivos, puentes o días previos y posteriores
- Periodo vacacional

Es habitual transformar variables como la hora del día o el día de la semana en una representación cíclica. Esto se debe a que, si se usan números enteros, el modelo podría interpretar que las 23:00 y las 00:00 están muy alejadas entre sí, cuando en realidad son horas consecutivas. Con esta transformación, el modelo entiende mejor que el tiempo es circular y no lineal, lo que mejora su capacidad para capturar patrones periódicos.

3.4.2 Variables clínicas y operativas

Las variables clínicas y operativas aportan información sobre el estado real del servicio en cada momento. En urgencias, la carga de trabajo no depende solo de cuántos pacientes llegan, sino también de cuántos están siendo atendidos, cuánto tiempo llevan esperando o qué recursos están ocupados en ese instante. *(Predicción En Tiempo Real de los Ingresos del Servicio de Urgencias Mediante un Modelo de Inteligencia Artificial y Su Integración En la Planificación de las Camas de Hospitalización – Revista Emergencias, s. f.-b)* Entre este tipo de variables se incluyen:

- Niveles de triaje
- Tiempos de espera
- Duración de estancia
- Proporción de ingresos
- Número de pruebas diagnósticas solicitadas
- Volumen de pacientes presentes en el servicio

Sin embargo, su uso depende de con cuánta antelación se quiere predecir. Cuando se predice a corto plazo, estas variables son muy útiles porque reflejan lo que está ocurriendo en el servicio en ese momento. En cambio, cuando se consulta una fecha futura, esta información simplemente no existe todavía, por lo que no puede utilizarse sin convertir el modelo en algo poco realista o inaplicable en la práctica.

3.4.3 Variables exógenas: meteorología y contexto

Las variables exógenas son aquellas que no forman parte del propio servicio de urgencias, pero que pueden influir en cuántos pacientes acuden. Entre ellas, la meteorología es una de las variables más relevantes. *(Burlotos et al., 2025)* También encontramos:

- Temperatura media, mínima y máxima
- Precipitación

- Humedad
- Velocidad del viento
- Horas de sol
- Presión atmosférica
- Olas de calor o frío

3.4.4 Variables históricas, retardos y medias móviles

Los retardos permiten que el modelo tenga en cuenta lo que ha ocurrido recientemente. Por ejemplo, saber cuántos pacientes llegaron en la hora anterior ayuda a detectar si la demanda está subiendo o bajando, mientras que el dato de la misma hora de la semana pasada permite identificar si ese comportamiento se repite de forma habitual. Las medias móviles complementan esta información resumiendo la actividad de las últimas horas o días en un único valor, lo que suaviza las fluctuaciones puntuales y facilita detectar tendencias más generales. *(Alenezi et al. 2026; Hu et al. 2023; Reboredo et al. 2023)* Encontramos:

- Medias móviles
- Retardos o lags
- Agregados históricos por patrón temporal

3.4.5 Calidad del dato: ruido, outliers y cambios de régimen

Los datos asistenciales suelen presentar problemas de calidad que deben abordarse antes de entrenar cualquier modelo predictivo. Entre los problemas más frecuentes se encuentran:

- Valores faltantes
- Valores duplicados
- Valores atípicos
- Registros inconsistentes

Por ello, una parte esencial del proceso es el preprocesamiento: donde problemas, como los valores atípicos no siempre deben eliminarse, ya que pueden reflejar picos reales de demanda provocados por situaciones excepcionales.

Por ello, es importante distinguir entre un error en el registro de datos y un episodio genuino de alta presión asistencial. De forma similar, cuando faltan datos, pueden estimarse a partir de valores cercanos, pero conviene dejar constancia de qué valores han sido imputados para no perder esa información. *(Syed et al. 2023; Shi et al. 2021)*

3.5 Series temporales y predicción

La predicción de la demanda en urgencias funciona como un problema de series temporales, ya que los datos están ordenados cronológicamente y el valor de cada registro está en una franja horaria que puede estar relacionada con valores anteriores. Por ello, el objetivo no consiste únicamente en estimar el número de llegadas, sino en aprovechar la estructura temporal de las series temporales para anticiparse.

En este tipo de problemas hay 3 conceptos clave que son relevantes:

- **Tendencia**: recoge cambios a largo plazo en el nivel medio de la demanda.
- **Estacionalidad**: representa ciclos repetitivos asociados a hora del día, días de la semana, meses o periodos concretos.
- **Componente irregular**: agrupa variaciones no explicadas por los patrones anteriores.

En el caso de urgencias, estos componentes aparecen de forma simultánea y se combinan con factores externos. *(Porto & Fogliatto, 2024; Alenezi et al., 2026)*

Otro concepto fundamental es la **autocorrelación**, mide la relación entre una observación y sus valores pasados. Si la demanda de una hora está relacionada con la demanda de horas anteriores, del día previo o de la misma hora de la semana anterior. Si existe esa relación, el modelo puede aprender de ella usando variables construidas a partir del historial de la serie. *(Alenezi et al. 2026)*

También es importante mencionar el **horizonte de predicción,** en una predicción a corto plazo pueden utilizarse datos recientes del propio servicio, como las llegadas de

las últimas horas, para estimar la hora siguiente. Pero para estimar la demanda de un día futuro concreto, solo puedes apoyarte en información que ya conoces de antemano: la hora, el día de la semana, si es festivo, la previsión meteorológica y los patrones históricos. *(Crisafulli et al., 2026).*

Por último, en los problemas de series temporales es fundamental evitar la **fuga de información tempora**l. Esta se produce cuando el modelo utiliza datos que no estarían disponibles en el momento real de realizar la predicción, lo que genera resultados artificialmente optimistas. Para evitarlo, tanto en la construcción de variables como en la validación de los modelos deben respetar el orden cronológico, entrenando con el pasado y evaluando sobre el futuro *(Porto & Fogliatto, 2024; Alenezi et al., 2026).*

3.6 Familias de modelos predictivos

La predicción de la demanda en urgencias puede abordarse mediante distintas familias de modelos, desde enfoques sencillos basados en patrones históricos hasta modelos estadísticos clásicos, algoritmos de machine learning y modelos de deep learning. Cada familia presenta ventajas y desventajas distintas, de forma que su elección dependerá directamente de la naturaleza de los datos y el objetivo de predicción.

Es por eso, por lo que el TFG, establece en primer lugar, modelos de referencia sencillos que permitan comprobar si cualquier modelo más complejo realmente mejora una predicción básica. Para seguir evaluando modelos estadísticos más flexibles capaces de incorporar variables temporales, meteorológicas, históricas y de calendario. Sin embargo, antes de explorar enfoques más sofisticados, es necesario establecer un punto de referencia claro: los modelos baseline.

3.6.1 Modelos base o baselines

Estos modelos proporcionan una referencia mínima de rendimiento y permiten valorar si un algoritmo más avanzado realmente está aportando una mejora predictiva. En problemas de demanda asistencial, son especialmente importantes porque muchas veces

la serie presenta patrones temporales fuertes que pueden ser capturados mediante reglas sencillas *(Tuominen et al., 2022; Alenezi et al., 2026).*

Modelo naive simple

Este tipo de modelo asume que el valor futuro de la serie será igual o muy parecido a un valor observado anteriormente. Es decir, no aprende patrones complejos ni utiliza variables externas, sino que toma como referencia inmediata el comportamiento más reciente de la demanda.

En una serie temporal sin estacionalidad, la predicción más simple sería tomar como estimación el último valor observado:

$$\hat{y}_{t+1} = y_t$$

Ecuación 1. Modelo naive simple.

Donde y^ t+1 representa la predicción para el siguiente instante temporal e yt representa el último valor observado.

Modelo naive estacional

En cambio, para la demanda de urgencias, es más adecuado usar la predicción tomando referencia del valor observado en el mismo momento del ciclo anterior. Usando un patrón semanal, puede utilizarse el valor registrado 168 horas antes (24 x 7 = 168 horas).

$$\hat{y}_t = y_{t-m}$$

Ecuación 2. Modelo naive estacional semanal aplicado a datos horarios.

Donde yt representa el número real de llegadas en la hora t, mientras que *yt−m (que sería el valor de 168)* representa el número de llegadas registrado en la misma hora de la semana anterior.

Media histórica

La predicción de la media histórica se obtiene calculando el promedio de las observaciones pasadas que pertenecen a una misma categoría temporal.

$$\hat{y}_{h,d} = \frac{1}{n_{h,d}} \sum_{i \in G(h,d)} y_i$$

Ecuación 3. Media histórica de llegadas por hora y día de la semana.

Donde h representa la hora del día, d el día de la semana, G(h,d) el conjunto de observaciones históricas pertenecientes a esa misma combinación temporal, y nh,d el número de observaciones incluidas en dicho grupo.

La media histórica tiene la ventaja de ser más estable que el modelo *naive*, ya que no depende de una única observación pasada, sino del comportamiento promedio observado en situaciones similares. Sin embargo, puede adaptarse peor a cambios recientes o situaciones excepcionales.

3.6.2 Modelos estadísticos clásicos

Los modelos estadísticos clásicos pertenecen a una de las familias de aproximaciones tradiciones para la predicción de series temporales. Su principal objetivo es definir y representar la dependencia de los datos, es decir, la relación estadística entre una observación y sus valores pasados. Además, permiten incorporar variables exógenas enriqueciendo el modelo. (*Porto & Fogliatto,2024; Petravić et al.,2025)*

Modelo ARIMA

Combina tres componentes: autoregresión, diferenciación y media móvil (errores pasados). La parte autorregresiva utiliza valores pasados de la serie, la diferenciación permite tratar tendencias o falta de estacionariedad, y la media móvil incorpora errores cometidos en predicciones anteriores.

$$ARIMA(p, d, q)$$

Ecuación 4. Notación general del modelo ARIMA.

Donde p indica el número de retardos autorregresivos, d el número de diferenciaciones aplicadas y q el número de errores pasados incluidos en el modelo.

Modelo SARIMA

SARIMA amplía el modelo ARIMA incorporando componentes estacionales. Esto permite modelar ciclos repetitivos, como los patrones diarios o semanales que pueden aparecer en la demanda asistencial.

$$SARIMA(p, d, q)(P, D, Q)_s$$

Ecuación 5. Especificación general del modelo SARIMA.

Donde (p,d,q) representa la parte no estacional, (P,D,Q) la parte estacional y s la longitud del ciclo.

Modelo SARIMAX

SARIMAX es una extensión de SARIMA que permite añadir variables exógenas. Estas variables pueden ser factores externos a la propia serie, como festivos, temperatura, precipitación o indicadores de calendario.

$$SARIMAX = SARIMA + X_t$$

Ecuación 6. Extensión del modelo SARIMA mediante variables exógenas.

Donde Xt representa el conjunto de variables externas utilizadas para mejorar la predicción.

Modelo Prophet

Prophet es un modelo diseñado para capturar tendencia, estacionalidades y efectos de calendario de forma flexible e interpretable. *(Yarubi et al., 2025)*

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t$$

Ecuación 7. Formulación aditiva del modelo Prophet para series temporales.

Donde g(t) representa la tendencia, s(t) las estacionalidades, h(t) los efectos de festivos o eventos especiales, y εt el error no explicado por el modelo.

Sin embargo, en problemas como la predicción horaria de urgencias, donde intervienen múltiples variables temporales, meteorológicas e históricas, estos modelos tienden a no ser suficientes para capturar relaciones no lineales. Por ello, también es habitual formular el problema con regresión supervisada.

3.6.3 Modelos de regresión y conteo

Los modelos de regresión temporal permiten transformar el problema de predicción de demanda en una tarea supervisada, donde la variable objetivo es el número de llegadas en una franja horaria concreta y las variables predictoras pueden incluir información temporal, meteorológica, de calendario e histórica. *(Jiang et al., 2022).*

Regresión Lineal

Estima la variable objetivo como una combinación lineal de las variables explicativas.

$$\hat{y}_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t}$$

Ecuación 8. Formulación general de la regresión lineal múltiple

Donde $\beta 0$ es el término independiente, $x1,t,x2,t,\ldots,xk$ son las variables predictoras, y $\beta 1,\beta 2,\ldots,\beta k$ representan el peso estimado de cada variable sobre la demanda.

Regresión de Poisson

Este modelo está diseñado para variables de conteo no negativas y estima la demanda esperada mediante una función exponencial.

$$\lambda_t = e^{\beta_0+\beta_1 x_{1,t}+\cdots+\beta_k x_{k,t}}$$

Ecuación 9. Modelo de regresión de Poisson con función de enlace logarítmica.

Donde λt representa el número esperado de llegadas en la franja horaria t. Esta formulación garantiza que las predicciones sean siempre positivas.

No obstante, la regresión de Poisson asume que la media y la varianza son similares. En la demanda asistencial esta condición puede no cumplirse, ya que suelen aparecer picos de afluencia y una variabilidad superior a la esperada *(Rostami-Tabar et al., 2024).*

3.6.4 Modelos de machine learning tabular

Los modelos de machine learning tabular abordan la predicción de demanda como un problema de aprendizaje supervisado: el modelo aprende a estimar el número de llegadas a urgencias en una franja horaria a partir de un conjunto de variables explicativas.

A diferencia de los modelos estadísticos clásicos o de regresión lineal, los modelos de machine learning pueden capturar relaciones no lineales e interacciones complejas entre variables. *(Porto & Fogliatto, 2024b)*

Árboles de decisión

Los árboles de decisión son modelos de aprendizaje supervisado que dividen los datos en grupos homogéneos mediante reglas sucesivas sobre las variables. En cada división se elige la variable y el umbral que mejor reduce el error de predicción. Este

proceso se repite de forma jerárquica hasta llegar a un nodo final, denominado hoja, donde se obtiene la predicción.

$$\hat{y}_t = \frac{1}{n_R} \sum_{i \in R} y_i$$

Ecuación 10. Predicción media en un nodo hoja del árbol de decisión.

Donde R representa la región final del árbol en la que cae la observación, y $n_R$es el número de observaciones históricas incluidas en esa región, y (yi) son los valores reales de la variable objetivo dentro de dicha región.

La principal ventaja de los árboles de decisión es que son modelos fáciles de interpretar, ya que su funcionamiento se puede entender como una secuencia de preguntas o reglas.

Además, los árboles de decisión permiten capturar relaciones no lineales entre las variables. Esto significa que no obligan a que el efecto de una variable sea siempre constante o proporcional.

Random Forest

El modelo combina múltiples árboles de decisión entrenados cada uno sobre una muestra distinta de los datos y con un subconjunto aleatorio de variables. Random Forest construye un conjunto de árboles independientes y combina sus predicciones para obtener un resultado más estable. Cada árbol se entrena sobre una muestra distinta del conjunto de datos, obtenida normalmente mediante muestreo con reemplazo.

La predicción final se obtiene haciendo un promedio de las predicciones individuales de todos los árboles. Al combinar múltiples árboles, Random Forest reduce la variabilidad que tendría un único árbol y produce predicciones más estables.

$$\hat{y}_t = \frac{1}{B} \sum_{b=1}^{B} f_b(X_t)$$

Ecuación 11. Predicción agregada de un modelo Random Forest.

Donde B es el número de árboles del bosque y $f_b(X_t)$representa la predicción del árbol b para el conjunto de variables (X_t). *(Tuominen et al., 2022)*

ExtraTrees

ExtraTrees, o Extremely Randomized Trees, es un modelo de aprendizaje supervisado basado en la combinación de múltiples árboles de decisión. Su funcionamiento es similar al de Random Forest, ya que ambos construyen varios árboles y combinan sus predicciones para obtener un resultado final más estable. Sin embargo, ExtraTrees introduce un mayor nivel de aleatoriedad en el proceso de construcción de los árboles.

Esto ayuda a reducir la varianza del modelo y lo hace más robusto frente al ruido, lo que puede resultar especialmente útil en problemas con muchas variables correlacionadas, como ocurre cuando se incorporan retardos, medias móviles y patrones históricos. No obstante, también presenta limitaciones. Al estar formado por muchos árboles, su interpretación es menos directa que la de un árbol de decisión individual.

Gradient Boosting

Los modelos de Gradient Boosting también combinan múltiples árboles de decisión, pero a diferencia de Random Forest, lo hacen de forma secuencial: cada nuevo árbol aprende de los errores cometidos por los anteriores, corrigiéndolos de forma progresiva. La predicción final se construye como una suma progresiva de modelos simples:

$$\hat{y}_t = \sum_{m=1}^{M} \eta f_m(X_t)$$

Ecuación 12. Predicción acumulativa en Gradient Boosting.

Donde M el número de árboles, $f_m(X_t)$representa el árbol añadido en la iteración m, y η es la tasa de aprendizaje, que controla cuánto aporta cada árbol al modelo final.

Este enfoque suele obtener buenos resultados en problemas tabulares porque permite capturar relaciones no lineales e interacciones complejas entre variables.

HistGradientBoosting

HistGradientBoosting es una variante eficiente de Gradient Boosting implementada en scikit-learn. Al igual que otros modelos de boosting, construye árboles de decisión de forma secuencial, de manera que cada nuevo árbol intenta corregir parte de los errores cometidos por los árboles anteriores.

Su principal característica es que transforma las variables continuas en intervalos o histogramas. En lugar de buscar puntos de corte exactos sobre todos los valores posibles de una variable, agrupa los valores en rangos y trabaja sobre esos grupos. Esta estrategia permite acelerar el entrenamiento y reducir el coste computacional, especialmente en conjuntos de datos medianos o grandes.

LightGBM

LightGBM es otro algoritmo basado en *Gradient Boosting*, especialmente diseñado para ser eficiente con grandes volúmenes de datos y muchas variables. LightGBM utiliza estrategias de crecimiento de árboles optimizadas que reducen el tiempo de entrenamiento y mejorar el rendimiento en problemas tabulares.

Una de sus principales ventajas es su capacidad para capturar relaciones no lineales e interacciones entre variables sin necesidad de especificarlas manualmente.

Además, LightGBM puede adaptarse a diferentes funciones de pérdida. Esto resulta muy útil en problemas donde la variable objetivo es un conteo, combinándola con una formulación basada en Poisson, se convierte más coherente con la naturaleza discreta y

no negativa de las llegadas de pacientes. También puede emplearse para estimar cuantiles, permitiendo construir intervalos de predicción en lugar de ofrecer únicamente una predicción puntual.

Modelos ensemble

Los modelos ensemble combinan las predicciones de varios modelos con el objetivo de obtener una predicción más estable. La idea es que distintos algoritmos pueden capturar aspectos diferentes del problema, por lo que su combinación puede reducir errores individuales y mejorar la generalización.

Una forma sencilla de ensemble consiste en calcular la media de las predicciones de varios modelos:

$$\hat{y}_t = \frac{1}{K} \sum_{k=1}^{K} \hat{y}_{k,t}$$

Ecuación 13. Predicción promedio de un modelo ensemble.

Donde K es el número de modelos combinados y $\hat{y}_{k,t}$ representa la predicción del modelo k en el instante t. (Peláez-Rodríguez et al.,2024*)*

3.6.5 Modelos de deep learning

Los modelos de deep learning ofrecen un enfoque diferente para abordar la predicción de series temporales mediante redes neuronales capaces de aprender patrones complejos directamente a partir de los datos. Estas técnicas no requieren definir manualmente las relaciones entre variables, sino que aprenden automáticamente patrones a partir del historial de la serie y de las variables disponibles.

En problemas de predicción de demanda resultan especialmente útiles cuando existen dependencias temporales, relaciones no lineales o grandes volúmenes de datos.

Dentro de esta familia, las arquitecturas más empleadas son LSTM y GRU. *(Zhao et al., 2022; Ming et al., 2025)*

LSTM (Long Short-Term Memory)

Las redes LSTM fueron desarrolladas para resolver las limitaciones de las RNN tradicionales. Su principal característica es la incorporación de mecanismos de memoria y compuertas internas que permiten decidir qué información debe mantenerse, actualizarse o descartarse a lo largo de la secuencia temporal.

Gracias a esta estructura, las LSTM pueden capturar dependencias tanto a corto como a medio plazo. En problemas de predicción horaria, suelen trabajar utilizando ventanas temporales de observaciones anteriores como entrada. Para el caso del TFG, el modelo aprende automáticamente la relación entre las observaciones previas y la predicción futura.

$$\hat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-n})$$

Ecuación 14. Formulación general de una red LSTM para predicción de series temporales.

Donde $\hat{y}_t$ representa la predicción en el instante temporal t; $y_{t-1}, y_{t-2}, \ldots, y_{t-n}$ corresponden a observaciones previas de la serie temporal utilizadas como secuencia de entrada; y f(·) representa la función no lineal aprendida por la red LSTM.

GRU (Gated Recurrent Unit)

Las redes GRU, o *Gated Recurrent Unit*, son una variante de las redes neuronales recurrentes diseñadas para trabajar con datos secuenciales. Al igual que las LSTM, permiten conservar información de observaciones anteriores y utilizarla para realizar predicciones futuras, lo que las hace adecuadas para problemas de series temporales.

La principal diferencia respecto a las LSTM es que las GRU presentan una estructura más sencilla. Mientras que las LSTM utilizan varios mecanismos internos para decidir qué información conservar, actualizar u olvidar, las GRU reducen esta

complejidad utilizando un menor número de compuertas. En concreto, suelen trabajar con una compuerta de actualización y una compuerta de reinicio, lo que permite controlar qué parte de la información pasada se mantiene y qué parte se modifica con la nueva observación.

3.6.6 Comparación y criterios de selección

La comparación entre familias de modelos permite analizar qué tipo de enfoque se adapta mejor a la naturaleza del problema de predicción de demanda en urgencias. Cada familia presenta ventajas y limitaciones diferentes en términos de precisión, interpretabilidad, complejidad y capacidad para trabajar con variables temporales.

La selección del modelo más adecuado no debe basarse únicamente en obtener el menor error promedio. Para el entorno sanitario también resulta importante considerar otros aspectos, como la estabilidad del modelo, su capacidad para evitar grandes errores en picos de demanda, el sesgo de las predicciones y la posibilidad de utilizar el modelo en situaciones reales de planificación.

Para evaluar el rendimiento de los modelos se utilizan métricas de error como MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) y BIAS (Sesgo del modelo).

MAE

El MAE mide el error medio absoluto entre las predicciones realizadas por el modelo y los valores reales observados. Se calcula como la media de las diferencias absolutas entre ambos valores.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Ecuación 15. Error Absoluto Medio (MAE)

y_i: representa el valor real observado

$\hat{y}_i$: representa la predicción realizada por el modelo

n: es el número total de observaciones

La principal ventaja del MAE es que su interpretación es directa, ya que se expresa en las mismas unidades que la variable objetivo.

RMSE

Mide la diferencia entre predicciones y valores reales, pero introduce una diferencia importante respecto al MAE: eleva los errores al cuadrado antes de calcular la media.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Ecuación 16. Raíz del Error Cuadrático Medio (RMSE).

y_i: representa el valor real observado

$\hat{y}_i$: representa la predicción realizada por el modelo

n: es el número total de observaciones

Al elevar los errores al cuadrado, los errores grandes tienen un peso mucho mayor en el resultado final. Esto significa que el RMSE penaliza especialmente las predicciones muy alejadas de la realidad.

BIAS (Sesgo del modelo)

Mide si el modelo tiende a sobreestimar o infraestimar sistemáticamente. Permite conocer la dirección del error, es decir, si las predicciones suelen quedar por encima o por debajo de los valores reales.

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$$

Ecuación 17. Sesgo medio de predicción (BIAS).

y_i: representa el valor real observado

$\hat{y}_i$: representa la predicción realizada por el modelo

n: es el número total de observaciones

La interpretación del **BIAS** depende de su signo:

- **BIAS positivo** indica que el modelo tiende a sobreestimar.
- **BIAS negativo** indica que el modelo tiende a infraestimar.
- **Valor cercano** a cero indica ausencia de sesgo sistemático.

3.7 Validación temporal y métricas de evaluación

Es importante recalcar que, en series temporales, no es suficiente calcular las métricas de error: el proceso de validación también debe respetar el orden cronológico de los datos.

A diferencia del aprendizaje supervisado tradicional, donde dividir los datos aleatoriamente entre entrenamiento y prueba es una práctica habitual, en este contexto esa estrategia no es válida, ya que podría introducir información futura en el entrenamiento y comprometer la integridad de la evaluación.

Por este motivo, la validación debe realizarse utilizando divisiones temporales:

- Periodo inicial para entrenamiento
- Periodo posterior para validación
- Bloque final independiente para prueba.

De este modo, el modelo aprende únicamente a partir del pasado y evalúa sobre datos futuros, consiguiendo una situación más cercana al uso real del sistema predictivo.

Además, con este enfoque evitamos la fuga de información temporal, que ocurriría si el modelo utiliza datos que en la práctica no estarían disponibles en el momento de predecir. Esto es importante porque, de ocurrir, los resultados parecerían mejores de lo que realmente son, llevando a conclusiones erróneas sobre el rendimiento del modelo.

4. METODOLOGÍA

Este apartado explica cómo se ha organizado el trabajo. Se describen las decisiones tomadas en cada fase: qué datos se han usado, cómo se han preparado, qué variables se han generado, qué modelos se han evaluado y cómo se ha validado el resultado. El objetivo es que quede claro por qué se ha seguido este camino y no otro.

4.1 Enfoque metodológico del trabajo

El presente TFG plantea un proyecto orientado a resolver un problema real del entorno sanitario: la predicción de la afluencia de pacientes en un servicio de urgencias hospitalarias. El objetivo principal es construir un sistema predictivo capaz de anticipar la demanda futura hora a hora, con el fin de servir de apoyo a la planificación de recursos asistenciales.

El trabajo se basa en el análisis de datos históricos reales, sin intervenir en el hospital ni manipular ninguna variable. Es por eso, que se descarta un enfoque de investigación cualitativo, como entrevistas al personal sanitario u observación directa, porque el objetivo es medir y predecir la demanda, no comprender percepciones ni procesos internos. Por ello, la metodología se apoya en registros ya existentes, combinando información asistencial con variables de calendario, temporales y meteorológicas, para identificar patrones históricos y entrenar modelos capaces de anticipar la afluencia futura.

El proyecto trabaja con **tres niveles** de datos. El primero es el **registro individual** de cada visita al servicio de urgencias. El segundo es una **serie temporal horaria**, construida a partir de esos registros y enriquecida con información de calendario, festivos y meteorología. El tercero es el dataset final de entrenamiento**,** que incorpora **variables históricas** calculadas siempre a partir de datos pasados para evitar el uso de información futura durante el modelado.

La unidad de análisis se define en torno a la hora. Esto permite capturar diferencias entre franjas del día, turnos, festivos y variaciones, identificando así momentos concretos de mayor presión asistencial que una agregación diaria no permitiría.

La variable principal del estudio es el número de llegadas al servicio de urgencias por hora. Porque para predecir la demanda, el sistema aprenderá de los datos del pasado, información histórica, del día y la hora, las condiciones meteorológicas y la actividad reciente del servicio, de esta manera el modelo podrá ser capaz de estimar cuántos pacientes llegarán en la siguiente hora.

Este enfoque resulta ser el más adecuado para el trabajo porque permite hacer tres cosas: analizar cómo ha evolucionado la demanda en el pasado, construir modelos que puedan evaluarse de forma objetiva, y trasladar los resultados a una herramienta útil para la planificación del servicio.

4.2 Estrategias metodológicas consideradas

A la hora de definir la metodología final del TFG, se valoraron distintas formas de abordar el problema, tanto desde el punto de vista del análisis de datos como desde el enfoque de desarrollo del proyecto. Esta comparación permitió elegir el enfoque más adecuado para el objetivo del trabajo y descartas otras alternativas menos adecuadas.

Entre las posibles opciones, se estudió el análisis descriptivo de la demanda histórica. Esta estrategia permitiría identificar patrones generales como las horas de mayor afluencia, las diferencias entre días laborables y festivos, o cómo varía la actividad a lo largo del año. Sin embargo, se descartó, ya que conocer lo que ha pasado en el pasado no es suficiente para predecir lo que ocurrirá en el futuro. El análisis descriptivo sí que se incorpora como una fase dentro de la metodología elegida, ya que ayuda a comprender los datos antes de modelarlos, pero no establece el objetivo final del trabajo.

La segunda alternativa que se planteó consistía en aplicar directamente modelos estadísticos o de aprendizaje automático, como ARIMA o Holt-Winters, sobre los datos originales del hospital. Y aunque este enfoque tenga la ventaja de ser simple y de estar bien establecido para problemas de predicción temporal, se descartó como estrategia principal porque estos modelos no permiten incorporar fácilmente variables externas. Por lo que se decidió utilizarlos como modelos de referencia o baseline, pero no como enfoque central del trabajo.

Es por eso por lo que, viendo las diferentes metodologías, se acabó desarrollando una metodología incremental basada en el ciclo de vida del dato. Esto permite transformar los datos originales paso a paso en una serie temporal horaria, construir distintas versiones del dataset y comparar los modelos de forma ordenada.

4.3 Diseño general del flujo de trabajo

Para el desarrollo del proyecto se diseñó un flujo de trabajo estructurado en nueve fases sucesivas que transforman los registros originales del hospital en un sistema capaz de estimar la demanda futura.

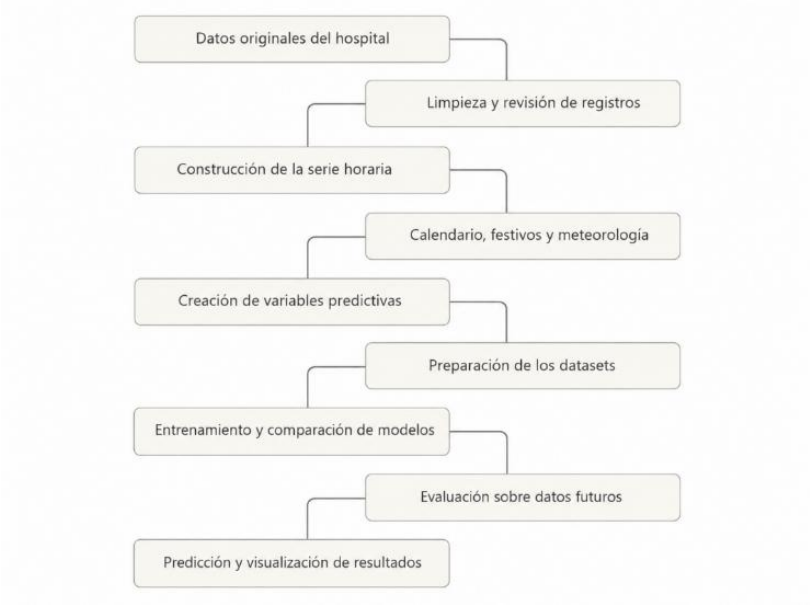

Figura 2. Flujo de las fases metodológicas seguidas en el proyecto.

La primera fase consiste en revisar los datos del hospital para identificar las variables más relevantes, verificar la coherencia de las fechas y detectar valores faltantes o errores. En la segunda fase los registros individuales se agregan por franjas horarias, transformando el dataset de visitas en una serie temporal.

A la serie horaria se le añade información externa procedente de tres fuentes: datos meteorológicos de AEMET, festivos locales, autonómicos y nacionales, y variables de calendario como el día de la semana o el mes. Se generan variables temporales, retardos de la demanda y medias móviles de la actividad reciente del servicio. Todas ellas se calculan usando exclusivamente datos del pasado. Con todo esto se entrenan y comparan los distintos modelos predictivos respetando el orden cronológico de los datos.

Donde finalmente el flujo termina con la generación de predicciones futuras y su presentación en un panel visual, con el objetivo de que los resultados sean interpretables y útiles para la planificación del servicio.

4.4 Fuentes de datos y periodo de análisis

El sistema predictivo combina tres fuentes de datos complementarias: los registros asistenciales del hospital, información de festivos y variables meteorológicas. Esta combinación permite construir un dataset que recoge la actividad del servicio, y además también factores externos con gran influencia sobre la demanda.

La fuente principal son los episodios de urgencias anonimizados proporcionados por el hospital, donde cada registro representa una visita al servicio. Estos datos son la base asistencial del trabajo que le permite construir la variable objetivo y variables derivadas. Las dos fuentes externas son el registro de festivos locales, autonómicos y nacionales de Cuenca, y los datos meteorológicos diarios de AEMET, que incluyen temperatura, precipitación, humedad, viento y presión atmosférica.

A partir de los registros hospitalarios originales se construye una serie temporal horaria, en la que cada observación representa una hora del periodo analizado. Esta transformación es necesaria porque el objetivo del trabajo es anticipar la demanda asistencial para intervalos horarios, no analizar episodios individuales de forma aislada.

El periodo del que se tiene información va desde el **01/12/2021 hasta el 30/11/2025**, lo que garantiza tener varios años de actividad para poder capturar patrones estacionales, mensuales, semanales y horarios. El dataset contiene **34.894 observaciones horarias**. La siguiente tabla resume las fuentes de datos originales empleadas en el TFG:

Fuente	Nivel	Uso dentro del proyecto
Registros Hospitalarios	Episodios	Fuente principal para construir las variables
Festivos de Cuenca	Día	Identificación de festivos, pre-festivos o post-festivos

Meteorología	Día	Información de temperaturas, precipitaciones, humedad, viento, sol y presión

Tabla 1. Fuentes de datos originales utilizadas en el proyecto.

A partir de estas fuentes se construyeron cuatro versiones sucesivas del dataset de modelado, incorporando variables de forma progresiva:

Dataset	**Filas**	**Variables**	**Periodo**	**Propósito**
Dataset_predictivo.csv	34.894	61	08/12/2021–30/11/2025	Primera versión predictiva
Dataset_predictivo_v2.csv	34.894	65	08/12/2021–30/11/2025	Versión enriquecida con información histórica
Dataset_predictivo_v3.csv	34.222	78	05/01/2022–30/11/2025	Versión avanzada final de modelos
Dataset_predictivo_v4.csv	34.894	102	08/12/2021–30/11/2025	Versión con todas las variables, para uso de predicción futura

Tabla 2. Resumen total de datasets utilizados en el proyecto.

En lugar de partir de un único conjunto cerrado, se construyeron versiones sucesivas que incorporan información adicional con posible valor predictivo.

4.5 Estrategia de modelado predictivo

La estrategia de modelado se planteó de forma progresiva, comenzando por modelos sencillos de referencia y avanzando hacia modelos más complejos. Esta decisión permite comprobar si el aumento de complejidad aporta una mejora real frente a alternativas básicas y evita seleccionar un modelo únicamente por sofisticación técnica.

En primer lugar, se utilizaron modelos de referencia o baselines, basados en patrones temporales simples, como repetir la demanda observada el día o la semana anteriores en la misma franja horaria. Estos modelos sirven como punto mínimo de comparación.

A continuación, se valoraron modelos estadísticos de series temporales y modelos de aprendizaje automático supervisado. Los modelos estadísticos permiten representar tendencia y estacionalidad, mientras que los modelos de machine learning resultan adecuados para trabajar con múltiples variables explicativas, como calendario, festivos, meteorología, retardos históricos y variables derivadas de la actividad reciente del servicio.

La elección final del enfoque de modelado no se basó únicamente en obtener el menor error sobre datos históricos, sino también en su posible uso operativo. Por ello, se priorizaron modelos capaces de generar predicciones futuras con variables disponibles o estimables en el momento de la consulta, evitando el uso de información que solo se conocería después. Este criterio es fundamental para que el sistema pueda aplicarse a escenarios reales de planificación hospitalaria.

Por eso, el modelo final debía mejorar a los baselines de referencia, comportarse de forma estable al predecir periodos no vistos durante el entrenamiento, y que visualmente facilite la interpretación de los resultados.

4.6 Diseño experimental y validación temporal

El diseño experimental del proyecto se definió teniendo en cuenta la naturaleza temporal de los datos. Al tratarse de una serie horaria de demanda asistencial, es importante no dividir los datos de forma aleatoria, ya que esto mezclaría observaciones

pasadas y futuras y podría generar una evaluación artificialmente optimista del rendimiento del modelo.

Por este motivo, es importante plantear la validación respetando el orden cronológico de la serie. Siendo así que, los modelos se entrenan sobre periodos históricos y se evalúan sobre periodos futuros, simulando las condiciones reales de uso del sistema.

La división se organizó en dos niveles. En primer lugar, se reservó el periodo correspondiente a enero-noviembre de 2025 como conjunto de test final. Este bloque no se utiliza durante el entrenamiento ni durante la selección de modelos, sino únicamente para la evaluación final. En segundo lugar, sobre el conjunto de entrenamiento, formado por los datos anteriores a 2025, se aplicó una validación temporal mediante TimeSeriesSplit con tres folds. Este procedimiento permite comparar configuraciones y modelos utilizando solo información pasada en cada partición.

De este modo, la evaluación mide no solo si el modelo funciona bien sobre datos pasados, sino si es capaz de predecir correctamente en condiciones similares a las de su uso real.

4.7 Métricas y criterios de evaluación

Para la evaluación de los modelos mediremos mediante métricas adecuadas para un problema de predicción numérica sobre conteos horarios de pacientes. El objetivo no era valorar únicamente si el modelo seguía la tendencia general de la demanda, sino medir de forma clara la diferencia entre las llegadas reales y las llegadas predichas.

La métrica principal seleccionada fue el **MAE** (*Mean Absolute Error*), ya que expresa el error medio en la misma unidad que la variable objetivo: número de pacientes. Esta métrica permite conocer cuántas llegadas de diferencia comete el modelo, de media, en cada franja horaria.

También se utilizó el **RMSE** (*Root Mean Squared Error*), que penaliza con mayor intensidad los errores grandes. Esta métrica resulta relevante porque, en un servicio de urgencias, los errores elevados pueden tener mayor impacto operativo, especialmente en momentos de alta presión asistencial.

Además, se analizó el **bias** o sesgo medio de predicción, con el objetivo de comprobar si el modelo tiende sistemáticamente a sobreestimar o infraestimar la

demanda. Este análisis es importante porque un modelo que infraestima de forma recurrente podría ser menos útil para la planificación de recursos, aunque su error medio global fuera aceptable.

Con estas métricas se valorará el rendimiento del modelo, buscando errores grandes, precisión general o posibles tendencias.

4.8 Recursos utilizados y reproducibilidad

Para el desarrollo del proyecto se han utilizado las siguientes herramientas:

- **Python**: lenguaje del TFG.
- **Pandas y numpy**: para la manipulación y preparación de los datos.
- **Scikit-learn:** librería principal.
- **XGBoost y LightGBM:** modelos de aprendizaje automático basados en gradient boosting para datos tabulares.

La organización del proyecto se planteará de forma modular, separando las distintas fases del flujo de trabajo: carga de datos, limpieza, construcción de la serie temporal, incorporación de fuentes externas, generación de variables, entrenamiento de modelos, evaluación y visualización. Esta separación permite revisar cada fase de forma independiente y revisar errores o ir modificando en cada fase.

4.9 Salida operativa: predicción futura y dashboard

El trabajo además contempla una salida operativa orientada a que las predicciones generadas sean útiles para apoyar la planificación del servicio de urgencias.

Para ello, el sistema generará predicciones futuras a partir de una fecha seleccionada. Esto exige que el modelo utilice únicamente variables disponibles o estimables en el momento de la consulta, como información de calendario, festivos, patrones históricos y previsión meteorológica. De esta forma, se evita depender de datos que solo se conocerían una vez transcurrida la hora predicha.

Las predicciones no se presentan como valores exactos sino como estimaciones orientativas acompañadas de un rango de incertidumbre, calculado a partir del error histórico del modelo en el periodo de validación. Esto permite que el usuario pueda interpretar la predicción sabiendo su fiabilidad aproximada.

La visualización de resultados se plantea mediante un dashboard o panel interactivo, donde puedan consultarse la demanda histórica, observar picos de afluencia o detectar patrones. Este formato ayuda a facilitar que los resultados técnicos del modelo sean comprensibles para perfiles no necesariamente especializados.

El panel actúa como puente entre el modelo predictivo y la toma de decisiones hospitalaria. Su finalidad es ayudar a identificar franjas horarias de mayor presión asistencial, anticipar posibles incrementos de demanda y ofrecer una base objetiva para la planificación de recursos.

5. DESARROLLO CIENTÍFICO-TÉCNICO

En el siguiente apartado se explica la metodología seguida en el TFG, en este punto se detalla cómo se ha llevado a cabo técnicamente cada fase del trabajo, desde la preparación de los datos hasta el entrenamiento, validación y evaluación de los modelos predictivos.

Este apartado permite comprender cómo se ha pasado de un conjunto de registros hospitalarios originales a un sistema completo de análisis, predicción y visualización de la demanda horaria en urgencias.

5.1 Visión general del desarrollo

Para el desarrollo del trabajo se han creado cinco fases secuenciales, que van desde la cesión de los datos por parte del hospital hasta la evaluación comparativa de los modelos predictivos. Las fases se dividen en:

1. Diseño del estudio y obtención de las fuentes de datos.
2. Construcción del dataset a través de seis pasos para su mejor estructura.
3. Ingeniería de características adicionales y eliminación justificada de observaciones,

4. Entrenamiento y optimización de cuatro familias de modelos sobre los datos.
5. Evaluación comparativa frente a baselines y análisis de los resultados.

De esta forma, cada fase se documenta detallando la metodología seguida, los procedimientos que se han aplicado, qué instrumentos se han empleado y las decisiones tomadas.

Mencionar que a lo largo de todo el desarrollo ha habido tres principios que se han seguido con el fin de mantener siempre la veracidad y la autenticidad de los datos y resultados.

• **No invención**: Todas las variables del dataset proceden de fuentes verificables: datos cedidos por el Hospital Virgen de la Luz, calendarios oficiales (*Ministerio de Trabajo, Junta de Comunidades de Castilla-La Mancha, Ayuntamiento de Cuenca*) y datos abiertos de la Agencia Estatal de Meteorología (*AEMET*). Ningún campo faltante se rellena con valores ficticios sin marcar explícitamente que se trata de una imputación.

• **Trazabilidad**. Siempre que se ha rellenado algún dato o se ha creado una nueva variable, se ha documentado el proceso. De esta forma, se puede saber qué datos son originales, cuáles han sido estimados y cómo se han generado las variables derivadas.

• **Ausencia de fuga de información (data leakage).** Las variables operativas observadas en el instante t (ocupación, ingresos, métricas de triaje) solo se incorporan al modelo predictivo a través de sus retardos (t−1, t−24, t−168 horas), de modo que el modelo nunca recibe información del presente o del futuro al predecir t+1.

5.2 Diseño del estudio

Con una visión de cómo se ha estructurado el TFG, en este apartado busca explicar el diseño del estudio.

5.2.1 Tipo de estudio y formulación del problema

El presente trabajo se plantea como un estudio observacional retrospectivo sobre todos los episodios atendidos en el Servicio de Urgencias del Hospital, entre el 1 de diciembre de 2021 y el 30 de noviembre de 2025, destacando un total de 1.461 días

naturales y 254.025 episodios individuales. A partir de esos registros se construye una serie temporal con resolución horaria y donde se busca estimar, para cada hora *t*, el número de llegadas que se producirán en la hora siguiente *t + 1*.

A cada hora se le agregan las variables con la información disponible en ese momento, creando un vector de características $\boldsymbol{X}_t$ con festivos, pre-festivos, hora del día, día de la semana, condiciones meteorológicas y llegadas pasadas junto con sus retardos y medias móviles. La variable objetivo que se usa para el TFG representa el número de pacientes que llegarán durante el intervalo *[t, t+1]*

El objetivo del modelo es aprender a estimar las llegadas de la hora siguiente a partir de la información disponible en la hora actual, expresado como $f(\boldsymbol{X}_t) \approx y_{t+1}$. Para que esta estimación sea válida en un entorno real, $\boldsymbol{X}_t$ solo puede contener variables ya conocidas en el momento de predecir, como la hora del día, los festivos, la meteorología o las llegadas de horas anteriores, pero nunca información generada dentro del propio intervalo que se quiere predecir.

5.2.2 Población, periodo y unidad de análisis

Los datos no se analizarán de forma individual para cada visita, sino que se agrupan por horas. Así, la unidad de análisis es cada hora del periodo estudiado. Cada hora tiene asociado su propio conjunto de variables y el número de llegadas de la hora siguiente, que es lo que el modelo intenta predecir.

Se eligió la hora como unidad de análisis por tres razones. La primera es que encaja con cómo funciona el servicio: el personal se organiza por turnos y las decisiones de refuerzo se toman por franjas horarias. La segunda es que hay patrones claros a lo largo del día, como el pico de mañana o el valle de madrugada, que se perderían si los datos se resumieran por días y la tercera es que agrupar por horas genera casi 35.000 observaciones, que permite entrenar los modelos con solidez.

5.3 Instrumentos y entorno tecnológico

Todo el desarrollo se ha realizado en Python 3 sobre un ordenador personal con Windows. La siguiente tabla resume las librerías empleadas y el papel que cumple cada una dentro del proyecto.

Librería	Versión / familia	Rol en el desarrollo
Pandas	Manipulación tabular	Carga, filtrado, agregación horaria, joins entre dataset hospitalario, festivos y meteorología
NumPy	Cálculo numérico	Operaciones vectorizadas, generación de codificaciones cíclicas, soporte de pandas
scikit-learn	ML clásico y utilidades	TimeSeriesSplit, métricas (MAE, RMSE), preprocesado y baselines
Statsmodels	Modelos estadísticos	SARIMAX (modelo 2) sobre la serie horaria
Prophet	Forecasting	Modelo 3 con doble estacionalidad (diaria y semanal)
LightGBM	Gradient Boosting	Modelo 4: árbol de decisión con boosting por gradiente
Optuna	Optimización bayesiana	Búsqueda de hiperparámetros de LightGBM
PyTorch	Redes neuronales	Modelo 5: LSTM con MC-Dropout para incertidumbre

matplotlib / seaborn	Visualización	Histogramas LOS, predicción vs real, MAE por hora, feature importance
openpyxl	Lectura de Excel	Lectura del fichero original del hospital (.xlsx, 15 MB)

Tabla 3. Librerías empleadas en el proyecto

5.4 Fuentes de datos

Como se ha mencionado anteriormente, el sistema predictivo combina tres fuentes de datos: los registros asistenciales del hospital, un calendario de festivos y datos meteorológicos de AEMET. En este apartado definiremos cada una.

5.4.1 Datos hospitalarios

Los datos basados en información del Hospital han sido facilitados en un fichero Excel (DatosOriginalesHospital.xlsx) con 254.025 registros, uno por cada episodio atendido en el Servicio de Urgencias durante el periodo de estudio.

Variable original	Tipo	Descripción
ID	Entero	Identificador único del episodio (sin valor fuera del fichero)
ACTO_CLINICO_FECHA_INICIO	Fecha-hora	Momento de inicio del episodio (= llegada del paciente)
ACTO_CLINICO_FECHA_FIN	Fecha-hora	Momento de finalización del episodio

EDAD	Texto	Rango etario en decenas (0-10, 11-20, …, >90)
SEXO	Texto	H, M, I (intersexo), O (otro)
PROCEDENCIA	Texto	Origen del paciente (iniciativa propia, AP, 112, etc.)
MOTIVO_CONSULTA	Texto	Motivo declarado de consulta
TRIAJE_CLASIFICACION	Texto	Nivel de triaje I-V (Manchester o equivalente)
TRIAJE_FECHA_HORA_INICIO	Fecha-hora	Momento de realización del triaje
ACTO_CLINICO_TIEMPO_ESTANCIA	Entero	Duración del episodio en minutos
ACTO_CLINICO_ESTADO	Texto	Estado del episodio (constante = CERRADO)

ACTO_CLINICO_DESTINO_FINAL	Texto	Destino final (domicilio, hospitalización, exitus, etc.)

Tabla 4. Variables originales del dataset hospitalario.

Además, de los datos sobre urgencias, el hospital facilitó datos sobre ocupación real y distribución de personal por horas.

5.4.2 Datos capacidad

Los datos correspondientes a la capacidad del hospital son los siguientes:

Recurso	**Valor declarado**
Boxes de urgencias	9 (capacidad constante)
Médicos por turno mañana y tarde	10
Médicos por turno noche	5 (asumido ya que el valor real era 5-6)
Enfermeras por turno	8
Camas hospitalarias	363
Radiología operativa	Sí (siempre)

Laboratorio operativo	Sí (siempre)

Tabla 5. Capacidad del Servicio de Urgencias.

Ante la indeterminación 5-6 médicos en turno de noche se ha optado por el valor 5. No se ha promediado a 5,5 para no introducir un valor inexistente en la operativa real. El horario establecido tal y como SESCAM ha informado consta: mañana 8:00-15:00, tarde 15:00-22:00, noche 22:00-8:00.

5.4.3 Datos meteorológicos (AEMET)

Todas las variables meteorológicas se han descargado del portal AEMET de OpenData correspondientes a la estación 8096 (Cuenca, altitud 949 m), para el mismo periodo correspondiente con el del dataset hospitalario.

Todo ha quedado unido en un dataset final (DatosAEMET.csv) que contiene 1.461 registros diarios con las siguientes magnitudes:

- Temperatura media, mínima y máxima
- Precipitación en milímetros y código “Ip” para precipitación inapreciable
- Velocidad media y racha máxima de Viento
- Horas de sol
- Presión atmosférica máxima y mínima
- Humedad relativa media, máxima y mínima.

Al incorporar los datos a través de la AEMET, los datos gratuitos y libres solo se publican de forma diaria no por horas, lo que limita la capacidad del modelo para capturar de forma más rigurosa los efectos del tiempo.

5.4.4 Calendario de festivos de Cuenca

El calendario de festivos se ha construido manualmente como fichero CSV (FestivosCuenca.csv) con los 68 días festivos correspondientes al periodo 2021-2025 en la ciudad de Cuenca. Para cada festivo se registran la fecha, el motivo (por ejemplo, "Día de la Constitución Española", "San Mateo") y el tipo de festivo (Nacional, Autonómico, Regional o Local). Las fuentes consultadas son los calendarios laborales oficiales del Ministerio de Trabajo, la Junta de Comunidades de Castilla-La Mancha y el Ayuntamiento de Cuenca.

5.4.5 Variables descartadas por no disponibilidad

La documentación del hospital incluye una serie de variables relativas al estado operativo y al contexto sanitario que no han sido proporcionadas en el fichero cedido. Por coherencia con el principio de no invención, estas variables no se incluyen en el modelo y quedan documentadas como vías de mejora para iteraciones futuras.

Variable no disponible	Tratamiento aplicado
Boxes disponibles en cada hora	Excluida; se usa la capacidad nominal fija = 9
Camas hospitalarias disponibles	Excluida; solo se conoce el total (363)
Radiología operativa	Excluida; se asume operativa siempre
Laboratorio operativo	Excluida; se asume operativo siempre
Activación de planes de contingencia	Excluida; no proporcionada
Declaración de epidemias / olas víricas	Excluida; reconstruible desde fuentes públicas en futuras iteraciones
Cierres o restricciones de atención primaria	Excluida; no proporcionada
Periodos vacacionales escolares	Excluida; reconstruible desde calendario JCCM si se desea ampliar

Tabla 6. Variables documentadas, pero no disponibles en los datos cedidos.

Con las variables ya definidas, se empezará a construir los datos y el procedimiento a seguir.

5.5 Arquitectura del procedimiento

El procesamiento de los datos se ha estructurado en seis pasos, cada uno encapsulado en un script Python independiente y orquestado por un script principal (pipeline.py) que los ejecuta.

La separación se ha elaborado para conseguir una mayor trazabilidad, modificación y reproducibilidad con el objetivo de que, si se produce un error en algún paso o hay que cambiar para que no afecte a todos los archivos.

5.6 Limpieza del Excel original

El primer paso será convertir el dataset cedido por el hospital y limpiarlo y prepararlo para trabajar. En este apartado se mencionará como se ha corregido el tipo de datos, juntar categorías o como se han gestionado los valores atípicos.

Este paso es crucial para garantizar que la calidad de los datos no contenga errores que puedan afectar a resultados finales.

5.6.1 Tratamiento de fechas y duraciones

El dataset cuenta con tres columnas de fecha-hora:

- **ACTO_CLINICO_FECHA_INICIO:** Hora de llegada de los pacientes a urgencias.
- **ACTO_CLINICO_FECHA_FIN:** Hora de salida de los pacientes de urgencias, es importante recalcar que esta fecha presentaba valores faltantes. Cuando no

existe una fecha de fin válida, el código crea una fecha de fin aproximada sumando el tiempo de estancia a la fecha de inicio.

- **TRIAJE_FECHA_HORA_INICIO:** Hora en la que los pacientes entran a triaje.

A partir de las fechas se construyen tres variables derivadas:

- **TIMESTAMP_HORA:** Es la fecha y la hora del registro redondeada hacia abajo a la hora exacta. Esta variable se usará después como clave para agrupar las llegadas por hora en el paso 2.

- **LOS_MIN:** Es la duración de la estancia del paciente en minutos. Se calcula como la diferencia entre la fecha de fin utilizada y la fecha de inicio del acto clínico.

- **TRIAGE_DELAY_MIN:** Mide el tiempo que pasa entre la llegada del paciente y la realización del triaje, siempre en minutos. Cuando aparecen valores negativos, se consideran errores de registro y se sustituyen por valores nulos (NaN).

5.6.2 Normalización de variables categóricas

En el proceso de transformación de los datos también se han depurado las variables categóricas, limpiando:

- **TRIAJE_CLASIFICACION:** Los valores del triaje se han unificado en mayúsculas: I, II, III, IV y V. En los datos originales había valores en minúsculas, como “iv” y “iii”, se han mantenido solo las mayúsculas para evitar tener variables repetidas.
- **SEXO:** Para la variable del sexo se normaliza también en mayúsculas. Se cambian los valores diferentes de H y M, (debido que solo aparecen en 37 casos) por OTRO, evitando así crear más categorías que representen poca información.
- **EDAD:** La edad se trabaja por rangos y no como una edad exacta (expresada así en los datos originales). Por eso, la variable se renombra como **EDAD_RANGO** y se crean **EDAD_MIN**, **EDAD_MAX** y **EDAD_ORD**, que permiten representar la edad de forma numérica.

- **PROCEDENCIA:** Se eliminan los espacios duplicados y espacios al inicio o al final del texto para dejar la variable en un formato homogéneo.

Hay variables que no se han modificado debido a que su uso no se va a ver implicado en el entrenamiento, como es el caso de **MOTIVO**, donde toda la información presenta el mismo valor de urgencias, y no aporta información diferenciadora para la predicción.

5.6.3 Tratamiento de outliers

Al crear las variables nuevas se ha podido observar que la distribución de **LOS_MIN** es fuertemente asimétrica con cola larga hacía la derecha. En este contexto eliminar los valores o hacer una imputación de la media, distorsionaría la distribución.

Es por eso por lo que, se ha aplicado una estrategia de capping (es un método que permite poner un límite máximo a una variable. Así, los valores que sean demasiado altos no se eliminan, sino que se sustituyen por ese límite para que no afecten demasiado al análisis), en este caso se ha realizado al percentil 99,95 sobre LOS_MIN.

1. El valor original de **LOS_MIN** se conserva intacto en su columna.
2. Se crea una variable binaria **LOS_OUTLIER_99_95** que indica los registros con estancias superiores al percentil 99,95.
3. Se crea una variable **LOS_MIN_CAP_99_95** con la estancia capada al percentil 99,95.
4. Se crea la variable **FECHA_FIN_CAP_99_95** calculada como, inicio + la estancia capada (LOS_MIN_CAP_99_95), de forma que se utiliza exclusivamente para el cálculo de la ocupación simultánea en el paso 2.

En este conjunto de datos, del umbral del percentil 99,95 es de **1.460 minutos**, es decir, algo más de **24 horas**. Con este valor solo se detectan **127 registros**, que superaban, lo que supone apenas el **0,05%** de los **254.025 registros** totales.

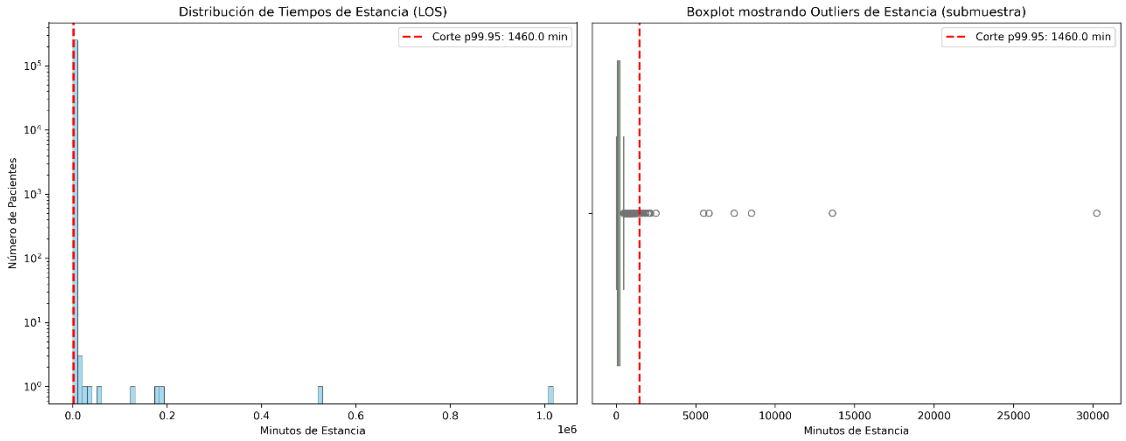

Figura 3. Distribución completa de LOS_MIN en escala logarítmica y boxplot con outliers extremos

En la Figura 3, se observa el histograma completo de LOS_MIN usando una escala logarítmica en el eje Y. Esta escala permite ver mejor la distribución, observando que la mayoría de los pacientes tienen estancias cortas, entre 0 y 500 minutos, mientras que existen algunos valores extremadamente altos.

Estos registros incluyen valores altos cercanos a 1.000.000 de minutos, lo que equivale casi a dos años. Lo que implica que hay casos en los que las estancias a urgencias no fueron cerrados, generando así errores de registro hospitalario-

Además, observando el boxplot de la derecha se confirma esta situación: la mayor parte de los datos queda concentrada al principio y aparecen algunos puntos muy alejados hacia la derecha, que representan esos valores extremos.

Para ver cómo se comportan estos valores se realiza un acercamiento al rango entre 0 y 2.000 minutos.

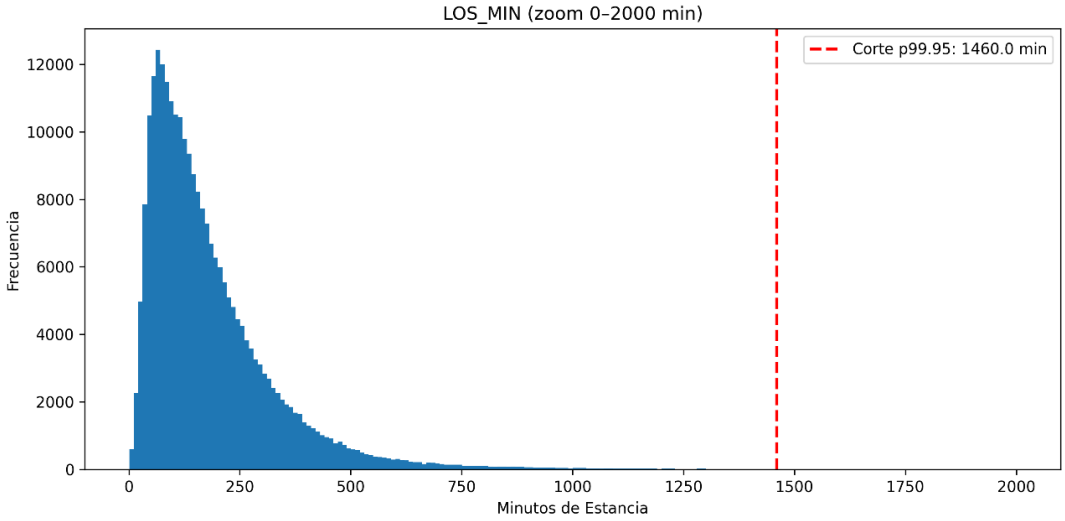

Figura 4. Zoom de la distribución de LOS_MIN en el rango 0-2.000 minutos en escala lineal.

En la figura 4, se observa un acercamiento al rango entre 0 y 2.000 minutos, donde se encuentra la mayoría de los episodios. La distribución normal mantiene la forma esperada donde para un servicio de urgencias: muchos pacientes tienen estancias cortas o medias, y pocos casos tienen estancias más largas.

Además, se observa la misma simetría hacia la derecha, con una mayor concentración en episodios entre 50-100 y una posible mediana cerca de los 120 minutos.

Además, tal y como se observa la **línea roja**, situada en **1.460 minutos**, aparece en una zona donde ya casi no hay registros. Lo significa que el límite elegido no modificará la parte principal de los datos, sino solo los casos más extremos.

Mencionar que esta corrección nunca afectará la variable objetivo del modelo, porque las llegadas se calculan con la fecha de inicio del episodio, no con la fecha de fin.

5.7 Construcción de la serie horaria

Para el paso 2 el dataset se transforma en una serie temporal generando una fila para cada hora. Primero se genera un marco temporal completo con todas las horas del periodo (1 de diciembre de 2021 a 30 de noviembre de 2025, lo que genera 35.064 horas teóricas). Posteriormente, tras crear las variables retardadas y eliminar el periodo inicial de calentamiento, el dataset predictivo final comienza el 8 de diciembre de 2021 y queda formado por 34.894 observaciones.

Sobre ese marco se calcula la variable objetivo y el resto de las métricas importantes utilizadas. Este paso es importante porque garantiza que no haya huecos en la serie, aunque alguna hora no tuviera ninguna visita.

5.7.1 Variable objetivo: LLEGADAS_HORA

Para cada hora t se cuenta el número de episodios cuyo **ACTO_CLINICO_FECHA_INICIO** cae en el intervalo [t, t+1]. De esta forma se obtiene el recuento de la variable **LLEGADAS_HORA**, que representa el número de llegadas registradas en cada franja horaria. Luego con esta variable se crea

TARGET_LLEGADAS_T_PLUS_1H, que será el objetivo predictivo del modelo y corresponde a las llegadas de la hora siguiente.

5.7.2 Ocupación simultánea

La **OCUPACION_SIMULTANEA** en la hora **t** se calcula como el número de pacientes que tienen un episodio activo en ese momento. Es decir, se cuentan los pacientes que su fecha de inicio es anterior o igual a **t** y cuya fecha de fin de ocupación es posterior a **t**.

Para evitar que posibles errores de registro en la fecha de fin inflen artificialmente, esta variable se crea como, **FECHA_FIN_OCUPACION = ACTO_CLINICO_FECHA_INICIO + LOS_MIN_CAP_99_95**.

Es importante recalcar que esta variable representa pacientes presentes simultáneamente en el servicio, no boxes físicos ocupados. Ya que como el dataset no contenía ese dato real de ocupación por box, la variable debe interpretarse como una medida aproximada de la carga.

5.7.3 Indicadores de presión asistencial

El hospital facilitó también una tabla con la capacidad del servicio (9 boxes, 10 médicos en turnos de día y 5 en noche, 8 enfermeras). Esta información no está registrada hora a hora como una medición real, sino que representa la capacidad habitual disponible del servicio. A partir de estos valores se construyen cuatro indicadores de presión por hora.

Variable	Definición	Interpretación
RATIO_BOXES	OCUPACION_SIMULTANEA / 9	Próximo a 1 → presión alta vs capacidad nominal
OVERFLOW_BOXES	max(0, OCUPACION_SIMULTANEA − 9)	Pacientes activos por encima de la capacidad nominal

PACIENTES_POR_MEDICO	OCUPACION_SIMULTANEA / MEDICOS_TURNO	Carga por médico según turno (10 día / 5 noche)
PACIENTES_POR_ENFERMERA	OCUPACION_SIMULTANEA / 8	Carga por enfermera

Tabla 7. Definición de presión asistencial

De este modo, con los indicadores se puede transformar la ocupación simultánea en medidas más interpretables de carga asistencial.

5.7.4 Métricas operativas adicionales

Con el fin de obtener mejores resultados se generan más métricas para ayudar a describir mejor el funcionamiento horario y del servicio de urgencias, se incluyen:

- **TRIAGE_DELAY_MEDIA** y **TRIAGE_DELAY_P95:** Retraso medio y percentil 95 del tiempo hasta triaje de los pacientes en esa hora.
- **INGRESOS_HORA** y **PCT_INGRESO:** Número de pacientes ingresados en esa hora.
- **TRIAJE_I** a **TRIAJE_V:** Número de pacientes en cada nivel de triaje.
- **PCT_TRIAJE_ALTO_I_II:** Pacientes con los niveles más altos de triaje (I, II Y III)

Con estas variables podemos analizar cada hora con más detalle, no solo por el número de llegadas sino la gravedad de los pacientes en esa hora de llegada.

Importante mencionar, que estas variables se calculan con información de la propia hora **t,** no servirían como variables para predecir horas futuras. Por eso en los pasos futuros se crean variables retardadas.

5.8 Calendario y festivos

A partir de la variable **TIMESTAMP_HORA** se generan las variables de calendario base:

- ANIO
- MES
- DIA
- HORA
- DIA_SEMANA (0 = lunes, 6 = domingo)
- ES_FIN_SEMANA
- SEMANA_ISO y TRIMESTRE.

Todas extracciones directas vienen del datetime sin modificaciones.

Aquí también se añade el dataset de FestivosCuenca.csv con los 68 festivos del periodo. Cuando en una misma fecha coinciden varios festivos, por ejemplo, uno nacional y otro autonómico, sus motivos y tipos se agrupan en un único registro para esa fecha.

Todas estas variables se unen con la serie horaria utilizando la fecha del día para generar las variables:

- FESTIVO_MOTIVO (texto)
- FESTIVO_TIPO (Nacional / Autonómico / Regional / Local)
- FESTIVO_TIPO_SIMPLE (versión normalizada) y ES_FESTIVO (binaria).

Además, no solo se tiene en cuenta el día del festivo, sino que se construyen tres variables para capturar posibles efectos de proximidad con estas fechas:

- **ES_PRE_FESTIVO:** 1 si el día siguiente es festivo. Anticipa cambios de comportamiento (visitas previas a un día sin atención primaria, por ejemplo).
- **ES_POST_FESTIVO:** 1 si el día anterior fue festivo. Suele asociarse a picos de demanda por acumulación de problemas no atendidos durante el festivo.
- **ES_PUENTE:** 1 si el día es laborable, pero está entre dos festivos, o es lunes con martes festivo, o viernes con jueves festivo.

5.9 Incorporación de meteorología

Para incorporar a los datos la información de la AEMET se hará utilizando la fecha del día, ya que la información meteorológica no está disponible a nivel diario.

Antes de unir ambos dataset, en el de meteorológica, la variable de precipitación "lp" (precipitación inapreciable), se crea una variable **PREC_IP,** que vale 1 cuando aparecen estos casos, consiguiendo así perder la menor información posible. (Luego este valor se sustituye por 0,0 en la variable numérica **prec_mm**). También se crea la variable **LLUVIA_FLAG,** que indica si hubo precipitación medible ese día.

Las demás variables meteorológicas se conservan en la misma forma numérica que traía el dataset. Teniendo:

- Temperaturas med, min, max
- Humedad
- Viento
- Horas del sol
- Presión atmosférica max, min

Y para el caso de las temperaturas se crea **TEMP_RANGO** = tmax – tmin, para ver la diferencia entre las temperatura máxima y mínima del día.

Como los datos de la meteorología fueron adquiridos de la AEMET de forma manual, se crea la variable **METEO_MISSING**, para comprobar que no haya ningún dato faltante. (Como el resultado de la variable es 0, porque se tienen todos los datos se acaba eliminando)

5.10 Retardos y medias móviles

Este paso es uno de los más importantes, porque aquí se generan todas las variables de lags (retardos) que el modelo usará, por lo que hay que tener especial cuidado con la fuga de información.

5.10.1 Imputación conservadora de meteorología restante

Aunque el cruce con AEMET fue completo y no quedaron días sin información meteorológica, algunas variables concretas sí presentaban valores ausentes de forma puntual. Entonces para mantener la trazabilidad se crean variables de ausencia.

- **PREC_MM_MISSING**
- **SOL_MISSING**
- **presMax_MISSING**
- **presMin_MISSING**

Después, los valores ausentes se rellenan de forma sencilla: la precipitación se completa con 0,0, las horas de sol con la mediana mensual y, si no fuera posible, con la mediana global; y la presión máxima y mínima con su mediana global.

5.10.2 Codificación cíclica de variables temporales

Es importante mencionar que alguna de las variables son cíclicas, como HORA y DIA_SEMANA, esto implica que, aunque la distancia entre la hora 23 y la hora 0 es pequeña, el modelo podría considerar una diferencia de 23. Para corregirlo se generan codificaciones con seno y coseno mediante las siguientes fórmulas:

```
if "HORA" in df.columns:
      df["HORA_SIN"] = np.sin(2 * np.pi * df["HORA"] / 24.0)
      df["HORA_COS"] = np.cos(2 * np.pi * df["HORA"] / 24.0)
                  if "DIA_SEMANA" in df.columns:
    df["DOW_SIN"] = np.sin(2 * np.pi * df["DIA_SEMANA"] / 7.0)
    df["DOW_COS"] = np.cos(2 * np.pi * df["DIA_SEMANA"] / 7.0)
```

5.10.3 Variables creadas a partir de valores anteriores (Lags del target)

Se generan lags de la variable **LLEGADAS_HORA** en los horizontes: 1h, 2h, 3h, 6h, 12h, 24h, 48h y 168h. Estas variables se añaden para que el modelo pueda usar información pasada y no haya fuga de información.

- **1h, 2h, 3h**: permiten saber cómo venía evolucionando la demanda en las horas más recientes.
- **6h, 12h:** ayudan a recoger patrones dentro del mismo día.
- **24h:** permite comparar con la misma hora del día anterior (estacionalidad diaria).
- **48h**: aporta información de los dos días anteriores.
- **168h:** corresponde a la misma hora de la semana anterior, ya que 168 horas equivalen a 7 días. (estacionalidad semanal).

5.10.4 Media móvil y desviación típica móvil

Además, de las llegadas por hora concreta pasadas, se generan medias y desviaciones típicas móviles sobre las llegadas para 6h, 24h y 168h.

Con la media móvil se resume el nivel reciente de demanda, poder saber si en las últimas horas llegaban muchos pacientes a urgencias.

Es importante que estos cálculos solo usen información pasada, por eso se utiliza **shift (1),** consiguiendo que el código ignore la hora actual y evitar fugas.

```
For w in ROLL_WINDOWS:

      df [f "LLEGADAS_ROLL_MEAN_{w}H"] =
      df["LLEGADAS_HORA"].shift(1).rolling(window= w ,
      min_periods = w).mean()

      df [f "LLEGADAS_ROLL_STD_{w}H"] =
      df["LLEGADAS_HORA"].shift(1).rolling(window= w ,
      min_periods = w).std()
```

5.10.5 Lags de variables operativas observadas

Para los lags es muy importante que la información solo sea del pasado es por eso por lo que hay variables que se descartan.

- OCUPACION_SIMULTANEA
- OVERFLOW_BOXES

- RATIO_BOXES
- INGRESOS_HORA
- PCT_INGRESO
- TRIAGE_DELAY_MEDIA
- TRIAGE_DELAY_P95
- PCT_TRIAJE_ALTO_I_II

Estas variables son observadas en t y por tanto no pueden usarse en t+1. De manera que se generan sus lags a 1h, 24h y 168h, para poder usarse en momentos pasados. Igualmente se generan medias móviles de pacientes presentes con shift(1) en la misma que anteriormente.

5.10.6 Eliminación del periodo de calentamiento

Para las primeras 168 horas del dataset (la primera semana, del 8 al 14 de diciembre de 2021), no se tiene suficiente información histórica para calcular los retardos, por lo que se excluirá del dataset.

El dataset, sale como salida (dataset_modelo.csv), que contiene 34.894 filas y 102 columnas.

5.11 Construcción del dataset predictivo final

En este paso se prepara el dataset final para el entrenamiento de los modelos. Para ello, se define qué variable se quiere predecir, qué variables se pueden usar como información de entrada y cuáles deben excluirse para evitar errores.

5.11.1. Feature autoregresiva en t (LLEGADAS_HORA_T)

Se incluye la variable **LLEGADAS_HORA_T**, que representa el número de pacientes que han llegado durante la hora actual **t**. Esta variable se utiliza como información de entrada para predecir las llegadas de la hora siguiente (**t+1**).

Además, esta variable aporta una señal auto regresiva importante, ya que la demanda de urgencias en una hora suele estar relacionada con la demanda registrada en las horas anteriores. De esta forma, **LLEGADAS_HORA_T** se considera una de las variables más importantes para anticipar la afluencia de la siguiente franja horaria.

5.11.2 Variables excluidas y justificación

Con las variables analizadas, encontramos variables que no aportan información relevante, es por eso por lo que se excluyen dejando muy bien recalcado el motivo.

Variable excluida	Motivo
CAP_BOXES_NOMINAL	Constante = 9 en todo el periodo. No aporta poder predictivo.
ENFERMERAS_TURNO	Constante = 8 en todo el periodo. No aporta poder predictivo.
MEDICOS_TURNO	Función determinística de TURNO (10 día / 5 noche).
OCUPACION_SIMULTANEA, OVERFLOW_BOXES, INGRESOS_HORA, PCT_INGRESO, TRIAJE_I a TRIAJE_V, TRIAGE_DELAY_MEDIA, TRIAGE_DELAY_P95, PCT_TRIAJE_ALTO_I_II (todas en t)	Observadas en t, no disponibles al predecir t+1. Solo se usan sus lags.
FESTIVO_MOTIVO	Texto libre de alta cardinalidad. Útil para descriptivo, no para modelo.
PREC_IP, METEO_MISSING	Detectadas como constantes y eliminadas automáticamente.

Tabla 8. Variables excluidas del dataset predictivo y motivo de exclusión.

5.11.3 Verificaciones automáticas y salida final

Tras volver a realizar las verificaciones automáticas, como en los pasos anteriores y las eliminaciones de las columnas.

La salida final es el dataset_predictivo.csv, que contiene 34.894 y 61 columnas, (3 columnas de identificación temporal y target, 58 features predictoras).

5.12 Caracterización del dataset y análisis exploratorio

Con el dataset predictivo ya finalizado, se analiza su comportamiento general antes de entrenar los modelos. El objetivo de este apartado es comprobar que la serie temporal está bien formada.

5.12.1 Distribución temporal y partición train/test

El dataset predictivo comprende datos desde el 8 de diciembre de 2021 hasta el 30 de noviembre de 2025, lo que serían casi cuatro años completos de observaciones. La distribución por años se muestra en la tabla junto con la partición temporal aplicada en la fase de entrenamiento.

Año	Filas (horas)	Cobertura	Conjunto
2021	575	Diciembre parcial	Train
2022	8.760	Año completo	Train
2023	8.760	Año completo	Train
2024	8.784	Año completo bisiesto	Train
2025	8.015	Enero-noviembre	Test
TOTAL	34.894	1.454 días completos	—

Tabla 9. Distribución temporal del dataset y partición train/test.

Con esta partición conseguimos tres objetivos claros:

1. El modelo se entrena sobre tres ciclos anuales completos (2022, 2023 y 2024) capturando estacionalidad anual.
2. La evaluación se realiza sobre un periodo independiente de 11 meses, suficientemente largo para capturar variabilidad estacional intra-test.
3. No hay solapamiento ni fuga temporal entre conjuntos.

La elección de 2025 como test, es por ser el año más reciente del dataset y reproduce fielmente el escenario de uso real más cercano a la realidad.

5.12.2 Distribución del target

La variable objetivo TARGET_LLEGADAS_T_PLUS_1H presenta los siguientes estadísticos descriptivos:

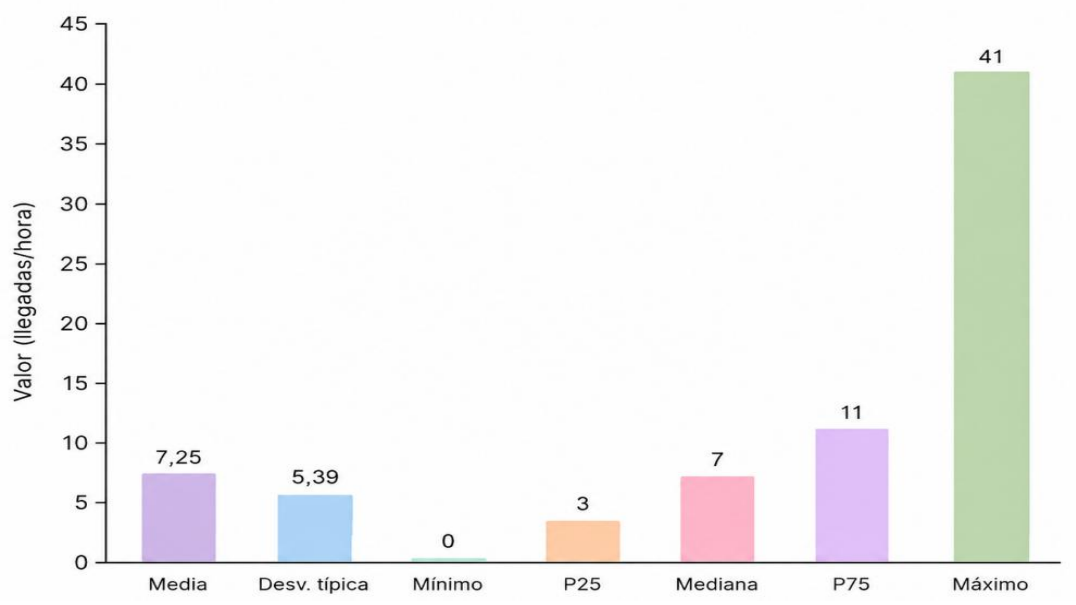

Figura 5. Estadísticos descriptivos del target.

5.12.3 Patrones temporales observados

A lo largo del análisis se han podido identificar los patrones.

- **Patrón diario:** pico de actividad entre las 9 y las 12 de la mañana (media de 14 llegadas/hora) y valle entre las 4 y las 6 de la madrugada (media de 1,5 llegadas/hora).
- **Patrón semanal:** el lunes es el día con mayor actividad (8,4 llegadas/hora de media), seguido por el resto de los laborables. Los fines de semana presentan menor actividad (6,2-6,3 llegadas/hora).
- **Festivos:** presentan ligeramente menos actividad que los días laborables (6,5 frente a 7,3), pero la diferencia es modesta y se ve compensada por la influencia del día de la semana en que cae el festivo.

5.13 Evolución del dataset predictivo

La construcción del dataset predictivo se planteó como un proceso incremental. En primer lugar, se diseñó una versión base, denominada v1, con variables temporales, meteorológicas, festivos, lags y medias móviles. Sobre esta primera versión se entrenaron los modelos iniciales y se analizaron sus resultados.

A partir de esta evaluación con el objetivo de mejorar los resultados se añadieron más variables históricas y es por lo que se crea v2. Posteriormente, se creó una versión v3 con variables adicionales más avanzadas, con el objetivo de comprobar si una mayor complejidad aportaba una mejora real.

Sin embargo, la comparación experimental mostró que v2 ofrecía el mejor equilibrio entre interpretabilidad, precisión y estabilidad.

5.13.1 Metodología de búsqueda de features

Una vez construido el dataset predictivo base (dataset_predictivo_v1, 61 columnas, 34.894 filas), se realizó una búsqueda de nuevas variables que pudieran mejorar la capacidad predictiva del modelo.

Se evaluaron más de 30 features candidatas agrupadas en ocho categorías: lags adicionales, rolling features, binarias de franja horaria, interacciones festivo × hora,

temperatura extrema, medias históricas por grupo, lags de semanas múltiples y features de tendencia.

Para poder aceptar una nueva variable, tenían que cumplir con:

- Correlación absoluta superior a 0,30 con el target
- Aportación de información no redundante respecto a las features ya existentes
- Posibilidad de cálculo en el instante t sin acceder a información de t+1 ni del futuro.

Cuando dos features candidatas estaban altamente correlacionadas entre sí, se conservaba la de mayor correlación con el target.

5.13.2 Features descartadas tras el análisis

Con este análisis encontramos que la mayoría de los candidatos evaluados fueron descartados.

Variable candidata	Correlación con target	Motivo del descarte
LAG_4H, LAG_5H	−0,06 / −0,16	Correlación negativa débil por ciclo diario; redundante con HORA_SIN/COS
LAG_7H a LAG_19H	−0,35 a +0,18	Correlación débil o negativa; ciclo ya capturado por LAG_24H
Temperatura extrema (frío/calor)	< 0,05	Señal insuficiente; temperatura ya incluida como variable continua
RATIO_VS_MEDIA	−0,038	No aporta información nueva sobre el resto de features
FIN_SEM × HORA_PICO	0,131	Marginal; ya capturado por ES_FIN_SEMANA y ES_HORA_PICO por separado

MEDIA_PP_HORA_DOW (ocupación)	0,222	Bajo impacto; ya cubierto por lags de OCUPACION_SIMULTANEA
SEMANA_DEL_MES	0,000	Sin correlación

Tabla 10. Variables candidatas descartadas y motivo del descarte

5.13.3 Dataset v2 variables aceptadas.

Tras analizar los resultados de v1, se buscó nuevas variables que pudieran mejorar los resultados obtenidos y se encontraron 4 variables, generando así un nuevo dataset_predictivo_v2 (con más información histórica). Donde las cuatro variables fueron:

- **MEDIA_HORA_DOW — media histórica por (hora, día de la semana).** Para cada combinación de hora del día (0-23) y día de la semana (0 = lunes, 6 = domingo), se calcula la media de llegadas observadas en el conjunto de train. Pudiendo así responder a cuántas llegadas se esperan en este hueco horario según el histórico completo de entrenamiento.
  Para la variable cada observación se calcula usando exclusivamente los datos de train del fold correspondiente en el TimeSeriesSplit, nunca incluyendo la observación propia ni datos del futuro.
- **MEDIA_HORA_MES — media histórica por (hora, mes).** Esta variable agrupa por mes en lugar de día de semana. Permitiendo capturar la estacionalidad anual horaria: la demanda esperada a las 10 h de enero no es la misma que la esperada a las 10 h de agosto.
- **DELTA_1H_2H — tendencia inmediata.** Diferencia LLEGADAS_LAG_1H − LLEGADAS_LAG_2H. Mide si la demanda está subiendo o bajando en el último tramo.
- **DELTA_24H_48H — tendencia día a día.** Diferencia LLEGADAS_LAG_24H − LLEGADAS_LAG_48H. Mide si la misma franja horaria fue más o menos concurrida ayer que anteayer.

5.13.4 Dataset v3 variables aceptadas.

Para comprobar si un dataset con variables más complejas podía variar mucho los resultados se incorporaron 13 variables, al dataset v2, creando así el dataset_predictivo_v3 con 78 columnas. Separando estas variables en tres grupos con correlaciones suficientemente altas:

Grupo A — Lags intra-ciclo diario: LAG_20H, LAG_21H, LAG_22H, LAG_23H.

Estos lags completan la información entre las 12 y las 24 horas anteriores. Aunque al principio pueden parecer menos intuitivos, ayudan a relacionar la demanda de la tarde del día anterior con la demanda de la mañana siguiente.

Si se quiere predecir la demanda de hoy a las 10:00, algunos de estos lags recogen información de las horas de ayer entre la mañana y la tarde. Esto puede ser útil porque la actividad del servicio sigue un ciclo diario, donde después de las horas de menor actividad, la demanda vuelve a aumentar por la mañana.

Grupo B — Binarias de franja horaria.

Además de los lags temporales, se incorporaron variables binarias que identifican franjas horarias especialmente relevantes dentro del ciclo diario

- **ES_HORA_PICO (horas 8-13):** Esta variable identifica las horas de mayor actividad del día. Aunque el modelo ya tiene variables que representan la hora de forma cíclica, como **HORA_SIN** y **HORA_COS**, añadir una variable binaria permite marcar de forma más directa si una hora pertenece o no al periodo punta.
- **ES_HORA_VALLE (horas 1-6):** Marca el periodo de mínima actividad. La correlación negativa es esperada y correcta.
- **ES_HORA_APERTURA (horas 8-9):** Marca las dos primeras horas del pico, donde el análisis de baselines mostró el mayor error. El modelo puede aprender un comportamiento específico para esta franja.

Grupo C — Lags semanales múltiples y sus medias.

El lag LAG_168H (una semana exacta) tiene la limitación de que una semana concreta puede ser atípica por festivos, eventos locales o condiciones meteorológicas excepcionales. Promediar las últimas N semanas suaviza ese ruido.

Para calcular estos promedios sin fuga de información, primero se generan los lags de 2, 3 y 4 semanas (LAG_336H, LAG_504H, LAG_672H) y después se calcula la media de los lags disponibles.

Variable	Descripción
LLEGADAS_LAG_336H	Llegadas exactas de hace 2 semanas (misma hora)
LLEGADAS_LAG_504H	Llegadas exactas de hace 3 semanas (misma hora)
LLEGADAS_LAG_672H	Llegadas exactas de hace 4 semanas (misma hora)
MEDIA_2SEM	Media de LAG_168H y LAG_336H
MEDIA_3SEM	Media de LAG_168H, LAG_336H y LAG_504H
MEDIA_4SEM	Media de los 4 lags semanales

Tabla 11. Variables grupo C

El hallazgo más importante de este grupo es que MEDIA_4SEM alcanza una correlación de 0,752 con el target, convirtiéndose en la tercera variable más correlacionada. Esto confirma que promediar cuatro semanas es más informativo que usar una sola porque amortigua la variabilidad de semanas atípicas.

5.13.5 Ranking final de variables por correlación con el target

Se recoge en una tabla el ranking completo de las 20 variables con mayor correlación.

Las variables nuevas (añadidas en v2 y v3) ocupan 13 de las 20 primeras posiciones, lo que confirma que el esfuerzo de buscar variables que mejoren el modelo v1 aportó valor real.

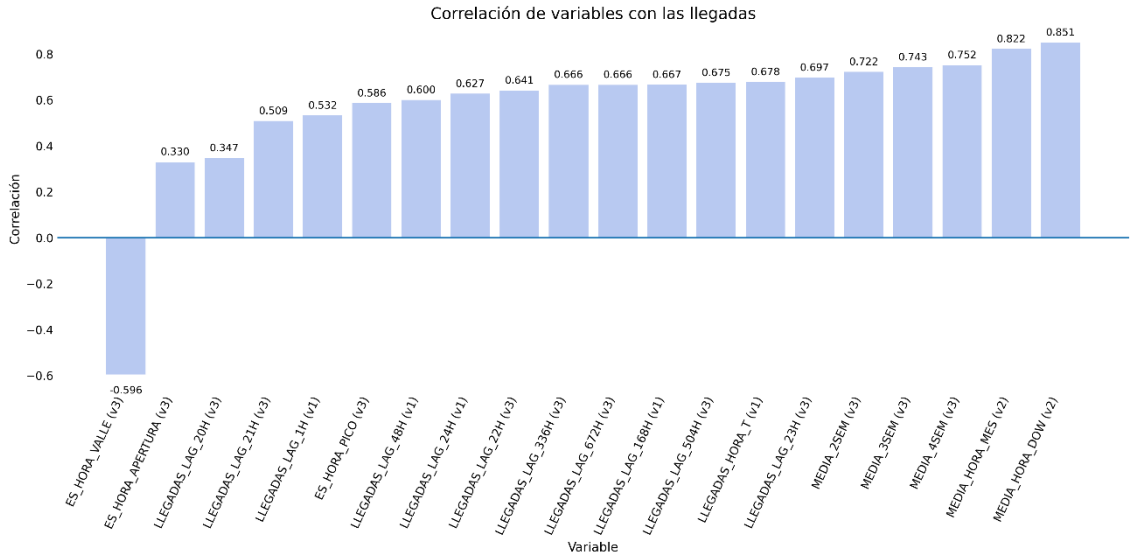

Figura 6. Ranking de features por correlación absoluta con el target

5.13.6 Estado final del dataset por versiones

Una vez definidas las tres versiones del dataset, estas se utilizaron como base para comparar el rendimiento de los modelos predictivos y analizar si la incorporación progresiva de variables suponía una mejora real en la predicción.

Versión	Columnas	Filas	Variables añadidas
v1 (original)	61	34.894	Dataset base del pipeline
v2	65	34.894	MEDIA_HORA_DOW, MEDIA_HORA_MES, DELTA_1H_2H, DELTA_24H_48H

V3	78	34.222	LAG_20-23H, ES_HORA_PICO/VALLE/APERTURA, LAG_336/504/672H, MEDIA_2/3/4SEM

Tabla 12. Versiones finales de los datasets.

5.14 Esquema de validación temporal

Para poder realizar una correcta validación de modelos en series temporales es muy importante no mezclar observaciones de distintos momentos. Para evitarlo hacemos un esquema a dos niveles.

1. Hold-out (conjunto reservado), final con los datos más recientes, reservándolos para la evaluación última de cada modelo

2. Validación interna con timeseriessplit sobre el conjunto de entrenamiento, usado para la búsqueda de hiperparámetros

5.14.1 Hold-out final (test 2025)

Se reserva como conjunto de test final para el periodo del 1 de enero al 30 de noviembre de 2025, con un total de 8.015 horas. Este conjunto no se usa durante el entrenamiento ni la búsqueda de hiperparámetros. Su única función es la evaluación final de cada modelo. Se selecciona 2025 como test al ser el año más reciente del dataset.

5.14.2 TimeSeriesSplit con 3 folds sobre el train

Para la validación de hiperparámetros y la estimación del rendimiento de cada modelo se emplea TimeSeriesSplit con 3 folds sobre el conjunto de train (diciembre 2021 – diciembre 2024).

Este esquema garantiza que en todo momento el modelo se entrena sobre datos anteriores al periodo de validación, sin filtración de información futura hacia el pasado.

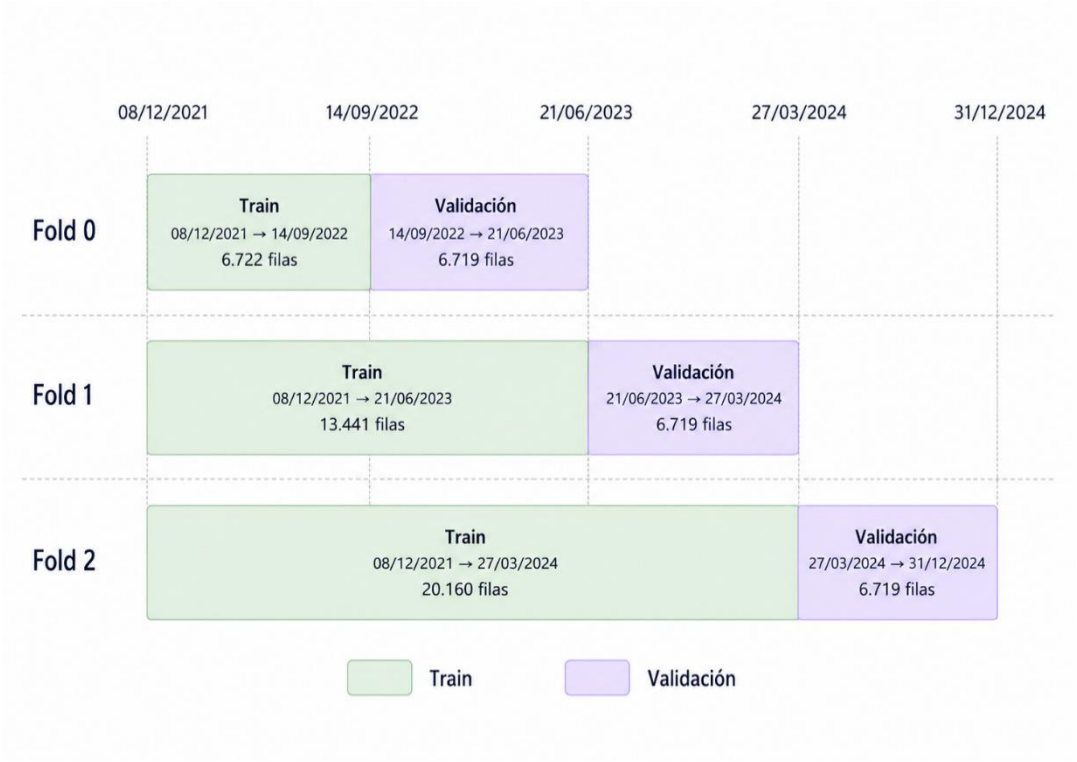

Figura 7. Gráfico de los folds de validación temporal (TimeSeriesSplit, 3 folds).

Se ha realizado el esquema con esta forma para que cumpla con dos propiedades importantes:

1. Que cada fold añade más datos al train, lo que permite estudiar el rendimiento del modelo y comprobar si mejora.
2. Conseguir que los tres periodos de validación cubran distintas épocas del año, pudiendo tener una estimación robusta de la estacionalidad anual del error.

5.15 Modelos de referencia (baselines)

Antes de evaluar modelos predictivos más complejos, es necesario establecer unos predictores sencillos que sirvan como referencia, para poder observar si el modelo está mejorando los resultados. Estos modelos son conocidos como baselines (modelo de referencia), y sirven como punto de comparación mínimo, para poder saber si un modelo avanzado mejora los resultados o no aporta una mejora real.

En este TFG se han definido tres modelos de referencia con distinto nivel de exigencia

5.15.1 Baseline 1: Naive estacional 168 h

Para el baseline 1, utilizamos la predicción más sencilla, un modelo naive estacional semanal. Para predecir para la hora *t +1*, las mismas llegadas que se observaron 168 horas antes (la misma hora del mismo día de la semana de la semana anterior).

Aquí no se necesita entrenamiento sino un retardo de 168h (ya tenemos creado en el dataset)

$$Pred\ (t+1) = LLEGADAS_LAG_168H(t)$$

Este es el baseline más fácil de superar, ya que no captura ninguna variación que no sea la estacionalidad semanal exacta.

5.15.2 Baseline 2: media histórica por (hora, día-semana, mes)

Para el baseline 2, se utilizará cada combinación de hora del día, día de la semana y del mes. Para poder calcular la media de llegadas observadas en el conjunto del train y como se usa para la predicción.

$$Pred\ (t+1) = media\ (llegadas \mid hora = h,\ dia_semana = d,\ mes = m)$$

Con este baseline se incorpora estacionalidad intra-día, semanal y anual, pero se mantiene estático y no se adapta a condiciones recientes.

5.15.3 Baseline 3: MEDIA_HORA_DOW directa

Para el tercer baseline se seleccionará la variable MEDIA_HORA_DOW, que recoge la media histórica de llegadas según la hora del día y el día de la semana. (esta variable la encontramos en el dataset v2). Al usarla directamente para la predicción permite establecer una referencia sencilla pero muy exigente.

5.15.4 Análisis de los baselines

Como se explicará más adelante en el apartado de resultados, los tres modelos de referencia tienden a predecir valores algo más bajos de los reales, sobre todo cuando hay picos altos de demanda.

Esto pasa cuando los modelos sencillos no captan bien los aumentos bruscos de llegadas a urgencias. Es por eso por lo que se espera que los modelos de aprendizaje automático pueden mejorar parcialmente este problema.

Con estos tres modelos de referencia se establecen distintos niveles de comparación, desde el más sencillo hasta el más exigente.

Los valores de MAE incluidos en la tabla proceden de la evaluación realizada en el apartado de resultados, pero se muestran aquí solo para tener una referencia inicial de los resultados obtenidos por estos baselines.

Nivel	Modelo de referencia	MAE test	Significado de superarlo
Suelo fácil	Naive 168h	3,2947	El modelo aporta algo respecto a repetir la semana pasada
Suelo medio	Media histórica	2,2188	El modelo supera la mejor heurística sin entrenamiento
Suelo duro	MEDIA_HORA_DOW directa	2,1718	El modelo supera la mejor señal individual del dataset

Tabla 13. Suelos de comparación para los modelos de ML.

5.16 Modelos de aprendizaje entrenados

Con los baseline ya definidos, tenemos una base para poder comparar con los modelos. Es por eso, que en este apartado se definirá las familias de modelos entrenados para la predicción de las llegadas a urgencias.

El proceso de modelado no se planteó como un único dataset de prueba, sino como una evolución progresiva. Primero se entrenaron modelos sobre el dataset v1, que incluía las variables iniciales. Después, al analizar sus resultados y detectar limitaciones para captar determinados patrones horarios y semanales, se construyó el dataset v2 incorporando variables históricas agregadas. Posteriormente, se desarrolló el dataset v3 con variables adicionales más avanzadas, con el objetivo de comprobar si una mayor complejidad aportaba una mejora real.

Al observar que el modelo v2, obtenía mejores resultados, se utilizó como base principal y se probaron más modelos de aprendizaje automático adicionales, basados en árboles, como LightGBM y Random Forest.

La comparación se realizó manteniendo el mismo esquema temporal de entrenamiento y evaluación, de forma que todos los modelos fueran valorados sobre los mismos periodos y bajo las mismas condiciones.

5.16.1 Familias de modelos y justificación de la selección

Para la selección de modelos se organizó en tres enfoques principales dentro de la predicción de series temporales:

- Modelos estadísticos clásicos
- Modelos basados en árboles de decisión
- Modelos secuenciales de deep learning

La inclusión de modelos clásicos nos ayuda ya que son fácilmente interpretables. **SARIMAX y Prophet** exponen los coeficientes autorregresivos y estacionales, lo que facilita la discusión de los patrones temporales detectados. Además, en horizontes cortos

(1 h) y con datos limitados, los modelos clásicos suelen ser competitivos con los modelos profundos.

LightGBM y Random Forest se incorporaron porque trabajan bien con datos tabulares y variables creadas manualmente, como lags, medias móviles, variables de calendario y variables meteorológicas.

También se incluyó **LSTM** ya que responde a la necesidad de evaluar si un modelo secuencial puede aprender patrones que las variables manualmente construidas no capturan.

Modelo	Familia	Justificación de inclusión
SARIMAX	Estadístico clásico	Modelo de referencia clásico; interpretabilidad; control de estacionalidad estacional 24h explícita
Prophet	Estadístico moderno	Doble estacionalidad (diaria y semanal); manejo nativo de festivos; benchmark habitual en literatura sanitaria
LightGBM	Gradient boosting	Mejor familia para forecasting tabular con features ingenierizadas; soporta quantile regression para intervalos
HistGradientBoosting	Gradient Boosting	Prueba adicional sobre v2 con boosting eficiente de scikit-learn
ExtraTrees	Ensemble de árboles aleatorizados	Prueba adicional sobre v2 para evaluar árboles con mayor aleatoriedad
Random Forest	Bagging de árboles	Prueba adicional sobre v2 para comprobar si mejora LightGBM
LSTM	Red neuronal recurrente	Modelo secuencial estado del arte; captura dependencias largas que features manuales pueden no representar

Tabla 14. Modelos entrenados y su justificación.

5.16.2 SARIMAX

Se entrenó un modelo **SARIMAX (1,0,1) (1,0,1) [24]** sobre la serie LLEGADAS_HORA, con un periodo estacional de 24 horas, para conseguir capturar el ciclo diario de la demanda.

Las variables exógenas incluidas fueron el día de la semana, la identificación de fin de semana, la presencia de festivo y, cuando estaban disponibles en la versión del dataset correspondiente, variables históricas como **MEDIA_HORA_DOW**, **MEDIA_HORA_MES** y variables de franja horaria como **ES_HORA_PICO** y **ES_HORA_VALLE**. De esta forma, SARIMAX combina la dinámica propia de la serie con información externa útil para explicar cambios en la demanda.

5.16.3 Prophet

También se utilizó el modelo Prophet que se ha empleado con su configuración nativa para doble estacionalidad (ciclo de 24 horas y ciclo de 7 días).

La estacionalidad anual se desactivó para mantener el modelo centrado en los ciclos que están más relacionados con la predicción a corto plazo, variación por hora del día y por día de la semana.

Se incorporó el calendario de festivos del fichero FestivosCuenca a través del mecanismo holidays de Prophet, que aprende un efecto específico para cada festivo.

El modelo se entrenó sobre la serie de entrenamiento completo y generó predicciones para el conjunto de test final.

5.16.4 LightGBM con optimización bayesiana

El modelo principal de la fase de modelado es LightGBM (Gradient Boosting Machine). Fue elegido porque es un algoritmo adecuado para problemas tabulares con muchas variaciones. Sus hiperparámetros se han optimizado mediante Optuna sobre el primer fold del TimeSeriesSplit, ejecutando alrededor de 50 trials de búsqueda bayesiana.

Buscando con la función objetivo minimizar el MAE sobre el conjunto de validación del fold 0.

A diferencia de los modelos estadísticos clásicos, LightGBM no necesita asumir una estructura lineal ni una forma concreta de la serie temporal. Su ventaja es que puede aprender relaciones no lineales e interacciones entre variables.

Los hiperparámetros ajustados fueron: num_leaves, max_depth, learning_rate, min_child_samples, feature_fraction, bagging_fraction, bagging_freq y reg_alpha.

Una vez fijados los hiperparámetros óptimos, el modelo se reentrenó sobre el train completo y se evaluó sobre el test final. Construyendo un intervalo de predicción aproximado del 80% mediante el rango P10-P90, así el modelo no ofrece solo una estimación puntual sino un rango de incertidumbre.

5.16.5 LSTM con MC-Dropout

La red recurrente LSTM se ha implementado en PyTorch como modelo secuencial de Deep learning (aprendizaje profundo). A diferencia de los modelos tabulares, que trabajan solo con una fila independiente para cada hora de variables, la LSTM recibe como entrada una secuencia temporal de observaciones anteriores y aprende de los patrones a partir de esa evolución.

La LSTM se entrenó sobre las distintas versiones del dataset (v1, v2, v3), manteniendo la arquitectura base, pero adaptándolo para las variables exógenas de cada versión. Para ambos datos se usa una entrada de 168 horas.

La arquitectura utilizada consta de 2 capas LSTM, un tamaño oculto de 128 unidades y una tasa de dropout de 0,2.

El entrenamiento se realizó con un tamaño de batch (lote) de 256, un máximo de 50 épocas, optimizador Adam con tasa de aprendizaje de 0,001 y función de pérdida MSELoss. Además, se aplicó early stopping (parado antes de tiempo) con paciencia de 7 épocas, de forma que el entrenamiento se detenía si la pérdida de validación no mejoraba.

Durante el entrenamiento final, el conjunto de test no se utilizó para ajustar el modelo ni para aplicar early stopping. Para ello se reservó el último fold como validación interna, manteniendo el test final únicamente para la evaluación definitiva del modelo.

5.16.6 Modelos adicionales sobre v2

Tras observar que el dataset v2 ofrecía los mejores resultados, se probaron modelos adicionales basados en árboles. El objetivo de esta fase fue comprobar si otros algoritmos podían mejorar los resultados obtenidos por LightGBM manteniendo como base el mismo conjunto de variables.

En esta fase se entrenaron tres modelos adicionales: HistGradientBoosting, ExtraTrees y Random Forest.

HistGradientBoosting se utilizó como alternativa dentro de la familia de Gradient Boosting. A diferencia de LightGBM, este modelo pertenece a la implementación de scikit-learn y utiliza una estrategia basada en histogramas para acelerar el entrenamiento.

En este trabajo se configuró con learning_rate = 0.04, max_iter = 600, max_leaf_nodes = 31, min_samples_leaf = 20, l2_regularization = 0.1 y early_stopping = True. Su inclusión permitió comprobar si otro enfoque de boosting podía mejorar o igualar el rendimiento de LightGBM sobre el dataset v2.

ExtraTrees se entrenó como modelo de ensemble de árboles con mayor aleatoriedad. A diferencia de Random Forest, introduce más aleatoriedad en la construcción de los árboles, lo que puede ayudar a reducir la varianza y mejorar la robustez ante variables correlacionadas.

El modelo se configuró con n_estimators = 250, max_depth = 22, min_samples_leaf = 4, min_samples_split = 8, max_features = sqrt y bootstrap = False.

Random Forest se entrenó con el objetivo de comprobar si podía mejorar los resultados obtenidos por LightGBM. A diferencia del otro modelo, que construye árboles de forma secuencial corrigiendo errores anteriores, Random Forest entrena múltiples árboles independientes y combina sus predicciones mediante promedio.

Se utilizó con una configuración específica de hiperparámetros: n_estimators = 180, max_depth = 20, min_samples_leaf = 5, min_samples_split = 10, max_features = sqrt y bootstrap = True.

Los tres modelos se integraron en un pipeline de preprocesamiento común, con imputación de variables numéricas mediante la mediana, imputación de variables categóricas mediante el valor más frecuente y codificación one-hot para las variables categóricas. Esto permitió comparar los modelos adicionales bajo una preparación de datos homogénea.

5.17 Modelo futuro para predicción manual

Además de los modelos entrenados para predecir la hora siguiente a partir de información reciente, se desarrolló una versión complementaria v4_ModeloFuturo. Con el objetivo de permitir que el usuario introduzca una fecha y una hora futura y obtenga una estimación de las llegadas en esa franja horaria con el error promedio y el rango de posibles llegadas.

La principal diferencia respecto a los datasets v1, v2 y v3 es que el modelo futuro no puede utilizar variables que solo se conocerían una vez llegada esa fecha. Donde este modelo no puede usar las llegadas de la hora anterior real, la ocupación reciente del servicio, los lags inmediatos ni las medias móviles calculadas a partir de actividad reciente. Ya que, en una consulta futura, esa información todavía no existe.

Es por eso, por lo que este modelo se construye con las variables disponibles y preguntando por terminal las variables exógenas.

5.17.1 Construcción del dataset futuro

Lo primero que se hizo fue crear un dataset específico, denominado **dataset_futuro_base.csv**. Este dataset parte de la serie horaria original, pero elimina las variables que no estarían disponibles en una predicción manual futura.

El objetivo de esta versión ya no es predecir la hora siguiente a partir de la hora actual, sino estimar directamente las llegadas esperadas en una hora futura concreta. Por ello, la variable objetivo se redefine como:

TARGET_LLEGADAS_HORA_FUTURA

Esta variable representa el número de llegadas registradas en la hora seleccionada. En entrenamiento y evaluación se conoce su valor real, pero en una predicción futura manual queda vacío y es el valor que el modelo debe estimar.

Las variables utilizadas en el modelo futuro se agrupan en varios bloques:

- **Variables temporales y de calendario**, hora, día de la semana, mes, trimestre, semana ISO y día del año.
- **Codificaciones cíclicas**, seno y coseno de la hora, del día de la semana, del mes y de la semana del año.
- **Variables de franja horaria**, turno, franja del día, estación, hora pico, hora valle, hora de apertura y cambio de turno.
- **Variables de festivos**, festivo, pre-festivo, post-festivo, puente, tipo de festivo y tipo de día.
- **Variables meteorológicas**, temperatura, precipitación, humedad, viento, presión y horas de sol.
- **Variables meteorológicas derivadas**, rango térmico, frío extremo, calor extremo, lluvia intensa, viento fuerte y anomalías respecto a la media mensual.
- **Patrones históricos agregados**, calculados por grupos temporales, como hora-día de semana, hora-mes, hora-día-mes, hora-tipo de día, hora-estación, hora-festivo, hora-franja y hora-turno.

Estas variables históricas agregadas incluyen medias, medianas, desviaciones típicas y percentiles, de forma que el modelo pueda conocer el comportamiento habitual de la demanda en situaciones similares sin utilizar información futura.

5.17.2 Prevención de fuga de información temporal

Las variables históricas se calculan evitando la fuga de información temporal. Para ello, el proceso se divide en tres bloques:

- **Train:** datos anteriores al 1 de julio de 2024.
- **Validación:** del 1 de julio de 2024 al 31 de diciembre de 2024.
- **Test final:** desde el 1 de enero de 2025 en adelante.

En la fase de validación, los estadísticos históricos se calculan únicamente con el conjunto de entrenamiento. Para la fase de test, los estadísticos se recalculan usando train + validación, y después se aplican al test final. De esta manera, el modelo nunca utiliza información posterior al periodo que está intentando predecir.

Además, se guardan los valores calculados durante el entrenamiento (medias, máximos, mínimos, etc.) para poder reutilizarlos después. De esta forma, cuando el usuario introduzca una fecha nueva, esa observación se transforma exactamente igual que los datos con los que se entrenó el modelo.

5.17.3 Modelos entrenados en la versión futura

Sobre el dataset v4 se entrenaron varios modelos basados en árboles, ya que este tipo de algoritmos funciona bien con datos tabulares y permite combinar variables temporales, meteorológicas, categóricas e históricas. Los modelos utilizados fueron:

Modelo	**Finalidad dentro de v4**
Baseline histórico	Referencia basada en medias y medianas históricas
LightGBM_Futuro	Modelo de gradient boosting para datos tabulares
ExtraTrees_Futuro	Ensemble de árboles con mayor aleatoriedad
RandomForest_Futuro	Modelo de árboles independientes mediante bagging
HistGradientBoosting_Futuro	Alternativa eficiente de boosting de scikit-learn

RandomForestTuned_Futuro	Versión ajustada de Random Forest mediante búsqueda ligera
LightGBMPoisson_Futuro	Versión de LightGBM adaptada a variable de conteo
Ensemble_Futuro	Combinación de varios modelos individuales

Tabla 15. Modelos utilizados dataset futuro.

5.17.4 Modelos adicionales y ajustes específicos

Para el modelo v4 se introducen algunos elementos nuevos respecto a los otros modelos.

En primer lugar, se entrenó una versión **RandomForestTuned_Futuro**, donde se probaron varias configuraciones de hiperparámetros sobre el conjunto de validación. El objetivo era comprobar si una versión más ajustada de Random Forest podía mejorar los resultados del Random Forest inicial.

En segundo lugar, se entrenó **LightGBMPoisson_Futuro**, una variante de LightGBM con objetivo Poisson. Porque como la variable objetivo es un conteo de pacientes,y por lo tanto, nunca puede ser negativo, de esta manera, aunque no garantiza mejores resultados, permite probar una función de pérdida.

Por último, se generaron modelos ensemble, combinando las predicciones de varios modelos individuales. Se probaron tres enfoques:

- **Ensemble_Media_Futuro:** media simple de las predicciones.
- **Ensemble_Ponderado_Futuro:** combinación ponderada según el MAE obtenido en validación.
- **Ensemble_CalibradoHora_Futuro:** ensemble ponderado con una corrección adicional por hora del día.

También se añadió una calibración por hora porque algunos modelos presentaban sesgos diferentes según la franja horaria. Ya que infraestimaban más en horas pico que

en horas valle. Con esto se permitía ajustar la predicción según el error medio observado para cada hora en el conjunto de validación.

5.17.5 Selección del modelo final futuro

Para poder predecir las llegadas en un día y hora concreto, se necesitaba elegir el modelo que mejores resultados tuviera, fijándose en las mejores métricas.

Pero para este caso no se basó únicamente en el menor **MAE** global, sino también en el sesgo del modelo. Ya que en los resultados se observó que la diferencia entre los dos MAE finales era muy poca (2.1812, 2.1860), sin embargo, el **BIAS** mejoraba más el modelo con un MAE un poco más bajo (-0,3696 a -0,2338).

Entonces, en un contexto hospitalario, un modelo que infraestima de forma sistemática la demanda puede ser menos útil para la planificación, aunque tenga un error medio bajo.

Es por eso por lo que, en la comparación final, aunque, **RandomForestTuned_Futuro** obtuvo el menor MAE en validación y un rendimiento muy competitivo en test. Cuando se probó el **Ensemble_CalibradoHora_Futuro**, se observó que se reducía el sesgo negativo muy notablemente.

Por este motivo, el modelo final recomendado para la predicción manual fue el **Ensemble_CalibradoHora_Futuro**, al ofrecer un mejor equilibrio entre error global y sesgo operativo.

5.17.6 Predicción manual de una fecha futura

El último paso del modelo v4 consiste en permitir una consulta manual. El usuario introduce una fecha futura y, opcionalmente, datos meteorológicos previstos. Si alguna variable meteorológica no se introduce, el sistema la imputa utilizando valores históricos aprendidos durante el entrenamiento.

A partir de esa información, el sistema genera automáticamente todas las variables necesarias: hora, día de la semana, festivos, codificaciones cíclicas, franja horaria, estación, variables meteorológicas derivadas y patrones históricos agregados.

```
======================================================================
PREDICCIÓN MANUAL — INTRODUCCIÓN DE DATOS
======================================================================
Introduce la fecha futura, por ejemplo 2026-06-15: 22026-05-25
Formato no válido. Usa un formato como: 2026-06-15
Introduce la fecha futura, por ejemplo 2026-06-15: 2026-05-25

--- Calendario y festivos ---
¿Es festivo? (s/n) [n]: n
¿Es pre-festivo? (s/n) [n]: n
¿Es post-festivo? (s/n) [n]: n
¿Es puente? (s/n) [n]: n

--- Meteorología prevista ---
Puedes pulsar ENTER si no conoces algún dato. Se imputará con valores históricos.
Precipitación en mm [ENTER si no lo sabes]:
Temperatura mínima [ENTER si no lo sabes]:
Temperatura máxima [ENTER si no lo sabes]:
Temperatura media no calculada: falta la temperatura mínima o máxima.
El modelo tratará tmed como dato faltante y aplicará la imputación aprendida.
Humedad media [ENTER si no lo sabes]:
Humedad máxima [ENTER si no lo sabes]:
Humedad mínima [ENTER si no lo sabes]:
Velocidad media del viento [ENTER si no lo sabes]:
Racha máxima de viento [ENTER si no lo sabes]:
Horas de sol [ENTER si no lo sabes]:
Presión máxima [ENTER si no lo sabes]:
Presión mínima [ENTER si no lo sabes]:
```

Figura 8. Ejemplo de salida por consola de las variables añadidas manualmente.

El resultado final incluye:

```
================================================================================
RESULTADO — MODELO FUTURO: PREDICCIÓN DEL DÍA COMPLETO
================================================================================
Modelo usado: Ensemble_CalibradoHora_Futuro
Fecha seleccionada: 2026-05-25
Temperatura mínima introducida: nan ºC
Temperatura máxima introducida: nan ºC
Temperatura media calculada: no disponible; se imputará en el modelo.

Predicciones por hora:
       FRANJA  PREDICCION_PUNTUAL                          LLEGADAS_APROX  MAE_ORIENTATIVO RANGO_ORIENTATIVO
00:00 - 00:59               3.62   se esperan unas 3-4 llegadas en esa hora.            1.57             2 - 5
01:00 - 01:59               2.30   se esperan unas 2-3 llegadas en esa hora.            1.33             1 - 4
02:00 - 02:59               1.68   se esperan unas 1-2 llegadas en esa hora.            1.05             1 - 3
03:00 - 03:59               1.34   se esperan unas 1-2 llegadas en esa hora.            1.08             0 - 2
04:00 - 04:59               1.52   se esperan unas 1-2 llegadas en esa hora.            1.10             0 - 3
05:00 - 05:59               1.31   se esperan unas 1-2 llegadas en esa hora.            1.03             0 - 2
06:00 - 06:59               1.72   se esperan unas 1-2 llegadas en esa hora.            0.96             1 - 3
07:00 - 07:59               2.76   se esperan unas 2-3 llegadas en esa hora.            1.34             1 - 4
08:00 - 08:59               8.01   se esperan unas 8-9 llegadas en esa hora.            2.17            6 - 10
09:00 - 09:59              16.51 se esperan unas 16-17 llegadas en esa hora.            3.03            13 - 20
10:00 - 10:59              18.18 se esperan unas 18-19 llegadas en esa hora.            3.36            15 - 22
11:00 - 11:59              16.90 se esperan unas 16-17 llegadas en esa hora.            3.29            14 - 20
12:00 - 12:59              16.34 se esperan unas 16-17 llegadas en esa hora.            3.27            13 - 20
13:00 - 13:59              14.77 se esperan unas 14-15 llegadas en esa hora.            3.02            12 - 18
14:00 - 14:59               9.25  se esperan unas 9-10 llegadas en esa hora.            2.69             7 - 12
15:00 - 15:59              11.86 se esperan unas 11-12 llegadas en esa hora.            2.65             9 - 15
16:00 - 16:59              12.00 se esperan unas 11-12 llegadas en esa hora.            2.72             9 - 15
17:00 - 17:59              11.10 se esperan unas 11-12 llegadas en esa hora.            2.84             8 - 14
18:00 - 18:59              11.52 se esperan unas 11-12 llegadas en esa hora.            2.69             9 - 14
19:00 - 19:59              10.20 se esperan unas 10-11 llegadas en esa hora.            2.63             8 - 13
20:00 - 20:59               9.95  se esperan unas 9-10 llegadas en esa hora.            2.58             7 - 13
21:00 - 21:59               7.41   se esperan unas 7-8 llegadas en esa hora.            2.30             5 - 10
22:00 - 22:59               6.02   se esperan unas 6-7 llegadas en esa hora.            2.12              4 - 8
23:00 - 23:59               3.92   se esperan unas 3-4 llegadas en esa hora.            1.65             2 - 6

Suma orientativa de predicciones horarias del día: 200.19 llegadas.
```

Figura 9. Resultado mostrado por pantalla sobre la hora prevista.

Este rango no debe interpretarse como un intervalo de confianza exacto, sino como una referencia práctica construida a partir del error medio histórico del modelo en el conjunto de test.

5.17.7 Utilidad operativa del modelo futuro

El modelo v4 permite trasladar el trabajo desde una comparación técnica de modelos hacia una herramienta más cercana al uso real. Mientras que los modelos v1, v2 y v3 sirven para evaluar la capacidad predictiva usando información histórica y reciente, el modelo futuro permite estimar la demanda para una fecha concreta en la que todavía no existen datos operativos recientes.

Por tanto, esta versión es especialmente útil para escenarios de planificación anticipada, donde el hospital necesita una estimación previa de la demanda esperada en una determinada hora, día o franja horaria.

5.18 Desarrollo panel interactivo.

Como complemento al desarrollo del sistema predictivo, se ha creado un dashboard en Power BI orientado a facilitar la consulta e interpretación de los datos del servicio de urgencias.

Su finalidad es poder presentar de forma visual los principales patrones, volumen de atenciones, distribución temporal, horas de mayor afluencia, días críticos, tiempos de triaje y procedencia y destino de los pacientes.

Entre las variables utilizadas en el panel, se incluye la fecha y hora de llegada, el número de atenciones, los niveles de triaje, los tiempos hasta el inicio de triaje, la procedencia de los pacientes y el destino asistencial tras la atención.

El panel no sustituye al modelo predictivo, sino que actúa como una herramienta de apoyo en la toma de decisiones, permitiendo comprender el comportamiento histórico del servicio y detectar las franjas en las que puede existir una mayor presión asistencial.

5.18.1 Estructura del panel

El panel se ha organizado en seis vistas principales, cada una centrada en un aspecto concreto del análisis. Esta estructura permite ordenar la información de forma progresiva, comenzando por una visión general del servicio y avanzando hacia análisis más específicos.

Vista del dashboard	Información mostrada	Utilidad
Resumen ejecutivo de urgencias	Atenciones totales, media diaria, tiempo mediano hasta triaje, hospitalización, registros sin hora de triaje y máximo diario	Obtener una visión general del servicio
Horas pico y patrón semanal	Máximo horario, franja de mayor demanda y mapa de calor por día y hora	Identificar cuándo se concentran más llegadas
Evolución mensual y estacionalidad	Comparación mensual por años y mes de mayor afluencia	Detectar variaciones estacionales
Días críticos y picos extraordinarios	Máximo diario, umbral del percentil 95 y número de días críticos	Localizar jornadas especialmente exigentes
Triaje y tiempo hasta inicio	Mediana y percentil 90 del tiempo hasta triaje por prioridad	Analizar la rapidez asistencial según nivel de urgencia
Procedencia y destino asistencial	Origen de los pacientes y destino final tras la atención	Comprender el flujo asistencial del servicio

Tabla 16. Resumen estructura panel.

5.18.2 Utilidad dentro del proyecto

La incorporación del panel al TFG ayuda de forma que la interpretación de los resultados sea más fácil para un usuario no técnico. Además, permite que la información no quede limitada a tablas o métricas numéricas, sino que pueda consultarse de forma más intuitiva.

Desde el punto de vista operativo, esta herramienta puede contribuir a mejorar la comprensión de la demanda asistencial y servir como apoyo para la planificación de recursos. Su utilidad principal reside en ofrecer una visión clara del comportamiento del servicio de urgencias, permitiendo identificar periodos de mayor actividad y posibles situaciones de saturación.

5.19 Código GitHub

Todo el código desarrollado en el marco de este TFG ha sido documentado, organizado y publicado en un repositorio accesible a través de GitHub.

En él se incluye tanto el código de preparación y tratamiento de los datos, como la conversión de los archivos originales, el procesamiento de la información hospitalaria, la incorporación de variables relacionadas con festivos y meteorología, el tratamiento de valores atípicos y la construcción de los conjuntos de datos empleados posteriormente en el modelado.

Además, se incorporan los archivos correspondientes al desarrollo y actualización de los modelos predictivos, así como el archivo de Power BI utilizado para la elaboración del panel informativo de urgencias. La documentación y los archivos asociados al desarrollo técnico del trabajo pueden consultarse en el siguiente enlace: https://github.com/LeyreVerduu/TFG

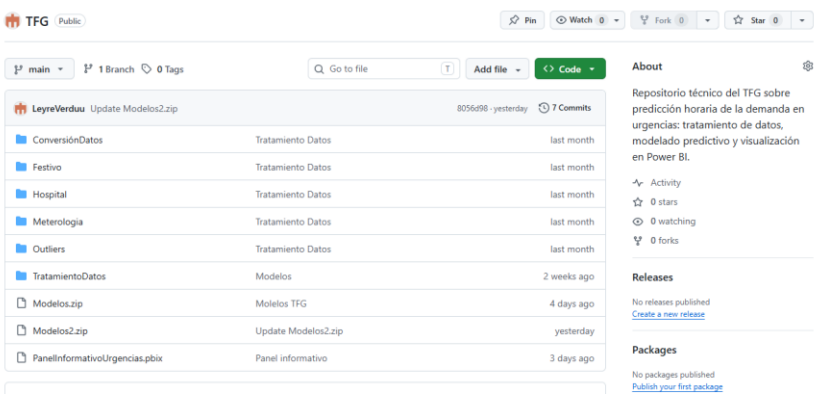

Figura 10. Captura de pantalla del repositorio de GitHub

6. RESULTADOS

Con el desarrollo finalizado, el objetivo de este apartado es presentar los resultados obtenidos tras entrenar y evaluar los modelos predictivos sobre el conjunto de test. Para ello, se dividirán en dos partes diferentes: la predicción de la hora siguiente a partir de información reciente disponible (dataset v1, v2, v3) y, por otro lado, la predicción manual de una fecha futura (dataset v4).

Es importante realizar una distinción debido a la diferente finalidad de ambas partes.

6.1 Distribución horaria y semanal

Para poder valorar una predicción real, es importante entender cómo se comporta la llegada de pacientes a urgencias. A través de los datos, se ha demostrado que la distribución no es constante, sino que durante la madrugada se registran menos llegadas, mientras que aumenta durante la mañana. Observando un punto clave de mayor afluencia, los lunes a las 11:00h con 17.9 llegadas por hora.

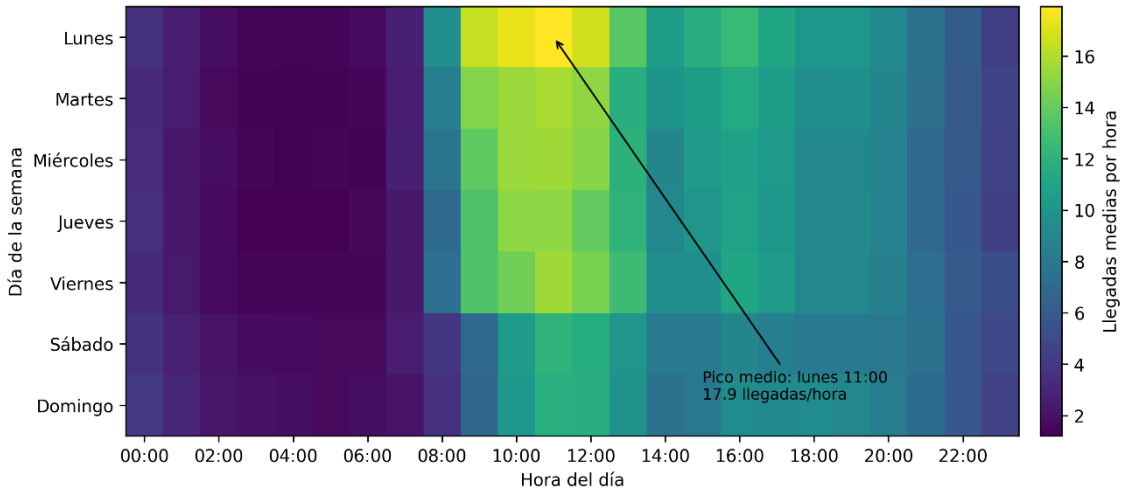

Figura 11. Demanda media de pacientes según el día de la semana y la hora del día.

En el mapa de calor se observa de manera directa, como las franjas de mañana, sobre todo al inicio de la semana presentan la mayor carga de llegadas. Siendo estas donde es más importante poner una estimación anticipada.

6.2 Resultados evolución: v1, v2 y v3.

En todo momento el TFG, evaluó si la incorporación de nuevas variables permitía mejorar la predicción de llegadas en la hora siguiente. Por ello, se crearon 3 dataset. El v1, el primer dataset con variables más sencillas de temporal, de calendario, de meteorología y de actividad previa del servicio. Al probar todos los modelos con el primer dataset, se obtuvo el mejor MAE mediante el modelo LSTM, con un error medio de 2,1323 pacientes por hora.

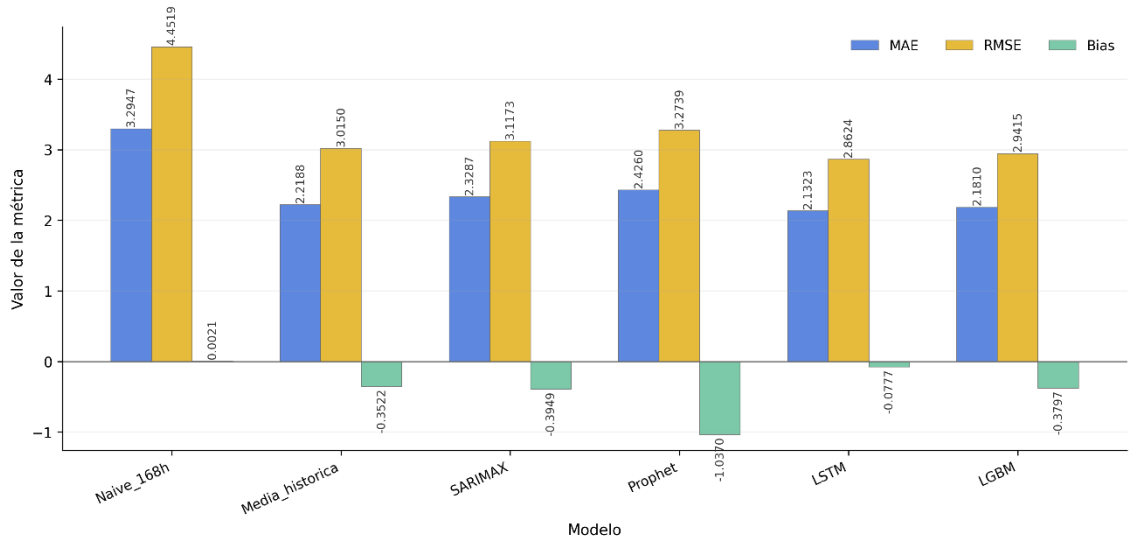

Figura 12. Resultados dataset v1.

Con el fin de mejorar el rendimiento predictivo, se creó el dataset v2, incorporando patrones históricos de demanda, como medias de llegadas según la hora y el día de la semana o la hora y el mes. Con estas nuevas variables se volvieron a probar los modelos, y los resultados de v2 mejoraron, pero no tanto como se esperaba.

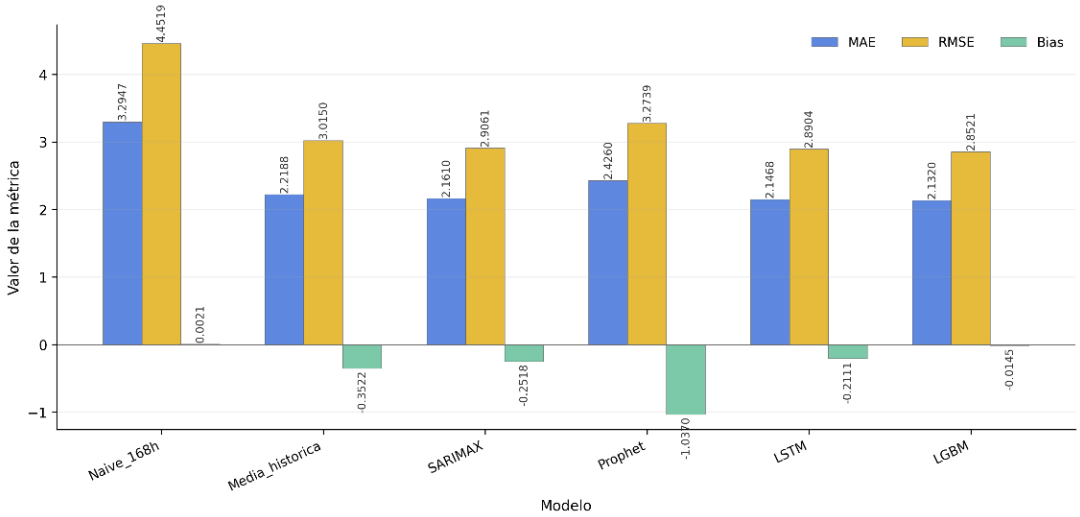

Figura 13. Resultados dataset v2

Es por eso por lo que surge v3, una versión más compleja que añadía variables temporales nuevas. Tras probar los mismos modelos para este dataset, se observó que los resultados eran peores que los de v1, entendiendo así que no por añadir variables más complejas el modelo las captaría mejor.

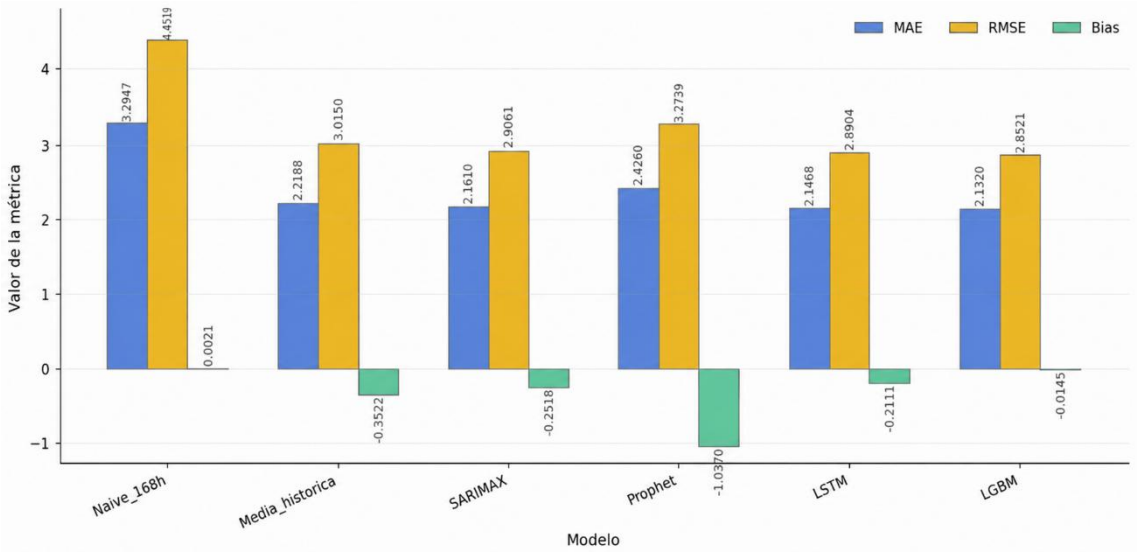

Figura 14. Resultados dataset v3

Tras observar los resultados y por ser el dataset con variables que más valor pueden aportar. Se elige el dataset v2, como el mejor de los 3, y se prueban 3 modelos más con el fin de intentar mejorar los resultados.

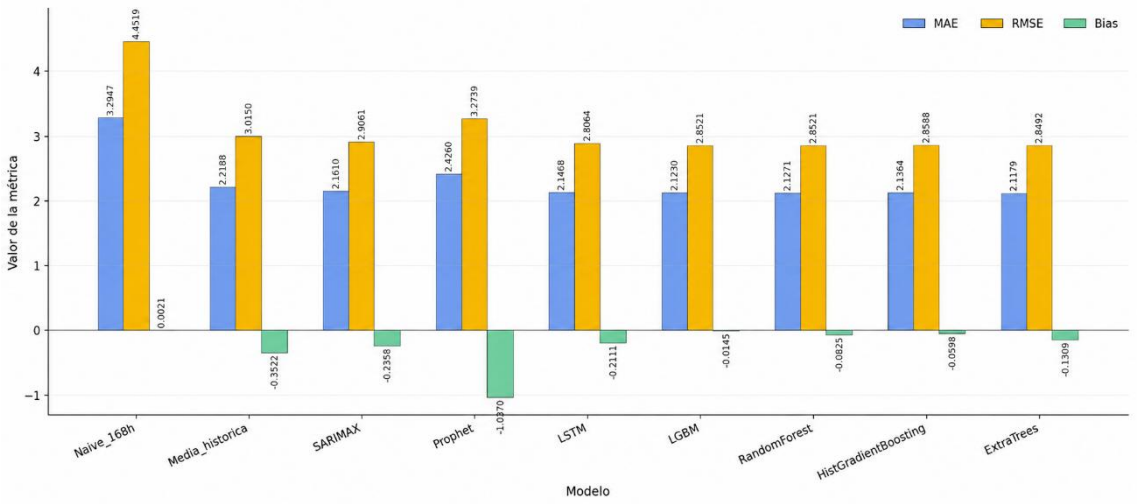

Figura 15. Resultados dataset v2 nuevos modelos añadidos.

Con los nuevos resultados de los modelos probados, se selecciona ExtraTrees, como el modelo que mejores resultados obtiene. Su error es de 2,1179 pacientes por hora, siendo el valor más bajo de todos los modelos probados sobre el dataset v2. Además, también obtiene el menor RMSE, lo que indica que controla mejor los errores más grandes.

Aunque LightGBM presenta un bias más cercano a cero, esto solo significa que sus predicciones están más equilibradas entre quedarse por encima o por debajo de la realidad. No significa necesariamente que acierte más.

En la siguiente tabla se recoge los resultados de los modelos que mejores resultados han dado.

Versión	Mejor modelo	MAE	RMSE	Bias
v1 Base	LSTM	2,1323	2,8624	−0,0777
v2 Features históricas	ExtraTrees	2,1179	2,8492	−0,1509

v3 Features avanzada	LightGBM	2,1324	2,8539	−0,0303

Tabla 17. Mejor resultado obtenido en cada dataset.

6.3 Comparación real-predicción v2

Para comprobar que el resultado sea realista, se compararon las llegadas registradas con las estimadas por el modelo durante el mes de enero de 2025.

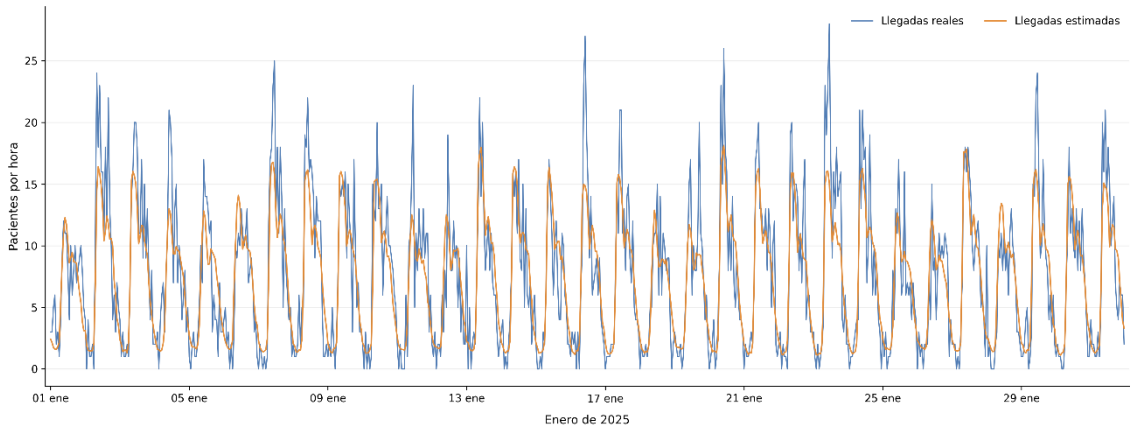

Figura 16. Comparación entre llegadas reales y predicciones durante el mes de enero 2025.

El modelo reproduce el comportamiento general de la demanda, identificando las horas de menor actividad durante la madrugada y el aumento habitual de llegadas durante las franjas diurnas. No obstante, las predicciones son más suaves que los valores reales y presentan mayor dificultad para anticipar algunos picos puntuales de afluencia. Por ello, el resultado debe interpretarse como una estimación del nivel esperado de demanda y no como una predicción exacta de episodios extraordinarios.

Además, el error no es igual en todas las franjas horarias. La mañana concentra simultáneamente una mayor llegada de pacientes y una mayor dificultad de predicción. En la siguiente figura se compara el error del modelo final seleccionado, ExtraTrees sobre el dataset v2, con una referencia básica basada en repetir la demanda registrada en la misma hora de la semana anterior (Naive 168h).

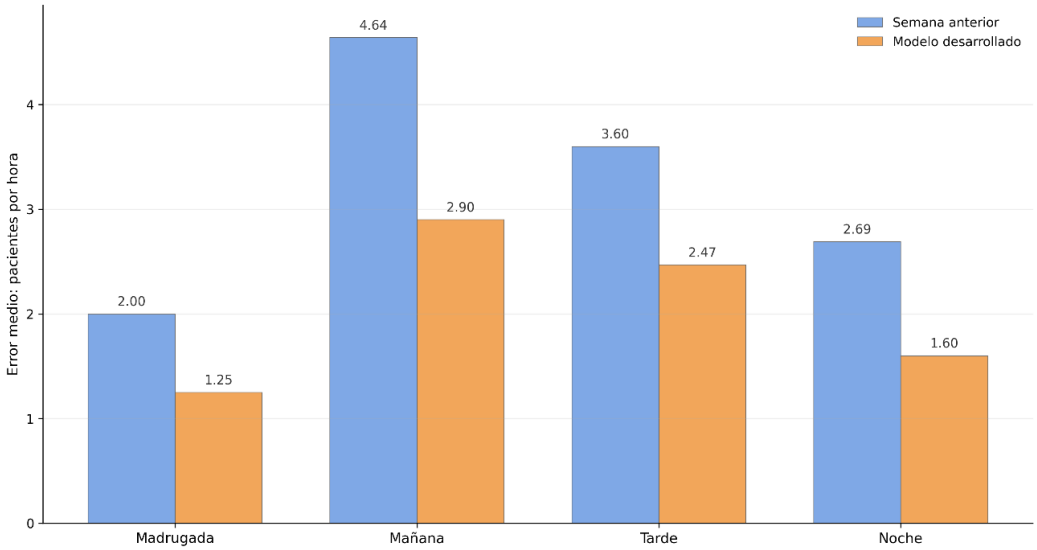

Figura 17. Error medio por franja horaria: estimación basada en la semana anterior frente al modelo desarrollado

En el gráfico se observa que franja comprendida entre las 8:00 y las 14:00 horas es la más difícil de estimar. En ella, repetir la demanda de la semana anterior genera un error medio de 4,64 pacientes por hora, mientras que el modelo final lo reduce a 2,90 pacientes por hora. Este resultado indica que el sistema mejora especialmente la estimación en uno de los periodos de mayor interés asistencial.

6.4 Utilidad del modelo v2 para el servicio de urgencias

Con los resultados del modelo ExtraTrees, podemos estimar cuántos pacientes podrían llegar al servicio de urgencias en la hora siguiente. Esto serviría como ayuda a corto plazo, utilizando información reciente de la actividad.

Con este modelo podríamos ayudar a detectar picos en las franjas en la que se concentran más llegadas de pacientes y donde resulta más difícil realizar una estimación precisa.

Aunque el modelo no permite saber con exactitud cuántos pacientes acudirán, sí puede estimar con un error de 2 pacientes, un rango de las posibles personas que pueden acudir a urgencias en la hora siguiente.

De esta forma, el sistema podría servir como apoyo para que el hospital preste una mayor atención a determinados momentos del día y pueda anticiparse a posibles situaciones de mayor presión asistencial.

Sin embargo, con este modelo no es posible calcular cuantas personas llegarán a urgencias un día concreto, ya que utiliza información reciente del servicio que todavía no existiría en ese momento.

Es por eso, que se desarrolló una segunda versión, denominada v4, orientada a estimar la demanda de una jornada futura utilizando únicamente información conocida de antemano.

6.5 Modelo v4 para predicción manual futura

Además del modelo anterior que estima la llegada de pacientes en la hora siguiente, el objetivo de este TFG también quería comprobar si con los datos de los que se disponía sería posible encontrar un modelo que pudiera estimar de manera real sin inventarse información, una versión donde el usuario pudiera introducir una fecha y automáticamente desglosar la franja de pacientes diarios que acudirían a urgencias ese día.

Por ello, se creó el dataset v4, que permitía cumplir con esta función. Esta versión utiliza únicamente información disponible antes de que transcurra la jornada, como calendario, festivos, meteorología prevista y patrones históricos de demanda. Es por eso, que se excluyen variables como las llegadas inmediatamente anteriores o la ocupación observada en tiempo real, y utiliza información disponible o estimable de antemano.

Para el dataset v4, se prueban los mismos modelos que para los anteriores dataset, pero adaptados a las nuevas necesidades de las variables y buscando más modelos que garanticen un mejor resultado.

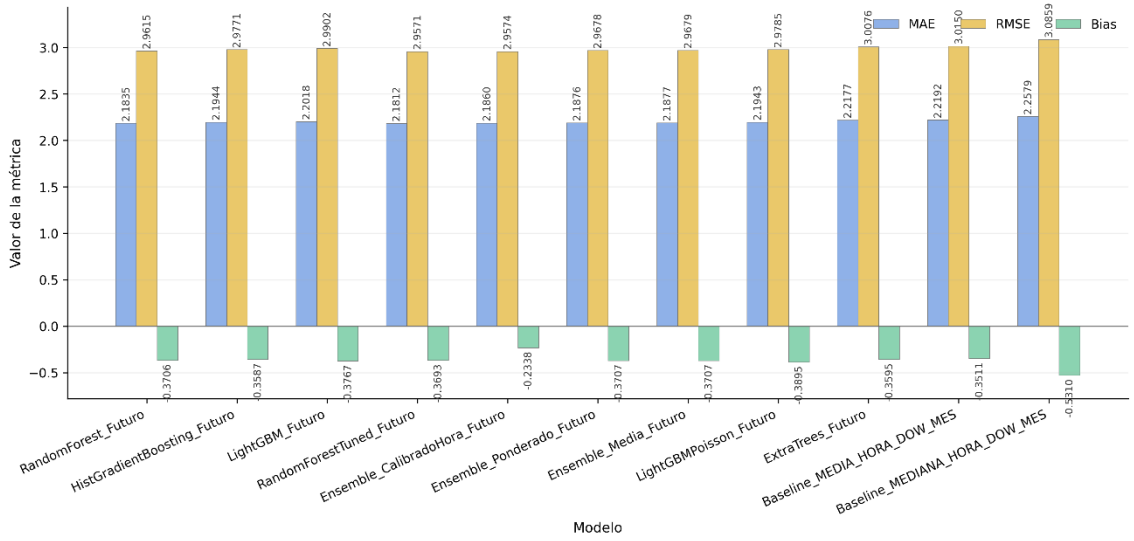

Figura 18. Resultados modelos v4

Se observa como modelo con menor MAE fue RandomForestTuned_Futuro, con un error medio de 2,1812 llegadas por hora. Sin embargo, el modelo Ensemble_CalibradoHora_Futuro obtuvo un MAE prácticamente idéntico (2,1860) y redujo de forma notable la infraestimación media de la demanda.

Por este motivo, se selecciona, porque RandomForestTuned_Futuro presenta una mayor tendencia a predecir menos pacientes de los que realmente llegan, mientras que Ensemble_CalibradoHora_Futuro reduce esta infraestimación. Esto es muy importante en el servicio de urgencias, ya que una estimación baja podría hacer que se anticipase una demanda inferior a la real.

Es por este motivo, que, aunque Ensemble_CalibradoHora_Futuro no obtiene el MAE más bajo, se considera la opción más adecuada para la predicción anticipada de una jornada futura. Ya que mantiene prácticamente la misma precisión que el mejor modelo y, al mismo tiempo, ofrece una estimación más prudente al reducir el riesgo de quedarse corto en la demanda prevista.

Modelo	MAE	RMSE	Bias	Decisión
RandomForestTuned Futuro	2,1812	2,9571	−0,3693	Menor MAE
Ensemble CalibradoHora Futuro	2,1860	2,9574	−0,2338	Modelo recomendado

Tabla 18. Resumen comparación de los dos modelos principales para la selección final de v4.

6.5.1 Ejemplo de uso rango de estimación anticipada

Para ver cómo sería un ejemplo de uso, la figura siguiente muestra la salida obtenida al introducir como ejemplo el día 25 de mayo de 2026.

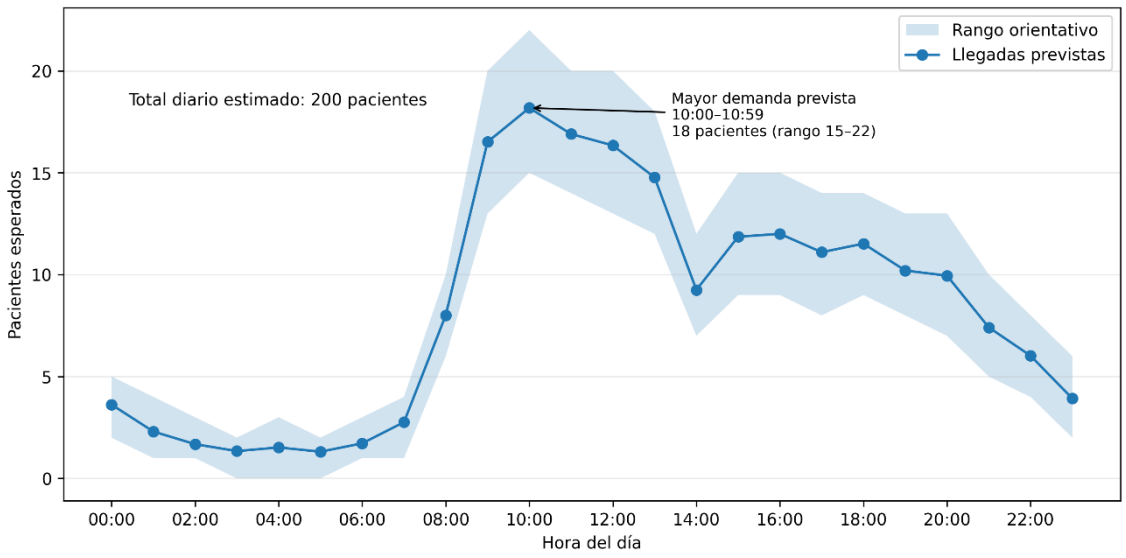

Figura 19. Ejemplo de estimación anticipada de la demanda por horas para un día completo.

Se observa como con el modelo seleccionado, el sistema estima un total aproximado de 200 llegadas. La mayor demanda prevista aparece entre las 10:00 y las 10:59 horas, con unas 18 llegadas esperadas y un rango orientativo de entre 15 y 22 pacientes.

Esta salida facilita localizar visualmente las horas en las que podría ser necesario prestar una mayor atención a la planificación del servicio.

Y de esta manera, aunque no es un intervalo de confianza exacto, el modelo permite obtener una estimación orientativa de los pacientes que podrían llegar al servicio de urgencias a lo largo de un día concreto.

6.6 Resultados del dashboard

En este apartado se comentarán los principales resultados obtenidos del panel visual, usado como herramienta para facilitar la comprensión.

6.6.1 Vista 1. Resumen ejecutivo de la actividad

En la primera vista del dashboard se observa una visión general de la actividad registrada en el servicio de urgencias durante el periodo analizado, del 1/12/2021 al 30/11/2025. El resumen ejecutivo recoge cerca de 254.000 atenciones, con una media de 173,87 pacientes al día y un máximo diario de 272.

Esto nos demuestra que la actividad no es estable, ya que existen jornadas donde la demanda supera el comportamiento medio. La diferencia entre la media diaria y el máximo diario son casi 100 pacientes, lo que permite identificar la existencia de días con mayor presión asistencial.

En el panel también se muestra un tiempo mediano hasta el inicio del triaje de 6 minutos. Esto significa que, entre los registros con hora de triaje disponible, la mitad de los pacientes inició el proceso en 6 minutos o menos. No obstante, también se observa que un 6,14 % de los registros no dispone de hora de triaje, por lo que este indicador debe interpretarse teniendo en cuenta esta limitación en la calidad del dato.

Así mismo, la tasa de hospitalización fue del 17,01 %, lo que indica que aproximadamente uno de cada seis pacientes atendidos en urgencias necesitó continuar su atención mediante ingreso hospitalario. Este dato es importante porque la presión sobre urgencias no depende únicamente del número de llegadas, sino también de cuántos pacientes requieren recursos hospitalarios posteriores, como camas.

Finalmente, el gráfico de atenciones por año permite observar la evolución del volumen asistencial durante el periodo estudiado. Los años completos muestran un volumen anual elevado y relativamente estable, con un ligero aumento entre 2022 y 2024. Sin embargo, la comparación debe realizarse con cautela, ya que 2021 solo incluye un mes y 2025 no contiene el mes de diciembre.



Figura 20. Vista de resumen ejecutivo de la actividad registrada en el servicio de urgencias durante el periodo analizado.

6.6.2 Vista 2. Patrón horario y semanal

En la vista del panel se analizan las horas pico y el patrón semanal de llegada de pacientes a urgencias. Con esta visualización se observa que la demanda no se reparte de manera uniforme a lo largo del día ni de la semana, sino que mantiene una estructura.

En primer lugar, se aprecia que el pico horario máximo registrado en el periodo fue de 41 pacientes en una única hora, concretamente el 25/08/2025 a las 12:00, lo representa una demanda elevada.

Además, las llegadas se concentran sobre todo en las horas centrales de la mañana, especialmente los lunes, con el pico medio situado los lunes a las 11:00, con una media de 17,89 llegadas por hora. Esto encaja con el análisis del error por hora de los modelos, donde estas franjas resultaban más difíciles de predecir por su mayor volumen y variabilidad.

El mapa de calor permite visualizar de forma clara este comportamiento. Las horas de madrugada presentan una menor afluencia, mientras que la demanda aumenta progresivamente a partir de las primeras horas de la mañana y alcanza sus valores más elevados entre las 9:00 y las 13:00. Este patrón se aprecia especialmente al inicio de la semana, donde los lunes concentran los valores medios más altos.



DiaSemana	00:...	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:(
lunes	3,64	2,61	1,82	1,52	1,45	1,43	1,68	2,67	9,35	16,43	17,38	17,89	16,67	13,51	10,50	11,73	12,45	11,12	10,06	10,(
martes	3,19	2,46	1,59	1,37	1,32	1,34	1,40	2,45	8,43	14,85	15,43	15,84	15,13	11,79	9,88	10,58	11,44	10,55	9,39	9,5
miércoles	3,31	2,33	1,84	1,44	1,28	1,51	1,39	2,68	7,65	13,78	15,40	15,51	15,02	11,97	8,85	10,38	10,80	10,22	9,13	8,9
jueves	3,41	2,33	1,68	1,20	1,25	1,23	1,54	2,36	7,06	13,28	15,06	15,09	13,90	12,11	8,99	9,91	10,82	10,02	8,95	9,7
viernes	3,23	2,43	1,70	1,51	1,47	1,41	1,45	2,33	7,27	13,34	14,33	15,55	14,56	12,66	9,42	9,62	11,03	10,32	8,96	9,(
sábado	3,63	2,73	2,11	1,74	1,67	1,64	1,64	2,46	3,82	7,00	10,44	12,10	11,60	9,96	8,00	8,19	8,80	8,46	8,11	8,7
domingo	4,05	3,11	2,27	2,00	1,91	1,64	1,81	2,09	3,43	6,70	10,13	11,82	11,72	9,82	7,44	7,94	9,38	9,15	9,34	9,(

Figura 21. Distribución horaria y semanal de la demanda asistencial mediante mapa de calor.

6.6.3 Vista 3. Evolución mensual y estacionalidad

La tercera vista del panel permite seguir la evolución mensual y observar posibles patrones estacionales. Esta visualización resulta útil porque permite comprobar si existen meses en los que se concentra una mayor actividad asistencial.

En el periodo analizado, agosto destaca como el mes con más atenciones acumuladas, alrededor de 24.397 atenciones. Además, al comparar con el grafico los años, se aprecia diferencias en el volumen asistencial registrado en los distintos meses.

Al comparar los años completos y los periodos comparables, se observa un incremento del 8,37 % entre 2022 y 2025, situándose 2025 como el año con mayor demanda comparable. Este resultado sugiere una tendencia de crecimiento en la actividad del servicio durante el periodo estudiado.

Esta información puede ser útil para identificar meses que históricamente presentan una mayor carga de trabajo y que, por tanto, podrían requerir una mayor vigilancia en futuras planificaciones. Como los meses de agosto, abril o julio con una mayor carga asistencial.

No obstante, esta interpretación debe realizarse con prudencia, ya que el volumen mensual puede estar influido por factores que no se han analizado de forma específica en este trabajo, como cambios poblacionales temporales, vacaciones, eventos locales o variaciones en el acceso a otros servicios sanitarios.



Figura 22. Evolución mensual de las atenciones y comparación de la demanda entre años.

6.6.4 Vista 4. Días críticos y picos extraordinarios

La cuarta vista del panel se centra en la identificación de días críticos y picos extraordinarios de demanda. Para ello se establece como referencia el percentil 95 de la demanda diaria.

Con este criterio, se fija el umbral de día crítico en 218 pacientes atendidos en una misma jornada. Esto significa que los días que igualan o superan este valor se sitúan dentro del 5 % de jornadas con mayor afluencia del periodo analizado. En total, se identifican 77 días críticos, lo que muestra que las situaciones de elevada presión asistencial no son únicamente episodios aislados, sino eventos que se repiten a lo largo del periodo estudiado.

El máximo registrado fue 272 pacientes, correspondientes al lunes 22/04/2024. Este valor está por encima del umbral crítico, por lo que representa una jornada especialmente exigente para el servicio de urgencias.

El gráfico de dispersión permite observar visualmente cómo la mayoría de los días se concentran por debajo del umbral del percentil 95, mientras que algunos puntos aparecen por encima de la línea roja, identificando las jornadas de mayor presión. Esta información resulta útil porque permite diferenciar entre la actividad habitual del servicio y aquellos días en los que la demanda alcanza niveles extraordinarios.

Desde un punto de vista operativo, esta vista puede ayudar a conocer con qué frecuencia se producen jornadas críticas y qué volumen de pacientes puede considerarse especialmente elevado para el hospital.



Figura 23. Identificación de días críticos mediante el umbral del percentil 95 de demanda diaria.

6.6.5 Vista 5. Triaje y tiempos de espera

En la quinta vista se analiza el tiempo transcurrido desde la llegada del paciente hasta el inicio del triaje, diferenciando además por nivel de prioridad asignado. A través de esta información se valora no solo el volumen de pacientes atendidos, sino parte del proceso inicial de atención en urgencias.

Durante el periodo analizado, se registraron 238.438 pacientes con hora de inicio de triaje disponible. La mediana global hasta el inicio del triaje fue de 6 minutos, lo que significa que la mitad de los pacientes con registro válido inició el triaje en 6 minutos o menos. Además, el percentil 90 se sitúa en 14 minutos, indicando que el 90 % de los pacientes con datos disponibles comenzó el triaje en un tiempo igual o inferior a ese valor.

Además, también se observa que un 6,14% de los registros no dispone de hora de inicio de triaje, por lo que estos casos no pueden utilizarse para calcular este indicador temporal.

Al analizar el tiempo hasta el inicio del triaje según el nivel de prioridad, se observan valores relativamente similares entre los niveles II, III, IV y V, con medianas entre 5 y 6 minutos y percentiles 90 entre 14 y 15 minutos. En cambio, el nivel I presenta una mediana de 8 minutos y un percentil 90 de 26 minutos, valores superiores al resto de categorías.

Este resultado debe interpretarse con cautela, ya que puede estar influido por el bajo número de casos, por diferencias en el registro administrativo de la hora de inicio del triaje o por el hecho de que, en situaciones de mayor gravedad, la atención clínica pueda iniciarse antes de que el triaje quede registrado en el sistema. Por tanto, este indicador no debe entenderse necesariamente como un mayor tiempo real de espera para los pacientes más urgentes.

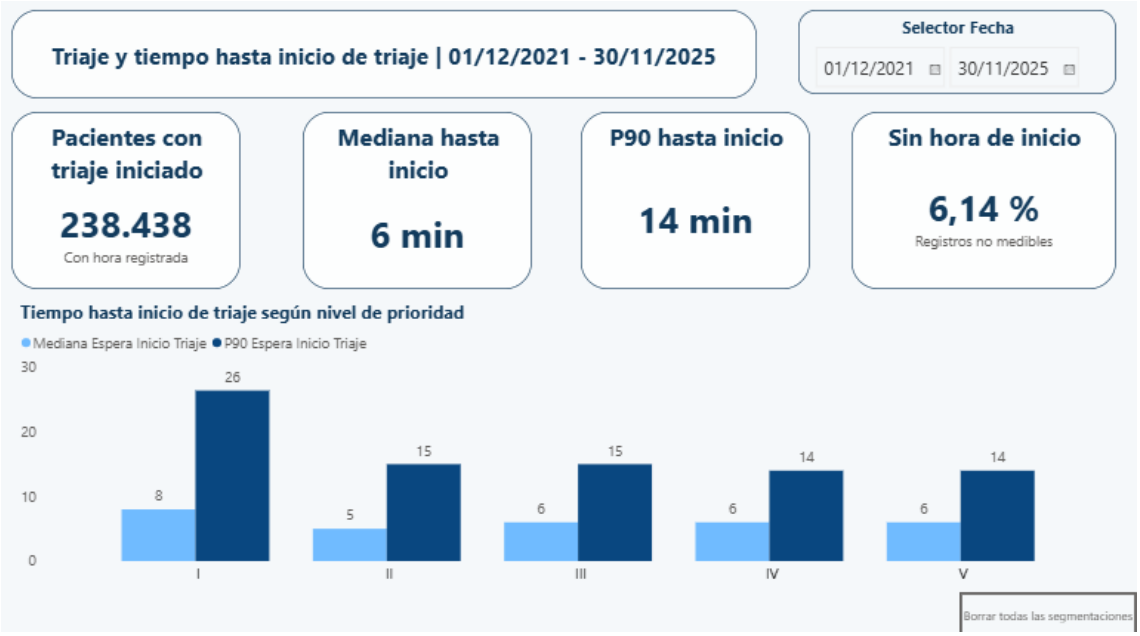

Figura 24. Tiempo hasta el inicio del triaje según el nivel de prioridad asignado al paciente.

6.6.6 Vista 6. Procedencia y destino asistencial

En la última vista del panel, se analizará la procedencia de los pacientes atendidos en urgencias y su destino final tras ser atendidos. Esto es importante ya que no solo interesa conocer cuántos pacientes acuden al servicio, sino también cómo acceden a él y qué ocurre después de la valoración médica.

En relación con la procedencia, la principal llegada es por propia iniciativa del paciente, que representa el 66,11 % de las llegadas. Esto indica que la mayoría de los pacientes acuden directamente al servicio de urgencias. En segundo lugar, aparece la Atención Primaria, con aproximadamente 56.000 pacientes, seguida de Emergencias 112, que representa el 9,82 % de las llegadas.

Luego se observa cómo el resultado más frecuente es el alta a domicilio, con alrededor de 206.000 pacientes. Esto indica que una parte mayoritaria de los episodios se resuelve dentro del propio servicio de urgencias. Sin embargo, también se observa que aproximadamente 43.000 pacientes finalizaron con hospitalización, lo que representa una tasa del 17,01 % sobre el total de atenciones.

Este dato es relevante porque la presión asistencial no termina en la llegada del paciente a urgencias. Cuando una proporción importante de pacientes requiere ingreso

hospitalario, la demanda también afecta a otros recursos del hospital, como camas, personal de planta y circuitos de hospitalización.



Figura 25. Distribución de la procedencia de los pacientes y del destino final tras la atención en urgencias.

En conjunto, el dashboard refleja de forma visual los principales patrones detectados en el análisis: la concentración de la demanda en ciertas franjas, los días de mayor presión, el comportamiento mensual y las características básicas del flujo de pacientes. Cabe matizar que, en su versión actual, el panel tiene una función principalmente descriptiva. Como línea de mejora futura, podría integrarse con el modelo v4 para mostrar predicciones de fechas concretas junto con sus rangos orientativos de error.

7. DISCUSIÓN

Los resultados demuestran que sí es posible predecir las llegadas por hora a urgencias usando los datos históricos del hospital junto con información del calendario y meteorológica.

Pero también que su resultado no solo es su error medio, sino que también debe encajar con la información realmente disponible en cada momento y ofrecer resultados que sean útiles e interpretables para planificar la atención.

7.1 Interpretación de los principales hallazgos

Tras realizarse un análisis de los resultados de los modelos, se obtienen hallazgos importantes:

- Se observa que, la demanda tiene una estructura temporal muy clara. Al pasar de v1 a v2 demuestra que las variables históricas agregadas aportan información muy valiosa, hasta el punto de que los modelos de árboles se colocan en las primeras posiciones. Saber cómo se comporta la demanda en una hora y tipo de día concretos es muy útil.

- El segundo hallazgo es que añadir complejidad no siempre mejora los resultados. Las variables nuevas de v3 no redujeron el MAE de v2, aunque algunas tenían una correlación alta con la variable objetivo.

- La importancia entre ser preciso y ser útil en la práctica. En v2, ExtraTrees consiguió el menor MAE, pero LightGBM tuvo un bias casi nulo y generó intervalos de predicción razonables. Lo mismo ocurrió en v4, RandomForestTuned fue más preciso, pero el ensemble calibrado redujo la infraestimación en las franjas críticas. En un hospital, esta diferencia es muy importante.

7.2 Relación con investigaciones anteriores

Los resultados de este TFG coinciden con varios estudios previos.

1. **Porto y Fogliatto (2024)** compararon modelos de machine learning en once servicios de urgencias de tres países y observaron que el rendimiento mejoraba al incluir variables de calendario, meteorología e ingeniería de características. Lo que se puede comparar directamente con la versión v2, que añade patrones históricos, es la que consigue el menor error, y los mejores modelos son los basados en árboles.

2. También se encuentran coincidencias con **Sudarshan et al. (2021)**, que probaron modelos como Random Forest y LSTM con información meteorológica y de calendario para predecir llegadas a urgencias.

3. La revisión de **Petravić et al. (2025)** indica que el calendario y la meteorología son las variables externas más usadas en este tipo de modelos, aunque también critica la falta de uniformidad en las metodologías y la necesidad de validar y presentar mejor los resultados.

4. Por último, el uso de intervalos de predicción coincide con **Rostami-Tabar et al. (2024)**, que defienden las predicciones probabilísticas horarias como apoyo para tomar decisiones en urgencias. Los intervalos P10-P90 de LightGBM no sustituyen una evaluación operativa real, pero aportan una idea de la incertidumbre más útil que una única predicción puntual.

7.3 Resultados discrepantes con la literatura

Al igual que hay estudios que defienden ideas comunes del TFG, también se encuentran otros estudios que defienden otras ideas.

Choudhury y Urena (2020), por ejemplo, compararon modelos de llegadas horarias a urgencias y seleccionaron ARIMA como el enfoque de mejor ajuste en su contexto. Cuando en este TFG hemos probado SARIMAX y los resultados han sido superados por otros modelos.

Esto se puede deber por varios motivos, y es que los modelos no usan la misma información siempre. Hay veces donde se incluyen variables históricas como en este TFG y otras donde se lleva un enfoque univariante, dependiendo casi únicamente de la propia serie de llegadas.

También, que cada hospital tiene sus propios patrones, así que un algoritmo que funciona bien en un centro no tiene por qué hacerlo en otro. Al igual que con los criterios de selección también pueden cambiar, ya que en este TFG no solo se valora el MAE, sino también el bias, la estabilidad por franjas horarias y la utilidad para la planificación futura.

8. OBJETIVOS DE DESARROLLO SOSTENIBLE (ODS)

Este TFG se ha planteado con el propósito de contribuir al cumplimiento de los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas. En concreto, las contribuciones más directas del proyecto se relacionan con los siguientes objetivos:

8.1 ODS 3. Salud y bienestar

El ODS 3 busca mejorar la salud y garantizar una atención de calidad para todas las personas. Este objetivo está directamente relacionado con el TFG, ya que el proyecto busca ayudar a mejorar la organización del servicio de urgencias del Hospital de Cuenca.

El modelo permite anticipar cuántos pacientes pueden llegar a urgencias en un día concreto. Esta información puede ser útil para planificar mejor los turnos del personal, preparar los recursos necesarios y prestar especial atención a las franjas con mayor demanda.

Garantizando así una mejora de la atención a futuros pacientes, ya que una mejor previsión de la demanda puede ayudar a reducir situaciones de saturación y disminuir los tiempos de espera. Favoreciendo así una atención más organizada y de mayor calidad para los pacientes.

8.2 ODS 9. Industria, innovación e infraestructura

El ODS 9 busca fomentar la innovación y el uso de nuevas tecnologías para mejorar los servicios y las infraestructuras. En el trabajo, se aplican técnicas de ciencia de datos y predicción a un servicio público. Transformando los datos disponibles del hospital en una herramienta útil para anticipar la demanda futura de pacientes.

Contribuyendo así a la modernización de los servicios sanitarios, mediante el uso de la tecnología para apoyar la toma de decisiones y mejorar la planificación de los recursos disponibles.

8.3 ODS 10. Reducción de las desigualdades

El ODS 10 busca reducir las desigualdades, también en el acceso a servicios públicos básicos como la sanidad. Este objetivo se relaciona de forma indirecta con el TFG, ya que una mala planificación en urgencias puede hacer que algunos pacientes tengan que esperar más tiempo dependiendo del momento en el que acudan al hospital.

El proyecto puede contribuir a que la atención recibida dependa menos de si el paciente acude en una hora con mucha afluencia o en un momento más tranquilo.

8.4 ODS 11. Ciudades y comunidades sostenibles

Además de los objetivos anteriores, el TFG también puede relacionarse de forma complementaria con el ODS 11.

El ODS 11, busca mejorar la calidad de vida de las personas mediante servicios públicos más eficientes. En este caso, el proyecto contribuye a mejorar la planificación de un servicio esencial para la población de Cuenca y sus alrededores, como es la atención de urgencias.

9. CONCLUSIONES

El presente Trabajo Fin de Grado se ha centrado en el desarrollo de un sistema capaz de predecir la demanda de pacientes en el servicio de urgencias del Hospital Virgen de la Luz de Cuenca. Utilizando datos reales del hospital junto con información de calendario y meteorología.

Además, se ha creado un panel interactivo, con el objetivo de facilitar su consulta y servir de apoyo para la planificación de recursos.

En este apartado se presentan las conclusiones del trabajo, su cumplimiento con los objetivos planteados, además, de encontrar las limitaciones durante el desarrollo y las posibles líneas de trabajo futuro.

9.1 Conclusiones respecto al objetivo general

Se puede afirmar que el objetivo principal del TFG se ha cumplido, ya que se ha conseguido desarrollar un sistema capaz de estimar cuántos pacientes pueden llegar al servicio de urgencias en una hora concreta.

Para construir este sistema se han utilizado datos históricos reales del hospital, combinados con información de calendario, meteorología y patrones anteriores de demanda.

Los modelos que mejores resultados obtuvieron fueron el ExtraTrees aplicado al dataset v2, para predecir la t +1 hora, y el modelo de Ensemble aplicado al dataset v4, para predecir una hora futura de un día futuro cualquiera.

Por tanto, se afirma que el uso de modelos predictivos puede ser útil para anticipar la demanda de urgencias y ofrecer información de apoyo para una mejor planificación del servicio.

9.2 Conclusiones respecto a los objetivos específicos

En relación con los objetivos específicos planteados, el trabajo ha partido de los datos originales de urgencias, llevando a cabo una limpieza de ellos. Arreglando fechas, unificando categorías y revisando tiempos de estancia. Con todo esto, se consiguió un total de 254.025 episodios, garantizando una buena base para los modelos de entrenamiento.

Siguiendo con otro de los objetivos, se agruparon los datos por horas para convertirlos en una serie temporal. Obteniendo un dataset final, formado por 34.894 observaciones, una por cada hora del periodo estudiado. De esta forma se pudo analizar la evolución de las llegadas hora a hora e identificar patrones de demanda.

Al enriquecer el dataset con información externa, variables de festivos, calendario y datos meteorológicos de la AEMET, se cumplió con otro de los requisitos del trabajo. Aunque la información meteorológica disponible presenta la limitación de tener resolución diaria, su inclusión permitió enriquecer el contexto de la predicción.

Asimismo, también se generaron variables orientadas en el comportamiento histórico de la demanda, como medias móviles y retardos Estas variables resultaron especialmente relevantes, ya que los resultados mostraron que la demanda de urgencias presenta patrones temporales repetitivos que pueden aprovecharse para mejorar la capacidad predictiva del sistema.

También se cumplió con el objetivo de comparar distintas versiones del dataset para ver cuál funcionaba mejor. El punto de partida fue el dataset v1, después se añadieron más variables al v2, y por último v3, con características más complejas.

En el TFG se evaluaron numerosos modelos de distintas naturalezas: modelos de referencia, modelos estadísticos como SARIMAX y Prophet, modelos basados en árboles como LightGBM, HistGradientBoosting, ExtraTrees y Random Forest, y un modelo secuencial de deep learning basado en LSTM.

Además, se añadió una cuarta versión, la v4, pensada para predecir manualmente una fecha y hora concretas en el futuro. La diferencia con los modelos anteriores es que esta versión solo usa información que ya se conoce o se puede estimar antes del momento de la consulta, sin depender de datos operativos recientes que en una planificación anticipada todavía no existirían.

Por último, el diseño del dashboard permite trasladar los resultados del análisis a una herramienta visual más accesible. El panel se organiza en varias vistas que recogen los grandes números del servicio, el patrón semanal y las horas pico, la evolución mensual y estacional, los días críticos por encima del percentil 95, los tiempos de triaje por nivel de prioridad y la procedencia y destino de los pacientes. Pudiendo en un futuro incorporar los datos futuros y hacer un análisis.

Se concluye este apartado, afirmando que se han cumplido todos los objetivos específicos que se plantearon al inicio del TFG. Se han preparado y analizado los datos del servicio, construido varios datasets predictivos, comparado distintos modelos con un esquema temporal, se ha desarrollado una versión pensada para consultas futuras y presentado los resultados en un dashboard visual.

9.3 Limitaciones del estudio

Como en todos los estudios, siempre hay limitaciones que te impiden mejorar los resultados o alcanzar el objetivo principal que se establecía. Para este TFG las principales limitaciones encontradas han sido:

- El estudio se basa en datos de un único hospital. Los resultados reflejan el comportamiento del modelo en el Hospital Virgen de la Luz de Cuenca, pero no se sabe cómo se comportarán los modelos para otros hospitales.
- Los datos de AEMET tienen resolución diaria, es decir, el mismo valor se repite para las 24 horas del día. Esto impide representar bien fenómenos meteorológicos puntuales que podrían influir en las llegadas en una franja concreta.
- También falta de información más concreta por parte del hospital, ya que muchos de los datos no estaban registrados por horas, sino un valor general para toda una franja. Lo que ha generado que se tuvieran que excluir dichos valores del estudio por falta de información útil.
- Una limitación directa al modelo v4, ha sido no poder utilizar información de las horas anteriores o la ocupación observada. Esto reduce la información disponible.

9.4 Implicaciones prácticas y líneas futuras

A pesar de estas limitaciones, el trabajo demuestra que un sistema predictivo puede aportar información útil para organizar la planificación de urgencias.

El modelo no debe interpretarse como una herramienta de decisión ni como un sustituto del personal responsable del servicio, sino como un apoyo para anticipar franjas de mayor afluencia y disponer de una referencia cuantitativa antes de planificar recursos.

De cara al futuro, hay líneas de mejora que permitirían convertir este prototipo en una herramienta realmente utilizable en el día a día del servicio de urgencias, e incluso comercializable para otros hospitales.

Lo más importante sería incluir nuevos periodos temporales, permitiendo así poder validar el modelo con más años, captando cambios.

También como se ha indicado anteriormente, en este TFG, muchas variables del centro estaban registradas a nivel diario o sin detalle suficiente, lo que obligó a

descartarlas o simplificarlas. Si se pudiera contar con información detallada por horas (ingresos, altas, tiempos de espera, número de pacientes en sala, ocupación de boxes en tiempo real, traslados desde otros centros, ambulancias en camino, etc.), el modelo ganaría mucha precisión y podría adaptarse mejor a la situación real de cada momento. Como también convendría incorporar meteorología horaria en lugar de diaria.

Otra mejora clave sería integrar variables operativas en tiempo real, como la ocupación actual del hospital, el estado de las camas, el personal disponible en cada turno o las salidas previstas. Esto convertiría al sistema en una herramienta verdaderamente dinámica, capaz de ajustar sus predicciones en función de lo que está ocurriendo en cada momento.

También sería útil implementar un sistema de reentrenamiento automático, de manera que el modelo se actualice periódicamente con los datos más recientes y se adapte a cambios en el comportamiento de la demanda sin requerir intervención manual.

Con todas estas mejoras, el sistema podría pasar de ser un prototipo académico a convertirse en una herramienta verdaderamente útil para la gestión diaria de los servicios de urgencias. En conjunto, este trabajo demuestra que la aplicación de técnicas de ciencia de datos a registros hospitalarios reales puede aportar información valiosa para anticipar la demanda asistencial, apoyar la planificación operativa y avanzar hacia una gestión sanitaria más eficiente y basada en datos.

BIBLIOGRAFÍA

Aldrich, J. M., Hiscock, R., & Boyd, J. (2022). Emergency department demand forecasting and resource planning using time series models. BMJ Open, 12(4), e056017. https://bmjopen.bmj.com/content/12/4/e056017

Alenezi, A. M., Sameh, M., Aljohani, M., & Alharbi, H. (2026). Predicting Emergency Department Patient Arrivals at Hospitals Using Machine Learning Techniques. Healthcare, 14(9), 1191. https://www.mdpi.com/2227-9032/14/9/1191

Arora, S., Li, Y., Youtie, J., & Shapira, P. (2022). Digital health innovation and emergency care systems: A data-driven approach. NPJ Digital Medicine, 5(1), 1–10. https://www.nature.com/articles/s41746-022-00649-y

Burlotos, A., Stenson, B. A., Grossestreuer, A. V., & Dresser, C. (2025). Weather information improves a predictive model of emergency department arrivals. Internal And Emergency Medicine. https://doi.org/10.1007/s11739-025-04062-7

Choudhury, A., & Urena, E. (2020). Forecasting hourly emergency department arrival using time series analysis. British Journal Of Healthcare Management, 26(1), 34-43. https://doi.org/10.12968/bjhc.2019.0067

Crisafulli, P., del Río Mangada, A., Segura Sampedro, J. J., Mirasso, C. R., Toral, R., & Galla, T. (2026). Forecasting emergency department visits in the reference hospital of the Balearic Islands: The role of tourist and weather data. PLOS ONE, 21(3), e0343713. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0343713

El Hospital Universitario de Cuenca toma hoy el relevo del Hospital Virgen de la Luz para seguir atendiendo la Salud de los conquenses | Servicio de Salud de Castilla-La Mancha. (s. f.). https://sanidad.castillalamancha.es/saladeprensa/notas-de-prensa/el-hospital-universitario-de-cuenca-toma-hoy-el-relevo-del-hospital

Evaluación de modelos de series temporales para la previsión de la demanda de emergencias sanitarias – Revista Emergencias. (s. f.). https://revistaemergencias.org/numeros-anteriores/volumen-24/numero-3/evaluacion-de-modelos-de-series-temporales-para-la-prevision-de-la-demanda-de-emergencias-sanitarias/

Hernández-Suárez, D., & Montesinos-López, O. A. (2020). Time series forecasting for emergency department demand: A comparative study. Revista Emergencias, 24(3), 203–210. https://revistaemergencias.org/numeros-anteriores/volumen-24/numero-3/evaluacion-de-modelos-de-series-temporales-para-la-prevision-de-la-demanda-de-emergencias-sanitarias/

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. https://pubmed.ncbi.nlm.nih.gov/36669911/

Jiang, S., Liu, Q., & Ding, B. (2022). 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

Jones, S. S., Thomas, A., Evans, R. S., Welch, S. J., Haug, P. J., & Snow, G. L. (2008). Forecasting daily patient volumes in the emergency department. Academic Emergency Medicine, 15(2), 159–170. https://pmc.ncbi.nlm.nih.gov/articles/PMC7252622/

Khaldi, M., Ben Ahmed, M., & Bensmail, H. (2024). Advanced time series approaches for healthcare demand forecasting. Decision Support Systems, 176, 114023. https://www.sciencedirect.com/science/article/pii/S0169260724000294

Martínez-López, M., Gómez-Rodríguez, R., & Sánchez-García, J. (2012). Modelo predictor de ingreso hospitalario a la llegada al servicio de urgencias. Anales del Sistema Sanitario de Navarra, 35(2), 209–220. https://recyt.fecyt.es/index.php/ASSN/article/view/15420/10898

Ming, C., Lee, G. J. W., Teo, Y. N., Teo, Y. H., Zhou, X., Ho, E. S. Y., Toh, E. M. S., Ong, M. E. H., Tan, B. Y. Q., & Ho, A. F. W. (2025). Deep learning modelling to forecast emergency department visits using calendar, meteorological, internet search data and stock market price. Computer Methods and Programs in Biomedicine, 267, 108808. https://www.sciencedirect.com/science/article/pii/S0169260725002251

Ministerio de Sanidad. (2025). Informe Anual del Sistema Nacional de Salud 2024. Sistema de Información Sanitaria del Sistema Nacional de Salud. https://www.sanidad.gob.es/estadEstudios/estadisticas/sisInfSanSNS/tablasEstadisticas/InfAnual2024/INFORME_ANUAL_2024.pdf

Ministerio de Sanidad, Consumo y Bienestar Social - Ciudadanos - Prestaciones y centro sanitarios - Centros y servicios del SNS - Instituto de Informacion Sanitaria - Centros y servicios del SNS - Centros. (s. f.). https://www.sanidad.gob.es/va/ciudadanos/centros.do?metodo=realizarDetalle&numero=160014&tipo=hospital

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://www.sciencedirect.com/science/article/pii/S0169260724000294

Petravić, L., Rogina, K. G., Albreht, T., Kukec, A., & Žibert, J. (2025). Prognostic models for predicting patient arrivals in emergency departments: an updated systematic review and research agenda. BMC Emergency Medicine, 25(1), 106. https://doi.org/10.1186/s12873-025-01250-8

Porto, B. M., & Fogliatto, F. S. (2024). Enhanced forecasting of emergency department patient arrivals using feature engineering approach and machine learning. BMC Medical Informatics And Decision Making, 24(1), 377. https://doi.org/10.1186/s12911-024-02788-6

Predicción en tiempo real de los ingresos del servicio de urgencias mediante un modelo de inteligencia artificial y su integración en la planificación de las camas de hospitalización – Revista Emergencias. (s. f.-b). https://revistaemergencias.org/numeros-anteriores/volumen-38/numero-2/prediccion-en-tiempo-real-de-los-ingresos-del-servicio-de-urgencias-mediante-un-modelo-de-inteligencia-artificial-y-su-integracion-en-la-planificacion-de-las-camas-de-hospitalizacion

Reboredo, J. C., Barba-Queiruga, J. R., Ojea-Ferreiro, J., & Reyes-Santias, F. (2023). Forecasting emergency department arrivals using INGARCH models. Health Economics Review, 13(1), 51. https://doi.org/10.1186/s13561-023-00456-5

Revista Emergencias. (2015, agosto). Pere Tudela, Josep Maria Mòdol. https://revistaemergencias.org/wp-content/uploads/2023/08/Emergencias-2015_27_2_113-120-120.pdf

Rivas-Ruiz, R., Silva-Jivaja, K. M., Galicia-Solórzano, J. L., Ríos-Brito, K. F. R., Zamora, P., Pérez-Rodríguez, M., & Roy-García, I. A. (2021). Pertinencia de los modelos multivariantes en medicina de urgencias. Revista de Educaci N E Investigaci N En Emergencias, 3(2). https://doi.org/10.24875/reie.21000048

Rostami-Tabar, B., Browell, J., & Svetunkov, I. (2023). Probabilistic forecasting of hourly emergency department arrivals. Health Systems, 13(2), 133-149. https://doi.org/10.1080/20476965.2023.2200526

Sanidad atiende todas las demandas del comité de huelga dentro de sus competencias y garantizando la unidad del Sistema Nacional de Salud. (s. f.). https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/sanidad14/paginas/2025/111225-demandas-comite-huelga.aspx

Serrano, A. (2022a, junio 2). Modelización y predicción de series temporales en salud. Instituto de Ingeniería del Conocimiento. https://www.iic.uam.es/lasalud/modelizacion-y-prediccion-de-series-temporales-en-salud/

Serrano, A. (2022b, junio 2). Series Temporales en salud: cómo procesar los datos. Instituto de Ingeniería del Conocimiento. https://www.iic.uam.es/lasalud/series-temporales-en-salud-procesar-datos/

Shi, X., Prins, C., Van Pottelbergh, G., Mamouris, P., Vaes, B., & De Moor, B. (2021). An automated data cleaning method for Electronic Health Records by incorporating clinical knowledge. BMC Medical Informatics and Decision Making, 21, 267. https://link.springer.com/article/10.1186/s12911-021-01630-7

Sudarshan, V. K., Brabrand, M., Range, T. M., & Wiil, U. K. (2021). Performance evaluation of Emergency Department patient arrivals forecasting models by including meteorological and calendar information: A comparative study. Computers In Biology And Medicine, 135, 104541. https://doi.org/10.1016/j.compbiomed.2021.104541

Syed, R., Eden, R., Makasi, T., et al. (2023). Digital Health Data Quality Issues: Systematic Review. Journal of Medical Internet Research, 25, e42615. https://pubmed.ncbi.nlm.nih.gov/37000497/

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., Kanniainen, J., Oksala, N., Palomäki, A., & Roine, A. (2024). Forecasting emergency department occupancy with advanced machine learning models and multivariable input. International Journal of Forecasting, 40(4), 1410–1420. https://www.sciencedirect.com/science/article/pii/S0169207023001346

Vollmer, M. A. C., Glampson, B., Mellan, T., et al. (2021). A unified machine learning approach to time series forecasting applied to demand at emergency departments. BMC Emergency Medicine, 21, 9. https://link.springer.com/article/10.1186/s12873-020-00395-y

Yarubi, M. A., Kamisan, N. A. B., & Norrulashikin, S. M. (2025). Modeling and forecasting emergency department crowding using SARIMA, Holt Winter method, and Prophet models. International Journal Of Health Sciences, 9(1), 230-243. https://doi.org/10.53730/ijhs.v9n1.15567

Zhang, Y., Zhang, J., Tao, M., Shu, J., & Zhu, D. (2022). Forecasting patient arrivals at emergency department using calendar and meteorological information. Applied Intelligence, 52(10), 11232-11243. https://doi.org/10.1007/s10489-021-03085-9

Zhang, Z., Zheng, K., & Hanauer, D. A. (2023). Emergency department demand prediction using real-world hospital data. JMIR Medical Informatics, 11, e73960. https://medinform.jmir.org/2025/1/e73960

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