



FACULTAD DE HÁBITAT, INFRAESTRUCTURA Y CREATIVIDAD
INGENIERÍA EN SISTEMAS

MAESTRÍA EN SISTEMAS DE INFORMACIÓN, MENCIÓN DATA SCIENCE

TEMA

Diseño y evaluación comparativa de modelos predictivos para estimar la demanda mensual de consultas externas por especialidad médica en hospitales de segundo nivel

AUTOR: Lic. Steven Leonardo Godoy Castillo

TUTOR: Mgtr. Sebastián Felipe Tamayo Proaño

QUITO – 2026

Plan Final del Anteproyecto

1. Título del proyecto

Diseño y evaluación comparativa de modelos predictivos para estimar la demanda mensual de consulta externa en especialidades médicas de un hospital de segundo nivel de Quito.

2. Planteamiento del problema

La consulta externa especializada constituye un servicio crítico para la gestión hospitalaria, debido a que concentra una parte relevante de la atención programada y requiere coordinar recursos humanos, consultorios, agendas médicas, disponibilidad de especialistas y capacidad operativa institucional. Cuando la demanda de consultas externas no se estima con suficiente anticipación, pueden generarse desequilibrios entre la oferta disponible y el volumen de pacientes atendidos, lo que incrementa los tiempos de espera, las reprogramaciones, la sobrecarga del personal y el uso ineficiente de los recursos hospitalarios.

En los hospitales de segundo nivel, esta problemática adquiere especial relevancia porque estas instituciones atienden una demanda diversa de especialidades médicas y deben organizar su capacidad asistencial de acuerdo con patrones cambiantes de atención. Dichos patrones pueden variar según la especialidad, el mes, la estacionalidad, los días laborables, los feriados, los rezagos de atención, la disponibilidad médica y otros factores asociados al comportamiento histórico de la demanda.

En el hospital objeto de estudio, la planificación de las consultas externas por especialidad médica puede beneficiarse del análisis sistemático de los registros históricos almacenados en el sistema HIS-Oracle. Sin embargo, el simple conteo retrospectivo de atenciones no permite anticipar con suficiente precisión los cambios futuros de la demanda ni identificar de manera robusta patrones temporales relevantes para la toma de decisiones operativas. Por ello, resulta necesario evaluar si los modelos predictivos basados en series de tiempo y aprendizaje automático pueden aportar estimaciones más útiles para la planificación mensual de consultas externas.

El problema de investigación se centra en determinar qué modelos predictivos permiten estimar con mejor desempeño la demanda mensual de consultas externas por especialidad médica, a partir de registros históricos institucionales debidamente preparados, validados y analizados. Esta estimación no busca reemplazar la planificación administrativa o médica, sino proporcionar evidencia cuantitativa que apoye la asignación de recursos, la programación de agendas y la anticipación de posibles variaciones en la demanda.

3. Pregunta de investigación

¿Qué modelo predictivo presenta mejor desempeño para estimar la demanda mensual de consultas externas por especialidad médica en un hospital de segundo nivel?

4. Objetivos generales y específicos

4.1. Objetivo general

Diseñar, implementar y evaluar comparativamente modelos predictivos para estimar la demanda mensual de consultas externas por especialidad médica en un hospital de segundo nivel.

4.2. Objetivos específicos

- Caracterizar el comportamiento histórico de la demanda de consulta externa por especialidad médica.
- Seleccionar las especialidades médicas que serán incluidas en el modelado, de acuerdo con criterios de volumen, continuidad histórica y disponibilidad de datos.
- Preparar el conjunto de datos mediante técnicas de preprocesamiento de datos.
- Implementar modelos predictivos para estimar la demanda de consulta externa por especialidad médica.
- Evaluar y comparar el desempeño de los modelos mediante métricas de error adecuadas para problemas de pronóstico.
- Presentar los resultados mediante tablas, gráficos y un tablero de visualización que facilite la interpretación de la demanda histórica, los pronósticos y el desempeño de los modelos.

5. Justificación del proyecto

La predicción de la demanda de consulta externa en especialidades médicas constituye una necesidad relevante para la gestión hospitalaria, debido a que las variaciones en la atención dificultan la planificación eficiente de recursos humanos, agendas médicas, turnos, consultorios y capacidad operativa. En Ecuador, el acceso oportuno a consultas especializadas continúa siendo un desafío, ya que un paciente puede esperar hasta 102 días para recibir atención médica especializada y, en algunos casos, los tiempos de espera han alcanzado entre tres y seis meses (Primicias, 2025). Asimismo, el IESS informó que, pese al incremento en la oferta de citas médicas, el tiempo promedio de espera en especialidades se mantiene alrededor de los 85 días. Estas condiciones evidencian la necesidad de fortalecer los mecanismos de planificación y gestión de la demanda de servicios especializados.

La falta de previsión de la demanda puede generar desequilibrios entre la oferta disponible y el volumen de pacientes atendidos, ocasionando sobrecarga del personal de salud, incremento en los tiempos de espera, cancelación o reprogramación de citas y, en determinados periodos, subutilización de recursos humanos y materiales. De hecho, estudios realizados en Ecuador señalan que la saturación de la consulta externa, la limitada disponibilidad de citas y las deficiencias en los procesos administrativos afectan negativamente la satisfacción de los pacientes y la eficiencia institucional (Gómez, 2019). Además, la organización del personal y la gestión de las referencias hacia especialidades influyen significativamente en la percepción de los tiempos de espera y en la calidad del servicio.

En este contexto, las herramientas predictivas y las técnicas de analítica avanzada representan una alternativa pertinente para apoyar la planificación de los servicios de salud. Según Avaunza (2024), los modelos predictivos contribuyen a una distribución más equilibrada de los recursos, disminuyen los tiempos de atención y favorecen una respuesta más oportuna a las necesidades de los pacientes. Asimismo, las técnicas de aprendizaje automático y series de tiempo pueden ser útiles para identificar patrones de comportamiento en los datos históricos, analizar variaciones temporales y generar estimaciones que apoyen la toma de decisiones operativas en instituciones de salud.

Desde la perspectiva social y organizacional, una mejor estimación de la demanda mensual de consulta externa puede contribuir a mejorar la planificación de agendas, optimizar la utilización de recursos médicos y favorecer una atención más organizada y eficiente. No obstante, aunque los modelos predictivos no resuelven por sí solos los problemas estructurales de acceso a la atención especializada, sí pueden aportar evidencia cuantitativa para anticipar periodos de mayor demanda y apoyar decisiones relacionadas con la programación de servicios.

Por otra parte, el estudio contribuirá a ampliar la evidencia empírica sobre el uso de modelos predictivos en el contexto hospitalario ecuatoriano, donde aún son limitadas las investigaciones orientadas a la estimación de la demanda de consulta externa por especialidad médica mediante técnicas de aprendizaje automático. Por ello, el presente trabajo busca generar un procedimiento basado en datos históricos institucionales que pueda servir como referencia para futuras investigaciones y para la implementación de estrategias de gestión hospitalaria apoyadas en evidencia.

En consecuencia, la investigación propone diseñar, implementar y evaluar comparativamente modelos predictivos basados en datos históricos institucionales, técnicas de aprendizaje automático y series de tiempo, con el propósito de estimar la demanda mensual de consulta externa por especialidad médica. Los resultados esperados buscan apoyar la planificación y la gestión eficiente de los recursos hospitalarios, fortaleciendo la toma de decisiones basada en datos en un hospital de segundo nivel.

6. Marco teórico

6.1. Consulta Externa y Demanda de Especialidades Médicas

La consulta externa constituye uno de los principales servicios asistenciales de los hospitales, ya que permite la atención programada de pacientes sin requerir hospitalización. En los hospitales de mediana y alta complejidad, este servicio concentra una gran cantidad de atenciones y presenta una elevada variabilidad en función de factores temporales, organizacionales y poblacionales.

La demanda de especialidades médicas se refiere al volumen de consultas requeridas por los pacientes en un período determinado. Su comportamiento puede verse influenciado por diversos factores, entre ellos la especialidad médica, la época del año, los días festivos y los patrones históricos de atención. En Ecuador, se han reportado tiempos prolongados de espera para

consultas especializadas, alcanzando en algunos casos hasta 102 días, situación que evidencia desafíos en la planificación y gestión de los servicios de salud (Primicias, 2025).

La adecuada gestión de la demanda requiere anticipar el volumen futuro de consultas para optimizar la asignación de médicos especialistas, la disponibilidad de consultorios y la planificación de agendas médicas. En este contexto, los modelos predictivos constituyen una herramienta que permite transformar datos históricos en información útil para la toma de decisiones.

6.2. Analítica Predictiva en Salud

La analítica predictiva es una disciplina que utiliza técnicas estadísticas, matemáticas y computacionales para estimar eventos futuros a partir de datos históricos. Su objetivo principal es identificar patrones y tendencias que permitan anticipar comportamientos y apoyar la toma de decisiones.

En el sector salud, la analítica predictiva se emplea para estimar la demanda de servicios, predecir ingresos hospitalarios, gestionar recursos asistenciales y optimizar procesos operativos. Según Avaunza (2024), la aplicación de modelos predictivos contribuye a mejorar la planificación institucional mediante una distribución más eficiente de los recursos disponibles.

La incorporación de técnicas de machine learning ha ampliado las capacidades de la analítica predictiva, permitiendo identificar relaciones complejas entre variables y generar modelos con mayor capacidad de adaptación a grandes volúmenes de datos.

6.3. Machine Learning en la Predicción de la Demanda

El machine learning o aprendizaje automático es una rama de la inteligencia artificial que permite a los sistemas aprender patrones a partir de datos históricos y realizar predicciones sin ser programados explícitamente para cada situación.

Dentro de esta disciplina, el aprendizaje supervisado utiliza datos históricos compuestos por variables de entrada y una variable objetivo conocida, con el propósito de construir modelos capaces de estimar valores futuros. En esta investigación, la variable objetivo corresponde al número de consultas externas por especialidad médica.

Entre los algoritmos más utilizados para problemas de regresión se encuentran XGBoost, Random Forest y Gradient Boosting. Estos modelos pertenecen a la categoría de métodos de ensamble, los cuales combinan múltiples árboles de decisión para mejorar la precisión predictiva y reducir errores de estimación.

XGBoost es un algoritmo de boosting optimizado que construye árboles de decisión de forma secuencial, corrigiendo progresivamente los errores del modelo anterior. Random Forest genera múltiples árboles de decisión independientes utilizando muestras aleatorias de los datos y combina sus resultados para obtener una predicción más robusta. Por su parte, Gradient Boosting construye modelos secuenciales orientados a minimizar una función de pérdida mediante técnicas de optimización.

6.4. Series de Tiempo y Pronóstico

Una serie de tiempo es una secuencia ordenada de observaciones registradas a intervalos de tiempo. En el ámbito hospitalario, las series de tiempo permiten analizar la evolución de la demanda de consultas y detectar patrones de comportamiento a lo largo del tiempo.

Las series de tiempo suelen estar compuestas por elementos como tendencia, estacionalidad, ciclos y variaciones aleatorias. La identificación de estos componentes facilita la construcción de modelos de pronóstico capaces de estimar valores futuros de una variable.

Entre los métodos tradicionales de pronóstico destacan los modelos ARIMA (Autoregressive Integrated Moving Average) y SARIMA (Seasonal Autoregressive Integrated Moving Average), ampliamente utilizados para modelar fenómenos con dependencia temporal y patrones estacionales.

6.5. Metodología CRISP-DM

CRISP-DM (Cross Industry Standard Process for Data Mining) es una metodología ampliamente utilizada en proyectos de minería de datos y ciencia de datos. Su objetivo es proporcionar un proceso estructurado y reproducible para el desarrollo de soluciones analíticas.

La metodología se compone de seis fases: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. Estas etapas permiten transformar datos en conocimiento útil mediante un proceso sistemático y orientado a objetivos.

En esta investigación, CRISP-DM proporciona el marco metodológico para la construcción y evaluación del modelo predictivo de demanda hospitalaria.

6.6. Modelos candidatos y justificación metodológica

Tabla 1. Modelos candidatos para la investigación

Modelo	Rol en el estudio	Ventajas	Riesgos o limitaciones
Modelo ingenuo / promedio móvil	Línea base de comparación	Permite verificar si los modelos avanzados realmente mejoran un método simple.	Puede ser insuficiente ante cambios abruptos o patrones complejos.
ARIMA / SARIMA	Modelo clásico de series de tiempo	Adecuado para dependencia temporal y posible estacionalidad.	Requiere historia suficiente; puede rendir mal con múltiples variables externas.
Regresión regularizada	Modelo interpretable de referencia	Facilita analizar el peso de variables temporales y rezagos.	Puede subrepresentar relaciones no lineales.
Random Forest Regressor	Modelo de ensamble no lineal	Robusto ante interacciones entre variables y útil para importancia de características.	Puede extrapolar mal tendencias futuras y requiere validación temporal estricta.
Gradient Boosting Regressor	Modelo de ensamble secuencial	Suele ofrecer buen desempeño en datos tabulares.	Puede sobreajustarse si no se controlan hiperparámetros.

XGBoost Regressor	Modelo avanzado de boosting	Alta capacidad predictiva y manejo eficiente de variables tabulares.	Mayor complejidad; requiere ajuste cuidadoso e interpretación complementaria.
-------------------	-----------------------------	--	---

6.7. Métricas de evaluación

La evaluación de modelos de pronóstico debe cuantificar la diferencia entre valores observados y valores estimados. Las métricas principales serán MAE, RMSE y MAPE o sMAPE, según la presencia de valores bajos o cercanos a cero. MAE facilita interpretar el error promedio en número de consultas; RMSE penaliza errores grandes; MAPE permite una lectura porcentual, aunque puede ser inestable cuando existen conteos pequeños.

El coeficiente R^2 puede reportarse como métrica secundaria, pero no debe ser el único criterio de selección, especialmente en series temporales. Para aumentar el rigor, el desempeño de los modelos debe compararse contra una línea base simple y evaluarse por especialidad, no solo mediante un promedio global.

7. Metodología de investigación

7.1. Enfoque, alcance y diseño metodológico

La investigación tendrá un enfoque cuantitativo, aplicado y predictivo-comparativo. Es cuantitativa porque utilizará registros estructurados de consulta externa y métricas numéricas de evaluación. Es aplicada porque busca apoyar un problema real de planificación hospitalaria. Es predictivo-comparativa porque implementará y comparará distintos modelos para estimar demanda futura.

El estudio no realizará inferencias causales; las asociaciones identificadas serán interpretadas como patrones útiles para predicción y planificación, no como relaciones causa-efecto.

7.2. Fuente de datos y unidad de análisis

La fuente principal será secundaria interna: registros históricos de consulta externa extraídos del sistema HIS-Oracle del hospital objeto de estudio, previa autorización institucional y aplicación de medidas de confidencialidad, es decir, no se utilizarán datos personales identificables de pacientes.

La unidad primaria de análisis será el registro de atenciones por especialidad y fecha. Para el modelado, los datos podrán agregarse a nivel diario o mensual según su disponibilidad y calidad.

Para que el pronóstico del segundo semestre de 2026 tenga validez temporal, se va a trabajar con datos históricos desde del 2023 a mayo de 2026, con la finalidad de realizar una predicción más robusta.

7.3. Consideraciones éticas y de confidencialidad

El estudio utilizará datos institucionales secundarios y no requerirá información personal identificable de pacientes, además, se trabajará con datos agregados por especialidad y período, o con registros anonimizados cuando sea necesario.

Los resultados se reportarán de forma agregada y no se utilizarán para evaluar individualmente a pacientes, médicos o personal administrativo. El modelo será presentado como herramienta de apoyo a la planificación, no como mecanismo automático de decisión clínica o administrativa.

8. Cronograma de actividades

Fase	Semana	Fecha Inicio	Fecha Fin	Etapas	Responsable
Fase 1	Semana 1	8/6/26	14/6/26	Planificación de la Investigación	Maestrante / director
	Semana 2	15/6/2026	21/6/2026		
Fase 2	Semana 3	22/6/26	5/7/26	Comprensión negocio y datos	Maestrante / Hospital
Fase 3	Semana 4	06/7/2026	9/7/2026	Preparación de los datos	Maestrante / director
	Semana 5	10/7/2026	14/7/2026		
Fase 4	Semana 6	15/7/2026	20/7/2026	Modelado	Maestrante / director
	Semana 7	21/7/2026	26/7/2026		
Fase 5	Semana 8	27/7/2026	2/8/2026	Evaluación	Maestrante / director
Fase 6	Semana 9	03/8/2026	10/8/2026	Redacción y Presentación	Maestrante / director
	Semana 10	11/8/2026	14/8/2026		

9. Recursos

Tabla 6. Recursos requeridos

Categoría	Recurso	Uso previsto
Humanos	Maestrante	Extracción, preparación, modelado, análisis y redacción de la investigación.
Humanos	Director de tesis	Orientación metodológica, revisión académica y validación del avance.
Humanos	Jefatura de TI del hospital	Facilitación del acceso a HIS-Oracle y explicación de la estructura de datos.
Técnicos	Python 3.10+, pandas, NumPy	Procesamiento, limpieza y estructuración de datos.
Técnicos	scikit-learn, XGBoost, statsmodels	Entrenamiento y evaluación de modelos predictivos.
Técnicos	Matplotlib, Seaborn, Power BI	Visualización de resultados y construcción de dashboard.
Técnicos	Google Colab o entorno local	Experimentación y ejecución de modelos.
Técnicos	Zotero	Gestión bibliográfica y control de referencias.

Financieros	Transporte, impresión, contingencias	Visitas institucionales, entrega física y gastos menores.
-------------	--------------------------------------	---

10. Alcances y limitaciones

- El estudio se limita a un hospital de segundo nivel de Quito; sus resultados no podrán generalizarse automáticamente a otros hospitales sin validación externa.
- El modelo estimará demanda observada en los registros institucionales, no demanda insatisfecha ni necesidad epidemiológica real de la población.
- La calidad de los resultados dependerá de la completitud, consistencia y extensión temporal de los datos extraídos del HIS-Oracle.
- El estudio no realizará inferencia causal sobre los factores que determinan la demanda; su objetivo es predictivo y aplicado.
- Las predicciones deberán interpretarse como apoyo a la planificación, no como sustituto de decisiones administrativas, clínicas o presupuestarias.