

**OPTIMIZACIÓN DE TIEMPOS TRIAGE:
Diseño de tableros BI y Predicción con machine learning para mejorar la
atención de los pacientes**

Diana Carolina Ussa Ruiz y Víctor Julio Gomez Carvajal

Especialización en BigData, Fundación Universitaria Compensar
Desarrollar una propuesta de Consultoría

Ingeniero(a) Katherine Serrano León

27 de Noviembre de 2023



Esta obra está bajo una [Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0/).

OPTIMIZACIÓN DE TIEMPOS TRIAGE:

Diseño de tableros BI y Predicción con machine learning para mejorar la atención de los pacientes

Diana Carolina Ussa Ruiz y Víctor Julio Gomez Carvajal

Trabajo de grado presentado como requisito parcial para optar al título de:
Especialización en Big Data

Director (a):

Ingeniero(a) Katherine Serrano León

Fundación Universitaria Unipanamericana – Compensar
Facultad de Ingeniería, Ingeniería de (Especialización en Big Data)
Bogotá DC, Colombia
2023

Resumen

La situación actual que se presenta en la Subred Sur ubicada al Sur de Bogotá, se evidencia cómo las áreas de urgencias; según las quejas de PQRS donde se expresa “Tiempos de atención prolongados, generando frustración por las pérdidas de tiempos en los puntos de atención”.

Esta consultoría se enfocará en la Subred Sur ubicada al Sur de Bogotá, donde hay una elevada demora en la atención de los pacientes en las áreas de urgencias. Esta investigación es un producto de trabajo de grado, que se relaciona con línea de investigación; a través de un enfoque mixto que combina métodos cuantitativos.

En la Subred Sur, La primera acción de cara al usuario y su familia es el acompañamiento en las Respuestas a las PQRS por trato deshumanizado con relación a procesos de atención: el núcleo de humanización adopta un enfoque personalizado con cada individuo propenso a involucrarse en comportamientos que no contribuyen al proceso de humanización de los servicios. Este impulsa un compromiso centrado en mejorar y modificar actitudes, basado en el fortalecimiento de las conexiones entre el usuario y el colaborador.

Palabras claves: Triage, tiempos de atención, sector Salud, pacientes, áreas de urgencias.

Abstract

In the current situation that occurs in the Southern Subnet located south of Bogotá, it is evident how the emergency areas; according to PQRS complaints where it is expressed "Prolonged attention times, generating frustration due to loss of time at the attention points."

This consultancy will focus on the Southern Subnet located south of Bogotá, where there is a high delay in patient care in emergency areas. This research is a degree work product, which is related to the line of research; through a mixed approach that combines quantitative methods.

In the Southern Subnet, the first action for the user and their family is support in the Responses to the PQRS due to dehumanized treatment in relation to care processes: the humanization core adopts a personalized approach with each individual prone to engaging in behaviors that do not contribute to the process of humanization of services. This promotes a commitment focused on improving and modifying attitudes, based on strengthening connections between the user and the collaborator.

Keywords: Triage, attention times, Health sector, patients, emergency areas.

Contenido

	PÁG.
RESUMEN	III
INTRODUCCIÓN	1
ANTECEDENTES Y JUSTIFICACIÓN.....	1
FORMULACIÓN DEL PROBLEMA	2
<i>Pregunta de Investigación</i>	2
OBJETIVOS	2
<i>Objetivo General</i>	2
<i>Objetivos Específicos</i>	3
ALCANCES Y LIMITACIONES	3
1. MARCO TEÓRICO	4
2. DISEÑO METODOLÓGICO	5
2.1. REQUERIMIENTOS	6
2.2. PROCESAMIENTO	6
2.2.1. TRATAMIENTO DE LOS DATOS.....	7
2.2.2. CONSTRUCCIÓN DE PROGRAMAS PYTHON PARA VALIDACIÓN DATOS	7
2.2.3. TÉCNICAS DE LIMPIEZA DE DATOS (VALORES ATÍPICOS).	8
2.2.4. TÉCNICAS DE LIMPIEZA DE DATOS (IMPUTACIÓN VALORES NULOS).	9
2.3. DESARROLLO DE LOS MODELOS.....	9
2.3.1. TIPOS DE MODELOS DE PRONÓSTICOS DE SERIES DE TIEMPO	10
2.3.2. DEFINICIÓN DE MÉTRICAS ADECUADAS DE DESEMPEÑO.....	10
2.3.3. MODELOS USADOS	10
2.4. DISEÑO E IMPLEMENTACIÓN BI	11
2.5. IMPLEMENTACIÓN DEL SISTEMA DE MONITOREO.....	12
3. RESULTADOS – DISCUSIONES	13
3.1. IMPLEMENTACIÓN DEL MODELO SELECCIONADO.....	13
4. CONCLUSIONES Y RECOMENDACIONES.....	15
4.1. CONCLUSIONES	15
4.2. RECOMENDACIONES	15
5. BIBLIOGRAFÍA.....	16

Lista de tablas

Pág.

TABLA 1: REQUERIMIENTOS DE NEGOCIO	6
TABLA 2: IDENTIFICACIÓN VALORES NULOS	9

Lista de figuras

Pág.

FIGURA 1 – GRAFICAS DE DENSIDAD VARIABLES (TIEMPO TOTAL, HORA, DÍA, MES)	7
FIGURA 2 DISEÑO INICIAL DE TABLEROS EN POWER BI	12
FIGURA 3 – COMPARACIÓN GRÁFICA DE LOS DIFERENTES MODELOS DE SERIES DE TIEMPO	13
FIGURA 4 – COMPARACIÓN EVALUACIÓN MÉTRICAS DE LOS DIFERENTES MODELOS DE SERIES DE TIEMPO ...	13
FIGURA 5 – GRÁFICAS DE APLICACIÓN WEB – TIEMPOS DE ATENCIÓN DE URGENCIAS – PREDICCIÓN	14
FIGURA 6 – GRÁFICAS DE POWER BI – HISTÓRICO DE ATENCIONES	14

Introducción

El sector de la salud en Colombia juega un papel crucial en la prestación de servicios médicos y el bienestar de la población. El sistema de salud colombiano está conformado por diferentes actores, políticas y entidades, y su principal objetivo es garantizar el acceso equitativo a la atención médica y la promoción de la salud en todo el país

La atención médica de urgencia desempeña un factor importante en la preservación de la vida y la salud de los pacientes en situaciones críticas. El área de urgencias de los hospitales y centros de salud es un espacio dinámico y esencial para brindar atención inmediata a personas que requieren atención médica urgente. Sin embargo, la sobrecarga de pacientes y la necesidad de una priorización efectiva en la atención pueden generar tiempos de espera prolongados y, en consecuencia, impactar negativamente en la calidad y eficiencia del servicio. (Rodríguez-Páez, F. G., Jiménez-Barbosa, W. G., & Palencia-Sánchez, F. (2018)

En el contexto de la atención médica, la optimización de los tiempos de espera en el área de urgencias es decisivo para garantizar una atención eficiente y de calidad a los pacientes. En este sentido, el presente proyecto tiene como objetivo principal diseñar tableros de Business Intelligence que representen datos históricos de tiempos de espera, y desarrollar un modelo predictivo de regresión mediante técnicas de Machine Learning. La integración de estas herramientas potenciará la toma de decisiones en el servicio médico, mejorando su eficiencia y adaptabilidad. (Raschka & Mirjalili, 2017)

Antecedentes y Justificación

La Subred Sur, cuenta con una cifra total de 1,241,784 residentes hasta el año 2022, según las estimaciones demográficas por zonas para Bogotá entre 2005 y 2035. Aproximadamente el 99% de esta población está distribuida en áreas urbanas. Ciudad Bolívar lidera con la proporción más significativa, alcanzando el 52.8%, seguida por Usme con un 32.3%. Tunjuelito representa el 14.6%, mientras que Sumapaz, que es predominantemente rural, constituye el 0.3% de la población total.

La Resolución No. 5596 de 2015 del Ministerio de Salud y Protección Social tiene como objetivo definir los parámetros técnicos para el Sistema de Selección y Clasificación de pacientes en los

servicios de urgencias, denominado "Triage", que será aplicado en las Instituciones Prestadoras de Servicios de Salud, los tiempos de espera para la atención se rigen por la clasificación de Triage, de acuerdo con lo establecido en la mencionada Resolución.

Lo más importante de esta consultoría es poder aportar herramientas y conclusiones a la Subred Sur E.S.E para que sean aplicadas con el objetivo que puedan identificar el impacto como la automatización de procesos por medio de analítica de datos puede optimizar estos tiempos de atención en el área de urgencias.

Formulación del Problema

En el ámbito de la salud, se ha observado un incremento en los tiempos de atención en las áreas de Triage de urgencias, generando una creciente insatisfacción entre los pacientes y un aumento en las quejas relacionadas con la demora en la atención inicial. Este fenómeno plantea la necesidad de comprender a fondo las causas subyacentes y las consecuencias vinculadas a los lapsos de espera en las áreas de urgencias, con el fin de perfeccionar la calidad y eficacia en la prestación de servicios de urgencias. (Restrepo-Zea, J. H., Jaén-Posada, J. S., Espinal Piedrahita, J. J., & Zapata Flórez, P. A. (2018).

Pregunta de Investigación

¿Cómo podemos identificar y utilizar de manera efectiva las variables clave que influyen en los tiempos de atención de pacientes en las áreas de Triage de urgencias de la Subred Sur ESE, para diseñar unos Dashboard que mejore la toma de decisiones mediante técnicas de Machine Learning y optimice la gestión de recursos en tiempo real?.

Objetivos

Objetivo General

Diseñar tableros de Business Intelligence para representar datos históricos de tiempos de espera en atención médica. Utilizando técnicas de Machine Learning, desarrollaremos un modelo predictivo de regresión para predecir tiempos de espera en tiempo real. Esta integración potenciará la toma de decisiones, mejorando la eficiencia y adaptabilidad del servicio médico.

Objetivos Específicos

- Identificar las variables relevantes que impactan los tiempos de atención de los pacientes en las áreas de urgencias de la Subred Sur ESE. Esto implica la obtención de datos históricos de pacientes, tiempos de espera, información médica y otros datos pertinentes para la adecuada preparación de los datos para un correcto análisis y modelado.
- Desarrollar y validar un modelo de regresión con técnicas de machine learning que pueda predecir los tiempos de espera con base en las variables identificadas en el Objetivo Específico 1.
- Diseñar un sistema de monitoreo en tiempo real para ajustar las operaciones de urgencias en función de la demanda y los patrones detectados en el área de urgencias de la Subred Sur ESE. Esto permitirá una gestión más adaptativa y eficiente de los recursos.

Alcances y Limitaciones

El proyecto se centra en identificar variables clave que inciden en los tiempos de espera en las áreas de Triage de urgencias de la Subred Sur ESE. La recopilación, preparación y análisis de datos históricos, incluyendo información sobre pacientes, tiempos de espera y factores clínicos, demográficos y operativos relevantes, constituirán parte integral de este trabajo.

Se desarrollará un modelo de regresión utilizando técnicas de machine learning para predecir los tiempos de espera en función de estas variables identificadas. Además, se diseñará e implementará un sistema de monitoreo en tiempo real para ajustar las operaciones de urgencias según la demanda y patrones detectados.

El proyecto no abarca la intervención directa en la atención médica ni la implementación de cambios en la infraestructura hospitalaria.

No contar con la calidad o disponibilidad limitada de datos históricos puede afectar la precisión del modelo de predicción.

Factores no contemplados que podrían influir en los tiempos de espera, limitando la exactitud de las predicciones del modelo.

1.Marco Teórico

En el marco de esta consultoría, se adopta un enfoque basado en Big Data, específicamente haciendo uso de la metodología de Extracción, Transformación y Carga (ETL). Se procedió a recopilar datos históricos de pacientes atendidos a lo largo de un periodo temporal, abordando tanto aspectos cualitativos como cuantitativos. Este enfoque se seleccionó con el propósito de satisfacer los objetivos planteados y abordar de manera integral el problema objeto de estudio.

Para garantizar una observación precisa y orientada, se llevó a cabo un análisis del procedimiento establecido por la Subred en la atención de urgencias. Este análisis permitió la identificación de variables cruciales para la medición de los tiempos asociados a cada operación realizada durante la atención de urgencias. Posteriormente, esta información fue cargada en una base de datos, y se emplearon diversas herramientas, como Power BI, programas en Python, DAX y Power Query, para iniciar el procesamiento de datos.

Todos los archivos generados en las distintas fases del proyecto fueron almacenados en un repositorio de datos accesible en la siguiente ruta: <https://github.com/Vitotoju/Compensar>.

Este enfoque facilita la visualización ágil de los datos, ofreciendo un contexto global de la información recopilada en el ámbito de las urgencias. Además, este repositorio sirve como fuente de referencia para determinar las variables más relevantes que inciden en los tiempos de espera, las cuales se utilizarán en la construcción del modelo de machine learning. Este proceso de almacenamiento y accesibilidad de datos contribuye a la transparencia y reproducibilidad de los resultados obtenidos en el estudio.

2. Diseño Metodológico

En el contexto de la investigación, se adopta un enfoque de Big Data mediante la aplicación de la metodología de Extracción, Transformación y Carga (ETL). La recopilación de datos históricos de pacientes atendidos durante un periodo determinado constituye un paso crucial para abordar los objetivos propuestos y entender a fondo la problemática en estudio.

La primera fase del proceso metodológico consistió en analizar en detalle el procedimiento establecido por la Subred para la atención de urgencias. Este análisis se centró en la identificación de variables críticas para medir los tiempos asociados a cada operación realizada durante la atención. A partir de este análisis, se procedió a la selección de variables relevantes que serán fundamentales en las fases subsiguientes del estudio.

Una vez seleccionadas las variables, se llevó a cabo un proceso exhaustivo de limpieza de datos. Esto implicó la identificación y corrección de posibles errores, la gestión de valores faltantes y la normalización de datos para garantizar la coherencia y calidad de la información a utilizar.

En cuanto a la construcción de modelos de predicción, se implementaron técnicas avanzadas de machine learning. Estos modelos se diseñaron con el objetivo de prever los posibles tiempos de atención de los pacientes en las áreas de los servicios de urgencias, utilizando las variables previamente seleccionadas. Se exploraron algoritmos de aprendizaje supervisado y no supervisado para maximizar la precisión y la relevancia de las predicciones.

Simultáneamente, se diseñaron Dashboard interactivos mediante herramientas como Power BI, Python y Streamlit. Estos tableros se estructuraron de manera intuitiva para visualizar de manera efectiva los resultados de los modelos y proporcionar una comprensión más accesible de los datos.

Este enfoque metodológico, que abarca desde la selección de variables hasta la implementación de modelos de predicción y la creación de Dashboard, se diseñó con el propósito de ofrecer una solución integral y efectiva para abordar la problemática en el área de urgencias.

2.1. Requerimientos

El proyecto pretende entregar una herramienta que, al aplicarla sobre las atenciones de los pacientes de urgencias, permita identificar los tiempos de atención históricos y posibles problemas en sus características actuales y predicciones sobre si terminarían con dificultades en su cierre, conllevando a generar mejoras en la atención de los tiempos de pacientes. Estas herramientas se generarán a partir del entendimiento de los datos históricos, donde se realizarán comparaciones entre lo realmente ejecutado y la capacidad instalada en esas áreas de urgencias; con el propósito de discernir las particularidades presentes en los datos, se busca detectar las razones subyacentes del incremento en los lapsos de atención de los pacientes en el sector de urgencias.

A continuación, se definen los requerimientos del negocio:

Tabla 1: Requerimientos de negocio

Aspecto	Nombre	Requerimiento
R1	Herramienta descriptiva	Desarrollar una herramienta que proporcione a los usuarios información de los tiempos de espera históricos en Triage por meses, días, semanas y turnos el cual se podría utilizar para mejorar la atención de los pacientes en minutos.
R2	Herramienta de predicción de tiempos de atención	Desarrollar una Herramienta que permita predecir los tiempos de espera mínimo a 1 mes, el cual se podría utilizar como alerta para tomar decisiones para mejorar la eficiencia de las filas y tiempos de espera.

2.2. Procesamiento

Para realizar todo el procesamiento de la información se utilizan varias herramientas como son: “Visual Studio Code”, Python, Power BI, Streamlit, GitHub; para así realizar todos los temas de cargue, validación, transformación, creación de pantallas, visualización de datos, creación de informes y gráficas, creación de tableros y creación de un aplicativo online para ser visualizado en la intranet.

2.2.1. Tratamiento de los datos.

A partir de los requerimientos planteados y validando las fuentes de datos (Servidor de Dinámica Gerencial – Software institucional ERP), se decide tomar como fuente de información la base de datos entregada por Subred Sur ESE.

A partir de esta base de datos se crean varias sentencias SQL en el motor de Base de Datos y se realizan varios procesos ETL que después de varios procesos se define las variables definitivas y se crea esta sentencia SQL final llamada : “Tiempos Total de tiempos de Triage”, el cual se decidió debido a la necesidad del proyecto y así usar solamente las requeridas para el caso que estamos llevando (10 variables) :

"FECHA_LLEGADA", "TIEMPO_TOTAL_FINAL", "CENTRO_ATENCION", "CLASIFICACION_TRIAGE", "PACIENTE_EDAD", "PACIENTE_#_DOCUMENTO", "EDAD_RANGO", "NOMBRE_ENTIDAD", "SEXO", "Month"

2.2.2. Construcción de programas Python para validación datos

Inicialmente se decide crear varios programas en Python para poder validar los datos y su relación con las demás variables como son: (Ciencia de datos, s.f)

- Programa cargue y validación de la información (Cargue Información Python.ipynb):
- Proceso validacion de datos y variables.ipynb
- Ejercicios Practicos_exploracion de datos.ipynb
- Modelo Base time_series_forecasting_arima_lstm_Random_forest_prophet.ipynb
- Modelo de teoría de colas - 3 oct.ipynb
- Modelo de bosque aleatorio(18oct).ipynb
- Varios_Modelos_de_series_de_tiempo_8_novV3.ipynb
- Varios_modelos_14nov version9 - Final.ipynb

El código Python se consolida en un archivo PDF de forma cronológica : [Compensar/documentos/Codigo_python_-_Tiempos_de_espera.pdf_at_main_-_Vitotoju/Compensar\(github.com\)](#)

Se desarrolla como versión final el código Python : [TiemposUrgFinal.py](#) , que es la aplicación cargada a publicar en Streamlit.

En el repositorio de GitHub se detalla cada uno de los desarrollos de los programas Python:
<https://github.com/Vitotoju/Compensar/tree/main/code>

2.2.3. Técnicas de limpieza de datos (Valores atípicos).

Inicialmente, sin procesamiento de los datos, la base de datos contenía **1.375.710** filas y 19 columnas o variables como se mencionó en el capítulo 2, después del tratamiento o procesamiento de datos se eliminaron columnas duplicadas que podrían generar datos erróneos al correr los diferentes modelos o que no aportan información a la pregunta de investigación.

Producto del procesamiento de los datos realizado por medio de los programas de python y como resultado se evidencia franjas de datos o valores atípicos que afectan el resultado de los modelos a predecir ver (Figura 1 – Gráficas de densidad variables (Tiempo total, Hora, Día, Mes) a los archivos planos generados de las consultas de la base de datos, teniendo en cuenta el gran volumen del dataset y con la intención de enfocar el esfuerzo del trabajo para que los resultados puedan ser interpretables, se filtran y seleccionan los tiempos en minutos mayores a 0 y menores a 420 eliminando del análisis otros tiempos en minutos como los minutos menores a 0(son registros que no contienen fechas,) y mayores a 420 minutos (ósea a 7 horas de espera) que creemos que se ajusta a los modelos que queremos predecir, estos datos se deben que no fueron ingresadas en el sistema en mismo momento de atención del proceso.

Este proceso reduce los registros a 82.416 de las atenciones realizadas y seleccionando solamente datos desde el año 2023 de la SUBRED SUR ESE. Por mejor ajuste de los datos y validación de los modelos.

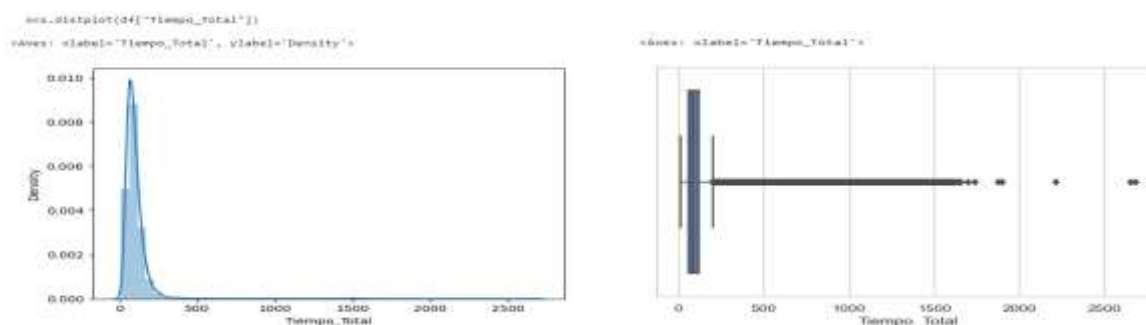


Figura 1 – Gráficas de densidad variables (Tiempo total, Hora, Día, Mes)

2.2.4. Técnicas de limpieza de datos (Imputación valores nulos).

Producto del procesamiento de los datos realizado por medio de los programas de python y como resultado de las técnicas realizadas se evidencian los siguientes registros con imputaciones para valores nulos (ver Tabla 2).

Algunas de las técnicas realizadas para detectar campos null en la data creado en python es:

En la tabla siguiente se presenta el proceso de identificación de valores nulos en el dataframe y además se presenta el tratamiento realizado a los valores nulos encontrados.

Tabla 2: identificación valores nulos

Columna	Non-Null	Tratamiento
TIEMPO_TOTAL	5200	Se reemplaza por la media de la misma columna
FECHA_LLEGADA	7830	Se reemplaza con los datos FECHA_TRIAGE
FECHA_TRIAGE	57320	Se reemplaza con los datos FECHA_INGRESO
FECHA_INGRESO	777	Se reemplaza con los datos FECHA_ATENCION
FECHA_ATENCION	15	Se reemplaza con los datos FECHA_INGRESO
MEDICO	15	Se reemplaza por la media de la misma columna
CENTRO_ATENCION	35	Se reemplaza por el centro de atención del MEDICO tratante
NOMBRE_ENTIDAD	7830	Se reemplaza por el valor más común de la sede encontrada.

2.3. Desarrollo de los modelos

Machine Learning (ML) representa un ámbito de investigación centrado en la creación de algoritmos y modelos que faculden a las computadoras para adquirir conocimientos a partir de los datos, mejorando su eficacia y generando predicciones que fundamenten la toma de decisiones basada en dicho aprendizaje. Dentro del campo de machine learning hay diferentes áreas de especialidad entre las que se encuentran el aprendizaje supervisado y el aprendizaje no supervisado (Kuhn & Johnson, 2013, 41).

En este capítulo se pretende dar respuesta al requerimiento R2 “Herramienta de predicción de tiempos de atención” (Desarrollar una herramienta que proporcione a los usuarios información de los tiempos de espera históricos en Triage por meses, días, semanas y turnos el cual se podría utilizar para mejorar la atención de los pacientes en minutos.), por lo que se busca, por medio de un modelo

entender los posibles tiempos de atención futuros, filtrados por días, meses, sedes, de tal manera que se puedan presentar de una forma de fácil interpretación para el usuario final.

2.3.1. Tipos de Modelos de pronósticos de Series de Tiempo

La resolución de problemas relacionados con la predicción de series temporales se presenta como un desafío considerable en el ámbito del modelado predictivo. En contraste con los modelos predictivos de regresión, las series temporales introducen una complejidad adicional al incorporar una dependencia secuencial entre las variables de entrada. Para abordar esta complejidad, se emplea un tipo de red neuronal especialmente diseñada para gestionar secuencias de dependencia, conocida como Redes Neuronales Recurrentes.

En el presente proyecto se utilizarán varios modelos de predicción para definir cuál se adapta más a nuestro proyecto con los modelos : Arima, LSTM, Random Forest, Prophet, Teoría de Colas , Bosque aleatorio.

2.3.2. Definición de métricas adecuadas de desempeño

Al desarrollar modelos de aprendizaje automático, generalmente comparamos varios modelos durante la fase de construcción. Luego, estimamos los rendimientos de esos modelos y seleccionamos el modelo que considera que tiene mayores probabilidades de funcionar bien.

Necesitamos medidas objetivas con técnicas de desempeño para poder decidir qué pronóstico conservar como su pronóstico real, y por el uso de las siguientes métricas: MSE, MAE, MAPE, RMSE, R2

2.3.3. Modelos Usados

A partir de programas realizados en Python, y las variables definidas se crearon varias funciones que sirvieron para iniciar los entrenamientos o aprendizajes de los algoritmos para cada uno de los modelos a probar. (Müller & Guido, 2016)

Se crearon funciones generales para iniciar los entrenamientos de los modelos (Función para evaluar Métricas, Función para División de para entrenamiento y prueba, Función para agrupamiento de variables de tiempo (hora, día de la semana, semana, turno, mes).

Con las funciones anteriores queremos aplicar la misma metodología en cada de las modelos a probar, en específico la función de métricas nos va ayudar a comparar la mayoría de los modelos ya que define cual se acerca más a la tendencia en tiempo de los datos a probar en este caso, vamos a entrenar los algoritmos con datos desde el año 2017 que representan 1.375.710 registros.

Detalle completo de los resultados de cada modelo Ver el link : <https://github.com/Vitotoju/Compensar/blob/main/documentos/Detalle%20de%20los%20modelos.pdf>

A continuación, en forma general se da una conclusión de cada modelo después de ejecutar cada sentencia de Python: (Naren, 2023)

Modelo Prophet : se acerca mucho al dato histórico, y además el R2 está muy cerca de 0.

Modelo Teoría de Colas : es un modelo completo, y largo para ejecutar en este proyecto debido que pide 3 variables importantes (servidores, médicos, estaciones) el cual hace complejo con las variables seleccionadas ejecutar o entrenar este modelo de forma correcta.

Modelo Bosque Aleatorio: este modelo no se ajusta a nuestro proyecto, pide unas variables que no se ajusta al proyecto.

Modelo Arima: se acerca mucho al dato histórico cuando se selecciona la línea de tiempo menor a 3, pero cuando se cambia la variable por más tiempo el modelo difiere mucho del histórico y además el R2 está más lejos del 0.

Modelo LSTM Predictions : este modelo no se acerca en nada a la tendencia del dato histórico, y además el R2 está más lejos del 0.

2.4. Diseño e Implementación BI

A partir de la información definida anteriormente, se genera una base de datos depurada, el cual desde la herramienta Power BI se incorpora esta Data, lo cual se diseña las diferentes páginas

inicialmente para ir creando el tablero adecuado según la necesidad de la subred, adicionando filtros en cada tablero como son : (Mes, Sede, tipo de clasificación Triage) y gráficas en barras, tales como : Día de la semana, Hora día, turno y mes).

Además, para mejorar las gráficas y ajustar los datos de acuerdo a la necesidad se realizan varias técnicas de transformación de datos como son Power Query y Python para optimizar estos tableros o DashBoard.

A continuación, se relaciona la ruta GitHub de la construcción de tableros -las fases (Repositorio) :

<https://github.com/Vitotoju/Compensar/tree/main/powerbi>

Grafica del Diseño inicial de tableros o páginas de Power BI (ver figura 2)



Figura 2 – Diseño inicial de Tableros en Power BI

2.5. Implementación del Sistema de Monitoreo

Se puede evidenciar el diseño final del sistema monitoreo en :

- Aplicación web : <https://compensar-bxzvtymht9gdkqhxpnmnlb.streamlit.app/>
- Power BI Intranet : <https://app.powerbi.com/groups/me/reports/95c6bea4-ce54-4cde-822e-7150742ab1dd/ReportSection26379441c7904a83988e?experience=power-bi>

3.Resultados – Discusiones

Al analizar los datos recolectados en la función de evaluación de Métricas se evidencia que el modelo Prophet está más cercano al cero (0) y es más cercano a los datos históricos y al verificar también la gráfica de Predicción de los diferentes Modelos se evidencia la línea de Prophet es más parecida a la línea original ósea a la realidad de los datos históricos.

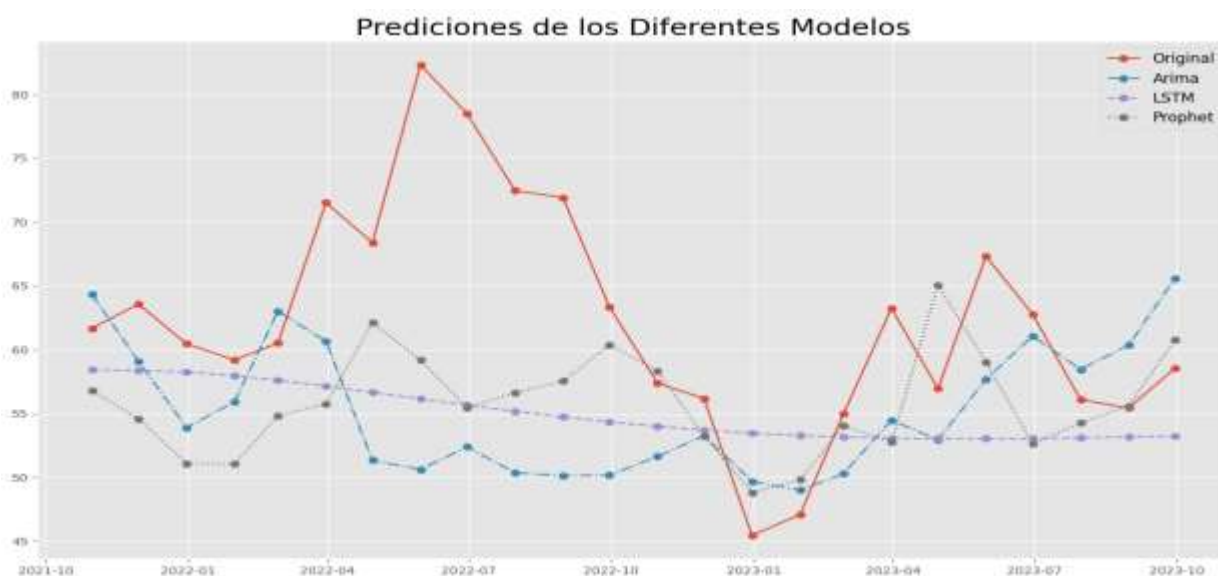


Figura 3 – Comparación Gráfica de los diferentes modelos de series de tiempo

	TimeSeries	MSE	MAE	RMSE	MAPE	R2
0	Prophet	104.123758	7.949468	10.204105	11.858376	-0.397853
1	ARIMA	153.197501	9.182392	12.377298	13.574875	-1.056664
2	LSTM	119.324184	8.501765	10.923561	12.784076	-0.601918

Figura 4 – Comparación Evaluación Métricas de los diferentes modelos de series de tiempo

3.1. Implementación del modelo seleccionado

De acuerdo a los resultados generados por cada uno de los modelos de predicción se da como conclusión el usar el **Modelo Prophet** para nuestro proyecto de predicción de tiempos de atención de los usuarios de acuerdo a Horas, días, semanas, turnos, meses.

Y así por medio de varias herramientas tales como GitHub (repositorio), Visual Studio Code (para realizar programas de Python), Power BI (para diseñar Dashboard) y Streamlit (herramienta para publicar proyecto como página web), se desarrolla diferentes tableros, aplicaciones, gráficas y como solución final para la visualización de datos históricos y predicción de datos así :

- Tablero de Predicción de Tiempos (Power Bi – Intranet) (ver Figura 6)
- Aplicación Web :Ver aplicación web en la página de Streamlit : <https://compensar-bxztvymht9gdkqhxpmbalb.streamlit.app/> (ver figura 7)



Figura 5 – Gráficas de aplicación Web – Tiempos de atención de Urgencias – Predicción

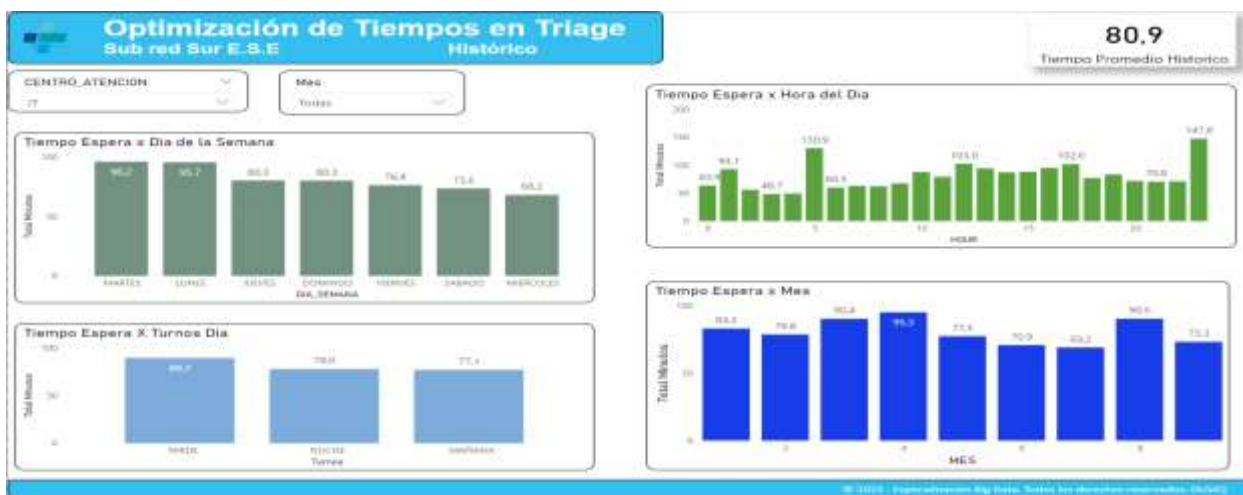


Figura 6 – Gráficas de Power BI – Histórico de atenciones

4. Conclusiones y recomendaciones

4.1. Conclusiones

La minuciosa investigación y el análisis exhaustivo llevado a cabo posibilitaron la identificación de las variables clave que ejercen un impacto significativo en los tiempos de espera, lo cual ha servido como base para la concepción de soluciones adaptadas a las exigencias específicas de la Subred Sur ESE en el ámbito de urgencias.

Al realizar una comparativa gráfica entre diversos modelos, se constata que la trayectoria predictiva del Modelo Prophet se asemeja de manera más precisa a la realidad manifestada en los datos históricos. Este respaldo se refuerza mediante la evaluación de métricas generadas por cada modelo, consolidando así la elección del Modelo Prophet como la herramienta más apropiada para la predicción de tiempos de atención.

La implementación de Dashboard en tiempo real, configurados como tableros de gestión, facilita a los profesionales de la salud la toma de decisiones ágiles en el contexto de la atención a pacientes. Estos tableros no solo posibilitan ajustes proactivos, sino que también contribuyen a la mejora continua de las condiciones de atención, alineándose de manera dinámica con las necesidades operativas del entorno de urgencias.

4.2. Recomendaciones

Para maximizar el beneficio del modelo de predicción implementado, se sugiere adicionar variables nuevas al modelo creado que intervienen en los tiempos de espera: observación, ayudas diagnósticas, registro de historia clínica.

Incorporar esta solución en el APP institucional para que sea accesible desde cualquier dispositivo móvil por parte de los profesionales de la salud y tomar decisiones sobre la marcha sin estar limitados a un escritorio

Establecer un sistema de monitorización incorporando alertas y notificaciones que adviertan a los profesionales de la salud sobre cambios significativos en los datos.

5. Bibliografía

- Naren8520. (2023). "Time Series Forecasting: ARIMA, LSTM, Random Forest, Prophet." GitHub. Recuperado de https://github.com/Naren8520/Python/blob/main/time_series_forecasting_arima_lstm_Random_forest_prophet.ipynb
- Ciencia de Datos. (s.f.). "Modelos ARIMA y SARIMAX en Python." Ciencia de Datos. Recuperado de <https://cienciadedatos.net/documentos/py51-modelos-arima-sarimax-python>
- Müller, A. C., & Guido, S. (2016). "Introducción al aprendizaje automático con Python." O'Reilly Media.
- Raschka, S., & Mirjalili, V. (2017). "Python Machine Learning." Packt Publishing.
- Strome, T. L. (2013). "Healthcare Analytics for Quality and Performance Improvement." John Wiley & Sons.
- Kuhn & Johnson, 2013, 41 (2013). " Applied Predictive Modeling." Max Kuhn & Kjell Johnson.
- Rodríguez-Páez, F. G., Jiménez-Barbosa, W. G., & Palencia-Sánchez, F. (2018). Uso de los servicios de urgencias en Bogotá, Colombia: Un análisis desde el Triage. Universidad y Salud, 20(3), 215-226.
- COBO BORRERO, A. J., & VARGAS CARRILLO, A. L. F. O. N. S. O. (2016). Triage en Emergencias y Desastres en Colombia:¿ cómo, cuándo y dónde utilizarlo?. Revista Gastrohnp, 18.
- Martínez, P., Martínez, J., Nuño, P., & Cavazos, J. (2015). Mejora en el tiempo de atención al paciente en una unidad de urgencias mediante la aplicación de manufactura esbelta. Información tecnológica, 26(6), 187-198.

Restrepo-Zea, J. H., Jaén-Posada, J. S., Espinal Piedrahita, J. J., & Zapata Flórez, P. A. (2018). Saturación en los servicios de urgencias: Análisis de cuatro hospitales de Medellín y simulación de estrategias. *Revista Gerencia y Políticas de Salud*, 17(34), 130-144.