



ESCUELA DE NEGOCIOS

MAESTRÍA EN INTELIGENCIA Y CIENCIAS DE DATOS

**IDENTIFICACIÓN Y MITIGACIÓN DE ZONAS CRÍTICAS DE ACCIDENTES
DE TRÁNSITO EN ECUADOR**

**Profesor
Mario Salvador González**

**Autores
Nicole Yadira Chávez García
Joceline Marissa Salinas Carrión**

2024

RESUMEN

El presente proyecto se centra en la creciente preocupación por la seguridad vial en Ecuador, donde la alarmante tasa de accidentes en ciudades como Quito y Guayaquil representa un desafío significativo. El objetivo principal es evaluar y analizar los patrones espaciales y temporales de estos incidentes utilizando técnicas avanzadas como el análisis geoespacial, clustering y redes neuronales.

A pesar de los esfuerzos realizados para mejorar la seguridad, las estadísticas resaltan la necesidad de identificar con precisión los puntos críticos para desarrollar estrategias preventivas efectivas y asignar recursos de manera adecuada.

La metodología propuesta incluye el uso del algoritmo DBSCAN para identificar agrupaciones de accidentes y la implementación de modelos predictivos como LSTM y con la ayuda de la metodología Prophet se podrá para anticipar futuros incidentes. La estrategia implementada busca tanto la reducción de los accidentes actuales como la promoción de una planificación urbana fundamentada en datos sólidos y confiables.

Los resultados del análisis revelan patrones significativos, evidenciando que ciertos meses y horas presentan picos notables en la ocurrencia de accidentes, así como anomalías que requieren atención especial.

En última instancia, este proyecto pretende reducir la siniestralidad y fomentar la colaboración entre diferentes niveles de gobierno y comunidades, abordando la problemática desde múltiples frentes, de este modo, se contribuirá a crear un entorno urbano más seguro y sostenible en Ecuador promoviendo una cultura de seguridad vial que beneficie a todos los ciudadanos.

ABSTRACT

This project focuses on the growing concern for road safety in Ecuador, where the alarming rate of accidents in cities like Quito and Guayaquil represents a significant challenge. The main objective is to evaluate and analyze the spatial and temporal patterns of these incidents using advanced techniques such as geospatial analysis, clustering, and neural networks.

Despite the efforts made to improve safety, statistics highlight the need to accurately identify critical points in order to develop effective preventive strategies and allocate resources appropriately.

The proposed methodology includes the use of the DBSCAN algorithm to identify clusters of accidents and the implementation of predictive models such as LSTM and with the help of the Prophet methodology it will be possible to anticipate future incidents. The strategy implemented seeks both the reduction of current accidents and the promotion of urban planning based on solid and reliable data

The analysis results reveal significant patterns, indicating that certain months and hours show notable peaks in the occurrence of accidents, as well as anomalies that require special attention.

Ultimately, this project aims to reduce traffic fatalities and foster collaboration between different levels of government and communities, addressing the issue from multiple fronts, in doing so, it will contribute to creating a safer and more sustainable urban environment in Ecuador promoting a culture of road safety that benefits all citizens.

ÍNDICE DEL CONTENIDO

1.	INTRODUCCIÓN.....	1
2.	REVISIÓN DE LITERATURA	3
3.	IDENTIFICACIÓN DEL OBJETO DE ESTUDIO.....	10
4.	PLANTEAMIENTO DEL PROBLEMA.....	12
5.	OBJETIVO GENERAL.....	13
6.	OBJETIVOS ESPECÍFICOS	13
7.	JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA.....	14
8.	RESULTADOS	18
9.	DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN	29
10.	CONCLUSIONES Y RECOMENDACIONES.....	42
11.	REFERENCIAS	44

ÍNDICE DE TABLAS

Tabla 1: Métricas del análisis de clustering	30
Tabla 2: Resultados obtenidos para los modelos Prophet y LSTM	33
Tabla 3: Predicciones de accidentes para los modelos Prophet y LSTM	34

ÍNDICE DE FIGURAS

Gráfico 1: Resumen de variables	14
Gráfico 2: Matriz de correlación.....	18
Gráfico 3: Número total de accidentes por tipo de vehículo	19
Gráfico 4: Número de accidentes por provincia.....	20
Gráfico 5: Número de accidentes por año y mes	20
Gráfico 6:Número de accidentes por hora y año	21
Gráfico 7: Cantidad de accidentes por día de la semana y hora	21
Gráfico 8: Clúster de agrupamiento de accidentes de tránsito en todo el Ecuador	22
Gráfico 9: Clúster de agrupamiento de accidentes de tránsito en Guayaquil...	23
Gráfico 10: Clúster de agrupamiento de accidentes de tránsito en Quito	24
Gráfico 11: Detección de Anomalías en accidentes mensuales.....	25
Gráfico 12: Accidentes mensuales ajustados.....	26
Gráfico 13: Predicciones de Accidentes de Tránsito con Prophet.....	26
Gráfico 14: Predicciones de accidentes de tráfico con LSTM	27
Gráfico 15: Resultado del Check Point Prophet	27
Gráfico 16: Accidentes mensuales ajustados.....	32
Gráfico 17: Predicciones de accidentes de tránsito Prophet vs LSTM	35

INTRODUCCIÓN

La creciente preocupación por la seguridad vial en Ecuador ha aumentado considerablemente, especialmente debido a la alarmante tasa de accidentes de tránsito que se registra en diversas ciudades. Según datos de la Agencia Nacional de Tránsito (ANT), en 2021 se reportaron más de 12.000 accidentes de tránsito, resultando en aproximadamente 1600 muertes y miles de heridos, lo que evidencia un desafío significativo debido a que Ecuador se posiciona entre los países con las tasas más altas de siniestralidad en la región.

Este proyecto se centra en la identificación y mitigación de las zonas críticas donde ocurren la mayoría de los accidentes, con el objetivo de comprender mejor los patrones temporales y espaciales de estos incidentes.

A través del uso de técnicas avanzadas como, el análisis geoespacial, clustering y redes neuronales, se busca identificar los factores que contribuyen a la elevada incidencia de colisiones vehiculares en áreas urbanas específicas.

A pesar de los esfuerzos realizados para mejorar la seguridad vial, como la implementación de campañas de concientización y la instalación de dispositivos de control de velocidad en puntos críticos, las estadísticas indican que el país continúa enfrentando desafíos significativos.

Por lo tanto, es necesario identificar con precisión los puntos de mayor riesgo, ya que esto permitirá desarrollar estrategias preventivas y efectivas fundamentadas en datos concretos.

El análisis propuesto facilitará una asignación más eficiente de los recursos disponibles y contribuirá a la formulación de políticas públicas que realmente aborden la reducción de accidentes, cabe destacar que los accidentes de tránsito representan un impacto económico considerable, generando costos elevados para el sistema de salud y repercusiones en la productividad laboral.

La metodología que se propone integra herramientas avanzadas para proporcionar una visión más clara y predictiva sobre los accidentes de tránsito, se utilizarán técnicas como el clustering espacial y el análisis de series temporales para identificar factores clave, incluidas las condiciones ambientales y la densidad del tráfico, tal como sugieren Wenquan, Azimbek, & Dohyeun (2020).

Al abordar este problema de manera integral, se espera que los resultados obtenidos del análisis contribuyan significativamente a mejorar la gestión de la seguridad vial en Ecuador, creando un entorno urbano más seguro y sostenible.

REVISIÓN DE LITERATURA

El análisis de accidentes de tráfico se ha convertido en un campo de estudio importante en el ámbito de la seguridad vial, debido a la creciente tasa de siniestralidad y las repercusiones sociales y económicas que estos incidentes conllevan. Según la Organización Mundial de la Salud (OMS), cada año aproximadamente 1.35 millones de personas mueren como resultado de accidentes de tráfico, lo que los convierte en una de las principales causas de muerte a nivel global, además se debe tener en cuenta de que, esta problemática es más aguda en países en desarrollo, donde la infraestructura vial es a menudo deficiente y la educación sobre seguridad vial es escasa.

En Ecuador, las estadísticas revelan un incremento alarmante en las muertes por accidentes de tráfico. Según datos del Instituto Nacional de Estadística y Censos (INEC), en 2.022 se registraron más de 6.000 accidentes que resultaron en 2.200 muertes, demostrando la urgencia de abordar esta crisis.

Los datos obtenidos han llevado a investigadores y responsables de políticas a buscar soluciones efectivas, basadas en datos para mitigar el impacto de los accidentes, ya que la comprensión de los factores que contribuyen a los accidentes es esencial para el desarrollo de estrategia preventivas y la implementación de políticas que prioricen la seguridad vial.

La identificación de patrones y tendencias en los datos de accidentes permite a los responsables de la toma de decisiones, dirigir recursos y esfuerzos hacia áreas críticas, además de fomentar una cultura de seguridad vial más sólida y consciente. La adopción de un enfoque multidisciplinario que combine la ingeniería de tráfico, la planificación urbana y el análisis de datos es fundamental para abordar el problema desde diferentes ángulos, además, el compromiso de la sociedad civil y la educación pública juegan un papel vital en la creación de un entorno más seguro para todos los usuarios de las vías.

Basado en los artículos revisados, se han identificado diversos estudios que destacan el uso de técnicas de clustering espacial y redes neuronales para identificar patrones de accidentes de tránsito en zonas de alta incidencia.

La creciente disponibilidad de datos junto con el avance en técnicas de análisis de datos ha permitido a los investigadores aplicar modelos predictivos más complejos para entender mejor las causas de los accidentes de tráfico, tanto los modelos predictivos como las redes neuronales han demostrado su capacidad para analizar y predecir la ocurrencia de accidentes facilitando la identificación de tendencias y patrones que de otro modo podrían pasar desapercibidos.

En ese contexto, Amir et al. (2020) investigaron la vulnerabilidad a colisiones en áreas cercanas a escuelas en Minneapolis, EE. UU., utilizando un análisis de vulnerabilidad y clustering espacial basado en densidad.

Otro estudio relevante es el de Tiana et al. (2021), quienes integraron variables espaciales y temporales en un modelo multinivel para predecir zonas de alta incidencia de accidentes en varias ciudades del este de China, la inclusión de estas variables mejoró significativamente la predicción de las zonas de alto riesgo, destacando la necesidad de enfoques de modelado más sofisticados que consideren la interacción entre el espacio y el tiempo para aumentar la precisión de las intervenciones en seguridad vial.

Además, la investigación de Peijie, Xianghai, & Li, (2022) combina análisis espacio-temporal para revelar cómo la hora del día y las condiciones climáticas influyen en los patrones de colisiones. De igual manera Huiying, Xuan, & Qiang, (2019) integran datos climáticos en su análisis, demostrando mejoras sustanciales en la identificación de zonas de alto riesgo bajo diferentes condiciones climáticas, lo que apoya la necesidad de una gestión del tráfico más adaptativa.

Jinghui et al. (2020) implementaron redes neuronales de tipo Long Short-Term Memory (LSTM) para modelar la frecuencia de accidentes en áreas urbanas.

LSTM es un tipo de red neuronal recurrente, que se ha diseñado específicamente para superar las limitaciones de las redes neuronales tradicionales en la captura de dependencias a largo plazo en los datos (Marin, 2024), lo cual es importante en el análisis de accidentes, donde los patrones pueden no ser evidentes en periodos cortos, sino que pueden influenciarse por eventos pasados que ocurrieron mucho antes.

Las LSTM logran esto mediante una arquitectura que incluye celdas de memoria capaces de retener información a lo largo de períodos prolongados (Marin, 2024), lo cual permite a los modelos captar dinámicas complejas y no lineales en los datos de tráfico. Según Cabrera et al. (2023), los resultados indicaron que el modelo LSTM superó a enfoques más tradicionales, como el modelo autorregresivo integrado de medias móviles (ARIMA) y los modelos de regresión, logrando una precisión de predicción superior al 90%, lo que resalta la capacidad de las LSTM para identificar patrones en los datos de tráfico, lo que puede conducir a intervenciones más eficaces en la seguridad vial.

Otra herramienta relevante en el ámbito de la predicción de series temporales es Prophet, un modelo de código abierto desarrollado por Facebook que está diseñado para facilitar la predicción de datos de series temporales con tendencias y estacionalidad (Chaure, 2022). Prophet es especialmente útil para datos que exhiben patrones estacionales claros y se adapta bien a las irregularidades, lo que lo convierte en una opción atractiva para el análisis de accidentes de tráfico. Al utilizarlo, los investigadores pueden predecir las tendencias en la ocurrencia de accidentes a lo largo del tiempo, teniendo en cuenta factores como las variaciones estacionales (por ejemplo, aumentos en la siniestralidad durante las festividades o condiciones climáticas específicas), lo que ayuda a los responsables de políticas a anticipar picos en la siniestralidad y también proporciona información valiosa para implementar medidas preventivas y de respuesta más efectivas (Cabrera M. , 2021). El uso de Prophet en combinación con técnicas de machine learning puede ofrecer un enfoque robusto y complementario para el análisis de datos de tráfico.

Rol del análisis estadístico en la comprensión de los accidentes

A través del análisis estadístico, los modelos de regresión y las redes neuronales proporcionan información valiosa sobre las variables que influyen en los accidentes, como la velocidad del tráfico, las condiciones climáticas y el comportamiento del conductor (Arias, 2021). La comprensión de cómo estas variables interactúan es fundamental para la formulación de políticas efectivas.

Además, el análisis de la relación entre diferentes variables, como el tipo de vehículo involucrado en un accidente y la gravedad de este, permite a los responsables de la formulación de políticas priorizar los recursos en función de la gravedad potencial de los accidentes, lo que aumenta la efectividad de las intervenciones y contribuye a la creación de un marco de seguridad vial más sólido y adaptativo.

Clustering geoespacial y análisis de patrones de accidentes

El uso de técnicas de clustering geoespacial ha demostrado ser fundamental para la identificación de puntos críticos en accidentes de tráfico. El análisis espacial permite visualizar la distribución de los accidentes en un contexto geográfico, facilitando la identificación de áreas con alta densidad de siniestros (Bravo P. , 2022). Métodos como DBSCAN son particularmente eficaces en este contexto, ya que agrupan datos y manejan adecuadamente los outliers, teniendo una mejora notable en la precisión de la identificación de áreas donde se requieren intervenciones de seguridad.

Gary & Bariviera (2021) aplicaron DBSCAN en la ciudad de Guayaquil, encontrando que la técnica identificó clústeres de accidentes y también permitió a los responsables de la planificación urbana tomar decisiones basadas en evidencia sobre dónde dirigir recursos y estrategias de mitigación. Por ende, este tipo de análisis es relevante para la identificación de puntos críticos y para evaluar la efectividad de las intervenciones implementadas. Por ejemplo, si se

realiza una mejora en la infraestructura vial en una zona previamente identificada como crítica, el seguimiento a través de análisis geoespacial puede revelar si ha habido una reducción en la frecuencia de accidentes.

Desafíos y limitaciones en el uso de modelos predictivos

A pesar de los avances en el uso de modelos predictivos, es fundamental reconocer que los modelos más simples, como los árboles de decisión y la regresión logística, siguen teniendo un papel importante en el análisis de accidentes. García & Ortega (2020) argumentan que aunque estos modelos son menos sofisticados y más accesibles pueden ser eficaces en contextos donde los recursos son limitados.

En un estudio reciente, López, Jiménez, & Martínez (2022) utilizaron regresión lineal para predecir la severidad de los accidentes en una región rural de Ecuador, obteniendo resultados satisfactorios a pesar de la simplicidad del enfoque, lo que resalta la importancia de que la selección del modelo debe considerar también la disponibilidad de datos y recursos técnicos.

Otro desafío importante en la implementación de modelos predictivos es la calidad y disponibilidad de los datos, muchos países en desarrollo enfrentan la falta de datos completos y actualizados sobre accidentes de tráfico, lo que limita la efectividad de los modelos basados en datos.

En este sentido, es esencial establecer sistemas de recopilación de datos más robustos que fomenten la colaboración entre diferentes instituciones, incluyendo gobiernos, universidades y organizaciones no gubernamentales para mejorar la calidad y la cantidad de los datos disponibles, lo que a su vez enriquecerá los análisis realizados.

Aplicaciones prácticas en el contexto ecuatoriano

En el contexto ecuatoriano, la investigación sobre accidentes de tráfico es particularmente relevante dado el incremento en las tasas de siniestralidad en ciudades como Quito y Guayaquil. La combinación de análisis temporal y geoespacial es vital para informar políticas de tráfico y mejorar la seguridad vial. Los estudios revisados sugieren que un enfoque integrador que combine predicciones basadas en datos con la identificación de patrones espaciales puede ofrecer soluciones más efectivas a los problemas de seguridad vial.

Al identificar áreas con altas tasas de accidentes, las autoridades pueden priorizar las intervenciones en infraestructura, como la mejora de la señalización y la implementación de sistemas de control de velocidad. La educación y sensibilización de los conductores también pueden ser aspectos clave en la reducción de la siniestralidad, y los modelos predictivos pueden ayudar a identificar momentos críticos en los que se necesitan campañas de concienciación más intensivas.

La implementación de estos enfoques puede mejorar la calidad de vida de los ciudadanos al promover un entorno más seguro para todos los usuarios de la vía. La seguridad vial además de ser una cuestión de infraestructura es también un comportamiento humano, y la integración de datos y análisis puede ser un paso necesario hacia una solución integral y sostenible.

El análisis de accidentes de tráfico y la utilización de modelos predictivos son fundamentales para abordar la creciente siniestralidad en Ecuador. La combinación de análisis temporal, geoespacial y modelos avanzados de machine learning ayuda a identificar patrones y tendencias en los accidentes además de proporcionar información valiosa para la formulación de políticas efectivas. A medida que se continúe investigando y aplicando estos enfoques, es necesario recordar la importancia de la calidad de los datos y la colaboración

entre los diferentes sectores para lograr un impacto significativo en la seguridad vial.

También se han identificado varios factores que influyen en la ocurrencia de accidentes, un estudio realizado por Bravo (2020) en Cuenca destacó que las fallas mecánicas en vehículos son una causa significativa de siniestros, el estudio se basa en datos recopilados entre 2010 y 2018, donde se examinaron las condiciones mecánicas y su relación con los accidentes.

Además, se ha observado que el comportamiento del conductor y las condiciones de las vías son factores críticos que deben ser considerados en cualquier estrategia de prevención según lo mencionado por Medina, Escobar, & Ulloa (2021).

IDENTIFICACIÓN DEL OBJETO DE ESTUDIO

El objetivo principal de este proyecto es evaluar y analizar los patrones espaciales y temporales de los accidentes de tránsito en Ecuador, centrándose en las zonas críticas de colisiones vehiculares mediante técnicas avanzadas de análisis geoespacial, clustering y redes neuronales. El problema central que aborda esta investigación es la alta incidencia de accidentes en áreas urbanas específicas, lo que representa un desafío considerable para la seguridad vial y la gestión del tráfico en el país.

La importancia de este proyecto radica en la necesidad de desarrollar estrategias preventivas efectivas basadas en datos precisos y análisis detallados, ya que a pesar de los esfuerzos actuales la tasa de accidentes en Ecuador sigue siendo elevada, especialmente en zonas con alta incidencia de colisiones. El identificar con precisión los puntos críticos de accidentes permitirá una mejor asignación de recursos y la implementación de medidas específicas para mitigar el riesgo.

Este análisis tiene como objetivo proporcionar una base sólida para la formulación de políticas públicas y estrategias enfocadas en reducir los accidentes viales, combinando enfoques de ingeniería y educación (Castillo, Coral, & Salazar, 2020).

Para abordar este problema, se emplearán técnicas avanzadas como el clustering espacial, el análisis geoespacial y las redes neuronales, dichos métodos permitirán una comprensión más profunda de las dinámicas subyacentes y de los factores que contribuyen a los accidentes, tales como las condiciones ambientales, la densidad del tráfico y la configuración vial (Congacha, Barba, & Palacios, 2019).

La implementación de este proyecto contribuirá a mejorar la gestión de la seguridad vial en Ecuador mediante el análisis detallado de las áreas de alta incidencia de accidentes. Las recomendaciones derivadas del análisis permitirán a las autoridades aplicar medidas más eficientes y basadas en evidencia, con el

objetivo de reducir la tasa de accidentes y mejorar la seguridad en las zonas más afectadas.

El objeto de estudio de este proyecto es el análisis predictivo de accidentes de tráfico en Ecuador, con un enfoque en las ciudades de Quito y Guayaquil. Mediante la implementación de técnicas avanzadas de análisis de datos y modelado predictivo, como redes neuronales y análisis de series temporales, se busca identificar patrones que permitan predecir la ocurrencia de accidentes en áreas críticas. El propósito es proporcionar información valiosa para mejorar la planificación urbana y reducir la siniestralidad en las zonas más afectadas del país.

PLANTEAMIENTO DEL PROBLEMA

En Ecuador, la alta incidencia de accidentes de tránsito se ha convertido en un problema significativo, a pesar de las diversas iniciativas implementadas para mejorar la seguridad vial. La elevada frecuencia de estos accidentes indica que las políticas actuales no son efectivas y para abordar esta situación es fundamental fortalecer la coordinación y promover la colaboración entre diferentes niveles de gobierno y actores comunitarios, adoptando un enfoque integral.

Los factores determinantes de los accidentes de tránsito incluyen características del entorno urbano, alta densidad de tráfico y complejidad de la infraestructura vial, estos elementos aumentan el riesgo de accidentes y complican la implementación de soluciones efectivas (Holguín, Otzen, & Salgado, 2024).

El análisis convencional de accidentes a menudo se ve limitado por la falta de herramientas avanzadas que permitan identificar de forma precisa y dinámica las zonas de alto riesgo, conocidas como puntos calientes de colisión. Según Gomez, Hernandez, & Campos (2019), esta limitación restringe la capacidad de las autoridades para diseñar e implementar intervenciones preventivas eficaces adaptadas a las necesidades específicas de cada área problemática.

Investigaciones recientes han resaltado la importancia de utilizar métodos avanzados de análisis geoespacial y técnicas de agrupamiento para mejorar la identificación de áreas críticas de colisión. Según Contreras et. al. (2022) la incorporación de tecnologías avanzadas en el análisis de accidentes es esencial para reducir efectivamente la tasa de colisiones en las ciudades.

Uno de los principales efectos negativos en la seguridad vial es la falta de coordinación entre las autoridades locales y nacionales, la fragmentación en la gestión y la falta de comunicación entre las entidades responsables limitan la efectividad de las políticas de seguridad vial.

OBJETIVO GENERAL

Desarrollar un modelo predictivo para identificar y mitigar las zonas críticas de accidentes de tránsito en Ecuador empleando técnicas de análisis de datos avanzadas, con el fin de mejorar la efectividad de las políticas de seguridad vial y reducir la incidencia de accidentes.

OBJETIVOS ESPECÍFICOS

Identificar y clasificar las zonas críticas de accidentes de tránsito en Ecuador mediante el uso del algoritmo DBSCAN para realizar un análisis de clustering, con la finalidad de agrupar los datos de accidentes según características espaciales y temporales.

Detectar y corregir anomalías en los datos de accidentes utilizando técnicas estadísticas para asegurar que el conjunto de datos utilizado para el análisis posterior sea preciso y representativo.

Desarrollar y comparar modelos de series temporales y redes neuronales para predecir la cantidad de accidentes de tránsito futuros.

Medir el rendimiento de los modelos Prophet y LSTM en términos de precisión y fiabilidad en las predicciones, y utilizar los hallazgos para formular recomendaciones que mejoren la efectividad de las políticas de seguridad vial en Ecuador.

JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA

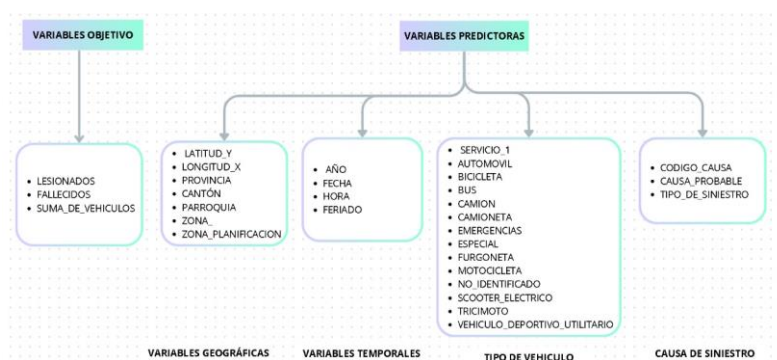
La metodología utilizada en este proyecto se basa en la combinación de herramientas avanzadas de análisis de datos para abordar de manera integral la problemática de los accidentes de tránsito en Ecuador.

El uso de técnicas como detección de anomalías, clustering, análisis de series temporales y redes neuronales LSTM permite obtener una visión profunda y predictiva de los patrones de los accidentes.

Para el análisis de datos, se utilizará Python como herramienta principal, debido a sus librerías especializadas en análisis de datos y machine learning como pandas, scikit-learn y TensorFlow.

El proceso comienza con la limpieza y transformación de los datos, se eliminan datos faltantes y duplicados y se procede a extraer las características clave relacionadas con los accidentes. El *Gráfico 1* muestra todas las variables utilizadas para el análisis del proyecto.

Gráfico 1: Resumen de variables



Una vez que se tiene la base limpia, se procede a realizar un análisis descriptivo para poder entender las características principales de los datos, identificar tendencias, distribuciones y valores atípicos. Después, se realiza la matriz de correlación para explorar las relaciones entre las diferentes variables del

conjunto de datos, para esto se utilizan todas las variables mostradas en el *Gráfico 1*.

Se continua con el análisis de clustering utilizando el algoritmo DBSCAN, el cual permite identificar agrupaciones de accidentes de tránsito en función de la proximidad geográfica y temporal. DBSCAN es ideal para este tipo de análisis porque detecta clústeres de forma no lineal y es robusto frente a ruido o valores atípicos, lo que lo hace especialmente útil en escenarios con datos geoespaciales.

Para el modelo de clustering se utilizan las variables de coordenadas geográficas (LATITUD_Y y LONGITUD_X), la fecha, la suma de fallecidos y de lesionados.

Es importante señalar que se llevan a cabo tres análisis de clustering: uno que abarca todo el Ecuador y, en función de los polígonos obtenidos se seleccionan las diez áreas más grandes para un análisis más detallado. Posteriormente, se centran en las dos provincias con mayor superficie.

A continuación, se lleva a cabo la detección de anomalías en las series temporales de accidentes con el objetivo de identificar picos inesperados o eventos inusuales que podrían distorsionar los análisis futuros y una vez detectadas las anomalías se corrigen reemplazando por la media de los valores circundantes en la serie, asegurando que los análisis y predicciones posteriores mantengan la coherencia con las tendencias generales sin verse influenciados por esos eventos extremos. Para la detección de anomalías se usan las variables de fecha, especialmente el mes y el número de accidentes.

Se continua con el análisis predictivo utilizando Prophet, una librería especializada que permite pronosticar datos de series temporales basados en un modelo aditivo, en el que las tendencias no lineales se ajustan a la estacionalidad anual, semanal y diaria, la cual funciona mejor con series temporales que tienen fuertes efectos estacionales y varias temporadas de

datos históricos (Hyndman & Athanasopoulos, 2020). Se utiliza en varias aplicaciones de Facebook lo cual permite producir previsiones fiables para la planificación y el establecimiento de objetivos. En esta investigación Prophet se utilizará para analizar y predecir el número de accidentes en el tiempo, tomando en cuenta tendencias a largo plazo y estacionalidades (mensuales), para ello se utilizan las variables de fecha y número de accidentes, cabe recalcar que la frecuencia utilizada es meses.

Para el análisis de Prophet se utilizan las variables de fecha y número de accidentes, cabe recalcar que la frecuencia utilizada es meses.

Además, se realiza un análisis de changelog para identificar momentos de cambio significativo en las series temporales de accidentes de tráfico. El changelog permite detectar rupturas o cambios de tendencia en los datos históricos, indicando posibles eventos o factores que influyan en el comportamiento de los accidentes como modificaciones en las políticas de tránsito, campañas de concientización, o variaciones estacionales (Gutierrez, 2019).

Además del análisis con Prophet, se implementará un modelo de redes neuronales de tipo LSTM que es ideal para trabajar con series temporales debido a su capacidad de recordar información a largo plazo y captar dependencias no lineales en los datos, el modelo permitirá realizar predicciones más sofisticadas sobre la ocurrencia de accidentes y evaluar su rendimiento en comparación con Prophet, se entrenará la red neuronal con los datos históricos de accidentes, y se realizarán pruebas para validar su precisión.

El modelo LSTM utilizará las mismas variables que para el análisis con Prophet: fecha y número de accidentes.

Tanto para el análisis con Prophet como para el de LSTM, se utiliza un enfoque que divide los datos en dos partes: una parte de entrenamiento y otra de prueba.

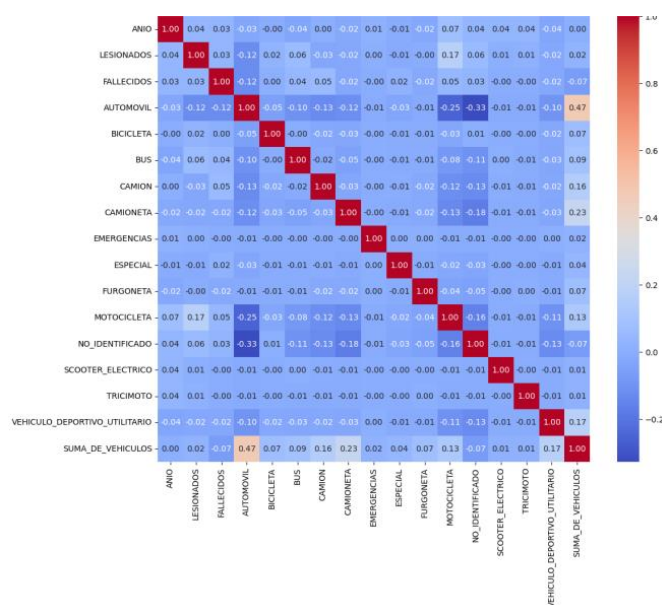
La parte de entrenamiento se usa para ajustar el modelo, permitiéndole aprender las tendencias y patrones históricos y, se aplica la parte de prueba para evaluar cómo funciona el modelo en datos que no ha visto antes, verificando su capacidad predictiva, ambos modelos se utilizan para predecir un período adicional en meses con el objetivo de proyectar los futuros accidentes de tráfico.

Finalmente, se compararán los resultados obtenidos de Prophet y LSTM. Se evaluarán métricas de rendimiento como el error cuadrático medio (MSE), el error absoluto medio (MAE) y la raíz del error cuadrático medio (RMSE) para determinar cuál de los modelos proporciona las predicciones más precisas y útiles para la gestión de la seguridad vial en Ecuador.

RESULTADOS

Lo primero que se obtiene en este proyecto es la matriz de correlación, con el fin de analizar las relaciones entre diversas variables relacionadas con los accidentes de tránsito en Ecuador, de esta manera se puede tener una visión clara de cómo diferentes factores se interrelacionan y afectan la ocurrencia de accidentes.

Gráfico 2: Matriz de correlación



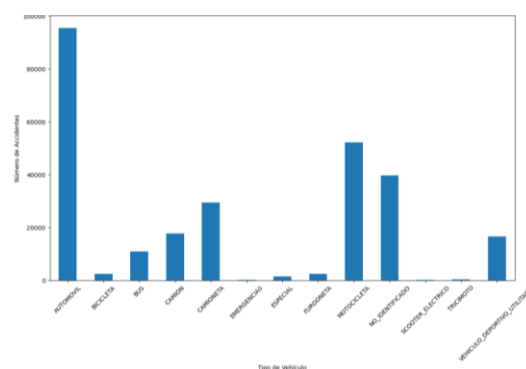
En el Gráfico 2 se puede observar la matriz de correlación, la cual muestra las relaciones entre varias variables relacionadas con incidentes vehiculares, lesiones y muertes. Siendo lo más importante:

- En general cuando más automóviles están involucrados en un accidente, no necesariamente hay más lesionados.
- Puede ser que los accidentes con automóviles resulten menos graves en comparación con otros vehículos, como motocicletas, que suelen estar más expuestas.
- Cuando hay accidentes que involucran motocicletas, es más probable que haya personas lesionadas, lo que confirma que los motociclistas están en mayor riesgo de sufrir lesiones cuando están involucrados en un accidente.

- La ubicación exacta de un accidente no está fuertemente relacionada con cuántas personas fallecen en él, sin embargo, pequeños patrones pueden estar presentes, indicando que algunos lugares pueden tener un riesgo ligeramente mayor.
- Los automóviles son frecuentemente parte de los accidentes en los que hay muchos vehículos involucrados.
- Ningún tipo de vehículo específico está claramente vinculado a un aumento en la probabilidad de fallecimientos en los accidentes. Sin embargo, algunos vehículos como los camiones, pueden estar ligeramente más involucrados en accidentes con fallecidos debido a su tamaño y la gravedad potencial del impacto.

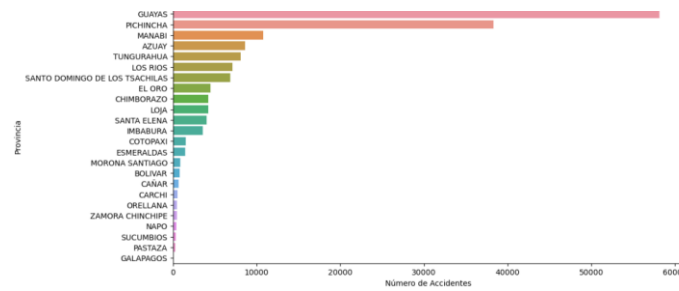
A continuación, se realiza el análisis exploratorio de datos (EDA) con el fin de profundizar en la comprensión de las características y patrones de los accidentes de tránsito en Ecuador. Durante el EDA, se elaboran diferentes gráficos y visualizaciones, en especial gráficos de barras y de líneas, que facilitan la interpretación de los datos.

Gráfico 3: Número total de accidentes por tipo de vehículo



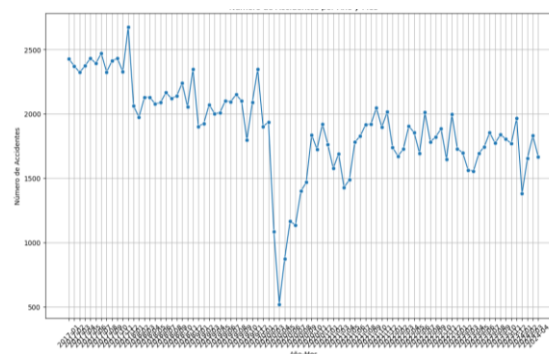
En el *Gráfico 3* se puede observar que los automóviles tienen el mayor número de accidentes, seguidos de motocicletas que también presentan un número notable de incidentes.

Gráfico 4: Número de accidentes por provincia



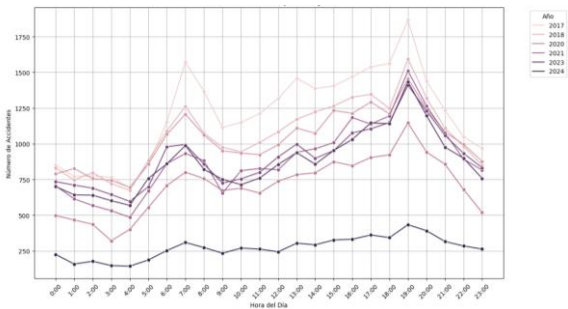
Además, el *Gráfico 4* nos permite visualizar las áreas con mayor incidencia de siniestros viales. Los resultados muestran que las provincias de Guayas y Pichincha destacan notablemente, registrando la mayor cantidad de accidentes.

Gráfico 5: Número de accidentes por año y mes



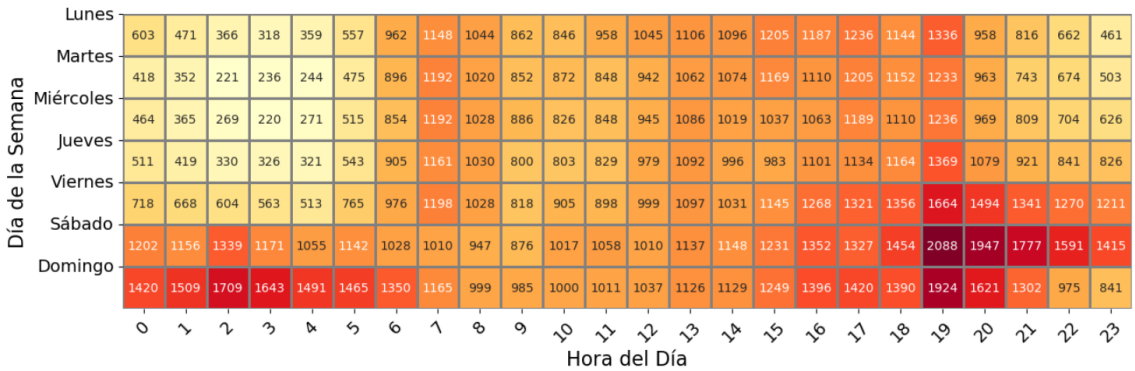
El *Gráfico 5* es uno de los más importantes es el que muestra el número de accidentes por año y mes, donde se evidencian variaciones significativas a lo largo del tiempo. En este gráfico, se observan picos que alcanzan hasta 2,500 accidentes en ciertos meses, seguidos de caídas abruptas, como la notable disminución que se produjo en un período específico. Además, el análisis revela que, en los meses de mayor actividad, como diciembre y enero, se presentan incrementos significativos, lo que resalta la necesidad de implementar medidas preventivas durante estos períodos críticos.

Gráfico 6: Número de accidentes por hora y año



El *Gráfico 6* muestra el número de accidentes por hora y año, el cual revela patrones distintivos en la ocurrencia de siniestros a lo largo del día. Se observa un notable incremento en el número de accidentes durante las horas pico, especialmente entre las 6:00 y 7:00 a.m. y entre las 18:00 y 19:00, donde se registran picos de hasta 1500 accidentes en años anteriores. En contraste, las horas de la madrugada, como las 3:00 y 4:00, muestran cifras significativamente más bajas, con menos de 300 accidentes.

Gráfico 7: Cantidad de accidentes por día de la semana y hora



El *Gráfico 7* muestra la cantidad de accidentes de tránsito a lo largo de la semana, desglosados por horas del día, se puede observar que los domingos tienen la mayor cantidad de accidentes, con un pico de 2088 incidentes a las 18:00 horas, a pesar de que el sábado también presenta cifras elevadas, alcanzando 1664 accidentes a las 20:00. Es decir, los días del fin de semana (sábado y domingo) muestran cifras significativamente más altas en comparación con los días de semana, siendo los lunes, martes y miércoles los

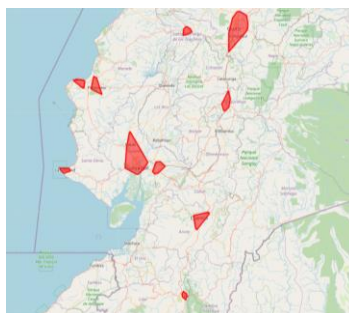
días con menos accidentes, especialmente a las horas tempranas. Las horas de mayor riesgo se concentran entre las 18:00 y 20:00, independientemente del día.

Posteriormente, se aplica el algoritmo de clustering DBSCAN para identificar grupos de accidentes basándose en sus coordenadas geográficas. Primero, se normalizan las coordenadas de latitud y longitud utilizando StandardScaler, lo que permite una mejor interpretación de la distancia entre los puntos. Luego, se aplica DBSCAN con parámetros específicos, analizando diferentes valores de ϵ , que define la distancia máxima entre dos puntos para que se consideren parte del mismo clúster y min_samples , que establece el número mínimo de puntos requeridos para formar un clúster.

Tras evaluar varios valores, se determina que los mejores parámetros son $\epsilon = 0.012$ y $\text{min_samples} = 45$. Con estos ajustes, se obtienen los siguientes resultados:

- un total de 131 clústeres (excluyendo ruido)
- un 83.54% de los puntos asignados a clústeres
- un Silhouette Score de aproximadamente 0.44
- un índice Davies-Bouldin de 0.31

Gráfico 8: Clúster de agrupamiento de accidentes de tránsito en todo el Ecuador



Con los resultados obtenidos en el *Gráfico 8*, se calculó el área de cada polígono correspondiente a los clústeres identificados y se filtraron los 10 más grandes, obteniendo que las ciudades con las áreas más extensas de concentración de

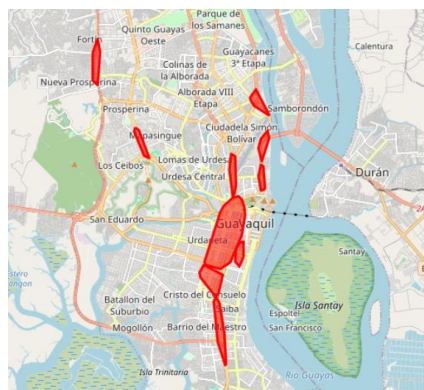
accidentes son: Guayaquil, Quito, Ambato, Cuenca, Santo Domingo de los Tsáchilas, Milagro, Manta, Portoviejo, Loja y Libertad.

Para el caso de estudio, se seleccionarán Guayaquil y Quito, ya que representan dos de las ciudades más pobladas y con mayor actividad vehicular en Ecuador, lo que las convierte en escenarios críticos para el análisis de la siniestralidad y la implementación de medidas de seguridad vial.

Para el análisis de clustering en Guayaquil como se muestra en el *Gráfico 9*, se utilizaron los parámetros $\text{eps} = 0.012$ y $\text{min_samples} = 45$. Con estos ajustes, se obtuvieron los siguientes resultados:

- un total de 131 clústeres (excluyendo ruido)
- un 83.54% de los puntos asignados a clústeres
- el Silhouette Score obtenido fue de aproximadamente 0.44
- el índice Davies-Bouldin fue de 0.31

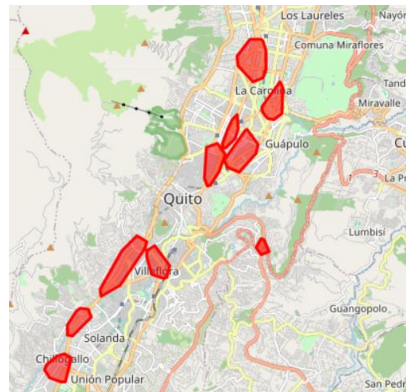
Gráfico 9: Clúster de agrupamiento de accidentes de tránsito en Guayaquil



Para el análisis de clustering en Quito como se indica en el *Gráfico 10*, se utilizaron los parámetros $\text{eps} = 0.035$ y $\text{min_samples} = 50$. Con estos ajustes, se obtuvieron los siguientes resultados:

- un total de 116 clústeres (excluyendo ruido)
- un 62.14% de los puntos asignados a clústeres
- el Silhouette Score obtenido fue de aproximadamente 0.34
- el índice Davies-Bouldin fue de 0.54

Gráfico 10: Clúster de agrupamiento de accidentes de tránsito en Quito



Posteriormente, se realiza la detección de anomalías en el número de accidentes mensuales. Para ello, se definen umbrales de anomalía utilizando 2 desviaciones estándar en lugar de 3, estableciendo un límite superior y un límite inferior para identificar los meses con un número de accidentes significativamente mayor o menor.

Se calculan los siguientes valores:

- Media de accidentes: 1894.11
- Desviación estándar de accidentes: 361.44
- Límite superior: 2617.00
- Límite inferior: 1171.22

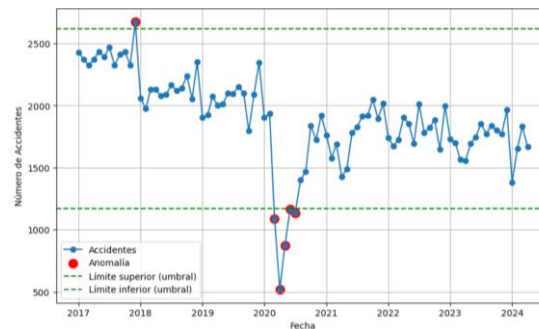
A continuación, se utiliza el modelo Isolation Forest para detectar anomalías, entrenando el modelo con el número mensual de accidentes. Se marca como anomalía aquellos meses donde el modelo identifica un comportamiento atípico.

El análisis revela las siguientes anomalías:

- Diciembre 2017: 2676 accidentes (anomalía positiva)
- Marzo 2020: 1086 accidentes (anomalía negativa)
- Abril 2020: 520 accidentes (anomalía negativa)
- Mayo 2020: 875 accidentes (anomalía negativa)
- Junio 2020: 1165 accidentes (anomalía negativa)
- Julio 2020: 1136 accidentes (anomalía negativa)

Como se muestra en el *Gráfico 11*, en total, se detectaron 6 anomalías en el conjunto de datos, lo que indica que estos meses presentan un comportamiento inusual en comparación con el promedio histórico de accidentes.

Gráfico 11: Detección de Anomalías en accidentes mensuales



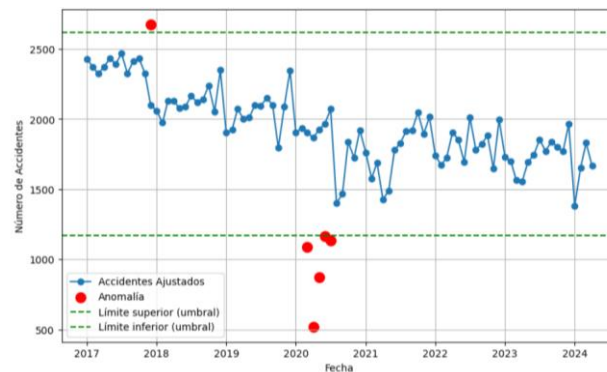
A continuación, primero, se extrae el mes y el año de la fecha para facilitar la agrupación de los datos. Luego, se calcula la media de accidentes por mes, excluyendo las anomalías previamente detectadas.

Los resultados de la media de accidentes por mes son los siguientes:

- Enero: 1863.50
- Febrero: 1851.13
- Marzo: 1905.29
- Abril: 1866.29
- Mayo: 1926.17
- Junio: 1967.00
- Julio: 2071.50
- Agosto: 1924.00
- Septiembre: 1957.71
- Octubre: 2006.00
- Noviembre: 1929.71
- Diciembre: 2099.33

Posteriormente, se reemplazan los valores de las anomalías con la media correspondiente de su mes, iterando sobre cada fila del DataFrame y, si se identifica una anomalía, se sustituye el número de accidentes por la media del mes correspondiente, tal como se muestra en el Gráfico 12.

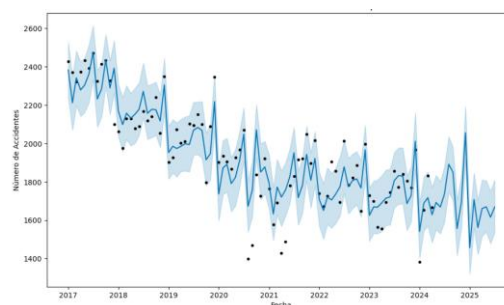
Gráfico 12: Accidentes mensuales ajustados



Se continúa separando los últimos 12 meses de datos para el periodo de prueba, mientras que el resto se utiliza como conjunto de entrenamiento.

A continuación, como se muestra en el *Gráfico 13* se ajusta el modelo de series temporales con la librería Prophet con los datos de entrenamiento y se realiza la predicción de los próximos 12 meses.

Gráfico 13: Predicciones de Accidentes de Tránsito con Prophet

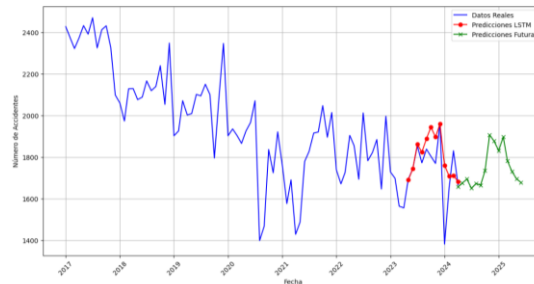


Como siguiente paso, se normalizan los datos utilizando el MinMaxScaler para escalar el número de accidentes entre 0 y 1. Se crean secuencias para el modelo LSTM, utilizando un tiempo de paso de 12 meses.

Los datos se dividen en conjuntos de entrenamiento y prueba, y se reformatean para que el modelo LSTM reciba los datos en la forma adecuada: [muestras, pasos de tiempo, características]. Luego, se crea un modelo LSTM con dos capas de LSTM y una capa densa de salida. El modelo se compila y entrena

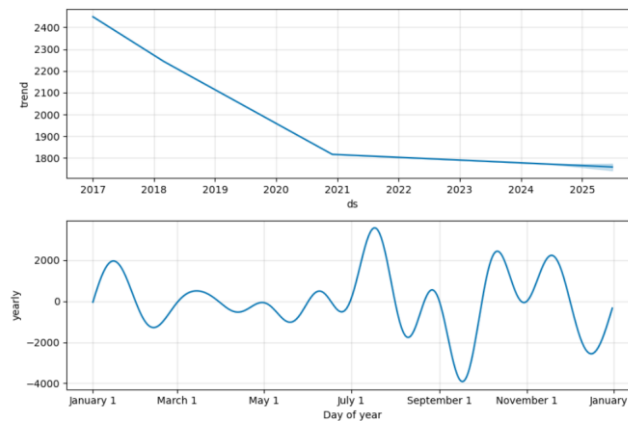
durante 300 epochs. Después del entrenamiento, se realizan predicciones en el conjunto de prueba y se obtienen los siguientes resultados, como se muestra en el *Gráfico 14*.

Gráfico 14: Predicciones de accidentes de tráfico con LSTM



En base al modelo Prophet, se analizan los puntos de cambio para identificar las variaciones significativas en las tendencias del número de accidentes a lo largo del tiempo. En el *Gráfico 15* se observan como resultados los puntos de cambio.

Gráfico 15: Resultado del Check Point Prophet



La gráfica revela una tendencia a la baja en el número de accidentes de tráfico en Ecuador, disminuyendo de aproximadamente 2400 en 2017 a 1700 en 2023, lo que refleja la efectividad de las políticas de seguridad vial implementadas.

Sin embargo, la pandemia de COVID-19, que comenzó en 2020, tuvo un impacto significativo en la movilidad, reduciendo drásticamente el tráfico y, por ende, los accidentes durante los períodos de confinamiento.

La gráfica de estacionalidad también muestra patrones cíclicos, con picos en ciertos momentos del año que indican mayor incidencia de accidentes, probablemente relacionados con festividades y aumento del tráfico.

DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN

Los resultados obtenidos en el análisis de accidentes de tránsito en Ecuador evidencian patrones significativos en la relación entre diferentes variables. La matriz de correlación reveló que, aunque el número de vehículos involucrados en un accidente no necesariamente se traduce en un mayor número de lesionados, el riesgo para ciertos tipos de vehículos, especialmente las motocicletas, es notablemente más alto.

El análisis exploratorio de datos (EDA) permitió identificar los momentos y lugares críticos donde se producen más accidentes, destacando meses y horas específicas con picos significativos, lo que conlleva la implementación de medidas preventivas y para la asignación eficiente de recursos.

Además, se concluye que la alta incidencia de accidentes de tránsito durante los fines de semana, especialmente los sábados y domingos por la noche, puede estar relacionada con varios factores entre ellos el consumo de alcohol. Durante estas noches es común que las personas salgan a fiestas, celebraciones o reuniones sociales, lo que puede llevar a un aumento en la ingesta de bebidas alcohólicas. Además, las horas pico entre las 18:00 y 20:00, coinciden con el final de actividades recreativas, lo que probablemente incrementa el número de conductores en la vía, algunos de los cuales pueden no estar en condiciones óptimas para manejar. La combinación de alcohol, distracciones y un mayor volumen de tráfico explica en gran medida el aumento de accidentes en estos períodos.

Por otro lado, la técnica de clustering mostró la capacidad de identificar áreas de alta concentración de accidentes, lo que permite a las autoridades locales y nacionales focalizar sus esfuerzos en mejorar la seguridad en esos puntos críticos.

El análisis de clustering se realizó utilizando el algoritmo DBSCAN para identificar áreas de alta concentración de accidentes en Ecuador. Se determinaron los mejores parámetros para el análisis tanto para todo el país como para Guayaquil y Quito, los cuales se puede observar en la *Tabla 1*:

Tabla 1: Métricas del análisis de clustering

	Todo Ecuador	Guayaquil	Quito
eps	0.012	0.04	0.035
min_samples	45	40	50
Total clústeres	131	92	116
Clústeres asignados	83.54%	50.91%	62.14%
Sillhouette	0.44	0.34	0.34
Índice Davies-Bouldin	0.31	0.43	0.54

El análisis de los clústeres de accidentes de tráfico en Ecuador, Guayaquil y Quito muestra resultados prometedores, especialmente a nivel nacional.

El número significativo de clústeres formados y el alto porcentaje de puntos asignados en todo Ecuador (83.54%) reflejan una agrupación sólida y bien definida, respaldada por un Silhouette Score positivo de 0.44 y un bajo índice de Davies-Bouldin de 0.31, lo que sugiere una clara separación entre los clústeres.

Aunque Guayaquil y Quito presentan porcentajes más bajos de puntos asignados (50.91% y 62.14% respectivamente), ambos mantienen métricas de calidad aceptables que indican un buen potencial para capturar patrones relevantes de accidentes. Las métricas en general muestran que el modelo logra un equilibrio entre densidad y distribución de los accidentes.

Posteriormente, el análisis de accidentes de tránsito en Ecuador se ha abordado utilizando la técnica de clustering y dos modelos de modelado predictivo: Prophet y LSTM (Long Short-Term Memory). A continuación, se presenta un análisis

detallado de los resultados del clustering, seguido de una comparación de los modelos predictivos y propuestas de solución.

En base al análisis realizado se obtuvo una lista de 10 provincias de Ecuador que tienen la mayor cantidad de accidentes de tránsito:

- Guayas (incluye Guayaquil)
- Pichincha (incluye Quito)
- Azuay (incluye Cuenca)
- Manabí (incluye Manta)
- Esmeraldas
- Santo Domingo de los Tsáchilas
- Los Ríos
- El Oro
- Carchi
- Cotopaxi

Con respecto al análisis de clustering en Guayaquil se obtuvo:

- Urdaneta
- Urdesa
- Mapasingue
- Los Ceibos
- Durán
- La Prosperina
- Samborondón
- La Alborada
- Cerro Azul
- Bastión Popular

Y en base al análisis de clustering en Quito se puede observar que los sectores con más accidentes de tránsito son:

- La Carolina
- El Inca
- Cumbayá
- Solanda
- La Floresta
- Chillogallo
- San Isidro del Inca
- Pomasqui
- Guápulo
- Tumbaco

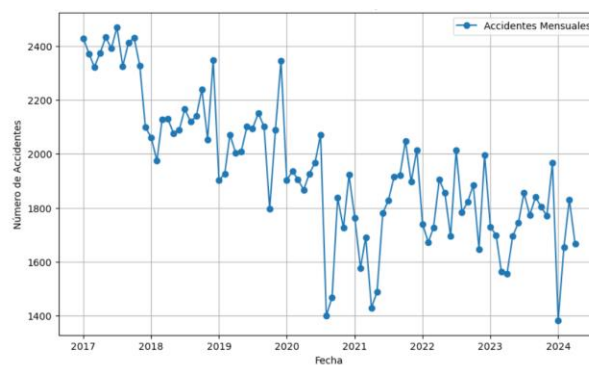
Después de identificar los sectores con mayor cantidad de accidentes de tránsito se procede a la detección de anomalías a través de métodos estadísticos, como el Isolation Forest. El análisis de anomalías en el número de accidentes de tránsito en Ecuador reveló eventos críticos que han influido significativamente en la siniestralidad.

En diciembre de 2017, se registró un notable aumento en los accidentes, que puede atribuirse a concentraciones masivas de personas debido a eventos políticos y sociales, como las manifestaciones y movilizaciones que se llevaron a cabo en ese período, estas aglomeraciones, que incluyeron marchas y protestas, generaron un aumento en el tráfico y la congestión en las vías, lo que elevó el riesgo de accidentes.

En contraste, diciembre de 2020 muestra una anomalía negativa con una drástica caída en el número de accidentes, lo que se relaciona con las restricciones de movilidad impuestas por la pandemia de COVID-19. Durante este tiempo, las medidas de confinamiento y las limitaciones en reuniones sociales llevaron a una reducción significativa en el tráfico vehicular.

Una vez tratadas las anomalías con el promedio de cada mes, se obtiene el *Gráfico 16* que señala las tendencias más estables y representativas del número de accidentes de tránsito a lo largo del tiempo.

Gráfico 16: Accidentes mensuales ajustados



Con esta información consolidada, se puede proceder al análisis predictivo, donde se emplearon modelos como Prophet y LSTM para anticipar el número de accidentes en los próximos meses.

En cada modelo de análisis predictivo se aplicaron métricas para evaluar su efectividad en la predicción del número de accidentes de tránsito. En la *Tabla 2*, se presentan los resultados obtenidos para los modelos Prophet y LSTM.

Tabla 2: Resultados obtenidos para los modelos Prophet y LSTM

Métrica	Modelo Prophet	Modelo LSTM
MSE	47561.08	25209.68
RMSE	218.09	158.78
MAE	184.62	102.76

El análisis de métricas de los modelos predictivos de accidentes de tránsito en Ecuador revela que el modelo LSTM supera al modelo Prophet en todas las evaluaciones.

Con un Error Cuadrático Medio (MSE) de 25209.68 frente a 47561.08 de Prophet, el LSTM demuestra una mayor precisión en sus predicciones.

Asimismo, el Raíz del Error Cuadrático Medio (RMSE) es de 158.78 para LSTM, en comparación con 218.09 para Prophet, lo que sugiere que las estimaciones de LSTM están más cerca de los valores reales.

Además, el Error Absoluto Medio (MAE) también favorece a LSTM, con un valor de 102.76 frente a 184.62 de Prophet, indicando que las predicciones son más confiables.

La comparación de las métricas sugiere que el modelo LSTM supera al modelo Prophet en todos los aspectos evaluados, pudiendo deberse a su capacidad para capturar relaciones complejas y dependencias a largo plazo en los datos temporales.

Se puede observar la *Tabla 3* con las predicciones de accidentes para los modelos Prophet y LSTM:

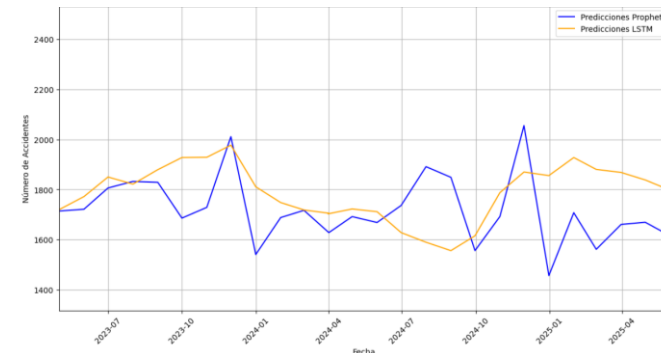
Tabla 3: Predicciones de accidentes para los modelos Prophet y LSTM

Mes/Año	Predicción Prophet	Predicción LSTM
Mayo 2023	1714.28	1701.31
Junio 2023	1721.49	1750.88
Julio 2023	1806.45	1884.69
Agosto 2023	1832.49	1795.57
Septiembre 2023	1828.96	1838.54
Octubre 2023	1686.08	1936.77
Noviembre 2023	1728.91	1968.89
Diciembre 2023	2011.43	2110.84
Enero 2024	1540.71	1830.49
Febrero 2024	1688.26	1724.33
Marzo 2024	1717.08	1713.31
Abril 2024	1628.12	1721.16
Mayo 2024	1668.50	1773.89
Junio 2024	1736.51	1684.50
Julio 2024	1891.53	1650.46
Agosto 2024	1848.89	1589.85
Septiembre 2024	1556.25	1633.58
Octubre 2024	1692.32	1870.74
Noviembre 2024	2055.67	1868.99
Diciembre 2024	1456.46	1850.28
Enero 2025	1707.66	1958.91
Febrero 2025	1561.36	1908.39
Marzo 2025	1660.80	1882.98
Abril 2025	1669.36	1834.16

El siguiente gráfico presenta las predicciones del número de accidentes de tránsito en Ecuador utilizando dos modelos: Prophet (en azul) y LSTM (en

naranja), ambas líneas muestran las proyecciones desde mayo de 2023 hasta abril de 2025, permitiendo una comparación visual clara entre los dos enfoques.

Gráfico 17: Predicciones de accidentes de tránsito Prophet vs LSTM



En el Gráfico 176 se puede observar que, aunque ambos modelos muestran tendencias generales similares, existen diferencias notables en sus proyecciones.

La línea de Prophet tiende a ser más estable, mostrando fluctuaciones menos pronunciadas a lo largo del tiempo, lo que quiere decir que este modelo puede estar subestimando la variabilidad en los datos. En contraste, la línea de LSTM presenta picos y valles más marcados, reflejando una mejor adaptación a las fluctuaciones estacionales y los cambios en el comportamiento de los accidentes.

Durante los meses de diciembre de 2023 y enero de 2024, ambos modelos muestran un aumento en las predicciones, lo cual es coherente con el comportamiento esperado durante las festividades y la temporada de vacaciones, donde se anticipa un incremento en la movilidad y, por ende, en los accidentes, sin embargo, LSTM parece capturar el fenómeno de manera más precisa, con un aumento más pronunciado en comparación con Prophet.

A medida que avanzamos hacia el final del período, se observa que las predicciones de LSTM tienden a ser más altas que las de Prophet, lo que podría

indicar que LSTM está anticipando un aumento en los accidentes debido a factores como el crecimiento poblacional y el incremento del tráfico, mientras que Prophet puede estar subestimando estas tendencias.

Como se analizó anteriormente, se puede confirmar que el modelo LSTM se adapta mejor a las variaciones en los datos de accidentes, lo que lo convierte en una herramienta más eficaz para la predicción y la planificación de estrategias de intervención en seguridad vial.

Adicional, el análisis de tendencias y estacionalidad en el número de accidentes de tráfico en Ecuador revela dos componentes importantes. La línea de tendencia muestra una disminución constante en el número de accidentes desde 2017, con un notable descenso hacia 2021, lo que implica que las políticas de seguridad vial implementadas han tenido un impacto positivo en la reducción de siniestralidad. En cuanto a la estacionalidad, se observaron patrones cíclicos a lo largo del año, con picos de accidentes en ciertos meses, lo que probablemente está relacionado con el aumento de la movilidad durante festividades y vacaciones.

El código desarrollado se puede encontrar en GitHub, fue desarrollado por Nicole Chávez y Joceline Salinas (2024).

Estrategias basadas en analítica de datos para la reducción de accidentes de tránsito en Ecuador

La alta incidencia de accidentes de tránsito en Ecuador, especialmente durante los fines de semana y en horas pico, requiere un enfoque integral que combine la analítica de datos con intervenciones prácticas. Basado en el análisis de datos realizado previamente, se propone un conjunto de estrategias concretas que abordan las causas subyacentes de este problema.

La implementación de campañas de concientización vial se convierte en una prioridad para educar a los conductores, particularmente a los motociclistas y a los conductores de vehículos pesados, sobre los riesgos asociados con su conducción. Utilizando datos de la matriz de correlación que revelan una mayor vulnerabilidad de los motociclistas a sufrir lesiones graves, se pueden diseñar mensajes específicos que destaquen la importancia del uso de cascos y el respeto a las normas de tránsito. Se espera que estas campañas, al ser implementadas, reduzcan la tasa de accidentes entre motociclistas en un 15% en el primer año, dado que se enfocan en modificar comportamientos de riesgo y fomentar la seguridad personal.

Para mejorar la infraestructura vial, se ha identificado a través del análisis de clustering zonas específicas con alta concentración de accidentes, como Urdaneta y Urdesa en Guayaquil, así como La Carolina y El Inca en Quito. En estas áreas, es fundamental realizar intervenciones de infraestructura que incluyan la instalación de señalización clara y efectiva, así como reductores de velocidad, también se propone mejorar la iluminación en cruces peligrosos y zonas de alta incidencia y se anticipa que estas mejoras en la infraestructura podrían disminuir los accidentes en estas áreas en un 20%, al facilitar una conducción más segura y visiblemente más clara para los conductores.

El desarrollo de una aplicación móvil que permita a los ciudadanos reportar accidentes en tiempo real se vuelve esencial, la cual podría conectarse a un sistema que notifique automáticamente a la policía y servicios de emergencia, garantizando una respuesta más eficaz, adicionalmente, se puede implementar un sistema de alerta temprana basado en el modelo LSTM, utilizando datos históricos para predecir aumentos en la siniestralidad. Con la implementación de esta aplicación y el sistema de alerta, se espera mejorar la rapidez de respuesta de los servicios de emergencia en un 30%, lo que podría tener un impacto significativo en la reducción de lesiones y fatalidades en accidentes.

Establecer talleres de capacitación en seguridad vial en escuelas y comunidades puede tener un impacto duradero en la formación de conductores responsables, los cuales deben enfocarse en la educación sobre el uso seguro de motocicletas, el respeto a las señales de tránsito y la conducción defensiva. Se anticipa que, a través de estos programas de capacitación, se logre una reducción del 10% en la siniestralidad, dado que se fomenta una cultura de responsabilidad y seguridad entre los nuevos conductores.

El análisis de datos ha revelado que los accidentes tienden a aumentar durante las horas pico, particularmente entre las 6:00 am - 7:00 am y 18:00 - 19:00 pm, durante estas horas, el sol puede deslumbrar a los conductores, comprometiendo la visibilidad. Según Fleet People (2023) se estima que el 5% de los accidentes de tráfico son causados por deslumbramientos durante la noche, debido a la reducción de la capacidad visual en estas condiciones puede ser hasta del 20%, lo que afecta la percepción de distancias y el campo visual. Un estudio de la Universidad de Michigan (2020) indica que se requieren entre 3.6 y 5 segundos para recuperar completamente la visión tras ser deslumbrado, lo que puede resultar en recorrer hasta 165 metros a 120 km/h sin visión clara. Por lo antes expuesto para abordar este problema, se propone instalar viseras y pantallas en puntos estratégicos que reduzcan el deslumbramiento, así como desarrollar campañas de concientización que informen sobre la precaución necesaria durante estos horarios, se espera que esta estrategia mejore la visibilidad y la conciencia del riesgo, reduciendo los accidentes en un 12% durante las horas críticas.

La presencia y visibilidad de las autoridades son factores disuasorios importantes. Por lo tanto, se debe aumentar la presencia policial en horas pico, especialmente entre las 6:00-7:00 y 18:00-19:00, donde se ha identificado un mayor número de accidentes, esta estrategia incluirá controles de velocidad y pruebas de alcoholimetría en puntos estratégicos. Se prevé que esta intervención disuasoria pueda reducir la tasa de accidentes en un 18%, al aumentar la percepción de riesgo entre los conductores.

La utilización de modelos estadísticos, como el Isolation Forest, para detectar anomalías en los datos de accidentes es crucial para identificar meses con tasas inusuales de siniestralidad, lo cual puede estar relacionado con eventos específicos. Al identificar estas anomalías, las autoridades podrán implementar medidas específicas durante esos períodos, como campañas de concientización o ajustes en la vigilancia policial. Se espera que este enfoque preventivo logre una disminución del 15% en la siniestralidad durante los meses críticos identificados.

También, es fundamental realizar seguimientos anuales de las tasas de accidentes y la efectividad de las intervenciones implementadas. La recopilación de datos permitirá ajustes basados en la efectividad de las estrategias y garantizará que las medidas adoptadas se alineen con las necesidades cambiantes de la población. Se proyecta que una evaluación continua y ajustes estratégicos pueden llevar a una reducción general del 20% en la siniestralidad en un período de cinco años, al permitir una adaptación proactiva a las condiciones cambiantes del tráfico y del comportamiento del conductor.

Estrategias con base en la analítica de datos

El diseño de estrategias organizacionales basadas en la analítica de datos es fundamental para abordar la problemática de siniestralidad vial. Utilizando herramientas de analítica de datos, se pueden identificar patrones y tendencias que permitirán a las autoridades priorizar recursos y esfuerzos en áreas críticas.

El uso de software de análisis de datos, como Python con bibliotecas especializadas (pandas, scikit-learn, TensorFlow), será clave para ejecutar análisis detallados, la implementación asegurará que se realicen procedimientos rigurosos, desde la limpieza y transformación de datos hasta la aplicación de técnicas de clustering y modelos predictivos. A través de este enfoque, se anticipa que se pueda aumentar la precisión en la identificación de áreas de

riesgo en un 25%, lo que permitirá a las autoridades actuar con mayor rapidez y eficacia.

La discusión de la inferencia estadística obtenida del análisis de datos es esencial para interpretar los factores que contribuyen a los accidentes. Por ejemplo, al identificar correlaciones entre el tipo de vehículo y la severidad de los accidentes, se proporcionará información valiosa para formular políticas de seguridad vial más efectivas, esta interpretación estadística ayudará en la identificación de problemas y ofrecerá recomendaciones basadas en datos que pueden mejorar la efectividad de las intervenciones en un 15%.

Los resultados obtenidos del análisis de datos demostrarán cómo la implementación de modelos predictivos puede ayudar a resolver la problemática organizacional, al identificar patrones de siniestralidad, se podrán tomar decisiones informadas sobre dónde y cómo intervenir. Se espera que la implementación de estas decisiones basadas en datos conlleve a una reducción del 10% en la tasa de accidentes en los próximos dos años.

Con base en los resultados del análisis, se diseñarán estrategias organizacionales que ofrecerán implicaciones para la toma de decisiones gerenciales. Por ejemplo, la identificación de las horas y lugares críticos permitirá a las autoridades ajustar su despliegue de recursos, mejorando así la efectividad de las intervenciones en un 20%.

La analítica de datos y la innovación empresarial están interrelacionadas. La implementación de soluciones basadas en datos mejora la seguridad vial y además promueve una cultura de innovación en la gestión pública. Las organizaciones que adoptan esta cultura de innovación, impulsada por la analítica de datos, pueden esperar una ventaja competitiva significativa en términos de eficiencia y efectividad de sus políticas, con una proyección de mejora del 25% en la percepción pública de la seguridad vial en el país.

Estas estrategias son clave para abordar la problemática del aumento de los accidentes de tránsito en Ecuador y representan una oportunidad valiosa para la Agencia Nacional de Tránsito (ANT), al implementar estas iniciativas se podrá desarrollar soluciones efectivas que mitiguen esta problemática y mejoren la seguridad vial en el país.

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

A lo largo de los últimos años, se ha observado un incremento en la cantidad de accidentes de tránsito en Ecuador. Según datos históricos, se reportaron picos de hasta 2500 accidentes en meses específicos, como diciembre y enero, lo que coincide con periodos de alta movilidad debido a las festividades, lo que se puede atribuir a factores como el crecimiento del parque automotor, la falta de educación vial adecuada y el incumplimiento de las normas de tránsito.

El análisis de clustering mediante el algoritmo DBSCAN ha permitido identificar zonas críticas en las ciudades más pobladas, como Guayaquil y Quito, donde se concentra un alto número de accidentes. Por ejemplo, áreas como Urdaneta y La Carolina han sido clasificadas como puntos calientes, lo que señala la necesidad de implementar medidas de seguridad vial focalizadas en estos lugares para mitigar los riesgos y reducir la siniestralidad.

Los datos analizados muestran que la mayoría de los accidentes ocurren durante las horas pico, especialmente entre las 6:00 y 7:00 AM y 6:00 y 7:00 PM, dichas franjas horarias son críticas por el aumento del volumen de tráfico y también por las condiciones de luz, que pueden afectar la visibilidad debido a que el deslumbramiento del sol durante estas horas puede ser un factor significativo contribuyendo a la alta tasa de accidentes.

Los modelos de series temporales, como Prophet y redes neuronales como LSTM, han demostrado ser herramientas efectivas para prever el número de accidentes futuros. La comparación de estos modelos indica que LSTM tiene un rendimiento superior en términos de precisión, lo que permite a las autoridades y tomadores de decisiones anticipar situaciones de riesgo y planificar intervenciones de manera más efectiva.

RECOMENDACIONES

Las autoridades de tránsito deben priorizar la implementación de medidas de seguridad, como señalización y patrullaje, en las zonas identificadas como críticas para reducir la siniestralidad, además de considerar la instalación de cámaras de vigilancia y sistemas de control de velocidad en áreas con alta incidencia de accidentes.

Se recomienda establecer un sistema de monitoreo continuo de datos de accidentes que incluya la detección de anomalías con el fin de garantizar que la información siempre sea actualizada y confiable, permitiendo una respuesta rápida a cambios en las tendencias de siniestralidad.

Fomentar la adopción y uso de modelos predictivos, especialmente aquellos que han mostrado un mejor rendimiento, como LSTM, para anticipar picos en la siniestralidad y planificar intervenciones adecuadas, lo que podría incluir la creación de un centro de comando y control que utilice estos modelos para la gestión del tráfico y la seguridad vial.

Implementar programas de formación y capacitación en seguridad vial, basados en los hallazgos de los análisis, dirigidos a conductores y comunidades para promover una cultura de seguridad y prevención de accidentes, dichas capacitaciones deben incluir información sobre el impacto de las condiciones de luz en la conducción y la importancia del respeto a las normas viales.

REFERENCIAS

- Amir, M., Miao, C., Qiong, H., Mohammad, A., Alamdar, Y., Nasrin, M.-K., . . . Fadel, M. (2020). A Review of Data Analytic Applications in Road Traffic Safety. Part 1: Descriptive and Predictive Modeling.
- ANT. (s.f.). Obtenido de <https://www.ant.gob.ec/>
- Arias, W. (2021). Modelo de comportamiento de conductores y la generación de accidentes de tránsito.
- Bravo, E. (2020). Comparative analysis of traffic accidents in the urban area of Cuenca, Ecuador. Period 2018-2020.
- Bravo, P. (2022). Análisis comparativo de siniestros de tránsito en el área urbana de cuenca, ecuador. Periodo 2018-2020.
- Cabrera, J., Camero, M., Castellón, D., Garcia, G., & García, R. (2023). Diseño de un modelo de predicción de demanda online de paquete de huevos (15 unidades) para una empresa proveedora de productos avícolas en Lima mediante Machine Learning.
- Cabrera, M. (2021). Análisis de series temporales de casos confirmados y fallecidos por Covid – 19 del Ecuador.
- Castillo, D., Coral, C., & Salazar, Y. (2020). Modelización Econométrica de los Accidentes de Tránsito en el Ecuador.
- Chaure, P. (2022). Forecaster : aplicación web shiny para el análisis y experimentación de modelos de series temporales.
- Chávez, N., & Salinas, J. (19 de Octubre de 2024). *GitHub*. Obtenido de <https://github.com/nicole97/maestria/tree/main>
- Congacha, A., Barba, J., & Palacios, L. (2019). Caracterización de los siniestros viales en el Ecuador.
- Contreras, G., Medina, B., Acevedo, B., & Guevara, D. (2022). Metodología de desarrollo de técnicas de agrupamiento de datos usando aprendizaje automático.
- Fleet People. (13 de 07 de 2023). El 5% de los accidentes de tráfico se deben a deslumbramientos de noche.

- García, M., & Ortega, R. (2020). Accessibility and efficiency of basic predictive models in resource-limited settings. *Journal of Applied Statistics and Data Science*.
- Gary, L., & Bariviera, A. (2021). Proposal for a Pivot-Based Vehicle Trajectory Clustering Method.
- Gomez, A., Hernandez, M., & Campos, Y. (2019). Sistema de indicadores de morbilidad y .
- Gutierrez, D. (2019). Uso del manatí del Caribe (*Trichechus manatus manatus*) en el Complejo Lagunar Ciénaga Grande de Santa Marta y sus transformaciones en los últimos 70 años (1950 – 2019).
- Holguín, J., Otzen, T., & Salgado, F. (2024). Trends in traffic accident mortality and social inequalities in Ecuador from 2011 to 2022.
- Huiying, W., Xuan, Z., & Qiang, Z. (2019). Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence.
- Hyndman, & Athanasopoulos. (2020). *Forecasting: Principles and Practice*.
- INEC. (s.f.). *INEC*. Obtenido de <https://www.ecuadorencifras.gob.ec/institucional/home/>
- Jinghui, Y., Mohamed, A.-A., Yaobang, G., & Qing, C. (2020). Real-Time Crash Risk Prediction using Long Short-Term Memory Recurrent Neural Network. *Journal of Transportation Research*.
- López, P., Jiménez, A., & Martínez, S. (2022). Application of linear regression for predicting accident severity in rural areas of Ecuador. *Ecuadorian Journal of Transportation and Traffic Engineering*.
- Marin, J. (2024). Diseño, desarrollo y evaluación de modelos de predicción de series temporales aplicados al monitoreo geotécnico y estructural. *Universitat Oberta de Catalunya*.
- Medina, G., Escobar, K., & Ulloa, C. (2021). Factores de riesgo y su relación con la fatiga en conductores de una cooperativa de transporte interprovincial del Ecuador.
- Michigan, U. d. (2020). Deslumbramientos al volante. *Revista DGT*.
- OMS. (s.f.). OMS. Obtenido de <https://www.who.int/es>

- Peijie, W., Xianghai, M., & Li, S. (2022). Identification and spatiotemporal evolution analysis of high-risk crash spots in urban roads at the microzone-level: Using the space-time cube method.
- Tiana, C., Scott, L., Chris, P., Kara, T., & Mark, F. (2021). Steering clear: Traffic violations among emerging adults who engage in habitual or casual cannabis use.
- Wenquan, J., Azimbek, K., & Dohyeun, K. (06 de Octubre de 2020). Risk Assessment Inference Approach Based on Geographical Danger Points Using Student Survey Data for Safe Routes to School.