UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA EN BIG DATA Y CIENCIA DE DATOS

Resolución: RPC-SO-32-No.536-2023.

PROYECTO DE TITULACIÓN EN OPCIÓN AL GRADO DE MAGISTER

Título del proyecto:
Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias
Línea de Investigación:
Ciencias de la ingeniería aplicadas a la producción, sociedad y desarrollo sustentable
Campo amplio de conocimiento:
Tecnologías de la Información y la Comunicación (TIC)
Autor/a:
Hans Gonzalo Muñoz Zapata
Tutor/a:
PhD. Renato Mauricio Toasa Guachi Mg Mario Ruben Pérez Cargua

Quito – Ecuador

2025

APROBACIÓN DEL TUTOR

Yo, Renato Mauricio Toasa Guachi con C.I: 1804724167 en mi calidad de Tutor del proyecto de investigación titulado: Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias

Elaborado por: Hans Gonzalo Muñoz Zapata, de C.I: 1717227852 estudiante de la Maestría: Big data y Ciencia de Datos de la **UNIVERSIDAD TECNOLÓGICA ISRAEL (UISRAEL)**, como parte de los requisitos sustanciales con fines de obtener el Título de Magister, me permito declarar que luego de haber orientado, analizado y revisado el trabajo de titulación, lo apruebo en todas sus partes.

Quito D.M., septiembre de 2025

Firmado electrónicamente por:
RENATO MAURICIO
TOASA GUACHI
Validar únicamente con FirmaEC

_____

Firma

APROBACIÓN DEL TUTOR

Yo, Mario Ruben Pérez Cargua con C.I: 0603251984 en mi calidad de Tutor del proyecto de investigación titulado: Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias

Elaborado por: Hans Gonzalo Muñoz Zapata, de C.I: 1717227852 estudiante de la Maestría: Big data y Ciencia de Datos de la **UNIVERSIDAD TECNOLÓGICA ISRAEL (UISRAEL)**, como parte de los requisitos sustanciales con fines de obtener el Título de Magister, me permito declarar que luego de haber orientado, analizado y revisado el trabajo de titulación, lo apruebo en todas sus partes.

Quito D.M., septiembre de 2025

Firmado electrónicamente por:
MARIO RUBEN PEREZ CARGUA
Validar únicamente con FirmaEC

_____

Firma

DECLARACIÓN DE AUTORIZACIÓN POR PARTE DEL ESTUDIANTE

Yo, Hans Gonzalo Muñoz Zapata con C.I: 1717227852, autor/a del proyecto de titulación denominado: Técnicas de clasificación para identificar patrones de comportamiento y determinar la tipología delito que tiene más fuerza en México vs Ecuador Previo a la obtención del título de Magister en Big Data.

1. Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar el respectivo trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.
2. Manifiesto mi voluntad de ceder a la Universidad Tecnológica Israel los derechos patrimoniales consagrados en la Ley de Propiedad Intelectual del Ecuador, artículos 4, 5 y 6, en calidad de autor@ del trabajo de titulación, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente. En concordancia suscribo este documento en el momento que hago entrega del trabajo final en formato impreso y digital como parte del acervo bibliográfico de la Universidad Tecnológica Israel.
3. Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de prosperidad intelectual vigentes.

Quito D.M., septiembre de 2025

HANS GONZALO MUÑOZ ZAPATA
Firmado digitalmente por HANS GONZALO MUÑOZ ZAPATA
Fecha: 2025.09.16 05:03:30 -05'00'

Firma

Tabla de contenidos

Índice de tablas

Índice de figuras

INFORMACIÓN GENERAL

Contextualización del tema

Considerando los acontecimientos actuales, en America Latina según Gómez Albarello y Corzo Salamanca (2021) explica Una confusión entre conceptos, por ende, no se marca una diferencia entre la violencia y criminalidad, pero estos se podría dar por diferentes factores: economía en crecimiento, calidad de vida, estatus del Estado de Derecho, equidad, el microtráfico, la disponibilidad de armas de fuego entre otros.

Considerando lo anterior se podría afirmar que la criminalidad y la violencia causas o factores similares, que podrían ser el índice de un homicidio, es por ello que se buscará la implementación de herramientas tecnológicas para determinar patrones de comportamiento delictivo.

Granados Garzón,L (2024) menciona que el patrón delincuencial Series de Delitos (Series), es:" Conjunto de delitos análogos que se piensa son cometidos por el mismo individuo o grupo de individuos que actúan organizadamente", es decir, cuando el o los delincuentes actúan de una manera organizada sobre un mismo lugar determinado se podría entender como un patrón de delito en este caso series.

Pineda, A.L (2022) argumenta que la violencia tiene efectos nocivos para la sociedad y las personas. Esto es especialmente cierto en América Latina, una región que se destaca a nivel mundial por su alta tasa de homicidios. Pero permite denotar como la sociedad se ha ido denigrando, además se destaca un gran porcentaje de delitos en Latinoamérica, es por ello que identificar este tipo de delitos es una promesa práctica para determinar ciertos tipos y patrones de comportamiento, lo que llevará en un futuro posiblemente a desarrollar estrategias aplicables a la ley que se basan en evidencia y en estudios previos.

En un Informe a nivel mundial sobre el estupefaciente llamado cocaína de la UNODC, de 2023, señala que desde 2021 Brasil se encuentra entre países de salida de la cocaína, sólo destaca por Ecuador (Malamud y Núñez ,2024). Ecuador ha llegado a estar entre los países que en la última década se ha posicionado entre los primeros lugares para el tráfico de sustancias sujetas a fiscalización, pero esto es una realidad que es uno de los factores que causa los homicidios.

Según García (2021) explica que un algoritmo se establecería como un proceso de sucesiones lógicas de manera necesaria, tiene el fin de efectuar una actividad determinada, como la solución de problemas en diversos ámbitos o clasificar elementos. Lo que se busca es mediante ayuda de herramientas tecnológicas se realizará un análisis de zonas de riego y tipología de crímenes con la visión de formular posibles políticas públicas y reducir el índice de delitos en zonas determinadas buscando factores de patrones de comportamiento.

Problema de investigación

El verdadero desafío de esta investigación no se basa en escoger un algoritmo como lo es Random Forest, sino más bien resaltar las limitaciones intrínsecas de los datos delictivos como lo son la naturaleza dinámica, irregularidad asociada a series temporales y datos con ruido.

Se debe considerar que las limitaciones afectan a la capacidad del modelo para aprender patrones confiables, esto tiene relación directa con la construcción de un algoritmo de predicción Random Forest, se deberá enfrentar algunos retos como: Evaluación de desempeño, comparación entre contextos nacionales e internacionales, procesamiento inteligente, entre otros factores.

Para realizar un análisis del patrón de comportamiento series mediante un dashboard comparativo entre Ecuador y México se deberá entrenar al modelo de datos adecuadamente, determinar la tipología más frecuente y sus contextos característicos en México con relación a Ecuador.

¿Cómo analizar los datos de delitos en Latinoamérica mediante el uso de algoritmos de predicción, con el fin de identificar patrones de comportamiento y determinar los tipos de delitos más frecuentes en Ecuador en relación con México ?

Objetivo general

Realizar un análisis Predictivo sobre el patrón Series de delitos en México y Ecuador mediante el Algoritmo Random Forest para obtener una comparativa entre países.

Objetivos específicos

- Contextualizar los fundamentos teóricos relacionados con el análisis de datos criminales, las técnicas de clasificación supervisada y los enfoques criminológicos orientados a la comprensión del comportamiento homicida en contextos latinoamericanos.

- Diagnosticar la situación actual de los delitos en México y Ecuador mediante la recopilación, limpieza y análisis exploratorio de datos oficiales, y determinar las variables más relevantes para su estudio.
- Desarrollar un dashboard basado en clasificación que permita identificar patrones de comportamiento vinculados a los tipos de delitos más frecuentes registrados en México vs Ecuador.
- Validar cuán efectivo es el dashboard construido mediante la aprobación de especialistas en analistas de datos.

Vinculación con la sociedad y beneficiarios directos:

Se llevó a cabo un Webinar en el Banco General Rumiñahui mediante Teams de manera remota, orientado a arquitectos de datos, desarrolladores señor y desarrolladores estándar, entre otros con el fin de poner en contexto conceptos clave de Big Data, ciencia de datos, algoritmos de datos, en específico se presentó una explicación de los algoritmos de predicción y cómo estos podrían mejorar la calidad de vida, además se debería contextualizar en cuales son las posibles áreas en las que se podría implementar en el área financiera así como consta en el Anexo 1.

Además, considerará que la presente investigación tiene relevancia en el ámbito social ya que está directamente relacionado con el diario vivir, así como en lo tecnológico, responde a la necesidad urgente de comprender con mayor profundidad el fenómeno de los delitos en Latinoamérica mediante el uso de técnicas analíticas avanzadas. En un contexto donde la violencia letal constituye una de las principales preocupaciones de seguridad ciudadana, el aprovechamiento de los datos disponibles a través de algoritmos de clasificación representa una posible estrategia innovadora en base a evidencia para respaldar la formulación de decisiones enfocadas en la prevención de los homicidios.

Desde una visión más práctica, esta investigación tiene como fin el uso de un proceso analítico, el cual ayudará a identificar patrones de comportamiento y determinar los tipos de delitos más frecuentes, que a posterior beneficiará al establecimiento de políticas y tomar acciones a favor de la población Latinoamericana, el uso del dashboard como ayuda visual de clasificación beneficiará a la observación temprana de datos permitiendo la utilización de estos para su estudio.

CAPÍTULO I: DESCRIPCIÓN DEL PROYECTO

1.1. Contextualización general del estado del arte

El Derecho define legalmente los delitos como aquellos patrones de comportamiento que están establecidos en el Código penal indicando que: son delitos o faltas las acciones y omisiones dolosas o imprudentes penadas por la Ley (García, 2021), en base a la definición anterior se podría considerar que los delitos son una forma de incumplir las normativas o reglas impuestas por la sociedad bajo un contexto, adicional debe ser denunciado.

Fernández Pérez (2022), entre algunos de los delitos se encuentran: Robo a personas: acción que podría darse en la vía pública, es lo más comúnmente reportado o denunciado, robo a domicilios: ocurre cuando los propietarios no se encuentran en el lugar o son sorprendidos por los delincuentes, robo de vehículos y motocicletas: sigue siendo un problema que se ha visto en crecimiento en los últimos años, además de la proliferación de los asaltos armados: muchas veces cometidos por grupos organizados o GDO's, como se autodenominan.

Considerando según Martínez Pérez (2024), los Grupos Armados Organizados (GAO) y los Grupos Delincuenciales Organizados (GDO), son los nuevos actores criminales formados en la evolución del accionar del estado, generando fenómenos como el terrorismo, el tráfico de estupefacientes, el tráfico de personas y el tráfico de material nuclear, estableciendo así implicaciones a la seguridad estatal y provocando un cambio paradigmático en la seguridad nacional, es decir, estos grupos se han ido desarrollando enmarcados en actividades delincuenciales además existen más delitos relacionados con los GDO´S como: la estafa: engaño con fines económicos mediante tecnología o artículos digitales, amenazas: intimidación a las personas de forma escrita o verbal, falsedad documental: uso o creación de documentos falsos o impresión de documentos estatales falsificados, cohecho y corrupción: delitos cometidos por funcionarios públicos, que son en ocasiones amenazados por los grupos para que se tomen decisiones, pero el mayor delito que se atribuyen estas bandas son los homicidio y asesinato.

Considerando que el estado ha realizado cambios en búsqueda de una mejora en la calidad de vida, proponiendo una mayor fluidez, en especial al momento que se realiza una denuncia, debido a que al parecer continúa siendo un problema que aún atormenta a Ecuador, Ecuavisa (2024), menciona que las víctimas de delitos viven un suplicio al presentar las denuncias en la Fiscalía y Policía, esto no sólo generará desconfianza además podrá ser aprovechado por los grupos organizados delincuenciales para consolidar su poder buscando actuar con

premeditación, alevosía, se podrían asociar a delitos cada vez más inimaginables que la sociedad no necesariamente conocía pero que se podrían ver en las noticias del día a día.

De acuerdo a Hidalgo Guayaquil (2024), considera que el terrorismo en el Ecuador se podría establecer a una persona o grupo de personas los cuales realizan acciones con premeditación, posiblemente, provocando un estado de terror a un segmento o conjunto de la población, con lo que buscan conseguir una provocación de miedo, o en ocaciones se encuentre en peligro la vida o derechos que se encuentran normados, no solo considerando daños a la propiedad sino también daños emocionales o afectivos, esto quiere decir el terrorismo no solo determinado para un conjunto de personas o requiere que de estatus social, más bien habla acerca de las acciones que generen terror a una persona o conjunto de personas donde mediante actos que ponen en riesgo la vida o los derechos que son protegidos por la ley, estas actividades o conductos son sancionadas de acuerdo a derecho y lo que se busca es una convivencia pacífica entre las personas.

La falta de acciones concretas y políticas públicas para la prevención investigación, sanción de las personas que participan en este tipo de actos en Ecuador refleja una problemáticas mucho más amplia, donde inclusive se ve afectada las seguridad social, además que en los últimos años se ha visto una problemática en el ámbito penitenciario, allí se denota una falta de control, se ha visto manifestada por una serie de fugas, amotinamientos, linchamientos entre otros, donde no solamente se busca únicamente la violencia, sino tiene un propósito de infundir terror, en la población interna así como la sociedad en generral Hidalgo Guayaquil (2024), lo que implica que la sociedad viva en un estado de zozobra debido a este alto índice de crecimiento delictivo que no solo estaría limitado a realizar daños a propiedades o materiales sino también a las afectaciones emocionales y psicológicas. Aunque existen sanciones para este tipo de actos la ausencia de políticas públicas dirigidas a un segmento de la población específico que rompe la convivencia tranquila de acuerdo a lo descrito por el autor no son suficientes en un contexto nacional, es decir, En últimos años, se ha proliferado la delincuencia es en el sector penitenciario donde se describe un crecimiento en el nivel de violencia.

En 2022, Barroso y Soria desarrollaron una investigación llamada “La perfilación criminal como herramienta forense en la investigación de delitos contra la vida. Ecuador. Sociedad & Tecnología”, donde se tenía como propósito el análisis de perfilación e identificación criminal cuyo objetivo es la localización de las personas de cometen delitos graves, donde se busca establecer patrones psicológicos utilizado patrones comportamentales que guían a acciones que cometerá el perpetrador mediante un análisis de perfil geográfico, estableciendo un

modus operando bien llamado firma criminal, además la investigación advierte que En el Ecuador el sistema de justicia penal carece de mecanismos para la identificación de perfiles delictivos en base a comportamientos delictivos o criminales, lo mismo que limita la respuesta antes esta problemática.

En 2024, Ecuavisa mencionó que"*ciudades como Guayaquil, Durán, Esmeraldas, y en general en la costa, los hechos de violencia extrema (extorsiones, vacunas, asesinatos, etc) son frecuentes, hasta el punto de que ya se han convertido en noticia común*", en consecuencia al observar cada día las noticias se denota que el índice de violencia crecer a una escala nacional a tal punto que se ha vuelto un problemas de seguridad nacional, donde el gobierno ha establecido ciertos parámetros para calificar a los grupos denominados GDO como terroristas con la finalidad de endurecer las medidas legales, buscando frenar en algo el avance de los delitos y la violencia además que merme en algo la situación. La normalización del problema solo causará un freno en el avance de las políticas esto debido a que no solo se requiere de acciones represivas sino también de generar acciones de prevención que aborden las temáticas estructurales del problemas como lo es el desempleo y la falta de oportunidades. Se menciona que las ciudades de la costa tienen hechos de violencia extrema, esto podría ser debido a que los grupos delictivos, buscan estar cerca del puerto donde se generan actividades económicas de comercio y transporte fluvial.

Por otro lado, la relación que existe entre factores socioeconómicos y la delincuencia ha sido explorada con anterioridad por Guzmán Sánchez (2024) menciona que: Existe una correlación significativa entre el desempleo y los índices delictivos. Lo que equivale a decir que, las organizaciones delincuenciales encuentran un lugar ideal para ejercer control, violencia aprovechando la falta de recursos económicos, el desempleo prolongado, la falta de recursos. Esto causa un círculo vicioso de inseguridad, la violencia y delitos. Lo que dificulta la prolongación de políticas públicas, amenazando la estabilidad financiera y emocional de las personas.

Esto quiere decir, aunque no se podría afirmar causalidad directa, sí se evidencia que existe una conexión entre a violencia y la economía, la misma que actúa como un incentivo para las mafias donde buscan reclutar personas por el miedo y de bajos recursos económicos.

1.2. Proceso investigativo metodológico

1.2.1 Enfoque investigación

En la presente investigación se guía por un enfoque metodológico cualitativo para lograr una comprensión del problema de manera integral, existe predominancia en la fase inicial:

Enfoque cualitativo:

Se iniciará de la observación directa la misma que busca contextualizar el origen, la calidad de los datos de fuentes oficiales, diccionarios de datos, análisis crítico y perspectiva del impacto social de los datos.

1.2.2 Tipo de Investigación

La investigación es de tipo aplicada, porque busca resolver un problema concreto relacionado con la seguridad pública, mediante el uso de herramientas tecnológicas. Tiene componente descriptivo debido a su tipo:

Describe y caracteriza el estado actual de los datos así como las variables que influyen en la ocurrencia de delitos, así como los patrones comportamentales relacionados.

1.2.3 Población y muestra

Población: Está conformada por los registros oficiales de delitos de Ecuador y México durante un período histórico significativo (2014-2025) Esto incluye algunos metadatos variables como: delito, geografía y contexto. Además de toda la documentación que describirá cómo se obtuvieron los datos.

Muestra: Representa los datos históricos en un lapso de tiempo además de fuentes secundarias como manuales, diccionarios de datos(al menos 2 millones de datos).

1.2.4 Métodos, técnicas e instrumentos

Método cualitativo: en la investigación se basa en la observación directa y crítica de registros oficiales. Además de una clasificación perspectiva (positiva, negativa, neutra) sobre la utilidad del análisis.

Técnicas e instrumentos

Técnicas

- Observación directa de fuentes principales y secundarias
- Clasificación de la percepción en un gráfico tipo pastel.

Los instrumentos son:

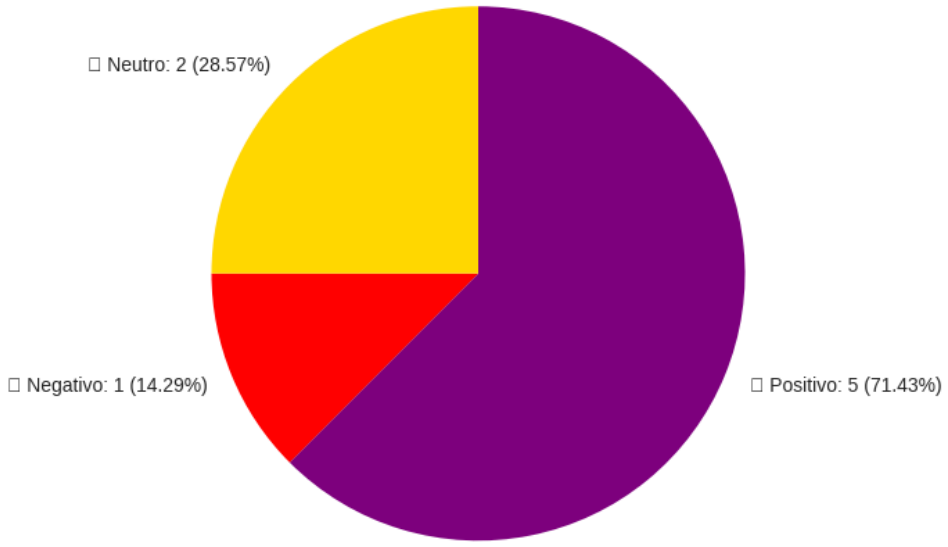
- Sistema de categorización (Positivo, Negativo, Neutro)

1.3. Análisis de resultados

A continuación se analizará el enfoque cualitativo que se llevó a cabo mediante observación directa mediante la síntesis del anexo 5, además se revisará un análisis de acuerdo con la Figura 3 ya que muestra un gráfico de tipología pastel, donde se representa la observación directa clasificada en 3 categorías que son: positivo, negativo y neutro. Las proporciones fueron estimadas de acuerdo a lo revisado con anterioridad.

Figura 3

Distribución de Observación Directa



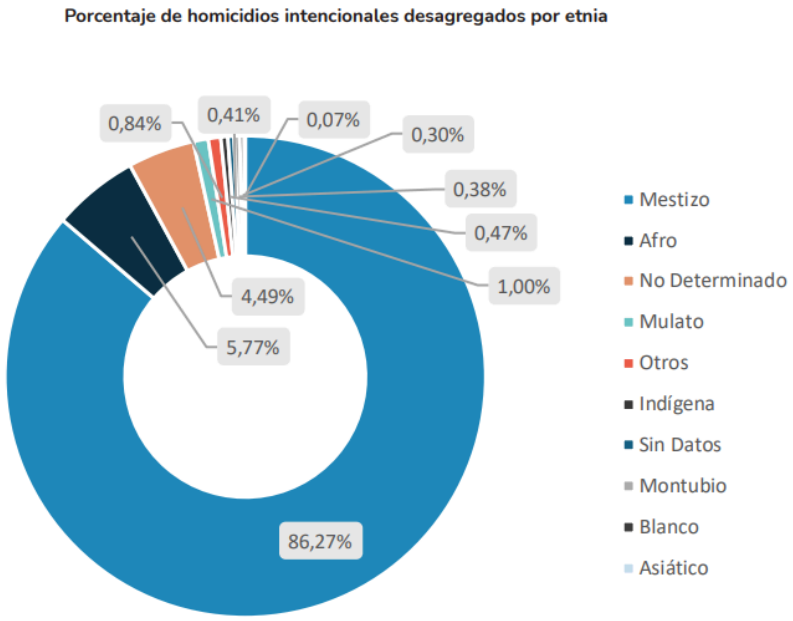
🟣 Positivo (71.43%): Se destaca el aprendizaje, la utilidad del análisis exploratorio, la construcción del dashboard y el valor aportado al análisis criminal.

🔴 Negativo (14.29%): Se mencionan errores cometidos, datos mal estructurados y la necesidad de ajustes en los modelos.

🟡 Neutro (28.57%): Incluye observaciones técnicas, como la revisión manual de registros y la descripción del método deductivo sin juicio de valor.

Figura 4

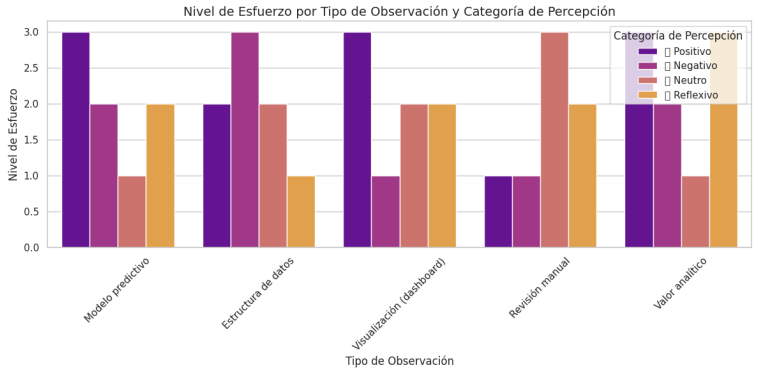
Porcentaje de homicidios desagregados por etnia



De acuerdo Rivera (2025) en su analista estadístico planteado acerca del porcentaje de homicidios intencionales desagregados por etnia(figura 4), mediante la observación se denotará el crecimiento a nivel de violencia en este caso de homicidios a personas mestizas teniendo un total de 26.87 en el 2025, esto no solo a nivel de un segmento de la sociedad, se da en diferentes variables pero de acuerdo a la observación directa se podría afirmar que es el segmento más grande de la población, además de ser uno de los posiblemente afectados.

Figura 5

Nivel de esfuerzo por tipo de observación y categoría de percepción



En el gráfico anterior tiene como finalidad realizar una representación gráfica del nivel de esfuerzo en varios casos ejemplificar donde se podría concentrar la percepción positiva, es decir, que conlleva más trabajo, pero se reconoce una mayor satisfacción, al contrario, en la revisión manual, aunque requiere mayor nivel de esfuerzo la percepción de esfuerzo es negativa o neutra ya que es una tarea pesada poco valorada.

Se podría interpretar que no todas las tareas son igualmente valoradas, es decir, hay actividades repetitivas que tienden a ser rechazadas mientras hay otras que generan más impacto.

CAPÍTULO II: PROPUESTA

2.1. Fundamentos teóricos aplicados

1.1.1 Big data.

Uribe (2024). Conjunto de datos de gran volumen, variedad y velocidad de producción, no podrían ser analizados utilizando procesos o herramientas tradicionales. Es decir que el Big Data es el análisis de grandes cantidades de datos, por ejemplo, los datos que producen las redes sociales son millones. Considerando lo anterior no se podría realizar un análisis utilizando herramientas normales buscando la velocidad, como la variedad de datos ya que al ser de diferentes tipos o estructuras de datos, lo que se buscará es mantener la veracidad entre otras características.

Además, se debe considerar que la terminología de Big Big data hace referencia a la gestión de grandes conjuntos de datos que son complejos de procesar con las herramientas que frecuentemente utiliza. López (2024). Lo que no solo se refiere al procesamiento sino a todo el marco de gestión, para lo cual se requiere herramientas avanzadas que sean capaces de analizar, almacenar y extraer valor de un conjunto de datos.

1.1.2 Ciencia de datos.

Padilla y Contreras (2024). Estudio científico de la creación, validación y transformación de datos para crear significado. Es decir, la ciencia permitirá obtener conocimiento y valor de los datos.

1.1.3 Google Colab.

Es una aplicación Web que permite tener un entorno donde ejecutar algoritmos, permite también la posibilidad de modificar estos algoritmos y crearlos de forma colaborativa a través de la organización de diferentes grupos de trabajo. Amiconi, D. F. (2023), es decir, es una aplicación Web que no solo permitirá ejecutar tareas,sino también podría trabajar de forma colaborativa y procesar grandes volúmenes de datos.

1.1.4 Delitos.

Se considera como un hecho, acto, conducta, comportamiento humano inadecuado y en Ecuador el querer o pertenecer a un grupo delictivo denominado GDO, también podrían considerarse como terroristas. Arreaga, K. M. P. (2021), esto quiere decir que un delito está ligado netamente al comportamiento humano, además se denota que en ecuador es posible la

existencia de los llamados Grupos Delictivos Organizados que atemorizan a las personas, lo cual también podría ser considerado como un un delito.

1.1.5 Random Forest.

Se trata de una estrategia de aprendizaje controlado que produce diversos árboles de decisión a partir de un grupo de información para entrenamiento. Los resultados son logrados debido a la fusión para crear un modelo único, donde se contrastan los resultados individuales de cada árbol Vilema Lara (2022), lo que quiere decir, que se centra en ser un algoritmo de predicción basado en árboles.

1.1.6 Python.

En 1989 fué cuando un joven holandés de 24 años comenzó sus primeras implementaciones del lenguaje de programación que se conoce como Python. Pinargote-Zambrano (2024). Es decir, que Python es un lenguaje de programación que tiene cerca de 34 años muy robusto que se utiliza para Big Data y análisis de datos por sus diferentes librerías.

Python es un lenguaje que posee una sintaxis intuitiva y sencilla que resulta fácil de utilizar, esto inclusive para personas que recién están iniciando. Resulta atractivo para realizar análisis de datos para personas que no necesariamente tengan una formación académica sólida en programación Holguín-Londoño (2024), es decir , este lenguaje ha ganado popularidad debido a su versatilidad y simplicidad al momento de utilizar la sintaxis además que permite a las personas sin experiencia que interactúan y sea factible aprenderlo y aplicarlo en análisis de datos.

1.1.7 CBPR.

Con base en el texto de García Aniño et al. (2023)tiene como significado Investigación Participativa Basada en la Comunidad no solo es framework para realizar investigación, esto debido a que predice resultados, es decir, que también anticipa o guía hacia resultados exitosos debido a su diseño colaborativo.

1.1.8 CMMI (Capability Maturity Model Integration).

Con base en el texto de Gökalp y Martinez (2021) tiene como significado Modelo de integración de Madurez de Capacidades es decir, es un modelo que mide en base a las capacidades, estima un valor numérico.

El Capability Maturity Model Integration (CMMI), fue creado por Software Engineering Institute (SEI), donde se ha establecido como un marco de trabajo reconocido por la evaluación y mejora de procesos de software. Es un modelo, que su aplicabilidad se orienta a sectores ores industriales y empresariales, donde se tiende a evaluar no solo la madurez de los procesos sino también una escala progresiva Cedeño Delgado y Rodríguez Véliz (2025), Es decir que el CMMI no solo destaca por su flexibilidad sino también por la adaptación, aunque no es el único modelo que realiza este tipo de métricas, no posee muchas limitaciones en su capacidad de adaptación en algunos contextos de acuerdo a lo que se requiera.

1.1.9 Formatos de Archivos

Según Holguín-Londoño (2024), tanto en ciencia de datos como en el análisis de datos, es normal encontrarse con tipos de archivos distintos, donde cada uno posee sus propias características y usos. Es decir, que en la búsqueda y extracción de datos se podrá encontrar la información en diferentes tipos de archivos sin que estos se encuentren bajo un estándar.

1.1.10 Modelo de FURPS

Según Caraveo, Payró, Y De los Santos Torres, G. (2025) establece que: FURPS, fue creada por Hewlett-Packard en el año 1987, establece algunos factores de calidad que deben cumplir las herramientas o softwares, los cuales son: Funcionalidad (Functionality), Usabilidad (Usability), Confiabilidad (Reliability), Desempeño (Performance) y Soportabilidad (Supportability) o Capacidad de Soporte, en otras palabras quiere decir que es un conjunto de factores que proporcionan una serie de métricas las cuales se evalúan las herramientas, estableciendo tanto características específicas, así como tiempo de respuestas, la eficiencia, la confiabilidad entre otros.

1.1.11 McCall

Es un modelo basado en Estrategias de pruebas para software donde se permite una mejora en calidad de proyectos de software o algoritmos especialmente en su desarrollo bajo ciertas métricas Zacarías (2025). En otras palabras, McCall es como un faro que sirve de guía para medir ciertas métricas en el software donde se buscará no solo funcione bien sino que sea útil y confiable para quienes lo usan, permite con ello proponer una serie de factores que generan métricas con lo que ayudará a una posible mejora en el desempeño de los sistemas informáticos, algoritmos en búsqueda de una perspectiva más integral acercada a la realidad orientada a una evaluación continua de mejora.

2.2. Descripción de la propuesta

Es importante considerar que el CPBR en esta investigación se plantea como un marco de trabajo dentro de la metodología para el proceso investigativo. Tiene como finalidad la implementación mediante las fases metodológicas que son 4 según López Rios, Merino y Marulan (2021): diagnóstico participativo, formulación de microproyectos, etapa de ejecución de los microproyectos, etapa de evaluación de proceso. Es decir, de acuerdo con cada fase se establecerá cada una de ellas y cómo se incorporará de acuerdo al proyecto.

Diagnóstico Participativo

En la fase se identifica la necesidad de esta investigación donde busca contextualizar indicadores técnicos bajo los principios de CMMI que establecen madurez de modelos de acuerdo con Gökalp y Martinez (2021).

Formulación de microproyectos

Se deberá denotar indicadores que conformarán las métricas en CMMI esto busca que el proceso de creación de los modelos tengan madurez así como capacidades técnicas e inclusive se alineará con las necesidades de la investigación. Para determinar los indicadores se ha utilizado se basen en el Modelo de FURPS que según Villafranca, Sánchez, Fernández-Medina y Piattini(2017): Se encuentra basado en McCall, además considera atributos como funcionalidad, usabilidad, confiabilidad, desempeño y capacidad de soporte (Functionality, Usability, Reliability, Performance, Supportability). En el caso de la investigación se establecen 3 atributos(Confiabilidad, Desempeño, Usabilidad), la elección se llevó a cabo debido a los datos que se requieren cuantificar son de gran medida.

Etapa de ejecución de los microproyectos

Al indicar las ponderaciones que se ha establecido en base al contexto, así como lo representación la tabla 2 que representa un cuadro de mando integral, alineados a un valor numérico de cada característica donde el valor máximo (5, Adapta totalmente) y en mínimo (1, No incorpora), mejorará el rendimiento de los modelos al ejecutarse.

Tabla 2.
Escala numérica

Valoración	Descripción Valoración M
1	No incorpora
2	Incorpora superficialmente
3	Incorpora parcialmente
4	Integra consistentemente
5	able totalmente

La investigación Participativa Basada en la Comunidad (CBPR) se tomó 3 principios, para realizar la ponderación de los algoritmos, transparencia e Interoperabilidad, robustez con datos reales, validación participativa, teniendo en cuenta que estos son aplicados por el contexto así como lo considera Chen al. (2021), los algoritmos deben ser utilizados para funciones prácticas, es decir, no solo deben ser teóricos.

Además, se buscó realizar breve análisis de los resultados obtenidos de la Tabla 2 que posiciona al Algoritmo Random Forest como el más adecuado, para ello se han validado algunas características como:

Tabla 3.
Ponderación de algoritmos de aprendizaje

Principio CBPR	Random Forest	Árbol de Decisión	Regresión Logística	SVM	Red Neuronal
Confiabilidad	4	4	5	2	1
Desempeño	5	2	3	3	2
Usabilidad	5	3	4	2	1

Confiablidad

La Confiabilidad no solo será una propiedad, sino una característica que facilita la interacción humana. Vilema Lara (2022), considerando que existen datos con ruido, busca mostrar la veracidad de los mismos. Es necesario acotar que existen personas que no poseen gran experticia en el campo propuesto, para ello, es factible visualizar coeficientes en búsqueda de la información y cómo interpreta la misma.

Desempeño

Mediante la utilización de CBPR se tiene como objetivo corregir errores de muestreo y recolección, los algoritmos de aprendizaje poseen robustez, además los modelos variados poseen características a las cuales deben adaptarse, buscando estar lo más cerca de la validez de los mismos.

Usabilidad

Este enfoque asegura que los datos puedan ser validados, socialmente útiles para y por los usuarios, además no se pretenderá realizar un proceso aislado, sino una interpretación que colabore con la comunidad.

Evaluación de proceso

La técnica de aprendizaje automático acuerdo a los resultados de la tabla 2 es **Random Forest** para tareas de regresión en el proyecto y permitió modelar la relación entre el año y el número de ocurrencias para realizar predicciones futuras, considerando que es una técnica de análisis la cual será aplicada a un conjunto de datos.

De acuerdo con una investigación realizada acerca delitos se establece que Random Forest es un modelo predictivo con menor error absoluto, Según Vera Alarcón (2024) considerando tres posibles algoritmos. La Tabla 3 representa los valores de resultado de la evaluación ejecutada a 3 modelos. El modelo Random Forest tiene un rendimiento más preciso, tiene el menor error absoluto medio (MAE) de aproximadamente 312,57 con una desviación estándar de 126,23. En promedio, las predicciones del modelo Random Forest difieren en alrededor de 312,57 unidades de la cantidad real de delitos, con una variabilidad de aproximadamente 126,23 unidades.

Tabla 3.
Algoritmo, Error absoluto medio, Desviación estándar

Algoritmo	Error Absoluto Medio (MAE)	Desviación Estándar
Random Forest	312.57	126.23
Linear Regression	3223.43	330.06
Gradient Boosting	424.60	102.71

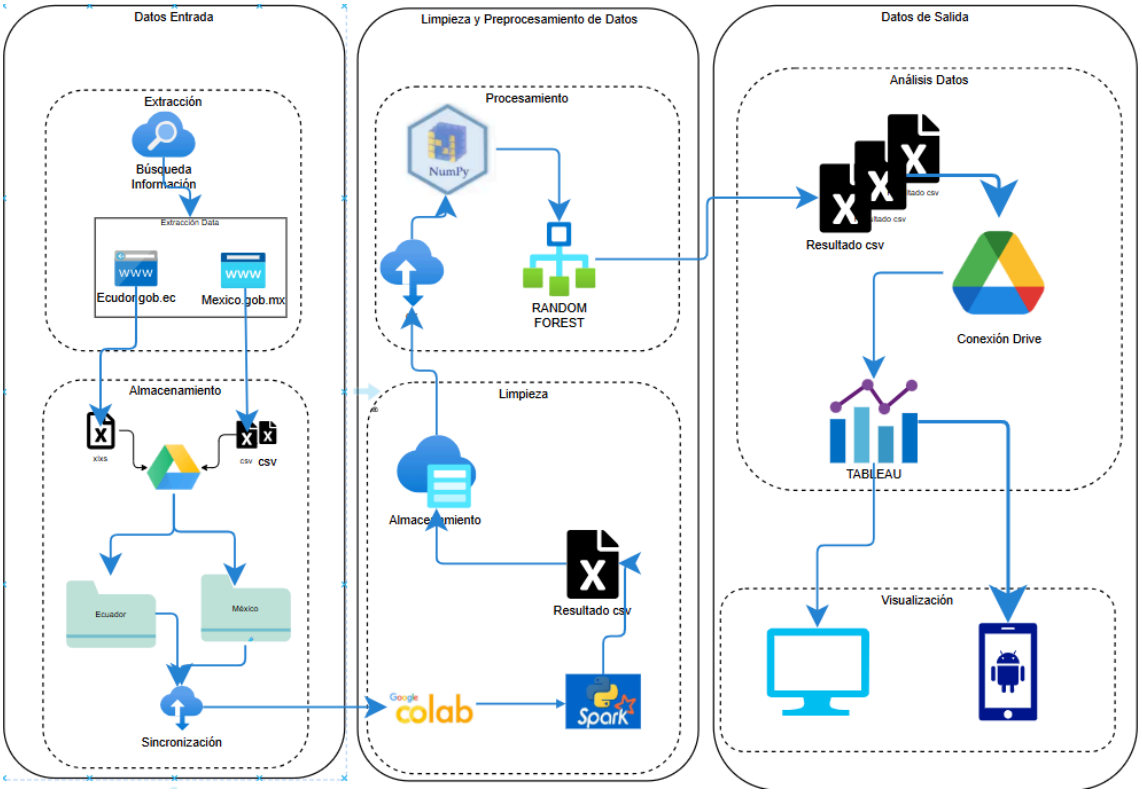
Se planteará utilizar herramientas como Google Colab para procesar líneas de programación en lenguaje Python, buscando que el desarrollo agregue, procese datos, utilizado la nube con

notebooks y bibliotecas especializadas en ciencia de datos así como la implementación de librerías como es Pyspark, para la manipulación y limpieza de datos, scikit-learn para la implementación del modelo Random Forest y otras utilidades de machine learning. Además, se utilizará matplotlib para una visualización previa de los resultados se verán ya establecidos en un tablero Tableau realizando la comparativa como en el gráfico del Anexo 2 donde se encuentra la comparativa acerca de los patrones de comportamiento, se espera que exista una prolongación y una leve disminución del patrón comportamental tipo series. Los datos CSV fueron almacenados y accedidos desde Google Drive, aprovechando el almacenamiento en la nube una vez que los DataSet estaban consolidados, lo que busca no tener reprocesos.

a. Estructura general

Figura 6

Estructura General



b. Explicación del aporte

Considerando que este proyecto busca realizar un análisis predictivo, el principal aporte de la propuesta es la capacidad de generar proyecciones cuantitativas, esto de acuerdo con la posible cantidad futura de contravenciones y de patrones delictivos específicos ('Lugar caliente', 'Otro', 'Serie'). Esta información, derivada del modelado con Random Forest y el

análisis de datos históricos previamente cargados, podría ser un insumo valioso para la toma de decisiones en materia de seguridad. Permite a las autoridades y planificadores anticipar tendencias delictivas, patrones específicos. Por otro lado, se mide la eficiencia del algoritmo en base a la proyección de datos comportamentales. Esto para otros posibles usos en diferentes campos de aplicabilidad.

c. Estrategias y/o técnicas

1. Datos de Entrada (extracción):

* Funcionamiento: Los Datos fueron cargados a un repositorio de Google Drive, donde se almacenaban diferenciados por carpetas. En estos datos se encuentran históricos de años pasados con un sin número de variables tanto para Ecuador como para México. Lo que se realizará mediante el uso del año como variable independiente y el total de delitos (conteo de cada patrón delictivo por año) como variable independiente. Pero para ello se debe normalizar los datos y en la primera fase se realizará una unificación de todos los registros y generar un solo archivo donde contenga los datos que se desea recabar como año, patrón delictivo, entre los más importantes.

* Empleo: Estos datos se usarán para "entrenar" el modelo, lo que permitirá que el modelo aprenda la relación entre el año y el número de delitos/patrones, los datos serán consolidados en un solo Dataset, para luego estar en un archivo como se ve en el Anexo 2, se observará el número de archivos que se tiene por país y la organización que se ha dado para continuar con el proceso.

2. Limpieza y Preprocesamiento de Datos:

* Funcionamiento: en esta segunda fase se ve la ejecución de la información para ello se usará Google Colab junto con algunas de sus librerías que permita organizar los datos de acuerdo a la necesidad y realizada una limpieza, en esta herramienta también se encargará de ejecutar el algoritmo que se usa es Random Forest que sirve para el aprendizaje automático porque construye múltiples árboles de decisión durante su entrenamiento lo que realiza es que genera una predicción en base a un promedio de predicción de los árboles individuales. Esto es lo que hace que Randon Forest sea robusto en un sobreajuste además de capturar relaciones no lineales en datos que fue entrenado.

* Empleo: se utiliza Google Colab y se entrena una instancia con el modelo de datos históricos, para limpiarlos. Una vez realizada la actividad de limpieza se procede a entrenar el

modelo con tendencias y patrones en cómo cambian los delitos año tras año. Una vez entrenado, el modelo está listo para hacer predicciones sobre datos no vistos (años futuros).

3. Datos de Salida (Output)/Proyección:

* Funcionamiento: El producto de salida del modelo es la predicción o proyección generada mediante el modelo entrenado. Para un año futuro dado (por ejemplo, 2026), el modelo utilizará lo que aprendió de los datos históricos con la finalidad de realizar una estimación del total patrón delictivo específico para ese año.

* Empleo: La proyección (el valor 2026) se utilizará como una estimación de una posibilidad con un margen de error aceptable, en búsqueda del posible comportamiento de los delitos o patrones de delitos en base a datos históricos. Esta información es el resultado principal del modelado, servirá como base para la toma de decisiones así también como para realizar la planificación. Además, el conocimiento y sociabilización del análisis sería mediante un tablero en tableau.

2.3. Validación de la propuesta

Este trabajo fue validado por dos Especialistas:

- Hernan Alejandro Arteaga Calispa => Arquitecto en Datos (Anexo 3)
- Laura Estefania Landeta Sandoval => Magister en Neuropsicología (Anexo 4)

2.4. Matriz de articulación de la propuesta

Tabla 4.
Matriz de articulación

EJES O PARTES PRINCIPALES	SUSTENTO TEÓRICO	SUSTENTO METODOLÓGICO	ESTRATEGIAS / TÉCNICAS	DESCRIPCIÓN DE RESULTADOS	INSTRUMENTOS APLICADOS
Datos de Entrada (extracción)	Basado en el enfoque de minería de datos.	Metodología cuantitativa con enfoque exploratorio.	Normalización de datos. Almacenamiento Consolidado en Google. Colan	Se genera un archivo unificado que contiene variables clave.	Repositorio digital (Google Drive), hoja de cálculo estructurada, scripts de extracción.
Limpieza Procesamiento	Basado en aprendizaje automático, específicamente en el algoritmo Random Forest,	Se aplica una técnica supervisada de machine learning	Implementación del algoritmo Random Forest, limpieza de datos inconsistentes	El modelo logra identificar cómo varían los delitos a lo largo del tiempo.)	Librerías de machine learning (scikit-learn, pandas).
Datos de Salida proyección	Se apoya en modelos predictivos que permiten estimar comportamientos	Se utiliza una proyección estadística basada en el modelo entrenado.	Predicciones de delitos por año 2026	Se obtendrá una estimación del número de delitos para años futuros,	Modelo entrenado, visualizadores de datos (tableau),

Fuente: Elaboración propia

CONCLUSIONES

En base a la investigación se concluye que la integración de bases teóricas permiten comprender que el análisis de datos criminales no solo es un aspecto técnico , sino además una herramienta para interpretar fenómenos sociales como lo son los delitos, es decir que mediante los análisis criminales es posible establecer características para entender la sociedad actual, lo cual es de suma importancia ya que los datos resumen los acontecimientos actuales, criminalidad, patrones y tendencias actuales..

El diagnóstico concluyó que ambos países enfrentan retos similares en cuanto a delitos, aunque existen diferencias entre los factores que los impulsan, se ha destacado que los patrones comportamentales en los delitos se disparan bajo factores sismilares, es así como se podría validar que en Ecuador de acuerdo a los datos en los últimos años ha tenido un mayor índice delictivo,es importante denotar ya que en países como México en cambio se denota una variante uniforme de crecimiento, mientras que en ecuador existen picos de criminalidad considerando diferencias marcadas acerca del comportamiento de la delincuencia.

El dashboard permitió identificar cómo los patrones de comportamiento avanzan, así mismo tener una visión clara de un solo vistazo, considerando las tendencias de registradas entre México vs Ecuador, con la finalidad de tener una observación detallada sobre los datos de forma inmediata y en tiempo real. Donde se enmarca la importancia de conocer cómo avanzan los datos sobre una tendencia, así como los patrones comportamentales sobre delitos seriales para brindar un enfoque más específico sobre la realidad actual.

Mediante la validación por parte de expertos se permite concluir que existen áreas de mejora, además de validar que se cumplan aspectos técnicos, considerando el manejo de los datos, las herramientas que se utilizó entre otros factores los cuales deberían ser considerados para próximas investigaciones, además de la importancia de la validación de un experto sobre el comportamiento para determinar qué series comportamentales no solamente se aplican en una ciudad sino en varias realizado un aporte práctico en el desarrollo de análisis predictivos acerca de la seguridad.

RECOMENDACIONES

Se recomienda fortalecer el vínculo de la ciencia de datos y los patrones de comportamiento, promoviendo espacios donde exista una contextualización interdisciplinaria con expertos de seguridad, psicólogos y personas de tecnología con la finalidad de colaborar en la generación de modelos predictivos precisos, socialmente aceptables.

Por otro lado, está la puerta siempre abierta para realizar investigaciones comparativas con otros algoritmos,para ello se recomienda que las instituciones públicas fortalezcan la recolección y transparencia de los datos, para tener un fácil acceso a información confiables que permita a los investigadores tomar decisiones más acercadas a la realidad con base a evidencia más no a suposiciones.

Se recomienda que se realice una divulgación más amplia sobre el dashboard y que se lo mantenga actualizado además se integren a plataformas accesibles públicamente, con la finalidad que cualquier persona pueda enterarse acerca de las diferentes realidades y como avanza el Ecuador y México en materia de delitos.

Se recomienda la creación de un comité de especialistas que cada cierto tiempo ejecute una revisión acerca del desempeño del dashboards así como de los datos que este muestra. Además se sugiere mantener una retroalimentación continua basada en la observación de especialistas del área.

BIBLIOGRAFÍA

Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine learning applications based on SVM classification review. Qubahan Academic Journal, 1(2), 81-90.

Abeliuk, A., & Gutiérrez, C. (2021). Historia y evolución de la inteligencia artificial. Revista Bits de Ciencia, (21), 14-21.

Amiconi, D. F. (2023). Enseñanza y aplicación de métodos numéricos en problemas de ingeniería, utilizando la potencialidad de la inteligencia artificial (chatgpt) combinada con programación en python y entorno de trabajo colaborativo en google colab. AJEA (Actas de Jornadas y Eventos Académicos de UTN), (AJEA 30).

Arreaga, K. M. P. (2021). Evolución de la teoría del delito. Revista Diversidad Científica, 1(1), 97-104.

Barroso, F. B. G., & Soria, Y. L. (2022). La perfilación criminal como herramienta forense en la investigación de delitos contra la vida. Ecuador. Sociedad & Tecnología, 5(2), 365-378.

Caraveo, Z. K. G., Payró, M. P. S., & De los Santos Torres, G. (2025). Análisis comparativo de herramientas para explorar grandes volúmenes de datos: propuesta metodológica. Dilemas contemporáneos: Educación, Política y Valores.

Cedeño Delgado, I., & Rodríguez Véliz, M. (2025). Evaluación de la calidad del software en sistemas de educación superior utilizando el modelo CMMI: Evaluation of Software Quality in Higher Education Systems Using the CMMI Model. Revista Científica Multidisciplinar G-Nerando, 6(1), Pág. 192 –. https://doi.org/10.60100/rcmg.v6i1.402

Chen, H., Gomez, C., Huang, C.-M., & Unberath, M. (2021). Explainable Medical Imaging AI Needs Human‐Centered Design: Guidelines and Evidence from a Systematic Review.

Chitarroni, H. (2002). La regresión logística.

Fernández Pérez, A. O. Pobreza, desigualdad y delitos en Ecuador 2010-2022 (Master's thesis, Quito, Ecuador: Flacso Ecuador).

García, Ó. (2021). Genealogía forense. Implicaciones sociales, éticas, legales y científicas. Revista Española de Medicina Legal, 47(3), 112-119

García Añino, E., Bonell, C., Dutto, I., Catalfamo Formento, P. (2023). Evaluación de ámbitos clínicos para el desarrollo de herramientas de Análisis del Movimiento Humano.

Gómez Albarello, J. G., & Corzo Salamanca, J. A. (2021). Criminalidad homicida, capitalismo y democracia. Análisis Político, 34(102), 23-53.

Granados Garzón, L. (2024). Análisis criminológico de grupos delincuenciales y sistema integrado de alerta temprana para detectar patrones de ocupación criminal (Doctoral dissertation, Universidad de Granada).

Guzmán Sánchez, A. K. (2024). Factores económicos y la delincuencia en la provincia de Tungurahua.

Hidalgo Guayaquil, A. E. (2024). Enfoque criminológico del delito de terrorismo en el Ecuador (Master's thesis, Quito: Universidad de las Américas, 2024).

Holguín-Londoño, M. (2024). Introducción a la ciencia y análisis de datos con Python: una visión empresarial. Universidad Tecnológica de Pereira. Disponible en: https://hdl.handle.net/11059/15599

Martínez Pérez, J. A., & Pérez Martín, P. S. (2024). Regresión logística. SEMERGEN, Soc. Esp. Med. Rural Gen.(Ed. Impr.), e102086-e102086.

Malamud, C., & Núñez, J. (2024). Informe sobre el tráfico de cocaína en América Latina: implicaciones para la seguridad regional. En UNODC (2023), Informe Mundial sobre las Drogas 2023. Oficina de las Naciones Unidas contra la Droga y el Delito. https://www.unodc.org/res/WDR-2023/Special_points_S.pdf

López, C. P. (2024). Ciencia de datos a través de Python: Técnicas de aprendizaje supervisado: clasificación. Ibergarceta.

López Ríos, J. M., Mejía Merino, C. M., Frías Epinayú, C. E., & Marulanda, S. C. (2021). Estrategias comunitarias para la seguridad alimentaria en indígenas wayuu, La Guajira, Colombia. *Revista Española de Nutrición Comunitaria*, *27*(1)

Pineda, A. L. P. (2022). Violencia intrafamiliar. Huella de la Palabra, (16), 30-41.

Secretaría Nacional de Planificación. (2021). Plan de Creación de Oportunidades 2021-2025. Creación de oportunidades, 17-34. https://acortar.link/1unBK7

Padilla, J., & Contreras, L. (2024). Ciencia de datos con python: Transformación y selección de variables. Ediciones de la U.

Pinargote-Zambrano, J. J., Lino-Calle, V. A., & Vera-Almeida, B. J. (2024). Python en la enseñanza de las Matemáticas para estudiantes de nivelación en Educación Superior. MQRInvestigar, 8(3), 3966-3989.

Rivera, R., Lovato, K., Vallejo, M. F., & Chaves Olaya, L. (2025). Boletín anual de homicidios intencionales en Ecuador: Análisis estadístico 2024. Observatorio Ecuatoriano de Crimen Organizado (OECO). Fundación Panamericana para el Desarrollo (PADF).

Sholeh, M., Lestari, U., & Andayati, D. (2024). Comparison of Feature Selection with Information Gain Method in Decision Tree, Regression Logistic and Random Forest Algorithms. *Journal of Applied Business and Technology*, *5*(3), 146-153.

Tascón, M. (2013). Introducción: Big data. Pasado, presente y futuro. Telos: Cuadernos de comunicación e innovación, (95), 47-50.

Uribe, J. J. M. (2024). Big data e inteligencia de negocios aplicado en el sector gastronómico. Infometric@-Serie Ingeniería, Básicas y Agrícolas, 7(1).

Vera Alarcón, M. J. (2024). Modelo predictivo de las tendencias de delitos en el Ecuador [Tesis de Maestría, Universidad Técnica Estatal de Quevedo]. Repositorio UTEQ. https://repositorio.uteq.edu.ec/server/api/core/bitstreams/9334ee09-50e9-4ab4-a903-eb354c5acfc0/content.

Vegas Gabriel, J. (2024). Redes neuronales diferenciales: Fundamentos y Aplicaciones.

Vilema Lara, P. H. (2022). Detección de fallos en mantenimiento predictivo utilizando el método de aprendizaje de máquina Random Forest.

Zacarías, E. E. (2025). Diseño de la investigación de implementación de pruebas automatizadas para certificaciones de control de calidad en el ciclo de desarrollo de software de freelance,https://biblio.ingenieria.usac.edu.gt/protocolos/2025/TGP1673.pdf

ANEXOS

Anexo 1

Formato Webinar

Webinar: *"Random Forest y Validación Predictiva en IA"*

Objetivo

Presentar de forma clara y práctica que es IA, cómo funciona el algoritmo Random Forest dentro del contexto de la inteligencia artificial, y cómo se valida su desempeño en aplicaciones reales.

Duración Total: 30 minutos

Minuto	Tema	Detalles
0–5	**Bienvenida e introducción a la IA predictiva**	Breve explicación de qué es la IA, qué significa "predictivo" y por qué es relevante hoy. Ejemplos cotidianos (diagnóstico médico, predicción de fraudes, clima).
5–15	**¿Qué es Random Forest?**	

- Explicación sencilla del algoritmo: árboles de decisión, votación, robustez.
- Ventajas frente a otros modelos (precisión, resistencia al overfitting).
- Ejemplo visual: predicción de abandono escolar o riesgo crediticio. | | 15–25 | **Validación del modelo** |
- Métricas clave: precisión, recall, F1-score, matriz de confusión.
- Validación cruzada y división de datos.
- ¿Cómo saber si el modelo es confiable? ¿Qué errores puede cometer? | | 25–30 | **Cierre reflexivo y preguntas** |
- ¿Qué impacto tiene un modelo mal validado en la vida real?
- Espacio para 2–3 preguntas rápidas.
- Invitación a seguir aprendiendo y aplicar con responsabilidad.

Pendiente Cargar la hoja de firmas

URL

:

https://bgrecuador-my.sharepoint.com/personal/hgmz7852_bgr_com_ec/_layouts/15/stream.aspx?id=%2Fpersonal%2Fhgmz7852%5Fbgr%5Fcom%5Fec%2FDocuments%2FGrabaciones%2FPaso%20a%20Producci%C3%B3n%2D20250826%5F221603%2DGrabaci%C3%B3n%20de%20la%20reuni%C3%B3n%2Emp4&referrer=StreamWebApp%2EWeb&referrerScenario=AddressBarCopied%2Eview%2E82de79b7%2D5176%2D4c31%2D93a8%2D14bb30ebb641&isDarkMode=false

UNIVERSIDAD TECNOLÓGICA ISRAEL
ESCUELA DE POSGRADOS "ESPOG"
REGISTRO DE PARTICIPACIN

Tema: Inteligencia Artificial, Big Data y Algoritmos Predictivos

Fecha: _____ Lugar: Banco General Rumiñahui

N°	NOMBRE COMPLETO	CARGO	FIRMA
1	Cristian Hernández	Desarrollador	
2	Andrea Caicedeas	Oficial de Gestión de servicios de TI	
3	Omar Andino	Scrum Master	
4	Mauricio Tompanta	Desarrollador SV	
5	Jenny Terán	Desarrollador SW	
6	CHRISTIAN CATACTA	Scrum Master	
7	Cristina Morales	Desarrollador Junior	
8	Erick Gamboa	Arquitecto de Soluciones de TI	
9	Verónica Recalde	Jefe de Aplicaciones	
10	Angel Carrillo	Arq Soluciones TI	

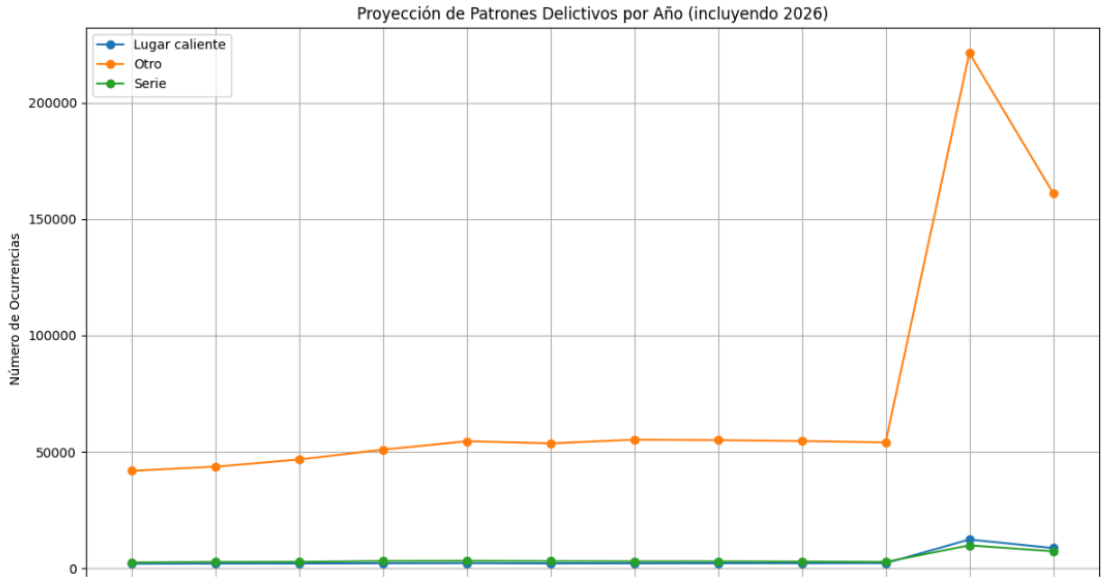
Firma Responsable:

_____

Hans Muñoz

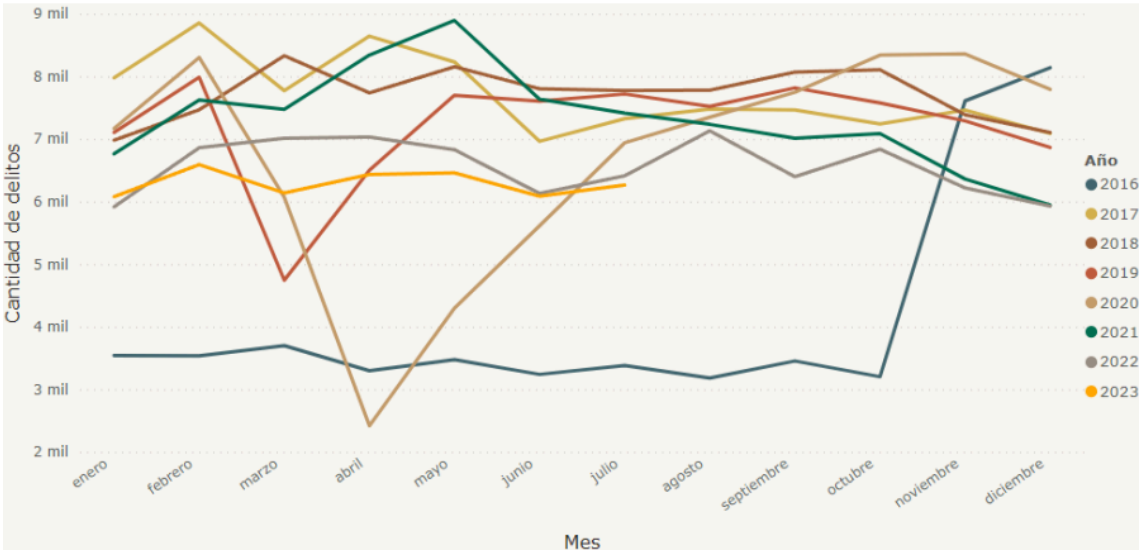
ANEXO 2

Proyección de patrones delictivos por Año (incluida proyección)

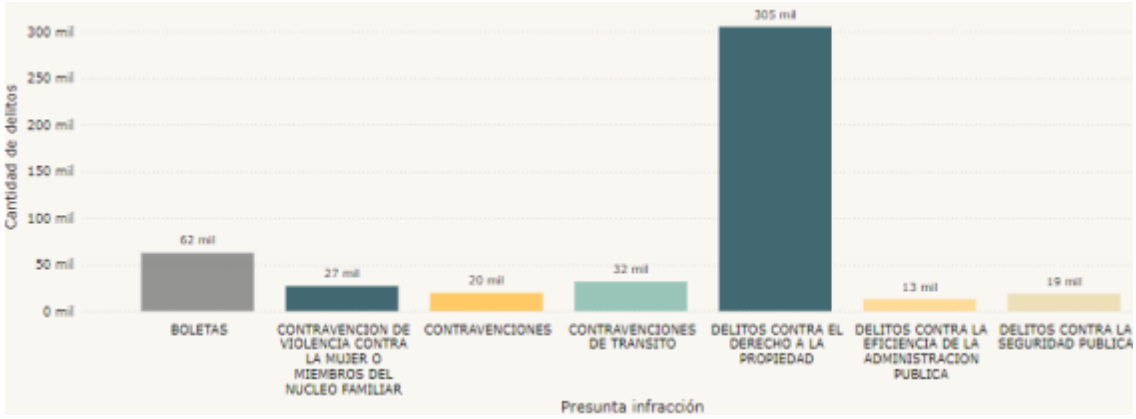


OTROS GRÁFICOS

DISTRIBUCIÓN POR MES



Distribución por tipo infracción



ANEXO 3

Perfil Especialista

Universidad Israel
ESPOG | Escuela de Posgrados

UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA BIG DATA Y CIENCIA DE DATOS

INSTRUMENTO PARA VALIDACIÓN DE LA PROPUESTA

Estimado colega:

Se solicita su valiosa cooperación para evaluar la calidad del siguiente contenido digital **"**Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias**"**. Sus criterios son de suma importancia para la realización de este trabajo, por lo que se le pide que brinde su cooperación contestando las preguntas que se realizan a continuación.

Datos informativos

Validado por: Hernan Alejandro Arteaga Calispa
Título obtenido: MASTER EN ANALISIS Y VISUALIZACION DE DATOS MASIVOS ANALISIS Y VISUALIZACION DE DATOS MASIVOS
C.I.: 1713979514
E-mail: harteaga@bgr.com.ec
Institución de Trabajo: Banco General Rumiñahui
Cargo: ARQUITECTO DE SOLUCIONES TI
Años de experiencia en el área: 10

Universidad Israel

ESPOG | Escuela de Posgrados

Instructivo:

- Responda cada criterio con la máxima sinceridad del caso.
- Revisar, observar y analizar la propuesta de la plataforma virtual, blog o sitio web.
- Coloque una X en cada indicador, tomando en cuenta que Muy adecuado equivale a 5, Bastante Adecuado equivale a 4, Adecuado equivale a 3, Poco Adecuado equivale a 2 e Inadecuado equivale a 1.

Tema: "Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias"

Indicadores	Muy adecuado	Bastante Adecuado	Adecuado	Poco adecuado	Inadecuado
Pertinencia	x				
Aplicabilidad	x				
Factibilidad	x				
Novedad		x			
Fundamentación pedagógica	x				
Fundamentación tecnológica	x				
Indicaciones para su uso	x				
TOTAL	34				

Observaciones: El uso de Random Forest implementado para el análisis de patrones predictivos de comportamiento en serie, requiere una gran cantidad de datos para entrenar el algoritmo y que su error absoluto sea menor a lo esperado. Además, al Integrar PYSpark en Colab permite escalar el análisis lo que hace un estudio más completo.

Recomendaciones: Se recomienda que el proyecto de titulación detalle la infraestructura computacional y los componentes donde se procesa debido al enfoque de nube presentado, esto con el objetivo de proponer una arquitectura escalable y parametrizable que sea el punto de partida para el inicio de otras investigaciones.

Lugar, fecha de validación: Quito, 1 septiembre 2025

HERNAN ALEJANDRO ARTEAGA CALISPA

Firmado digitalmente por HERNAN ALEJANDRO ARTEAGA CALISPA
Fecha: 2025.09.01 17:22:46 -05'00'

Firma del especialista
Hernan Alejandro Arteaga Calispa

ANEXO 4

Perfil Especialista

UNIVERSIDAD TECNOLÓGICA ISRAEL

ESCUELA DE POSGRADOS "ESPOG"

MAESTRÍA BIG DATA Y CIENCIA DE DATOS

INSTRUMENTO PARA VALIDACIÓN DE LA PROPUESTA

Estimado colega:

Se solicita su valiosa cooperación para evaluar la calidad del siguiente contenido digital "Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias". Sus criterios son de suma importancia para la realización de este trabajo, por lo que se le pide que brinde su cooperación contestando las preguntas que se realizan a continuación.

Datos informativos

Validado por: Laura Estefania Landeta Sandoval
Título obtenido: MÁSTER EN ANALISIS Y VISUALIZACION DE DATOS MASIVOS ANALISIS Y VISUALIZACION DE DATOS MASIVOS
C.I.: 1722314034
E-mail: tefa.93jm@hotmail.com
Institución de Trabajo: Centro Neuropsicorehabilitacion "Cente"
Cargo: Directora
Años de experiencia en el área: 8

Instructivo:

- Responda cada criterio con la máxima sinceridad del caso.
- Revisar, observar y analizar la propuesta de la plataforma virtual, blog o sitio web.

Universidad Israel

ESPOG | Escuela de Posgrados

- Coloque una X en cada indicador, tomando en cuenta que Muy adecuado equivale a 5, Bastante Adecuado equivale a 4, Adecuado equivale a 3, Poco Adecuado equivale a 2 e Inadecuado equivale a 1.

Tema: "Análisis predictivo de series temporales de delitos en México y Ecuador mediante el algoritmo Random Forest: una comparación de patrones y tendencias"

Indicadores	Muy adecuado	Bastante Adecuado	Adecuado	Poco adecuado	Inadecuado
Pertinencia	x				
Aplicabilidad	x				
Factibilidad	x				
Novedad	x				
Fundamentación pedagógica	x				
Fundamentación tecnológica	x				
Indicaciones para su uso	x				
TOTAL	35				

Observaciones: Se puede identificar patrones conductuales que podrían estar vinculadas a estructuras cognitivas repetitivas o disfuncionales de acuerdo con lo expuesto en el proyecto. Desde la neuropsicología, esto sugiere que ciertos perfiles delictivos podrían responder a esquemas mentales automatizados, reforzados por contextos sociales y emocionales específicos.

Recomendaciones: Se recomienda incluir en el proyecto una dimensión interpretativa que relacione los patrones detectados con modelos neuroconductuales. Esto permitiría no solo predecir delitos, sino también entender los mecanismos cognitivos que los sustentan, facilitando intervenciones preventivas más efectivas y personalizadas.

Lugar, fecha de validación: Quito, 8 septiembre 2025

Firma del especialista
Laura Estefania Landeta Sandoval

ANEXO 5

Observación Directa

DESARROLLO DEL TRABAJO	OBSERVACIONES	CATEGORÍA
HABILIDADES DE ORGANIZACIÓN Y PLANIFICACIÓN	Se observa una planificación clara que distingue y ejecuta de forma paralela dos metodologías (cualitativa y cuantitativa). La estructura del informe evidencia una organización deliberada.	Positivo
CONOCIMIENTOS BÁSICOS EN BÚSQUEDA Y TRATAMIENTO	Conocimiento sólido en identificación de fuentes oficiales, uso de Python para limpieza y transformación de datos, y aplicación de algoritmos avanzados (Random Forest).	Positivo
HABILIDADES EN INVESTIGACIÓN Y BÚSQUEDA	Habilidad para recurrir a fuentes primarias oficiales y confiables. Se complementa la data cuantitativa con investigación cualitativa mediante observación directa y fuentes secundarias.	Neutro
HABILIDADES DE GESTIÓN DE LA INFORMACIÓN	Excelente gestión al integrar datos de múltiples fuentes, transformarlos, modelarlos y visualizarlos en dashboard. Se categoriza la información cualitativa de forma efectiva.	Positivo
HABILIDADES DE ANÁLISIS Y SÍNTESIS	Capacidad analítica fuerte al cruzar hallazgos cuantitativos con observaciones cualitativas para generar insights contextuales.	Positivo
HABILIDADES DE RESOLUCIÓN DE PROBLEMAS	Se identifican problemas (datos mal estructurados, necesidad de ajustes en modelos) y se aplican soluciones técnicas. La carga de trabajo pesada se percibe como un punto negativo.	Negativo
HABILIDADES EN LA GESTIÓN	El trabajo parece estar bien dividido, permitiendo el avance simultáneo de técnicas cualitativas y cuantitativas. La coherencia del informe sugiere buena coordinación grupal.	Neutro