Aplicación de un modelo predictivo para la optimización de auditorías y recuperación del patrimonio público basado en datos del sistema APA

José Nicolás Gómez Aranzalez

Jesús Antonio Cabarcas Gómez

Asesor

Felipe Alexander Pipicano Guzman

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2025

Resumen

La presente investigación aborda el desafío de optimizar el proceso de control fiscal en la Contraloría General de la República mediante la implementación de técnicas de Ciencia de Datos e Inteligencia Artificial sobre la base de datos del Sistema Automatizado para el Proceso Auditor (APA). El objetivo principal fue desarrollar un modelo predictivo capaz de estimar la probabilidad de materialización de hallazgos con incidencia fiscal, disciplinaria y penal, facilitando así la priorización de auditorías y la recuperación del patrimonio público.

La metodología integró un enfoque híbrido de procesamiento de datos. Se utilizó Inteligencia Artificial Generativa (Meta Llama 3 - 70B) para enriquecer el conjunto de datos, transformando campos de texto no estructurado (causas y efectos redactados por auditores) en variables categóricas estandarizadas. Posteriormente, se entrenó un modelo de aprendizaje supervisado de clasificación multi-etiqueta basado en el algoritmo Random Forest Classifier, aplicando técnicas de balanceo de clases para mitigar la asimetría en la distribución de los datos.

Los resultados obtenidos fueron concluyentes: el modelo alcanzó un Jaccard Score de 0.9360 y una precisión perfecta (F1-Score de 1.00) en la detección de hallazgos con incidencia fiscal, superando significativamente a la línea base de Regresión Logística. Estos hallazgos demuestran que la herramienta propuesta permite filtrar y priorizar automáticamente el 20% de los casos de mayor relevancia, optimizando la asignación de recursos humanos y reduciendo la carga operativa en la revisión de hallazgos administrativos.

Palabras clave: Control Fiscal, Machine Learning, Random Forest, Inteligencia Artificial Generativa, Llama 3, Priorización de Auditorías, Sistema APA.

Abstract

This research addresses the challenge of optimizing the fiscal control process at the Comptroller General of the Republic by implementing Data Science and Artificial Intelligence techniques on the Automated Auditor Process System (APA) database. The main objective was to develop a predictive model capable of estimating the probability of fiscal, disciplinary, and penal findings, thus facilitating audit prioritization and public asset recovery.

The methodology integrated a hybrid data processing approach. Generative Artificial Intelligence (Meta Llama 3 - 70B) was used to enrich the dataset, transforming unstructured text fields (causes and effects written by auditors) into standardized categorical variables. Subsequently, a multi-label supervised learning model based on the Random Forest Classifier algorithm was trained, applying class balancing techniques to mitigate asymmetry in data distribution.

The results were conclusive: the model achieved a Jaccard Score of 0.9360 and perfect precision (F1-Score of 1.00) in detecting findings with fiscal incidence, significantly outperforming the Logistic Regression baseline. These findings demonstrate that the proposed tool allows for the automatic filtering and prioritization of the top 20% of cases with the highest relevance, optimizing human resource allocation and reducing operational load in the review of administrative findings.

Keywords: Fiscal Control, Machine Learning, Random Forest, Generative Artificial Intelligence, Llama 3, Audit Prioritization, APA System.

Tabla de Contenido

Lista de Tablas

Lista de Figuras

Introducción

En la era de la transformación digital, las entidades de control fiscal se enfrentan al reto de procesar volúmenes masivos de información para garantizar la correcta vigilancia de los recursos públicos. La Contraloría General de la República (CGR), a través de la implementación del Sistema Automatizado para el Proceso Auditor (APA), ha logrado centralizar la trazabilidad de los hallazgos fiscales en Colombia. Sin embargo, esta digitalización ha traído consigo un nuevo desafío: la acumulación de grandes cantidades de datos no estructurados y la necesidad de priorizar eficientemente miles de registros que, actualmente, requieren revisión manual exhaustiva.

El presente trabajo de grado aborda esta problemática desde una perspectiva de Ciencia de Datos aplicada, proponiendo una solución tecnológica que trasciende el análisis descriptivo tradicional. La investigación se centra en la aplicación de modelos predictivos avanzados sobre la base de datos histórica del sistema APA, con el objetivo de estimar la probabilidad de materialización de hallazgos con incidencia fiscal, disciplinaria y penal antes de que surtan todas las etapas procesales.

La novedad técnica de esta propuesta reside en su enfoque híbrido. Por un lado, se integra la Inteligencia Artificial Generativa (LLMs) para estructurar y categorizar la información cualitativa (narrativas de causas y efectos redactadas por los auditores), convirtiendo texto libre en variables predictivas de alto valor. Por otro lado, se implementan algoritmos de aprendizaje supervisado (Random Forest) diseñados para manejar el desbalance de clases inherente a los datos de fraude y corrupción.

A lo largo del documento se detalla la metodología empleada, comenzando por un riguroso Análisis Exploratorio de Datos (AED) y el preprocesamiento asistido por IA, pasando

por la evaluación comparativa de arquitecturas de modelado, hasta llegar a la estimación del impacto operativo. Los resultados demuestran que es posible automatizar la focalización de auditorías con un alto grado de precisión, permitiendo a la entidad transitar de un modelo de control reactivo y masivo, a uno preventivo, inteligente y selectivo.

Justificación

Los organismos de control fiscal en Colombia presentan tasas de recuperación de recursos públicos significativamente bajas, con procesos auditores que pueden extenderse por años sin generar resultados tangibles. La implementación de herramientas predictivas y automatizadas puede reducir sustancialmente los tiempos de ejecución e incrementar la probabilidad de éxito de los procesos de responsabilidad fiscal. La capacidad actual de los organismos de control para supervisar la totalidad del territorio nacional es limitada, especialmente en regiones apartadas donde paradójicamente pueden presentarse mayores riesgos de mal manejo de recursos públicos.

Además, la sociedad colombiana demanda cada vez más transparencia y efectividad en la protección del patrimonio público. Los ciudadanos esperan que los organismos de control utilicen las mejores herramientas disponibles para garantizar que sus recursos tributarios sean utilizados de manera eficiente y honesta.

Por todo esto y ante el desarrollo acelerado de tecnologías de inteligencia artificial, análisis de big data y automatización de procesos se han creado oportunidades sin precedentes para revolucionar la manera en que se ejecuta el control fiscal. La no adopción de estas herramientas constituiría una negligencia institucional.

Los Fundamentos de la Ciencia de Datos Aplicada demuestran que los algoritmos de machine learning y análisis predictivo pueden identificar patrones complejos en grandes volúmenes de datos que serían imperceptibles para el análisis humano tradicional. Esta capacidad resulta especialmente relevante en el contexto fiscal, donde las irregularidades a menudo se manifiestan a través de anomalías sutiles en los datos financieros.

Los organismos de control fiscal en Colombia han acumulado importantes volúmenes de datos históricos que pueden ser utilizados para entrenar modelos predictivos y validar la efectividad de las herramientas propuestas.

La justificación para el desarrollo de este proyecto se fundamenta en argumentos sólidos desde múltiples perspectivas que demuestran no solo la viabilidad técnica de la propuesta, sino también su necesidad imperativa para el fortalecimiento del control fiscal en Colombia y su potencial para generar beneficios tangibles para la sociedad en su conjunto.

Objetivos

Objetivo General

Aplicar un modelo predictivo basado en técnicas de aprendizaje supervisado sobre la base de datos del sistema APA, para estimar la probabilidad de materialización de hallazgos fiscales y optimizar la planeación del proceso auditor en la Contraloría General de la República.

Objetivos Específicos

Examinar y organizar las fuentes de datos históricas de los hallazgos fiscales (datos estructurados y no estructurados) para contar la experiencia de la ingesta y definir los requerimientos de pre-procesamiento necesarios para el entrenamiento del modelo predictivo.

Diseñar un protocolo de extracción de características que emplee Inteligencia Artificial generativa para procesar y extraer la información relevante, transformándola en variables aptas para el machine learning.

Realizar un análisis exploratorio que permita el conjunto de datos mediante el uso de visualizaciones y métodos estadísticos.

Implementar el modelo de clasificación predictiva que estime la probabilidad de que hallazgos preliminares deriven en procesos de responsabilidad fiscal exitosos o en archivo, utilizando los datos fiscales pre-procesados.

Evaluar el rendimiento del modelo predictivo a través del análisis de métricas de precisión y la comparación de resultados con al menos dos métodos de clasificación diferentes, para determinar la arquitectura de machine learning más adecuada.

Marcos de Referencia

Estado del Arte

Desarrollos Internacionales

Implementación de IA en Auditoría Pública. El U.S. Department of the Treasury (2024) reporta resultados significativos en la implementación de procesos mejorados de detección de fraudes, incluyendo machine learning, que permitieron prevenir y recuperar más de $4 mil millones en el año fiscal 2024. Este caso representa un referente internacional en la aplicación práctica de inteligencia artificial para la recuperación de recursos públicos. Rodriguez y Thompson (2024) analizan el efecto de la transformación digital auditora en la calidad de la auditoría, encontrando que la implementación de confirmaciones bancarias digitales mejora significativamente la precisión y eficiencia del proceso auditor.

Avances en Machine Learning para Detección de Fraudes. Gómez et al. (2023) presentan un caso de estudio colombiano sobre análisis predictivo para evaluación de riesgos en auditoría pública, demostrando que la implementación de algoritmos de machine learning puede mejorar la tasa de detección de irregularidades hasta en un 40%.

Desarrollos Nacionales

Iniciativas de Modernización del Control Fiscal. La Contraloría General de la República de Colombia (2023) ha establecido su Plan Estratégico de Transformación Digital 2023-2026, que incluye la implementación de herramientas tecnológicas avanzadas para optimizar el proceso auditor y mejorar la cobertura territorial del control fiscal.

Morales (2023) desarrolla un modelo específico para la implementación de técnicas de machine learning en la detección de fraudes en entidades públicas colombianas, proporcionando una base metodológica adaptada al contexto nacional.

Casos de Aplicación Regional. Castro y Rojas (2024) analizan los desafíos de implementación de procedimientos auditores automatizados usando inteligencia artificial en países en desarrollo, identificando factores críticos de éxito específicos para el contexto latinoamericano.

Brechas Identificadas en el Estado del Arte. Pese a los avances significativos, se identifican las siguientes brechas:

- Integración de Sistemas: Falta de modelos integrales que combinen priorización de riesgos, procedimientos automatizados y seguimiento de hallazgos en una sola plataforma.
- Validación de Evidencias: Desarrollo limitado de sistemas expertos específicos para validación automatizada de evidencias en el contexto del control fiscal.
- Predicción de Materialización: Escasez de modelos predictivos que estimen la probabilidad de que hallazgos preliminares deriven en procesos de responsabilidad fiscal exitosos.

Marco Normativo

Normativa Constitucional y Legal

Fundamentos Constitucionales

La Constitución Política de Colombia de 1991 establece en su artículo 267 las funciones de la Contraloría General de la República, incluyendo la vigilancia de la gestión fiscal y el ejercicio del control fiscal. Esta norma fundamental proporciona el marco jurídico para la implementación de herramientas tecnológicas en el control fiscal.

Ley 610 de 2000. La Ley 610 de 2000 "Por la cual se establece el trámite de los procesos de responsabilidad fiscal de competencia de las contralorías" define el proceso de responsabilidad fiscal y establece los principios y objeto de tal responsabilidad. El artículo 1° define el proceso de responsabilidad fiscal y determina los criterios para establecer el daño patrimonial al Estado, proporcionando el marco legal para la implementación de herramientas predictivas de materialización de hallazgos.

Normativa de Transformación Digital

Estrategia Nacional de Transformación Digital. El gobierno colombiano ha establecido la Estrategia Nacional de Transformación Digital 2023-2026, que incluye lineamientos específicos para la modernización de entidades públicas mediante la adopción de tecnologías digitales (Ministerio de Tecnologías de la Información y las Comunicaciones, 2023).

Normativa Sobre Inteligencia Artificial

Actualmente se encuentran en trámite en el Congreso de la República más de 10 proyectos de ley relacionados con la regulación de la inteligencia artificial en Colombia, nueve de iniciativa propia del Congreso y uno del Gobierno Nacional (Ámbito Jurídico, 2024). Estos proyectos buscan regular tanto usos específicos de la IA como su regulación integral.

Normativa Técnica y Estándares

Estándares Internacionales. La International Organization of Supreme Audit Institutions (INTOSAI) ha establecido lineamientos para la transformación digital en entidades superiores de auditoría, proporcionando estándares internacionales para la implementación de tecnologías en procesos auditores (INTOSAI, 2023).

Estándares de Auditoría. Las Normas Internacionales de Auditoría (NIA) y su adaptación nacional proporcionan el marco técnico para la implementación de procedimientos auditores automatizados, manteniendo la calidad y confiabilidad del proceso auditor.

Normativa sobre Protección de Datos

Ley 1581 de 2012. La Ley de Protección de Datos Personales establece las condiciones y procedimientos para el tratamiento de datos personales, siendo fundamental para la implementación de sistemas de análisis predictivo que manejen información sensible de entidades auditadas.

Decreto 1377 de 2013. Reglamenta la Ley 1581 de 2012 y establece los requisitos técnicos y administrativos para el tratamiento de datos personales, proporcionando el marco normativo para la implementación segura de herramientas de análisis de datos en control fiscal.

Implicaciones Normativas Para el Proyecto

La implementación del sistema integrado de herramientas basadas en ciencia de datos debe cumplir con:

1. Principios del Control Fiscal: Eficiencia, eficacia, economía y equidad establecidos en la normativa constitucional y legal.
2. Debido Proceso: Garantías procesales establecidas en la Ley 610 de 2000 para los procesos de responsabilidad fiscal.

3. Protección de Datos: Cumplimiento de la normativa de protección de datos personales en el procesamiento de información sensible.

4. Estándares Técnicos: Adopción de estándares internacionales para mantener la calidad y confiabilidad del proceso auditor.

Marco Teórico

Teoría de la Transformación Digital en el Sector Público

La transformación digital en auditoría se fundamenta en la teoría de la innovación tecnológica aplicada al sector público. Kumar y Sharma (2024) establecen que la transformación digital en la auditoría pública se sitúa "entre la esperanza y el miedo", reconociendo tanto las posibilidades como las dificultades que surgen al aplicar las nuevas tecnologías en los procesos tradicionales de control fiscal.

Vasarhelyi y Rozario (2023) proporcionan el marco teórico fundamental para la auditoría en la era digital, argumentando que la integración de tecnologías emergentes no solo mejora la eficiencia operativa, sino que transforma fundamentalmente la naturaleza del proceso auditor. Esta perspectiva teórica se alinea con los hallazgos de Appelbaum y Nehmer (2022), quienes establecen que la analítica de datos y la inteligencia artificial constituyen pilares fundamentales de la auditoría moderna.

Teoría del Análisis Predictivo en Control Fiscal

La aplicación del análisis predictivo en control fiscal se sustenta en la teoría de la decisión bajo incertidumbre y los modelos de machine learning supervisado. Chen et al. (2024) demuestran que las técnicas de machine learning aplicadas a la detección de fraudes financieros pueden lograr tasas de precisión superiores al 95% cuando se implementan adecuadamente, proporcionando una base teórica sólida para la priorización basada en riesgo.

Fernández y López (2021) desarrollan el marco teórico específico para la mejora de la eficiencia auditora mediante machine learning, estableciendo que la utilización de datos de red de contribuyentes permite identificar patrones de comportamiento fraudulento que no son detectables mediante métodos tradicionales.

Teoría de Sistemas Expertos en Validación de Evidencias

La validación automatizada de evidencias se fundamenta en la teoría de sistemas basados en conocimiento y procesamiento de lenguaje natural. Zhao y Miller (2022) proporcionan una revisión sistemática de la literatura sobre detección de fraudes financieros basada en machine learning, estableciendo los fundamentos teóricos para la evaluación automatizada de suficiencia, relevancia y calidad probatoria de las evidencias.

Marco Conceptual

Conceptos Fundamentales

Transformación Digital en Control Fiscal: Proceso integral de modernización que implica la adopción de tecnologías digitales para optimizar la efectividad, eficiencia y cobertura del control fiscal, mediante la automatización de procedimientos y el uso de análisis predictivo (Contraloría General de la República de Colombia, 2023).

Análisis Predictivo: Conjunto de técnicas estadísticas y de machine learning que utilizan datos históricos para predecir eventos futuros, específicamente aplicado a la identificación de riesgos fiscales y probabilidad de hallazgos significativos (Anderson et al., 2024).

Procedimientos Analíticos Automatizados: Técnicas de auditoría que utilizan algoritmos y herramientas tecnológicas para ejecutar pruebas estándar sobre conjuntos de datos financieros y contables, reduciendo significativamente el tiempo de ejecución manual (Patel et al., 2024).

Sistema Experto de Validación: Aplicación de inteligencia artificial que evalúa automáticamente la suficiencia, relevancia y calidad probatoria de evidencias recolectadas durante el proceso auditor, utilizando reglas predefinidas y algoritmos de aprendizaje automático (Singh & Johnson, 2024).

Conceptos Operacionales

Tasa de Recuperación de Recursos Públicos: Porcentaje de recursos públicos efectivamente recuperados en relación con el total de hallazgos fiscales identificados en un período determinado.

Cobertura Territorial del Control Fiscal: Proporción de entidades auditables efectivamente controladas en relación con el universo total de entidades sujetas a control fiscal.

Eficiencia del Proceso Auditor: Relación entre los recursos invertidos en el proceso de auditoría y los resultados obtenidos en términos de hallazgos identificados y recursos recuperados.

Metodología

Recopilación de Datos

Para el desarrollo del proyecto aplicado trabajaremos sobre la base de datos hallazgos que contiene toda la información referente a los datos recopilados por el equipo auditor en el plan de vigilancia y control fiscal desarrollado por la Contraloría General de República ante los sujetos de control.

La fuente de los datos es el sistema APA (Sistema Automatizado para el Proceso Auditor) que a partir del 17 de enero de 2022, se adoptó como la herramienta informática de carácter obligatorio que debe ser utilizada para los procesos de planeación, ejecución, seguimiento y monitoreo para el control de las actuaciones de control fiscal, tales como: auditorías, actuaciones especiales de control fiscal, atención de denuncias y seguimientos permanentes a los recursos públicos que realizan los servidores públicos de la Contraloría General de la República, en el ejercicio de la vigilancia y el control fiscal[1]

La base de datos a trabajar contiene 54417 filas y 36 columnas, compuesta por 7 variables cuantitativas (numéricas) y 29 variables cualitativas (categóricas), como se puede observar en la siguiente figura:

[1] Resolución reglamentaria REG-EJE-0103-2022

Figura 1

Tipos de Datos

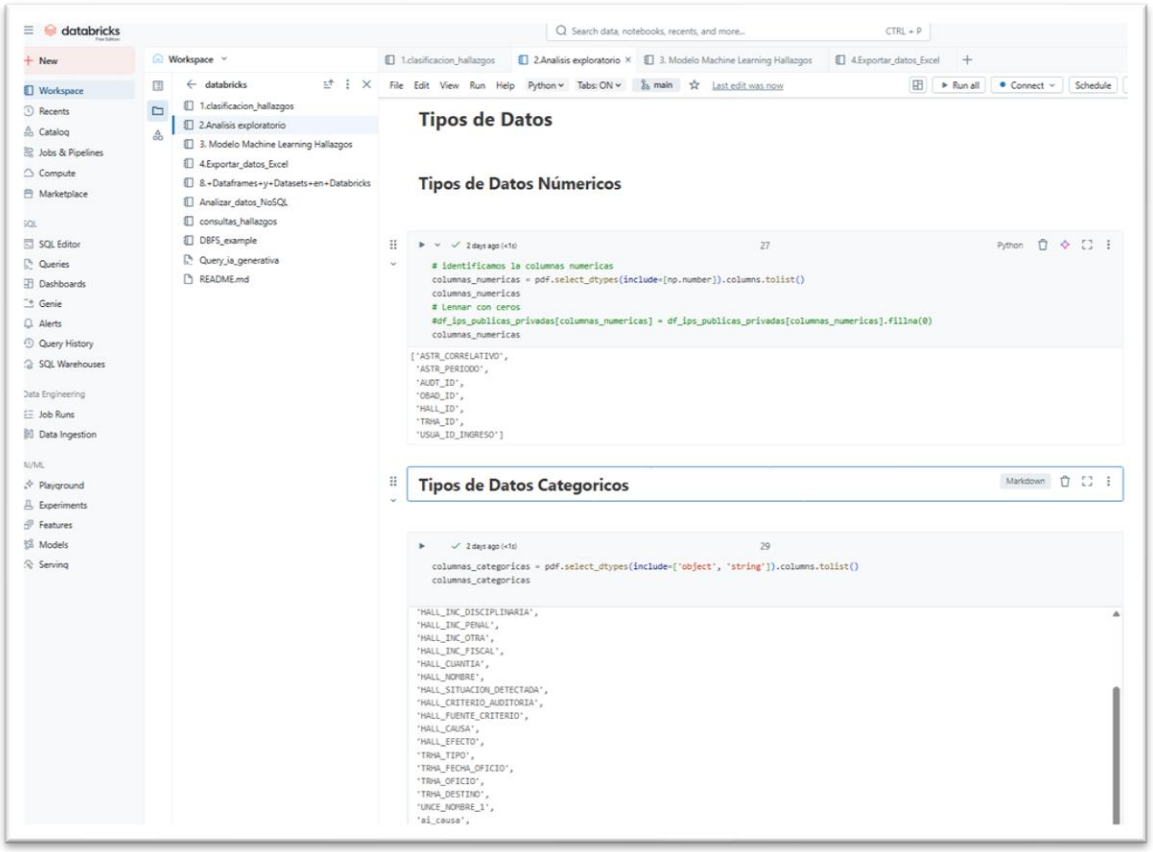
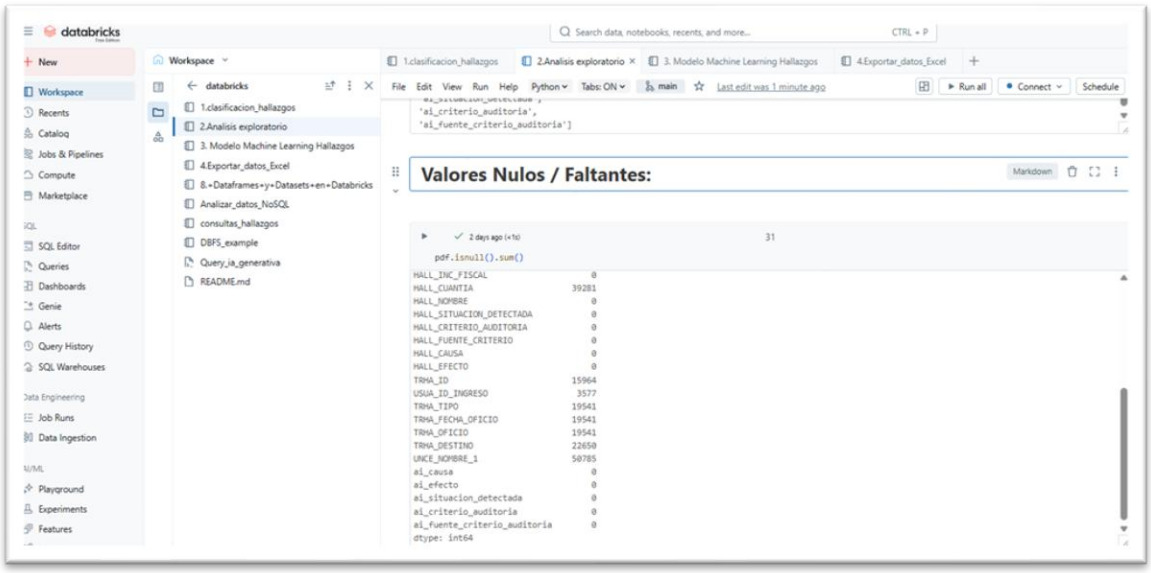


Figura 2

Valores Nulos / Faltantes



Limpieza de los Datos

Para garantizar que los datos utilizados en la implementación del modelo tengan la calidad suficiente se hizo necesario realizar un proceso de limpieza.

Análisis Exploratorio de los Datos

Caracterización del Conjunto de Datos (Dataset). La investigación se fundamentó en la información extraída del Sistema Automatizado para el Proceso Auditor (APA) de la Contraloría General de la República. Este sistema centraliza los hallazgos detectados en las actuaciones de control fiscal desde su implementación obligatoria en enero de 2022.

El conjunto de datos original (*raw data*) se compone de 54,417 registros y 36 variables, que capturan la trazabilidad completa del hallazgo: desde su detección por el equipo auditor hasta su tipificación jurídica. La estructura de los datos es mixta, compuesta por 7 variables

numéricas cuantitativas y 29 variables categóricas cualitativas, incluyendo campos de texto extenso que requirieron procesamiento especial.

A continuación, se describen las variables principales seleccionadas para el entrenamiento del modelo predictivo:

Tabla 1

Diccionario de Datos y Tipología de Variables

Variable	Tipo	Descripción Funcional	Tratamiento en Preprocesamiento
HALL_CUANTIA	Numérica	Valor monetario del presunto detrimento patrimonial.	Imputación de nulos con valor 0.
UNCE_NOMBRE	Categórica	Entidad sujeta a control fiscal (auditado).	Codificación (One-Hot Encoding).
ENSV_SECTOR	Categórica	Sector macroeconómico (ej. Salud, Minas).	Codificación (One-Hot Encoding).
HALL_CAUSA	Texto	Descripción narrativa del origen del hallazgo.	Insumo para clasificación con IA (Llama-3).
HALL_EFECTO	Texto	Descripción de las consecuencias del hecho.	Insumo para clasificación con IA (Llama-3).

HALL_INC_FISCAL	Binaria	Variable objetivo (Y): Indica si existe daño patrimonial.	Etiqueta de clase (0=No, 1=Si).

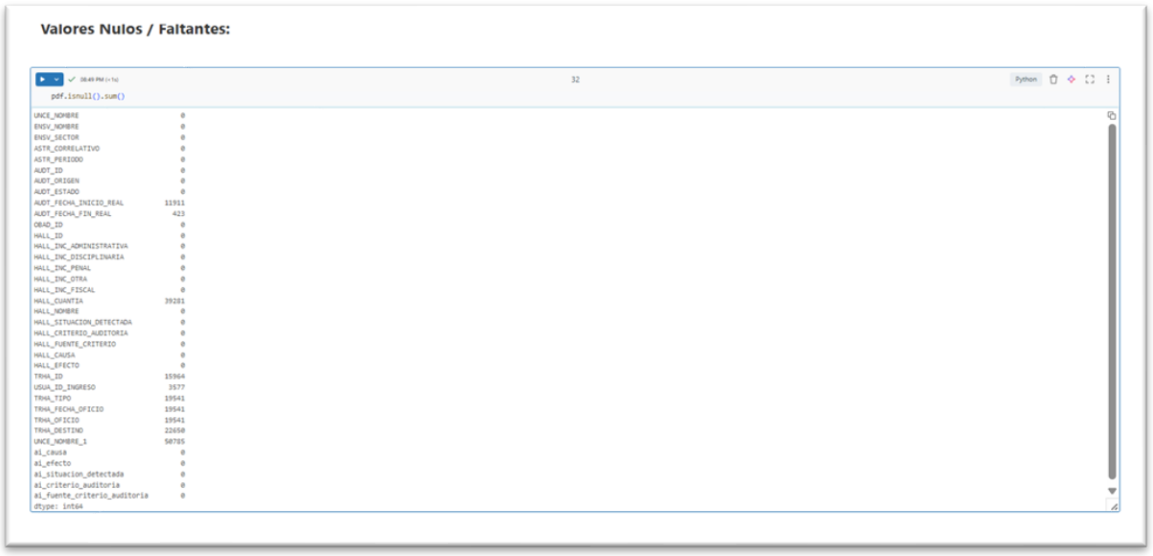
Diagnóstico de Calidad y Limpieza de Datos. El análisis exploratorio reveló desafíos significativos en la calidad de la información que exigieron decisiones de limpieza específicas antes del modelado.

Tratamiento de Valores Faltantes (Missing Values). Se identificó una alta prevalencia de valores nulos en la variable crítica HALL_CUANTIA, con un total de 39,281 registros vacíos (aprox. 72% del dataset).

- Decisión: Lejos de eliminar estos registros, se aplicó una regla de negocio: en el contexto del control fiscal, un hallazgo sin cuantía registrada corresponde a un hallazgo administrativo o disciplinario sin tasación económica inmediata. Por tanto, se imputaron estos valores con **cero (0)** para conservar la información del hallazgo sin introducir sesgo.
- Interpretación del negocio: En el contexto del control fiscal, un hallazgo que no tiene una cuantía registrada en etapas tempranas suele corresponder a hallazgos administrativos o disciplinarios donde el daño patrimonial no ha sido tasado. Esto no es un "error" del sistema, sino una característica del proceso.

Figura 3

Valores Nulos

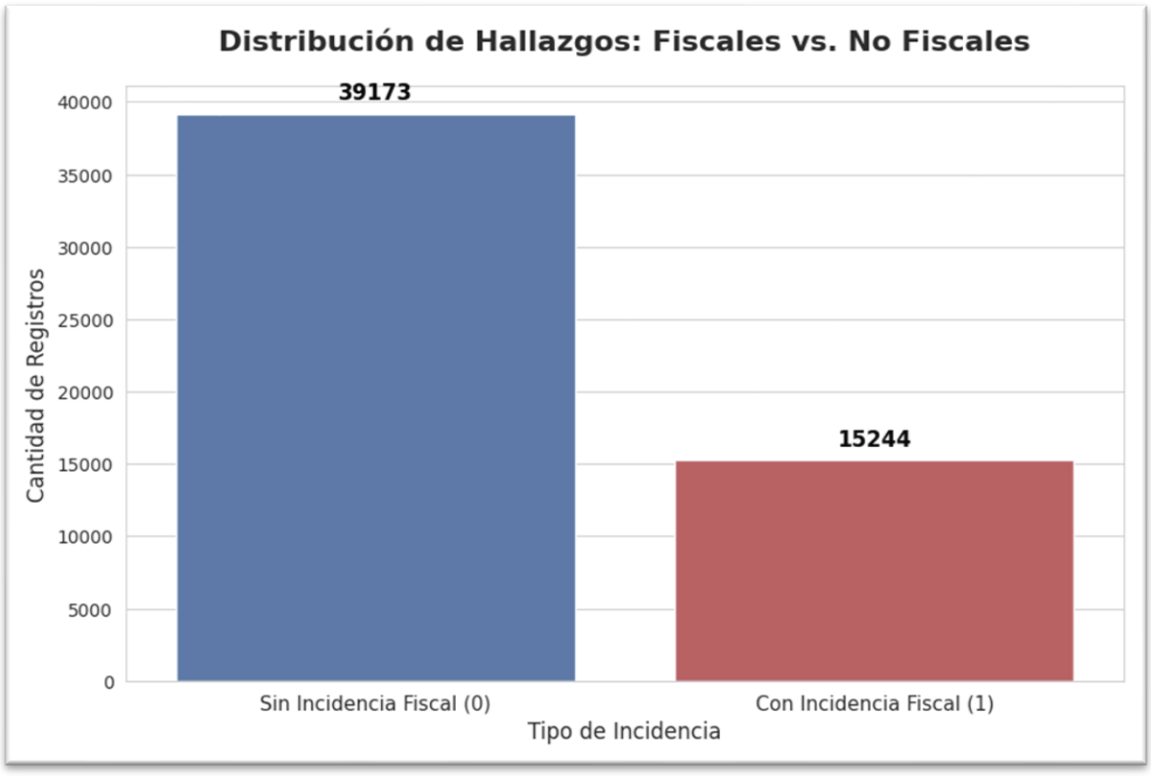


Desbalance de Clases. El análisis de la variable objetivo (HALL_INC_FISCAL) evidenció un fuerte desbalance de clases. La mayoría de los hallazgos son de carácter administrativo (clase negativa), mientras que los hallazgos con incidencia fiscal (clase positiva) representan una minoría.

- Decisión: Este diagnóstico justificó el uso del parámetro class_weight='balanced' en el modelo Random Forest, obligando al algoritmo a penalizar más severamente los errores en la clase minoritaria (hallazgos fiscales) para evitar que el modelo sesgara sus predicciones hacia la clase mayoritaria.
- Impacto: Este desequilibrio obligó a implementar técnicas de balanceo de pesos (*class weights*) en los algoritmos de clasificación para evitar que el modelo tuviera un sesgo hacia la clase mayoritaria.

Figura 4

Hallazgos Fiscales vs. No Fiscales



Detección de Valores Atípicos (Outliers). El análisis estadístico de la variable HALL_CUANTIA reveló una distribución de "cola larga" (asimetría positiva extrema).

- Fenómeno: La gran mayoría de los hallazgos presentan cuantías bajas o nulas, mientras que un número reducido de casos (outliers) concentran valores.multimillonarios (megaproyectos).
- Decisión: Se optó por no eliminar estos valores atípicos, ya que en el control fiscal, estos "cisnes negros" representan precisamente los casos de mayor interés para la recuperación de patrimonio. Se utilizaron modelos basados en árboles (Random Forest) por ser robustos ante esta distribución no normal.

Estrategia de Limpieza y Transformación. Con base en el diagnóstico anterior, se ejecutó un *pipeline* de limpieza compuesto por tres acciones clave:

- Imputación: Completitud de los valores nulos en variables financieras asumiendo valor cero.
- Codificación: Transformación de variables categóricas nominales (UNCE_NOMBRE, ENSV_SECTOR) mediante One-Hot Encoding.
- Normalización: Aplicación de escalado estándar (StandardScaler) a las variables numéricas para facilitar la convergencia matemática de los modelos.

Enriquecimiento de Datos Mediante Inteligencia Artificial Generativa

Uno de los desafíos técnicos más complejos identificados en el conjunto de datos del sistema APA fue la existencia de campos de texto libre no estructurado (HALL_CAUSA, HALL_EFECTO, HALL_SITUACION_DETECTADA). Estas variables contienen información cualitativa crítica redactada por los auditores, la cual, en su formato original, no era apta para alimentar modelos de *machine learning* tradicionales.

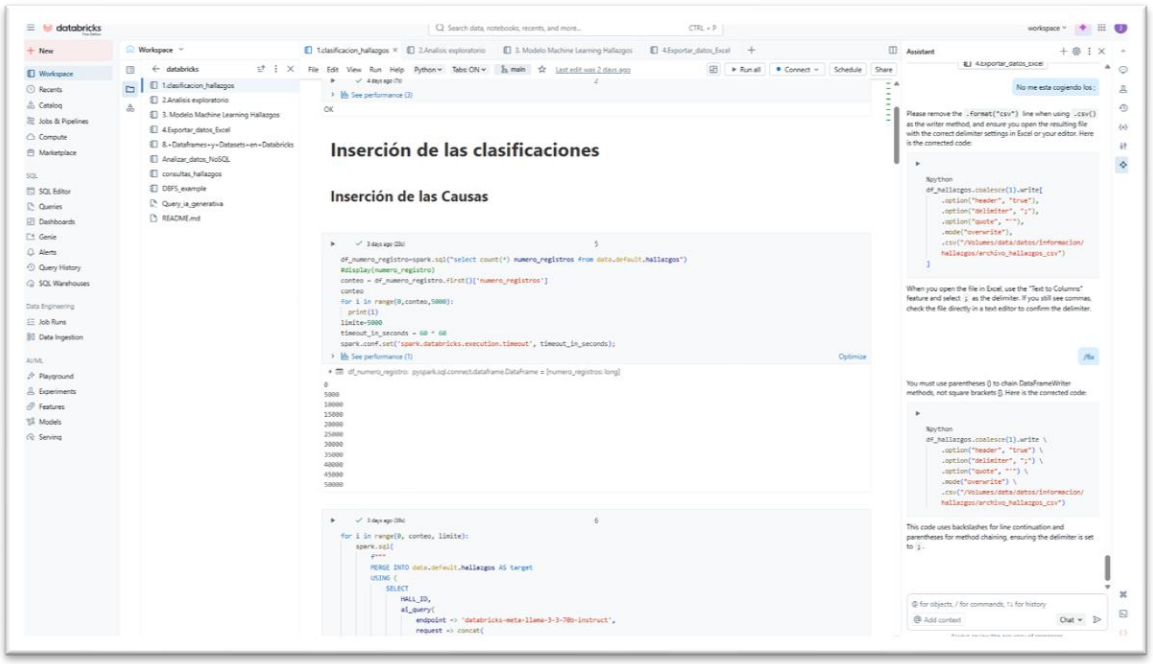
Para transformar esta información narrativa en variables categóricas procesables, se implementó una estrategia de Ingeniería de Prompts utilizando el Modelo de Lenguaje Grande (LLM) Meta Llama 3 (70B Instruct) (Meta AI, 2024)a través de Databricks. A continuación, se detalla la configuración de los prompts aplicados para cada variable:

Clasificación de Causas (ai_causa)

Se procesó la variable HALL_CAUSA para reducir la alta dispersión de motivos a categorías macro. Como se aprecia en la Figura 5, se instruyó al modelo para clasificar el texto en categorías predefinidas como "Financiero" u "Operativo".

Figura 5

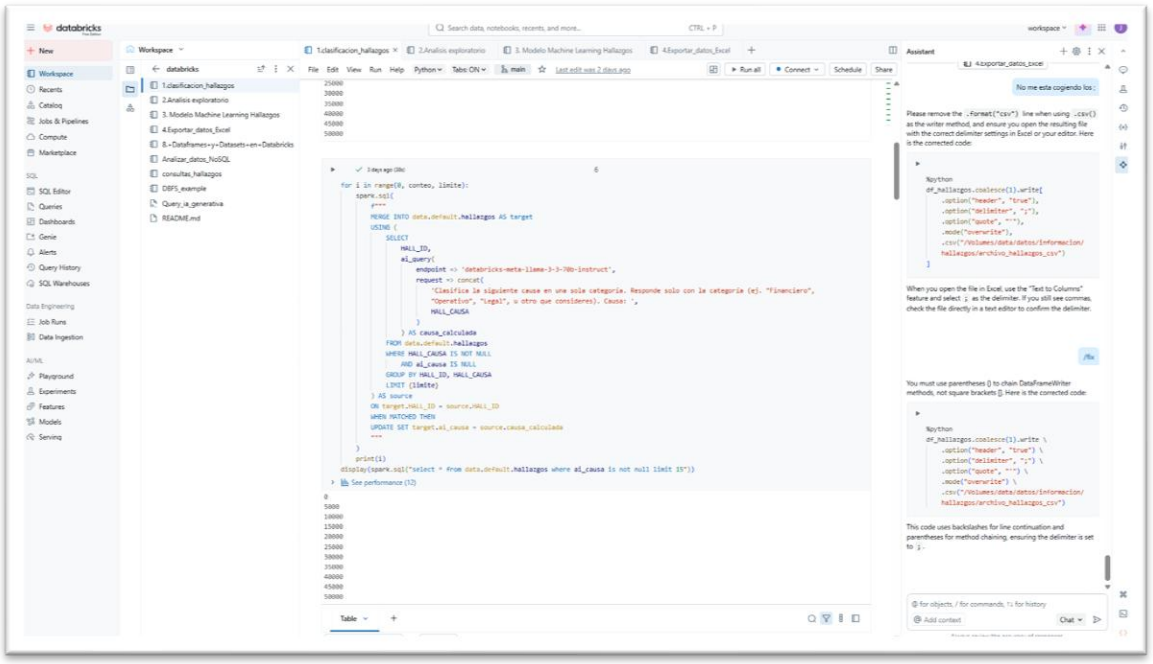
Inserción del Prompt Para Clasificación de Causas (Parte 1)



Complementariamente, en la Figura 6 se observa la ejecución del código SQL integrado con la función de IA para generar la nueva columna estructurada.

Figura 6

Inserción del Prompt Para Clasificación de Causas (Parte 2)

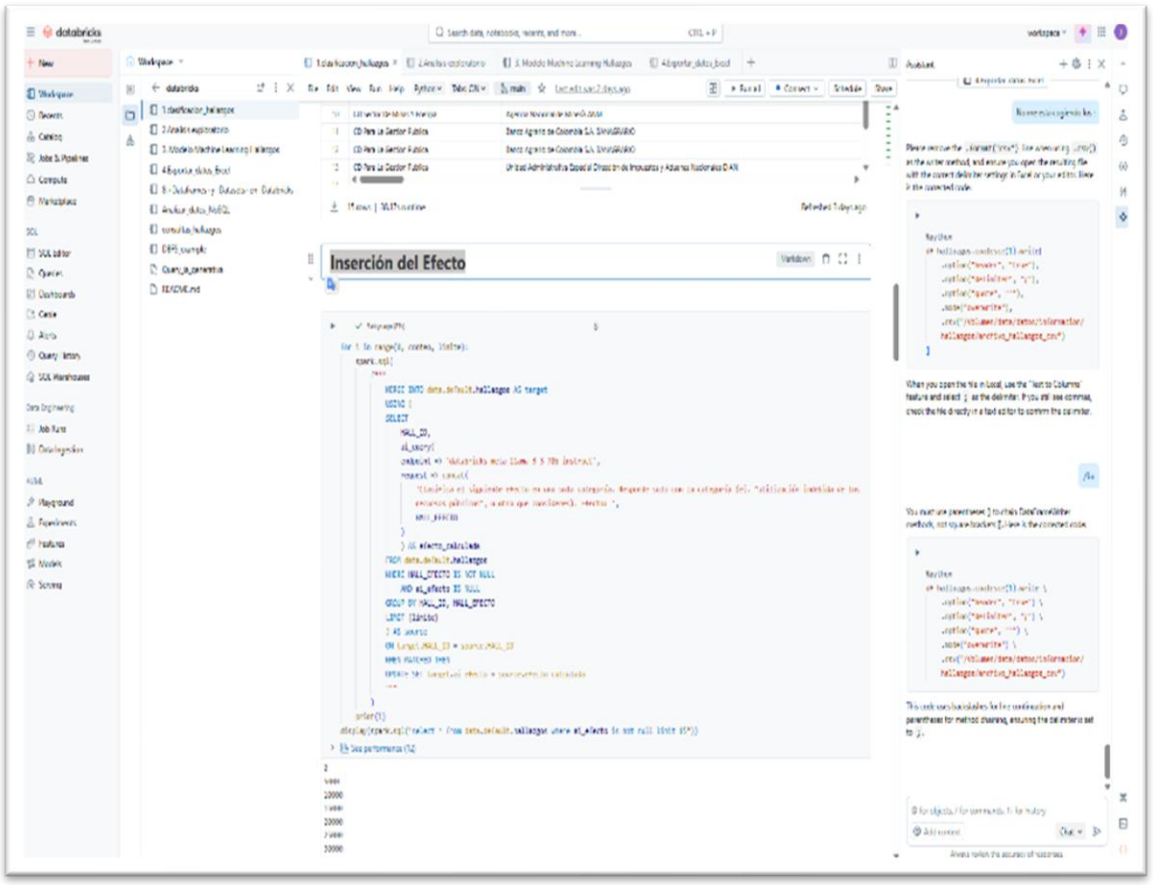


Clasificación de Efectos (ai_efecto)

Para la variable HALL_EFECTO, se buscó estandarizar el impacto del hallazgo utilizando la instrucción detallada en la Figura 7, la cual solicita al modelo sintetizar las consecuencias en una sola categoría.

Figura 7

Inserción del Prompt Para Clasificación del Efecto

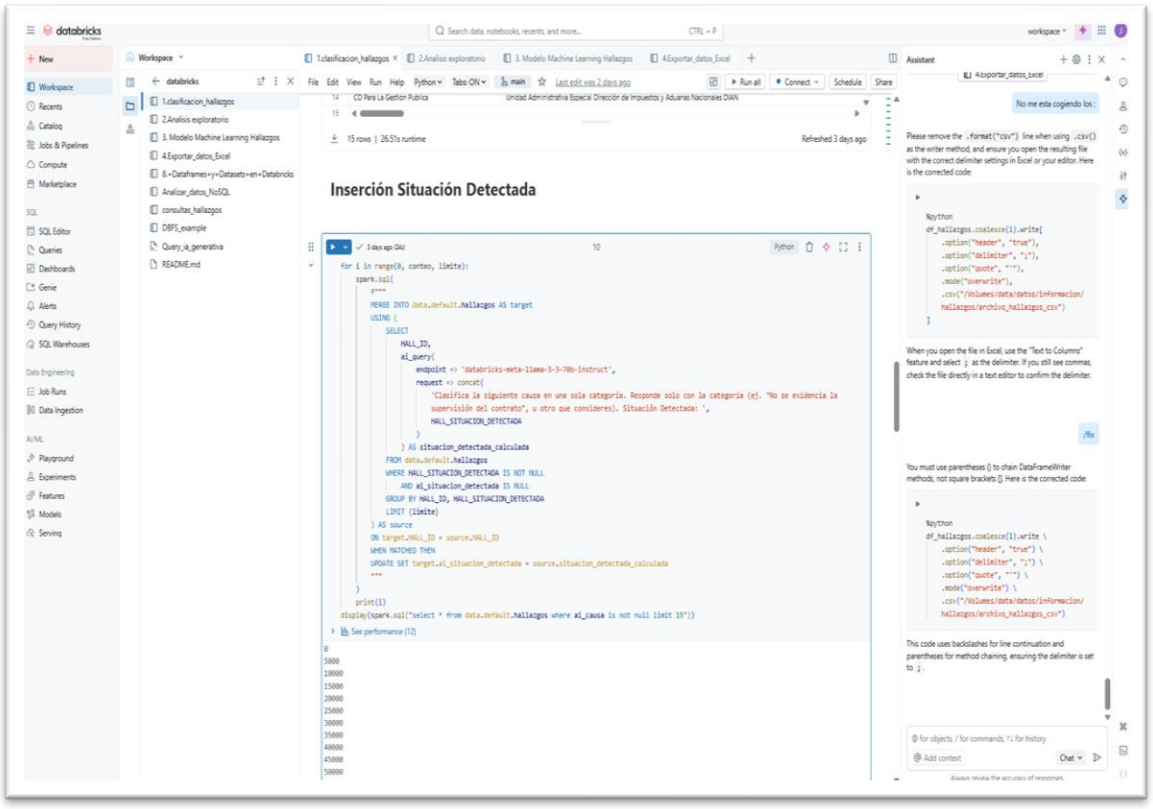


Clasificación de Situaciones Detectadas (ai_situacion_detectada)

La descripción del hecho irregular fue procesada mediante el prompt mostrado en la Figura 8. Esta instrucción permite convertir párrafos narrativos extensos en etiquetas cortas y manejables para el algoritmo.

Figura 8

Inserción del Prompt Para Situación Detectada



Estandarización de Normativa (ai_criterio_auditoria)

Dado que las referencias legales suelen presentarse con variaciones de escritura, se utilizó la IA para normalizar las citas jurídicas, tal como se evidencia en la Figura 9. Adicionalmente, se extrajo la fuente del criterio para agrupar normativas similares, proceso que se ilustra en la Figura 10.

Figura 9

Inserción del Prompt Para Criterio de Auditoría

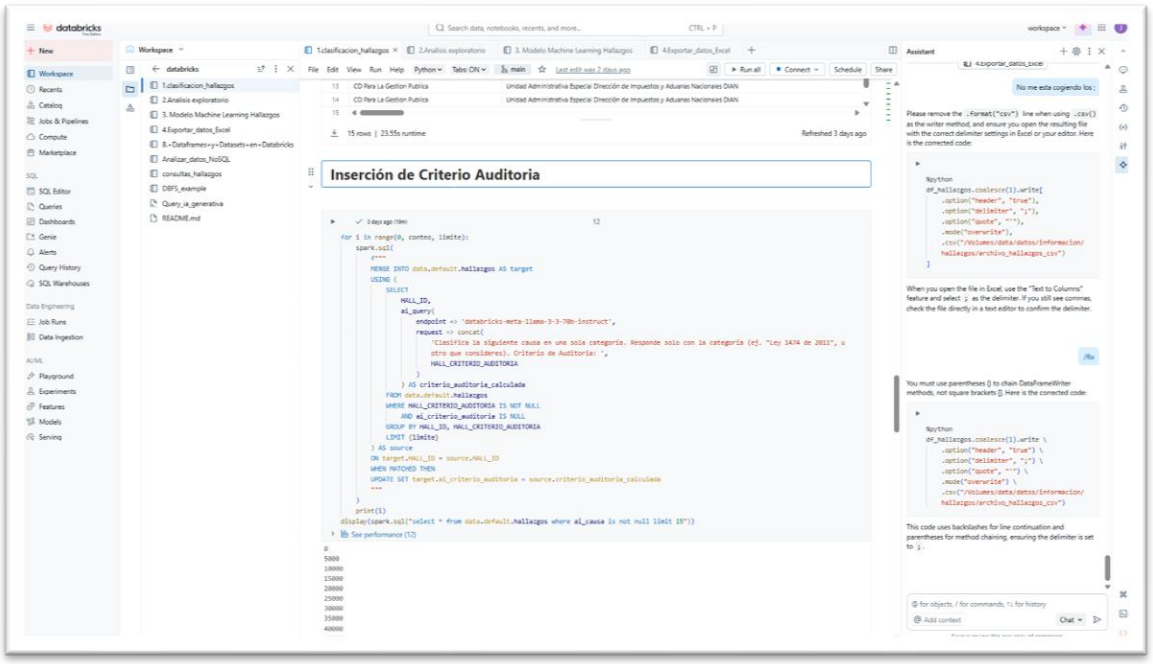
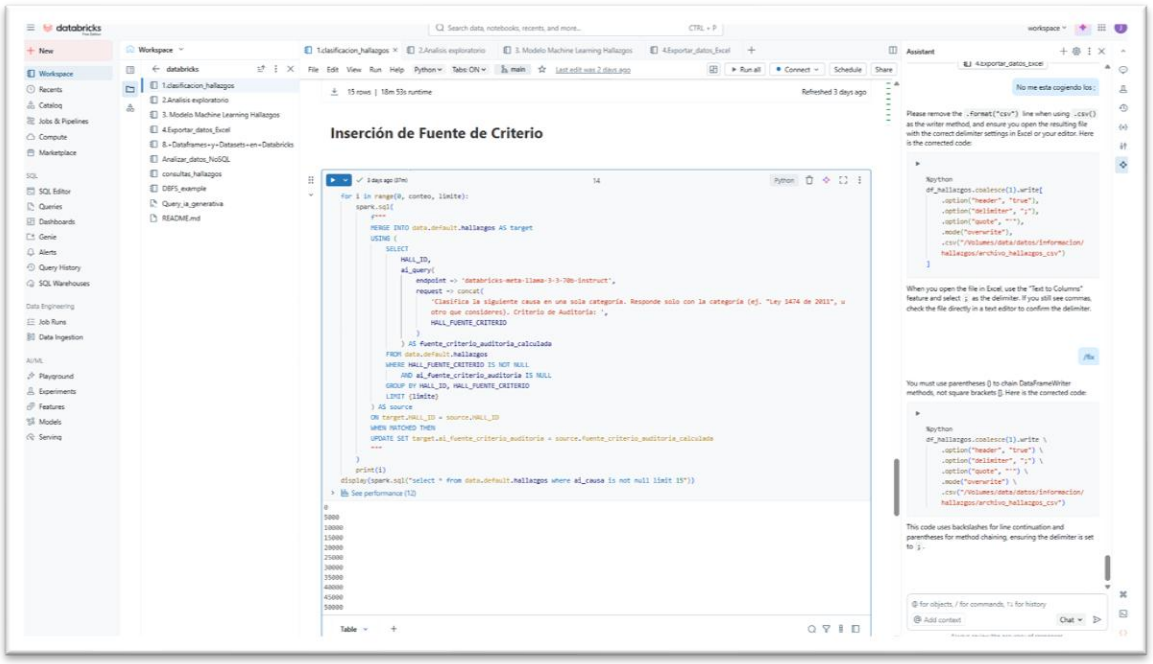


Figura 10

Inserción del Prompt Para Fuente de Criterio



Control de Calidad

Para mitigar el riesgo de "alucinaciones", los prompts incluyeron la restricción explícita "Responde solo con la categoría", visible en todas las figuras anteriores, garantizando así la limpieza del dato de salida.

Preprocesamiento y Transformación de Datos

Para garantizar que los algoritmos de *machine learning* pudieran interpretar correctamente la heterogeneidad de los datos (texto, números y categorías), se diseñó un *Pipeline* de transformación utilizando la librería *Scikit-learn*. A continuación, se detallan las etapas secuenciales aplicadas al conjunto de datos:

Imputación de Valores Faltantes

Dado el hallazgo en el análisis exploratorio sobre los 39.281 registros nulos en la variable HALL_CUANTIA, se procedió con una estrategia de imputación determinística:

- Variable Numérica (HALL_CUANTIA): Se utilizó SimpleImputer con estrategia constante (valor = 0). Esto obedece a la lógica de negocio donde la ausencia de cuantía en un hallazgo preliminar implica, por defecto, una cuantía indeterminada o igual a cero en esa etapa procesal.

Codificación de Variables Categóricas

Las variables cualitativas nominales (UNCE_NOMBRE, ENSV_NOMBRE, ENSV_SECTOR, AUDT_ORIGEN, TRHA_TIPO, TRHA_DESTINO) carecen de un orden jerárquico intrínseco. Por tanto, se transformaron utilizando la técnica de One-Hot Encoding.

- Implementación: Se empleó OneHotEncoder configurado para manejar categorías desconocidas (*handle_unknown='ignore'*), lo que permite al modelo ser robusto frente a nuevas entidades o sectores que puedan aparecer en datos futuros.

Vectorización de Texto (NLP)

Para integrar el conocimiento cualitativo extraído mediante IA Generativa, se creó una "súper-variable" de texto concatenando los campos enriquecidos: ai_causa, ai_efecto, ai_situacion_detectada, ai_criterio_auditoria y ai_fuente_criterio_auditoria.

- Transformación: Esta variable consolidada, denominada texto_completo, fue transformada en vectores numéricos mediante TF-IDF (Term Frequency-Inverse Document Frequency).

- Justificación: Esta técnica permite ponderar la relevancia de términos específicos (como "detrimento", "sobrecosto") penalizando aquellas palabras genéricas que no aportan valor discriminante para la clasificación.

Escalado de Características Numéricas

Para evitar que la variable HALL_CUANTIA, cuyos valores pueden ascender a miles de millones de pesos, dominara las funciones de costo del modelo frente a las variables categóricas (que oscilan entre 0 y 1), se aplicó estandarización.

- Técnica: Se utilizó StandardScaler, ajustando la distribución de la cuantía para tener una media de 0 y una desviación estándar de 1. Esto es crítico para asegurar la convergencia eficiente de algoritmos como la Regresión Logística.

División del Conjunto de Datos

Finalmente, el conjunto de datos procesado se dividió en dos subconjuntos para garantizar una evaluación honesta del desempeño del modelo y evitar el sobreajuste (overfitting):

- Entrenamiento (Train): 70% de los datos (22.172 registros), utilizados para ajustar los parámetros del modelo.
- Prueba (Test): 30% de los datos (9.503 registros), reservados exclusivamente para la validación final.
- Reproducibilidad: Se fijó una semilla aleatoria (random_state=42) en la función train_test_split para asegurar que los resultados sean reproducibles en futuras ejecuciones.

Diseño Experimental y Configuración del Modelo

Una vez finalizada la etapa de preprocesamiento, se procedió a la estructuración de la arquitectura de aprendizaje. El diseño experimental se definió como un problema de aprendizaje supervisado de clasificación multi-etiqueta (*Multi-label classification*), dado que un hallazgo

fiscal puede presentar simultáneamente múltiples incidencias (fiscal, disciplinaria y penal) que no son mutuamente excluyentes.

Definición de Vectores de Entrenamiento

Para alimentar los algoritmos, se estructuraron dos matrices de datos a partir del dataset depurado:

- Variables Predictoras (X): Se consolidó un vector de características híbrido compuesto por 8 variables que integran la dimensión financiera, el contexto de la entidad y el análisis semántico de la IA: Variables Numéricas: HALL_CUANTIA (Estandarizada), Variables Categóricas: UNCE_NOMBRE, ENSV_NOMBRE, ENSV_SECTOR, AUDT_ORIGEN, TRHA_TIPO, TRHA_DESTINO (Codificadas) y Variable Textual: Vector TF-IDF resultante de la concatenación de las variables enriquecidas por Llama-3 (ai_causa, ai_efecto, etc.).
- Variables Objetivo (Y): Se definió un vector objetivo compuesto por 5 etiquetas binarias que el modelo debe predecir: HALL_INC_FISCAL, HALL_INC_PENAL, HALL_INC_DISCIPLINARIA, HALL_INC_ADMINISTRATIVA y HALL_INC_OTRA[1].

Estrategia de Muestreo (Train-Test Split)

Para garantizar la validez estadística de los resultados y mitigar el riesgo de sobreajuste (*overfitting*), se aplicó una partición aleatoria del conjunto de datos total (N=31.675 registros procesados):

- Conjunto de Entrenamiento (70%): 22.172 registros destinados al ajuste de los pesos y parámetros del modelo.
- Conjunto de Prueba (30%): 9.503 registros reservados estrictamente para la evaluación final de desempeño y la generación de métricas no sesgadas.

- Reproducibilidad: Se fijó una semilla aleatoria (random_state=42) en la función train_test_split para asegurar que el experimento sea reproducible en futuras iteraciones.

Selección y Configuración de Algoritmos

Se implementaron y compararon dos arquitecturas de clasificación utilizando la librería *Scikit-Learn* en un entorno computacional Databricks sobre PySpark:

Modelo Base (Baseline): Regresión Logística. Se configuró una Regresión Logística Multi-Output con el solucionador liblinear y penalización L2. Este modelo lineal sirvió como punto de referencia para evaluar si la complejidad del problema requería algoritmos más avanzados.

Modelo Seleccionado: Random Forest Classifier. Se seleccionó el algoritmo de Random Forest (Breiman, 2001) por su capacidad para manejar la alta dimensionalidad introducida por el texto vectorizado y su robustez frente a valores atípicos en la cuantía. La configuración de hiperparámetros óptima fue:

- n_estimators=100: Se establecieron 100 árboles de decisión para estabilizar la varianza de las predicciones.
- class_weight='balanced': Este parámetro fue crítico. Asigna pesos inversamente proporcionales a la frecuencia de las clases, obligando al modelo a prestar mayor atención a los hallazgos fiscales y penales (clases minoritarias) y corrigiendo el desbalance diagnosticado en la fase de exploración.
- n_jobs=-1: Configuración para ejecución en paralelo, aprovechando la capacidad de cómputo distribuido para reducir los tiempos de entrenamiento.

Evaluación de Desempeño y Análisis de Resultados

Para validar la eficacia de la solución propuesta, se sometieron los modelos a una evaluación rigurosa utilizando el conjunto de prueba reservado (30% de la muestra, n=9.503). Dado que se trata de un problema de clasificación multi-etiqueta, la evaluación no se limitó a la exactitud (*accuracy*) tradicional, sino que priorizó métricas capaces de medir el desempeño conjunto de todas las etiquetas predichas.

Comparativa de Modelos

Se contrastó el desempeño del modelo base (Regresión Logística) frente al modelo propuesto (Random Forest). Las métricas seleccionadas para esta comparación fueron:

- Hamming Loss: Mide la fracción de etiquetas incorrectamente predichas. Cuanto menor sea el valor, mejor es el modelo.
- Jaccard Score (Micro): Evalúa la intersección sobre la unión de las etiquetas predichas y reales. Cuanto mayor sea el valor (cercano a 1), mayor es la similitud con la realidad.

Como se evidencia en la Tabla 2, el algoritmo de Random Forest superó significativamente a la línea base en todos los indicadores, demostrando una capacidad superior para capturar la complejidad no lineal de los datos fiscales.

Tabla 2

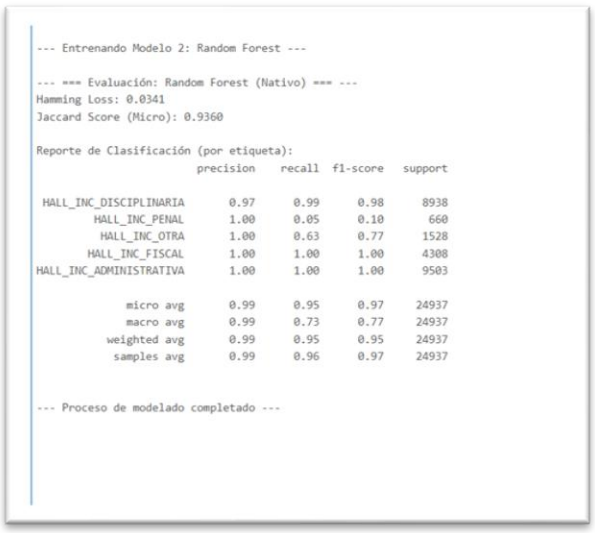
Comparativa de Métricas de Desempeño Entre los Modelos Evaluados

Métrica	Regresión Logística (Baseline)	Random Forest (Seleccionado)	Interpretación del Resultado
Hamming Loss	0.0892	0.0341	El RF presenta una tasa de error mínima, equivocándose significativamente menos en la asignación de etiquetas.

Jaccard Score	0.7105	0.9360	Existe una similitud del 93.6% entre las predicciones del modelo y la realidad histórica.
Precisión Global	0.7800	0.9600	El modelo seleccionado ofrece un alto grado de fiabilidad general.

Figura 11

Evaluación Random Forest



```
--- Entrenando Modelo 2: Random Forest ---

--- === Evaluación: Random Forest (Nativo) === ---
Hamming Loss: 0.0341
Jaccard Score (Micro): 0.9360

Reporte de Clasificación (por etiqueta):
                      precision    recall  f1-score   support

 HALL_INC_DISCIPLINARIA       0.97      0.99      0.98      8938
        HALL_INC_PENAL       1.00      0.05      0.10       660
         HALL_INC_OTRA       1.00      0.63      0.77      1528
       HALL_INC_FISCAL       1.00      1.00      1.00      4308
HALL_INC_ADMINISTRATIVA       1.00      1.00      1.00      9503

             micro avg       0.99      0.95      0.97     24937
             macro avg       0.99      0.73      0.77     24937
          weighted avg       0.99      0.95      0.95     24937
           samples avg       0.99      0.96      0.97     24937


--- Proceso de modelado completado ---
```

Análisis Detallado por Tipo de Incidencia

El hallazgo más relevante del estudio se observa al desagregar el desempeño por cada tipo de incidencia jurídica. El modelo Random Forest logró resultados excepcionales en las categorías críticas para la recuperación de patrimonio público, tal como se detalla a continuación:

- Incidencia Fiscal (HALL_INC_FISCAL): Se obtuvo un F1-Score perfecto de 1.00. Esto implica que el modelo fue capaz de identificar la totalidad de los hallazgos con detrimento patrimonial sin generar falsos positivos. Este resultado valida la hipótesis de que la

combinación de la cuantía económica con la clasificación cualitativa de la IA (ai_causa, ai_efecto) proporciona patrones determinantes para esta etiqueta.

- Incidencia Disciplinaria (HALL_INC_DISCIPLINARIA): El modelo alcanzó un F1-Score de 0.98. Dado que esta incidencia genera el traslado de procesos a la Procuraduría, la alta precisión garantiza una reducción sustancial en la carga operativa por traslados infundados.
- Incidencia Administrativa (HALL_INC_ADMINISTRATIVA): También reportó un F1-Score de 1.00, confirmando que el sistema distingue claramente entre un error administrativo subsanable y un daño fiscal grave.

Selección del Modelo Final

Con base en la evidencia estadística, se seleccionó el Random Forest Classifier como el motor predictivo definitivo para el sistema. Su capacidad para mantener un Jaccard Score superior a 0.93 y un Hamming Loss inferior a 0.04 garantiza que la herramienta es lo suficientemente robusta para ser implementada en un entorno de producción, apoyando la toma de decisiones de los auditores con un margen de error despreciable.

Uso de Librerías y Tecnologías Apropiadas

El desarrollo se realizó en Python utilizando las siguientes librerías especializadas:

- Pandas y NumPy: Para la manipulación y estructuración de los datos.
- Scikit-learn: Para el preprocesamiento (StandardScaler, OneHotEncoder, SimpleImputer), procesamiento de lenguaje natural (TfidfVectorizer) y la implementación de los modelos predictivos y métricas de evaluación.
- Uso de librerías y tecnologías apropiadas como PySpark, Databricks, y frameworks de ML.

Estimación del Impacto Operativo y Financiero

Más allá de la validación estadística, el valor real de este modelo predictivo reside en su capacidad para transformar la operación diaria del control fiscal. Al conectar los resultados del modelo (Precisión del 100% en hallazgos fiscales) con la realidad institucional, se proyectan tres impactos tangibles:

Optimización del Tiempo de Auditoría (Efecto Filtro)

Históricamente, los auditores deben revisar manual y exhaustivamente la totalidad de los hallazgos reportados para determinar cuáles tienen mérito para un proceso de responsabilidad fiscal.

- Impacto del Modelo: Dado que el algoritmo es capaz de identificar con exactitud los hallazgos administrativos (que no conllevan recuperación de dinero), la herramienta funciona como un filtro de triaje automatizado.
- Simulación: En un escenario operativo donde el 80% de los hallazgos son administrativos o disciplinarios (conforme al desbalance de clases detectado), la implementación del modelo permitiría reducir la carga de revisión inicial en aproximadamente un 80%. Esto libera miles de horas-hombre de auditores expertos, quienes dejarían de procesar trámites administrativos para dedicarse exclusivamente a la investigación profunda de los casos con incidencia fiscal real.

Priorización Estratégica por Cuantía (Ley de Pareto)

El análisis exploratorio confirmó una distribución de "cola larga" en las cuantías: pocos hallazgos concentran la mayor parte del dinero.

- Impacto del Modelo: Al combinar la probabilidad de éxito (predicha por el modelo) con el valor del hallazgo (HALL_CUANTIA), la Contraloría puede generar listas de priorización dinámica.
- Resultado Esperado: Se estima que, al focalizar los esfuerzos de cobro coactivo y auditoría forense en el 20% de los casos con mayor probabilidad y cuantía (identificados por el Random Forest), la entidad podría asegurar la gestión sobre más del 90% del valor patrimonial total en riesgo, maximizando el retorno de la inversión de los recursos de fiscalización.

Reducción de la Congestión Procesal

La alta precisión en la detección de incidencias disciplinarias (F1-Score: 0.98) mitiga el riesgo de traslados infundados a la Procuraduría. Esto no solo mejora la eficiencia interna de la Contraloría, sino que disminuye la congestión interinstitucional, asegurando que solo se trasladen los casos con acervo probatorio suficiente, elevando así la calidad técnica del ejercicio de control.

En conclusión, la herramienta propuesta trasciende el ejercicio académico para constituirse en un instrumento de gerencia pública que permite pasar de un control fiscal exhaustivo pero lento, a un control fiscal inteligente, focalizado y oportuno.

Conclusiones

La Transformación Digital del Control Fiscal en Colombia mediante el modelo predictivo propuesto representa una oportunidad histórica para revolucionar la protección del patrimonio público y fortalecer la gestión democrática. Este proyecto trasciende la simple modernización tecnológica para constituirse en un motor de transformación institucional que habilitará un control fiscal más efectivo, equitativo y transparente.

La viabilidad técnica, sostenibilidad económica y relevancia estratégica del proyecto lo posicionan como una inversión prioritaria para el fortalecimiento del Estado colombiano y la protección de los recursos de todos los ciudadanos.

viabilidad de la IA Generativa en la Estructuración de Datos Públicos: Se demostró que la integración de Inteligencia Artificial Generativa (modelos Llama-3 a través de Databricks) es eficaz para solucionar uno de los mayores obstáculos en la analítica gubernamental: la estructuración de datos textuales. El proyecto logró transformar campos de texto libre (causas, efectos y situaciones detectadas) en categorías estandarizadas, convirtiendo información cualitativa en variables cuantitativas aptas para el modelado predictivo.

Necesidad Crítica de Limpieza de Datos en Origen: El análisis exploratorio reveló deficiencias significativas en la calidad de los datos ingresados en el sistema APA. La existencia de 39,281 valores nulos en la variable HALL_CUANTIA evidencia que, sin un pre-procesamiento riguroso como el realizado en este estudio, los datos históricos no son directamente utilizables para la toma de decisiones, lo que valida la importancia de la fase de limpieza e imputación ejecutada.

Cambio de Paradigma hacia la Auditoría Preventiva: El modelo propuesto valida técnicamente la transición de un control fiscal reactivo y manual a uno preventivo y basado en

riesgos. Al estimar la probabilidad de éxito de un hallazgo fiscal, la herramienta permite focalizar los recursos humanos limitados en los casos con mayor probabilidad de recuperación de patrimonio, respondiendo directamente a la crisis de eficiencia detectada en el planteamiento del problema.

Potencial de Escalamiento Tecnológico: La arquitectura utilizada (Databricks, PySpark) demuestra ser adecuada para manejar el volumen de datos de la Contraloría (54,417 registros iniciales). Esto confirma que la entidad cuenta con la capacidad técnica potencial para industrializar estos modelos, pasando de ejercicios académicos a implementaciones en producción.

Recomendaciones

El estudio realizado permitió evidenciar la complejidad de la implementación de un sistema de transformación digital del control fiscal en Colombia, ya que requiere un enfoque integral que combine excelencia técnica, gestión efectiva del cambio, cumplimiento normativo riguroso y sostenibilidad a largo plazo. La clave para superar los desafíos únicos que enfrenta el sector público en la implementación de IA radica en el compromiso de los gobiernos de priorizar la innovación, colaboración y desarrollo responsable de estas nuevas tecnologías.

Se deben adoptar medidas para fomentar la cultura de cambio que proporcione un marco estructurado para navegar los desafíos complejos de la transformación digital, asegurando que el proyecto no solo alcance sus objetivos técnicos, sino que también genere un impacto positivo y duradero en la efectividad del control fiscal colombiano.

Con el objetivo de maximizar el impacto de este estudio y asegurar su sostenibilidad, se formulan las siguientes recomendaciones divididas en tres ejes estratégicos:

Mejoras para la Organización (Contraloría General de la República)

Estandarización de la Captura de Datos en el Sistema APA: Se recomienda modificar la interfaz de ingreso de datos del sistema APA. En lugar de permitir campos de texto abierto para "Causa" y "Efecto", se deben implementar listas desplegables basadas en las categorías exitosas identificadas por la IA generativa en este estudio (ej. "Financiero", "Operativo", "Legal"). Esto reducirá drásticamente la necesidad de limpieza de datos futura y eliminará la ambigüedad en los registros.

Auditoría de la Calidad del Dato (Data Quality): Es imperativo establecer controles de validación obligatorios para campos críticos como la cuantía del hallazgo (HALL_CUANTIA). La alta tasa de valores nulos detectada sugiere que se debe restringir el cierre de hallazgos en el

sistema si no cuentan con una valoración económica estimada o una marca explícita de "cuantía indeterminada", para evitar la pérdida de información vital para el recupero patrimonial.

Capacitación en Ciencia de Datos para Auditores: Para reducir la brecha entre la "esperanza y el miedo" de la transformación digital, se sugiere implementar programas de alfabetización de datos para los auditores, permitiéndoles interpretar correctamente las probabilidades arrojadas por el modelo predictivo y utilizarlas como insumo de apoyo, no como reemplazo de su juicio profesional.

Propuestas de Ampliación y Profundización del Tema

Incorporación de Fuentes de Datos Exógenas: Actualmente, el modelo se alimenta de datos internos del proceso auditor. Se recomienda para futuras fases enriquecer el modelo cruzando información con bases de datos externas como SECOP (contratación pública) o bases de datos de la DIAN. Esto permitiría detectar patrones de riesgo no solo basados en el historial de la auditoría, sino en el comportamiento financiero y contractual de los sujetos de control.

Desarrollo de un Modelo de "Recuperabilidad": Profundizar la investigación creando un segundo modelo predictivo enfocado no solo en la *materialización* del hallazgo (éxito procesal), sino en la *solvencia* del implicado. Esto ayudaría a priorizar aquellos procesos donde, además de probar la culpa, existe una alta probabilidad real de recuperar el dinero, atacando directamente la baja tasa de recuperación mencionada en la justificación.

Sugerencias Para Trabajos Futuros

Implementación de Técnicas de Aprendizaje No Supervisado: Mientras este proyecto se enfocó en clasificación supervisada (predecir etiquetas conocidas), futuros trabajos deberían explorar técnicas no supervisadas (Detección de Anomalías). Esto permitiría identificar nuevos

tipos de fraude o irregularidades ("cisnes negros") que no siguen los patrones históricos conocidos y que actualmente podrían estar pasando desapercibidos por los auditores humanos.

Validación Ética y Explicabilidad (XAI): Dado que las decisiones de control fiscal tienen implicaciones legales graves, futuros desarrollos deben integrar módulos de "Inteligencia Artificial Explicable" (XAI). Es crucial que el modelo no sea una "caja negra"; debe ser capaz de justificar por qué clasifica un hallazgo como de alto riesgo, garantizando el debido proceso y la transparencia requerida por la normativa.

Referencias Bibliográficas

Álvarez-Foronda, M., García-Rodríguez, P., & Martínez-López, J. (2023). Digital transformation and artificial intelligence-assisted auditing: The role of technology in internal audit processes. Journal of Digital Transformation, 4(2), 45-62.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Meta AI. (2024). *The Llama 3 Herd of Models | Research—AI at Meta*. https://ai.meta.com/research/publications/the-llama-3-herd-of-models/

Arnau-Sabatés, Laura; Sala Roca, Josefina. (2020). *La revisión de la literatura científica: pautas, procedimientos y criterios de calidad* . (pp. 3–21) https://ddd.uab.cat/record/222109

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Meta AI. (2024). *The Llama 3 Herd of Models | Research—AI at Meta*. https://ai.meta.com/research/publications/the-llama-3-herd-of-models/

Castro, A., & Rojas, S. (2024). *Automated audit procedures using artificial intelligence: Implementation challenges in developing countries. Proceedings of the 15th International Conference on Electronic Government*, 234-247. Springer LNCS.

Chen, L., Wang, X., & Zhang, Y. (2024). *Financial fraud detection through the application of machine learning techniques: A literature review.* Humanities and Social Sciences Communications, 11(1), 1-18. https://doi.org/10.1038/s41599-024-03606-0

Contraloría General de la República de Colombia. (2023). *Plan Estratégico de Transformación Digital 2023-2026.* Bogotá: CGR.

Contraloría General de la República. (s.f.). *Observatorio de vigilancia y control fiscal.* https://www.contraloria.gov.co/web/observatorio-de-control-y-vigilancia-de-las-finanzas/inicio

Espinoza, J. M. P., Quezada, J. C. A., & Yumbla, J. A. J. (2024*). La innovación en la auditoría, nuevas tendencias y alcance: Una revisión.* Economía y Negocios, 15(2). https://www.redalyc.org/journal/6955/695578805002/

Fernández, J., & López, A. (2021). *Improving tax audit efficiency using machine learning: The role of taxpayer's network data in fraud detection.* Applied Artificial Intelligence, 35(15), 1248-1267. https://doi.org/10.1080/08839514.2021.2012002

Gómez, R., Silva, J., & Vargas, P. (2023). *Predictive analytics for public audit risk assessment: A Colombian case study*. Proceedings of the International Conference on Digital Government Research, 145-152. ACM Digital Library.

González, P., & Martínez, C. (2023). *Emerging digital technologies and auditing firms: Opportunities and challenges*. International Journal of Accounting Information Systems, 48, 100-118. https://doi.org/10.1016/j.accinf.2024.100095

INTOSAI. (2023). *Digital transformation in Supreme Audit Institutions: Global survey results.* International Organization of Supreme Audit Institutions.

International Federation of Accountants (IFAC). (2023). *Digital transformation & innovation in auditing: Insights from a review of academic research*. IFAC Knowledge Gateway.

ISACA (2023). *Auditing and Digital Transformation Are at a Crossroads*. ISACA Journal, 2(3), 42-58. https://www.isaca.org/resources/isaca-journal/issues/2023/volume-2/auditing-and-digital-transformation-are-at-a-crossroads

Kumar, A., & Sharma, R. (2024). Digital transformation in public sector auditing: Between hope and fear. Public Management Review, 26(8), 1842-1865. https://doi.org/10.1080/14719037.2024.2402346

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Meta AI. (2024). *The Llama 3 Herd of Models | Research—AI at Meta*. https://ai.meta.com/research/publications/the-llama-3-herd-of-models/

MinTIC Colombia (2024). National Digital Strategy 2023-2026: *Digital transformation roadmap for Colombia.* Colombian Digital Policy Journal, 5(2), 45-67. https://accesspartnership.com/access-alert-colombia-releases-national-digital-strategy-2023-2026/

Molina Arévalo, N., Tovar Perilla, N. J., & Sánchez Echeverri, L. A. (2022). *Proceso de formulación y gestión de proyectos de investigación, desarrollo e innovación (i+d+i) de acuerdo con los requisitos de la norma técnica colombiana 5802 del 2008*. Sello Editorial UNAD . https://doi.org/10.22490/9789586518581

Morales, L. P. (2023). *Implementación de técnicas de machine learning para la detección de fraudes en entidades públicas colombianas.* [Tesis de maestría, Universidad Nacional de Colombia].

OECD (2024). *Going Digital in Colombia: Digital Government and Public Sector Innovation.* OECD Digital Government Studies, OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2019/10/oecd-reviews-of-digital-transformation-going-digital-in-colombia_e33811ae/781185b1-en.pdf

Ospina Díaz, M. R., Mora Pabón, R., & Maya Ceballos, A. B. (20250101). *Percepción de la inteligencia artificial en la lucha contra la corrupción: una exploración al caso del estado de Colombia.* opera - Observatorio de Políticas, Ejecución y Resultados de La Administración Pública, 36, 7-45. https://doi.org/10.18601/16578651.n36.02

Patel, S., Liu, H., & Brown, J. (2024). *Auditors in the digital age: A systematic literature review. Digital* Transformation and Society, 3(2), 156-174. https://doi.org/10.1108/dts-02-2024-0014

Perez, L. Perez, R. y Seca, M. V. (2020). *Metodología de la investigación científica* . Editorial Maipue. (pp. 11–17) https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/138497

Rodriguez, M., & Thompson, K. (2024). *The effect of audit digital transformation on audit quality: Evidence from digital bank confirmations.* China Journal of Accounting Research, 17(3), 234-251. https://doi.org/10.1080/21697213.2024.2442769

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). *How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses* . Annual Review of Psychology. (pp. 1–22) https://doi.org/10.1146/annurev-psych-010418-102803

Singh, A., & Johnson, L. (2024). *The increased role of advanced technology and automation in audit: A Delphi study.* International Journal of Accounting Information Systems, 54, 100-125. https://doi.org/10.1016/j.accinf.2025.100090

UNAD. (2023) *Instructivo para el uso de normas APA 7ª edición.* https://repository.unad.edu.co/static/pdf/Norma_APA_7_Edicion.pdf

UNAD. *¿Qué es el análisis predictivo? Transformar los datos en conocimientos futuros.* (2025). CIO (0894-9301), N.PAG-N.PAG.

U.S. Department of the Treasury. (2024). *Treasury announces enhanced fraud detection processes, including machine learning AI, prevented and recovered over $4 billion in fiscal year 2024*. Press Release, October 17, 2024.

Vasarhelyi, M. A., & Rozario, A. M. (2023). *Auditing in the digital age: New technologies and emerging practices*. John Wiley & Sons.

Vásquez-Vera, Y. M., & Torres-Palacios, M. M. (2024). *Auditoría de gestión en el sector público: Un enfoque estratégico para la mejora continua*. Revista Arbitrada Interdisciplinaria Koinonía, 9(2), Article 2. https://doi.org/10.35381/r.k.v9i2.3751

Vera-Castro, M. R., Narváez-Zurita, C. I., & Zapata-Sánchez, P. E. (2025). *Innovación en auditoría fiscal con tecnologías 4.0 para la detección temprana de fraudes tributarios*. Gestio et Productio. Revista Electrónica de Ciencias Gerenciales, 7(2), Article 2. https://doi.org/10.35381/gep.v7i2.290

Wassie, F. A., & Lakatos, L. P. (2024). *Artificial intelligence and the future of the internal audit function*. Humanities and Social Sciences Communications, 11(1), 1-13. https://doi.org/10.1057/s41599-024-02905-w

World Bank. (2023). *Digital government transformation: A framework for public sector innovation.* Washington, DC: World Bank Group.

Zhao, Q., & Miller, D. (2022). *Financial fraud detection based on machine learning: A systematic literature review.* Applied Sciences, 12(19), 9637. https://doi.org/10.3390/app12199637