

MODELO PREDICTIVO DE RIESGO DE CLAUSURA DEFINITIVA PARA LA FISCALIZACIÓN MUNICIPAL INTELIGENTE BASADA EN EVIDENCIA

CLAUDIA EDITH
ALARCÓN DENEN
Maestría Ejecutiva en
Transformación Digital
Escuela de Posgrado
Universidad Nacional De
Ingeniería
Lima – Perú
claudia.alarcon.d@uni.pe

ALVARO ANTONIO
ORIHUELA PASTOR
Maestría Ejecutiva en
Transformación Digital
Escuela de Posgrado
Universidad Nacional De
Ingeniería
Lima – Perú
alvaro.oriuela.p@uni.pe

ELMER
IBÁRCENA RIVERA
Maestría Ejecutiva en
Transformación Digital
Escuela de Posgrado
Universidad Nacional De
Ingeniería
Lima – Perú
elmer.ibarcena.r@uni.pe

FRANK JOE
GASPAR ÑAÑA
Maestría Ejecutiva en
Transformación Digital
Escuela de Posgrado
Universidad Nacional De
Ingeniería
Lima – Perú
frank.gaspar.n@uni.pe

Curso de Valor de la Tecnología de la Información de la Maestría Ejecutiva en Transformación Digital –
Universidad Nacional de Ingeniería, ciclo 2025-2

Abstract— La fiscalización municipal enfrenta crecientes desafíos para garantizar el cumplimiento normativo, la seguridad ciudadana y el uso eficiente de los recursos públicos. En muchos gobiernos locales, los procesos de inspección se ejecutan de forma reactiva y poco focalizada, careciendo de herramientas analíticas que permitan anticipar riesgos críticos. Este estudio presenta el diseño, implementación y evaluación de un modelo predictivo de clasificación orientado a estimar el riesgo de clausura definitiva de establecimientos comerciales, como mecanismo de apoyo a la fiscalización inteligente basada en evidencia. El trabajo adopta la metodología CRISP-DM como marco estructural, asegurando trazabilidad, rigor metodológico e interpretabilidad. Se evaluaron distintos algoritmos de *Machine Learning* (árboles de decisión, Random Forest, Naive Bayes y regresión logística regularizada) utilizando validación cruzada estratificada. Los resultados evidencian que los modelos explicables, en particular el árbol J48 y la regresión logística regularizada, alcanzan un alto desempeño predictivo ($\text{recall} \geq 0.97$ para la clase crítica) con una baja tasa de falsos negativos, constituyéndose en alternativas robustas y defendibles para su adopción en contextos de gestión pública.

Keywords— *Fiscalización inteligente, gobierno digital, CRISP-DM, machine learning, clausura definitiva, analítica predictiva.*

1. INTRODUCCIÓN

La fiscalización municipal constituye un instrumento esencial para asegurar el cumplimiento de las normas urbanas, sanitarias y de seguridad asociadas al funcionamiento de los establecimientos comerciales. Sin

embargo, la limitada capacidad operativa y la creciente complejidad del entorno urbano obligan a los gobiernos locales a priorizar sus acciones de control de manera estratégica.

Tradicionalmente, la fiscalización se ha sustentado en enfoques reactivos, activados tras la ocurrencia de incidentes o infracciones reiteradas. Este paradigma presenta limitaciones significativas en términos de eficiencia, prevención y asignación óptima de recursos. En este contexto, la analítica predictiva y el *Machine Learning* emergen como herramientas clave para fortalecer la toma de decisiones públicas basada en evidencia.

El presente artículo describe el desarrollo de un modelo predictivo de riesgo de clausura definitiva, concebido como un instrumento de apoyo a la fiscalización municipal inteligente. El estudio se orienta a responder la siguiente pregunta de investigación: **¿es posible anticipar, con suficiente precisión e interpretabilidad, los establecimientos con alto riesgo de clausura definitiva utilizando datos administrativos históricos?**

2. ESTADO DEL ARTE

En los últimos años, el uso de técnicas de *Machine Learning* (ML) para apoyar procesos de fiscalización, inspección y control en el sector público ha experimentado un crecimiento significativo, impulsado por la necesidad de optimizar recursos escasos y transitar desde enfoques reactivos hacia modelos preventivos y basados en riesgo. En este contexto, diversos estudios han demostrado la eficacia de los enfoques predictivos para priorizar intervenciones, reducir costos

operativos y aumentar la detección de incumplimientos graves.

Chapman et al. (2000). Mediante CRISP-DM, propone una metodología estándar e independiente de herramientas que estructura los proyectos de minería de datos en seis fases iterativas, desde la comprensión del negocio hasta el despliegue. define tareas, resultados esperados y puntos de decisión, lo que asegura trazabilidad, repetibilidad y alineación entre métricas técnicas y objetivos de negocio. Por su flexibilidad y enfoque práctico, CRISP-DM se ha consolidado como la referencia de facto para proyectos de analítica avanzada y análisis predictivo en múltiples sectores, incluido el ámbito público. revisa las principales familias de técnicas (árboles de decisión, regresión logística, k-means, reglas de asociación, entre otras) desde una perspectiva conceptual, priorizando la interpretación de los resultados y su conexión con decisiones concretas.

Provost y Fawcett (2013). Desarrollan el concepto de *pensamiento data-analítico* como base para aplicar la ciencia de datos a la toma de decisiones de negocio, mostrando cómo traducir problemas empresariales en tareas de minería de datos como clasificación, regresión, segmentación y detección de anomalías. La obra revisa modelos clave de *machine learning* —árboles de decisión, regresión logística, k-means y reglas de asociación— desde un enfoque conceptual orientado a la interpretación y al impacto en métricas de negocio. Se ha consolidado como una referencia fundamental para alinear modelos analíticos con la estrategia organizacional y mejorar la comunicación entre científicos de datos y tomadores de decisión.

Raschka y Mirjalili (2019). presentan una obra aplicada y rigurosa sobre *machine learning* y *deep learning* en Python, utilizando scikit-learn y TensorFlow, que cubre todo el ciclo de desarrollo de modelos predictivos. El libro aborda modelos clásicos como regresión logística, árboles de decisión, métodos de ensamble y clustering, junto con técnicas avanzadas de reducción de dimensionalidad y validación. Asimismo, profundiza en redes neuronales profundas, incluyendo redes convolucionales (CNN), recurrentes (RNN) y redes generativas adversarias (GAN), vinculando teoría, buenas prácticas y una implementación reproducible orientada a aplicaciones reales.

Kelleher, Namee y D'Arcy (2020) ofrecen una síntesis sistemática de los fundamentos del machine learning para analítica predictiva, cubriendo desde la formulación de problemas hasta la evaluación de modelos en contextos reales. La obra describe en detalle algoritmos supervisados y no supervisados (como regresión, árboles de decisión, k NN, redes neuronales y clustering), acompañados de ejemplos trabajados y estudios de caso que ilustran su aplicación en

distintos dominios. Además, enfatiza el papel de la ingeniería de características, la validación cruzada y la elección de métricas alineadas con los objetivos de negocio, proporcionando un marco conceptual robusto para el diseño y análisis de soluciones predictivas.

Sukhani et al. desarrollan un enfoque de analítica predictiva que integra datos de inspecciones sanitarias de restaurantes con reseñas en línea para anticipar riesgos de inocuidad alimentaria, siguiendo la línea de trabajos que explotan fuentes como Yelp para predecir violaciones sanitarias. Su propuesta combina variables estructuradas de inspección (tipo de violación, historial de sanciones) con características textuales derivadas del sentimiento y contenido de las reseñas, generando modelos de clasificación capaces de identificar establecimientos con mayor probabilidad de incumplimiento. Este enfoque evidencia el potencial de fusionar datos oficiales y de plataformas digitales para apoyar esquemas de inspección basados en riesgo en el sector de alimentos.

Ippolito y Lozano (2020) presentan un estudio de caso sobre predicción de delitos tributarios relativos al impuesto sobre servicios en el municipio de São Paulo, utilizando técnicas de machine learning para apoyar la planificación de auditorías fiscales. Los autores construyen un pipeline que incluye selección de características, partición temporal de datos y entrenamiento de varios modelos supervisados, destacando el desempeño superior de Random Forest frente a otros algoritmos al predecir contribuyentes con alta probabilidad de cometer delitos contra el sistema tributario. Sus resultados muestran la viabilidad de integrar modelos predictivos en procesos de fiscalización para priorizar auditorías y mejorar la eficiencia operativa de la administración tributaria.

Walter (2013) aborda la predicción de resultados de inspecciones sanitarias de restaurantes en la ciudad de Nueva York utilizando técnicas de machine learning sobre datos históricos de inspección. El trabajo, desarrollado como proyecto de curso de CS229 en Stanford, emplea modelos de clasificación para anticipar la probabilidad de violaciones graves y cierres, a partir de variables como historial de infracciones, tipo de establecimiento y ubicación. Este estudio constituye un antecedente relevante en el uso de datos de inspección administrativa para construir modelos de riesgo que permitan orientar de forma más eficiente las visitas de fiscalización en el ámbito de salud pública.

El caso “**Analytics and AI Solutions for Risk Based Inspections**”, documentado por el **Observatory of Public Sector Innovation (2023)**, describe cómo el Ministerio de Asuntos Municipales y Vivienda de Arabia Saudita (MoMRAH) implementó soluciones de analítica avanzada e

IA para priorizar inspecciones municipales con enfoque basado en riesgo. La iniciativa ClearVision combina clustering, motores de despacho basado en riesgo y modelos de priorización de casos para optimizar las rutas de los inspectores, logrando reducciones de hasta 30% en visitas y una mejor focalización de áreas críticas. Este ejemplo ilustra cómo la integración de datos operativos, rediseño organizacional y modelos predictivos puede transformar la supervisión regulatoria en grandes jurisdicciones.

Bernasconi (2018), es uno de los primeros trabajos en describir el uso de machine learning por la policía italiana, particularmente en Venecia, para predecir la ocurrencia de delitos a partir de patrones histórico espaciales de criminalidad. Según referencias secundarias, el estudio analiza cómo los cuerpos policiales utilizan modelos predictivos para identificar zonas y franjas horarias con mayor probabilidad de robos y otros crímenes, apoyando la asignación táctica de recursos. Este caso se suele presentar como ejemplo de “predictive policing” europeo, complementando experiencias de Estados Unidos en la aplicación de algoritmos a la prevención del delito.

Cinelli (2020) analiza la prevención del delito y el análisis predictivo en el contexto italiano, revisando herramientas y proyectos que utilizan modelos de riesgo espacial y temporal para apoyar la labor policial y fiscal. El trabajo discute experiencias como Delia® y X Law, así como enfoques de Risk Terrain Modeling aplicados a ciudades como Milán, Roma y Nápoles, destacando su uso para identificar “hotspots” delictivos y orientar patrullaje y controles. Al situar el caso italiano dentro de la evolución global del predictive policing, el estudio aporta una perspectiva comparada sobre beneficios, limitaciones y desafíos ético jurídicos del uso de IA en seguridad pública.

Fingas (2017), en una nota de Engadget, reporta que el Departamento de Policía de Chicago observó reducciones significativas en tiroteos (39%) y homicidios (33%) en un distrito tras implementar herramientas de “predictive policing” basadas en software como HunchLab, listas de sujetos de alto riesgo y sensores de disparos. Aunque se trata de una fuente periodística y no de un estudio académico, este reporte se cita frecuentemente como evidencia temprana del impacto potencial de los modelos predictivos en la asignación de recursos policiales. El caso de Chicago se ha convertido así en referencia clave para discutir tanto las promesas como las controversias del uso de analítica avanzada en la seguridad ciudadana.

McKenney (2017), en un informe de la Treasury Inspector General for Tax Administration, evalúa el Return Review Program (RRP) del IRS como sistema de detección de fraude en declaraciones de impuestos, diseñado para superar las

capacidades del antiguo Electronic Fraud Detection System. El reporte muestra que en un piloto de 2014 el RRP incrementó la detección de fraude en 59,4% y redujo la tasa de falsos positivos frente a sistemas previos, consolidándolo como herramienta central para identificar robo de identidad y fraude previo al reembolso. Este estudio institucional es ampliamente citado en la literatura sobre uso de machine learning y analítica avanzada en administraciones tributarias.

Olavsrud (2019), en un artículo de CIO, describe cómo el IRS utiliza el RRP como sistema principal para detectar robo de identidad y fraude previo a la devolución mediante técnicas y modelos predictivos que explotan patrones sutiles en los datos de declaraciones. El texto detalla que el programa combina análisis de datos masivos, scoring de riesgo y modelos de clustering para vincular declaraciones sospechosas, ejemplificando la transición del organismo hacia un enfoque intensivo en analítica. Esta fuente periodística complementa los informes oficiales al ofrecer una visión de alto nivel sobre la arquitectura analítica del IRS y su impacto en la lucha contra el fraude fiscal.

Anichukwueze, Osuji y Oguntegbe (2023), proponen modelos de analítica predictiva para la detección temprana de brechas de cumplimiento y violaciones regulatorias en organizaciones, utilizando técnicas de machine learning para cuantificar el riesgo de incumplimiento. El estudio evalúa algoritmos como random forests, gradient boosting y redes neuronales, reportando precisiones superiores al 80% en distintos tipos de violaciones y mostrando que los modelos predictivos pueden reducir significativamente la incidencia de incumplimientos y sanciones cuando se integran en sistemas de monitoreo continuo. Este trabajo extiende el uso de la analítica avanzada más allá de la detección de fraude hacia una gestión proactiva del riesgo regulatorio.

González y Velásquez (2013), presentan un caso pionero de la administración tributaria chilena en el que se aplican técnicas de minería de datos para caracterizar y detectar contribuyentes que utilizan facturas falsas, un esquema relevante de evasión del IVA. Los autores comparan el rendimiento de árboles de decisión, redes neuronales y redes bayesianas sobre datos del Servicio de Impuestos Internos, encontrando que las redes neuronales alcanzan tasas de detección cercanas al 92% en la identificación de contribuyentes fraudulentos. Este estudio se ha convertido en referencia obligada para trabajos posteriores sobre detección de fraude fiscal mediante machine learning en América Latina.

3. MARCO METODOLÓGICO

3.1 Metodología CRISP-DM

El estudio adopta la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining) como marco metodológico principal, dada su amplia aceptación y aplicabilidad en proyectos de ciencia de datos tanto en el sector privado como en el sector público. CRISP-DM estructura el proceso analítico en seis fases: comprensión del negocio, comprensión de los datos, preparación de los datos, modelamiento, evaluación y despliegue, permitiendo un enfoque iterativo y sistemático.

Cuadro resume el uso de modelos de Machine Learning en predicción de riesgos, destacando su aplicación. Incluye referencias a estudios clave que han demostrado la eficacia de estos enfoques.

Modelo ML	Aplicación Predictiva	Referencias
Regresión Logística (baseline interpretable)	Usada como baseline en modelos de riesgo y priorización; en inspecciones sanitarias y predicción de resultados	Michael Walter (2013), (Fingas, 2017)
Árboles de decisión (J48/C4.5)	En fiscalización municipal y tareas de clasificación es común compararlo con otros modelos; en São Paulo se evaluaron árboles en predicción de ilícitos	(Bernasconi, 2018), Ippolito, A., & Lozano, A. C. G. (2020),
Random Forest (ensamble de árboles)	Mostró mejor desempeño frente a varios algoritmos para apoyar planes de fiscalización/auditoría	Ippolito, A., & Lozano, A. C. G. (2020), (McKenney, 2017)
Naive Bayes (baseline probabilístico)	Clasificación rápida con muchos atributos categóricos; útil como benchmark y cuando se requiere simplicidad	Ippolito, A., & Lozano, A. C. G. (2020)
Gradient Boosting (XGBoost/GBTs)	maximiza desempeño predictivo en dominios de inspección, salud pública y predicción de resultados de inspecciones se reporta uso efectivo de modelos de boosting	Sukhani, S., Tang, S., Mitra, C., Zhang, J., Patel, A., & Yetukuri, C. (2013)
Inspección” (RBI) orientados a priorización	Transformar inspección reactiva en inspección preventiva : asignar recursos donde la probabilidad/impacto del incumplimiento es mayor	Analytics and AI Solutions for Risk Based Inspections (Observatory of Public Sector Innovation, 2023)
Auditoría/fiscalización “inteligente” en gobiernos (casos municipales)	Identificar contribuyentes/establecimientos con mayor probabilidad de infracción grave y optimizar planes de fiscalización	(Olavsrud, 2019), Ippolito, A., & Lozano, A. C. G. (2020)

4. COMPRENSIÓN DEL NEGOCIO

El problema de negocio identificado se relaciona con la ausencia de mecanismos analíticos objetivos para priorizar las inspecciones municipales en función del riesgo real de clausura definitiva de los establecimientos comerciales. Este vacío genera inspecciones poco focalizadas, sobrecostos operativos y una limitada capacidad preventiva.

El objetivo analítico del estudio fue desarrollar un modelo de clasificación binaria capaz de predecir la variable **risk_high**, asociada a la ocurrencia de clausura definitiva, minimizando especialmente los falsos negativos, dado su alto costo institucional y social.

Dado el marcado desbalance de la variable objetivo, donde los casos de clausura definitiva representaban una proporción significativamente menor respecto del total de establecimientos, se aplicó la técnica **Synthetic Minority Over-sampling Technique (SMOTE)** como estrategia de balanceo de clases. SMOTE permitió generar instancias sintéticas de la clase minoritaria a partir de sus vecinos más cercanos en el espacio de características, incrementando su representación sin recurrir a la duplicación directa de observaciones. Este procedimiento contribuyó a mitigar el sesgo del modelo hacia la clase mayoritaria, mejorar la capacidad de aprendizaje de los algoritmos supervisados y favorecer una evaluación más equitativa del desempeño predictivo, particularmente en términos de *recall* y reducción de falsos negativos para la clase crítica.

5. COMPRENSIÓN Y PREPARACIÓN DE LOS DATOS

5.1 Dataset y variables

Se utilizó un dataset administrativo real compuesto por 1,856 registros, donde cada observación corresponde a un establecimiento comercial. Las variables incluyen características físicas del local, información normativa, historial de inspecciones e infracciones, condiciones territoriales y situación administrativa.

5.2 Construcción del label

La variable objetivo **risk_high** se definió de forma binaria a partir de la condición de clausura definitiva, garantizando una correspondencia directa con el objetivo de negocio y evitando ambigüedades conceptuales.

5.3 Preprocesamiento de Datos

El preprocesamiento de los datos incluyó un manejo diferenciado de valores faltantes basado en el significado institucional de cada variable, así como una detección conservadora de valores atípicos, privilegiando la

preservación de casos extremos plausibles por su relevancia en modelos de riesgo. Adicionalmente, se llevó a cabo una auditoría exhaustiva para prevenir la inclusión de variables con *data leakage*, asegurando que el modelo utilizara exclusivamente información disponible antes del proceso de fiscalización. Este enfoque garantizó la validez operativa del modelo y evitó la sobreestimación artificial de su desempeño predictivo.

El preprocesamiento se diseñó bajo los principios de CRISP-DM (fase Data Preparation) y de ciencia de datos responsable, con tres objetivos principales:

1. Garantizar calidad y consistencia de los datos.
2. Evitar sesgos artificiales que inflen el desempeño del modelo.
3. Asegurar validez operativa, es decir, que el modelo utilice únicamente información disponible antes del acto de fiscalización.

5.3.1. Manejo de valores faltantes (Missing Values)

Diagnóstico inicial

Se realizó un análisis exhaustivo de valores faltantes por variable, identificando tres patrones:

Tipo de variable	Patrón observado	Interpretación
Variables estructurales (área, aforo)	Faltantes escasos (<5%)	Error de registro
Variables de historial sancionador	Faltantes frecuentes	Ausencia real de eventos
Variables normativas	Faltantes sistemáticos	No aplicabilidad

5.3.2. Detección y tratamiento de valores atípicos (Outliers)

Enfoque conceptual

Se adoptó una estrategia conservadora:

- Detectar outliers
- Interpretarlos, no eliminarlos automáticamente

5.4 Eliminación de data leakage

Se realizó un análisis exhaustivo para identificar y excluir variables que incorporaran información posterior al evento de clausura definitiva o que constituyeran indicadores directos del resultado, con el fin de preservar la validez predictiva del modelo.

5.5 Dataset model-ready

El conjunto final de datos incluyó 42 variables predictoras, correctamente tipificadas y preparadas para su uso en algoritmos de *Machine Learning*, y fue exportado en formato compatible con herramientas de minería de datos.

5.6 Balanceo del set de datos

Dado el marcado desbalance de la variable objetivo, donde los casos de clausura definitiva representaban una proporción significativamente menor respecto del total de establecimientos, se aplicó la técnica Synthetic Minority Over-sampling Technique (SMOTE) como estrategia de balanceo de clases. SMOTE permitió generar instancias sintéticas de la clase minoritaria a partir de sus vecinos más cercanos en el espacio de características, incrementando su representación sin recurrir a la duplicación directa de observaciones. Este procedimiento contribuyó a mitigar el sesgo del modelo hacia la clase mayoritaria, mejorar la capacidad de aprendizaje de los algoritmos supervisados y favorecer una evaluación más equitativa del desempeño predictivo, particularmente en términos de *recall* y reducción de falsos negativos para la clase crítica.

6. MODELAMIENTO

Se entrenaron y evaluaron los siguientes algoritmos de clasificación supervisada:

- Árbol de decisión J48
- Random Forest
- Naive Bayes
- Regresión Logística con regularización ridge

Todos los modelos fueron evaluados mediante validación cruzada estratificada de 10 pliegues, considerando el desbalance natural de la clase objetivo.

7. EVALUACIÓN Y RESULTADOS

La evaluación se centró en métricas relevantes para la gestión del riesgo público: recall de la clase crítica, falsos negativos, precisión, F1-score y área bajo la curva ROC.

Los resultados muestran que:

- La regresión logística regularizada alcanzó el mayor desempeño predictivo, con un recall de 0.989 y solo dos falsos negativos.
- El árbol de decisión J48 presentó un desempeño ligeramente inferior en términos de recall, pero destacó por su alta interpretabilidad y claridad en la generación de reglas.

- Naive Bayes exhibió un número elevado de falsos negativos, resultando inadecuado para un contexto de fiscalización crítica.

Modelo	Recall (1)	Falsos Negativos (FN)	Precision (1)	F1 (1)	ROC AUC
J48	0.972	5	0.978	0.975	1.000
Random Forest	0.956	8	0.994	0.975	1.000
Naive Bayes	0.889	20	0.800	0.842	0.994
Logistic (R = 1e-4)	0.989	2	0.994	0.992	1.000

“La comparación de modelos evidencia que la Regresión Logística regularizada (R=1e-4) y el árbol de decisión J48 presentan el mejor desempeño para la predicción de riesgo crítico, al minimizar falsos negativos (2 y 5 respectivamente), mantener altos valores de recall (≥ 0.97) y alcanzar un área bajo la curva ROC cercana a 1.0, garantizando tanto efectividad operativa como explicabilidad institucional.”

Matriz de comparación con otras investigaciones

Estudio	Problema (objetivo) / Modelo	Métricas reportadas	Resultado
MoMRAH, Ministerio de Asuntos Municipales y Vivienda de Arabia Saudita (Arabia Saudita) – “Analytics and AI Solutions for Risk Based Inspections”	Motores analíticos para priorización y despacho basado en riesgo (inspección municipal) / “Case clustering”, “case prioritization”, “risk-based dispatching”	Recall/Precision/F1/ROC AUC: NR	Reportan impacto operativo: clustering con hasta 30% de ahorro en visitas de inspectores y focalización de zonas (“1200+ areas”) , pero sin métricas clásicas de ML.
Modelo predictivo de riesgo de clausura definitiva para la fiscalización municipal (San Isidro, Weka, SMOTE, 10-fold CV)	Predicción de riesgo de clausura definitiva para fiscalización inteligente / Probamos varios modelos; mejor desempeño global con Logistic Regression	Logistic Regression (ridge 1e-8): Recall 0.989, Precision 1.000, F1 0.994, ROC AUC 0.998 (clase 1 = “riesgo alto”)	Desempeño muy alto en identificación de alto riesgo (minimiza FN), lo que es crítico para fiscalización preventiva basada en evidencia.

Estudio	Problema (objetivo) / Modelo	Métricas reportadas	Resultado
Ippolito & Lozano (Municipio de São Paulo) – “Tax Crime Prediction”	Predicción de “crímenes tributarios” para orientar auditorías / Comparan varios clasificadores; Random Forest destaca	Entrenamiento/validación (RF): Recall 0.516. Precision 0.870. F1 0.648. Validación en 2018: Recall 0.889. Precision 0.640. F1 0.744.	Buen desempeño en precision (menos falsos positivos), útil cuando la inspección es costosa; no reportan AUC.
Modelo predictivo de riesgo de clausura definitiva para la fiscalización municipal (San Isidro, Weka, SMOTE, 10-fold CV)	Predicción de riesgo de clausura definitiva para fiscalización inteligente / Probamos varios modelos; mejor desempeño global con Logistic Regression	Logistic Regression (ridge 1e-8): Recall 0.989, Precision 1.000, F1 0.994, ROC AUC 0.998 (clase 1 = “riesgo alto”)	Desempeño muy alto en identificación de alto riesgo (minimiza FN), lo que es crítico para fiscalización preventiva basada en evidencia.

Frente a MoMRAH, la comparación es principalmente conceptual/operativa, porque su documento enfatiza ahorro y focalización territorial más que métricas predictivas estándar.

Evaluación del impacto del modelo en la entidad

La implementación del modelo predictivo de riesgo de clausura definitiva permitiría una reducción estimada del 30% al 40% de inspecciones no prioritarias, lo que se traduce en un ahorro anual aproximado de entre 4,500 y 6,100 horas-hombre. Este impacto equivale a liberar entre 2.5 y 3.4 inspectores a tiempo completo por año, permitiendo la reasignación de recursos hacia fiscalizaciones de mayor complejidad y riesgo, sin comprometer la capacidad de control ni la cobertura operativa.

Impacto comparativo

São Paulo: impacto en *eficiencia de auditoría*.

MoMRAH: impacto directo en *productividad institucional*.

San Isidro: impacto combinado:

- eficiencia,
- prevención,
- priorización,
- reducción de riesgo extremo (clausura definitiva).

Frente al caso de São Paulo, nuestro modelo muestra Recall y F1 considerablemente superiores (aunque son dominios distintos: fraude tributario vs clausura definitiva). El estudio brasileño prioriza precisión y validación temporal (2018).

NIVEL DE CRITICIDAD DEL EVENTO

Caso	Evento predicho
MoMRAH- Ministerio de Asuntos Municipales y Vivienda de Arabia Saudita	Ahorro en visitas de inspectores y focalización de zonas
São Paulo	Riesgo tributario
San Isidro	Clausura definitiva (evento extremo, alto costo social)

8. DISCUSIÓN

Los hallazgos evidencian que el riesgo de clausura definitiva se encuentra fuertemente asociado a variables relacionadas con la reincidencia, el historial sancionador y los resultados de inspecciones técnicas previas. Asimismo, se confirma que modelos explicables pueden ofrecer niveles de desempeño comparables —e incluso superiores— a modelos más complejos, cuando el fenómeno subyacente presenta patrones administrativos consistentes.

Desde una perspectiva de gobierno digital, la interpretabilidad del modelo emerge como un criterio tan relevante como la precisión predictiva, especialmente en entornos sujetos a auditoría, control y rendición de cuentas.

9. CONCLUSIONES

El estudio demuestra que es factible implementar un modelo predictivo robusto y explicable para apoyar la fiscalización municipal inteligente basada en evidencia. La aplicación rigurosa de CRISP-DM permitió estructurar un proceso analítico sólido, desde la definición del problema hasta la evaluación de modelos.

Considerando que el modelo predictivo reduce entre 30% y 40% las inspecciones de bajo valor, el requerimiento operativo de fiscalizadores disminuiría de 60 a un rango estimado entre 36 y 42 funcionarios. No obstante, se propone reorientar la capacidad liberada hacia fiscalizaciones especializadas, prevención y seguimiento de reincidencias, en lugar de reducir personal, alineando el modelo con un enfoque de maximización del valor público.

Se concluye que una estrategia híbrida, combinando árboles de decisión para la generación de reglas y regresión logística para el scoring continuo de riesgo, representa una solución óptima para contextos municipales.

10. LÍNEAS FUTURAS DE INVESTIGACIÓN

Futuras investigaciones podrían explorar la integración de estos modelos en sistemas de información geográfica, el uso de técnicas de *Explainable AI* más avanzadas y la evaluación

del impacto real del modelo en la reducción de incidentes y riesgos urbanos.

11. REFERENCIAS

- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. IBM.
- Provost, F., & Fawcett, T. (2013). *Data science for business*. O'Reilly.
- Raschka, S., & Mirjalili, V. (2019). *Python machine learning*. Packt.
- Kelleher, J. D., Namee, B. M., & D'Arcy, A. (2020). *Fundamentals of machine learning for predictive data analytics*. MIT Press.
- Sukhani, S., Tang, S., Mitra, C., Zhang, J., Patel, A., & Yetukuri, C. (n.d.). Analyzing Restaurant Health Inspections and Reviews to Predict Food Safety Risks.
- Ippolito, A., & Lozano, A. C. G. (2020). Tax Crime Prediction with Machine Learning: A Case Study in the Municipality of São Paulo. In *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020) - Volume 1 (pp. 452–459)*. SCITEPRESS- Science and Technology Publications, Lda.
- Walter, M. (2013). *Prediction of NYC Restaurant Health Inspection Results*.
- Analytics and AI Solutions for Risk Based Inspections (*Observatory of Public Sector Innovation, 2023*).
- Bernasconi, M. (2018). *The use of Machine Learning by the Italian Police with the goal of predicting crimes*.
- Cinelli, V. (2020), *CRIME prevention and predictive analysis: the Italian case*.
- Fingas, J. (2017, 5 de agosto). *Chicago police see less violent crime after using predictive code*. Engadget.

- McKenney, M. E. (2017). *The Return Review Program increases fraud detection; however, full retirement of the Electronic Fraud Detection System will be delayed* (Report No. 2017-20-080). Treasury Inspector General for Tax Administration.
- Olavsrud, T. (2019, 24 de junio). *IRS combats fraud with advanced data analytics*. CIO.
- Anichukwueze, C. C., Osuji, V. C., & Oguntegbe, E. E. (2023). *Predictive analytics models for early detection of compliance breaches and regulatory violations*. Multidisciplinary Frontiers Research Journal, 3(3), 1–15.
- González, P. C., & Velásquez, J. D. (2013). *Characterization and detection of taxpayers with false invoices using data mining techniques*. Expert Systems with Applications, 40(5), 1427–1436.