

**Escuela Superior Politécnica del Litoral**  
**Facultad de Ciencias Sociales y Humanísticas**

**Mejora del proceso de debida diligencia y gestión de riesgos**

**PLA-FT en aseguradora mediante ML**

ADMI-1305

**Proyecto Integrador**

Previo la obtención del Título de:

**Licenciado en Auditoría y Control de Gestión**

Presentado por:

Rugel Velasco Madeline

Carpio Ortiz Jean Carlos

Guayaquil – Ecuador

2025

## Dedicatoria

---

A **Dios**, quien va delante de mí en cada etapa de mi vida y de manera especial, en este peldaño que hoy me permite avanzar. A Él sea toda la gloria, porque ha transformado mi corazón, fortalecido mi carácter y guiando mi actitud y esfuerzo a lo largo de todo este proceso académico. Su gracia ha sido sustento constante y su propósito, mi dirección.

*Fíate de Jehová de todo tu corazón, y no te apoyes en tu propia prudencia. Reconócelo en todos tus caminos, y él enderezará tus veredas. (Proverbios 3:5,6)*

A mi **familia**, por su apoyo incondicional y por ser un ejemplo de esfuerzo y perseverancia. Sus consejos, bendiciones y amor han orientado mi camino y me han impulsado a seguir adelante.

- **Madelline Marian Rugel Velasco**

## Agradecimientos

---

A **Dios**, fuente de todo conocimiento y sabiduría,  
en quien he puesto mi confianza para  
alcanzar cada objetivo conforme a su  
voluntad.

A mis hermanas, **Angie R., Mia R., Lidia R.**,  
además de ser mi familia son mis amigas  
especiales, gracias por su amor, comprensión  
y por acompañarme con aliento en cada  
etapa. A **Eduardo V.** por su apoyo sincero y  
la motivación permanente que fueron de gran  
valor.

A mi compañero de tesis Jean Carlos C. por su  
gran dedicación y compromiso, piezas clave  
para la culminación exitosa de este trabajo.

A mis docentes, en especial **Mgtr. Diana  
Montalvo** y **Ph.D. Benigno Alfredo  
Armijos De la Cruz** por compartir sus  
conocimientos y orientación, aportando  
significativamente a mi formación  
profesional.

- **Madelline Marian Rugel Velasco**

## Dedicatoria

---

A **Dios**, por ser ese faro eterno que ilumina mis noches más oscuras, por ser la fortaleza y el apoyo en los momentos difíciles, y por ser esa guía constante que dirige mi camino.

A mi **familia**, por sostenerme con su cariño incondicional y ese apoyo firme que nunca me permitió darme por vencido. Gracias, porque su amor, sacrificio y fuerza han sido los cimientos de todo lo que soy, y porque son en el motor inagotable que me impulsa a seguir adelante hasta alcanzar mis metas.

Y, muy especialmente, a **mí mismo**, por no rendirme cuando todo parecía difícil y las fuerzas fallaban. Gracias por esa fuerza para seguir adelante en los momentos duros, que probaron mi voluntad y me enseñaron que solo con constancia se llega al éxito.

- **Jean Carlos Carpio Ortiz**

## Agradecimientos

---

A **Dios** por ser mi guía y fuente de serenidad en cada etapa de este camino y a **mis padres**, por su amor sin fin y confianza inquebrantable en mí.

A mis amigos de la **ESPOL**, por las risas, los momentos compartidos y las batallas libradas en las aulas, su compañía hizo que este camino fuera más humano y memorable. En especial a, Jesús T., Nahiska M., Melanny C. y a mi amiga de tesis Madelline R. por su esfuerzo y dedicación.

A mis docentes, por compartir su conocimiento con una vocación que trasciende las aulas. Expreso mi más profundo agradecimiento a la **Mgtr. Diana Montalvo** y a la **Mgtr. Jessica Menéndez** por su confianza y apoyo constante.

- **Jean Carlos Carpio Ortiz**

## Declaración Expresa

---

Nosotros Rugel Velasco Madeline Rugel y Carpio Ortiz Jean Carlos, acordamos y reconocemos que:

La titularidad de los derechos patrimoniales de autor (derechos de autor) del proyecto de graduación corresponderá al autor o autores, sin perjuicio de lo cual la ESPOL recibe en este acto una licencia gratuita de plazo indefinido para el uso no comercial y comercial de la obra con facultad de sublicenciar, incluyendo la autorización para su divulgación, así como para la creación y uso de obras derivadas. En el caso de usos comerciales se respetará el porcentaje de participación en beneficios que corresponda a favor del autor o autores.

La titularidad total y exclusiva sobre los derechos patrimoniales de patente de invención, modelo de utilidad, diseño industrial, secreto industrial, software o información no divulgada que corresponda o pueda corresponder respecto de cualquier investigación, desarrollo tecnológico o invención realizada por mí/nosotros durante el desarrollo del proyecto de graduación, pertenecerán de forma total, exclusiva e indivisible a la ESPOL, sin perjuicio del porcentaje que me/nos corresponda de los beneficios económicos que la ESPOL reciba por la explotación de mi/nuestra innovación, de ser el caso.

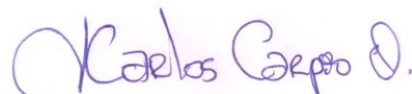
En los casos donde la Oficina de Transferencia de Resultados de Investigación (OTRI) de la ESPOL comunique al/los autores/es que existe una innovación potencialmente patentable sobre los resultados del proyecto de graduación, no se realizará publicación o divulgación alguna, sin la autorización expresa y previa de la ESPOL.

Guayaquil, 03 de febrero del 2026.



---

Madeline Marian Rugel Velasco



---

Jean Carlos Carpio Ortiz

## **Evaluadores**

---

**Ing. Alfredo Armijos Benigno**

Profesor de Materia

---

**Ing. Diana Montalvo Barrera**

Tutor de proyecto

### **3. Resumen**

El proyecto diseña un sistema de calificación de riesgo dinámico que integra la automatización de la matriz de riesgos y un modelo de Aprendizaje automático para determinar el perfil de riesgo del cliente. El estudio se justifica ante la necesidad de transformar la gestión manual en un proceso predictivo y estandarizada que mitigue el lavado de activos bajo normativas como la UAFE. Para el desarrollo, se adoptó una metodología híbrida que combinó la gestión de procesos (BPM), para el rediseño operativo y el enfoque CRISP-DM para el ciclo de vida de los datos. Se realizó un diagnóstico inicial en el área de cumplimiento y se construyó un prototipo alineado con la Ley de Protección de Datos Personales. Los resultados se evidenciaron en la optimización operativa, reduciendo el tiempo de procesamiento por cliente de 45 a 10 minutos. Asimismo, la tasa de falsos positivos descendió del 85% al 20% por la precisión del modelo.

En conclusión, la transición de un modelo manual a un predictivo valida la mejora en la eficiencia y confiabilidad del proceso. La automatización del 80% de la debida diligencia fortalece la postura regulatoria y optimiza la experiencia del analista.

#### **Palabras Clave:**

Aprendizaje automático, Automatización, Gestión de riesgos, Puntaje de riesgo y Debida diligencia

#### **4. Abstract**

The project designs a dynamic risk rating system that integrates risk matrix automation and a machine learning model to determine the customer's risk profile. The study is justified by the need to transform manual management into a predictive and standardized process that mitigates money laundering under regulations such as the UAFE. For the development, a hybrid methodology was adopted that combine process management (BPM) for operational redesign and the CRISP-DM approach for the data life cycle. An initial diagnosis was made in the area of compliance, and a prototype was built in line with the Personal Data Protection Law. The results were evident in operational optimization, reducing processing time per customer from 45 to 10 minutes. Likewise, the false positive rate dropped from 85% to 20% due to the accuracy of the model.

In conclusion, the transition from a manual model to a predictive model validates the improvement in the efficiency and reliability of the process. The automation of 80% of due diligence strengthens the regulatory position and optimizes the analyst's experience.

***Keywords:***

*Machine Learning, Automation, Risk Management, Risk Scoring, and Due Diligence*

## Índice General

<b>3. Resumen .....</b>	<b>8</b>
<b>4. Abstract.....</b>	<b>9</b>
<b>1. Introducción .....</b>	<b>16</b>
<b>1.1. Descripción de la problemática .....</b>	<b>17</b>
<b>1.2. Justificación del problema .....</b>	<b>19</b>
<b>1.3. Alcance .....</b>	<b>20</b>
<b>1.4. Objetivos .....</b>	<b>21</b>
1.4.1. Objetivo general.....	21
1.4.2. Objetivos específicos .....	21
<b>1.5.1. Marco teórico .....</b>	<b>22</b>
1.5.1. Marco Conceptual.....	22
1.5.1.1. El proceso de debida diligencia .....	22
1.5.1.2. El aprendizaje automático o Machine learning.....	22
1.5.1.3. La automatización de procesos .....	22
1.5.1.4. La gestión de riesgo .....	23
1.5.2. Marco legal .....	23
1.5.2.1. Ley de prevención de lavado de activos y del financiamiento de delitos.....	23
1.5.2.2. Reglamento de la Superintendencia de Compañías, Valores y seguros (SCVS)..	23
1.5.2.3. Recomendaciones del Grupo de acción financiera internacional (GAFI) .....	23
1.5.3. Marco referencial .....	24
1.5.4. Marco metodológico .....	25
<b>2. Metodología .....</b>	<b>27</b>

<b>2.1. Descripción de las alternativas de solución .....</b>	<b>27</b>
2.1.1. Optimización manual y estandarización de procedimientos con capacitación .....	27
2.1.2. Integración de fuentes de datos externos para una debida diligencia mejorada .....	28
2.1.3. Digitalización de formularios y documentos con validación asistida.....	28
2.1.4. Automatización de tareas repetidas con RPA.....	28
2.1.5. Reestructuración de la matriz de riesgo y metodología EBR .....	28
2.1.6. Modelo de Machine Learning (ML) con perfilamiento de riesgo dinámico .....	29
<b>2.2. Evaluación de alternativas con metodologías comparativas.....</b>	<b>29</b>
2.2.1. Matriz de Pugh.....	29
2.2.2. Matriz Impacto - Esfuerzo .....	32
<b>2.3. Enfoque metodológico .....</b>	<b>33</b>
<b>2.4. Diseño metodológico .....</b>	<b>34</b>
2.4.1. Fase 1: Diagnostico operativo (BPM) – Comprensión del negocio (CRISP-DM).....	35
2.4.1. Fase 2: Análisis de datos (BPM) - Comprensión de los datos (CRISP-DM) .....	36
2.4.3. Fase 3: Preparación de los datos .....	36
2.4.4. Fase 4: Motor de decisión (BPM) - Modelado (CRISP-DM).....	37
2.4.5. Fase 5: Validación Normativa (BPM) - Evaluación (CRISP-DM) .....	37
2.4.6. Fase 6: Implementación (BPM) - Despliegue (CRISP-DM) .....	38
<b>2.5. Consideraciones éticas, legales y de seguridad de la información.....</b>	<b>38</b>
<b>3. Resultados y análisis .....</b>	<b>40</b>
<b>3.1. Evaluación del desempeño operativos: comparativa de línea base vs. resultados de prueba piloto.....</b>	<b>41</b>
3.1.1. Resultados de la evaluación inicial (línea base As- Is).....	41

3.1.2. Resultados de prueba piloto .....	42
<b>3.2. Factores de riesgos y variables del modelo predictivo.....</b>	<b>45</b>
3.2.1. Muestra de estudio .....	45
3.2.2. Importancia de variables estratégicas .....	46
<b>3.3. Diseño y arquitectura de la solución .....</b>	<b>47</b>
3.3.1. Arquitectura de la solución automatizada y predictiva.....	47
<b>3.4. Evaluación de la eficacia predictiva del modelo (To-be).....</b>	<b>48</b>
3.4.1. Análisis cuantitativo.....	48
<b>3.5. Evaluación cualitativa de los analistas de cumplimiento .....</b>	<b>48</b>
3.5.1. Calidad percibida en la toma de las decisiones.....	49
3.5.2. Reducción drástica de la carga cognitiva.....	49
<b>3.6. Análisis costo beneficio de la solución.....</b>	<b>49</b>
3.6.1. Análisis de costos.....	50
3.6.2. Análisis de beneficios .....	51
<b>4. Conclusiones y recomendaciones.....</b>	<b>54</b>
<b>4.1. Conclusiones .....</b>	<b>54</b>
<b>4.2. Recomendaciones .....</b>	<b>55</b>
<b>5. Referencia .....</b>	<b>56</b>

## Abreviaturas

<b><i>ML</i></b>	Machine Learning
<b><i>PLA/FT</i></b>	Prevención de Lavado de Activos y Financiamiento del terrorismo
<b><i>ALA/CFT</i></b>	Anti Lavado de Activos y Contra Financiamiento del Terrorismo
<b><i>API</i></b>	Application Programming Interface
<b><i>DDC</i></b>	Debida Diligencia del Cliente
<b><i>GAFI</i></b>	Grupo de Acción Financiera Internacional
<b><i>GAFILAT</i></b>	Grupo de Acción Financiera Internacional de Lationamérica
<b><i>SPARLAFTD</i></b>	Sistema de Prevención y Administración de Riesgos de Lavado de Activos, Financiamiento del Terrorismo y Otros Delitos
<b><i>SCVS</i></b>	Superintendencia de Compañías, Valores y Seguros
<b><i>LOPDEDLAD</i></b>	La Ley Orgánica de Prevención, Detección y Erradicación del Delito de Lavado de Activos y del Financiamiento de Delitos
<b><i>UAFE</i></b>	Unidad de Análisis Financiero y Económico
<b><i>LOPDP</i></b>	Ley Orgánica de Protección de Datos Personales
<b><i>KYC</i></b>	Know Your Customer
<b><i>RPA</i></b>	Automatización robótica de procesos
<b><i>EDA</i></b>	Análisis Exploratorio de Datos
<b><i>OCR</i></b>	Optical Character Recognition

## Índice de figuras

<b>Figura 1:</b> Análisis de causa raíz.....	20
<b>Figura 2:</b> Matriz Impacto - Esfuerzo .....	32
<b>Figura 3:</b> Metodologías aplicadas .....	34
<b>Figura 4:</b> Mapa Jerárquico de Factores y Variables .....	45
<b>Figura 5:</b> Características de variables con SHAP .....	46
<b>Figura 6:</b> Diagrama de flujo.....	47

## Índice de tablas

<b>Tabla 1:</b> Boletines de la FGE, casos de LA en el sector asegurador .....	17
<b>Tabla 2</b> Definición de criterios.....	29
<b>Tabla 3</b> Matriz Pugh, evaluación comparativa de alternativas .....	31
<b>Tabla 4</b> Clasificación de propuestas según impacto y esfuerzo.....	32
<b>Tabla 5:</b> Configuración del modelo .....	37
<b>Tabla 6:</b> Análisis comparativo del proceso de debida diligencia (As-Is vs To-Be) .....	44
<b>Tabla 7:</b> Descripción de APIs .....	50
<b>Tabla 8:</b> Análisis de beneficios cualitativos .....	51
<b>Tabla 9:</b> Análisis de beneficios cuantitativos .....	52
<b>Tabla 10:</b> Desgloce de puntuación por características de Impacto y Esfuerzo . <b>¡Error! Marcador no definido.</b>	
<b>Tabla 11:</b> Matriz de hallazgos operativos .....	61

# Capítulo 1

## **1. Introducción**

La prevención contra el lavado de activos y el financiamiento del terrorismo (PLA/FT) constituye un pilar fundamental para la estabilidad y transparencia del sistema económico y financiero global. En el contexto ecuatoriano, las empresas aseguradoras operan como sujetos obligados (SO) ante la Unidad de Análisis Financiera y Económico (UAFE), encontrándose sujeta a un marco regulatorio riguroso que exige la implementación, ejecución y monitoreo de un Sistema de prevención y administración de riesgo de lavado de activos, financiamiento del terrorismo y otros delitos (SPARLAFTD) (Resolución No. JPRF-V-2022-024, 2022)

El punto central de este sistema (SPARLAFTD) es el enfoque basado en riesgo (EBR) una estrategia que permite a las entidades a identificar, medir, controlar y monitorear los riesgos de lavado de activo y financiamiento al terrorismo, considerando factores asociados a clientes, productos, servicios, canales de distribución y jurisdicción territorial. La herramienta fundamental en este proceso es la debida diligencia del cliente (DDC) la cual busca verificar y actualizar la identidad de los clientes y detectar transacciones sospechosas, por lo que permite establecer sus perfiles de riesgos y facilitar el cumplimiento normativo.

Según un informe emitido por el GAFILAT sobre Ecuador señala recomendaciones de implementar herramientas tecnológicas para fortalecer el cumplimiento de las medidas de ALA/CFT (GAFILAT, 2023) el presente proyecto se orienta en abordar las deficiencias existentes en el proceso de debida diligencia del cliente y en la gestión de riesgo PLA/FT dentro de una aseguradora ecuatoriana, dichas deficiencias se evidencian en limitaciones operativas y en la falta de objetividad de los métodos actualmente utilizados, los cuales se sustentan en reglas estáticas para la evaluación de riesgos, verificaciones manuales y errores de transcripción lo que conlleva a resultados no deseados, entre ellos: elevada tasa de falsos positivos, lentitud en la

detección de patrones sospechosos y dependencia de la subjetividad de los analista de cumplimiento para la calificación del riesgo.

Bajo este escenario, la propuesta del presente estudio se centra en el desarrollo de un modelo inteligente sustentado en técnicas de Machine Learning (ML), cuyo objetivo es transformar el proceso de debida diligencia del cliente de una gestión manual y subjetiva hacia un sistema automatizado, estandarizado y predictivo. La implementación de algoritmos de ML permite optimizar la revisión del asegurado integrando el análisis de los datos a través de la automatización de la matriz de riesgo que define con precisión un perfil de riesgo dinámico, de ese modo se cumple con la normativa y se reduce la subjetividad en la evaluación del riesgo.

### 1.1.Descripción de la problemática

De acuerdo con la Evaluación Nacional de Riesgos (ENR), el Ecuador presenta un nivel de riesgo “medio-alto” frente al lavado de activos (UAFE, 2024). Este análisis identifica a la corrupción y la evasión fiscal como las principales amenazas en este ámbito.

A continuación, la siguiente tabla resume casos reales a la contratación de pólizas de seguros que exponen las debilidades críticas de los actuales métodos de evaluación de riesgos.

**Tabla 1:** Boletines de la FGE, casos de LA en el sector asegurador

N. de casos	Descripción	Resumen
158-DC-2025	Caso Petroecuador	La investigación se originó por irregularidades en la contratación de una póliza de seguro de ramo técnico que cubría riesgos petroleros y responsabilidad civil por más USD 98 millones de dólares. Se identificó pagos anticipados y sobrevaloración de activos asegurables sin respaldo documental (FGE, 2025)
360-DC-2021	Casos Seguros Sucre I	La Fiscalía presentó testigos y evidencias de operaciones financieras irregulares asociadas a pólizas de seguros, que incluye transferencias, adquisición de bienes y vinculación de exdirectivos de la entidad Juan Xavier R con esquemas de sobornos de reaseguradoras. (FGE, 2021)

814-DC-2021	Casos Seguros Sucre II	El tribunal sentenció a diez años al exgerente de Seguros Sucre, José Luis R. y su esposa por lavado de activos vinculados a una póliza de USD 35,2 millones de dólares, tras comprobarse movimientos financieros injustificados, recibieron transferencias desde el exterior y adquirieron bienes con fondos de origen ilícito. (FGE, 2021)
-------------	------------------------------	--

Los casos de sector asegurador ecuatoriano reflejan la vulnerabilidad de corrupción y lavado de activos, pone en manifiesto la importancia de fortalecer los mecanismos de supervisión, promover mayor transparencia en la gestión de seguros y aplicar de manera estricta los procesos de debida diligencia. En el contexto de la aseguradora en estudio, el proceso de debida diligencia y gestión de riesgo PLA/FT presenta nudos críticos que comprometen tanto la eficiencia operativa como la efectividad del sistema de administración de riesgos.

En primer lugar, se identifica las siguientes deficiencias principales: ineficiencia y falta de automatización en la DDC, la revisión y validación de la información de los clientes se realiza de manera manual, lo cual incrementa el riesgo de errores humanos y genera lentitud en el proceso. La revisión exhaustiva de los asegurados y beneficiarios requiere que el oficial de cumplimiento consulte y verifique información en múltiples portales oficiales externos entre ellos: OFAC, SRI, Función Judicial, Fiscalía, así como la recopilación de documentos de respaldo, esta intensiva revisión demanda un tiempo considerable deteniendo la emisión de pólizas y afectando la capacidad operativa de la entidad.

Por consiguiente, la subjetividad e inconsistencia en la evaluación de riesgo, pese a que el enfoque basado en riesgo se basa en metodología que requieren asignar pesos o ponderaciones a factores clave, tales como el tipo de cliente, actividad económica y jurisdicción depende de alto nivel de subjetividad del analista, por lo que afecta la consistencia en la toma de decisiones. Esto

se debe, a una limitada disponibilidad de herramientas tecnológicas que permitan gestionar de forma eficiente la verificación y cuantificación del riesgo.

Por lo tanto, la problemática radica en la necesidad de transformar el proceso de DDC actual, mediante la automatización de la recopilación de información del cliente proveniente del formulario, de modo que la matriz de riesgo se complete de automáticamente y, de forma paralela, un modelo inteligente basado en Machine Learning genere niveles de riesgo ponderados más precisa y con capacidad predictiva, fortaleciendo así la gestión del riesgo PLA/FT y la eficiencia operativa de la entidad.

## **1.2. Justificación del problema**

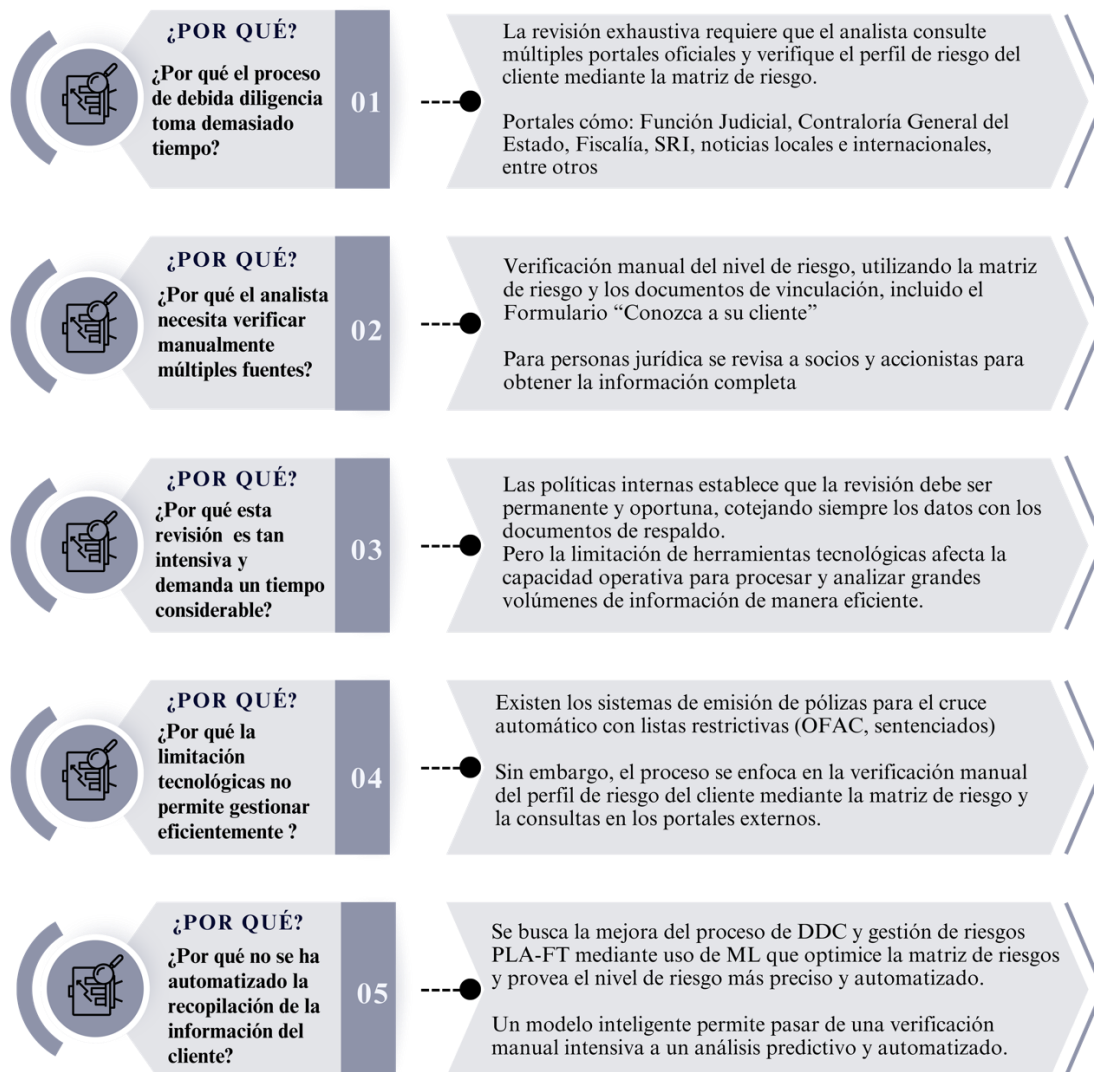
La presente investigación aportará un mecanismo de control preventivo de lavado de activos en el sector asegurador, ante las ineficiencias del proceso actual.

Mediante la aplicación de las metodologías de los “5 Por qué” se identifica como causa raíz la usencia de herramientas tecnológicas que permitan procesar grandes volúmenes de información de manera eficiente, generando cuellos de botellas operativos que afectan la emisión de pólizas. Por lo tanto, la dependencia de verificaciones manuales y procesos extensos generan cuellos de botella operativos, incrementa la carga de trabajo del oficial de cumplimiento y limita la capacidad de la aseguradora para gestionar riesgos de PLA/FT de manera oportuna y objetiva.

En este contexto, la implementación de un modelo inteligente basado en Machine Learning se justifica como una solución estructural y sostenible, al permitir la automatización de la matriz de riesgo, la obtención de perfiles de riesgo más precisos y predictivos, alineado con la norma ISO 31000.

Por lo que permitirá a la asegurado reducir costo operativos asociados a la revisión manual, disminuir el riesgo legal y reputacional ante los entes de control UAFE, SCVS, y mejorar la experiencia de los clientes al agilizar los tiempos para la emisión de pólizas.

**Figura 1: Análisis de causa raíz**



### 1.3. Alcance

El estudio se desarrollará en el área de cumplimiento y control interno de una aseguradora que opera en Ecuador, abarcando el proceso actual de debida diligencia del cliente, la revisión de los manuales de control interno de la compañía y el análisis de la estructura de la matriz de riesgo empleada, el análisis se fundamenta en el cumplimiento de la Ley Orgánica de Prevención, Detección y Erradicación del Lavado de Activos (LOPDEDLAD), incluyendo sus regulaciones y los criterios de un enfoque basado en riesgo (EBR). El presente proyecto busca identificar áreas de fortalecimiento y proponer la aplicación de técnicas de Machine learning

como una alternativa para para optimizar la eficiencia y precisión del proceso de gestión de riesgos, permitiendo un perfilamiento de riesgo dinámico y la generación de alertas tempranas.

## **1.4.Objetivos**

### ***1.4.1. Objetivo general***

Diseñar un modelo de calificación de riesgo que integre la automatización de la matriz institucional con modelos de Machine Learning, bajo el enfoque basado en riesgo (EBR), para optimizar la eficiencia del proceso de debida diligencia de cliente y fortalecer la prevención LA/FT en una aseguradora ecuatoriana.

### ***1.4.2. Objetivos específicos***

- Analizar los controles aplicados en la etapa de vinculación de clientes, mediante entrevistas y revisión documental, para identificar las vulnerabilidades del control interno e ineficiencias operativas que afectan la mitigación del riesgo de LA/FT.
- Determinar los factores de riesgo y variables críticas para la parametrización de la matriz de riesgo y el entrenamiento de modelos de ML asegurando la calidad, integridad y privacidad de la información, en cumplimiento de la normativa vigente y la protección de datos personales.
- Diseñar un modelo conceptual de EBR que integre algoritmos de ML para el cálculo del perfil de riesgo y automatización para el llenado de la matriz de riesgo mediante un script desarrollado en Python.
- Establecer y evaluar la eficacia operativa del modelo ML diseñado, comparando métricas de desempeño (tasa de falsos positivos y reducción de tiempos de procesamientos) frente al sistema actual.

## **1.5.1. Marco teórico**

### ***1.5.1. Marco Conceptual***

#### ***1.5.1.1. El proceso de debida diligencia***

Es un procedimiento fundamental en las entidades financieras o aseguradoras para identificar, verificar y evaluar el riesgo que representan sus clientes antes de establecer una relación comercial o realizar una inversión. Este proceso es crucial, ya que ayuda a las instituciones a blindarse contra el LA-FT, garantizando no solo el cumplimiento normativo, sino una comprensión profunda del perfil de su cliente (Gaviyau & Sibindi, 2023).

#### ***1.5.1.2. El aprendizaje automático o Machine learning***

Se define como una subcategoría fundamental de la inteligencia artificial que permite a los sistemas identificar patrones complejos y optimizar su desempeño de forma autónoma a partir de los datos, requiriendo una intervención humana mínima, una vez desplegado el modelo. En el sector financiero, su implementación ha revolucionado la gestión de riesgos, convirtiéndose en una herramienta crítica para la calificación crediticia y la detección de anomalías (Milojevic & Redzepagic, 2021).

#### ***1.5.1.3. La automatización de procesos***

El objetivo central radica en optimizar la eficiencia de la gestión de riesgo mediante la simplificación y automatización de los procesos operativos (Xiangting Shi, 2025), transformando tareas manuales complejas en funciones rápidas y dinámicas, permitiendo a las instituciones concentrar sus recursos hacia el análisis de valor añadido y la inteligencia de datos, facilitando la toma de decisiones estratégicas fundamentada en datos y no solo en intuición (Prisznyák, 2022)

#### ***1.5.1.4. La gestión de riesgo***

El un componente crítico en el entorno financiero contemporáneo, cuyo propósito principal es la mitigación de riesgos y la garantía de estabilidad financiera (Xiangting Shi, 2025). En este contexto, la integración de tecnologías disruptivas como la Inteligencia artificial y el Machine learning representan una mejora significativa en la capacidad analítica de dicha gestión (Milojevic & Redzepagic, 2021)

#### ***1.5.2. Marco legal***

##### ***1.5.2.1. Ley de prevención de lavado de activos y del financiamiento de delitos***

El objetivo fundamental de este marco normativo es prevenir, detectar y erradicar tres delitos críticos: el lavado de activos, el financiamiento del terrorismo y el financiamiento de la proliferación de armas de destrucción masiva (Ley Orgánica de Prevención, 2025) Asimismo, la ley también detalla las obligaciones de debida diligencia que consiste en la identificación del cliente y la debida diligencia reforzada en escenarios de alto riesgo, especialmente cuando se involucran las personas expuestas políticamente (PEP) o actividades económicas de alta vulnerabilidad, garantizando un monitoreo proporcional de la amenaza identificada.

##### ***1.5.2.2. Reglamento de la Superintendencia de Compañías, Valores y seguros (SCVS)***

La Superintendencia de Compañías, Valores y seguros (SCVS), es su (Resolución No.SCVS-INC-DNCDN-2023-0002, 2023) establece las funciones del oficial de cumplimiento, los procedimientos de verificación de clientes, los criterios de actualización de la matriz de riesgo y la documentación de políticas y controles internos.

##### ***1.5.2.3. Recomendaciones del Grupo de acción financiera internacional (GAFI)***

Estos lineamientos son reconocidos como el estándar global para combatir el LA- FT. Funcionan como un marco integral que los países deben adaptar a su contexto jurídico y está

estructurado en 40 recomendaciones fundamentales (The FATF Recommendations, 2025). Las recomendaciones claves para el sector asegurador son la debida diligencia del cliente (Recomendación 10), la conservación de registros (Recomendación 11) y las nuevas tecnologías y mitigación de riesgos (Recomendación 15).

### ***1.5.3. Marco referencial***

La literatura académica reciente y los análisis de los expertos confirman las severas limitaciones de los sistemas y enfoques tradicionales de prevención del lavado de activos y financiación del terrorismo (PLA/FT). Estos sistemas tradicionales, a menudo basados en reglas, son incapaces de adaptarse a las tácticas criminales en evolución constante. Estas limitaciones justifican la urgencia de que las instituciones financieras adopten soluciones impulsadas por herramientas tecnológicas como (IA/ML) que superen las debilidades de los sistemas obsoletos en la lucha del LA-FT (Howard Chitimira, 2024)

Investigaciones recientes demuestran que los modelos de ML no solo mejoran la precisión en la detección de transacciones sospechosas, sino que también pueden proporcionar a los oficiales de cumplimiento una comprensión y explicación detrás de cada alerta (Vishakha & Vishwamitr, 2024). Asimismo, dichas tecnologías permiten superar las vulnerabilidades de los sistemas tradicionales, por lo que se observa un incremento en el uso de diversos algoritmos de ML para el análisis del riesgo y para la detección de patrones de fraude en tiempo real (Xu Tian, 2024)

Históricamente, la mayor parte de las investigaciones sobre la gestión de riesgos y la implementación de tecnológicas avanzadas, se han enfocado en el sector bancario, dada su relevancia y magnitud de operaciones (Howard Chitimira, 2024). Sin embargo, se observa un desplazamiento hacia nuevas líneas de investigación en sectores emergentes como Fintech y Venture capital, áreas que, aunque menos investigadas al principio, están ganando terreno

rápidamente debido a la urgencia de implementar procesos de debida diligencia más eficientes y precisos en estas empresas (Vito Christian Samudra, 2024)

#### ***1.5.4. Marco metodológico***

En esta investigación se optó por una metodología mixta que ha evolucionado a lo largo de los últimos 20 años y que también se conoce como el tercer paradigma de investigación (Sara Bagur, 2021) Este enfoque metodológico permite realizar una combinación de métodos cualitativo y cuantitativos, lo cual resulta fundamental para una comprensión más profunda y robusta del fenómeno del cumplimiento, de la que se podría obtener con un solo método.

Este método mixto proporciona el lenguaje metodológico para investigar la eficacia y validar la evidencia de algoritmos de ML en el sector financiero. El método cualitativo, explora el proceso actual de debida diligencia para identificar puntos críticos desde la perspectiva del jefe de cumplimiento, lo que permite comprender los desafíos y las necesidades reales. Los hallazgos en esta etapa cualitativa servirán de guía para la evaluación del método cuantitativo, lo que permitirá asegurar que las soluciones tecnológicas se adapten a dichas necesidades.

## **Capítulo 2**

## **2. Metodología**

El presente estudio adoptará un enfoque mixto, combinando métodos cualitativos y cuantitativos de un caso centrado en el área de cumplimiento de una aseguradora. A nivel cualitativo se explorará el proceso actual de debida diligencia del cliente (DDC) para identificar los puntos críticos y necesidades operativas, mientras que, a nivel cuantitativo se validará la eficacia del modelo inteligente de Machine Learning y el script de automatización propuestos para transformar la gestión del riesgo manual y subjetiva a un sistema automatizado y predictivo.

Para la estructuración del ciclo de vida del modelo ML, se implementará la metodología CRISP-DM garantizando el cumplimiento en la prevención de LA/FT de manera eficaz y oportuna, paralelamente para el mapeo y rediseño del flujo de trabajo operativo se utilizará el modelo BPM, facilitando la transición hacia el proceso de debida diligencia automatizada.

El alcance del estudio es en el área de cumplimiento de una aseguradora en calidad de sujeto obligado (SO) ante la UAFE y está sujeta a un marco regulatorio, la investigación utilizará como unidad de análisis los expedientes de clientes (identidad, actividad económica y perfil del riesgo) garantizando una solución que se alinea con las regulaciones nacionales y la innovación tecnológica, para una gestión de riesgo automatizada y eficiente que mitigue la exposición a delitos como lavado de dinero identificado como una amenaza de riesgo medio-alto en Ecuador.

### **2.1. Descripción de las alternativas de solución**

Para determinar la solución final, se evaluaron las distintas alternativas a través de la matriz de pugh y matriz de impacto-esfuerzo, con el fin de dar prioridad a la propuesta que ofrece mayor valor agregado para mejorar el trabajo operativo de debida diligencia.

#### ***2.1.1. Optimización manual y estandarización de procedimientos con capacitación***

Proponer una revisión integral de los procedimientos internos del SPARLAFTD, para identificar ineficiencias, redefinir factores y categorías de riesgo; así, estandarizar formularios y

flujos de trabajo incluyendo acción de capacitaciones y, de este modo, fortalecer el desempeño operativo y garantizar la conformidad normativa, sin incorporar nuevas tecnologías.

### ***2.1.2. Integración de fuentes de datos externos para una debida diligencia mejorada***

Desarrollar una conexión con bases de datos externas y fuentes especializadas, como listas PEP, registros de sanciones, noticias locales e internacionales y portales oficiales de Ecuador (Función Judicial, Contraloría General del Estado, Fiscalía, SRI, etc.), con el fin de enriquecer el perfil del cliente y obtener una visión más integral del riesgo potencial.

### ***2.1.3. Digitalización de formularios y documentos con validación asistida***

Migrar de formatos físicos a digitales, transformando el proceso de captura de información de cliente mediante uso de tecnologías como el reconocimiento óptico de caracteres (OCR) para la extracción de datos, y formularios electrónicos con validaciones en tiempo real, a fin de automatizar el flujo de trabajo y optimizar la recolección, almacenamiento y accesibilidad de la información del cliente.

### ***2.1.4. Automatización de tareas repetidas con RPA***

Implementar robots de software (bots) para automatizar tareas manuales, recopilar datos desde múltiples sistemas, cargar y consolidar expedientes del cliente, y generar reportes estandarizados. De esta manera, la automatización robótica de proceso (RPA) fortalece la eficiencia operativa y disminuye errores humanos derivados de tareas repetitivas.

### ***2.1.5. Reestructuración de la matriz de riesgo y metodología EBR***

Incorporar nuevos indicadores de riesgos tanto cuantitativos como cualitativos (como el volumen transaccional, los montos operados, la geografía de operaciones, el tipo del cliente, su reputación, etc.) permitiría estandarizar los criterios de ponderación, y en consecuencia definir umbrales de riesgos más precisos para optimizar la evaluación y gestión conforme a la metodología basada en riesgo (EBR).

### **2.1.6. Modelo de Machine Learning (ML) con perfilamiento de riesgo dinámico**

Para un perfilamiento de riesgo dinámico, desarrollar un modelo predictivo y supervisado con capacidad de analizar grandes volúmenes de datos integrando la información histórica interna y externa del cliente, con el fin de detectar anomalías y estimar la probabilidad de riesgo de cada cliente mediante un puntaje (scoring) que se actualiza según el comportamiento del asegurado. De este modo, se reduce la subjetividad y minimiza la tasa de falsos positivos.

## **2.2. Evaluación de alternativas con metodologías comparativas**

### **2.2.1. Matriz de Pugh**

Para asegurar que la solución final no se base en percepción subjetiva, se empleó la matriz de Pugh, como una herramienta de evaluación cuantitativa para la comparación objetiva de distintas alternativas. En primer lugar, se definieron seis criterios de evaluación alineados con los objetivos de la aseguradora, los cuales se detallan en la Tabla 2: Definición de criterios. A cada criterio se le asignó un nivel de impacto es una escala del 1 al 5, donde el valor 5 representa un impacto alto, el 3 un impacto medio y el 1 un impacto bajo.

**Tabla 2** *Definición de criterios*

<b><i>Criterios</i></b>	<b><i>Descripción</i></b>	<b><i>Nivel de impacto</i></b>
Eficiencia operacional	Reducción de tiempos de análisis y carga operativa	5
Reducción de sesgo humano	Minimizar la subjetividad del analista al evaluar alertas	3
Precisión en la detección de riesgos	Efectividad del sistema de PLA/FT para identificar operaciones inusuales del cliente	5
Adaptabilidad a cambios regulatorios	Flexibilidad para ajustarse a nuevas normativas de la UAFE o entes de control	5
Costo de implementación y mantenimiento	Inversión financiera requerida	1
Escalabilidad	Capacidad de solución para un alto volumen de asegurados sin perder eficiencia	3

Posteriormente, se procedió a la evaluación comparativa en la Matriz de Pugh (ver la Tabla 3: Matriz Pugh, evaluación comparativa de alternativas) la metodología de puntuación se realizó bajo criterios de los analistas de compliance, consistió en valorar cada alternativa asignando valores de (+1) solución superior, (0) igual, (-1) solución inferior, estos valores se multiplicaron por el nivel de impacto de cada criterio para obtener una puntuación ponderada total. Sobre el criterio de precisión en la detección de riesgos es una variable clave para el área y las alternativas de solución como integración de fuentes y la automatización con RPA se les asignó una calificación de (0) porque en sí agilizan el flujo de información, pero no mejoran la capacidad de análisis de riesgos, las otras alternativas como la optimización de procesos fueron valorada con (0) en el criterio de escalabilidad debido a su dependencia directa de los recursos humanos, en caso de grandes volúmenes de expedientes se requerirá el doble de personal. Por otro lado, la alternativa de reestructuración de la matriz de riesgo obtuvo una calificación (-1) en eficiencia, porque los analistas consideran que agregar más controles a la matriz sin herramientas tecnológicas que lo soporten, incrementaría la carga burocrática y lentitud en el proceso de alta de clientes, afectando la agilidad operativa.

Los resultados evidencian a la alternativa 6 el modelo de Machine Learning con perfilamiento de riesgo dinámico como la solución óptima con un puntaje ponderado de 21 puntos y como segunda solución la automatización con RPA con un puntaje ponderado de 10 puntos, estas propuestas tienen la capacidad de satisfacer a criterios de alto impacto como eficiencia operacional reduciendo la carga operativa, la detección de riesgo eliminando la subjetividad humana y adaptabilidad para mejorar la detección de lavadores de activos

**Tabla 3** Matriz Pugh, evaluación comparativa de alternativas

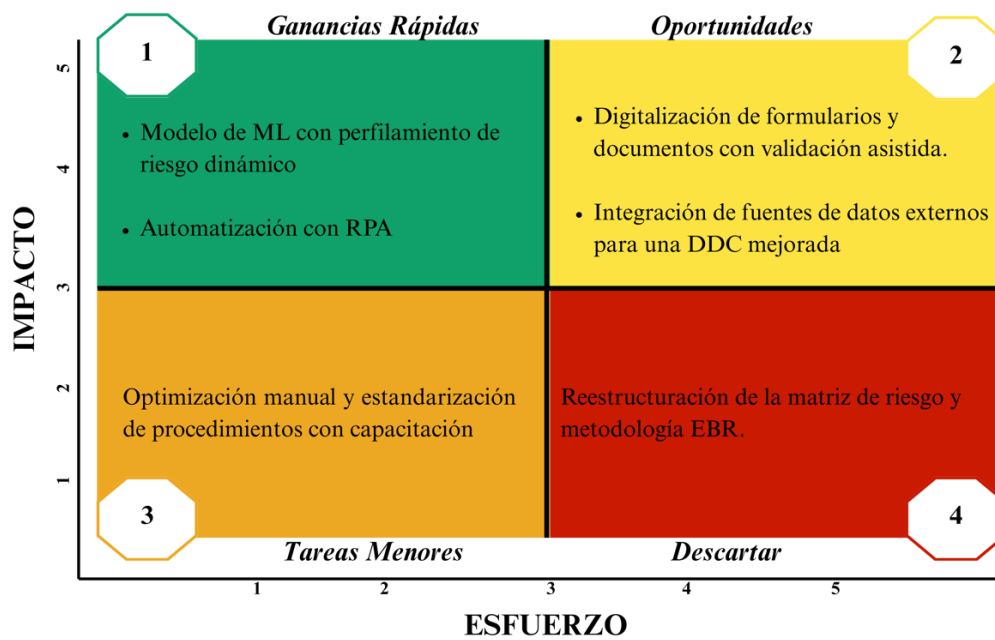
<i>Criterios</i>	<i>Impacto Puntuación</i>	Optimización manual y estandarización de procesos capacitación	Integración de fuentes de datos externos	Digitalización de formularios y documentos con validación asistida.	Automatización de tareas repetidas con RPA	Reestructuración de la matriz de riesgo y metodología EBR	Modelo de ML con perfilamiento de riesgo dinámico	
Eficiencia operacional	5	0	1	1	1	-1	1	
Reducción de sesgo humano	3	-1	1	1	1	-1	1	
Precisión en la detección de riesgos	5	0	0	0	0	0	1	
Adaptabilidad a cambios regulatorios	5	1	0	0	0	1	1	
Costo de implementación y mantenimiento	1	0	-1	-1	-1	0	0	
Escalabilidad	3	0	0	0	1	0	1	
	<i>Nro. de 1</i>	1	2	2	3	1	5	<i>Mejor</i>
	<i>Nro. de 0</i>	4	3	3	2	3	1	<i>Igual</i>
	<i>Nro. de -1</i>	1	1	1	1	2	0	<i>Peor</i>
Puntuaciones base de la solución		0	1	1	2	-1	5	
Puntuaciones ponderadas de la solución		2	7	7	10	-3	<b>21</b>	

*Nota: la puntuación ponderada resulta de la suma del producto ente el puntaje de la alternativa (-1,0,1) y el nivel de impacto del criterio*

### 2.2.2. Matriz Impacto - Esfuerzo

Como complemento al análisis multicriterio de Pugh, se utilizó la matriz de impacto-esfuerzo para evaluar la viabilidad de implementación de las alternativas, para la construcción de esta matriz, se desglosaron variables de impacto y esfuerzo detalladas en la Tabla 10: Desglose de puntuación por características de impacto y esfuerzo, ubicada en anexos, luego con base a los promedios resultantes las alternativas fueron ubicadas en cuatro cuadrantes.

**Figura 2:** Matriz Impacto - Esfuerzo



**Tabla 4** Clasificación de propuestas según impacto y esfuerzo

Nro.	Alternativas	Impacto estimado	Esfuerzo estimado
1	Optimización manual y capacitación	2,67	1,67
2	Integración de fuentes externas	4,00	4,00
3	Digitalización de formularios	4,33	4,00
4	Automatización con RPA	4,00	2,67
5	Reestructuración de matriz de riesgo	2,33	4,00
6	Modelo ML con perfilamiento de riesgo	4,67	2,00

Como resultado del análisis, se selecciona una solución híbrida compuesta por un modelo Machine Learning y automatización con RPA, ambas posicionadas en el cuadrante de “ganancias rápidas” donde el modelo ML destaca con el mayor impacto estimado de (4,67) y el menor índice de esfuerzo estimado (2) el bajo puntaje de esfuerzo no se refiere a una simplicidad técnica, sino a la eficiencia de implementación, al utilizar ecosistemas como Python y Google Colab sin licencias costosas. En cambio, la alternativa 5 “reestructuración de la matriz de riesgo” aunque de manera técnica es sencilla, implica un esfuerzo organizacional alto, dado a la oposición al proceso de cambio y la demanda de capacitación masiva.

La propuesta final consiste en desarrollar un código con RPA encargado de buscar automáticamente el formulario de vinculación “conozca a su cliente” y extraer la información estructurada, esos datos servirán como input para el modelo de ML, en el cual procesará la información para generar un perfilamiento de riesgo dinámico en tiempo real. Esta arquitectura garantiza que la aseguradora maximice la precisión en la detección de posibles casos LA/FT, al tiempo que minimiza la carga operativa manual y el sesgo por error humano.

### **2.3. Enfoque metodológico**

El presente estudio adoptó un enfoque metodológico mixto o híbrido, integrando la estructura estratégica de CRISP-DM (Cross-Industry Standard Process for Data Mining) como el marco principal que guía el desarrollo del proyecto. Paralelamente, se incorporan los principios de BPM (Business Process Management) con el propósito de asegurar de manera correcta la integración del modelo en el proceso operativo de la aseguradora.

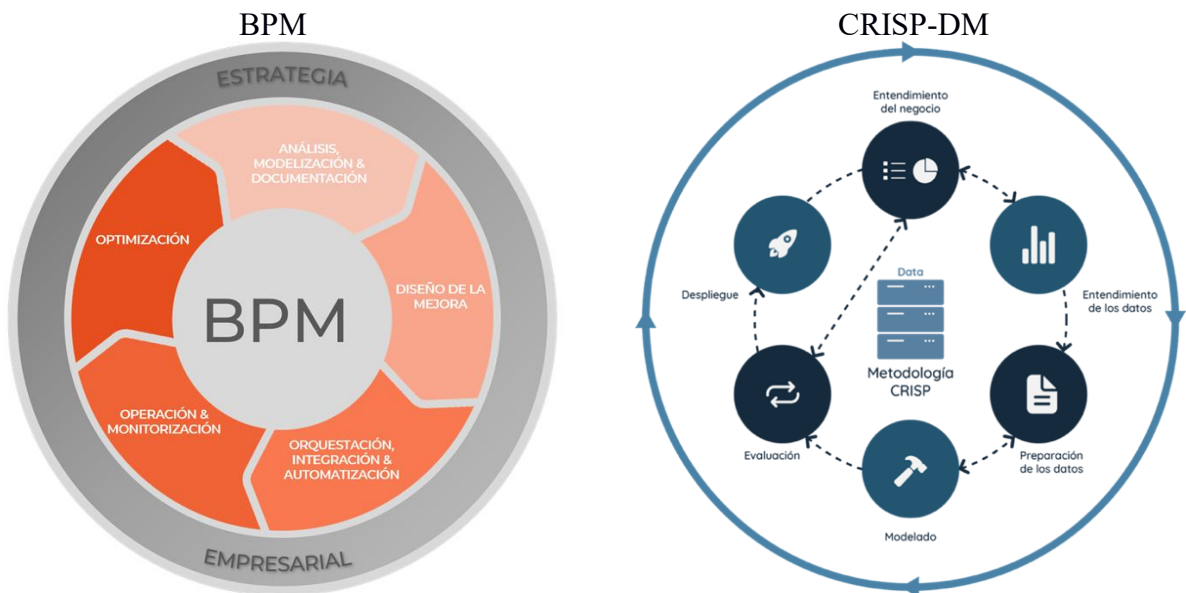
Esta combinación de metodología permitió que la propuesta, apoyada con herramientas tecnológicas avanzadas no solo respondiera a la complejidad operativa, sino también a los requisitos normativos como LOPDP y estándares aplicables al proceso de debida diligencia que garantizó un equilibrio óptimo entre su cumplimiento legal y funcionalidad práctica.

## 2.4. Diseño metodológico

La aplicación de este enfoque inició con la metodología BPM para realizar un diagnóstico inicial del proceso de debida diligencia y la gestión de riesgos PLA-FT. Mediante entrevistas con los analistas de cumplimiento se permitió identificar los puntos de dolor y cuellos de botellas que afronta esta área en sus operaciones diarias, identificando que el procesamiento de la información del cliente se realiza 100% de manera manual, lo cual resulta lento y propenso a errores humanos. Además, que existe una automatización insuficiente en el perfilamiento de riesgo de los clientes lo cual genera una elevada tasa de falsos positivos.

Dicho análisis, permitió diseñar un proceso To-Be optimizado que contemple dos mejoras claves: el uso de un modelo de Machine Learning para el perfilamiento del riesgo de los clientes y una automatización en el procesamiento de la información en la matriz de riesgo. Una vez definido los requerimientos desde la perspectiva del BPM, se procedió con la metodología CRISP-DM para guiar el desarrollo técnico del modelo de Machine Learning

**Figura 3:** Metodologías aplicadas



*fases de la metodología BPM*

*fases de la metodología CRISP-DM*

Este ciclo de vida se ejecutó dentro del ecosistema de Google (Drive, Sheets, Colab, etc.), con el fin de optimizar recursos y facilitar la realización de pruebas piloto. Dichas pruebas permitieron recopilar datos cuantitativos sobre el desempeño del modelo de Machine Learning, los cuales servirán como sustento para justificar una futura inversión en infraestructura tecnológica basada en este tipo de herramientas avanzadas.

El proceso se estructuró en varias fases, que incluyeron la comprensión del negocio, el análisis comparativo de los resultados obtenidos a partir de la matriz de riesgo tradicional frente a aquellos generados por el modelo predictivo, la preparación de los datos, el modelado, la evaluación y, finalmente, la implementación. A partir de esta metodología, se desarrolló un plan de pruebas comparativo que permitió contrastar los resultados obtenidos de la matriz de riesgo con aquellos generados por el modelo predictivo de ML.

#### ***2.4.1. Fase 1: Diagnostico operativo (BPM) – Comprensión del negocio (CRISP-DM)***

En esta fase se ejecutó el levantamiento del proceso actual (As-Is) de debida diligencia (DDC) a través de entrevistas con los dos oficiales de cumplimiento. A partir de este análisis, se identificaron las brechas operativas que limitan la eficiencia del control PLA/FT y se elaboró la Tabla 11: Matriz de hallazgos operativos detallada en anexos, la cual sintetiza los puntos de dolor detectados.

Los principales puntos críticos se relacionan con la subjetividad del perfilamiento de riesgo y la digitación manual de la información. En consecuencia, desde la perspectiva del enfoque CRISP-DM se definen metas duales, por un lado, automatizar el llenado de la información del cliente en la matriz de riesgo para reducir los tiempos de procesamiento y por el otro, proponer un modelo de Machine Learning que supere la eficacia del sistema actual, con el fin de disminuir los falsos positivos.

#### ***2.4.2. Fase 2: Análisis de datos (BPM) - Comprensión de los datos (CRISP-DM)***

En esta fase se procedió a la recolección de los expedientes digitales, formularios de vinculación y documentos PDF que fueron depositados en un repositorio seguro en Google Drive con el fin de garantizar un acceso centralizado. De manera adicional, se llevó a cabo un análisis exploratorio de datos (EDA) con el objetivo de evaluar las propiedades fundamentales y calidad de las variables.

En cumplimiento con la ley orgánica de protección de datos personales (LOPD), se aplicó un protocolo de pseudonimización que permitió disociar los datos identificativos de los clientes y una minimización de datos para solo elegir las variables de análisis (Ingresos, Geografía, etc.) para poder llevar a cabo el prototipo funcional. Asimismo, en esta fase se desarrolló una investigación técnica sobre el uso de las APIS de Google para evaluar la viabilidad de extraer la información necesaria desde los documentos PDFs para la automatización.

#### ***2.4.3. Fase 3: Preparación de los datos***

Durante esta etapa se desarrolla la solución del proyecto, la cual está compuesta por dos componentes claves. Inicialmente se desarrolla un script de automatización en el entorno de Google Colab empleando Python, el cual no solo extrae los datos del cliente del formulario KYC, sino que también gestiona de forma integral la consulta de APIs y el llenado estandarizado de la información en la matriz de riesgo en Google Sheets.

Posteriormente, se implementó un segundo script especializado en el preprocesamiento de los datos para transformar la información en un dataset depurado y coherente para alimentar el entrenamiento del modelo XGboost. Cabe destacar que esta tarea de limpieza no es solo por la eficiencia técnica, sino que también actúa como un control de integridad, al asegurar la fiabilidad

de los datos financieros y no financieros. Este proceso se alinea directamente con los principios de información y comunicación establecidos por el marco COSO III.

#### **2.4.4. Fase 4: Motor de decisión (BPM) - Modelado (CRISP-DM)**

En esta fase, con un conjunto de datos preparados se enfocó en seleccionar, entrenar y optimizar el algoritmo predictivo, dándole prioridad al uso de XGboost para el rediseño operativo (To-Be) este modelo es reconocido por su alto rendimiento con datos estructurados como los que componen la matriz de riesgo, que se busca minimizar el riesgo residual (ISO 31000) priorizando la detección de casos sospechosos. Asimismo, el entrenamiento se realizó en Google colab para aprovechar los recursos computacionales disponibles.

**Tabla 5:** Configuración del modelo

<b>Componente</b>	<b>Especificación técnica</b>	<b>Justificación</b>
Algoritmo seleccionado	XGboost	Equilibrio entre rendimiento y velocidad con matrices de riesgo estructuradas
Entorno de ejecución	Google Colab	Maximizar el uso de recursos disponibles

#### **2.4.5. Fase 5: Validación Normativa (BPM) - Evaluación (CRISP-DM)**

En esta fase el modelo entrenado fue sometido a una evaluación rigurosa para validar su rendimiento y asegurar que se estén cumpliendo los objetivos iniciales del área de cumplimiento. En esta etapa se utilizó un conjunto de datos para poder obtener las métricas necesarias y así justificar de manera cuantitativa su rendimiento. Además, se aplicó técnicas como SHAP para interpretar sus decisiones y así obtener resultados cualitativos que justifiquen la transparencia del modelo, para garantizar el cumplimiento del principio de transparencia de la LOPDP y el modelo COSO.

#### **2.4.6. Fase 6: Implementación (BPM) - Despliegue (CRISP-DM)**

La etapa final consistió en la integración de los distintos componentes en un prototipo plenamente funcional. Al tratarse de un proyecto de grado el alcance se delimitó a un entorno controlado, logrando combinar con éxito la automatización de la matriz de riesgo con el motor inteligente de perfilamiento de riesgo. Esta sinergia no solo permite tener un control más riguroso de la cartera de clientes, sino que también entrega un score de riesgo más preciso que fundamenta la toma de decisiones por parte del jefe de cumplimiento.

Este despliegue no solo optimiza la operatividad del proceso, sino que fortalece el ambiente de control interno por el modelo COSO al facilitar un monitoreo continuo y reportes estandarizados.

#### **2.5. Consideraciones éticas, legales y de seguridad de la información**

El tratamiento de los datos personales tanto para el entrenamiento predictivo como para el rediseño del proceso se ajustará a la LOPDP (2021) y a las directrices de la resolución SCVS-INC-DNCDN-2023-0002. Para garantizar el cumplimiento se aplicaron mecanismos de pseudonimización y minimización, reforzados por un control de accesos basado en roles y por políticas de la UAFE. Cabe destacar que la prueba piloto se desarrolla en un entorno segregado, utilizando datos históricos anonimizados que excluyen cualquier toma de decisión automatizada carente de revisión humana.

El marco metodológico propuesto combina con éxito la gestión por procesos (BPM) con la rigurosidad analítica de datos (CRISP-DM) para mejorar el proceso DDC. Integra el ecosistema de Google, métricas de evaluación y el cumplimiento de los marcos LOPDP y COSO III, lo que asegura que la solución final no sea solo técnicamente sólida, sino que también sea éticamente responsable y sostenible.

## **Capítulo 3**

### 3. Resultados y análisis

La implementación de este marco metodológico híbrido (BPM y CRISP-DM) permitió la creación de un modelo integral específicamente diseñado para transformar el proceso de debida diligencia en la aseguradora. Esta solución se traduce en dos componentes tecnológicos que operan de manera coordinada para dinamizar el flujo operativo: en primera instancia, un script de automatización para la migración e integración de datos desde el formulario KYC hacia la matriz de riesgo corporativa, y, en segunda instancia, la implementación de un algoritmo de Machine Learning configurado en la ponderación dinámica y precisa de riesgo de los clientes para su clasificación técnica (alta, media, baja).

El presente capítulo sistematiza el análisis de los resultados obtenidos durante la implementación de los dos componentes tecnológicos. Con el fin de evaluar la efectividad de la solución bajo criterios de objetividad e imparcialidad, se diseñó un esquema experimental compuesto por 15 iteraciones de pruebas piloto controladas con una muestra de 300 clientes que fue procesada de manera secuencial y proporcionada por el área de cumplimiento. Debido a las limitaciones de la APIs de Google, que restringe el procesamiento a un máximo de 20 registros diarias, esta fase se extendió a un período de dos semanas, lo que permitió validar la estabilidad del sistema y garantizar la integridad de los datos a través de múltiples ciclos de automatización.

La validación del modelo se fundamentó mediante un enfoque comparativo de “antes y después” (*As-Is vs To-Be*), lo que permitió contrastar indicadores clave de desempeño (KPIs) cuantitativos. Para la automatización operativa, se evaluaron métricas tales como: tiempos, tasa de errores de escritura y el nivel de autonomía operativa, por otro lado, para la eficacia del modelo predictivo los indicadores son: tasa de falsos positivos, precisión, accuracy y capacidad operativa diaria. Este diagnóstico se complementó con una dimensión cualitativa integrando la

valoración técnica de los analistas de cumplimiento para validar la utilidad de los nuevos artefactos tecnológicos

### **3.1. Evaluación del desempeño operativos: comparativa de línea base vs. resultados de prueba piloto**

#### ***3.1.1. Resultados de la evaluación inicial (línea base As- Is)***

El diagnóstico del estado inicial se realizó bajo la guía de la metodología BPM, lo que permitió cuantificar el flujo de trabajo e identificar parámetros críticos sobre la operatividad del proceso. Los hallazgos revelaron que el proceso depende en gran medida de intervención manual, lo que compromete la agilidad del proceso y la toma de decisiones estratégicas.

Esta carga operativa se traduce en la alta demanda por cliente debido a que, la recopilación de información y el llenado de datos en la matriz de riesgo tiene un tiempo promedio de 45 minutos. De hecho, tras el análisis exploratorio de registros histórico se reveló una tasa de error del 10% en transcripción de información. Sin embargo, el punto más crítico se identificó a nivel predictivo, en donde las reglas vigentes arrojan una tasa de falsos positivos superior al 85%, esta ineficiencia satura la capacidad del analista de cumplimiento, al obligarlo a realizar investigaciones exhaustivas sobre alertas que, en su mayoría de casos carecen de valor real para la gestión de riesgos. Como hallazgo final, se identificó la ausencia de trazabilidad en el almacenamiento y flujo documental, debido a la falta de registros del estado de procesamiento de cada expediente, esta situación genera riesgos de omisión o duplicidad de información, lo cual incrementa la vulnerabilidad del proceso frente a normativas de seguridad y protección de datos personales.

### ***3.1.2. Resultados de prueba piloto***

El primer componente desarrollado fue el script de automatización orientado al procesamiento de datos del cliente, esta herramienta fue diseñada para poblar la matriz y reducir los tiempos y aumentar la confiabilidad de los datos.

Durante las pruebas controladas, el script logro interactuar con la API de Géminis 1.5 pro, la cual cuenta con una capacidad de reconocimiento óptico de caracteres (OCR), que permitió la extracción de los datos del formulario KYC hacia la matriz de riesgo en un tiempo promedio de 60 segundos. Es importante señalar que en la matriz no solo se pondera datos del formulario, sino también datos del sistema interno como ramo del negocio, suma asegurada comercializada y canal de distribución. Este desempeño representa una reducción del tiempo de ejecución superior al 90% respecto a la línea base, además se logra una tasa del 100% en la eliminación de errores por transcripción manual, lo que asegura la consistencia necesaria para las etapas analíticas posteriores.

En paralelo, el segundo artefacto tecnológico se materializa en un modelo de Machine Learning basado en el algoritmo XGBoost, este motor inteligente de ponderación de riesgo fue sometido a una etapa de entrenamiento y validación rigurosa con registros de datos históricos de la compañía, lo que permite sustituir el sistema estático basado en reglas por un modelo predictivo dinámico de alta precisión, mitigando la subjetividad detectada en el análisis inicial. La implementación de la solución integral mitiga la falta de trazabilidad al centralizar el flujo documental en un entorno digital y automatizar la transición de expedientes desde la carpeta “01 formularios nuevos” hacia la carpeta “02 formulario procesados”, reduciendo el riesgo de que el analista olvide registrar un asegurado o se procese de manera duplicada. Con ello, se fortalece la auditabilidad de los expedientes y asegura el cumplimiento de la LOPDP y del marco COSO III.

La metodología validación se realizó mediante sesiones de 30 min por zoom, en la que se ejecutaron los nuevos artefactos tecnológicos a escenarios reales y como parte de la solución To-Be se rediseño la matriz de riesgo bajo una estructura de ponderación dinámica que evalúa los cuatro factores exigidos por estándares internacionales factor persona, producto, canal y jurisdicción (Las 40 Recomendaciones del GAFI, 2025) . A continuación, en la Tabla 6 análisis comparativo se sintetiza los indicadores de desempeño obtenidos tras la fase de validación, detallando el impacto y las mejoras alcanzadas.

**Tabla 6:** Análisis comparativo del proceso de debida diligencia (As-Is vs To-Be)

<b>Métrica</b>	<b>Proceso Manual (As-Is)</b>	<b>Solución Tecnológica (To-Be)</b>	<b>Impacto / Mejora</b>
<b>Tiempo de procesamiento</b>	45 minutos promedio por cliente.	10 minuto promedio por cliente	Reducción del 80% en tiempo para procesar
<b>Tasa de errores de escritura</b>	10% de error en transcripción manual.	90% de error de transcripción (Uso de APIs con OCR)	Eliminación del 100% de fallas por digitación manual e integridad en los datos Reducción de carga operativa con
<b>Nivel de autonomía operativa</b>	Dependencia total de intervención manual	80% automatización de la matriz de riesgo	intervención manual limitada a variables no digitalizadas como el ramo y la suma asegurada
<b>Capacidad operativa diaria</b>	Aproximadamente 20 expedientes dentro de la jornada laboral	Incremento del 50% de expedientes	Crecimiento de la capacidad de respuesta institucional ante el riesgo del asegurado, si se procede con la relación comercial
<b>Gestión de almacenamiento</b>	Carpetas compartidas sin orden ni estructura secuencial	Flujo automatizado por carpetas “01 nuevos / 02 procesados”	Trazabilidad del 100% del estado del expediente, reducción de riesgos por pérdida o duplicidad de la información.
<b>Nivel de subjetividad</b>	Ponderación basada en el criterio subjetivo del analista.	Ponderación objetiva y dinámica mediante algoritmo de aprendizaje supervisado.	Estandarización del perfilamiento de riesgo bajo metodología enfoque basado en riesgo (EBR)
<b>Tasa de faltos positivos</b>	80% - 85% de alertas de bajo valor.	Reducción al 20% mediante modelo predictivo XGBoost.	Mayor enfoque en riesgos reales y reducción de investigaciones exhaustivas.

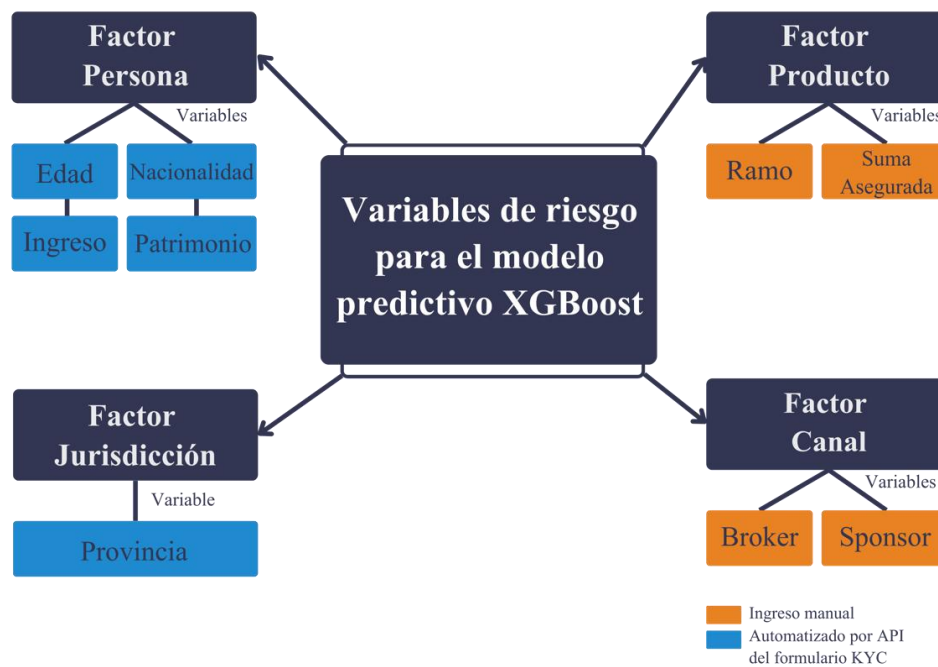
## 3.2. Factores de riesgos y variables del modelo predictivo

### 3.2.1. Muestra de estudio

El conjunto de datos seleccionados corresponde a 300 observaciones de registros de clientes, los cuales fueron sometidos a un proceso de pseudonimización y minimización para blindar la privacidad de la información sensible. La determinación de este volumen de datos no fue un proceso aleatorio, sino que corresponde a la necesidad de dotar al algoritmo XGBoost de una base suficientemente amplia para alcanzar un nivel predictivo alto sin incurrir en el sobreajuste.

Asimismo, el modelo se nutre de un conjunto de variables estratégicas que aportan la estabilidad estadística necesaria para mitigar el impacto de anomalías puntuales. Esta configuración no solo garantiza la efectividad del sistema, sino que garantiza la capacidad para generalizar los resultados hacia la población total de asegurados, asegurando la fiabilidad del modelo en escenarios de riesgos reales.

**Figura 4:** Mapa Jerárquico de Factores y Variables

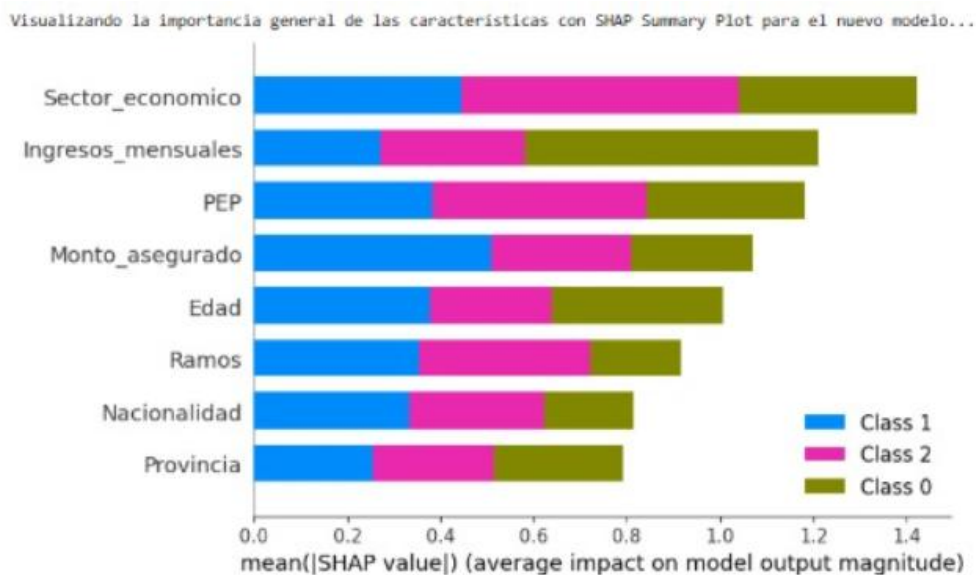


Como se ilustra en la figura 4, el modelo predictivo se fundamenta bajo una estructura de riesgo diseñada para la prevención del LA-FT que integra automáticamente los factores de cliente y zona geográfica desde el formulario KYC, mientras que los factores de producto y canal de distribución requieren de una intervención manual por parte del analista, dado que las variables de estos factores como ramo, suma asegurada, bróker no constan en el formulario. Esta configuración permite que el algoritmo XGBoost genere un score de riesgo que se fundamenta en información exhaustiva y veraz.

### 3.2.2. *Importancia de variables estratégicas*

Con el objetivo de mitigar el fenómeno de caja negra propiamente de XGBoost y de dotar de transparencia al sistema, se profundizó en la arquitectura interna del modelo mediante la técnica SHAP, la cual permitió identificar los factores con mayor peso específico en la generación de predicciones que se desglosan en la figura 5. Estos hallazgos son de suma importancia puesto que el algoritmo ha logrado asimilar y priorizar atributos que guardan una correlación directa con los pilares preventivos de lavado de activos.

**Figura 5:** *Características de variables con SHAP*



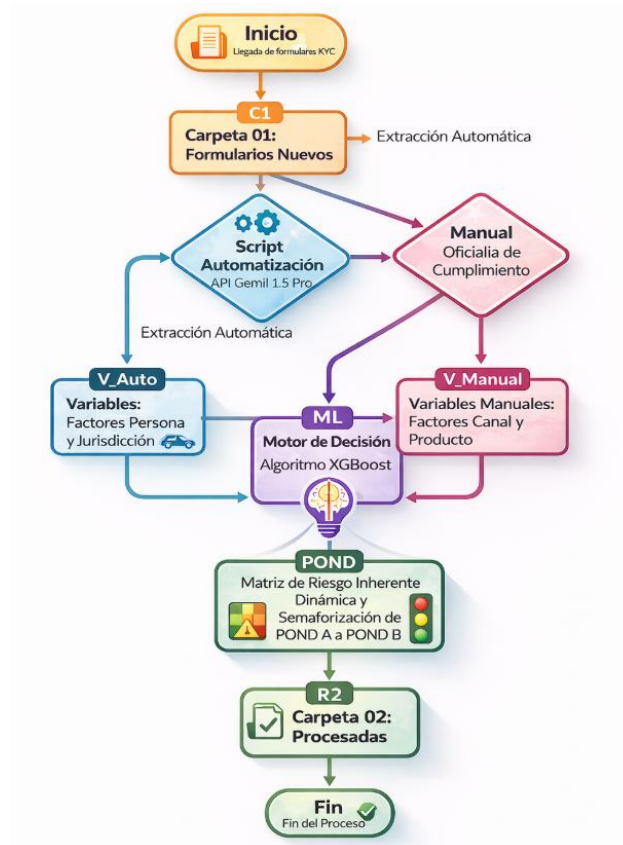
Bajo este análisis, la jerarquía de relevancia sitúa a la variable de sector económico y a la cantidad de ingresos mensuales como los vectores determinantes en la calificación del riesgo, seguidos con un peso significativo, la vinculación con figuras PEP y el monto asegurado. Esta configuración aporta robustez a la interpretabilidad del proceso y funciona como una validación empírica del sistema, por lo que se confirma de este modo que el sistema se encuentra alineado con estándares internacionales en gestión de riesgos y prevención de delitos financieros.

### 3.3. Diseño y arquitectura de la solución

#### 3.3.1. Arquitectura de la solución automatizada y predictiva

La arquitectura de la solución aprovecha las capacidades del entorno de Google para una implementación eficiente y escalabilidad, sin la necesidad de incurrir en costos adicionales de infraestructura física o de servidores. A continuación, se describe en el diagrama de flujo

Figura 6: Diagrama de flujo



Como se describe en el diagrama de flujo, el ciclo de vida de la información inicia con una captura estructurada de datos mediante el formulario KYC alojado en Google Drive; posteriormente, se ejecuta el script de automatización y ponderación de riesgo, tras lo cual los formularios pasan a nueva carpeta, mejorando la trazabilidad del proceso.

Este registro inicial actúa como un disparador para el Script alojado en Google Colab que desempeña el rol de una capa de automatización robótica de procesos, la cual centraliza la información, interactúa con APIs externas para el enriquecimiento de datos en tiempo real y consolida la matriz de riesgo que alimenta al motor predictivo XGBoost para generar un score de riesgo predictivo.

### **3.4. Evaluación de la eficacia predictiva del modelo (To-be)**

#### ***3.4.1. Análisis cuantitativo***

Con el propósito de evaluar la eficacia del modelo XGBoost, bajo condiciones de rigor científico, se reservó un conjunto de datos independiente de  $n=60$ , equivalentes al 20% de la muestra original. Este subset se mantuvo completamente aislado del proceso de entrenamiento, garantizando así una validación imparcial de su capacidad de perfilamiento de riesgo, al contrastar este desempeño con la línea base de reglas estáticas, se evidenciaron notables mejoras cuantitativas en precisión, lo que evidencia la superioridad técnica para identificar patrones de riesgo complejos que el método tradicional omitía.

### **3.5. Evaluación cualitativa de los analistas de cumplimiento**

Como complemento al análisis cuantitativo, se aplicaron entrevistas a los analistas de cumplimiento tras la culminación de la fase piloto. De este ejercicio surgieron dos dimensiones claves que sustentan la adopción del sistema desde la experiencia del usuario: la calidad percibida en la toma de decisiones y la mitigación de la carga cognitiva.

### ***3.5.1. Calidad percibida en la toma de las decisiones***

La retroalimentación de los participantes confirmó que la solución implementada impulsa la mejora analítica, al automatizar tareas repetitivas se liberó una cuota significativa de tiempo y capacidad cognitiva que se consumían en labores de bajo valor, este cambio no es solo operativo sino funcional puesto que, transforma la naturaleza del rol del analista, pasando de ser un recolector de información a un investigador estratégico.

Esta transición se materializa en el testimonio de un analista quien destaca que, al delegar el trabajo pesado al prototipo, el tiempo anteriormente consumido en la transcripción manual se redirige ahora hacia al análisis crítico. De este modo, su juicio de experto se centra en la interpretación profunda y síntesis final de hallazgos.

### ***3.5.2. Reducción drástica de la carga cognitiva***

La drástica reducción de falsos positivos tuvo un impacto positivo tangible en la moral y bienestar del equipo. Los analistas describieron que el sistema convencional como una fatiga de alertas, un fenómeno donde la exposición constante a alertas irrelevantes disminuye la atención y aumenta la probabilidad de pasar por alto un riesgo genuino. Sin embargo, gracias a la precisión del prototipo funcional, esta dinámica diaria se transformó por completo, sustituyendo el agotamiento por un enfoque analítico más certero.

## **3.6. Análisis costo beneficio de la solución**

Con el propósito de validar la viabilidad financiera y el valor estratégico de la solución propuesta, se desarrolló un análisis de costo-beneficio. Dicho análisis permitió conocer los recursos requeridos para una implementación a una escala corporativa e identificar los beneficios tangibles e intangibles derivados de la automatización, lo que demuestra el retorno de una inversión favorable para la aseguradora.

### 3.6.1. Análisis de costos

Los costos asociados al prototipo funcional se mantuvieron en niveles mínimos gracias al uso de herramientas del ecosistema de Google y pruebas gratuitas, pero la viabilidad del proyecto exige una proyección hacia una escala corporativa. Para este cálculo se toma como referencia la estructura actual del área, conformada por dos oficiales de cumplimiento quienes gestionan un volumen de 60 revisiones diarias. Bajo este escenario, la transición hacia el modelo automatizado no solo representa un ahorro en infraestructura tecnológica, sino una optimización en la capacidad actual, permitiendo que el equipo absorba una mayor carga operativa sin la necesidad de incrementar gastos en capital humano.

El factor crítico en la estructura de costos de la automatización es el uso de APIs especializadas en el reconocimiento de caracteres (OCR), este requerimiento surge de una limitación detectada en los archivos que se reciben por parte de los clientes, debido a que el 40% son documentos escaneado y al usar técnicas convencionales de extracción de texto, resultaría ineficaz ante estos archivos de imagen, invalidando el procesamiento automático. Por lo que, a continuación, se presenta en la Tabla #7, una comparación entre las capacidades nativas empleadas en el prototipo y las prestaciones de APIs especializadas en OCR

**Tabla 7:** Descripción de APIs

Descripción de APIs	Gemini 1.5 Pro (En uso)	Azure Document Intelligence	Google Document AI
Costo estimado por 1000 paginas	\$3,50	\$50	\$30
Precisión de OCR	Media	Alta	Muy alta
Cantidad de procesamiento diario	20	100	100
Tipo de procesamiento admitido	Sincrónico	Sincrónico - Asincrónico	Asincrónico

Dicha comparación no solo justifica la elección actual, sino que establece una ruta para una futura transición hacia servicios de reconocimiento de caracteres con mayor capacidad de procesamiento. Por lo tanto, la inversión en este tipo de servicio no es opcional, sino una condición necesaria para garantizar el funcionamiento óptimo del script y la captura de información de manera adecuada

### **3.6.2. Análisis de beneficios**

Los beneficios derivados del prototipo se pueden dividir en dos categorías. Por un lado, se identifican los beneficios cuantitativos, los cuales permiten realizar una medición directa del ahorro operativo y el retorno de la inversión. Por otro lado, también se identifican beneficios cualitativos, aunque menos tangibles, pero de alto valor estratégico porque son fundamentales para optimizar la calidad del servicio y fortalecer el cumplimiento normativo dentro de un entorno de gestión de riesgo cada vez más exigente

El análisis cualitativo revela que esta solución va más allá de la optimización operativa, sino que representa un salto de madurez organizacional del sistema de cumplimiento, al combinar tecnología, talento humano y marcos de gobernanza. Lo que lo convierte en un eje estratégico frente a riesgos regulatorios y reputacionales.

**Tabla 8:** *Análisis de beneficios cualitativos*

<b>Beneficios Cualitativos</b>	<b>Descripción</b>	<b>Valor para la organización</b>
Cumplimiento y riesgos regulatorios	Fortalecimiento del cumplimiento normativo con detección de riesgo más preci	Reduce la probabilidad de sanciones regulatorias, multas millonarias y daño reputaciona
Auditabilidad del proces	Evaluación de clientes con decisiones registradas y explicables	Simplificación de auditorías p evidencia solida de control y gobernanza

Capacidad de detección temprana	Detección proactiva de operaciones de lavado de activos.	Mayor confianza de clientes y entes reguladores
Enfoque en tareas estratégicas	Investigaciones más complejas y de carácter estratégico	Mayor calidad analítica

El análisis cuantitativo se obtuvo a partir de la comparación entre el proceso operativo manual y el flujo automatizado propuesto, para esta comparativa se utilizaron mediciones críticas de tiempo de ejecución, tiempo de procesamiento, carga operativa y tasa de falsos positivos. Por lo que se evidencia que la aplicación de esta solución genera una mejora sustancial en el proceso inicial de evaluación del cliente.

*Tabla 9: Análisis de beneficios cuantitativos*

<b>Beneficios Cuantitativos</b>	<b>Descripción</b>	<b>Valor para la organización</b>
Eficiencia comercial	Reducción del tiempo en emisión de pólizas	Agilización de nuevos clientes sin incremento proporcional de recursos
Tiempo operativo por cliente	Disminución en un 80% procesamiento, pasando de 45 a 10min.	Liberación de capacidad operativa para tareas más analíticas
Carga operativa mensual	Ahorro estimado de 100 horas mensuales de trabajo operativo.	Optimización del tiempo del personal Uso más eficiente de los recursos analítico y
Investigaciones improductivas	Reducción de la tasa de falsos positivos del 85% al 20%.	disminución del desgaste operativo

## **Capítulo 4**

## **4. Conclusiones y recomendaciones**

### **4.1. Conclusiones**

Tras el análisis integral de los resultados expuestos, se derivan las siguientes conclusiones, las cuales están relacionadas directamente con los objetivos que dieron origen a esta investigación, consolidando una propuesta de valor que responde a las necesidades actuales de la institución en materia de cumplimiento.

En primer lugar, se concluye que el proceso de debida diligencia manual es propenso a errores, como la recopilación de información y transcripción manual de datos. Este proceso manual no solo fue identificado como un cuello de botella, sino como una debilidad en el marco de control de riesgo de la institución.

En segundo lugar, se concluye que el modelo XGBoost es capaz de identificar y ponderar factores de riesgos de manera más dinámica y precisa a comparación del sistema convencional

En tercer lugar, se concluye que la viabilidad de diseñar soluciones de alto impacto no está condicionada a grandes presupuestos, sino a la integración de herramientas accesibles, la arquitectura desarrollada que integra un motor de automatización con un modelo predictivo ha demostrado su capacidad operativa capaz de rivalizar con sistemas más complejos. Por lo que este hallazgo es relevante, pues valida que la transformación digital en el área de cumplimiento puede iniciarse mediante una innovación de bajo coste.

Para finalizar, se concluye que la aplicación de esta solución híbrida supera de manera significativa el proceso tradicional en eficacia y en eficiencia. La evidencia cuantitativa demostró que se alcanzó automatización del 80% de la matriz de riesgo y perfilamiento dinámico del score de riesgo, lo que se traduce en tiempo de procesamiento de clientes más rápido, en la eliminación de los errores de transcripción y la reducción de falsos positivos. Demostrando que la solución

no solo optimiza recursos, sino que fortalece la capacidad de detección, haciendo que la función del área sea más rápida, fiable y precisa.

#### **4.2. Recomendaciones**

Se recomienda incluir a más analistas en el uso del prototipo funcional para recopilar retroalimentación cualitativa adicional, lo que va a permitir afinar la interfaz de usuario y optimizar el flujo de trabajo, para que la herramienta se adapte a las necesidades operativas.

Se propone iniciar la planificación técnica para la migración de la solución a entornos más robusto para garantizar la escalabilidad, la seguridad y gobernanza de los datos requeridos para operar en entornos más robustos como Google Cloud Platform o Microsoft Azure.

Se sugiere realizar un análisis exhaustivo de la estacionalidad de la carga operativa para optimizar el gasto en servicios de terceros. Por lo que, en lugar de tener una infraestructura de alto coste, se recomienda realizar una contratación de API escalables donde el volumen de procesamiento de caracteres se ajuste al flujo mensual de clientes.

Se recomienda establecer un protocolo formal para el monitoreo continuo del XGBoost, asegurando su fiabilidad a largo plazo. Este marco debería de incluir métricas precisión y reentrenamiento sistemático del modelo con nuevos datos ante nuevos tipos de riesgos identificados.

## 5. Referencia

GAFILAT. (2023). *INFORME DE EVALUACIÓN MUTUA DE LA REPÚBLICA DEL ECUADOR*.

UAFE, U. d. (2024). *UAFE, Económico Unidad de Análisis Financiero y*. Obtenido de <https://www.uafe.gob.ec/uafe-presento-los-resultados-de-la-evaluacion-nacional-de-riesgos-de-lavado-de-activos-y-financiamiento-del-terrorismo-del-ecuador/>

FGE, F. G. (2025). *6 personas son vinculadas a causa por presunto peculado en Petroecuador*. Obtenido de Fiscalía General del Estado: <https://www.fiscalia.gob.ec/accesibilidad/6-personas-son-vinculadas-a-causa-por-presunto-peculado-en-petroecuador/>

FGE, F. G. (2021). *Caso Seguros Sucre: 10 años de prisión para exgerente y su esposa por lavado de activos*. Obtenido de Fiscalía General del Estado: <https://www.fiscalia.gob.ec/accesibilidad/caso-seguros-sucre-10-anos-de-prision-para-exgerente-y-su-esposa-por-lavado-de-activos/>

FGE, F. G. (2021). *Fiscalía ha presentado 14 testigos en el juicio contra exgerente de Seguros Sucre*. Obtenido de Fiscalía General del Estado: <https://www.fiscalia.gob.ec/accesibilidad/fiscalia-ha-presentado-14-testigos-en-el-juicio-contr-exgerente-de-seguros-sucre/>

Gaviyau, G., & Sibindi, A. B. (2023). *Prevención del blanqueo de capitales y debida diligencia del cliente: evidencia empírica de Sudáfrica*. Obtenido de Revista de Control del Lavdo de Dinero: <https://www.emerald.com/jmlc/article/26/7/224/238224/Anti-money-laundering-and-customer-due-diligence>

Milojevic & Redzepagic, N. M. (2021). *Prospects of Artificial Intelligence and Machine Learning Application in Banking Risk Management*. Obtenido de Journal of Central

Banking Theory and Practice: <https://reference-global.com/article/10.2478/jcbtp-2021-0023?tab=abstract>

Xiangting Shi, Y. Z. (2025). *Deep learning for enhanced risk management: a novel approach to analyzing financial reports*. Obtenido de PeerJ Computer Science: <https://peerj.com/articles/cs-2661/>

Prisznyák, A. (2022). *Bankrobotics: Artificial Intelligence and Machine Learning Powered Banking Risk Management*. Obtenido de Public Finance Quarterly: <https://journals.lib.uni-corvinus.hu/index.php/penzugyiszemle/article/view/1194>

Ley Orgánica de Prevención, D. y. (2025). *LEY PREVENCIÓN DE LAVADO DE ACTIVOS Y DEL FINANCIAMIENTO DE DELITOS*. Obtenido de LEXIS: <https://procuraduria.utpl.edu.ec/NormativaExterna/LEY%20PREVENCION%20DE%20LAVADO%20DE%20ACTIVOS%20Y%20DEL%20FINANCIAMIENTO%20DE%20DELITOS.pdf>

Resolución No. JPRF-V-2022-024. (2022). *LA JUNTA DE POLÍTICA Y REGULACIÓN FINANCIERA*.

Resolución No.SCVS-INC-DNCDN-2023-0002. (2023). *Superintenden de compañías, valores y seguros*.

The FATF Recommendations. (2025). *INTERNATIONAL STANDARDS ON COMBATING MONEY LAUNDERING AND THE FINANCING OF TERRORISM & PROLIFERATION*.

Howard Chitimira, E. T. (2024). *Leveraging Artificial Intelligence to Combat Money Laundering and Related Crimes in the South African Banking Sector*. [https://www.researchgate.net/publication/383967945\\_Leveraging\\_Artificial\\_Intelligence](https://www.researchgate.net/publication/383967945_Leveraging_Artificial_Intelligence)

\_to\_Combat\_Money\_Laundering\_and\_Related\_Crimes\_in\_the\_South\_African\_Banking\_Sector

Xu Tian, Z. T. (2024). *Machine learning in internet financial risk management: A systematic literature review*. Obtenido de Plos one:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0300195>

Vito Christian Samudra, D. P. (2024). *Application of Startup Success Prediction Models and Business Document Extraction Using Large Language Models to Enhance Due Diligence Efficiency*. Obtenido de Research Gate:

[https://www.researchgate.net/publication/386269384\\_Application\\_of\\_Startup\\_Success\\_Prediction\\_Models\\_and\\_Business\\_Document\\_Extraction\\_Using\\_Large\\_Language\\_Models\\_to\\_Enhance\\_Due\\_Diligence\\_Efficiency](https://www.researchgate.net/publication/386269384_Application_of_Startup_Success_Prediction_Models_and_Business_Document_Extraction_Using_Large_Language_Models_to_Enhance_Due_Diligence_Efficiency)

Sara Bagur, M. R.-L. (2021). *El Enfoque integrador de la metodología mixta en la investigación educativa*. Obtenido de Revista de investigación y evaluación educativa.

## **ANEXOS**

**Tabla 10:** Desgloce de puntuación por características de Impacto y Esfuerzo

Características	Optimización manual y estandarización de procesos capacitación	Integración de fuentes de datos externos	Digitalización de formularios y documentos con validación asistida.	Automatización de tareas repetidas con RPA	Reestructuración de la matriz de riesgo y metodología EBR	Modelo de ML con perfilamiento de riesgo dinámico
<b>IMPACTO</b>						
Falsos Positivos	2	3	4	4	2	5
Eficiencia	3	5	5	5	1	4
Cumplimiento Normativo	3	4	4	3	4	5
PROMEDIO	2,67	4,00	4,33	4,00	2,33	4,67
<b>ESFUERZO</b>						
Costo	2	5	4	3	3	1
Calidad datos	1	3	3	2	4	2
Tiempo	2	4	5	3	5	3
PROMEDIO	1,67	4,00	4,00	2,67	4,00	2,00

**Tabla 11:** *Matriz de hallazgos operativos*

<b>Subprocesos</b>	<b>Causa raíz</b>	<b>Punto de dolor</b>	<b>Impacto</b>	<b>Métrica</b>
Recopilación de información	Digitación manual de datos del cliente desde formularios escaneados, electrónicos al archivo de Excel	Alta probabilidad de errores tipográficos, datos incompletos y no visibles	Riesgo operativo: datos sucios que afecta la calidad para reportes	15-20 aprox min por cliente; tasa de error de transcripción de 10%
Validación en listas	Búsqueda individual en distintos portales (OFAC, antecedentes penales, SRI, ect.) para verificar información legal, tributaria y de ingresos.	Proceso lento donde el analista debe navegar en múltiples interfaces, capturando evidencia manual	Riesgo legal: eleva la probabilidad de falsos negativos al omitir un cliente sancionado, y vulnerabilidad ante auditorías por evidencia no estructurada	8-15 aprox min por cliente
Perfilamiento de riesgo	Asignación de nivel de riesgo basado en criterio subjetivo del analista	Inconsistencia en la calificación del riesgo del cliente	COSO III, principio 7: ausencia de objetivos claros que permitan una evaluación de riesgo consistente y comparable	Tasa de falsos positivos mayor a 85%
Matriz de riesgo	Hoja de Excel con fórmulas estáticas y pesos desactualizados para ponderar variables	Dificultad para actualizar factores de riesgo, fórmulas o celdas ocultas	ISO 31000: falta de dinamismo y adaptabilidad en la matriz de riesgo, lo que limita la actualización de factores de riesgo	Actualización de parámetros, 1 vez al año
Almacenamiento	Gestión de expedientes en carpetas compartidas sin orden ni estructura secuencial definida	Falta de trazabilidad del estado de gestión si está pendiente o procesada	LOPDP: riesgo de integridad y pérdida de control del ciclo de vida del dato, afectando la seguridad y confidencialidad	Trazabilidad del flujo de gestión: 0%

## Desarrollo de la Matriz de Riesgo

MATRIZ DE RIESGO													
Tipo ID	Numero ID	Apellido	Apellido2	Nombre	Género	Fecha Nacimiento	Fecha revisión	Edad	POND1	Nacionalidad	POND2	Provincia	POND3
									Nivel de riesgo por edad		Nivel de riesgo nacionalidad		Nivel de riesgo de criminalidad
Cédula	0967123701	Carpio	Piguave	Pepe Carlos	M	25/08/2001	11/11/2025	24	5	Ecuador	1	Guayas	5
Cédula	0943767939	Rugel	Rugel	Madelline Rugel	F	17/4/2000	11/11/2025	26	5	Ecuador	1	Guayas	5
Cédula	1103985832	Jaramillo	Llivigañay	Andrea del Cisne	F	18/07/1986	12/11/2025	39	2	Ecuador	1	Loja	2
Cédula	0956378921	Macias	Guaman	Jose Juan	M	26/07/1998	19/08/2025	27	5	ECUADOR	1	GUAYAS	5

Ingresos	Rango	POND5	Total Activos	Total Pasivos	Otros Ingresos	Patrimonio	POND6	Ramo	POND7	Suma Asegurada	POND8
		Nivel de riesgo de ingresos					Nivel de riesgo de patrimonio		Nivel de riesgo del ramo		Nivel de riesgo de suma asegurada
3000	3	1	370000	3000	0.00	367,000.00	1	Vida	1	550,000.00	5
600	1	5	150000	3000	0.00	147,000.00	1	Vehículo	5	48,600.00	2
879	2	2	70000	400	0.00	69,600.00	2	Multirisgo	3	789,000.00	5
4555	4	3	45000	15750	1200	29,250.00	2	Casco	4	30,000.00	2

Edad, nacioid,ing.pat ramos, SA broker, bancos provincia escala de semaforización?

Factor Persona	Factor Producto	Factor Canal	Jurisdicción	Riesgo Inherente
0.63	3	2	5	2.49
1.63	3.5	2	5	2.94
0.76	4	2	2	2.23
1.16	3	2	5	2.65
1.51	2.5	2	4	2.40

## Sesiones con el cliente

MACHINE LEARNING

```

model.fit(X_train, y_train_adjusted)

print("Modelo XGBoost entrenado exitosamente.")

Inicializando y entrenando el modelo XGBoost...
Número de clases únicas en y_train: 5
Ajustando etiquetas de y_train y y_test a 0-Indexed...
Modelo XGBoost entrenado exitosamente.

Reasoning: The XGBoost model has been successfully trained. The
next logical step is to use the trained model to make predictions on the
test set (X_test) so that its performance can be evaluated.

print("Realizando predicciones sobre el conjunto de prueba...")

y_pred = model.predict(X_test)

print("Predicciones realizadas exitosamente.")
# Mostrar las primeras 5 predicciones y los valores reales con
print("Primeras 5 predicciones y valores reales (ajustados):")
print(pd.DataFrame({"Actual": (adjusted), "y_test": y_test_adjusted.head(5)}))

Realizando predicciones sobre el conjunto de prueba...
Predicciones realizadas exitosamente.
                    
```

Factor Canal	Factor Jurisdicción	Riesgo Inherente	Riesgo	Debita diligencia	cliente
2	5	2.49	1	SIMPLIFICADA	
2	5	2.94	2	REGULAR	
2	2	2.23	1	SIMPLIFICADA	
2	5	2.65	2	REGULAR	
2	4	2.40	3	MEJORADA	
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			
2	#N/A	#N/A			

**FASE**  
**MINERÍA DE DATOS**  
**ESTRATÉGICA (CRISP-DM)**

- 01

**Compresión del negocio**  
Exploración de expedientes digitales y formularios KYC
- 02

**Análisis de datos**  
Base de datos histórica del cliente
- 03

**Preparación de los datos**  
Automatización del poblamiento de la matriz
- 04

**Modelado Inteligente**  
Algoritmo de aprendizaje supervisado XGBoost.
- 05

**Evaluación/Validación**  
Métricas de desempeño (precisión, tasa y tiempos)

