

**IMPLEMENTACIÓN DE UN MODELO ANALÍTICO PARA LA
IDENTIFICACIÓN DE COMPORTAMIENTOS INUSUALES
EN LOS PAGOS ANTICIPADOS DE UNA ENTIDAD DE
ARRENDAMIENTO FINANCIERO**

FELIPE HERNÁNDEZ MARTÍNEZ

**Trabajo de grado para optar al título de
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

ANDRÉS FELIPE CANO CADAVID

**Jefe de Sección, Sección Servicios Analíticos
Cumplimiento, Bancolombia S.A.**



**UNIVERSIDAD EIA
INGENIERÍA DE SISTEMAS Y COMPUTACIÓN
ENVIGADO
2019**

A mi familia, mis maestros, amigos y demás seres queridos, que no solo me han enseñado como puedo crecer como persona, sino también cómo puedo contribuir positivamente a mi contexto y a la sociedad.

« On ne voit bien qu'avec le cœur. L'essentiel est invisible pour les yeux. »

Antoine de Saint-Exupéry

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

AGRADECIMIENTOS

Agradezco a mi familia, especialmente a mis padres, a mi abuela y a mi hermano por apoyarme en mi crecimiento personal y profesional, por brindarme su acompañamiento incondicional, por permitirme soñar en grande y por creer en mí, su *niño*.

A mis amigos, especialmente a mis *amigas del primer día* y a *las catarinas*, por brindarme su apoyo, por acompañarme en todas las experiencias vividas en estos años, por motivarme a hacer y a aprender cosas nuevas y principalmente por confiar en mí.

A mis maestros, no solo universitarios, sino a todos aquellos que me han transmitido sus conocimientos y me han apasionado sobre la ciencia, la tecnología, las matemáticas, la historia, los idiomas y las artes. Especialmente quiero agradecerle a Hermilia Molina, por apasionarme por las ciencias computacionales; y a Luz Marina Arias Cadavid, por potencializar mis conocimientos en las ciencias sociales y por permitirme hacer parte de LEO.

A la Universidad EIA y a sus profesores, por su excelencia académica y personal, a mi director de carrera, Johan Vélez Macías, por su papel en mi proceso académico en la universidad, por ayudarme a encontrarle soluciones a los desafíos de la carrera y por sus comentarios motivacionales que “te hacen estudiar sistemas”.

A mi tutor, Andrés Felipe Cano Cadavid, por retarme académica y profesionalmente, y a todos mis compañeros de Bancolombia, especialmente a mis *padrinos*, María Alejandra Gaviria y Guillermo Romero, por abrirme las puertas al mundo profesional y aportar a mi transición de la academia a mercado laboral.

Por último, a Ana María García, pues sin ella no hubiese conocido ni me hubiese enamorado del ambiente y los valores de la EIA.

CONTENIDO

	pág.
INTRODUCCIÓN.....	13
1. PRELIMINARES.....	14
1.1 Planteamiento del problema	14
1.1.1 Formulación del problema	14
1.1.2 Justificación.....	15
1.2 Objetivos del proyecto	16
1.2.1 Objetivo General.....	16
1.2.2 Objetivos Específicos	16
1.3 Marco de referencia.....	16
1.3.1 Antecedentes	16
1.3.2 Marco teórico.....	17
2. METODOLOGÍA.....	21
2.1 Entendimiento del Negocio y de los Datos.....	21
2.2 Preparación de los Datos	22
2.3 Modelado.....	22
2.4 Evaluación y Despliegue	23
3. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS.....	24
3.1 ENTENDIMIENTO DEL NEGOCIO Y DE LOS DATOS	24
3.1.1 Arrendamiento Financiero (Leasing).....	24
3.1.2 Tipologías y Señales de Alerta de LAFT en Leasing	26
3.1.3 Identificación y Exploración de las Fuentes de Información.....	29

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

3.1.4	Resultado de la Etapa: Definición del Alcance del Monitoreo	30
3.2	PREPARACIÓN DE LOS DATOS Y MODELADO	31
3.2.1	Etapa 1: Visualización	31
3.2.2	Etapa 2: Árboles de Clasificación	35
3.2.3	Etapa 3: Modelo Estadístico	40
3.2.4	Etapa 4: Refinamiento del Modelo con Otras Variables	41
3.2.5	Modelo Final	43
3.3	EVALUACIÓN Y DESPLIEGUE	44
3.3.1	Evaluación Transversal: Modelo Prescriptivo	45
3.3.2	Prueba del Monitoreo: ¿Son o no alertas?.....	45
3.3.3	Piloto del Monitoreo: Evaluación con Casos Reales	45
3.3.4	Despliegue del Monitoreo	46
4.	CONCLUSIONES Y CONSIDERACIONES FINALES	48
	REFERENCIAS	50

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

LISTA DE TABLAS

	pág.
Tabla 1: Datos de Entrada para el Árbol de Clasificación	36
Tabla 2: Vista de Prepagos.....	38
Tabla 3: Vista de Capitales	42

LISTA DE FIGURAS

	pág.
Figura 1: Modelo CRISP-DM	21
Figura 2: Leasing en la Entidad Financiera	25
Figura 3: Esquema de Ejemplo: Tipología de Cesión	27
Figura 4: Esquema de Ejemplo: Tipología de Prepagos	28
Figura 5: Esquema de Ejemplo: Tipología de Testaferrato en Leasing	28
Figura 6: Fuentes de Datos Disponibles	30
Figura 7: Porcentaje de Clientes por Tipo de Persona	32
Figura 8: Porcentaje de Participación Monetaria por Tipo de Persona	32
Figura 9: Cantidades de Contratos por Tipo de Leasing	33
Figura 10: Diagrama de Caja (Porcentaje de Prepago por Tipo de Persona).....	34
Figura 11: Diagrama de Caja (Porcentaje de Prepago por Tipo de Leasing)	34
Figura 12: Diagrama de Caja (Porcentaje de Prepago por Tipo de Activo)	35
Figura 13: Importancia del Predictor	37
Figura 14: Ejemplo de Distribución del Porcentaje de Prepago para Personas Naturales Aplicando las Variables de Clasificación.....	39
Figura 15: Ejemplo de Distribución del Porcentaje de Prepago para Personas Jurídicas Aplicando las Variables de Clasificación.....	40
Figura 16: Fórmulas de Cálculo de Valores Atípicos.....	41
Figura 17: Ejemplos de Alertas	46

GLOSARIO

ALMACÉN DE DATOS: en inglés *data warehouse*, se refiere a una colección de datos con el cual las organizaciones pueden generar reportes, analizar datos y en general, aplicar la inteligencia de negocios en su contexto, desde un único sistema o base de datos.

ÁRBOL DE CLASIFICACIÓN: es una aplicación de los árboles de decisión para clasificación, compuesto por un conjunto de condiciones organizadas jerárquicamente, de tal forma que la toma de una decisión pueda determinarse por medio de condiciones que se cumplen desde la raíz del árbol hasta sus hojas (Moujahid, Inza, & Larrañaga, 2008).

ARRENDAMIENTO FINANCIERO: en economía, se refiere a un contrato de arrendamiento en el que un arrendador pasa el derecho de uso de un activo a un locatario a cambio del pago de unas cuotas periódicas (o cánones) y en el cual se estipula que al final del contrato el arrendatario puede acceder a la compra del activo arrendado.

ARRENDAMIENTO OPERATIVO: "...es un acuerdo donde se cede el derecho de uso de un activo del propietario a otra persona [...] a cambio de un pago periódico [...] por un plazo determinado ..." (Westreicher, 2019).

BOXPLOT: expresión en inglés que se refiere a un diagrama de caja.

CANON: es un pago periódico que se hace por el uso o disfrute de un bien.

COLOCACIÓN: primera etapa del lavado de activos, en la cual el lavador introduce sus fondos en el sistema financiero, para darles apariencia de legalidad.

CRISP-DM: es el Cross Industry Standard Procces for Data Mining, un modelo de trabajo para la minería de datos otros procesos analíticos.

CUARTIL: es "uno de los tres puntos que dividen un conjunto de datos numéricamente ordenados en cuatro partes iguales." (Edu2000 America Inc, 1995-2006)

DIAGRAMA DE CAJA: es una representación gráfica de datos numéricos por medio de sus cuartiles, en el cual se pueden identificar visualmente los valores atípicos.

DISTRIBUCIÓN DE DATOS: también conocida como distribución de frecuencias, es una agrupación de datos basándose en el número de observaciones de cada valor.

ELT: en computación, es el proceso de extracción, carga y transformación (*extract, load and transform*) de datos que realizan las organizaciones desde sus sistemas o bases de datos hacia un almacén de datos para su integración y uso en procesos de apoyo a la operación. Es una variación del proceso ETL en la que, por las capacidades computacionales del almacén de datos, la carga de los datos se hace primero y después de ello la transformación.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

ESTRATIFICACIÓN: segunda etapa del lavado de activos, en la que se separan los fondos de origen ilícito mediante transacciones financieras sofisticadas, que buscan disimular el rastro de la transacción de colocación.

ETL: en computación, es el proceso de extracción, transformación y carga de datos, del inglés *extract, transform and load*, que realizan las empresas desde sus bases de datos o sistemas hacia un almacén de datos para su integración y uso en procesos de apoyo a la operación.

FINANCIACIÓN DEL TERRORISMO: es la dotación de recursos a organizaciones terroristas, con las que estos costean sus actividades.

GAFI: el Grupo de Acción Financiera Internacional, y en inglés *Financial Action Task Force (FATF)*, es un ente intergubernamental creado en 1989 y encargado de fijar políticas y promover internacionalmente medidas legales, regulatorias y operativas para combatir el lavado de activos, la financiación del terrorismo y otras amenazas globales contra el sistema financiero (FAFT-GAFI, 2019).

GAFILAT: Grupo de Acción Financiera de Latinoamérica, "...es una organización intergubernamental regional que agrupa a 17 países de América del Sur, Centroamérica, América de Norte y el Caribe para prevenir y combatir el lavado de activos, el financiamiento del terrorismo y el financiamiento de la proliferación de armas de destrucción masiva (LA/FT/FPADM)..." (Grupo de Acción Financiera de Latinoamérica, 2019).

HIPERPARÁMETRO: son los parámetros ajustables y elegibles que sirven para entrenar un modelo computacional (Microsoft, 2018).

INTEGRACIÓN: tercera etapa del lavado de activos, en la que se busca dar apariencia de legalidad a los bienes ilícitos mediante el reingreso en la economía con transacciones que aparentan ser normales. En este punto es más difícil diferenciar los bienes lícitos de los ilícitos.

LAFT: acrónimo para Lavado de Activos y Financiación del Terrorismo.

LAVADO DE ACTIVOS: es el proceso mediante el cual los delincuentes buscan darle apariencia de legalidad a sus recursos, bienes o dinero de origen ilícito. Está compuesto de tres etapas: colocación, estratificación e integración (García Gibson, 2015). También puede ser conocido como lavado de dinero.

LAVADOR: persona que lava activos.

LEASING: "...es un mecanismo de financiación mediante el cual una entidad financiera [...], por instrucción de un cliente solicitante (denominado arrendatario o locatario), adquiere un activo de capital, el cual está bajo propiedad de la entidad, y se lo entrega al locatario en

arrendamiento financiero u operativo para su uso y goce por un periodo de tiempo a cambio de un pago periódico de una suma de dinero, denominado canon.” (ASOBANCARIA, 2016)

LEASING FINANCIERO: tipo de leasing en el cual, “al final de la operación [...] el locatario tiene la potestad de ejercer una opción de adquisición sobre el mismo bien a un precio pactado desde el inicio –generalmente a su favor y que no es superior al 30% del valor comercial del bien– o restituirlo (devolverlo) a la entidad financiera.”

LEASING HABITACIONAL: “...es una modalidad particular de leasing financiero destinado a la adquisición de vivienda, el cual es celebrado entre un establecimiento de crédito o compañía de financiamiento y un locatario.”

LOCATARIO: es el cliente o beneficiario en una operación de leasing.

MODELO COMPUTACIONAL: es un modelo matemático usado en ciencias computacionales con el cual se busca estudiar el comportamiento de un sistema complejo por medio de la simulación computacional (Melnik, 2015).

MONITOREO: desarrollo computacional que permite monitorear las operaciones transaccionales de los clientes en una entidad financiera.

SEÑAL DE ALERTA: es una situación que al ser analizada está por fuera de los comportamientos usuales y por ello facilitan el reconocimiento de operaciones de lavado de activos o financiación del terrorismo.

TIPOLOGÍA: “... estudios que analizan fenómenos, sectores, tendencias o modalidades por las cuales se realizan operaciones [de LAFT, las cuales] se caracterizan por ser descriptivas y sirven para proponer políticas, cambios normativos o estrategias para fortalecer la lucha contra estos delitos.” (Unidad de Información y Análisis Financiero, 2014)

UIAF: Unidad de Investigación y Análisis Financiero.

VALOR ATÍPICO: en estadística, es una observación inusual o numéricamente distante de las demás observaciones de datos.

VISTA: en el contexto de los datos, se refiere a una consulta a bases de datos que los muestra de manera tabular, organizados por columnas y filas.

RESUMEN

Las entidades financieras en Colombia deben contar, por regulaciones legales, con un Sistema de Administración del Riesgo de Lavado de Activos y Financiación del Terrorismo (SARLAFT); dentro de los elementos que contiene dicho sistema se encuentra el monitoreo de las operaciones transaccionales de los clientes de la entidad para identificar y reportar operaciones de lavado de activos y financiación del terrorismo. Actualmente, una entidad financiera colombiana que ofrece servicios de arrendamiento financiero (*leasing*) tiene la necesidad de monitorear las operaciones a las obligaciones financieras de sus clientes con el fin de identificar comportamientos inusuales y a partir de ellos posibles alertas de Lavado de Activos o Financiación del Terrorismo. En este trabajo se presenta una solución a dicha necesidad por medio de la implementación de un modelo computacional, enmarcado en la metodología de desarrollo de proyectos de minería de datos, CRISP-DM y utilizando técnicas estadísticas, analíticas y conceptos de inteligencia de negocios. Con este monitoreo se identifican operaciones inusuales para los clientes que realizaron prepagos parciales o totales sobre sus obligaciones de Leasing Financiero o Arrendamiento Operativo, que pudiesen considerarse como valores atípicos dentro del segmento interno del cliente en la entidad, y generan alertas, para su posterior evaluación por modelos prescriptivos y expertos del negocio para determinar si son una posible fuente de Lavado de Activos.

Palabras clave: lavado de activos, arrendamiento financiero, valores atípicos, modelo computacional, monitorear.

ABSTRACT

Financial entities in Colombia must have, by legal regulations, a Risk Management System for Money Laundering and Terrorist Financing (in Spanish: *SARLAFT*); the transactional monitoring of the entity's clients is contained within the elements of this management system to identify and report money laundering and terrorist financing operations. Currently, a Colombian financial institution which offers leasing services has the need to monitor the transactions of its customers in order to identify unusual behaviors and from them possible alerts of Money Laundering or Terrorism Financing. This paper presents a solution to this need through the implementation of a computational model, framed in the methodology for the development of data mining projects, CRISP-DM, and by using statistical techniques, analytics and business intelligence concepts. With this monitoring, unusual operations are identified for customers who made partial or total prepayments on their Financial Leasing or Operating Lease obligations, which could be considered as outliers within the client's internal segment in the entity and alerts are generated for further evaluation by prescriptive models and business experts to determine if they are a possible source of Money Laundering.

Keywords: money laundering, leasing, outliers, computational model, monitoring.

INTRODUCCIÓN

En el sector financiero colombiano existe una preocupación común respecto a las consecuencias que tienen para la economía del país delitos como la financiación del terrorismo, el lavado de activos y sus delitos fuente, como el narcotráfico, la trata de personas, la extorsión, etcétera; debido a ello, existen entidades de control y regulaciones que buscan prevenir y detectar el Lavado de Activos y la Financiación del Terrorismo – LAFT en la administración u operación de las entidades financieras colombianas. Una de estas regulaciones es la implementación de un Sistema de Administración del Riesgo de Lavado de Activos y Financiación del Terrorismo – SARLAFT, con el que, entre otras cosas, se debe monitorear el riesgo LAFT y reportar operaciones sospechosas de sus clientes o usuarios, por lo que dentro de una organización financiera es necesario el desarrollo de *monitoreos* con los que se identifiquen señales de alerta de clientes que, por medio de sus operaciones financieras, estén tratando de dar apariencia de legalidad al origen o destino de sus recursos, provenientes de actividades ilícitas.

En el presente trabajo se plantea la construcción de uno de estos *monitoreos*, enfocado a los productos de arrendamiento financiero o *leasing* de una entidad financiera del mercado colombiano, apoyándose en herramientas de software, plataformas de datos y conceptos de ciencia de los datos, de estadística, de analítica y de inteligencia de negocios, así como de la metodología CRISP-DM para el desarrollo de proyectos de minería de datos; respondiendo al cuestionamiento ¿cómo puede una organización monitorear el comportamiento transaccional de los clientes de su línea de negocio de arrendamiento financiero, basándose en modelos de ciencia de los datos y técnicas analíticas?

En las siguientes secciones se profundiza en la problemática identificada y la relevancia de su investigación, en los objetivos del presente estudio, en las etapas que componen la metodología de trabajo y las razones de su selección, en los resultados obtenidos en cada una de estas etapas y la manera en que fueron alcanzados, en los resultados finales del trabajo, en las conclusiones respecto a las hipótesis planteadas y los hallazgos encontrados durante el desarrollo y la construcción del monitoreo, así como algunas consideraciones sobre futuras investigaciones dirigidas a la detección del Lavado de Activos y Financiación del Terrorismo en entidades financieras de arrendamiento financiero.

1. PRELIMINARES

1.1 PLANTEAMIENTO DEL PROBLEMA

1.1.1 Formulación del problema

En Colombia, el Lavado de Activos y la Financiación del Terrorismo son delitos que, por sus características y su relación con otros delitos como el narcotráfico, el secuestro extorsivo, la trata de personas, la extorsión, el tráfico de armas, el contrabando (Grupo Bancolombia, 2018), entre otros, “generan consecuencias negativas para la economía del país y para las empresas del sector real” (Superintendencia de Sociedades, 2017).

Teniendo en cuenta dicho argumento, y enfocado en el sector financiero, el gobierno colombiano ha constituido entidades reguladoras y entidades de inteligencia, dentro de las que se encuentran la Superintendencia Financiera de Colombia y la Unidad de Información y Análisis Financiero (UIAF), que entre otras funciones, se encargan de velar por la prevención y detección del Lavado de Activos y la Financiación del Terrorismo, en adelante LAFT (Unidad de Información y Análisis Financiero, 2018) en las entidades financieras.

De esta manera, las entidades financieras deben cumplir con la implementación de un Sistema de Administración del Riesgo de Lavado de Activos y Financiación del Terrorismo – SARLAFT, regulado por la Parte I del Título IV, Capítulo IV de la Circular Básica Jurídica C.E. 055/2016 (Superintendencia Financiera de Colombia, 2019), y que es un sistema para “...identificar, medir, controlar y hacer monitoreo de los riesgos de lavado de activos y financiación del terrorismo” (Lozano Consultores en Prevención de Riesgo LAFT, 2017) y reportar aquellas operaciones de los clientes que, basadas en ciertas variables o características, se consideren sospechosas.

Es así como dentro de una organización financiera deben diseñarse, implementarse y evaluarse monitoreos con los cuales se puedan identificar señales de alerta relacionadas a los riesgos de LAFT en relación con los clientes y sus movimientos financieros; para poder construir dichos monitoreos las entidades se apoyan en plataformas de datos y herramientas de software, como Hadoop o Python, en la ciencia de los datos, la analítica y los modelos computacionales y que implementan como insumo fuentes de información disponibles en sistemas gestores de bases de datos como Oracle y SQL Server, o bien en hojas de cálculo, documentos de texto y otros.

Con esta premisa, para una entidad financiera que brinda servicios de arrendamiento financiero en el mercado colombiano, es importante monitorear el comportamiento transaccional de sus clientes, de tal forma que pueda tomar decisiones estratégicas para mitigar el riesgo asociado a estos comportamientos en términos de Lavado de Activos y Financiación del Terrorismo y cumplir con las regulaciones establecidas para su operación legal en el territorio nacional, específicamente en relación con los pagos anticipados de sus obligaciones de arrendamiento financiero.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

Es así como surge la pregunta: ¿cómo puede una organización monitorear el comportamiento transaccional de los clientes de su línea de negocio de arrendamiento financiero, basándose en modelos de ciencia de los datos y técnicas analíticas?

1.1.2 Justificación

El Lavado de Activos es el medio por el que las organizaciones criminales tratan de darle una apariencia de legalidad a los recursos que adquieren como resultado de sus actividades ilícitas. Por otra parte, la Financiación del Terrorismo “es el apoyo financiero, en cualquier forma, al terrorismo o a aquellos que lo fomentan, planifican o están implicados en el mismo” (Unidad de Información y Análisis Financiero, 2014). En ambos casos, el proceso por el cual se comete la actividad ilícita consta de tres etapas (Perotti, 2009): la colocación o situación; la distribución, transformación, estratificación, diversificación u ocultamiento; y la integración de dichos recursos en la economía.

En la legislación colombiana el Lavado de Activos y el Financiamiento del Terrorismo están considerados como actividades delictivas en los Artículos 323 y 345 del Código Penal respectivamente (Unidad de Información y Análisis Financiero, 2014); estos delitos usualmente se relacionan con otros como lo son el “...tráfico de migrantes, trata de personas, extorsión, [...] secuestro extorsivo, rebelión, tráfico de armas, [...] tráfico de drogas tóxicas, estupefacientes o sustancias psicotrópicas, delitos contra el sistema financiero...” (Unidad de Información y Análisis Financiero, 2014), o bien estos dineros “...suelen invertirse en la compra de armas y equipos, pagos de nómina o sostenimiento de células terroristas, costeo de logística para acciones terroristas, inversión en adiestramiento y tecnología, pago de sobornos y mantenimiento de complicidades, entre otras” (Unidad de Información y Análisis Financiero, 2014).

De esta manera, el Lavado de Activos y la Financiación del Terrorismo afectan el sistema financiero y la economía, ya que pueden provocar distorsiones en la inversión, las tasas de crecimiento, las tasas de cambio, las tasas de interés, entre otros, y en general causar deterioro económico en las naciones; más aún, estos delitos afectan el Producto Interno Bruto (PIB) de las naciones (Grupo Bancolombia, 2018), como lo indica la Oficina de las Naciones Unidas contra la Droga y el Delito (2011), al afirmar que los activos lavados globalmente representaron en el 2011 entre el 2% y el 5% del PIB mundial.

Adicionalmente, de acuerdo con la publicación *Lo que debe saber sobre el lavado de activos y la financiación del terrorismo*,

En el esfuerzo del Estado para prevenir y contener el delito, se hacen necesarias las labores de inteligencia económica, la producción de conocimiento, la colaboración de los ciudadanos con las autoridades y la colaboración público-privada en general. (Unidad de Información y Análisis Financiero, 2014)

En conclusión, teniendo en cuenta el contexto colombiano, las consecuencias que provocan el Lavado de Activos y la Financiación del Terrorismo, así como los volúmenes de información existente en las entidades financieras, la velocidad de las herramientas analíticas actuales y la existencia de entidades de inteligencia y los lineamientos

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

establecidos por estas, es de suma importancia la aplicación de la ciencia de datos y la analítica para agregar valor a la administración del riesgo de LAFT, de tal forma que los datos se vuelven un insumo clave y efectivo para realizar monitoreos transaccionales de los clientes.

1.2 OBJETIVOS DEL PROYECTO

1.2.1 Objetivo General

Monitorear las operaciones transaccionales de los clientes de una entidad financiera que presta servicios de arrendamiento financiero, a través de técnicas y modelos de minería de datos y análisis de información, referenciándose en la metodología CRISP-DM, para identificar comportamientos inusuales.

1.2.2 Objetivos Específicos

Se definen cuatro objetivos específicos que le apuntan al desarrollo del objetivo general:

1. Recolectar el conocimiento e información clave para el entendimiento del negocio y los datos que podrían describir el comportamiento transaccional de sus clientes.
2. Preparar los datos recolectados por medio de una plataforma analítica para la consolidación de una vista única que sirva de base para la selección e implementación de un modelo analítico.
3. Modelar los datos para identificar los clientes que presentan comportamientos inusuales.
4. Evaluar el modelo implementado para su posterior despliegue dentro de la entidad.

1.3 MARCO DE REFERENCIA

1.3.1 Antecedentes

El análisis del comportamiento de los clientes por medio de la implementación de modelos computacionales ha sido un tema de investigación en diferentes sectores del mercado como en el caso de Roa Ferrada & Fuente Mella (s.f.), cuya investigación aborda la implementación de modelos de minería de datos para el desarrollo de estudios sobre la industria chilena, de tal forma que dichos puedan ser "...validados en base a información transaccional [...] para luego presentar conclusiones que permitan a la empresa tomar medidas futuras en base al conocimiento adquirido." (Roa Ferrada & Fuente Mella, s.f.).

Por otra parte, Torres Chero & Farroñay Julca (2017) aplicaron "...técnicas de minería de datos para detectar patrones de comportamiento de clientes morosos [en una empresa de créditos] y de esta forma reducir los índices de morosidad [...]." Adicionalmente,

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

desarrollaron “...una solución de inteligencia de negocios para detectar patrones de comportamiento en clientes morosos en la empresa de créditos...” (Torres Chero & Farroñay Julca, 2017).

Asimismo, en la investigación de Morelo Tapias & Puello Marrugo (2014) se aplica la metodología CRISP-DM para implementar un modelo de segmentación de los clientes de una empresa, generándose un nuevo algoritmo de minería de datos e información clave para la empresa en términos del conocimiento del consumidor.

Por último, Correa Chaparro (2015) analiza y rediseña “...el proceso de monitoreo transaccional [de una compañía de financiamiento]” con el objetivo de optimizar dicho proceso, enmarcado en el Sistema de Administración del Riesgo de Lavado de Activos y Financiación del Terrorismo de dicha compañía.

1.3.2 Marco teórico

- **Big Data:** De acuerdo con compañías como Dell Inc. (2018) y SAS (2018), basadas principalmente en la definición de Gartner, Big Data se refiere a un alto volumen de datos, tanto estructurados como no estructurados, que usualmente cumplen con lo que se denomina *las tres V del Big Data*: volumen, velocidad y variedad. El primer término se refiere a la cantidad de datos que se almacenan en las organizaciones, sobre los cuales se pretende generar valor; el segundo término se refiere al ritmo con el que se transmiten los datos (usualmente en tiempo real); y el último término se refiere a los diferentes tipos de datos que existen actualmente: no solo bases de datos, sino también documentos de texto, audio y video, entre otros.
- **Minería de Datos:** Se refiere a la implementación “...de algoritmos que extraen ‘conocimiento’ de grandes bases de datos que acumulan la historia de las actividades de las organizaciones...” (Martínez Luna, 2011) con la finalidad de tomar decisiones sobre temas o situaciones organizacionales.
- **Hadoop:** *Apache Hadoop es un framework de código abierto que permite el procesamiento paralelo de grandes conjuntos de datos y la extracción colectiva de diferentes fuentes de datos. Hadoop consiste en el Sistema de archivos distribuidos de Hadoop (HDFS), YARN (Yet Another Resource Negotiator) y otros componentes, como MapReduce. YARN actúa como el sistema operativo que gestiona otras aplicaciones como MapReduce, que es responsable de procesar grandes conjuntos de datos de forma paralela.*

El framework de código abierto de Apache Hadoop se extiende para incluir componentes de software adicionales, como Spark, Zookeeper, Pig y Hive, junto con cientos de otros. Estos componentes adicionales abordan la ingestión, seguridad, creación de scripts, procesamiento, visualización y monitoreo de datos. No todos los componentes son necesarios y el uso de múltiples componentes depende completamente de las necesidades individuales del flujo de trabajo. (Dell Inc., 2018) [Traducción por el autor]

- **Inteligencia de Negocios:** De acuerdo con autores como Dedić & Stanier (2017), la Inteligencia de Negocios es una compilación de habilidades, estudios, información, procesos, tecnologías y arquitectura de datos que se enfocan a la gestión y generación de conocimiento de un negocio, basado en la analítica de los datos existentes del mismo; en otras palabras, es el uso de los datos disponibles como facilitadores en la toma de decisiones estratégicas.
- **Extract, Transform and Load (ETL):** De acuerdo con SAS Institute Inc. (¿Qué es ETL? | SAS, 2018), son procesos de Inteligencia de Negocios que buscan la integración de datos por medio de tres pasos: extraer, transformar, cargar, usados para combinar los datos de diferentes orígenes, como bases de datos transaccionales, sistemas organizacionales o bases de datos libres. De esta manera, los datos se *extraen* de una fuente o sistema origen, se *transforman* en términos de formato y calidad y luego se *cargan* en un almacén de datos u otro sistema similar, con el fin de que los datos relevantes sean utilizados en los procesos de negocio correspondientes.
- **Extract, Load and Transform (ELT):** de acuerdo con Rouse y Preslar (2005-2019) y con Skinner (2018), es el proceso por medio del que se busca llevar los datos a un almacén de datos, similar a un ETL, pero en el cual se realizan las operaciones de transformación de los datos posteriormente a su carga en el almacén de datos, con la finalidad de optimizar recursos y tiempo.
- **Metodologías ágiles de desarrollo:** En las últimas décadas se ha generado un cambio drástico en la manera en que se desarrolla software debido a la complejidad del proceso como tal, a la rigidez de las metodologías de desarrollo tradicionales, como *Waterfall*, y a problemáticas con las mismas metodologías como el llamado “triángulo de hierro” (Leffingwell, 2011), en el cual solo se podía estimar el costo y el cronograma de desarrollo luego de estimar los requisitos. Es así como surge un énfasis de desarrollo basado en el establecimiento de los principios y valores con los cuales se puedan desarrollar proyectos relacionados a las TIC acertadamente en relación con las necesidades del consumidor y dar una mejor respuesta al cambio que se genera en todo el proceso de desarrollo (Alaimo, 2013). A partir de este enfoque surgen las metodologías de desarrollo ágiles, como Scrum, que pretenden cumplir con cuatro valores y doce principios estipulados en lo que se denomina el “Manifiesto Ágil” (Agile Alliance, 2001) y que buscan de manera general que exista un “...involucramiento del usuario y el empleo de periodos de tiempo más cortos [...] para incrementar las tasas de proyectos exitosos” (Alaimo, 2013).
- **Cross Industry Standard Process for Data Mining:** CRISP-DM “...es un método probado para orientar [...] trabajos de minería de datos” (IBM Corporation, 2012), que puede ser utilizado como una metodología o como un modelo de proceso, el cual contiene un ciclo vital compuesto por seis etapas con una secuencia no necesariamente estricta, cada una con unos entregables específicos y que giran en torno a los datos. Las seis etapas son: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y distribución.

- **Gestión de Datos:** De acuerdo con la Data Management Association *DAMA*, citada por Mosley & Brackett (2010) este proceso se refiere a "...el desarrollo, ejecución y supervisión de planes, políticas, programas y prácticas que controlan, protegen, entregan y mejoran el valor de los datos y los activos de información".
- **CRUD:** Este acrónimo se refiere, de acuerdo con Heller (2007), a las cuatro funciones básicas del almacenaje persistente, usadas comúnmente en las aplicaciones de gestión de datos relacionales, como lo afirma Martin (1983). Estas funciones son las de Crear, Leer, Modificar y Eliminar, por sus nombres en inglés *Create, Read, Update y Delete*.
- **Open Database Connectivity (ODBC):** De acuerdo con Microsoft (2015), la Conectividad Abierta de Bases de Datos (ODBC por sus siglas en inglés) "...es una interfaz [...] que permite que las aplicaciones tengan acceso a los datos de cualquier base de datos en la que haya un controlador ODBC [...permitiendo que...] la aplicación sea independiente del sistema de administración de bases de datos (DBMS) de origen."
- **Python:** De acuerdo con Python Software Foundation (2001-2018), "...es un lenguaje de programación que [...] permite trabajar más rápidamente e integrar [...] sistemas más efectivamente." Por medio de integraciones con este lenguaje de programación se pueden generar procesos de calendarización para consolidar información, dentro de los procedimientos de ELT mencionados anteriormente.
- **Sistema Gestor de Bases de Datos:** Son sistemas computacionales que permiten la gestión de bases de datos, es decir, "...la definición de bases de datos; así como la elección de las estructuras de datos necesarios para el almacenamiento y búsqueda de los datos, ya sea de forma interactiva o a través de un lenguaje de programación" (EcuRed, 2017). Uno de los sistemas gestores más conocidos de software libre es MySQL, el cual es relacional, multihilo y multiusuario; de acuerdo con su sitio web es "...la base de datos de código abierto más popular del mundo [que] con su probado rendimiento, fiabilidad y facilidad de uso [...] se ha convertido en la elección de base de datos líder para aplicaciones basadas en web..." (Oracle Corporation, 2017).
- **Señal de Alerta:** La UIAF (2014) define una señal de alerta como una situación que, cuando es analizada, se sale de los comportamientos típicos del mercado o de los individuos; estas situaciones facilitan el reconocimiento de operaciones de lavado de activos o financiación del terrorismo, luego de un análisis más exhaustivo y profundo. Es importante resaltar que no todas las señales de alerta identificadas son suficientes para identificar una operación de LAFT.
- **Tipología:** "Son estudios que analizan fenómenos, sectores, tendencias o modalidades por las cuales se realizan operaciones [de LAFT, las cuales] se caracterizan por ser descriptivas y sirven para proponer políticas, cambios normativos o estrategias para fortalecer la lucha contra estos delitos. El objetivo de una tipología es brindar herramientas que le permitan [a las partes interesadas]

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

contar con información que les facilite mejorar los mecanismos de control, así como conocer la forma en la que operan los delincuentes.” (Unidad de Información y Análisis Financiero, 2014)

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

2. METODOLOGÍA

Teniendo en cuenta el alcance del proyecto, es decir, el objetivo general y los objetivos específicos, y teniendo como referente la metodología para el desarrollo de proyectos de minería de datos CRISP-DM, se puede identificar el procedimiento a llevar a cabo para el desarrollo del monitoreo; dicho procedimiento aplica las etapas de dicha metodología, sin dejar atrás el enfoque de los objetivos específicos; de esta manera se considera que el desarrollo del monitoreo es un proceso incremental y en el cual pueden presentarse modificaciones que afectan las etapas de la metodología, siempre teniendo como punto de referencia centralizada a los datos con los que se cuenta para el desarrollo del modelo, como se observa en la Figura 1: Modelo CRISP-DM, de (Jensen, 2012).

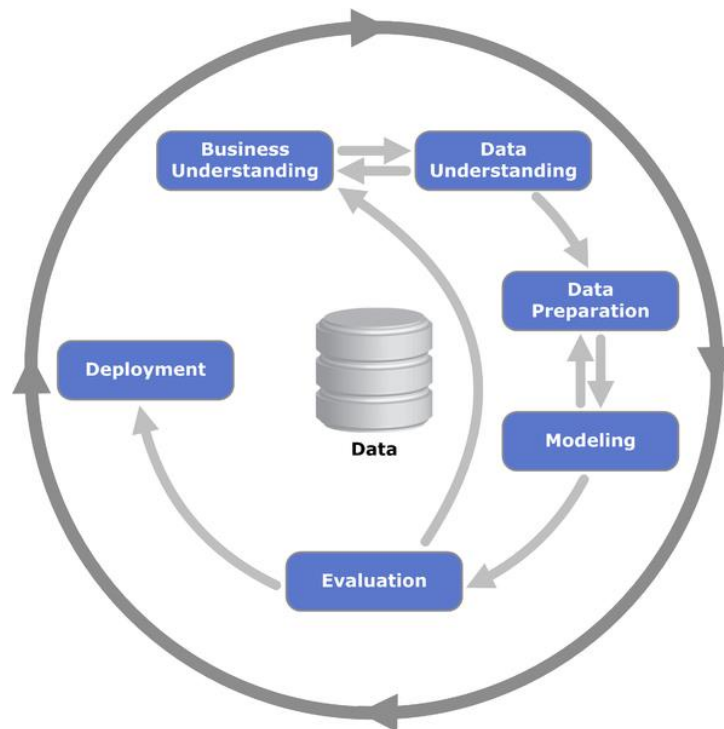


Figura 1: Modelo CRISP-DM

Es así como la metodología de trabajo se segmenta en cuatro etapas principales, las cuales están enfocadas a cada uno de los objetivos específicos del proyecto:

2.1 ENTENDIMIENTO DEL NEGOCIO Y DE LOS DATOS

En esta etapa se busca comprender el funcionamiento del arrendamiento financiero, el esquema de las fuentes de datos existentes en la entidad financiera e identificar las necesidades en términos comerciales y analíticos para la implementación del monitoreo, respondiendo a las preguntas: ¿cómo identificar los comportamientos inusuales en los

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

servicios de arrendamiento financiero? y ¿cuáles son las tipologías y señales de alerta LAFT en esta línea de negocio? Para responder a este cuestionamiento es necesaria la investigación y documentación sobre el arrendamiento financiero, su contexto, actores involucrados y ciclo de operación, tipologías y señales de alerta, así como la recopilación y exploración de los datos existentes en la entidad financiera y documentación de las características generales de los mismos en términos de metadatos, claves de asociación, descripción y calidad de los datos. Al terminar esta etapa, se deben determinar y listar las fuentes de información disponibles para el monitoreo y delimitar el alcance de este.

2.2 PREPARACIÓN DE LOS DATOS

Dentro de la preparación de los datos se debe generar una vista única que sea base para el monitoreo a implementar, teniendo en cuenta que los datos que incluya deben haber pasado por las siguientes etapas:

1. Seleccionar las variables o datos a monitorear en términos de los atributos o características a analizar.
2. Incluir o eliminar datos en términos de correlaciones, calidad o hipótesis que se formulen.
3. Limpiar los datos, para gestionar los datos ausentes, los errores o las incoherencias en estos, basándose en las conclusiones sobre calidad de datos de la etapa anterior.
4. Construir nuevos datos, como columnas calculadas o recodificadas que describan mejor las variables a considerar en el modelo.

2.3 MODELADO

En esta etapa del monitoreo se modelan los datos de manera iterativa, teniendo como base la vista única generada para las variables seleccionadas y a través de herramientas analíticas de software o de la aplicación de análisis estadísticos y reglas duras, con la finalidad de identificar las operaciones inusuales de los clientes y apuntando a una de las tipologías de LAFT en arrendamiento financiero, identificada en la primera etapa de la metodología y teniendo en cuenta la aplicación de los parámetros de entrenamiento o *hiperparámetros* adecuados al modelo y diseñando variables de evaluación de los resultados obtenidos de la técnica o técnicas empleadas, en términos de efectividad y confianza.

Para la selección del modelo se deben tener en cuenta características como (IBM Corporation, 2012) los tipos de los datos con que se dispone para la minería, el objetivo del monitoreo, los requisitos presentados por el negocio, documentando los supuestos que deban considerarse o aplicarse para la generación del modelo.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

El resultado de esta etapa es un el modelo funcional diseñado e implementado en una herramienta de software seleccionada y que contenga los análisis, *hiperparámetros* y reglas duras aplicadas a los datos; con el cual se generen alertas sobre personas cuyas operaciones sean consideradas inusuales o que, luego de un análisis a profundidad por parte de expertos, se puedan considerar como posibles operaciones de LAFT. Se espera que este tipo de análisis realizado por expertos sea el criterio para evaluar el modelo y sus alertas.

2.4 EVALUACIÓN Y DESPLIEGUE

En esta etapa se busca la valoración de criterios de rendimiento desde el punto de vista analítico y de expertos del negocio, de tal forma que se puedan evaluar los resultados obtenidos en el modelado, revisar y documentar las lecciones aprendidas en el proceso y determinar si se continua con el despliegue del monitoreo con los parámetros establecidos o si deben darse recomendaciones para refinar el modelo o las variables que influyen sobre este. El resultado de esta etapa es la decisión respecto al despliegue del monitoreo, la definición de una periodicidad en la ejecución, si aplica, y la documentación necesaria para la ejecución o los aspectos a considerar si se desea refinar o modificar el alcance.

3. PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS

Teniendo en cuenta la metodología iterativa de trabajo y las dependencias o relaciones clave encontradas en cada una de las etapas, y apuntando a los objetivos específicos planteados, la presentación y discusión de resultados se formula así:

1. Entendimiento del Negocio y de los Datos, explicando los resultados obtenidos en términos de conocimiento y definición de alcance del monitoreo a partir del entendimiento del arrendamiento financiero, de las tipologías y señales de alerta de LAFT en este producto y de los datos disponibles en la entidad;
2. Preparación de los Datos y Modelado, detallando los resultados iterativos de ambas etapas y las decisiones tomadas respecto al monitoreo, a los datos y al modelo; y
3. Evaluación y Despliegue, describiendo tanto los logros obtenidos a partir de la evaluación de las alertas generadas por el monitoreo como las etapas y decisiones respecto al monitoreo, posteriores a dicha evaluación. Aquí se presentan también los resultados finales, los cuales dan razón del cumplimiento del objetivo general de la investigación.

3.1 ENTENDIMIENTO DEL NEGOCIO Y DE LOS DATOS

Para comprender el funcionamiento del arrendamiento financiero se realizó una investigación en fuentes de información de entidades internacionales reconocidas, como la Oficina de las Naciones Unidas contra la Droga y el Delito, el Grupo de Acción Financiera de Latinoamérica (GAFILAT), y la Unidad de Investigación y Análisis Financiero de Colombia (UIAF); así como sesiones de capacitación con expertos la operación de este negocio dentro de la entidad financiera en estudio. Precisamente, se identifican conceptos clave para delimitar el alcance del modelo a implementar con la finalidad de determinar comportamientos sospechosos en las operaciones de esta línea de negocio.

3.1.1 Arrendamiento Financiero (Leasing)

El primer paso es el entendimiento del arrendamiento financiero como una línea de negocio de la entidad financiera, de esta forma se puede definir un Leasing o Arrendamiento Financiero, de acuerdo con el Programa de Asistencia Legal para América Latina y el Caribe de la Oficina de las Naciones Unidas contra la Droga y el Delito (2011), como:

...un producto financiero [...] donde mediante un contrato, una parte [o entidad de Leasing] entrega a la otra [llamada Locatario] un activo para su uso y goce, a cambio de un canon periódico (arrendamiento), durante un plazo convenido, a cuyo vencimiento, el bien se restituye a su propietario o se transfiere al usuario, si éste último decide ejercer una opción de adquisición [pactada].

Existen principalmente dos tipos básicos de leasing: leasing financiero y el arrendamiento u arrendamiento operativo; el primero se caracteriza por la opción de compra pactada desde la constitución del contrato, del locatario sobre el activo, y el segundo por ser un

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

arrendamiento sin opción de compra sobre el activo en cuestión, donde este siempre será de propiedad del arrendador o entidad de leasing. Adicionalmente, existen diferentes modalidades de leasing, que pueden ser utilizadas para la adquisición de bienes de importación o exportación, bienes inmobiliarios, infraestructura, medios de transporte o vehículos, e incluso elementos tecnológicos.

Dentro de los contratos de Leasing, existen diferentes modalidades de cánones, siendo las más generales (Oficina de las Naciones Unidas contra la Droga y el Delito, 2011) el canon fijo, el canon variable, el canon creciente y el canon decreciente. El canon fijo hace referencia a pagos iguales durante el tiempo estipulado en el contrato; en el canon variable los pagos fluctúan dependiendo de las condiciones del mercado o los acuerdos que se estipulen entre las partes; en el canon creciente los pagos aumentan conforme se va ejecutando el contrato, y en el canon decreciente los pagos disminuyen conforme se va ejecutando el contrato.

Por otra parte, y de manera general, al interior de la entidad financiera, el funcionamiento básico de leasing puede resumirse como en la Figura 2, donde se observan los eventos más importantes en el ciclo de vida de un contrato de leasing: la elección, por parte del cliente, de un activo y un proveedor de este; la venta del bien a la compañía de leasing; la entrega del bien desde la entidad financiera al cliente por medio de un contrato de Leasing con una duración y condiciones establecidas entre las partes; el pago del cliente a la entidad financiera de un canon periódico en el cual se paga un porcentaje de capital, de intereses y de otras variables relacionadas al activo, y definidas en el contrato; y la adquisición del activo por el cliente, al final del plazo definido entre la entidad financiera y este.



Figura 2: Leasing en la Entidad Financiera

Adicionalmente, la entidad financiera ofrece tres tipos principales de leasing para los diferentes segmentos internos de personas naturales y de personas jurídicas:

1. Leasing Financiero, con el cual se pueden adquirir vehículos, maquinaria y equipo, tecnología, inmuebles y otros bienes;

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

2. Leasing Habitacional, dedicado exclusivamente a la adquisición de bienes inmuebles como oficinas, apartamentos o casas; y
3. Arrendamiento Operativo, en el que no existe la opción de compra por parte del locatario y que, basado en la información recopilada, usualmente es usado con el objetivo (del locatario) de obtener un beneficio tributario.

Este hallazgo es importante ya que permite diferenciar la naturaleza de las operaciones que se realicen por el tipo de negociación que cada tipo conlleva: debido a que en algunos casos se tiene la opción de compra y en el otro no, esto puede ser aprovechado en la construcción del monitoreo para identificar comportamientos inusuales.

Luego de comprender el funcionamiento de una entidad de leasing y el flujo de las operaciones de un contrato de leasing, el siguiente cuestionamiento es, de manera general: ¿cómo emplean las organizaciones criminales este producto financiero para lavar dinero?

3.1.2 Tipologías y Señales de Alerta de LAFT en Leasing

Con el objetivo de identificar comportamientos inusuales en las operaciones de leasing dentro de la entidad financiera, definir un alcance al monitoreo y alertar una posible actividad de lavado de activos en este producto financiero, se deben conocer y comprender las tipologías y señales de alerta LAFT en este tipo de negocios y, adicionalmente, los datos disponibles en la entidad financiera.

De esta forma, basándose en los documentos del Grupo de Acción Financiera de Latinoamérica (2016) y de la Unidad de Investigación y Análisis Financiero – UIAF (Compilación de Tipologías de Lavado de Activos y Financiación del Terrorismo 2004 - 2013, s.f.), las operaciones que se realizan en un negocio de arrendamiento financiero pueden ser un medio usado por las organizaciones criminales para cometer hechos ilícitos que le apuntan a alguna de las etapas del lavado de activos (colocación, ocultamiento o integración) o bien a la financiación del terrorismo; así, se han identificado tres tipologías comunes para las operaciones de leasing:

1. La cesión del derecho de adquisición del activo financiado por medio del contrato de arrendamiento financiero, con la finalidad de lavar dinero de origen ilícito. En este caso el lavador aprovecha la posibilidad del locatario de ceder el derecho a la obtención del bien o activo financiado por medio del contrato de leasing, como se observa en la Figura 3 (GAFILAT, 2016), en donde un locatario cede el derecho de adquisición del inmueble, el cual es adquirido por la organización criminal por medio de un testaferro y, en este caso particular, con ayuda de un funcionario de la entidad de leasing.
2. La realización de prepagos parciales o totales a la obligación de arrendamiento financiero, en los cuales el dinero proviene de operaciones ilícitas. En este caso, y como se puede observar en el ejemplo de la Figura 4 (Unidad de Información y Análisis Financiero, s.f.), es el locatario quién realiza pagos anticipados a sus obligaciones de arrendamiento financiero, ya sean de manera parcial o total, con dinero de origen ilícito.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

3. El uso de apoderados o testaferros por parte de los proveedores de la operación de leasing, cuyos bienes fueron adquiridos con dineros de origen ilícito, para darle apariencia de legalidad a dichos bienes. En la Figura 5 (Unidad de Información y Análisis Financiero, s.f.) se puede observar el flujo de la tipología, identificándose la manera en que las organizaciones criminales podrían lavar activos en una operación de leasing por medio del uso de testaferros.

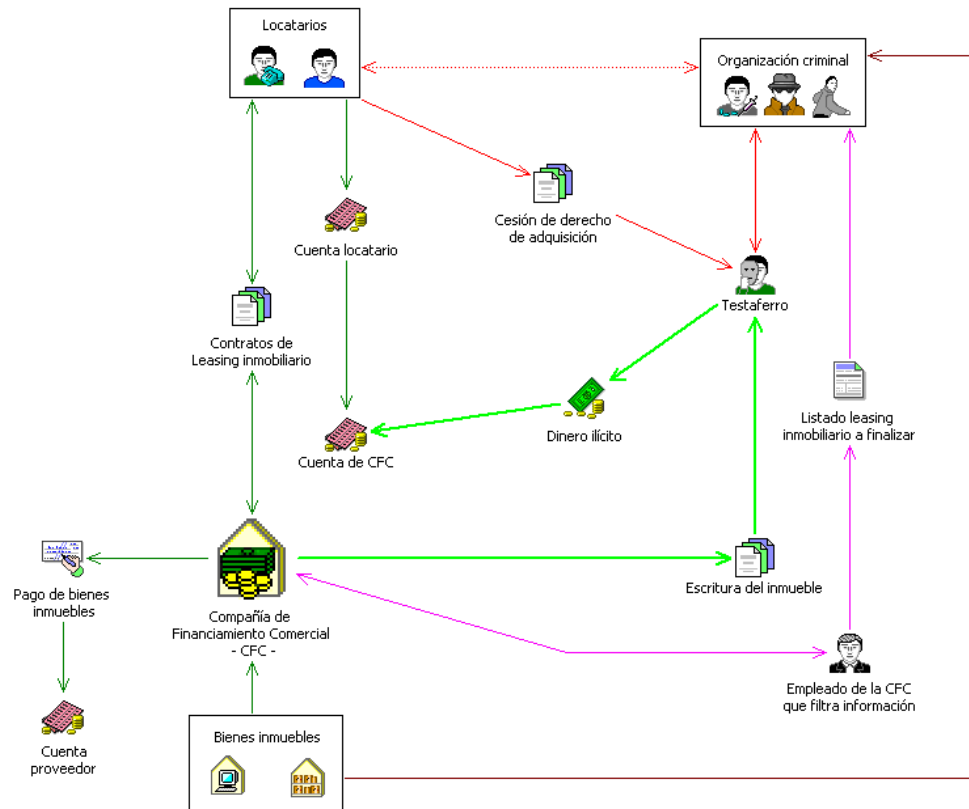


Figura 3: Esquema de Ejemplo: Tipología de Cesión

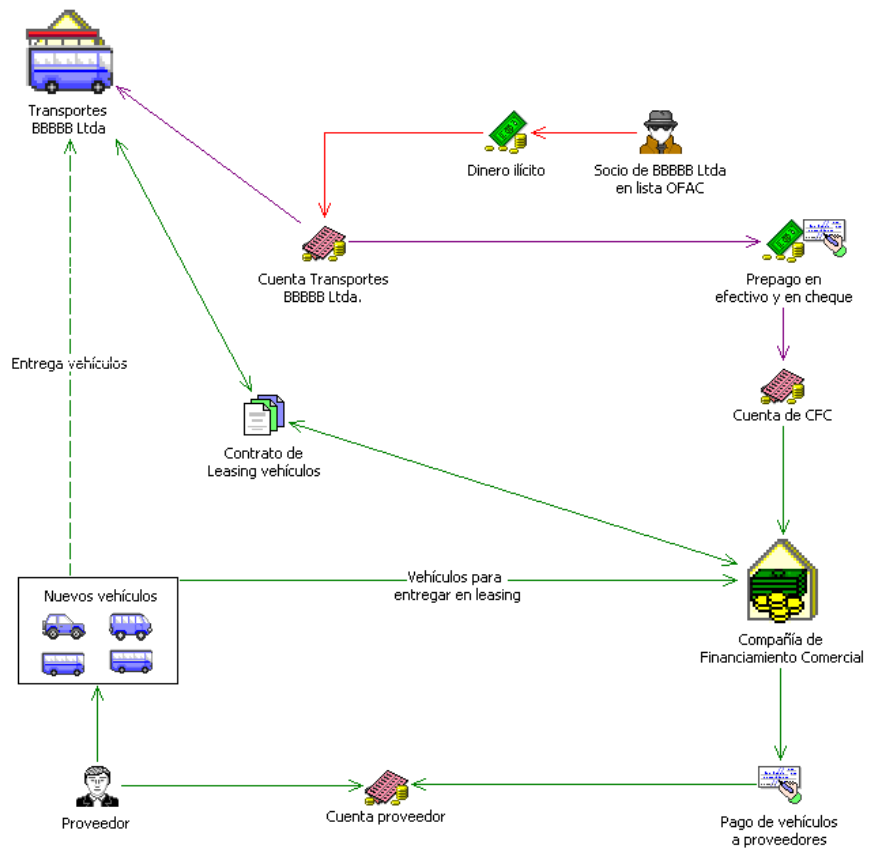


Figura 4: Esquema de Ejemplo: Tipología de Prepagos

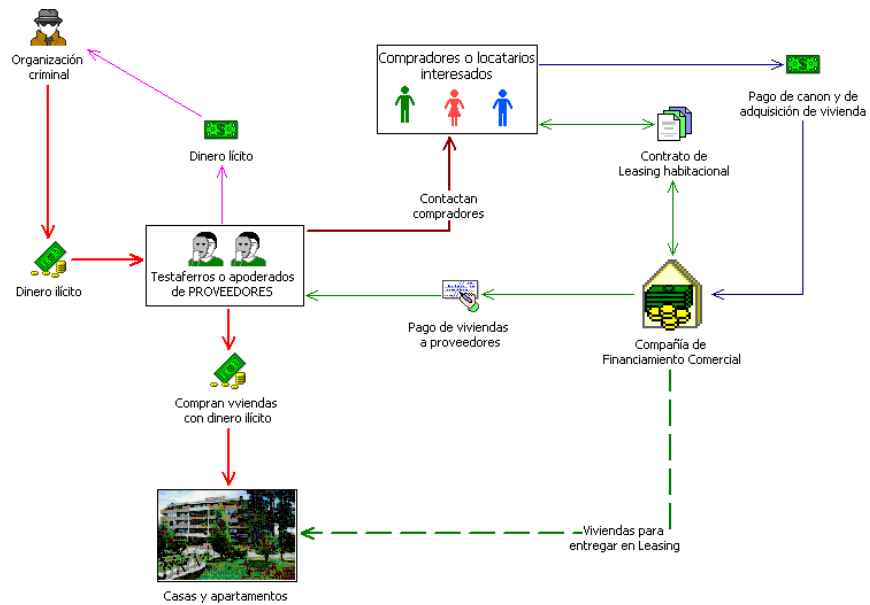


Figura 5: Esquema de Ejemplo: Tipología de Testaferrato en Leasing

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

Cada una de las tipologías tiene relacionada una cantidad de señales de alerta que podrían permitirle al evaluador, a la entidad financiera o a las entidades gubernamentales y de control, identificar posibles operaciones de lavado de activos en las operaciones de arrendamiento operativo, entre las que se destacan:

- Reiteradas operaciones de cesión de contratos de arrendamiento financiero.
- Clientes que le dan poca importancia a los costos asociados a las operaciones de exportación o importación relacionadas con un contrato de leasing.
- Prepagos totales o parciales a las operaciones de leasing sin justificación aparente.
- Incongruencias entre la información financiera del cliente y el dinero que entrega a la entidad financiera en los cánones, bien sea comunes o extraordinarios.
- Proveedores de bienes o activos que realizan la compraventa de este por medio de terceros o apoderados.
- Bienes adquiridos recientemente por montos inusuales.

A partir de esta información se puede delimitar el alcance del monitoreo a una o varias tipologías, enfocadas a una o varias señales de alerta; sin embargo, hay que conocer ¿qué información tiene la entidad financiera de sus operaciones y del cliente?, para poder determinar un alcance adecuado.

3.1.3 Identificación y Exploración de las Fuentes de Información

Luego de una investigación en los repositorios de metadatos internos y asesorías con el área de tecnología de la entidad financiera, se identifica que la entidad financiera tiene un sistema de datos para la línea de negocio Leasing, el cual está dividido en dos funcionalidades principales: información de los clientes e información de los contratos. En este sistema están disponibles entonces, en distintas librerías, la información correspondiente a los clientes: información socio demográfica y de contacto, relaciones entre clientes, maestro de proveedores, información de oficinas, parámetros de ubicación geográfica, entre otros; y la información correspondiente a los contratos: maestro de contratos, histórico de cánones, maestro de cobranzas, cuentas de cobro, transacciones, parámetros del sistema, entre otros.

Adicionalmente, se identificó la existencia de una fuente global de datos sociodemográfica de los clientes de todas las líneas de negocio, construida en el almacén de datos de la entidad con base en los principales sistemas de información de esta, la cual tiene como objetivo la consolidación de estos datos para su uso por parte de las áreas analíticas internas, de tal forma que dentro de este proyecto puede ser usada como un complemento a los datos con información de clientes identificadas en el sistema propio de la línea de negocio.

Basado en la información hallada, en los parámetros y metadatos existentes y con la finalidad de delimitar el alcance del monitoreo, se seleccionaron en total seis tablas de datos como base para el modelo, las cuales se exploraron en términos de contenido, columnas, cantidad de registros, porcentaje de datos nulos, consistencia de la información. De esta forma, se pudo identificar el modelo de datos que se puede emplear en las siguientes etapas del proyecto, el cual está compuesto como se muestra en la Figura 6, por las tablas del

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

sistema de Leasing así como la tabla del almacén de datos; cabe resaltar que, por razones de confidencialidad y cantidad de columnas, la figura se muestra de manera resumida.

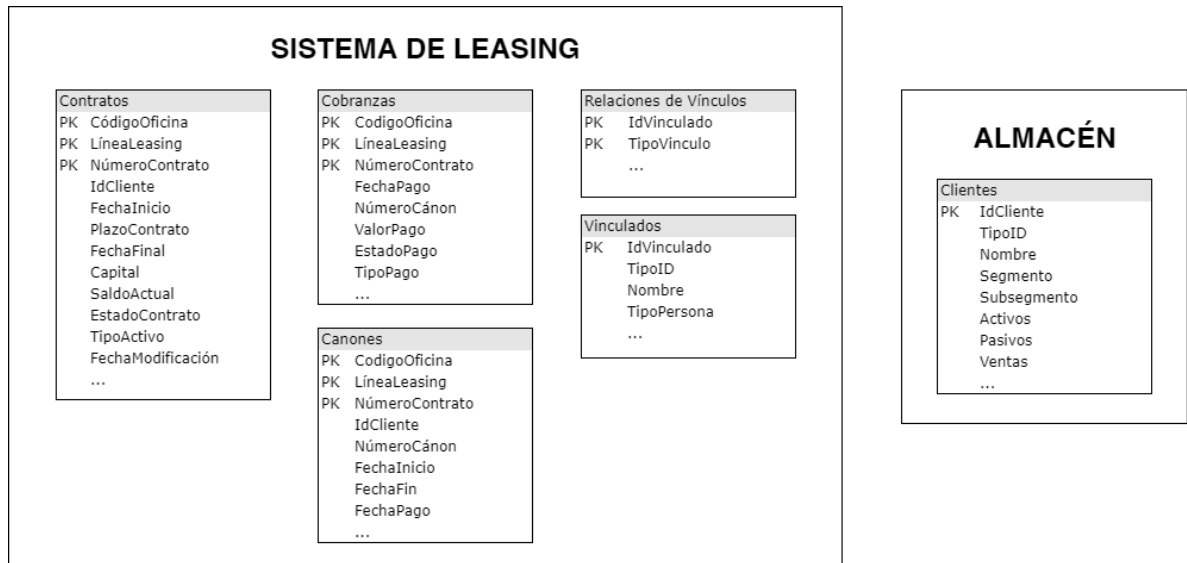


Figura 6: Fuentes de Datos Disponibles

3.1.4 Resultado de la Etapa: Definición del Alcance del Monitoreo

A partir de la información recolectada y los conocimientos adquiridos, se puede no solo comprender el funcionamiento del arrendamiento financiero, los actores involucrados en este y los tipos de leasing existentes dentro de la entidad financiera; sino también seleccionar una tipología de LAFT para esta línea de negocio y los datos necesarios para las siguientes etapas del monitoreo. De esta manera se define el alcance del monitoreo en términos de clientes a monitorear y la tipología LAFT, así: “Monitorear los pagos anticipados de los clientes sobre sus productos de Leasing activos y alertar aquellos que se consideren como potenciales casos de LAFT.”

La tipología seleccionada fue la de prepagos totales o parciales, debido a que en las fuentes de datos disponibles no existe una manera de identificar terceros que puede servirle a la organización criminal como apoderados o bien como beneficiarios de la cesión del contrato; mas, si se pueden identificar los comportamientos inusuales en los pagos de los cánones de los contratos, para lo cual es necesario seleccionar variables que permitan clasificar la población de tal forma que se puedan identificar las anomalías en los comportamientos de la misma; esta selección y clasificación hacen parte de las tareas para las siguientes etapas de la metodología: preparación de los datos y modelado.

Por último, en esta etapa se realizó una definición básica del enfoque de las señales de alerta a incluir, debido a que a medida que se avance en las etapas siguientes y como la metodología es iterativa e incremental, estas señales de alerta se van a refinar. La definición dada es: enfocar una señal de alerta al tipo de Arrendamiento Operativo, porque su comportamiento puede ser diferente en comparación al Leasing Financiero o Habitacional;

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

y enfocar otra a los otros dos tipos de arrendamiento financiero ofrecidos por la organización, los cuales comparten la opción de compra sobre el bien al final del contrato.

3.2 PREPARACIÓN DE LOS DATOS Y MODELADO

La metodología empleada para el desarrollo del monitoreo permite que existan entregas iterativas e incrementales, en las cuales permiten que al final de la construcción, se entregue un monitoreo efectivo, refinado y delimitado. Sin embargo, se presentan los resultados estructurados en cuatro *etapas* clave que se trabajaron para alcanzar el objetivo general del monitoreo y, asimismo, del proyecto; ya que, si bien existen dichos resultados generales en cada paso de la metodología, al aplicarla se identificó una alta correlación entre las mismas, como es de esperarse en un proyecto de minería de datos o bien, en los proyectos desarrollados por medio de metodologías ágiles de trabajo.

Puntualmente, en cada una de las entregas se preparó una vista de datos por medio de ELT con la interfaz de SQL de Cloudera, Impala (Cloudera, Inc., 2019); con cada vista se resolvieron los cuestionamientos de estas o bien se definió el uso de varias vistas de datos con el fin de construir un modelo estructurado y dinámico, como se explicará más adelante.

3.2.1 Etapa 1: Visualización

En la primera etapa de este ciclo de trabajo, se usaron herramientas de la inteligencia de negocios para identificar características del negocio por medio de los datos con el objetivo de detallar la información a modelar y refinar el alcance del monitoreo. Aquí se presentó la información existente y de manera general para responder a los cuestionamientos: ¿todos los clientes se comportan igual?, ¿todos los clientes representan el mismo nivel de riesgo?, ¿cómo se distribuyen los contratos?, ¿todos los tipos de contrato se comportan igual?, ¿qué tan común es que se realice un prepago?, entre otros. La información obtenida a partir de dichos cuestionamientos permitió plantear hipótesis para la definición de las señales de alerta para la tipología seleccionada y para la selección de las variables que permitan una correcta clasificación de los clientes y sus operaciones.

Para este proceso, las herramientas empleadas fueron Microsoft Power BI, un "...servicio de análisis empresarial que proporciona información detallada para permitir la toma de decisiones rápidas e informadas" (Microsoft, 2019), y SAS Enterprise Guide, una interfaz gráfica de SAS que permite el análisis y publicación de resultados de manera sencilla (SAS Institute Inc., 2019), disponibles en los equipos de la entidad financiera y seleccionadas por la necesidad existente de tener un objetivo construido correctamente para garantizar mejores resultados; con estas y aplicando conceptos de inteligencia de negocios, se halló y se diagramó la cantidad de clientes existentes en la entidad por tipo de persona, como se observa en la Figura 7, así como la participación monetaria por tipo de persona, como se observa en la Figura 8.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

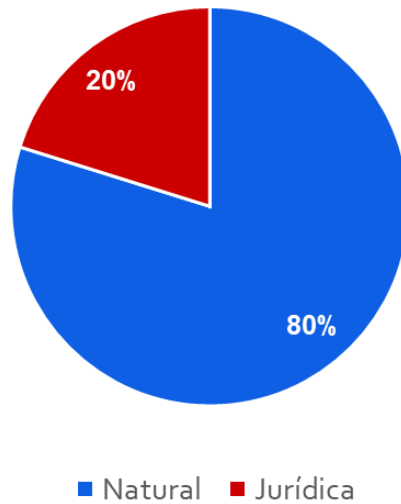


Figura 7: Porcentaje de Clientes por Tipo de Persona

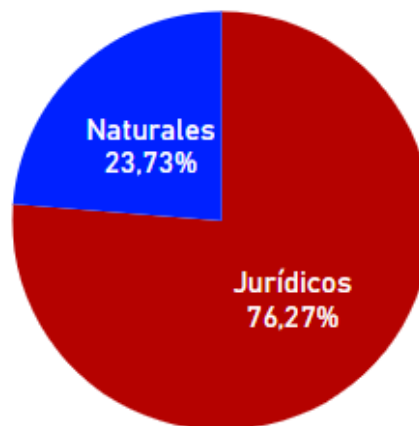


Figura 8: Porcentaje de Participación Monetaria por Tipo de Persona

Ambas figuras están representadas en porcentajes históricos, con un periodo de tiempo de 15 años, desde 2003 hasta 2018, y con ellas se concluye que para cubrir de manera adecuada toda la población a quienes la compañía de leasing presta sus servicios, es relevante separar el comportamiento de las Personas Naturales, las cuales representan la mayoría de los clientes, pero la minoría de participación monetaria; del comportamiento de las Personas Jurídicas, quienes representan la menor porción de clientes, pero tienen una mayor participación monetaria. También se pueden hacer este tipo de hipótesis respecto a la separación de los tipos de personas teniendo en cuenta la finalidad de la operación de leasing, que para personas naturales usualmente es la adquisición de vivienda familiar, mientras que para personas jurídicas puede ser la adquisición de activos especializados, la

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

importación o exportación de bienes, para resolver problemas de obsolescencia tecnológica o para obtener beneficios tributarios (ASOBANCARIA, 2016).

Del mismo modo se presentan gráficamente las cantidades de contratos activos para los tres tipos principales de arrendamiento financiero dentro de la entidad financiera; este resultado, que se observa en la Figura 9 es relevante para estructurar las señales de alerta enfocadas a la tipología seleccionada, ya que permite identificar el nicho del negocio en la entidad. Los datos están presentados en una gráfica TreeMap en términos de la cantidad en miles de contratos que ha emitido la entidad, para cada una de sus líneas de arrendamiento financiero.

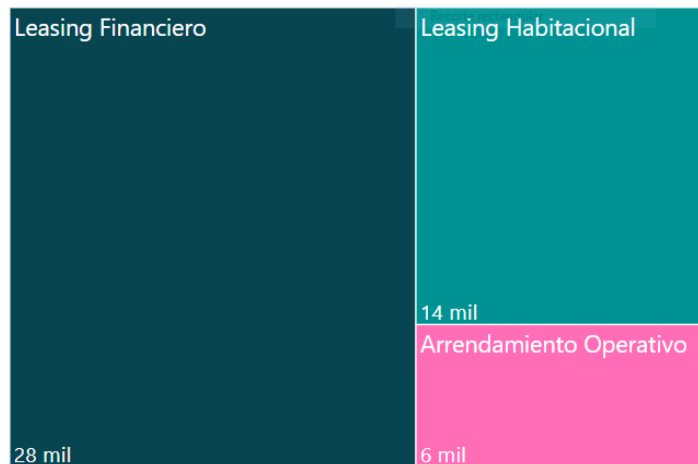


Figura 9: Cantidades de Contratos por Tipo de Leasing

Por último, se presentó gráficamente la información correspondiente al porcentaje de prepago por el Tipo de Persona (Figura 10), por el Tipo de Leasing (Figura 11) y por el Tipo de Activo (Figura 12); esta última categoría se refiere a una clasificación general disponible en los datos que guarda la entidad y que enmarca el bien o activo financiado dentro del contrato, como Maquinaria, Vehículos, Inmuebles, Tecnología u otros.

El conocimiento adquirido sobre el negocio y la exploración de los datos en la fase de Entendimiento del Negocio y Entendimiento de los Datos permitieron la generación de esta información, puesto que con estos es más sencillo identificar (o bien calcular) un prepago en la tabla de cobranzas del Sistema de Leasing (Figura 6: Fuentes de Datos Disponibles) y calcular el porcentaje de prepago respecto al valor total del contrato. Así, para la identificación de un prepago dentro de la tabla de cobranzas, se utilizan columnas de estado que identifican el tipo de cobranza como normal, abono o prepago y el estado del pago como total, parcial, final o extraordinario; y para calcular el porcentaje de prepago en el contrato para un periodo de tiempo mensual, se divide el valor total del pago de dicho periodo sobre el capital inicial financiado en el contrato.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

Adicionalmente al cálculo del porcentaje realizado sobre la tabla de cobranzas, se realizó un cálculo sobre la fecha del canon y fecha de pago disponibles en la tabla de cánones, para identificar posibles prepagos, sin embargo, el uso de esta fuente de datos fue descartado en esta etapa pues las variables disponibles en la misma no permiten la identificación de operaciones inusuales, debido a que el contenido de la tabla no referencia montos de pago, cantidad de operaciones ni tipo de pago, sino un plan de pagos del contrato.



Figura 10: Diagrama de Caja (Porcentaje de Prepago por Tipo de Persona)



Figura 11: Diagrama de Caja (Porcentaje de Prepago por Tipo de Leasing)

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

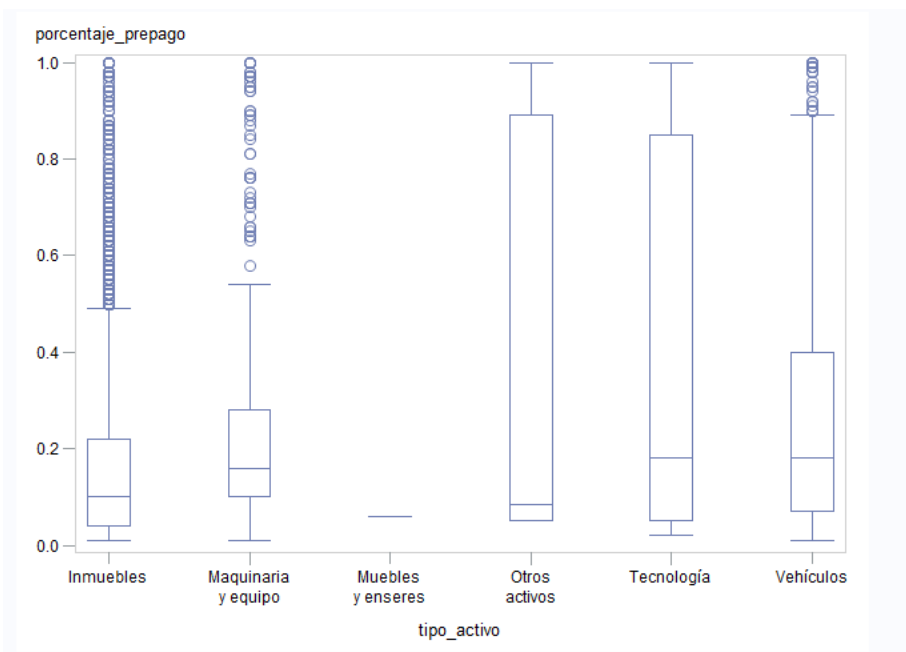


Figura 12: Diagrama de Caja (Porcentaje de Prepago por Tipo de Activo)

Para mostrar esta información se utilizaron diagramas de caja o *box plot*, en los que se observa una gran cantidad de valores atípicos, así como una alta dispersión en la distribución del porcentaje de prepagos, lo que significa que clasificar los datos teniendo en cuenta únicamente una variable (el tipo de persona, el tipo de leasing o el tipo de activo) en cada caso, podría generar una gran cantidad de falsos positivos luego del proceso de modelado y al generar las alertas.

Luego, y a partir de la información recolectada en esta etapa se planteó una hipótesis sobre el comportamiento de los prepagos: este comportamiento varía notoriamente dependiendo del Tipo de Persona, del Tipo de Leasing y del Tipo de Activo financiado; sin embargo, también surge la inquietud de ¿qué otras variables influyen en el comportamiento de los prepagos?

3.2.2 Etapa 2: Árboles de Clasificación

Luego de estudiar los resultados de la visualización, plantear la hipótesis de la variación en el comportamiento de los prepagos y cuestionarse sobre las demás variables que influyen sobre los prepagos, se tomó la decisión modelar la influencia de variables demográficas y económicas del cliente en los pagos anticipados a sus contratos u operaciones de arrendamiento financiero u operativo, con la finalidad de seleccionar aquellas que definan de mejor manera la normalidad en los prepagos de un cliente en específico o de un grupo de clientes con características comunes.

Este análisis se realizó por medio de la implementación de algoritmos de árboles de clasificación sobre una vista única de información en la herramienta de ciencia de datos

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

SPSS Modeler de IBM (IBM, 2015), con los que se concluyó el nivel de relevancia de dichas variables demográficas y económicas sobre la variable resultado: *prepagar* o no un contrato de leasing. Así, los datos de entrada seleccionados para su análisis se pueden observar en la Tabla 1, que relaciona el nombre dado a la variable en el modelo y la descripción de dicho atributo. Estas variables fueron seleccionadas teniendo en cuenta los datos disponibles tanto en el Sistema de Datos de Leasing, así como en el Almacén de Datos de la Entidad, en términos de calidad de datos (menor cantidad de valores nulos, cantidad de datos atípicos, actualización de los datos, entre otros).

Tabla 1: Datos de Entrada para el Árbol de Clasificación

Código	Descripción
edad_cli	Edad o Antigüedad
act_econom	Actividad Económica
cod_ciiu	Código CIIU
ric_total_cuantitativo	Riesgo del Cliente
segm	Segmento del Cliente
seg_sarlaft	Segmento SARLAFT
cod_ciudad_dirp	Ciudad
linea_leasing	Tipo de Leasing
cap_ini_finan	Capital Inicial Financiado
plazo_contrato	Plazo del Contrato
tipo_activo	Tipo de Activo
ventas	Total de Ventas
patrimonio	Patrimonio

Tras ejecutar el modelo, los resultados obtenidos con los algoritmos CHAID y C5.0, disponibles en la herramienta anteriormente mencionada, se observan en la Figura 13; de la que se puede concluir que de las variables de entrada seleccionadas anteriormente y que el árbol de clasificación analizó en una vista denominada “base de prepagos”, aquellas que pueden clasificar mejor los datos e identificar con más certeza un prepago son, además del tipo de persona, el tipo de leasing, el tipo de activo y el segmento del cliente, puesto que son aquellas con un porcentaje de importancia del clasificador mayor, es decir, con una mayor influencia sobre el resultado del modelo.

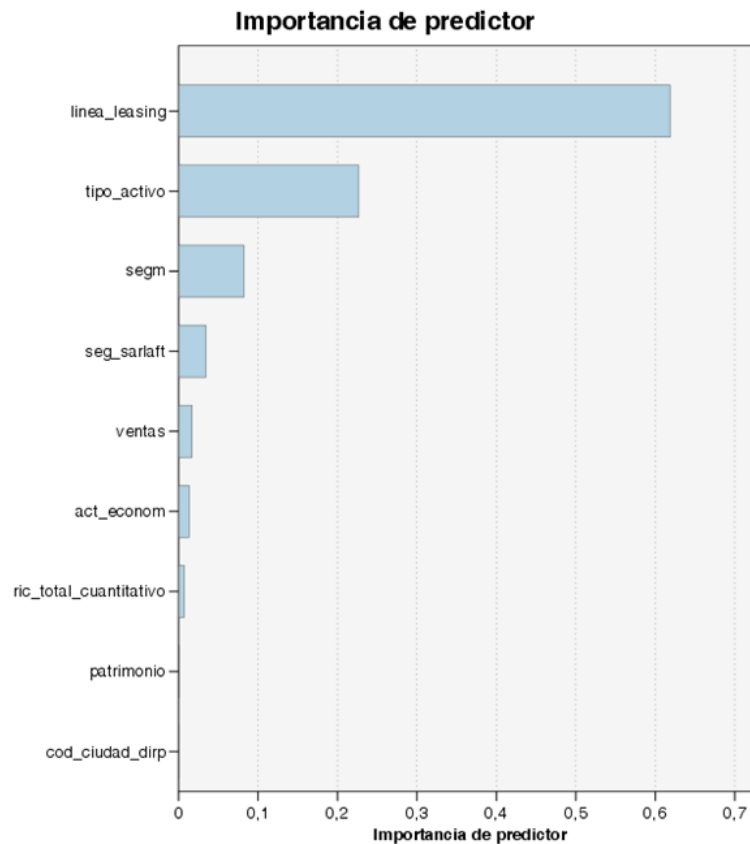


Figura 13: Importancia del Predictor

Al analizar los resultados obtenidos en el árbol de clasificación y teniendo en cuenta la opinión de expertos del negocio dentro de la entidad financiera, así como los resultados obtenidos en la etapa de visualización, se toma la decisión de modelar los datos teniendo en cuenta no solamente una variable de segmentación, sino tener en cuenta para esto las variables de entrada del tipo de persona, ya que la distribución y comportamiento de las Personas Naturales y las Personas Jurídicas es diferente por sus características propias; el tipo de leasing, el tipo de activo y el segmento del cliente (al interior de la entidad), ya que se acordó que estas son las variables más influyentes en el resultado de prepagar o no una obligación de Leasing.

Al finalizar esta etapa, se presentó la vista de prepagos con las variables de clasificación, es decir, aquellas columnas que permiten diferenciar adecuadamente las características de la población de clientes y sus contratos u obligaciones de arrendamiento financiero en la entidad respecto a la variable de análisis, la cual es el porcentaje de prepago sobre el capital inicial financiado de dichas obligaciones en un periodo especificado, e incluyendo información socioeconómica del cliente y de sus contratos, la cual es una base para estructura de las alertas a generarse en las siguientes etapas, como se observa en la Tabla 2.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

Tabla 2: Vista de Prepagos

Variable	Descripción	Variable de Clasificación	Variable de Análisis	Información para la Alerta
tipo_persona	Tipo de Persona	x		
tipo_id	Tipo de Identificación del Cliente			x
id_leasing	Número de Identificación del Cliente			x
apellido_nombre	Apellidos y Nombres del Cliente			x
act_econom	Actividad Económica			x
cod_ciiu	Código CIIU			x
total_ing	Ingresos			x
ventas	Ventas			x
tot_activos	Total de Activos			x
tot_pasivos	Total de Pasivos			x
riesgo_cliente	Riesgo del Cliente			x
segm	Segmento del Cliente	x		
subsegm	Subsegmento del Cliente			x
linea_leasing	Tipo de Leasing	x		
num_unico_contrato	Número del Contrato			x
f_inicio	Fecha de Inicio del Contrato			x
plazo_contrato	Plazo del Contrato			x
f_vencimiento	Fecha de Vencimiento del Contrato			x
cap_ini_finan	Capital Inicial Financiado			x
total_canon_pago	Total de Cánones Pagos			x
estado_contrato	Estado del Contrato			x
tipo_activo	Tipo de Activo	x		
periodo_prepago	Periodo del Prepago		x	
porcentaje_prepago	Porcentaje del Prepago		x	
total_prepago	Valor Total del Prepago			x
canitdad_pagos	Cantidad de Pagos			x

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

A partir de esta vista, de las variables de clasificación seleccionadas y de la exploración inicial del negocio y de los datos realizada anteriormente, se estructuró la primera señal de alerta, relativa al tipo de leasing Arrendamiento Operativo, en el cual por naturaleza no es común que se realicen prepagos parciales o totales: *Personas naturales o jurídicas, clientes de Leasing, que realicen pagos anticipados a sus obligaciones activas de Arrendamiento Operativo, dentro del periodo establecido.*

Por último, y trabajando sobre la misma vista de datos generada, se construyó un análisis estadístico descriptivo en el cual se estudió la variable de análisis (porcentaje de prepago sobre el capital inicial financiado de una operación de leasing) de manera histórica, es decir, sin tener en cuenta un periodo específico, y teniendo en cuenta las variables de clasificación seleccionadas anteriormente, con el objetivo de aplicar reglas estadísticas a la variable de análisis para así identificar los comportamientos inusuales por parte de los clientes de la entidad sobre sus operaciones que podrían considerarse como una señal de alerta.

Los resultados de este análisis se presentan de manera parcial en los ejemplos de la Figura 14 y de la Figura 15, que al ser comparados con resultados previamente hallados, como los de la Figura 10, la Figura 11, o la Figura 12, resultan en segmentos clasificados más acorde a los valores reales o al comportamiento del negocio, por lo que si se observan valores atípicos en los grupos, estos pueden entenderse como inusualidades en el porcentaje de prepago e incluso como comportamientos sospechosos, dado que el diagrama de caja y bigotes, refleja para cada uno de los grupos, el comportamiento usual de los clientes con un segmento particular que realizan prepagos sobre una operación de leasing con un tipo de leasing y tipo de activo específicos.

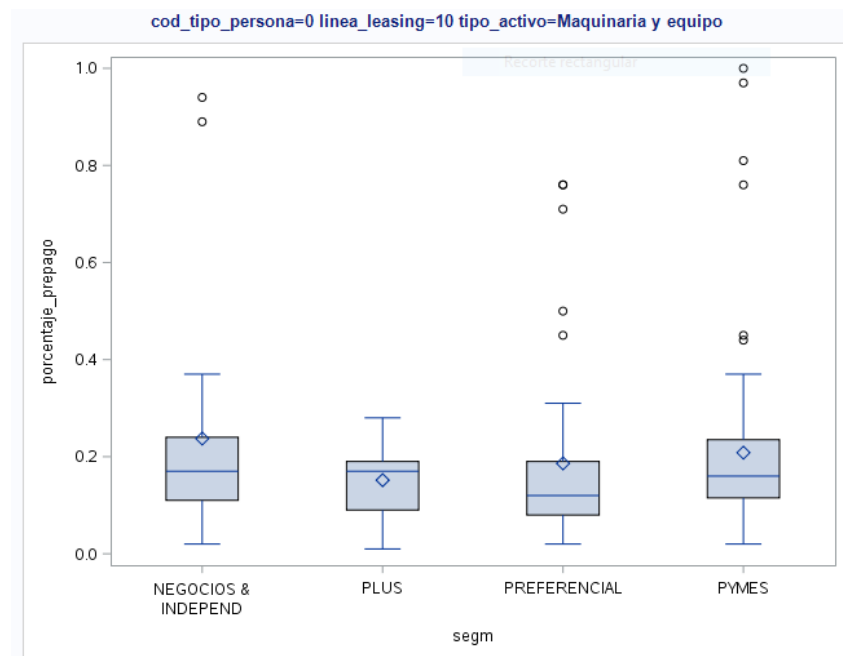


Figura 14: Ejemplo de Distribución del Porcentaje de Prepago para Personas Naturales Aplicando las Variables de Clasificación

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

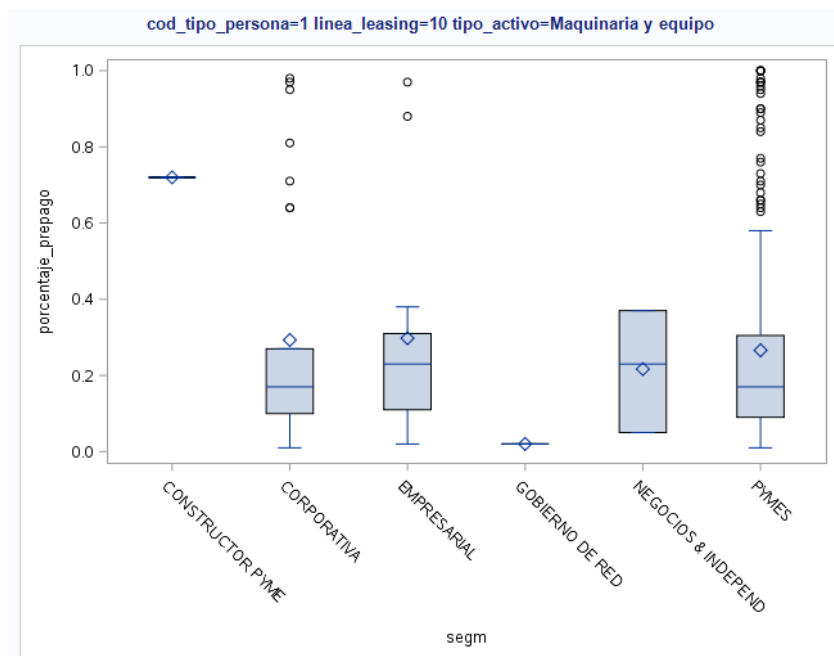


Figura 15: Ejemplo de Distribución del Porcentaje de Prepago para Personas Jurídicas Aplicando las Variables de Clasificación

3.2.3 Etapa 3: Modelo Estadístico

Partiendo del análisis estadístico descriptivo resultante de la etapa previa y sobre la base histórica de prepagos, se investigó la manera de identificar umbrales y límites para seleccionar dinámicamente las inusualidades presentes en el porcentaje de prepago clasificado por el tipo de persona, el tipo de leasing, el tipo de activo y el segmento del cliente. De esta investigación resultó la selección del método de Tukey (Lowry, 2008) para la clasificación de valores atípicos, para el cual es importante conocer las definiciones de valor atípico (*outlier*), que se refiere a aquellas observaciones que son distantes de los demás datos; puede entenderse como valores poco comunes o bien, basados en el interés de esta investigación, los valores inusuales en los pagos de obligaciones de leasing; y de rango intercuartil, referente a la diferencia entre el tercer y el primer cuartil de la distribución estadística; es usado en el diagrama de caja y bigotes, por ejemplo, para graficar la variabilidad de una distribución de datos.

Con el método de Tukey, entonces, se pueden calcular valores atípicos leves o internos y valores atípicos externos o extremos basados, como se observa en la Figura 16, cuya principal diferencia es la distancia a distribución usual de los datos.

Valor atípico interno > 1.5 * Diferencia
Intercuartil + Tercer Cuartil

Valor atípico externo > 3 * Diferencia
Intercuartil + Tercer Cuartil

Figura 16: Fórmulas de Cálculo de Valores Atípicos

Con este método, la investigación previa y las consideraciones estadísticas necesarias, se estructuró la segunda señal de alerta basada en la tipología seleccionada: *Personas naturales o jurídicas, clientes de Leasing, que realicen en el periodo establecido pagos anticipados a sus obligaciones de Leasing Financiero o Leasing Habitacional; cuyo porcentaje de prepago sea un valor atípico en la distribución de probabilidades histórica, segmentada por la línea de Leasing, el tipo de activo y el segmento comercial del cliente.*

Luego de ello, se aplicaron los cálculos a la vista de prepagos históricos, seleccionando únicamente aquellos clientes que, cuyos prepagos a obligaciones de leasing dentro de un periodo definido, y basándose en el método de Tukey, se consideraran como valores atípicos internos. Adicionalmente, se seleccionaron, para este mismo periodo, los clientes cuyos prepagos a obligaciones de leasing tuviesen solamente una observación histórica, es decir, aquellos prepagos que históricamente para cada tipo de persona, cada tipo de leasing, cada tipo de activo y cada segmento, solamente se hubiesen presentado una vez, ya que este comportamiento para la clasificación por grupos se considera inusual.

Con base entonces en las dos señales de alerta estructuradas hasta el momento, en los cálculos realizados sobre la vista base de prepagos, y para el periodo de octubre de 2018, se realizó una ejecución de prueba del monitoreo, con la que se identificaron 218 alertas, las cuales se pasaron a la fase de Evaluación a manera de prueba; esta fase será explicada en el numeral 3.3.2 de este documento con mayor detalle.

3.2.4 Etapa 4: Refinamiento del Modelo con Otras Variables

Posterior a la evaluación de prueba que resultó de la etapa anterior, y debido al alto número de alertas, al nivel de riesgo inherente a los montos de los contratos y a la capacidad del negocio para realizar las evaluaciones necesarias para tomar decisiones sobre los clientes (en etapas posteriores al alcance de este proyecto), se generaron acuerdos con los expertos del negocio para refinar las alertas generadas por el monitoreo, basándose en información propia de los contratos: las fechas de apertura y de cancelación, el canon en que se realiza el prepago y el capital inicial financiado en el contrato.

Para el caso de las fechas de apertura y cancelación, la hipótesis planteada fue la normalidad con la que se abren contratos con plazos mayores a un año y los mismos se cancelen dentro de los primeros once meses a través de un prepago total; concluyéndose que este tipo de comportamientos es inusual debido a las características inherentes del negocio de leasing, por lo que los casos que se presenten de este tipo, deben considerarse como una señal de alerta para un posible caso de LAFT.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

Para identificar este tipo de alertas se construyó en Impala una vista de contratos abiertos y cerrados en un periodo de tiempo definido por la periodicidad de ejecución del monitoreo (a este punto aún pendiente por definirse), partiendo de la información de cobranzas y la información de contratos del Sistema de Leasing de la entidad, y la inclusión de la señal de alerta relacionada a este tipo de comportamientos: *Personas naturales o jurídicas, clientes de Leasing, que abrieron contratos con plazo mayor a 12 meses en los últimos 11 meses y realizaron un pago anticipado parcial o total sobre las mismas con el cual cancelaron el contrato.*

Otra de las hipótesis identificadas con apoyo de la teoría y del negocio es qué tan comúnmente un cliente abre un contrato de leasing y realiza un prepago parcial en el primer canon que se considere un comportamiento inusual. Para analizar este planteamiento, se identificaron en la vista de prepagos la fecha de inicio del contrato y el periodo de pago, se analizó el comportamiento luego de seleccionar los valores atípicos internos y se llegó a la conclusión de que para ciertos tipos de activos, es común realizar prepagos en el primer canon, por lo que se reestructuró la señal de alerta relacionada a prepagos parciales de leasing financiero y leasing habitacional, con porcentajes definidos por los expertos, para los tipos de activo financiado que aplique esta excepción. De esta forma, se adicionaron las siguientes condiciones a la señal de alerta, válidas para los prepagos en el primer canon:

- *Los contratos con tipo de activo "Inmuebles" y cuyo porcentaje de pago sea mayor o igual al 30% del capital inicial financiado.*
- *Los contratos con tipo de activo "Vehículos" y cuyo porcentaje de pago sea mayor o igual al 40% del capital inicial financiado.*
- *Los contratos con tipo de activo "Maquinaria y equipo" y cuyo porcentaje de pago sea mayor o igual al 30% del capital inicial financiado.*

Asimismo, se construyó en Impala una vista con los campos que se observan en la Tabla 3 y partiendo de la información de clientes del Almacén de Datos de la entidad y de la información histórica de contratos del Sistema de Leasing, con la que se analizó el comportamiento del capital inicial financiado en los contratos para cada segmento de clientes, en los últimos tres años.

Tabla 3: Vista de Capitales

Variable	Descripción	Variable de Clasificación	Variable de Análisis
tipo_persona	Tipo de Persona	x	
linea_leasing	Tipo de Leasing	x	
tipo_activo	Tipo de Activo Financiado	x	
segm	Segmento del Cliente	x	
fecha_inicio	Fecha de Inicio del Contrato	x	
cap_ini_finan	Capital Inicial Financiado		x

Después se aplicó un análisis de estadística descriptiva a la variable de análisis, con el que se identificó la distribución de esta y se decidió, en compañía de los expertos del negocio,

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

por descartar las alertas generadas por el monitoreo en las que el capital inicial financiado en el contrato fuese menor al promedio, agrupado por las variables de clasificación mencionadas. Es relevante mencionar que para las ejecuciones del monitoreo este filtro se considera como dinámico, ya que varía en el tiempo sin necesidad de ser intervenido por un usuario en cada ejecución.

Luego de aplicar los filtros y excepciones resultantes de los análisis de esta etapa, se transformaron todas las alertas generadas para enfocarla a los clientes y no a los contratos, es decir, para evitar que se presente la situación en que, si un cliente realizó prepagos sobre varios contratos en el mismo periodo, se cuente cada uno de estos prepagos como una alerta, generando varias alertas para un mismo cliente, que se pueden resumirse en una sola alerta estructurada.

Por último, se construyó en Impala y por petición del negocio una vista histórica de recurrencia en el prepagado, parcial o total, por parte de los clientes sobre sus obligaciones de Leasing, la cual tiene el objetivo de complementar la información del cliente que es generada en las alertas.

3.2.5 Modelo Final

Al culminar el proceso de modelado y el proceso de refinamiento del modelo, el monitoreo resultante tiene las siguientes características:

1. Tipología LAFT a alertar: Prepagos parciales o totales de obligaciones de Leasing con dinero de origen ilícito: "Esta tipología hace referencia a locatarios que pagan anticipadamente las obligaciones de leasing con dinero proveniente de actividades ilícitas. Se presenta con mayor frecuencia en contratos de leasing relacionados con bienes necesarios para producir renta, por ejemplo: importaciones de maquinaria y equipo, bienes de capital, vehículos de transporte de carga o de pasajeros, etc." (Unidad de Información y Análisis Financiero, s.f.)
2. Señales de alerta modeladas:
 - a. Personas naturales o jurídicas, clientes de Leasing, que realicen pagos anticipados a sus obligaciones activas de Arrendamiento Operativo con capital inicial financiado mayor o igual al promedio de los últimos 3 años para el tipo de activo financiado y segmento del cliente, dentro del periodo establecido.
 - b. Personas naturales o jurídicas, clientes de Leasing, que abrieron contratos con plazo mayor a 12 meses en los últimos 11 meses y realizaron un pago anticipado sobre las mismas con el cual cancelaron el contrato; con capital inicial financiado en el contrato mayor o igual al promedio de los últimos 3 años para la línea de leasing, el tipo de activo financiado y el segmento del cliente.
 - c. Personas naturales o jurídicas, clientes de Leasing, que realicen en el periodo establecido pagos anticipados a sus obligaciones de Leasing Financiero o Leasing Habitacional con capital inicial financiado mayor o igual al promedio de los últimos 3 años para la línea de leasing, el tipo de activo financiado y el segmento del cliente; cuyo porcentaje de prepagado sea

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

un valor atípico en la distribución de probabilidades histórica, segmentada por la línea de Leasing, el tipo de activo y el segmento comercial del cliente. Adicionalmente, si el pago se realiza en el primer canon, se tienen en cuenta las siguientes características para considerarse una señal de alerta:

- i. Los contratos con tipo de activo "Inmuebles" y cuyo porcentaje de pago sea mayor o igual al 30% del capital inicial financiado.
 - ii. Los contratos con tipo de activo "Vehículos" y cuyo porcentaje de pago sea mayor o igual al 40% del capital inicial financiado.
 - iii. Los contratos con tipo de activo "Maquinaria y equipo" y cuyo porcentaje de pago sea mayor o igual al 30% del capital inicial financiado.
3. Vistas generadas para su ejecución:
- a. Vista de prepagos: información histórica de los prepagos, base para la generación de las alertas del tipo de arrendamiento operativo y del tipo de leasing financiero y leasing habitacional, sobre la cual se seleccionan los prepagos en el trimestre a analizar y a la que se aplica el método de Tukey para identificar valores atípicos, así como los filtros del capital inicial financiado, las excepciones del primer canon, el resumen de alertas generadas por cliente y la adición de la información de recurrencia sobre los prepagos, útil para los evaluadores de estas en el negocio.
 - b. Vista de capitales: información agrupada del capital inicial financiado en los contratos de leasing, con la que se establece un filtro para todos los tipos de alerta generados.
 - c. Vista de contratos abiertos y cerrados: información de los contratos abiertos en los últimos once meses y cancelados por medio de prepagos en el trimestre a analizar, y a la que se le aplican los filtros del capital inicial financiado, el resumen de alertas generadas por cliente y la adición de la información de recurrencia sobre los prepagos.
 - d. Vista de recurrencia: información histórica de la recurrencia en el prepago de obligaciones de arrendamiento financiero, tanto parciales como totales, por parte de los clientes.

El modelo final entonces pasa a una fase de evaluación en un entorno real, es decir, se realiza un piloto del monitoreo con casos y alertas reales, como se explicará en el numeral 3.3.3, con el objetivo de desplegarlo entonces en la entidad financiera.

3.3 EVALUACIÓN Y DESPLIEGUE

La fase de evaluación y despliegue del monitoreo se realizó en cuatro etapas clave, en las cuales se evaluaron los resultados obtenidos de manera iterativa e incremental en las etapas anteriores tanto por un modelo prescriptivo, así como por expertos del negocio. Esta etapa es fundamental en el desarrollo de proyectos de minería de datos, puesto que permite refinar los modelos construidos, de tal forma que se genere un alto valor para el usuario, que en este caso es la entidad financiera.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

3.3.1 Evaluación Transversal: Modelo Prescriptivo

La primera forma de evaluar el monitoreo en la tercera y en cuarta etapa de las fases de preparación de los datos y modelado de la metodología, presentadas en los numerales 3.2.3 y 3.2.4 del presente, fue a través de un modelo prescriptivo construido al interior de la entidad financiera, el cual se basa en información histórica y en variables de riesgo LAFT para decidir sobre la evaluación de cada alerta, o bien su cierre, su reporte como una operación sospechosa a los entes reguladores o la necesidad de ser analizada por un evaluador. Este tipo de evaluación se consideró como transversal ya que fue aplicado como una etapa previa a la evaluación de las alertas (en cada una de las iteraciones), por los expertos del negocio en la entidad financiera.

3.3.2 Prueba del Monitoreo: ¿Son o no alertas?

Al finalizar la Etapa 3: Modelo Estadístico de la fase anterior de la metodología, se generaron 218 alertas, las cuales se pasaron a la fase de Evaluación a manera de prueba, así: Las alertas se evaluaron con el modelo prescriptivo, que seleccionó en total 90 alertas para su evaluación por los expertos del negocio, 1 alerta con sugerencia de reporte a las entidades reguladoras, y 127 alertas descartadas. De las alertas que se consideraron para evaluación y para sugerencia de reporte a las entidades reguladoras se seleccionó una muestra de 35 alertas para un análisis básico por los expertos del negocio, con el cual se pudiese identificar si estos casos sí se consideran como alerta o no y poder así refinar el modelo en planteando hipótesis a partir de la información hallada en las alertas. Del mismo modo se seleccionaron 6 alertas de las que el modelo prescriptivo descartó, con las cuales se buscó identificar si la clasificación de estas fue la correcta, es decir, si debían ser descartadas en relación con las variables de riesgo usadas por el modelo prescriptivo.

A partir de la información obtenida en esta prueba respecto a las alertas generadas, se definieron los puntos con los que se refinó el modelo en la Etapa 4: Refinamiento del Modelo con Otras Variables.

3.3.3 Piloto del Monitoreo: Evaluación con Casos Reales

Posterior a la Etapa 4: Refinamiento del Modelo con Otras Variables e implementando el Modelo Final, se llevó a cabo un piloto en la ejecución del monitoreo, en el cual se analizó el periodo comprendido entre octubre y diciembre de 2018 y en el cual las alertas generadas por cada una de las señales de alerta construidas, fueron evaluadas por expertos de la entidad financiera luego de ser evaluadas por el modelo prescriptivo. Durante la ejecución del piloto se generaron 75 alertas, dentro de las cuales 42 correspondieron a la señal de alerta de contratos abiertos y cerrados y 33 a la señal de alerta de prepagos en leasing financiero y leasing habitacional; luego de la ejecución del modelo prescriptivo, pasaron a evaluación por expertos 5 alertas, con una estructura como la de los ejemplos presentados en la Figura 17.

« El cliente realizó un prepago en Leasing correspondiente al (los) contrato(s) 123456 en el (los) periodo(s) 201811 sobre el capital inicial financiado \$952,407,923 y por monto(s) de \$872,523,136, con el cual canceló este (estos) contrato(s) aperturados en los últimos 11 meses. Adicionalmente, ha tenido recurrencia de prepagos en 5 ocasiones. »

« El cliente realizó un prepago en Leasing correspondiente al (los) contrato(s) 345678 en el (los) periodo(s) 201810, que corresponde respectivamente a 72% del capital inicial financiado \$297,864,000 y por monto(s) de \$216,387,095. Adicionalmente, ha tenido recurrencia de prepagos en 0 ocasiones. »

Figura 17: Ejemplos de Alertas

Posteriormente, se les solicitó a los expertos concluir sobre las alertas evaluadas, con el objetivo de proceder con el despliegue del monitoreo, o bien con una etapa adicional de refinamiento; en general se concluyó que todas las alertas cumplieron las condiciones del monitoreo, apuntándole a la tipología seleccionada, que el porcentaje menor de prepago fue del 70% y el mayor de 97% del valor inicial, que en promedio para las alertas, el valor total prepagado corresponde al 91.74% del valor inicial, que si bien todas las alertas fueron justificadas y soportadas por los clientes y por ello no se detecta riesgo para la entidad financiera, las reglas, filtros y variables empleadas en el monitoreo permiten monitorear las operaciones de los clientes de una entidad financiera que presta servicios de arrendamiento financiero, para identificar comportamientos inusuales, por lo que la recomendación y decisión de esta evaluación es proceder con el despliegue del monitoreo al interior de la entidad financiera.

3.3.4 Despliegue del Monitoreo

Al finalizar el piloto para la evaluación de las alertas generadas por el monitoreo, se procede a desplegarlo al interior de la entidad financiera, así:

1. Objetivo del monitoreo: Monitorear los pagos anticipados de los clientes sobre sus productos de Leasing y alertar aquellos que se consideren como potenciales casos de LAFT.
2. Tipología LAFT a alertar: Prepagos parciales o totales de obligaciones de Leasing con dinero de origen ilícito.
3. Señales de alerta modeladas:
 - a. Personas naturales o jurídicas, clientes de Leasing, que realicen pagos anticipados a sus obligaciones activas de Arrendamiento Operativo con capital inicial financiado mayor o igual al promedio de los últimos 3 años para el tipo de activo financiado y segmento del cliente, dentro del periodo establecido.
 - b. Personas naturales o jurídicas, clientes de Leasing, que abrieron contratos con plazo mayor a 12 meses en los últimos 11 meses y realizaron un pago anticipado sobre las mismas con el cual cancelaron el contrato; con capital inicial financiado en el contrato mayor o igual al promedio de los últimos 3

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

años para la línea de leasing, el tipo de activo financiado y el segmento del cliente.

- c. Personas naturales o jurídicas, clientes de Leasing, que realicen en el periodo establecido pagos anticipados a sus obligaciones de Leasing Financiero o Leasing Habitacional con capital inicial financiado mayor o igual al promedio de los últimos 3 años para la línea de leasing, el tipo de activo financiado y el segmento del cliente; cuyo porcentaje de prepago parcial sea un valor atípico en la distribución de probabilidades histórica, segmentada por la línea de Leasing, el tipo de activo y el segmento comercial del cliente. Adicionalmente, si el pago se realiza en el primer canon, se tienen en cuenta las siguientes características para considerarse una señal de alerta:
 - i. Los contratos con tipo de activo "Inmuebles" y cuyo porcentaje de pago sea mayor o igual al 30% del capital inicial financiado.
 - ii. Los contratos con tipo de activo "Vehículos" y cuyo porcentaje de pago sea mayor o igual al 40% del capital inicial financiado.
 - iii. Los contratos con tipo de activo "Maquinaria y equipo" y cuyo porcentaje de pago sea mayor o igual al 30% del capital inicial financiado.
4. Periodicidad de ejecución: trimestral.
5. Herramientas en las que se debe ejecutar el monitoreo: Cloudera Impala para la generación de las vistas de datos y SAS Enterprise Guide 7.1 para la ejecución de las reglas y filtros estadísticos.

4. CONCLUSIONES Y CONSIDERACIONES FINALES

- La existencia y uso de herramientas analíticas, técnicas estadísticas y modelos computacionales permiten el desarrollo de modelos de detección de problemáticas como el Lavado de Activos y la Financiación del Terrorismo por las entidades financieras colombianas, ya que su enfoque se basa en la aplicación de la ciencia de los datos a grandes volúmenes de información y en parámetros objetivos y autónomos en su ejecución, conceptos que no son ajenos al sector real.
- La implementación de metodologías ágiles, iterativas e incrementales de trabajo, como CRISP-DM es efectiva para el desarrollo y afinamiento de modelos de detección del Lavado de Activos en el sector financiero colombiano, ya que sus etapas permiten la caracterización efectiva de las necesidades del negocio, así como la delimitación adecuada del alcance de estos modelos; debido a que es notoria la interconexión existente entre todas las etapas de la metodología, una capacidad de acción y de decisión efectivas mayores son identificadas durante la construcción de modelos de ciencia de los datos.
- La caracterización de los valores atípicos y la selección de umbrales dinámicos por medio de técnicas estadísticas para la clasificación de los datos son métodos útiles en la identificación de comportamientos inusuales de los clientes respecto a sus operaciones de arrendamiento financiero, debido a que su comprensión y aplicación son simples y los conocimientos que requieren son aquellos de estadística descriptiva básica y la determinación de las alertas varía dependiendo de las características de los datos.
- El uso de un algoritmo de clasificación para la selección de las variables que mejor agrupan la información es un enfoque acertado para construir un modelo más preciso ya que al estar compuesto de iteraciones y técnicas computacionales puede significar la reducción de la subjetividad en la selección de estas.
- El planteamiento y validación de hipótesis en las diferentes etapas de construcción de un monitoreo para la identificación de comportamientos inusuales es un factor importante en la definición de un alcance adecuado para el mismo, puesto que permite afianzar o refutar los supuestos existentes referentes al negocio, a los datos o al comportamiento de los clientes y usuarios involucrados en el entorno a monitorear.
- Es importante el acompañamiento de los usuarios en el desarrollo de un proyecto de minería de datos para la identificación de operaciones sospechosas, sobre todo en las etapas de exploración del negocio, exploración de los datos y en la evaluación de los resultados obtenidos, ya que su experticia en el negocio, sus conocimientos empíricos y teóricos y las hipótesis que plantean respecto al comportamiento de la población en el negocio son clave para alcanzar los resultados esperados.
- El conocimiento transversal sobre el negocio, los datos y la analítica aportan a la construcción de monitoreos útiles para las entidades financieras, porque simplifican el proceso de construcción y aumentan la confianza de monitorear las variables adecuadas del entorno.
- Puede ser de interés para el sector financiero aplicar esta metodología y las técnicas presentadas en este trabajo, así como otros modelos estadísticos y computacionales para la prevención y detección de la Financiación del Terrorismo.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

- Asimismo puede interesarle a las entidades financieras adquirir más información sobre la operación del negocio, bien sea partiendo de almacenes de datos internos de las entidades financieras o bien de fuentes de información públicas, para poder identificar otras señales de alerta y así aplicar modelos similares a otras tipologías de LAFT en el sector, debido a que con la información recolectada actualmente se hace más difícil la identificación de estas, como lo son el uso de testaferros o la cesión de contratos.
- El modelo construido durante la investigación podría servir de referencia en el futuro a otros sectores de la economía y teniendo en cuenta los demás verbos rectores del delito de Lavado de Activos, definidos en el Código Penal Colombiano.

REFERENCIAS

- Agile Alliance. (2001). *Manifiesto for Agile Software Development*. Obtenido de Manifiesto for Agile Software Development: <http://agilemanifiesto.org/iso/es/manifiesto.html>
- Alaimo, D. M. (2013). *Proyectos ágiles con Scrum: flexibilidad, aprendizaje, innovación y colaboración en contextos complejos* (Primera ed.). Ciudad Autónoma de Buenos Aires, Argentina: Kleer.
- ASOBANCARIA. (2016). *LEASING | ASOBANCARIA*. Recuperado el 17 de Mayo de 2019, de ASOBANCARIA: <https://www.asobancaria.com/leasing/leasing/>
- Cloudera, Inc. (7 de Mayo de 2019). *Cloudera Impala Overview | 5.3.x | Cloudera Documentation*. Obtenido de Cloudera Impala Overview | 5.3.x | Cloudera Documentation: https://www.cloudera.com/documentation/enterprise/5-3-x/topics/impala_intro.html
- Correa Chaparro, D. (2015). *Optimización del proceso de monitoreo de transacciones (SARLAFT)*. Trabajo de Grado Pregrado, Pontificia Universidad Javeriana, Facultad de Ciencias Económicas y Administrativas, Bogotá.
- Dedić, N., & Stanier, C. (17 de Marzo de 2017). Measuring the success of changes to Business Intelligence solutions to improve Business Intelligence reporting. *Journal of Management Analytics*(DOI: 10.1080/23270012.2017.1299048), 130-144.
- Dell Inc. (2018). *¿Qué es big data? | Dell EMC Colombia*. Recuperado el 15 de Noviembre de 2018, de Dell EMC Colombia: <https://www.dellemc.com/es-co/big-data/definitions.htm>
- Dell Inc. (2018). *What is Hadoop | EMC Glossary*. Recuperado el 12 de Noviembre de 2018, de EMC Glossary: <https://colombia.emc.com/corporate/glossary/hadoop.htm>
- EcuRed. (2017). *Sistema Gestor de Base de Datos - EcuRed*. Obtenido de EcuRed: https://www.ecured.cu/Sistema_Gestor_de_Base_de_Datos
- Edu2000 America Inc. (1995-2006). *Definición: cuartil*. Recuperado el 17 de Mayo de 2019, de Diccionario Visual de Matemáticas: <http://www.mathematicsdictionary.com/spanish/vmd/full/q/quartile.htm>
- FAFT-GAFI. (2019). *About - Financial Action Task Force (FATF)*. Recuperado el 17 de Mayo de 2019, de FATF-GAFI.ORG - Financial Action Task Force (FATF): <https://www.fatf-gafi.org/about/>
- GAFILAT. (2016). *Recopilación de Tipologías Regionales de GAFILAT 2009 – 2016*. Buenos Aires.

- García Gibson, R. (14 de Julio de 2015). Las 3 etapas del lavado de dinero. *Forbes México*. Recuperado el 17 de Mayo de 2019, de <https://www.forbes.com.mx/las-3-etapas-del-lavado-de-dinero/>
- Grupo Bancolombia. (2018). *Precauciones contra el lavado de activos*. Recuperado el 30 de Noviembre de 2018, de Grupo Bancolombia: <https://www.grupobancolombia.com/wps/portal/personas/aprender-es-facil/seguridad/lavado-activos>
- Grupo de Acción Financiera de Latinoamérica. (2019). *Organismo Internacional - GAFILAT*. Recuperado el 17 de Mayo de 2019, de Inicio - GAFILAT: <http://www.gafilat.org/index.php/es/gafilat/quienes-somos/organismo-internacional>
- Heller, M. (29 de Enero de 2007). REST and CRUD: the Impedance Mismatch. *InfoWorld from IDG*. Obtenido de Developer World: <https://www.infoworld.com/article/2640739/application-development/rest-and-crud--the-impedance-mismatch.html>
- IBM. (15 de Agosto de 2015). *SPSS Modeler - Visión General - Colombia | IBM*. Recuperado el Mayo de 2019, de SPSS Modeler - Visión General - Colombia | IBM: <https://www.ibm.com/co-es/products/spss-modeler>
- IBM Corporation. (2012). Manual CRISP-DM de IBM SPSS Modeler. Chicago, IL, USA.
- Jensen, K. (26 de Abril de 2012). *A diagram showing the relationship between the different phases of CRISP-DM and illustrates the recursive nature of a data mining project*. Recuperado el 1 de Febrero de 2019, de https://es.wikipedia.org/wiki/Archivo:CRISP-DM_Process_Diagram.png
- Leffingwell, D. (2011). *Agile software requirements: lean requirements practices for teams, programs, and the enterprise*. Boston: Pearson Education, Inc.
- Lowry, R. (17 de Octubre de 2008). One Way ANOVA – Independent Samples. Poughkeepsie, NY, USA. Recuperado el 15 de Mayo de 2019, de One Way ANOVA – Independent Samples: II: <https://web.archive.org/web/20081017161620/http://faculty.vassar.edu/lowry/ch14pt2.html>
- Lozano Consultores en Prevención de Riesgo LAFT. (21 de Julio de 2017). *¿Qué es Sarlaft? - Infolaft*. Obtenido de Infolaft: <https://www.infolaft.com/que-es-sarlaft/>
- Martin, J. (1983). *Managing the Data-base Environment*. Englewood Cliffs, New Jersey, Estados Unidos de América: Prentice-Hall.
- Martínez Luna, G. L. (Julio - Septiembre de 2011). Minería de datos: Cómo hallar una aguja en un pajar. *Ciencia*, 62(3), 55-63.

La información presentada en este documento es de exclusiva responsabilidad de los autores y no compromete a la EIA.

- Melnik, R. (2015). *Mathematical and Computational Modeling: With Applications in Natural and Social Sciences, Engineering, and the Arts*. Wiley.
- Microsoft. (2015). *Conceptos básicos de ODBC*. Recuperado el 15 de Noviembre de 2018, de Catálogo de referencia y API de Microsoft: <https://msdn.microsoft.com/es-es/library/thzzea08.aspx?f=255&MSPPEror=-2147217396>
- Microsoft. (3 de Diciembre de 2018). *Ajuste de los hiperparámetros de un modelo - Azure Machine Learning service | Microsoft Docs*. Recuperado el 17 de Mayo de 2019, de Microsoft Azure: plataforma y servicios de informática en la nube: <https://docs.microsoft.com/es-es/azure/machine-learning/service/how-to-tune-hyperparameters>
- Microsoft. (2019). *Power BI | Herramientas de BI para la visualización de datos interactivos*. Obtenido de Power BI | Herramientas de BI para la visualización de datos interactivos: <https://powerbi.microsoft.com/es-es/>
- Morelo Tapias, K. A., & Puello Marrugo, P. (2014). *Sistema para la caracterización de perfiles de clientes de la empresa zona T*. Universidad de Cartagena, Facultad de Ciencias e Ingenierías. Cartagena: Universidad de Cartagena.
- Mosley, M., & Brackett, M. (2010). *DAMA-DMBOK: Data Management Body of Knowledge*. Bradley Beach: Technics Publications.
- Moujahid, A., Inza, I., & Larrañaga, P. (2008). *Tema 8: Árboles de Clasificación*. Universidad del País Vasco, Departamento de Ciencias de la Computación e Inteligencia Artificial. Recuperado el 17 de Mayo de 2019, de <http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t8s-arboles>
- Oficina de las Naciones Unidas contra la Droga y el Delito. (2011). *Riesgo de Lavado de Activos en Instrumentos Financieros y Comerciales (Segunda Versión)*. Bogotá, Colombia: Impresol Ediciones.
- Oracle Corporation. (2017). *About MySQL*. Obtenido de MySQL: <https://www.mysql.com/about/>
- Perotti, J. (Mayo de 2009). LA PROBLEMÁTICA DEL LAVADO DE DINERO Y SUS EFECTOS GLOBALES: UNA MIRADA A LAS INICIATIVAS INTERNACIONALES Y LAS POLÍTICAS ARGENTINAS. *Revista UNISC*(20), 78-99. Recuperado el 30 de Noviembre de 2018, de <https://www.redalyc.org/articulo.oa?id=76711408007>
- Preslar, E., & Rouse, M. (2005-2019). *What is Extract, Load, Transform (ELT)? - Definition from WhatIs.com*. Recuperado el 17 de Mayo de 2019, de Data Management/Data Warehousing information, news and tips - SearchDataManagement: <https://searchdatamanagement.techtarget.com/definition/Extract-Load-Transform-ELT>

- Python Software Foundation. (2001-2018). *About Python™* | *Python.org*. Recuperado el 12 de Noviembre de 2018, de Python.org: <https://www.python.org/about/>
- Roa Ferrada, A. R., & Fuente Mella, H. d. (s.f.). *Estudio del comportamiento y los hábitos de compra de los clientes de una multitienda, mediante el uso de técnicas de minería de datos*. Universidad de Talca, Facultad de Ingeniería. Talca, Chile: Universidad de Talca (Chile).
- SAS Institute Inc. (2018). *¿Qué es el big data?* | SAS. Obtenido de Analytics, Business Intelligence and Data Management | SAS: https://www.sas.com/es_co/insights/big-data/what-is-big-data.html
- SAS Institute Inc. (2018). *¿Qué es ETL?* | SAS. Obtenido de SAS: https://www.sas.com/es_ar/insights/data-management/what-is-etl.html
- SAS Institute Inc. (2019). *Graphical User Interface for SAS, SAS Enterprise Guide* | SAS. Obtenido de Graphical User Interface for SAS, SAS Enterprise Guide | SAS: https://www.sas.com/en_us/software/enterprise-guide.html
- Skinner, C. (17 de Enero de 2018). *Using Redshift Spectrum to load data pipelines* | *Dativa*. Recuperado el 17 de Mayo de 2019, de Dativa | Consulting | Data Science | Data Analytics | Data Strategy: <https://www.dativa.com/using-amazon-redshift-spectrum-data-pipelines/>
- Superintendencia de Sociedades. (27 de Septiembre de 2017). *Prevención del Riesgo de Lavado de Activos y Financiación del Terrorismo*. Recuperado el 30 de Noviembre de 2018, de Superintendencia de Sociedades: https://www.supersociedades.gov.co/delegatura_aec/informes_publicaciones/lavado_activos/Paginas/default.aspx
- Superintendencia Financiera de Colombia. (3 de Mayo de 2019). *Circular Básica Jurídica (C.E. 029/14)*. Recuperado el 17 de Mayo de 2019, de Superintendencia Financiera de Colombia: <https://www.superfinanciera.gov.co/inicio/circular-basica-juridica-ce---10083443>
- Torres Chero, E., & Farroñay Julca, J. (2017). *Implementación de Minería de Datos para detectar Patrones de Comportamiento de Clientes Morosos en Empresa de Crédito Crediserv EIRL – Chiclayo/HICLAYO*. Tesis de Grado, Universidad Nacional Pedro Ruiz Gallo, Facultad de Ciencias Físicas y Matemáticas.
- Unidad de Información y Análisis Financiero. (2014). *Lo que debe saber sobre el lavado de activos y la financiación del terrorismo* (Segunda ed.). Colombia: Unidad de Información y Análisis Financiero. Recuperado el 30 de Noviembre de 2018, de http://www.uosario.edu.co/observatorio-de-lavado-de-activos/Archivos_Lavados/Lo-que-debe-saber-sobre-LAFT.pdf

- Unidad de Información y Análisis Financiero. (2018). *Prevención y Detección del LA/FT - UIAF*. Obtenido de UIAF: https://www.uiaf.gov.co/sistema_nacional_ala_cft_lavado_activos_financiacion_29271_prevencion_deteccion_la_ft_29737
- Unidad de Información y Análisis Financiero. (s.f.). *Compilación de Tipologías de Lavado de Activos y Financiación del Terrorismo 2004 - 2013*. Bogotá. Recuperado el 15 de Mayo de 2019, de https://www.urosario.edu.co/observatorio-de-lavado-de-activos/Archivos_Lavados/TIPOLOGIAS-DE-LAVADO-DE-ACTIVOS-Y-FINANCIACION-DEL/
- United Nations Office on Drugs and Crime (UNODC). (2011). *Estimating illicit financial flows resulting from drug trafficking and other transnational organized crimes*. Vienna: United Nations Office on Drugs and Crime (UNODC). Recuperado el 30 de Noviembre de 2018, de https://www.unodc.org/documents/data-and-analysis/Studies/Illicit_financial_flows_2011_web.pdf
- Westreicher, G. (2019). *Arrendamiento operativo - Definición, qué es y concepto | Economipedia*. Recuperado el 17 de Mayo de 2019, de Economipedia: <https://economipedia.com/definiciones/arrendamiento-operativo.html>