

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Gestión del riesgo LAFTFPADM de jurisdicciones para activos virtuales y
Proveedores de Servicios de Activos Virtuales (PSAV) a través de aprendizaje
de máquinas no supervisado

OSCAR HUMBERTO MORATTO SANTOS

Concurso de Monografías

Comité Latinoamericano para la Prevención del Lavado de Activos y el
Financiamiento del Terrorismo (COPLAFT)

Federación Latinoamericana de Bancos – FELABAN

2021

Sobre el autor

OSCAR HUMBERTO MORATTO SANTOS

Ingeniero civil y Especialista en Finanzas de la Universidad de los Andes, Colombia. Exsubdirector de Análisis de Operaciones de la Unidad de Información y Análisis Financiero – UIAF. Con experiencia en gestión de riesgos, en particular, del riesgo de lavado de activos, financiación al terrorismo y financiamiento de la proliferación de armas de destrucción masiva. Experiencia en diseño e implementación de modelos de aprendizaje de máquina (*machine learning*), primordialmente enfocados al cumplimiento normativo de la gestión del riesgo LAFTFPADM. Actualmente, director general de la firma Beyond Risk.

RESUMEN

El auge del uso de activos virtuales, incluyendo las criptomonedas, y su rol cada vez mayor en el ecosistema financiero, ha despertado el interés, no solo de inversionistas e instituciones financieras, sino también, de sus implicaciones en la estabilidad del sistema financiero (Auer, R., et al, 2022). Un aspecto de importancia fundamental que puede afectar dicha estabilidad es el relacionado con la gestión del riesgo de LAFTFPADM. En este sentido, contar con mecanismos que permitan la implementación de un enfoque basado en riesgos (EBR) con el objetivo de implementar medidas de mitigación acordes al nivel de exposición requiere la evaluación de diversas variables. Una de estas corresponde al factor de riesgo jurisdicción. En este sentido, este documento presenta estrategias para la implementación de técnicas de aprendizaje de máquina no supervisadas, como subconjunto de la inteligencia artificial, con el objetivo de identificar jurisdicciones potencialmente candidatas a una integración en medidas de gestión de riesgo LAFTFPADM. De esta forma, a través del uso de fuentes de información objetiva, puede reducirse el sesgo en el que se puede incurrir al implementar otras técnicas de evaluación basadas en juicio de experto.

Palabras clave: aprendizaje de máquinas no supervisado, análisis de cluster, gestión de riesgo LAFTFPADM, PSAV, AV.

Indice

Contenido

Introducción.....	1
Marco Teórico	2
Activos Virtuales y PSAV	2
Gestión de riesgo LAFTFPADM de AV y PSAV.....	6
La relevancia del factor de riesgo jurisdicción en la evaluación LAFTFPADM de AV y PSAV.....	24
Inteligencia artificial y Aprendizaje de máquinas	27
Uso de aprendizaje de máquinas en LAFTFPADM.....	29
Metodología	34
Comprender el negocio.....	35
Comprender los datos	36
Preparar los datos.....	44
Modelar	45
Evaluar	51
Implementación	52
Resultados y conclusiones	53
Bibliografía.....	57
Anexo 1 – Concentraciones de segmentos.....	64

LISTADO DE FIGURAS

Figura 1 – Ecosistema de PSAV. Fuente Cyphertrace	3
Figura 2 – Ciclo de vida típico de un AV. Fuente (CCAF, 2020)	4
Figura 3 - Proporción de Bitcoin relacionada con criminalidad. Fuente: Elliptic, 2020....	8
Figura 4 - origen y destino de fondos entrando y saliendo de darknet markets (DNM). Fuente: chainalysis 2020).....	9
Figura 5 – Concentración de lavado de activos con criptomonedas. Fuente: Chainalysis 2022).....	9
Figura 6 Niveles de riesgo de actores en el ecosistema cripto. Fuente: Chainalysis, 2020.....	10
Figura 7 – Porcentaje de transacciones ilícitas 2017 – 2020. Fuente: Chainalysis, 2022.....	13
Figura 8 – Esquema de lavado de dinero PlusToken. Fuente: Chainalysis, 2020).....	23
Figura 9 – Estructura transaccional. Sitio de abuso infantil. Fuente: Chainalysis, 2020	24
Figura 10 – Relación entre Inteligencia Artificial (AI) y Aprendizaje de Máquinas (ML). Fuente: Braincube	27

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Figura 11 Categorías de ML Fuente: Rainerwatl, 2021	29
Figura 12 – principales retos en el desarrollo e implementación de nuevas tecnologías. Fuente, FATF, 2021a.....	34
Figura 13 - Modelo CRISP-DM. Adaptado de Wirth y Hipp, 2020)	35
Figura 14 – Distribución de jurisdicciones informe Estrategia Internacional de Control de Narcóticos. Departamento de Estado de Estados Unidos	36
Figura 15 – Distribución de jurisdicciones. Lista Gris y Lista Negra de GAFI	37
Figura 16 – Distribución de jurisdicciones. Programa de sanciones OFAC.....	37
Figura 17 – Histograma de frecuencias de jurisdicciones. GPI	38
Figura 18 – Histograma de frecuencias de jurisdicciones. GTI.....	38
Figura 19 – Histograma de frecuencias de jurisdicciones. Indice de Gobernanza de Basilea.....	39
Figura 20 – Histograma de frecuencias de jurisdicciones. Indice de Percepción de Corrupción	39
Figura 21 – Histograma de frecuencias de jurisdicciones. Indice de Percepción	40
Figura 22 – Histograma de frecuencias de jurisdicciones. Indice de Adopción de criptomonedas.....	41
Figura 23 – Histograma de frecuencias de jurisdicciones. Indice de Adopción de DeFi	41
Figura 24 – Histograma de frecuencias de jurisdicciones. Nivel de KYC de PSAV	42
Figura 25 – Histograma de frecuencias de jurisdicciones. Número de ATM de cripto..	42
Figura 26 – Histograma de frecuencias de jurisdicciones. Resilencia.....	43
Figura 27 – Histograma de frecuencias de jurisdicciones. Legalidad de criptoactivos .	43
Figura 28 – Matriz de correlaciones	46
Figura 29 – Componentes principales	47
Figura 30 – Reducción de dimensiones. t-SNE.....	48
Figura 31 – Reducción de dimensiones. UMAP	48
Figura 32 – Criterio del codo	49
Figura 33 – Silueta	51

LISTADO DE TABLAS

Tabla 1 - Factores de riesgo asociados a AV y PSAV. Fuente: FATF, 2021B.....	16
Tabla 2 – Ejemplos de requerimientos normativos ALA/CFT que pueden ser abordados por estrategias de clusterización.....	31
Tabla 3 – Métodos de imputación de atributos	44
Tabla 4 – Varianza de atributos	46
Tabla 5 – Concentraciones por segmento.....	52

Introducción

El Bitcoin y otros AV, primordialmente, criptodivisasⁱ, fueron, inicialmente creados, como un mecanismo de substitución del dinero fiat y la banca comercial, por medio de la creación de un mercado descentralizado resistente al control y manipulación de instituciones existentes (Auler, 2022). Sin embargo, desde hace algunos años ha venido creciendo el interés de diversos actores en el mercado de AV. Incluso, algunos bancos centrales han venido desarrollando las denominadas divisas digitales de banco central – CBDC (*Central Bank Digital Currency*)ⁱⁱ y diversos bancos centrales han expresado preocupaciones relativas al impacto potencial de las denominadas monedas estables o *stablecoins*ⁱⁱⁱ en la estabilidad financiera y monetaria (Koss, Mattei, 2022).

Es así que el uso de nuevas tecnologías, incluyendo los Activos Virtuales (AV), generan importantes desafíos en el proceso de gestión de riesgos LAFTFPADM. En primera instancia, los países deben identificar los riesgos LAFTFPADM que los AV y los proveedores de servicios de activos virtuales (PSAV) generan, con el objetivo de implementar un Enfoque Basado en Riesgo (EBR) con y de esta forma que, las medidas diseñadas para mitigar los riesgos identificados sean proporcionales al nivel de exposición (GAFILAT, 2021b). Como segunda medida, y en función del marco regulatorio existente, aquellos agentes económicos, en cuyo apetito de riesgo está ya sea el operar con AV y/o establecer relaciones comerciales con contrapartes que operen con AV o con PSAV, deberán, en aplicación de un EBR, desarrollar mecanismos que permitan establecer el nivel de riesgo de la contraparte, llámese AV o PSAV, y de esta forma, establecer mitigantes particulares.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

A través de un modelo de aprendizaje de máquina se agrupan jurisdicciones incorporando información de diversas fuentes, relacionada con LAFTFPADM, riesgos de la industria cripto, legalidad de la industria por jurisdicción y fortaleza institucional.

De esta forma, este documento pretende cubrir dos (2) de los tres (3) temas propuestos en el año 2022 por parte de Felaban en el concurso de monografías.

Estos son: La tecnología en los programas de cumplimiento, en este caso una aplicación de machine learning (subconjunto de la Inteligencia Artificial) y la gestión del riesgo de lavado de activos y financiamiento del terrorismo de los criptoactivos.

Marco Teórico

En esta sección se presentan los principales componentes teóricos relacionados con la gestión de riesgo LAFTFPADM, en particular, relacionados con la gestión de riesgo de jurisdicción de los AV y de los PSAV.

Activos Virtuales y PSAV

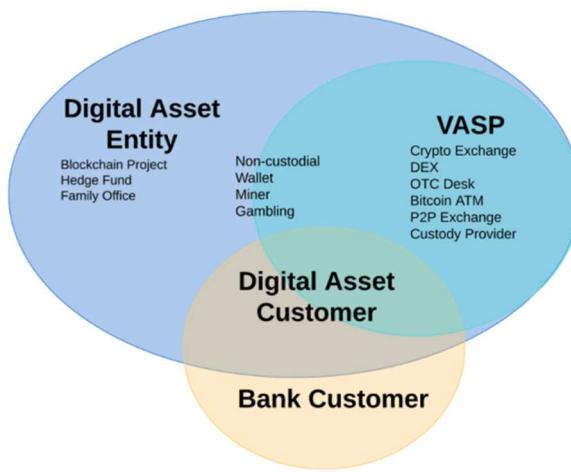
Un AV se define como “una representación digital de valor que se puede comercializar o transferir digitalmente y se puede utilizar para pagos o inversiones. Los activos virtuales no incluyen representaciones digitales de moneda fiat, valores y otros activos financieros que ya están cubiertos en otras partes de las Recomendaciones del GAFI” (GAFILAT, 2021b)

Por su parte, un PSAV es “cualquier persona física o jurídica que no esté cubierta en ningún otro lugar en virtud de las Recomendaciones y que, como negocio, realiza una o más de las siguientes actividades u operaciones para o en nombre

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

de otra persona física o jurídica: i. intercambio entre activos virtuales y monedas fiat; ii. intercambio entre una o más formas de activos virtuales; iii. transferencia de activos virtuales; iv. custodia y / o administración de activos virtuales o instrumentos que permitan el control sobre activos virtuales; y v. participación y provisión de servicios financieros relacionados con la oferta de un emisor y / o venta de un activo virtual" (GAFILAT, 2021 b). Ahora bien, el ecosistema de AV y PSAV es diverso, e incluye actores tales como los *exchanges*, los ATM de cripto, los custodios de billeteras y los fondos de riesgo (*hedge funds*) de cripto. Sin embargo, dada la definición de GAIFI relativa a PSAV, algunos participantes pueden ser o no ser denominados un PSAV en función de las actividades específicas que desarrollen. Por ejemplo, los mineros, por desarrollar la sola actividad de minar criptomonedas no los convierte en PSAV, sin embargo, un pool de mineros, si llegasen a custodiar los AV producto de la actividad de minería, pueden llegar a ser denominados un PSAV. De forma similar, un exchange descentralizado (DEX), es considerado un PSAV si facilita o realiza la transferencia o intercambio de valor, ya sea en AV o en moneda fiat. En la siguiente figura se presenta un panorama general del ecosistema de PSAV.

Figura 1 – Ecosistema de PSAV. Fuente Cyphertrace

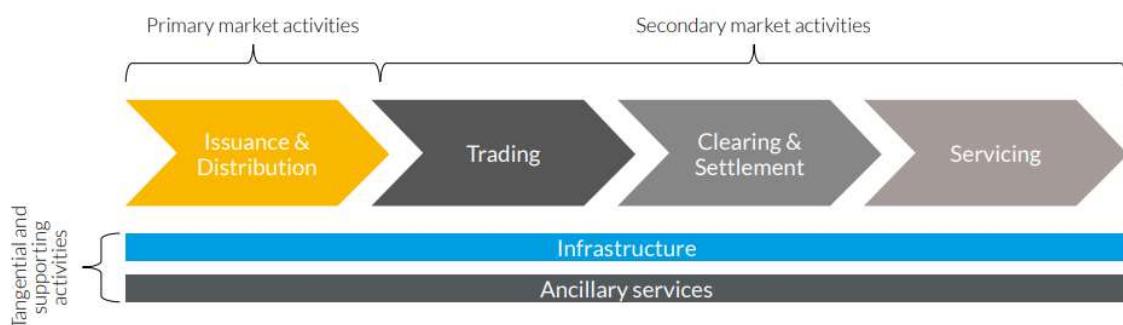


CipherTrace Cryptocurrency Intelligence

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

En este sentido, el ecosistema de AV incluye el desarrollo de diversas actividades que pueden ser desarrolladas con un AV (Choelo, et al, 2021). Estas actividades pueden ser mapeadas al ciclo de vida del AV, de forma tal que, se pueden asociar tres (3) actividades primordiales a este ciclo de vida (CCAF, 2020), las cuales se presentan gráficamente en la siguiente figura.

Figura 2 – Ciclo de vida típico de un AV. Fuente (CCAF, 2020)



Según la figura anterior, las fases del ciclo de vida de un AV en las cuales se puede clasificar un PSAV son (CCAF, 2020):

- Actividades del mercado primario. Se relacionan con la emisión y distribución del AV, entre las cuales se encuentran el denominado *underwriting* (vinculación del emisor e inversionistas, estructuración del negocio, evaluación de riesgos, mitigación de riesgos), emisión (definición de la naturaleza del AV, sostenibilidad de la emisión y emisión al mercado) registro del activo, depósito (almacenamiento seguro del AV), distribución del activo a los participantes del mercado, registro (mantenimiento de registros actualizados de los tenedores del activo), entre otras.
- Actividades del mercado secundario. Incluyen la negociación (trading) que embarca aspectos tales como la admisión del AV para negociación, descubrimiento de precio, punteo de órdenes, brokerage (alineamiento de órdenes entre compradores y vendedores con base en una comisión),

servicios de OTC (facilitar operaciones fuera dela plataforma formal de trading), market making (proveer liquidez al mercado), transferencia de propiedad del AV; compensación y liquidación y; servicios, tales como custodia (salvaguarda de AV), gestión de activos (gestionar portafolios a nombre de inversionistas), servicios de activos (servicios administrativos y operacionales tales como pago de intereses y dividendos, etc).

- Actividades tangenciales: encaminadas al soporte y aseguramiento que las actividades primarias y secundarias puedan ser conducidas de forma eficiente. Algunos ejemplos son los servicios de infraestructura, tales como operación de la red (gobierno y operación de la plataforma), gatekeeping (administración del acceso a la plataforma), desarrollo de plataforma, soporte técnico y mantenimiento, servicios de información (proveer información exógena y datos a la plataforma), desarrollo de aplicaciones sobre la plataforma, e integración (conexión de la plataforma con sistemas existentes) y; actividades auxiliares, tales como servicios bancarios (aceptar depósitos y emitir crédito a clientes basados en AV), proveer data y analítica del mercado, servicios LAFT (verificación de identidad de clientes/emisores, verificar fuente de fondos, identificación de PEP, screening), supervisión del mercado para prevenir, detectar e investigar prácticas abusivas en el mercado, servicios legales (estructuración de negocios, análisis regulatorio, cumplimiento normativo, etc.), ciberseguridad, asesoría, promoción de inversiones, calificación de riesgos basados en una metodología estándar (en el mundo tradicional, Moody's, S&P, Fitch; en el espacio cripto CryptoCompare, ICORatings), servicios de contabilidad, auditoría y aseguramiento. En el ecosistema de

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

AV, algunas actividades novedosas desde la perspectiva de infraestructura son la validación de transacciones (full nodes), el procesamiento de transacciones (confirmación de transacciones – consensus nodes), desarrollo de contratos inteligentes, tokenización de activos y servicios clave de administración (proveer una plataforma segura para la emisión, almacenamiento, transferencia y revocación de llaves criptográficas)

Entre los PSAV, el más conocido y uno de los de mayor relevancia es el denominado *exchange*. Según GAFI, un *Exchange* es una persona o entidad que, por una comisión, desarrolla dentro de sus actividades comerciales el intercambio de AV por moneda fiat, otros tipos de activos o AV, metales preciosos y viceversa; aceptando una amplia gama de medios de pago, incluyendo efectivo, transferencias, tarjetas de crédito y AV (GAFI, 2014). Debido al rol limitado de las instituciones financieras en la industria cripto, el rol dominante lo han tomado los *exchanges* (Auler, 2022).

Gestión de riesgo LAFTFPADM de AV y PSAV

Desde hace algún tiempo, los AV ha llamado la atención de reguladores y bancos centrales alrededor del mundo. Existe una amplia gama de preocupaciones de su adopción de manera general. Solo a manera de ejemplo, en el subconjunto de las denominadas *CBDC*, algunas de las mayores preocupaciones se centran en la facilitación de evasión fiscal, la pérdida de vigilancia de autoridades domóticas debido al uso de CBDC emitidas en el exterior, volatilidad en los mercados Forex y otras preocupaciones, las cuales incluyen riesgos asociados

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

a LAFTFPADM, ciberseguridad, desbalances en los flujos de capitales, paridad con el dólar americano, entre otras (Auler, et al, 2021).

En este sentido, en octubre de 2018, el Grupo de Acción Financiera Internacional (GAFI) actualizó sus recomendaciones, puntualmente, la Recomendación 15 R.15) y adicionó los conceptos de Activo Virtual (AV) y Proveedor de Servicios de Activos Virtuales (PSAV) a su glosario. Posteriormente, la nota interpretativa de la R.15 fue actualizada. En resumen, la R.15 y su nota interpretativa requieren que los PSAV sean objetos de regulación en términos de LAFTFPADM (incluyendo medidas coercitivas), cuenten con licencia o registro, cumplan con la denominada *Travel Rule*, reporten transacciones después de un umbral y sean monitoreados de manera efectiva, incluyendo a sus beneficiarios finales (BF).

Según las Recomendaciones GAFI, los AV deben ser considerados como “bienes”, “productos”, “fondos”, “fondos u otros activos” u otros activos de “valor equivalente” (GAFILAT, 2021b. R.15, Nota interpretativa).

Cabe anotar que el mercado de PSAV es heterogéneo y por ende, ciertos tipos de PSAV pueden llegar a tener mayor propensión a recibir AV de fuentes ilícitas, haciéndolos más riesgosos desde la perspectiva de LAFTFPAM. Se estima que, aunque billones de dólares fueron intercambiados en AV entre entidades ilícitas, el monto del lavado de activos a través de AV rondó los USD 8 billones, un monto relativamente inferior al presentado en el año 2019, el cual ha sido el máximo en términos de LA de fondos a través de AV, totalizando los USD 11 billones (Chanalys, 2022). A pesar de lo anterior, la proporción de fondos ilícitos relacionados con AV vinculados con actividad criminal, a medida que se adopta de forma masiva los AV, se ve cada vez menor. En el caso de Bitcoin, esta

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

influencia de transacciones ilícitas ha venido disminuyendo de forma importante en los últimos años.

Figura 3 - Proporción de Bitcoin relacionada con criminalidad. Fuente: Elliptic,

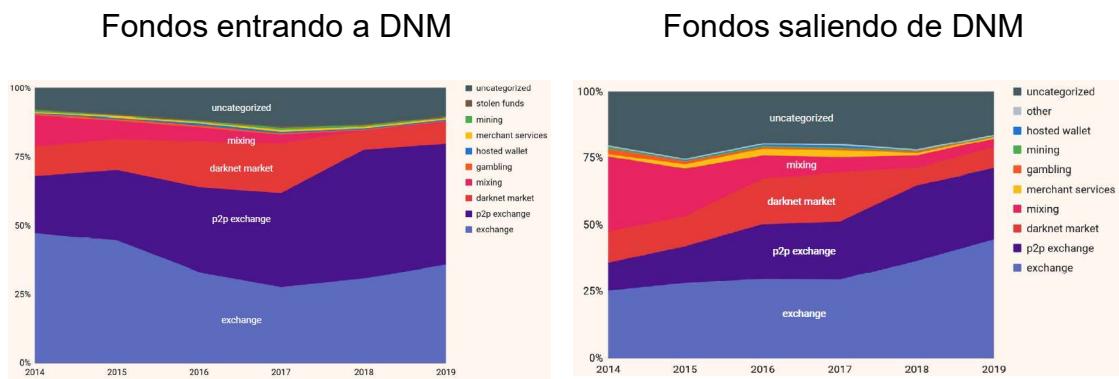
2020



En general, los servicios de mezclado de AV y los servicios de juego en línea pueden llegar a tener las mayores vulnerabilidades de exposición a LAFTFPAM (Fanusie y Robinso, 2018). Sin embargo, el Exchange sigue siendo un punto crucial en los métodos de LAFTFPADM, dado que es el mecanismo a través del cual se realiza el intercambio cripto-fiat. Para dar un ejemplo, en le caso de darknet markets, en el 2019, cerca del 80% de los fondos enviados a este tipo de sitios provenían de exchanges (combinando exchanges tradicionales y P2P) y los exchanges combinados representaron también el mayor mecanismo de cash-out de fondos ilícitos.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Figura 4 - origen y destino de fondos entrando y saliendo de darknet markets (DNM). Fuente: chainalasys 2020)



Estas actividades de LA usualmente se encuentran concentradas en un número reducido de direcciones, dado que en el año 2021, el 55% de las criptomonedas enviadas de direcciones ilícitas fueron a solo 270 direcciones de servicios de diversa índole, incluyendo PSAV (Chainalysis, 2022).

Figura 5 – Concentración de lavado de activos con criptomonedas. Fuente: Chainalysis 2022)



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Dada la multiplicidad de actores en el ecosistema de AV y las distintas actividades que diversos PSAV pueden desarrollar, estos pueden llegar a tener diversos niveles de riesgo intrínseco. A manera ilustrativa, se presentan los niveles de riesgo típicos, teniendo en cuenta un comportamiento natural y cumpliendo con estándares de gestión de riesgo acordes al tipo de productos y servicios ofrecidos por el PSAV, en términos del interés de un actor ilícito en utilizar dichos servicios. Es así por ejemplo, que un pool de mineros de cripto, si su actividad es exclusivamente la minería, dado que obtienen los AV a partir de los procesos de minado, con excepción de los mineros que obtienen los AV a través de cryptojacking o aquellos que el origen de recursos para la adquisición del hardware requerido para el proceso de minado sea de procedencia ilícita, este tipo de actores tendría un nivel de exposición bajo a operaciones LAFTFPADM.

Figura 6 Niveles de riesgo de actores en el ecosistema cripto. Fuente:
Chainalysis, 2020

Low	Medium	High	Severe
Hosted wallets	Exchanges	Mixers	Ransomware
Merchant services	ICOs	Gambling	Darknet markets
Mining pools	Cryptocurrency ATMs		Hacks / stolen funds
			Scams
			Terrorist financing
			Child abuse material
			Sanctions

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Es importante aclarar que la exposición de las instituciones financieras a criptomonedas se ha mantenido limitada (Auler, 2022), aunque, algunas instituciones evalúan un rol más activo por medio del ofrecimiento de servicios relacionados a criptomonedas en el futuro (Bloomberg, 2021).

A pesar de lo anterior, los PSAV, en particular, los *exchanges*, han permanecido ampliamente desregulados, estableciendo de esta forma un sistema financiero cripto en las sombras. En comparación con los exchanges regulados para activos financieros regulados, “la vigilancia de supervisión y regulatoria de los exchanges cripto – que incluye la protección del consumidor, integridad del mercado, trading, divulgación de información, prevención de LAFT – permanece irregular, en el mejor de los casos” (Auler, 2022, Choelo et al, 2021). En términos de supervisión de PSAV, el nivel de desarrollo es heterogéneo, variando desde la inexistencia, pasando por otras jurisdicciones que están aún en etapas de desarrollo de un marco regulatorio para esta industria hasta llegar a un pequeño grupo de países que realizan una supervisión más activa, tal como el llevar a cabo evaluaciones de los PSAV así como desarrollar procesos sancionatorios a los mismos (Choelo et al, 2021).

En este sentido, se argumenta que los intermediarios de criptomonedas, incluyendo los *exchanges*, deben ser sujetos al mismo tipo de regulación y vigilancia que otros intermediarios de otros tipos de activos económicamente equivalentes, incluyendo, desde la perspectiva de estabilidad financiera, protección del consumidor y del inversionista, estándares de ALA/CFT (conocimiento de clientes, reportes a la Unidad de Inteligencia Financiera (UIF), monitoreo transaccional, entre otras) (Auler, 2022; FATF, 2021B; Choelo et all, 2021). Es más, cada vez se observa que un reforzamiento de las reglas de

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado cumplimiento en materia de prevención de LAFTFPADM, particularmente en exchanges, debido entre otras razones, a la presión por parte de instituciones financieras que desean realizar integraciones con estos exchanges (Chainalysis, 2021a).

Finalmente, la naturaleza inherente rica en datos de los AV ofrece importantes oportunidades de supervisión, en la cual, se apoye de un lado la innovación financiera responsable asegurando una supervisión adecuada, a través del uso de nuevos métodos de supervisión y aplicaciones de suptech, haciendo uso intensivo de datos y herramientas tecnológicas como la analítica de blockchain con el objetivo de mejorar la efectividad de los marcos de supervisión (Choelo et al, 2021).

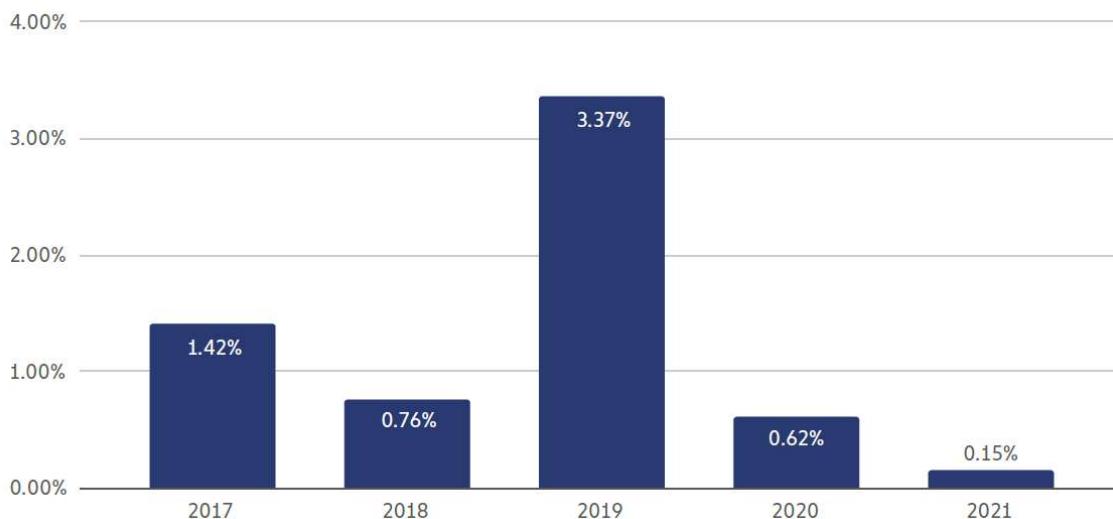
Como cualquier otro activo financiero y sector económico, los AV y los PSAV son vulnerables a ser utilizados como vehículos para la realización de operaciones LAFTFPAM. GAFI considera que los AV y los PSAV pueden llegar a tener un mayor grado de exposición a riesgos LAFTFPADM y por lo tanto, requerir de medidas apropiadas de mitigación (FAFT, 2021). Para el año 2021 se estima que el monto de los criptoactivos asociados a actividades ilícitas rondaron los USD 14 billones, lo que representa un incremento importante respecto al nivel de actividad ilícita del año 2020, el cual fue de 7.8 billones. A pesar de la magnitud de esta cifra, debe ser analizada en contexto. El mercado de criptodivisas ha venido expandiéndose de forma casi exponencial. En el año 2021 alcanzó los USD 1838 trillones, lo que representa un incremento del 567% respecto al monto total transado en el año 2020. Por lo tanto, este aumento en adopción hace que los AV sean más apetecidos por actores criminales. Sin embargo, cada vez la

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

proporción de actividad ilícita con criptoactivos disminuye a medida que su adopción se masifica, llegando en el año 2021, a representar solo el 0.15% del total de transacciones (Chainalysis, 2022).

Figura 7 – Porcentaje de transacciones ilícitas 2017 – 2020. Fuente:

Chainalysis, 2022.



Se han documentado diversas casuísticas y tipologías relacionadas con el uso de AV y/o PSAV en este tipo de situaciones. Entre las características primordiales del uso ilícito de los AV y PSAV se encuentran:

- Uso de AV como medio de pago en mercados de la red oscura – *darknet markets* – (Fanusie y Robinson 2018; Chainalysis, 2020, Chainalysis 2021a, Chainalysis, 2022)
- Uso de PSAV (exchanges, mezcladores, sitios de apuestas en línea, peel chains, entre otros) como puntos de conversión de AV a otros AV o moneda fiat. (Fanusie y Robinson (2018); US Department of Treasury, 2021, Chainalysis, 2020, Chainalysis, 2021a, Chainalysis, 2022, Elliptic, 2020)

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

- Uso de exchanges sin licencia o en jurisdicciones de alto riesgo. (Elliptic, 2020)
- Aprovechamiento de asimetrías regulatorias. La regulación ALA/CFT así como otras medidas, como por ejemplo, los controles de capitales, tales como la facilidad para movilizar divisas fuera de una jurisdicción (p.e. en China), influyen en el atractivo de dicha jurisdicción a ser utilizada en una operación de LAFTFPAM (Fanusie y Robinson, 2018; XX)
- Uso de AV y PSAV como mecanismos para burlar sanciones financieras (Popper et al, 2018, Chainalysis, 2020, Chainalysis, 2021a, Chainalysis, 2022, CipherTrace, 2021)
- La parición de nuevos AV de mayor opacidad y que permiten ocultar de mejor manera la actividad de los usuarios (Foley et al, 2019), dando mayor nivel de anonimato a las transacciones, lo cual, puede ser llamativo para organizaciones de crimen organizado (Elliptic, 2020).
- Uso de AV como medio de intercambio en sitios de abuso sexual de menores (Chainalysis, 2020).
- Uso de AV en estafas y esquemas Ponzi (Chainalysis, 2020, Chainalysis, 2021a, Chainalysis, 2022, CipherTrace, 2021). Esta categoría de actividades ilícitas incluye algunas tipologías novedosas, como las denominadas *rug pulls*, en las cuales, los cibercriminales desarrollan un proyecto aparentemente legítimo utilizando AV, solo para recibir fondos de inversionistas y desaparecer.
- Uso de AV como vehículo para lavar activos provenientes de crímenes cibernéticos, en especial, ransomware (Chainalysis, 2021a, Chainalysis, 2022, CipherTrace, 2021)

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

- Financiación al terrorismo por medio de AV (Chainalysis, 2021a, Chainalysis, 2022, Elliptic, 2020, CipherTrace, 2021)
- Surgimiento de las finanzas descentralizadas (DeFi) dentro del escenario de actividades ilícitas por medio de AV (Chainalysis, 2021a, Chainalysis, 2022, CipherTrace, 2021)
- Atomización de los mercados de darknet. Esto debido a varios factores, uno, las acciones por parte de autoridades judiciales en el desmantelamiento de este tipo de sitios y de otro, la creciente competencia entre estos actores, llegando incluso a desplegar ataques ciberneticos, principalmente, DoS, a competidores. (Chainalysis, 2021a)
- Fondos robados. En esta categoría se encuentran tanto AV que son robados de wallets de usuarios, o el hackeo a exchanges con wallets hosteadas, hasta el robo de tokens de plataformas DeFi (Chainalysis, 2022, Elliptic, 2020, CipherTrace, 2021).
- Uso de NFTs (Tokens no fungibles) en actividades de LAFTFPADM. Concentrados en manipulación del mercado por medio de incremento artificial del precio del NFT o el LA a través de la adquisición de NFTs (Chainalysis, 2022)
- Uso de exchanges descentralizados (DeX). (Elliptic, 2020)
- Intercambio de AV por divisas de juegos en línea. (Elliptic, 2020)
- Uso de tarjetas prepagadas asociadas a billeteras de AV para estratificar operaciones (Elliptic, 2020)

Es importante resaltar que, con la masificación del uso de los AV, la proporción de actividad ilícita de éstos han venido disminuyendo (Foley et al, 2019).

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Ahora bien, las principales preocupaciones relacionadas con LAFTFPADM asociadas los AV se centran en i. la adopción masiva de los AV y; ii adopción masiva de transferencias persona a persona (P2P^{iv}) sin necesidad de un intermediario regulado, lo cual puede ser llamativo para actores con vínculos con actividades ilícitas. Otras preocupaciones de GAIFI en relación al uso de AV son la capacidad de movilizar fondos con alcance global de manera casi instantánea e irreversible; así como facilitar una amplia gama de actividades financieras (transferencia de valor, actividades relacionadas con servicios, commodities, derivados y valores); el que los AV facilitan el anonimato o pseudoanonimato de transacciones, inhabilitando la posibilidad del PSAV de identificar al beneficiario final (BF); la falta de identificación de clientes y contrapartes; la regulación inexistente o desbalanceada entre jurisdicciones (FATF, 2021B);

Los principales factores de riesgo relacionados a AV y PSAV se presentan en la siguiente tabla (FATF, 2021B):

Tabla 1 - Factores de riesgo asociados a AV y PSAV. Fuente: FATF, 2021B

Factor	Aspectos a considerar
AV	Número, valor y volatilidad de las transacciones del AV. Capitalización de mercado del AV Valor en circulación Número de jurisdicción que utilizan el AV y número de usuarios por jurisdicción Cuota de mercado del AV en la jurisdicción

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Factor	Aspectos a considerar
	<p>Uso del AV como mecanismo de transferencia de valor transfronteriza</p> <p>El riesgo potencial de LAFTFPAM del AV que es intercambiado con o por moneda fiat o por otro AV y la extensión en las cuales transacciones/canales/plataformas basadas en AV interactúan con, o están más conectadas con transacciones/canales/plataformas basadas en fiat.</p>
	<p>La naturaleza y el alcance de los canales de pago del AV</p> <p>El número y valor de las transferencias realizadas con el AV y aquellas relacionadas con actividad ilegal entre i. PSAVs y sujetos obligados; ii. entre PSAVs y sujetos obligados y no obligados; iii. sujetos no obligados (p.e. P2P)</p>
	<p>El uso de técnicas de anonimización en las operaciones de AV (p.e. mixers, billeteras privadas, etc) y técnicas de desanonimato (p.e evaluación de riesgos de billeteras por medio de herramientas de análisis de blockchain)</p>
	<p>Uso de técnicas de anonimato de conexión a internet (p.e uso de ToR, o de I2P, dado que pueden ofuscar transacciones inhabilitando la capacidad del PSAV de identificar clientes y así, implementar medidas ALA/CFT eficaces.</p>
	<p>El tamaño del negocio base de clientes y la importancia de actividades transfronterizas del emisor</p>
PSAV	El número y tipos de PSAV ubicados en la jurisdicción.

Factor	Aspectos a considerar
	<p>El número y volumen de transacciones canalizadas por cada PSAV.</p>
	<p>La sofisticación del programa de prevención del riesgo LAFTFPADM del PSAV.</p> <p>El que el PSAV posea herramientas para evaluar y monitorear los AV con los que opera.</p>
	<p>El tamaño y tipos de la base de clientes del PSAV.</p> <p>El acceso a información de clientes que tenga el PSAVA.</p>
	<p>La naturaleza y características de los productos basados en AV que el PSAV ofrece.</p>
	<p>Las medidas adoptadas por el PSAV con el objetivo de reducir su nivel de exposición al riesgo LAFTFPADM</p>
	<p>El modo de operación y relacionamiento con clientes del PSAV (totalmente online o en persona – p.e ATM de cripto)</p>
	<p>Los vínculos del PSAV con diversas jurisdicciones y su potencial riesgo a violación de sanciones o a realizar operaciones LAFTFPAM</p>
	<p>Si el PSAV implementa el <i>travel rule</i></p>
	<p>Transacciones desde / hacia sujetos no obligados (p.e. con PSAV que se encuentren en jurisdicciones en las cuales no son objeto de regulación o supervisión),</p>
	<p>Los tipos específicos de AV que el PSAV ofrece (p.e AV de alto anonimato).</p> <p>El uso de métodos de ofuscación de transacciones</p>

En lo que respecta a señales de alerta, GAFI las ha clasificado en seis (6) categorías principales (FATF, 2020):

- Señales de alerta relacionadas con las transacciones.
 - Tamaño y frecuencia de las transacciones. Fraccionamiento de transacciones; ejecución de múltiples transacciones de alto monto en cortos periodos de tiempo o sucesivas, y/o por cuentas inactivas o de reciente apertura; transferencia de AV a múltiples PSAV, en particular, aquellos ubicados en jurisdicciones con regulación deficiente; depositar AV en un Exchange y posteriormente transferir el valor a uno o varios AV sin razón aparente o transferirlo a una billetera privada; aceptar fondos aparentemente robados o fraudulentos (estafas, hackeos, etc.)
- Señales de alerta relacionadas con los patrones de transacciones
 - Transacciones de nuevos usuarios. Depósitos iniciales de alto monto; transferencias de AV de alto monto.
 - Transacciones relacionadas a todos los usuarios. Uso de múltiples AV o cuentas sin razón aparente; transacciones frecuentes por más de una persona, desde la misma dirección IP por más de una persona o altos montos; transacciones provenientes de diversas billeteras no relacionadas en cantidades pequeñas y su posterior transferencia a otra billetera o intercambio a fiat; realizar conversiones VA-fiat con pérdidas financieras; conversión fita-VA o VA-VA sin explicación lógica.
- Señales de alerta relacionadas con el Anonimato

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

- Transacciones por un mismo usuario usando múltiples VA, particularmente, aquellos de alto anonimato, sin tener en cuenta los costos transaccionales; clientes que operan como un PSAV sin licencia y que cobran altos fees de transacción; conversión de un AV que opera en una blockchain pública (p.e Bitcoin) a un AV de alta opacidad (p.e ZCash); volúmenes altos de cash-out; transferencia desde o hacia billeteras que presentan patrones de uso de mecanismos de ofuscación (p.e mixers); fondos relacionados con billeteras con vínculos con fuentes sospechosas (p.e. darkweb marketplaces); uso de billeteras no hosteadas o en papel; usuarios con IP relacionadas con mercados oscuros o que proveen anonimato (ToR, VPNs, etc); alto número de billeteras controladas desde la misma dirección IP; uso de AV potencialmente relacionados con esquemas Ponzi o estafas; recepción de fondos desde PSAV con falencias evidentes en los procesos de debida diligencia y conocimiento del cliente.
- Señales de alerta respecto a remitentes y receptores.
 - Irregularidades en la creación de cuentas. Creación de cuentas bajo diferentes nombres de usuario para obviar controles; transacciones de direcciones IP ubicadas en jurisdicciones sancionadas o direcciones IP previamente identificadas como sospechosa; registro de dominios web en ubicaciones diferentes a la de operación del merchant o usuarios corporativos.
 - Irregularidades en el proceso de conocimiento del cliente. Información incompleta o insuficiente o renuencia a proporcionar

información; receptor o remisor sin conocimiento de la causa de la transacción, la fuente de ingresos o relación con la contraparte; documentación falsa.

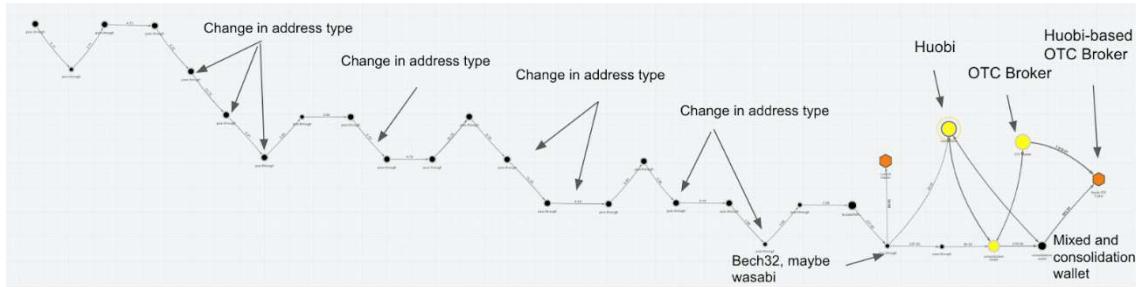
- Perfil. Información compartida entre cuentas (p.e., direcciones IP); discrepancias entre la IP del perfil del cliente y las IP relacionadas a las transacciones; información pública adversa del cliente.
- Perfil de cuentas mula o víctimas de fraude potenciales. El remitente no está familiarizado con operación de AV; clientes particularmente mayores al perfil promedio de clientes del PSAV; clientes que son personas vulnerables financieramente; compras de AV por parte del cliente que no corresponde con su capacidad financiera.
- Otros comportamientos inusuales. Cambios constantes de la información de identificación del cliente, direcciones IP o información financiera; sus de múltiples direcciones IP para acceder distintos PSAV; realización de transacciones recurrentes con otros usuarios en las cuales consistentemente presenta pérdidas o ganancias significantes.
- Señales de alerta de origen de fondos.
 - Transacciones con AV conectados a fraude, extorsión, ransomware, sanciones, mercados oscuros o sitios web ilegales; transacciones de AV desde o hacia sitios de apuesta en línea; uso de múltiples tarjetas debido o crédito para realizar intercambio cripto-plástico y de esta forma, realizar retiros importantes de fiat; falta de información o transparencia del origen de fondos.

- Señales de alerta de riesgo geográfico
 - Clientes que remiten o reciben fondos en un Exchange no registrado en la jurisdicción, ya sea del cliente o donde el Exchange opera; el cliente utiliza los servicios de un PSAV ubicado en una jurisdicción de alto riesgo debido a las deficiencias en su sistema ALA/CFT; el cliente traslada su operación a jurisdicciones sin regulación de AV.

En términos de medidas de mitigación de riesgo LAFTFPADM, se recomienda, además del registro y/o licenciamiento de PSAV, entre otras las siguientes: la aplicación de las Recomendaciones GAFI de la 10 a la 25 a los PSAV. En relación con los umbrales de reporte sobre los cuales los PSAV deben realizar el proceso de debida diligencia de clientes, se recomienda que sea de USD 1.000, en lugar de los USD 15.000 que aplica para las instituciones financieras (FATF, 2021B). Las reglas de transferencias electrónicas establecidas en la Recomendación 16, aplican a los PSAV y a las transferencias de AV de forma modificada (*travel rule*). El proceso de debida diligencia de clientes, el cual incluye, entre otros aspectos, la identificación de clientes, su BF, la verificación de identidad del cliente basada en información, datos o documentación confiable e independiente, soporte de fondos, mantenimiento de registros, monitoreo, así como el entendimiento del propósito de la relación comercial, debe ayudar al PSAV en la evaluación de riesgos LAFTFPADM relativas a las operaciones con AV, que el PSAV esté en capacidad de remitir ROS a la UIF, que el PSAV realice análisis sobre las direcciones IP de los usuarios y que haga uso de herramientas de análisis del blockchain, entre otras (FATF, 2021B). Respecto a este último

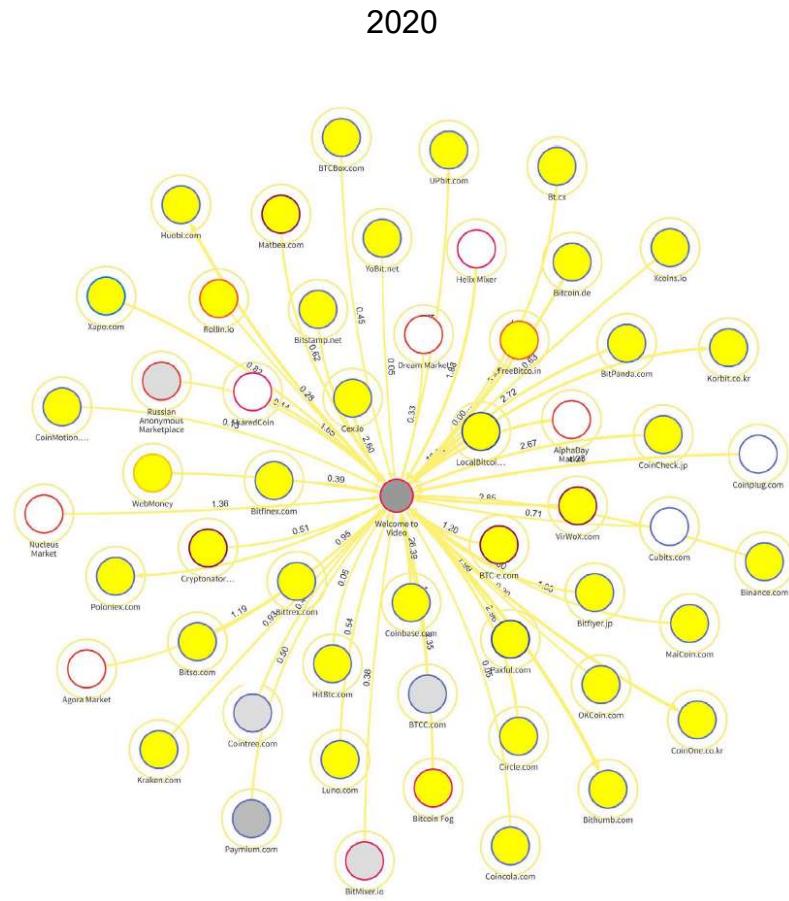
Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado punto, es difícil generalizar el nivel de exposición de un determinado tipo de PSAV sin realizar un análisis de las transacciones que canaliza. Los actores ilegales cada vez utilizan estrategias más sofisticadas con el objetivo de ofuscar operaciones, utilizando diversos actores del ecosistema de AV, lo cual, aumenta la complejidad de las mismas y exige del uso de herramientas de analítica del blockchain para su entendimiento. A manera de ejemplo, se presenta esquemáticamente el mecanismo de lavado de dinero del esquema ponzi PlusToken, el cual involucró no solo el fraccionamiento de transacciones, cambio de direcciones IP, uso de servicios de ofuscación como mixers, empleo de OTC brokers, consolidación de transacciones en billeteras y cashout a través de exchanges..

Figura 8 – Esquema de lavado de dinero PlusToken. Fuente: Chainalysis, 2020)



Cabe resaltar que, al igual que en las tipologías de LAFTFPADM tradicionales, en función del tipo de delito y mecanismo utilizado, la complejidad y tipo de operaciones para lavar recursos ilícitos se modifica. A manera de ejemplo, comprárese la estructuración de operaciones del caso de PlusToken con las transacciones típicas realizadas en una plataforma de venta de contenido de abuso infantil

Figura 9 – Estructura transaccional. Sitio de abuso infantil. Fuente: Chainalysis,



La relevancia del factor de riesgo jurisdicción en la evaluación

LAFTFPADM de AV y PSAV

La jurisdicción juega un papel relevante en la evaluación de riesgos LAFTFPADM tanto de AV como de PSAV. En este sentido, los actores criminales explotan las asimetrías regulatorias de AV y PSAV relativas a prevención de LAFFPADM, al mover los fondos provenientes de actividades ilegales hacia PSAV ubicados en jurisdicciones con regulación mínima o inexistente de LAFTFPADM. En dichas jurisdicciones, por ejemplo, el PSAV no requiere de una licencia o registro, no cuentan con obligación de remitir ROS, no cuentan con reportes transaccionales

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

objetivos luego de un umbral determinado, tienen políticas de conocimiento del cliente deficientes, entre otros aspectos.

Dentro de los aspectos relacionados al factor jurisdicción de importancia para esta evaluación se encuentran:

1. La Recomendación 19 requiere a los países que instruyan a sus sujetos obligados, incluyendo a los PSAV, en aplicar medidas de debida diligencia intensificada a relaciones comerciales y transacciones con personas naturales y jurídicas ubicadas en jurisdicciones de alto riesgo.
2. El alcance global y transfronterizo de las operaciones realizadas con AV, que puede hacer llamativo al AV como instrumento para realizar operaciones LAFTFPADM (FATF, 2021B).
3. El hecho que un PSAV localizado en una jurisdicción, p.e. con controles débiles o inexistentes, pueda ofrecer sus productos y servicios a clientes ubicados en otra jurisdicción, donde éstos sean sujetos a obligaciones ALA/CFT diferenciales (Fanusie y Robinson, 2018; FATF, 2021B).
4. El alcance casi global de los PSAV puede incrementar los riesgos LAFTFPADM asociados a los AV y a los PSAV debido a las brechas potenciales en información de clientes y transacciones (FATF, 2021B).
5. Tener en cuenta aquellos países o áreas geográficas identificadas que proveen fondos o apoyo a organizaciones terroristas o que tienen organizaciones terroristas designadas que operan en su territorio. (FATF, 2021B)
6. Considerar aquellos países identificados como de niveles significantes de crimen organizado, corrupción y otros tipos de actividad criminal, tales

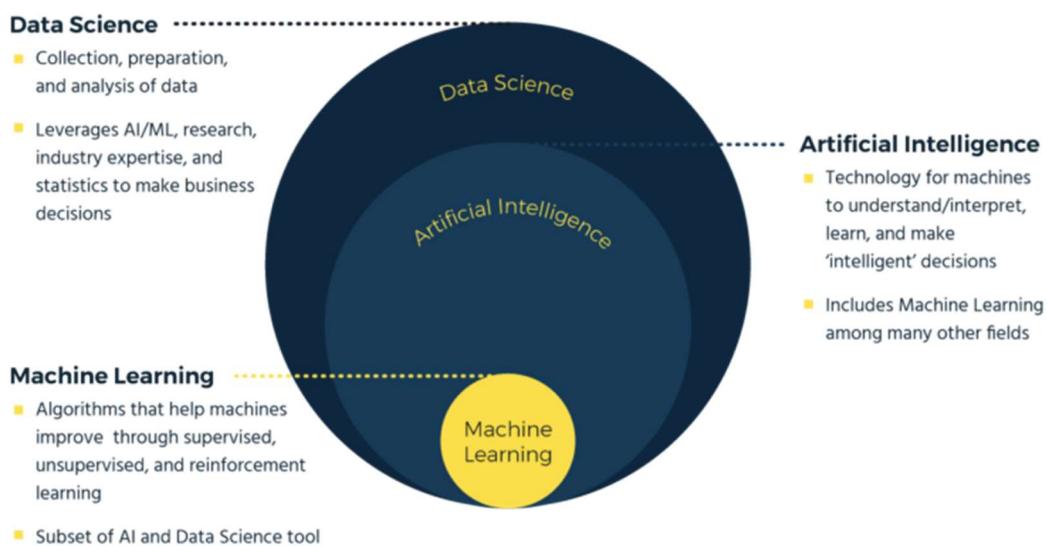
como el ser origen o tránsito de tráfico de estupefacientes, tráfico humano, contrabando y juego ilegal (FATF, 2021B).

7. Países sujetos de sanciones, embargos o medidas similares por parte de organizaciones internacionales, tales como las Naciones Unidas (FATF, 2021B)
8. Países con sistemas de gobernanza, legal, regulatorio y de aplicación de la ley débiles, incluyendo aquellos países identificados por GAFI como con sistemas ALA/CFT débiles, especialmente para PSAV (FATF, 2021B).
9. La ilegalidad / legalidad de las operaciones de AV en la jurisdicción. Algunas jurisdicciones han impuesto prohibiciones a todas (China considera todas las actividades relacionadas con AV como ilegales) o algunos tipos (p.e Bélgica prohíbe el ofrecer productos financieros a clientes retail de los cuales su retorno dependa, directa o indirectamente en un AV o de actividades relacionadas con AV) (Cohelo et al, 2021).
10. El licenciamiento y registro de PSAV varía entre jurisdicciones (Cohelo et al, 2021).
11. El nivel de supervisión. La supervisión de los PSAV se encuentra en etapas iniciales en casi todas las jurisdicciones (Cohelo et al, 2021).
12. El número y tipos de PSAV ubicados en la jurisdicción (FATF, 2021B)
13. Los vínculos del PSAV con diversas jurisdicciones y su potencial riesgo a violación de sanciones o a realizar operaciones LAFTFPAM (FATF, 2021B)

Inteligencia artificial y Aprendizaje de máquinas

La Inteligencia Artificial (IA) es la ciencia de replicar las habilidades del pensamiento humano para realizar tareas que típicamente requieren de inteligencia humana, tales como reconocer patrones, hacer predicciones, recomendaciones o tomar decisiones (FATF, 2021a). El aprendizaje de máquinas – *machine learning* – es un subconjunto de la IA en el cual, se entrena al sistema computacional para aprender de los datos, identificar patrones y hacer predicciones con mínima intervención humana (FATF, 2021a). La siguiente figura relaciona estos conceptos.

Figura 10 – Relación entre Inteligencia Artificial (AI) y Aprendizaje de Máquinas (ML). Fuente: Braincube



El aprendizaje de máquina se subdivide a su vez, típicamente en tres (3) categorías (Rainerwalt, 2021):

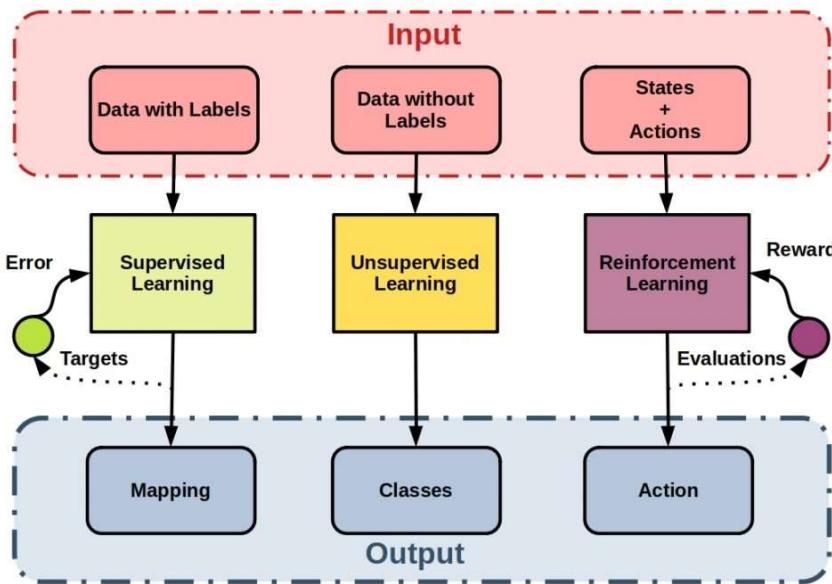
- Aprendizaje supervisado. En el aprendizaje supervisado, el sistema computacional aprende a partir de un set de datos sobre el cual se conoce

la respuesta (labeled). De esta forma, el algoritmo aprende iterativamente las dependencias entre los puntos de datos. El objetivo a ser aprendido es especificado de manera previa y el proceso de aprendizaje es supervisado a través de la comparación de las predicciones contra los resultados reales. Las principales categorías de aprendizaje supervisado son clasificación y regresión.

- Aprendizaje no supervisado. En el aprendizaje no supervisado, los patrones en los datos se determinan sin un set de datos marcados, sino que son extraídos a partir de la estructura propia de los datos. Las dos categorías principales de aprendizaje no supervisado son el clustering y la reducción de dimensiones.
- Aprendizaje por reforzamiento. En este escenario, un programa, o agente, debe de forma independiente, desarrollar una estrategia para desarrollar acciones en un ambiente controlado. Para este propósito, se diseñan refuerzos positivos y negativos, de la interacción del agente con su entorno. El agente debe entonces maximizar las recompensas y minimizar los castigos. El ambiente es entonces un escenario de simulación que el agente debe explorar.

Esquemáticamente las diferencias entre las categorías de aprendizaje supervisado se presentan a continuación

Figura 11 Categorías de ML Fuente: Rainerwatl, 2021



Uso de aprendizaje de máquinas en LAFTFPADM

El aprendizaje de máquinas, junto con el procesamiento de lenguaje natural, han sido identificados como las dos tecnologías de IA con mayores beneficios potenciales de uso en el entorno de prevención de LAFTFPADM, tanto para sujetos obligados como para reguladores (FATF, 2021a). El aprendizaje de máquinas, dada su capacidad de establecer patrones en la fuente de datos de manera automática sin necesidad que se encuentren “marcados” previamente, ofrece las mayores ventajas en los entornos ALA/CFT, al reducir los falsos positivos e identificar relaciones complejas de los datos (FATF, 2021a).

Algunos casos de uso en los cuales el aprendizaje de máquinas es de utilidad en prevención de LAFTFPADM son (FATF, 2021a):

- Identificación y verificación de clientes
- Monitoreo de relaciones comerciales y análisis transaccional y de comportamiento

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

- Identificación e implementación de actualizaciones regulatorias.
- Reportes automáticos de datos.

Algunas técnicas de aprendizaje de máquinas se han utilizado para la detección de fraude y el LA, con tasas de clasificación altas (altos verdaderos positivos y bajos falsos positivos) (UIAF, 2014b). A través de estas técnicas “es posible extraer patrones sobre los comportamientos típicos de los individuos, los cuales permiten caracterizar operaciones extrañas” (UIAF 2014a) o inusuales y de esta forma, apoyar al sujeto obligado en la identificación de ROS o a la UIF en la identificación de casos potencialmente judiciables,

Algunas aplicaciones específicas de aprendizaje de máquinas en el entorno LAFTFPADM son un sistema de clasificación automática de ROS. “En este caso, el aprendizaje de máquina encontrará reglas que le permitan definir qué combinaciones de valores en las variables económicas, financieras, judiciales, etc., de los individuos se asocian con un ROS de alta importancia”, combinándolas con otras reglas tales como el tipo de transacción, los montos, la periodicidad, la jurisdicción y, “a su vez, estas reglas podrán ser programadas para que el sistema continúe autónomamente con la calificación” (Gafilat, 2021a). Otro ejemplo es el uso de técnicas de clustering (agrupación), las cuales buscan generar grupos de datos lo más similares entre si, con diferencias entre los grupos (UIAF, 2014, b); así, para agrupar registros (p.e clientes) en grupos cohesivos basados en sus características, pueden generarse controles que pueden ser más adecuados desde un EBR, tales como umbrales transaccionales diferenciales por grupo, permitiendo un monitoreo personalizado y eficiente de la relación comercial (FATF, 2021 b). Las técnicas de cluster no solo son útiles

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

desde la perspectiva de uso de aprendizaje de máquinas en procesos de prevención y detección de operaciones LAFTFPADM, sino también, desde el punto de vista de cumplimiento normativo. Los procesos de cluster se denominan también como segmentación de forma indistinta (Hastie, et al, 2020). En este sentido, diversas regulaciones en Latinoamérica incluyen requerimientos normativos que pueden ser y deben ser cumplidos con la aplicación de técnicas de clusterización. Algunos ejemplos se presentan en la siguiente tabla.

Tabla 2 – Ejemplos de requerimientos normativos ALA/CFT que pueden ser abordados por estrategias de clusterización

Regulador	Requerimiento normativo
Superintendencia Financiera de Colombia. SARLAFT 4.0	Segmentación de factores de riesgo. Es el proceso por medio del cual se lleva a cabo la separación de elementos en grupos homogéneos al interior de ellos y heterogéneos entre ellos. La separación se fundamenta en el reconocimiento de diferencias significativas en sus características (variables de segmentación).
Superintendencia de Sociedades de Colombia. Política de supervisión de los Programas de Transparencia y Etica Empesarial PTEE	Que la Entidad Obligada cuente con metodologías de segmentación y clasificación de los Factores del Riesgo C/ST.
Superintendencia de Banca y Seguros SBS del Perú. Resolución SBS 2660-2015	Formación de segmentos y análisis de escenarios. Las empresas deben formar segmentos de mercado, estableciendo grupos que guardan una homogeneidad interna, pero una heterogeneidad entre ellos, de acuerdo con una o varias variables.
Superintendencia de Seguros República Dominicana. Norma que regula la prevención del lavado de Activos y financiamiento del terrorismo para el sector seguros	En el proceso de identificación de riesgos, el sujeto obligado debe definir la metodología aplicable para la segmentación de los factores de riesgo. Realizar una segmentación de los factores de riesgo, con base en la metodología definida.

Superintendencia del mercado de valores de Panamá. Guía para la implementación del EBR	El tratamiento estadístico debería permitir al SO demostrar que existe movilidad dentro de los clientes en el nivel de riesgo asignado y que siempre se cumple con la segmentación adecuada de los clientes
--	---

En este orden de ideas, con el objetivo de realizar un “agrupamiento” de jurisdicciones en función del riesgo LAFTFPADM, con el fin de aplicar un EBR, en este trabajo se desarrollara un modelo de cluster. Algunas ventajas del uso de una técnica de cluster sobre otras estrategias de conformación de grupos son:

- No se requiere conocer el número de grupos con antelación. De esta forma, se podrá seleccionar el número de grupos apropiado a las características propias del set de datos.
- Reducción de influencia externa: debido a que las técnicas de aprendizaje no supervisado no requieren de un set de entrenamiento. Al no tener set de entrenamiento, no se cuenta con una muestra de individuos, a los cuales previamente se les ha asignado una categoría o grupo, reduciendo de esta forma un posible sesgo exterior al asignar dichas categorías al set de datos o el que las categorías contempladas no reflejen todos los casos posibles en la estructura de la fuente de datos.
- El agrupamiento de elementos se da por las características propias de los mismos. De esta forma, el agrupamiento se realiza con base en el reconocimiento de diferencias significativas en las características de la base de datos;

Existen diversas técnicas de implementación de cluster (por particiones, densidad, jerárquico, etc.). Se ha seleccionado el algoritmo de K-Medias (MacQueen, 1967, Hartigan, J., & Wong, 1974), Esto debido a que para

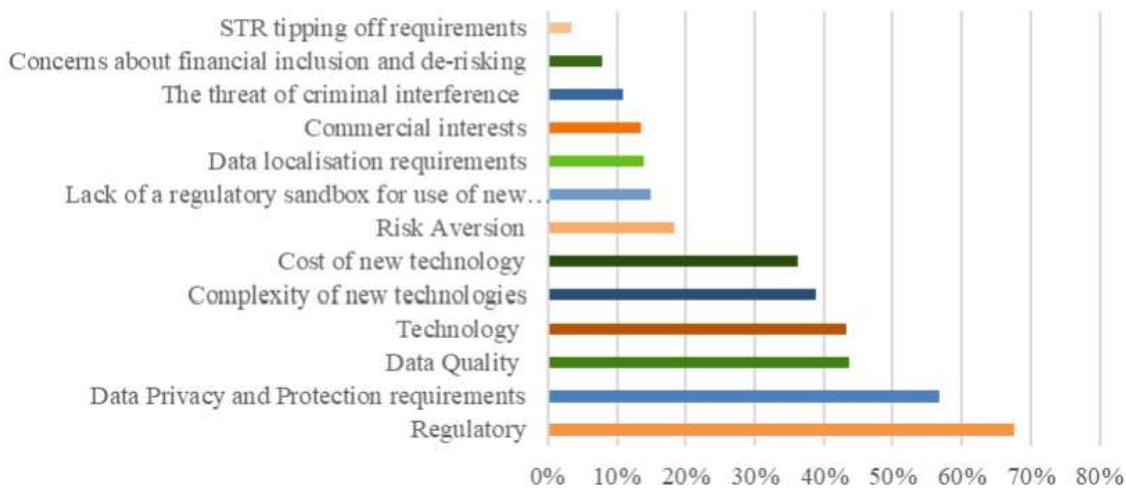
Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

volumetría y número de registros, ésta técnica, además de optimizar los tiempos de procesamiento, genera particiones apropiadas (Han, et. Al). En k-medias, Cada elemento se asigna al centroide que se encuentre más cerca, desde el punto de vista de distancia euclíadiana. Dicha distancia euclíadiana, entre los puntos $x_i, x_{i'}$ denominada $d(x_i, x_{i'})$ se calcula como $d(x_i, x_{i'}) = \|x_i - x_{i'}\|^2$.

- Según instituciones académicas líderes a nivel mundial, tales como la Universidad de Stanford y el MIT, es una de las técnicas de clúster más populares, sino la más popular que existe *Hastie, et al, 200, Ultman et al, 2014, Ng y Piech, Screehar et al, 2017) . Esto hace que las implementaciones de K-Medias en distintas plataformas, incluida la utilizada en este proyecto, sean consistentes entre sí. Adicionalmente, por su amplia adopción, es una técnica confiable. Debido al uso de una distancia euclíadiana como medida de disimilitud, K-medias requiere que todos los atributos sean de tipo numérico, por lo que es necesario realizar al menos una transformación de los atributos categóricos a numéricos.

La implementación de estrategias de innovación tecnológica en la prevención del LAFTFPADM presenta diversos retos, siendo los principales aquellos relacionados con los requisitos regulatorios y operativos. Los retos regulatorios se presentan primordialmente por la falta de apoyo explícito de supervisores a estas iniciativas, así como los retos asociados a la interpretabilidad y explicabilidad de resultados. Los retos operativos se concentran en la adaptación de los procesos existentes a nuevas aplicaciones tecnológicas, incluyendo, interconexión con sistemas legacy y entrenamiento de personal.

Figura 12 – principales retos en el desarrollo e implementación de nuevas tecnologías. Fuente, FATF, 2021a



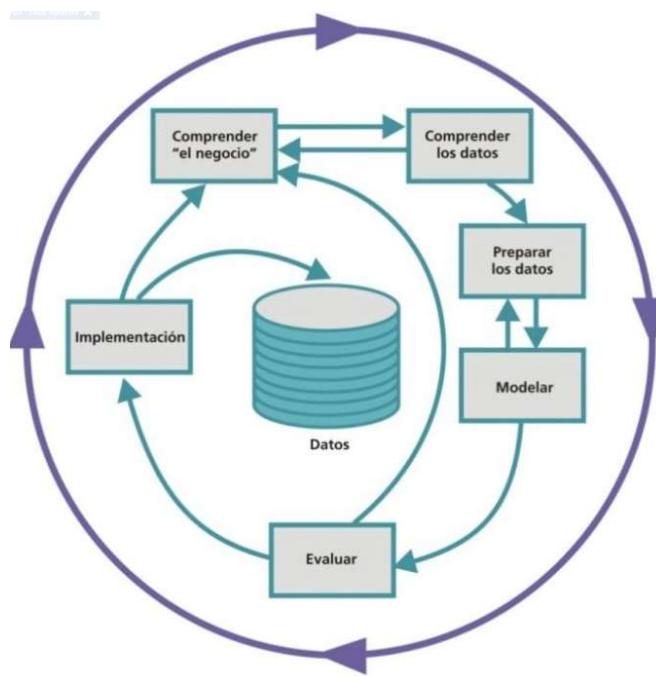
Otra preocupación relevante es aquella asociada a resultados no deseados de la implementación de este tipo de estrategias, como puede ser los sesgos adversos (p.e. reforzar el marginamiento de sectores, grupos étnicos o géneros), uso antiético de la tecnología, la falta de explicabilidad de los modelos (modelos caja negra),

Metodología

Para el desarrollo de modelos analíticos, se necesita de una metodología capaz de proporcionar una estrategia de orientación, que sea independiente de las tecnologías, los volúmenes de datos o los enfoques involucrados (IBM, 2015).

El marco metodológico adoptado corresponde al modelo CRISP-DM (*Cross Industry Standard Process for Data Mining*) (Wirth y Hipp, 2000), el cual es un estándar internacional ampliamente adoptado para el desarrollo de proyectos de minería de datos^v. Esquemáticamente las etapas del modelo se presentan en la siguiente figura

Figura 13 - Modelo CRISP-DM. Adaptado de Wirth y Hipp, 2020)



A continuación, se realiza una breve descripción de cada etapa en función de las acciones llevadas a cabo.

Comprensión del negocio

En el proceso de evaluación de riesgos LAFTFPADM de AV y PSAV, tal y como se expuso previamente, el factor de riesgo jurisdicción es fundamental para que, desde diversas perspectivas, pueda aplicarse un EBR en función de los riesgos identificados. Es por esto que, contar con un mecanismo de, asignación de nivel de riesgo LAFTFPADM a la jurisdicción en función de variables intrínsecas a la exposición de riesgo LAFTFPADM, así como a la operación de AV y PSAV es una primera etapa en esa tarea. Como segundo componente, el poder agrupar jurisdicciones en función de diversos atributos asociados al LAFTFPADM es una tarea que apoya la gestión de riesgos de contrapartes que operan con AV y de PSAV.

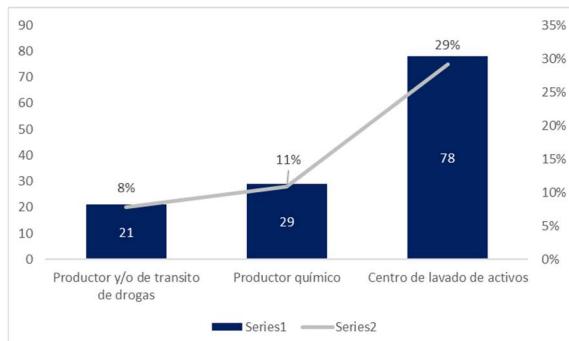
Comprender los datos

Se identificaron diversas fuentes de datos a ser inyectadas al modelo. Estas

fuentes se dividen en cuatro categorías:

- Fuentes relacionadas con LAFTFPADM
 - Productor y/o tránsito de drogas. Listado de países considerados como productores o tránsito de drogas por el Departamento de Estado de los Estados Unidos. 2021 International Narcotics Control Strategy Report. Volumen I: Drug and Chemical Control
 - Productor químico: Listado de países considerados como productores de insumos químicos para el procesamiento de narcóticos por el Departamento de Estado de los Estados Unidos. 2021 International Narcotics Control Strategy Report. Volumen I: Drug and Chemical Control
 - Centro de lavado de activos. Listado de países considerados como centro de lavado de activos por el Departamento de Estado de los Estados Unidos. 2021 International Narcotics Control Strategy Report. Volumen II: Money Laundering

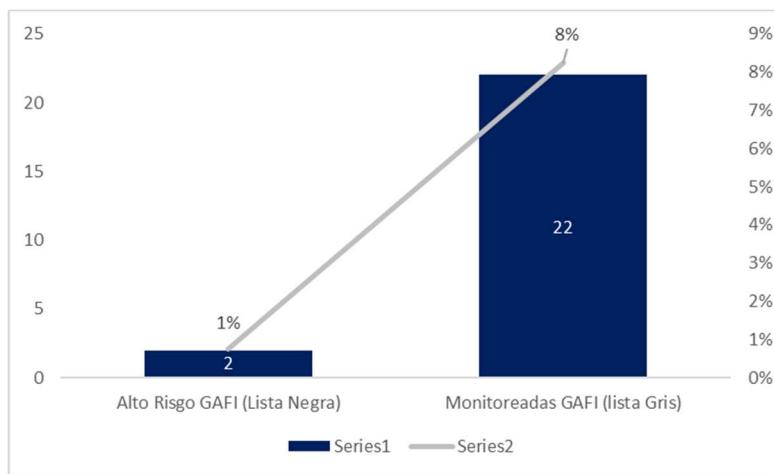
Figura 14 – Distribución de jurisdicciones informe Estrategia Internacional de Control de Narcóticos. Departamento de Estado de Estados Unidos



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

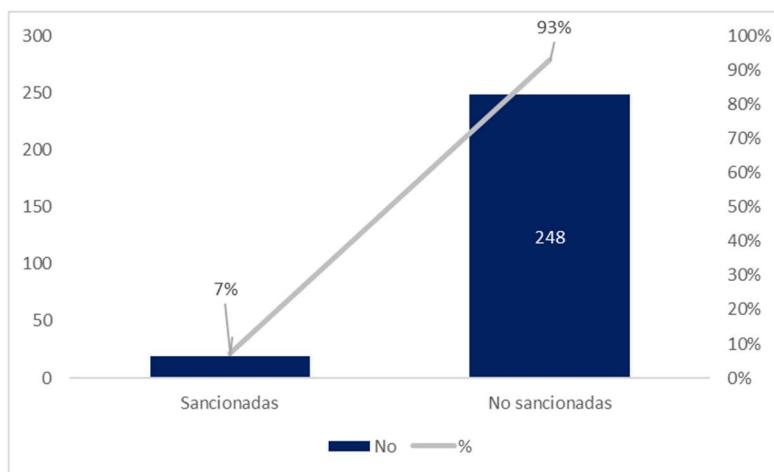
- Lista Negra de GAFI. Países que debido a sus limitaciones en los regímenes ALA/CFT han sido designados como jurisdicciones de alto riesgo. Actualizada a Marzo de 2022 (FATF, 2022, b)
- Lista Gris de GAFI. Jurisdicciones que cuentan con ciertas limitaciones en los regímenes ALA/CFT y que han sido designadas en lista de monitoreo, también denominada, lista gris. Actualizada a Marzo de 2022 (FATF 2022 a).

Figura 15 – Distribución de jurisdicciones. Lista Gris y Lista Negra de GAFI



- Programa de sanciones de OFAC. Jurisdicciones a las cuales OFAC les ha impuesto programa de sanciones.

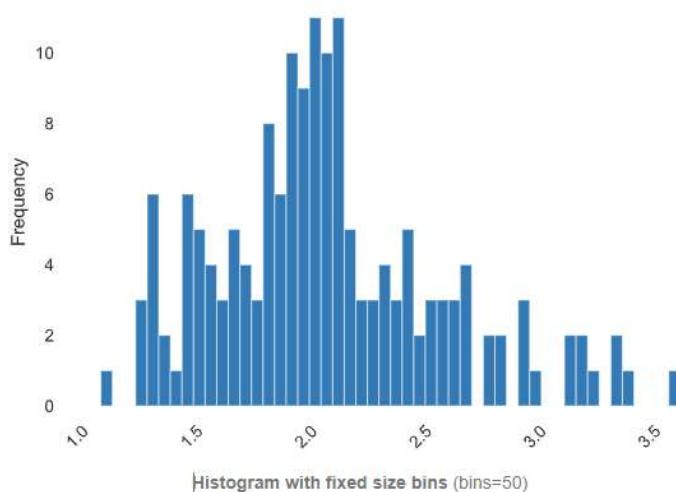
Figura 16 – Distribución de jurisdicciones. Programa de sanciones OFAC



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

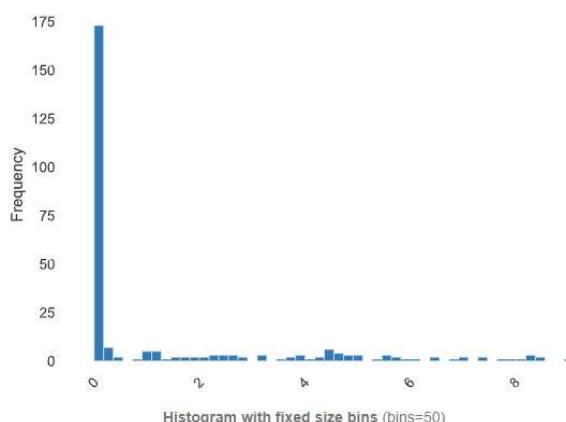
- Global Peace Index. Índice global de paz. (IEP, 2022 a) Es un indicador desarrollado por el Instituto de Economía y Paz que evalúa 23 indicadores cualitativos y cuantitativos que mide el nivel de paz en tres dimensiones. Primero, el nivel de seguridad de la sociedad, segundo, el nivel de conflicto nacional e internacional y, tercero, el grado de militarización.

Figura 17 – Histograma de frecuencias de jurisdicciones. GPI



- Global Terrorism Index. Índice Global de Terrorismo (IEP, 2022 b). Es un indicador desarrollado por el Instituto de Economía y Paz que provee un ranqueo ordinal de países basado en el impacto del terrorismo en la jurisdicción.

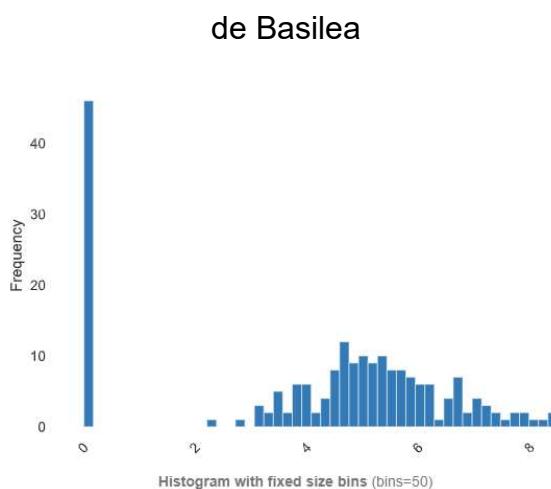
Figura 18 – Histograma de frecuencias de jurisdicciones. GTI



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

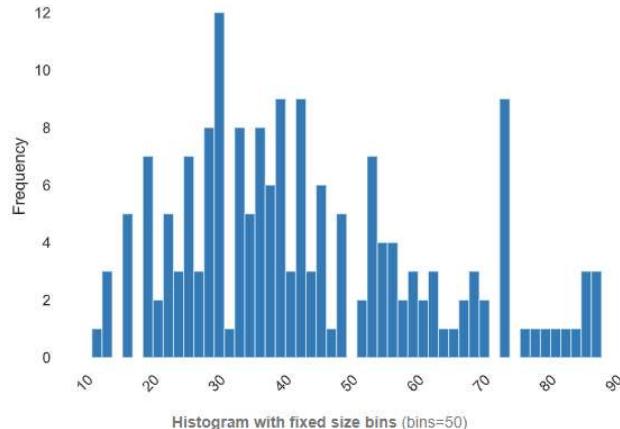
- Indice AML de Gobernanza de Basilea. Indicador independiente que provee scores de riesgo de LAFT, basado en 17 fuentes de información cubriendo 5 dominios. Calidad del marco ALA/CFT; soborno y corrupción; estándares de transparencia financiera; transparencia pública y; riesgo político y legal (BIG, 2022).

Figura 19 – Histograma de frecuencias de jurisdicciones. Indice de Gobernanza de Basilea



- Indice de Percepción de Corrupción. Indicador desarrollado por Transparencia Internacional que mide el nivel de percepción de corrupción de cada jurisdicción evaluada (Transparency International, 2022).

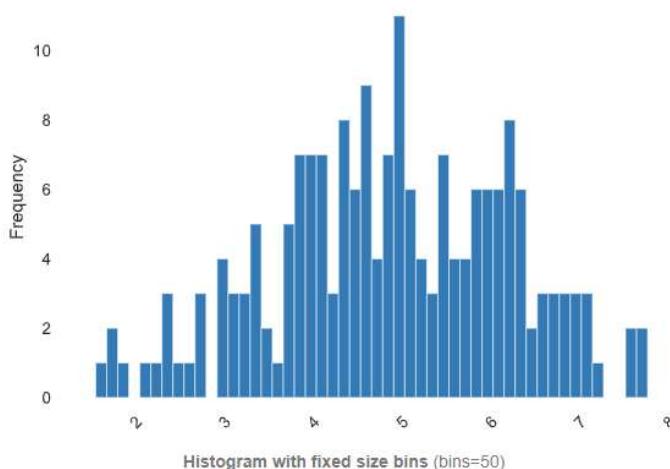
Figura 20 – Histograma de frecuencias de jurisdicciones. Indice de Percepción de Corrupción



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

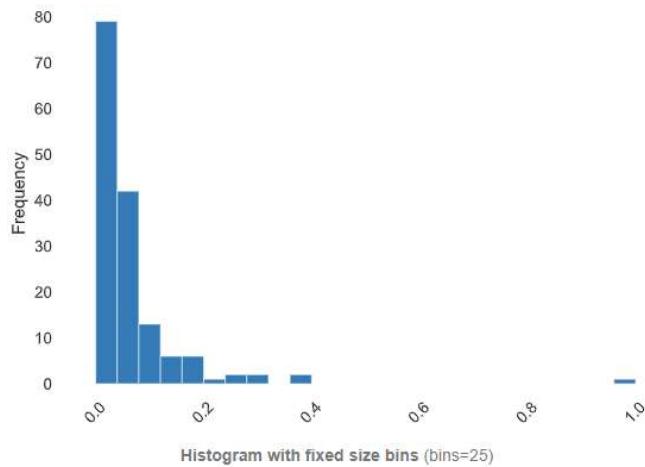
- Nivel de criminalidad (GITOC, 2021). Indicador desarrollado por GITOC que indica el nivel de criminalidad de cada jurisdicción (. El score evalúa dos dimensiones. En primera instancia, los mercados criminales (tráfico humano, contrabando de personas, tráfico de armas, crímenes ambientales, crímenes contra la fauna, crímenes contra recursos naturales no renovables, tráfico de heroína, cocaína, marihuana y drogas sintéticas. El segundo componente corresponde a los actores criminales, donde se incluyen grupos mafiosos, redes criminales, actores estatales, actores externos.

Figura 21 – Histograma de frecuencias de jurisdicciones. Índice de Percepción



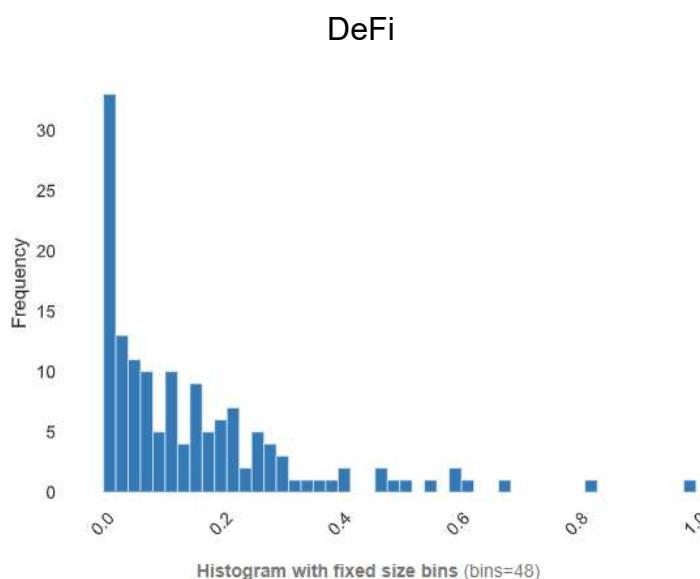
- Fuentes relacionadas con el mercado y el riesgo de AV y PSAV
 - Índice de adopción de criptomonedas (Chainalysis, 2021c). Desarrollado por Chainalysis, mide el nivel de penetración de uso de criptomonedas en la jurisdicción en particular. El índice mide tres aspectos, el valor en cadena de bloques recibido, el valor retail en cadena de bloques recibido y el volumen transado en exchanges P2P.

Figura 22 – Histograma de frecuencias de jurisdicciones. Índice de Adopción de criptomonedas



- Índice de adopción de DeFi (Chainalysis 2021c). Score desarrollado por Chainalysis que mide el nivel de penetración de plataformas DeFi por jurisdicción. El score se basa en tres mediciones, el valor en cadena de bloques recibido en plataformas DeFi, el número en cadena de bloques recibido en plataformas DeFi y, valor retail en cadena de bloques recibido en plataformas DeFi.

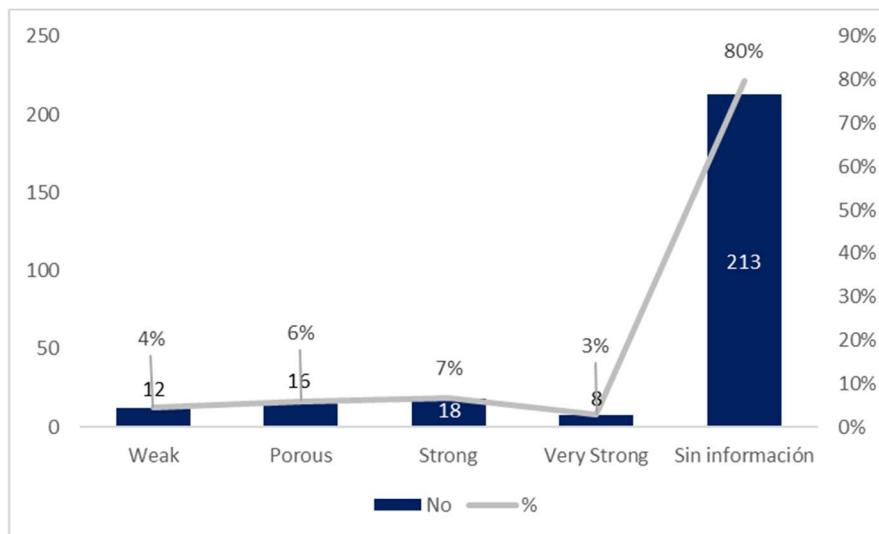
Figura 23 – Histograma de frecuencias de jurisdicciones. Índice de Adopción de DeFi



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

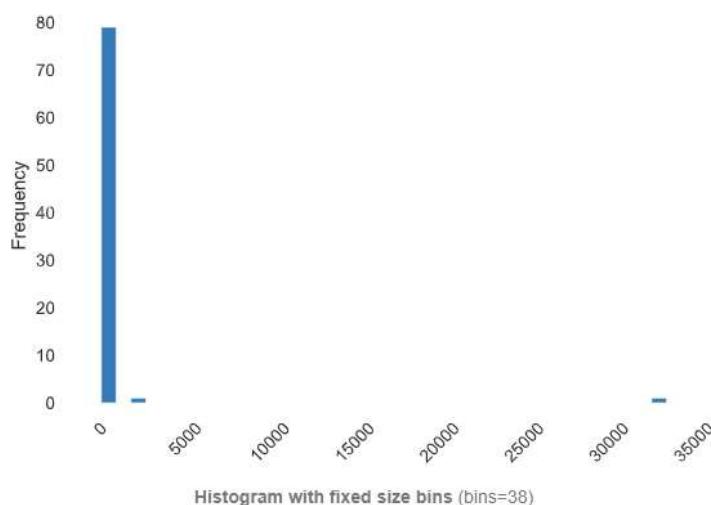
- Nivel de debida diligencia de clientes desarrollado por PSAV (CipherTrace, 2020). Categorización realizada por CipherTrace que determina el nivel de KYC que, en promedio, los PSAV realizan en determinada jurisdicción.

Figura 24 – Histograma de frecuencias de jurisdicciones. Nivel de KYC de PSAV



- Número de cajeros automáticos de criptomonedas (CoinMap).

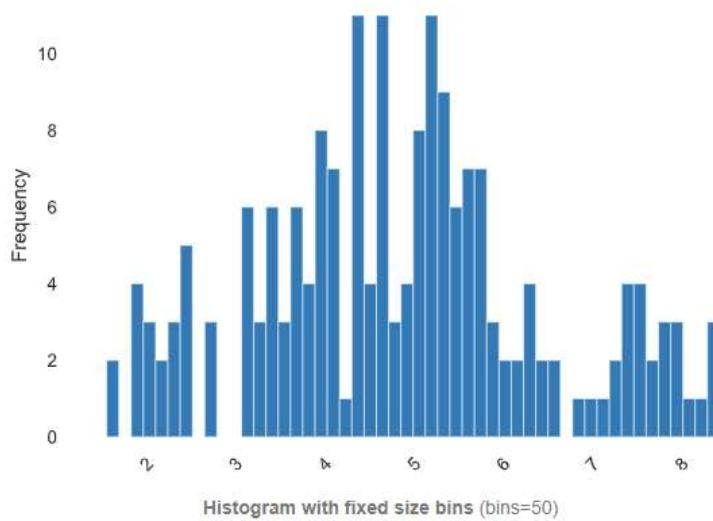
Figura 25 – Histograma de frecuencias de jurisdicciones. Número de ATM de cripto



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

- Resilencia (GITOC, 2021). Indicador diseñado por GITOC en el cual se evalúan el liderazgo político y gobernanza, la transparencia del gobierno, la cooperación internacional, políticas nacionales y leyes, sistema judicial, aplicación de la ley, integridad territorial, ALA, capacidad regulatoria, soporte a víctimas y testigos, prevención y actores no estatales.

Figura 26 – Histograma de frecuencias de jurisdicciones. Resilencia



Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Al unificar las fuentes de datos se construyó una base de datos de 267 registros, donde cada registro corresponde a una jurisdicción.

Preparar los datos

En esta etapa, se procesa la base de datos en función del tipo de modelo a implementar. En primera instancia, se identificaron 39 jurisdicciones, casi la totalidad correspondían a territorios de ultramar, que no contaban con información de ninguna de las variables recolectadas. Estas jurisdicciones fueron retiradas del análisis. Como consecuencia, se establecieron 228 jurisdicciones para el análisis. El segundo paso, dado que no todos los registros se encontraban completos, fue el determinar los métodos de imputación de datos.

Tabla 3 – Métodos de imputación de atributos

Atributo	Método de imputación
GTI	Ceros los vacíos. Se supone que no cuentan con información de actos terroristas
Ciphertrace VASP KYC	Nueva categoría. "No Info"
Bitcoin ATM	Ceros jurisdicciones de las que no se cuenta con información. Se supone que no tienen ATM Cripto
Crypto Legalty	Unknown

Los atributos no listados fueron imputados a través de un modelo de KNN. Este modelo corresponde a un clasificador de K vecinos más cercanos. Está basado en memoria y por lo tanto, no requiere de un modelo a ser ajustado. Dado un punto x_0 , se identifican los k datos de entrenamiento $x_{(r)}, r = 1, \dots, k$ más cercanos en distancia a x_0 y el dato faltante se imputa usando el voto mayoritario (mayor frecuencia) entre los k vecinos (Hastie et al, 2000). Para el modelo de imputación se utilizó un KNN con $k = 5$.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

Posteriormente, con base en las categorías de variables previamente enunciadas, se unificó la base de datos en función de los siguientes componentes:

- Nivel de riesgo LAFTFPADM. Variable categórica que consolida la información de las fuentes relacionadas con el riesgo LAFTFPADM.
- Nivel de riesgo de AV y PSAV. Variable categórica que consolida la información de las fuentes relacionadas con el mercado de AV y PSAV.
- Nivel de riesgo de Resilencia. Variable categórica que representa las debilidades en resiliencia de la jurisdicción
- Legalidad de criptoactivos. Variable categorica que representa los niveles de legalidad de uso de criptoactivos en la jurisdicción

Finalmente, debido a que el modelo de clustering será basado en distancias, las variables categóricas fueron transformadas a través del denominado *one-hot encoding*. De esta forma, se cuenta con una base de datos totalmente conformada por valor numéricos.

Modelar

El modelo seleccionado para realizar el agrupamiento de jurisdicciones en función de las variables previamente construidas, tal y como se expuso en el marco teórico, es el cluster de K-medias.

Antes de la implementación del modelo, es recomendable realizar algunos análisis adicionales sobre la fuente de datos.

El primero de estos análisis es el determinar si existen atributos cuya varianza sea cero. Esto implica que la variable contiene un valor constante en todos sus registros y por lo tanto, no aporta información que permita separar los registros

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

en grupos. Las varianzas de los atributos utilizados en el modelo se presentan a continuación

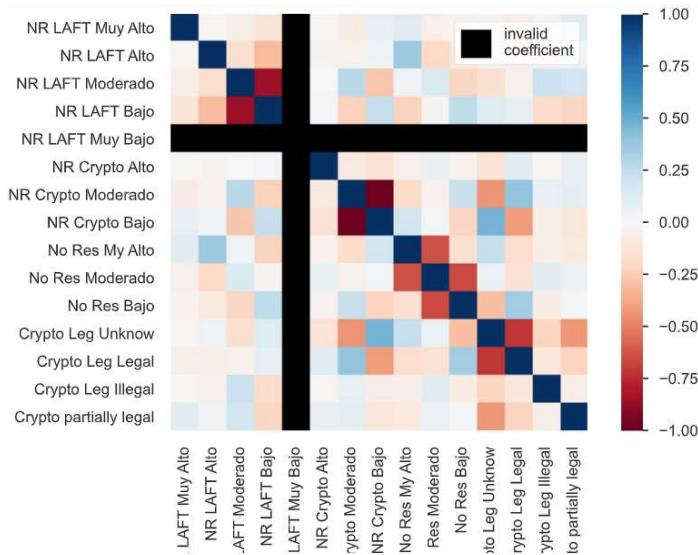
Tabla 4 – Varianza de atributos

Atributo		Varianza
Nivel de Riesgo	NR LAFT Muy Alto	0,00869498
	NR LAFT Alto	0,05376654
	NR LAFT Moderado	0,21275777
	NR LAFT Bajo	0,23382195
	NR LAFT Muy Bajo	0
	NR Crypto Alto	0,01298476
	NR Crypto Moderado	0,23382195
	NR Crypto Bajo	0,236996
Atributo		Varianza
Legalidad de Cripto	Crypto Leg Unknown	0,24068944
	Crypto Leg Legal	0,19180902
	Crypto Leg Illegal	0,02975916
	Crypto partially legal	0,10103109
NR Resiliencia	No Res My Alto	0,11101493
	No Res Moderado	0,19390582
	No Res Bajo	0,11747845

Del análisis de varianza se concluye que la variable Nivel de Riesgo LAFT Muy Bajo es retirada del análisis.

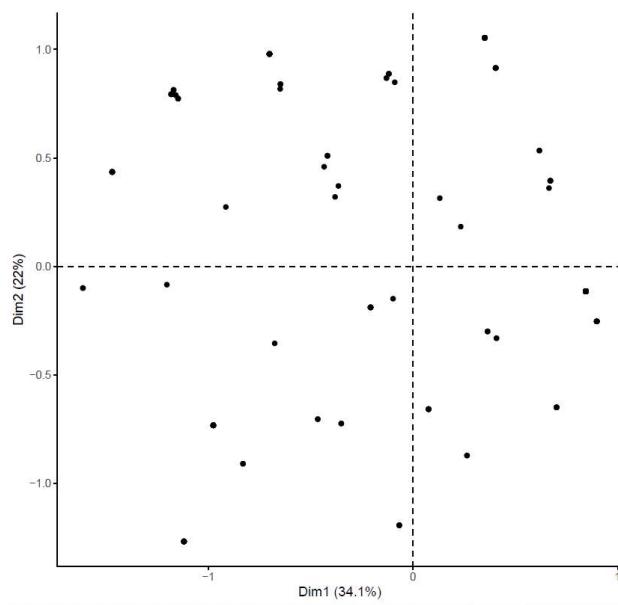
El segundo análisis preliminar implementado es del de correlaciones entre variables. Un grupo de atributos altamente correlacionados pueden ser representados en el modelo por un solo de dichos atributos. Ninguna variable fue removida del análisis por este concepto. A continuación se presenta la matriz de correlaciones.

Figura 28 – Matriz de correlaciones



El ultimo análisis previo a la implementación del modelo es determinar la “clustereabilidad” del set de datos. Esto implica determinar si los datos tienen una estructura intrínseca que conforma grupos o se encuentran distribuidos de forma aleatoria. Para esto, se utilizo inicialmente la técnica de reducción de dimensiones denominada Análisis de Componentes Principales *PCA – Principal Components Analysis). Esta consisten básicamente en proyectar los puntos sobre un nuevo sistema coordenado, donde cada dimensión corresponde a una combinación lineal de una o varias variables en la dimensión original. Por construcción, estos nuevos componentes son ortogonales entre si, lo que implica que la correlación entre ellos es cero.

Figura 29 – Componentes principales



En la representación grafica de los componentes principales, los dos primeros componentes agrupan el 54.1% de la varianza total de los datos. Aparentemente pueden establecerse grupos o segmentos en el set de datos. Con el fin de obtener mayor claridad respecto a la clustereabilidad del set de datos, se implementaron dos (2) técnicas de reducción de dimensiones, para combinaciones no lineales de variables. Estas fueron UMAP y t/SNE.

Figura 30 – Reducción de dimensiones. t-SNE

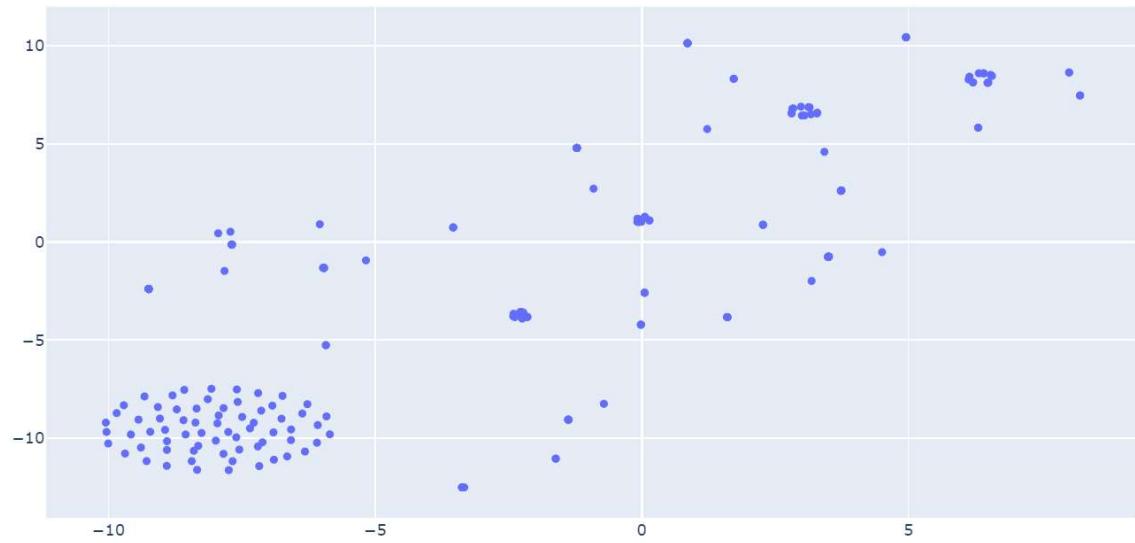
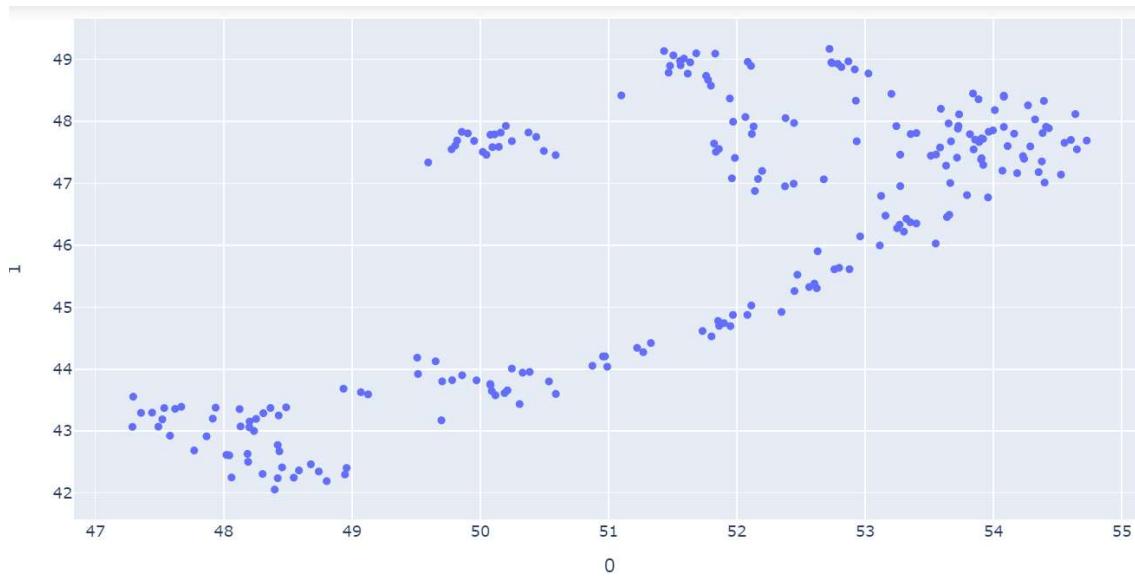


Figura 31 – Reducción de dimensiones. UMAP



Al observar los graficos resultantes de la aplicación de las técnicas de reducción de dimensiones UMAP y t-SNE, se observa con mayor claridad la estructura de grupos en el set de datos. Adicionalmente, visualmente se confirma que los grupos tienen tendencia hiperesférica, por lo que la técnica de k-medias es aplicable a los mismos.

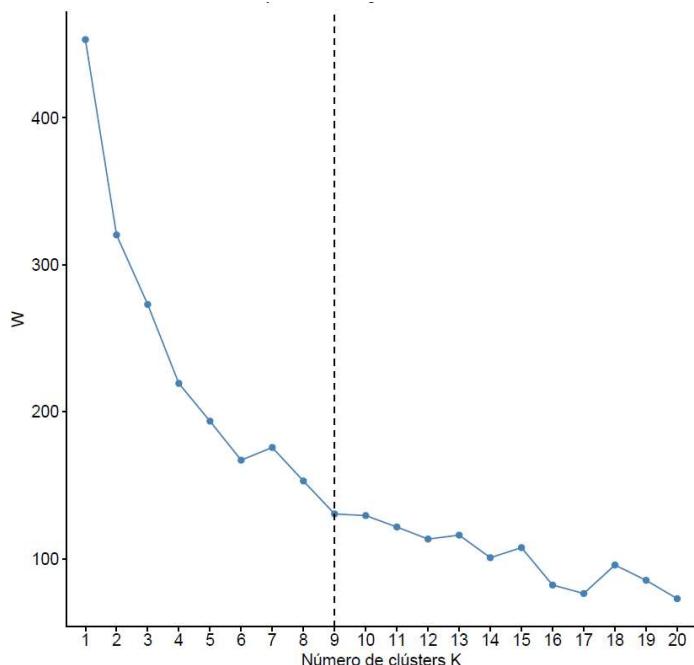
Una vez se determina que los datos son susceptibles a ser segmentados a través de cluster, se procede a determinar el numero optimo de segmentos. Para esto, a manera indicativa se utilizaron dos (2) técnicas.

La primera y mas tradicional, el criterio del codo (*elbow criterion*). Este método analiza el porcentaje de varianza como función del número de clusters. De esta forma, el número de clusters óptimo será aquel en el cual adicionar más clúster no proporciona un mejor modelamiento de los datos. De esta forma, se emplea el criterio de “Distancia promedio al intracentroide” (W) la cual es “*calculada promediando la distancia entre el centroide y todos los registros que pertenezcan a dicho clúster*”, según la siguiente fórmula:

$$W = \frac{1}{n} \sum_{i=1}^n \frac{1}{m} \sum_{j=1}^m d(c_i, p_j)$$

Donde n es el número de clústers, m el número de elementos que componen el clúster i , c_i denota el centroide del clúster i , p_j es el elemento j del clúster i y $d(c_i, p_j)$ es la distancia entre el centroide c_i y el punto p_j .

Figura 32 – Criterio del codo



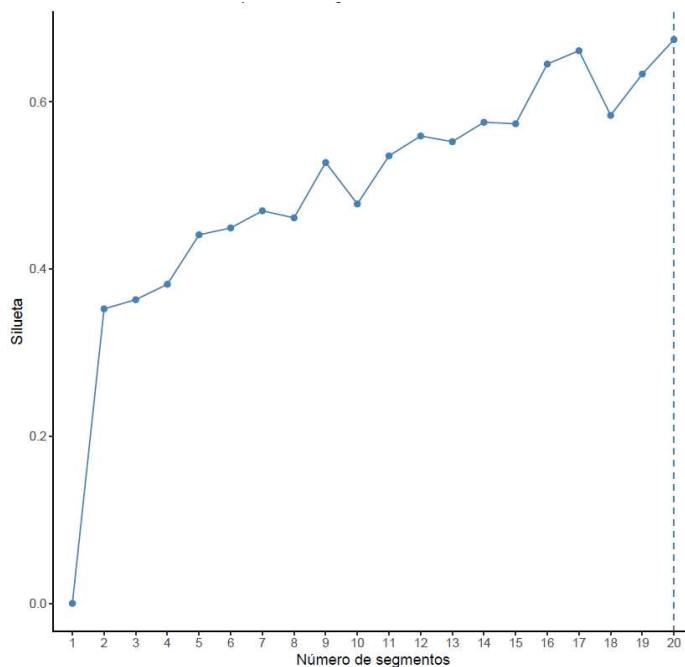
Bajo el criterio del codo, el numero optimo de segmentos esta cerca a los 14 segmentos.

El segundo criterio orientador para determinar el numero de segmentos es el uso del índice de Silueta (Peter, 1987). La silueta es “una medida de cohesión en el clúster y separación entre clústers” (Wendler, Groutup). La silueta puede ser calculada según la siguiente expresión (Rousseeuw, 1987):

$$s(i) = \frac{b(i) - a(i)}{\max \{a(i), b(i)\}}$$

Donde, para el datum i , sea $a(i)$ la distancia promedio entre i y todos los demás datos dentro del mismo clúster. De esta forma, $a(i)$ puede interpretarse como como una medida de qué tan bien i ha sido asignado a su clúster (entre menor el tamaño, ha sido asignado de mejor forma). La disimilaridad promedio del punto i a su clúster se define como c y es la distancia promedio de i a todos los puntos en c . Sea $b(i)$ la distancia promedio más baja de i a todos los puntos de cualquier clúster de los cuales i no pertenece. El clúster con la menor disimilaridad promedio se denomina clúster más cercano de i , dado que es el mejor siguiente clúster para el punto i . Para que $s(i)$ sea cercana a 1 se requiere que $a(i) \ll b(i)$. Dado que $a(i)$ es una medida de disimilaridad entre el punto i y su propio clúster, un valor pequeño implica que es una asignación correcta. Un valor alto de $b(i)$ implica que i ha sido asignado incorrectamente a su clúster. De esta forma, un valor de $s(i)$ cercano a uno significa que los datos fueron segmentados apropiadamente. De esta forma, el valor promedio $s(i)$ sobre todos los datos es una medida de qué tan bien agrupados se encuentran los datos en los clústers.

Figura 33 – Silueta



Los mayores valores de silueta se dan en 20 segmentos, seguido de 16 y luego 14.

Evaluar

En esta etapa del proyecto, se ha construido un modelo (o modelos) que aparentemente tienen alta calidad desde la perspectiva de análisis de datos. Antes de proceder a la etapa final (deployment) es importante evaluarlo(s) así como revisar las etapas necesarias para su generación, con el objetivo de asegurar que el modelo cumple con los objetivos de negocio. Un objetivo fundamental es el determinar si existen elementos de negocio que no fueron considerados. Al final de la etapa, una decisión respecto al uso del modelo de data mining es tomada

Con base en lo presentado en la sección anterior y al analizar los niveles de concentración mínima y máxima de segmentos resultantes (Vease Anexo 1), se determinó que el numero optimo de segmentos es de catorce (14). Este numero de segmentos asegura que no existe una alta concentración en un segmento

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado particular (inferior al 30%) pero también, que los segmentos con menor numero de registros no son tan pequeños que dejan de ser significantes. En la siguiente tabla se presentan los niveles de concentración de registros por segmento para el modelo seleccionado

Tabla 5 – Concentraciones por segmento

Cluster	n	%
1	30	13%
2	16	7%
3	10	4%
4	67	29%
5	13	6%
6	15	7%
7	10	4%
8	11	5%
9	12	5%
10	9	4%
11	4	2%
12	9	4%
13	15	7%
14	7	3%

Implementación

Debido al carácter teórico del ejercicio el modelo no fue desplegado en un ambiente de producción. Sin embargo, el modelo puede ser desplegado en las condiciones existentes por parte de un regulador, una UIF o un sujeto obligado, bajo la siguiente premisa. Esta corresponde a evaluar medidas de control diferenciales por segmento a jurisdicciones donde, requiera aplicar un EBR a jurisdicciones basado en el riesgo LAFTFPADM de AV o PSAV, con los cuales se encuentre en etapa preoperativa o de exploración.

Para esto, ante cambios en las variables utilizadas puede implementarse un modelo de pronóstico de segmentos que establezca si existen cambios estructurales que hagan que una jurisdicción cambie de segmento y por lo tanto,

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

pueda aumentar o disminuir las medidas de control diseñadas. Para esto, se aplica la siguiente expresión donde cada punto x es asignado a un cluster basado en la siguiente expresión

$$\arg \min_{c_i \in C} d(c_i, x)^2$$

Siendo c_i el centroide i y x el registro a evaluar.

Para entidades, tanto públicas como privadas que ya cuentan con información transaccional, por ejemplo, de moneda fiat en cash-in o cash-out,, o de ingreso y salida de AV (en el caso de un PSAV) el modelo puede ser ajustado o recalibrado para incorporar esta información transaccional y de esta forma, no solo determinar si hay cambios de segmento ante cambios en las variables utilizadas, sino también el diseñar señales de alerta diferenciales por segmento para cada una de los atributos transaccionales incorporados, con el objetivo de identificar jurisdicciones atípicas para el segmento y de esta manera, poder realizar una gestión de riesgos proactiva.

Es así que “el Sistema ALA/CFT se beneficia del análisis clúster porque le permite generar grupos homogéneos de individuos (segmentos) y, a partir de estos, identificar conductas atípicas que pueden llegar a asociarse con comportamientos sospechosos de delitos relacionados con el LA y la FT. En otras palabras, conociendo la habitualidad de un grupo de personas es posible reconocer cuando alguna de ellas se está comportando de manera extraña, en comparación con sus pares” (Gafilat, 2021 b).

Resultados y conclusiones

Las características de los segmentos identificados son>

- Segmento 1: Todos los registros del segmento cuentan con un nivel de riesgo (NR) LAFTFPADM Bajo NR de AV Bajo. En relación al NR de Resilencia institucional, el 53% son bajo y el 47% son moderado. Respecto a la legalidad de la industria cripto, en casi la totalidad (93%) es legal y en una pequeña proporción 7% parcialmente legal.
- Segmento 2. El segmento esta compuesto por jurisdicciones de NR LAFTFPADM Moderado, NR Cripto Bajo y NR de Resilencia institucional Moderado. En el 81% de las jurisdicciones del segmento no se cuenta con informacion respecto a la legalidad de la industria cripto y en el 19% se considera una actividad ilegal.
- Segmento 3. El 100% de las jurisdicciones del segmento tienen un NR LAFTFPADM Alto. En la mayor proporción (90%) el NR de la industria cripto es Bajo y en las jurisdicciones restantes (10%), es Moderado. En el segmento predomina (80%) el NR de Resilencia institucional Muy Alto, con una menor presencia (20%) de riesgo Moderado. En términos de legalidad, se presentan jurisdicciones de las cuales se desconoce esta informacion (80%), jurisdicciones donde el negocio cripto es legal (20%) y jurisdicciones donde es parcialmente legal (10%).
- Segmento 4. El segmento se conforma exclusivamente por jurisdicciones de NR LAFTFPADM bajo, NR de cripto Bajo, NR de Resilencia institucional moderado y legalidad de criptodivisas desconocido.
- Segmento 5. El segmento agrupa jurisdicciones mayoritariamente (92%) de NR LAFTFPADM Moderado, con una pequeña proporcion (8%) de NR Alto. Todos los registros tienen un NR de criptoactivos Moderado. Los NR

de resiliencia institucional presentes en el segmento son Moderado (85%)

y Bajo (15%). En todas las jurisdicciones, los criptoactivos son legales.

- Segmento 6. La totalidad de jurisdicciones del segmento tienen un NR LAFTFPADM Moderado, un NR de criptoactivos Moderado y se desconoce la legalidad de la industria cripto. El NR de Resilencia Institucional de mayor presencia en el segmento es Moderado (80%), seguido de Muy Alto (20%).
- Segmento 7. El segmento está conformado exclusivamente por jurisdicciones con NR LAFTFPADM Bajo, NR de cripto Bajo, NR de Resilencia Institucional Muy Alto y legalidad de criptoactivos desconocida.
- Segmento 8. Todas las jurisdicciones del segmento poseen un NR LAFTFPADM Bajo, un NR de cripto Bajo, NR de Resilencia Institucional Bajo. En términos de legalidad de la industria cripto, en el segmento se encuentran jurisdicciones donde se desconoce esta información (56%), donde es legal (36%) y donde es parcialmente legal (18%).
- Segmento 9. Los NR LAFTFPADM presentes en el segmento son Bajo (92%) y Alto (8%). Todos los registros tienen un NR de cripto Moderado. Los NR de Resilencia Institucional presentes en el segmento son Moderado (83%) y Bajo (17%). En la gran mayoría (92%) se desconoce el nivel de legalidad de la industria cripto mientras que en la proporción restante (8%) es ilegal.
- Segmento 10. Todos los registros del segmento poseen un NR LAFTFPADM Bajo, un NR de Resilencia Institucional Moderado y la industria cripto es legal. Los NR de criptoactivos presentes en el segmento son Bajo (78%) y Alto (22%).

- Segmento 11. El segmento congrega jurisdicciones de NR LAFTFPADM Moderado, NR Cripto Alto y donde la industria cripto es legal. En relación al NR de Resilencia Institucional, se presentan los niveles Moderado (75%) y Alto (25%).
- Segmento 12. En el segmento se ubican jurisdicciones con NR LAFTFPADM Bajo (56%), Moderado (33%) y Muy Alto (11%). Todas las jurisdicciones del segmento tienen un NR de criptoactivos Bajo, un NR de Resilencia Institucional Moderado y la industria cripto es parcialmente legal.
- Segmento 13. El NR LAFTFPADM preponderante en el segmento (93%) es el nivel Moderado, seguido del nivel Alto (7%). Los NR de la industria cripto presentes en el segmento son Moderado (93%) y Alto (7%). La totalidad de registros tienen un NR Moderado en Resilencia Institucional. La mayoría de jurisdicciones han legalizado parcialmente la industria cripto (80%), con presencia de jurisdicciones (20%) en las que se considera ilegal la industria.
- Segmento 14. En el segmento se ubican jurisdicciones con NR LAFTFPADM Moderado (86%) y Muy Alto (14%). Todos los registros del segmento tiene un NR Cripto bajo, NR de resilencia institucional Muy Alto y se desconoce si la industria cripto es legal o ilegal.

Finalmente, se puede concluir lo siguiente

- Es posible, a partir de fuente objetivas de información, desarrollar un modelo de aprendizaje de máquinas no supervisado que agrupe jurisdicciones en función de diversos atributos asociados a la exposición

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

en términos de LAFTFPADM, de características intrínsecas de la industria de AV y PSAV, de fortaleza de las instituciones y de legalidad de la industria de criptoactivos.

- Este modelo puede ser utilizado como insumo por sujetos obligados, organismo de regulación y UIF con el objetivo de analizar el mercado, determinar niveles de exposición desde la perspectiva de LAFT de AV y PSAV y de esta forma, diseñar mitigantes aplicando un EBR.
- El modelo puede ser de utilidad en procesos tales como: determinación de niveles de riesgo de contrapartes vinculadas con AV y/o PSAV; de evaluación de transacciones de contrapartes con AV y PSAV desde el componente de jurisdicción tanto entrante como saliente< de monitoreo de jurisdicciones con relación a AV y PSAV, de evaluación de cumplimiento normativo de contrapartes que operen con AV y PSAV, entre otros usos.
- El modelo puede recalibrarse para incluir información transaccional, ya sea de cash-in . cash-out de operaciones con AV, de transferencias desde y hacia PSAV, entre otras y de esta manera, diseñar señales de alerta diferenciales por segmento que permitan la identificación de atípicos transaccionales.

Bibliografía

1. Auer, R., Boar, C., Cornelli, G, Frost, J., Holden, H., Wehrli, A. (2021). CBDCs beyond borders:results from a survey of central banks. BIS Papers No 116. Junio 2021.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

2. Auer, R , Farag, M, Lewrick, U., Orazem, L., Zoss, M (2022) Banking in the shadow of Bitcoin? The institutional adoption of cryptocurrencies. BIS Working Papers No 1013. Mayo.
3. Basel Institute of Gobernance (2022). Basel AML Index 2021: 10th Public Edition. Ranking money laundering and terrorist financing risks around the world
4. Bloomberg (2021): “Banks chart cautious crypto plans with regulators taking aim”, Junio 21 de 2021.
5. BrainCube. AI vs. Machine Learning vs. Data Science for Industry. Key Definitions, Differences, and How They Work Together. Disponible en <https://braincube.com/resource/manufacturing-ai-vs-machine-learning-vs-data-science/>
6. CipherTrace (2021). Cryptocurrency crime and anto-money laundering report. Agosto 2021.
7. Coelho R, J Fishman y D Garcia Ocampo (2021): “Supervising cryptoassets for anti-money laundering”, FSI Insights, no 31.
8. CipherTrace. What Exactly is a Virtual Asset Service Provider (VASP)?
<https://ciphertrace.com/what-exactly-is-a-virtual-asset-service-provider-vasp/>
9. CipherTrace (2020). CipherTrace Geographic Risk Report: VASP KYC by Jurisdiction
10. CoinMap. BitCoin ATM Map. Disponible en <https://coinatmradar.com/>
11. Chainalysis (2020). Who`s who on the blockchains? The chainalysis guide to cryptocurrency typologies. Octubre 2020.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

12. Chainalysis (2021a). The 2021 crypto Crime Report. Everything you need to know about ransomware, darknet markets, and more. Febrero 2021.
13. Chainalysis (2021b). The 2021 Geography of cryptocurrency report. Analysis of geographic trends in cryptocurrency adoption and usage. Octubre 2021.
14. Chainalysis (2021c). The 2021 Geography of Cryptocurrency Report Analysis of Geographic Trends in Cryptocurrency Adoption and Usage
15. Chainalysis (2022). The 2022 crypto Crime Report. Original data and research into cryptocurrency-based crime.
16. Elptic (2020). Financial crime typologies in cryptoassets. The concise guide for compliance leaders.
17. Fanusie, Y, Robinson, T.(2018) Bitcoin Laundering: an analysis of illicit flows into digital currency services.
18. Financial Action Task Force (FATF) (2014). Virtual currencies key definitions and potential AML/CFT risks, Junio.
19. Financial Action Task Force (FATF) (2019). Money laundering risks from “stablecoins” and other emerging assets. Disponible en <https://www.fatf-gafi.org/publications/fatfgeneral/documents/statement-virtual-assets-global-stablecoins.html>
20. Financial Action Task Force (FATF) (2020). Virtual Assets Red Flag Indicators of Money Laundering and Terrorist Financing
21. Financial Action Task Force (FATF) (2021a). Opportunities and challenges of new technologies for AML/CTF.Julio 2021.
22. Financial Action Task Force (FATF) (2021b). Virtual Assets and Virtual Assets Service Providers. Octubre 2021.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

23. Financial Action Task Force (FATF) (2022a). High-Risk Jurisdictions subject to a Call for Action - Marzo 2022. Disponible en <http://www.fatf-gafi.org/publications/high-risk-and-other-monitored-jurisdictions/documents/call-for-action-march-2022.html>
24. Financial Action Task Force (FATF) (2022b). urisdicctions under Increased Monitoring - Marzo 2022. Disponible en <http://www.fatf-gafi.org/publications/high-risk-and-other-monitored-jurisdictions/documents/increased-monitoring-march-2022.html>
25. Financial Stability Board (FSB) (2020): "Regulation, Supervision and Oversight of 'Global Stablecoin' Arrangements, Final Report and High-Level Recommendations", Octubre..
26. Foley, S, J Karlsen and T Putnins (2019): "Sex, drugs, and bitcoin: how much illegal activity is financed through cryptocurrencies?", The Review of Financial Studies, vol 32(5), pp 1798–1853
27. GAFILAT (2021a). Análisis Estratégico sobre el uso de inteligencia artificial, minería de datos y análisis de Big Data en prevención y detección LA/FT
28. GAFILAT (2021b). Estándares Internacionales sobre la lucha contra el lavado de activos, el financiamiento del terrorismo y el financiamiento de la proliferación de armas de destrucción masiva.(UIF/MP).
29. Global Initiative against transnational organized crime (GTOC, 2021). Global Organized Crime Index 2021.
30. Han, J.; Kamber, M.; Pei, J. Data Mining, concepts and techniques. Elsevier. Third Edition.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

31. Hartigan, J., & Wong, M. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108. doi:10.2307/2346830
32. Hastie, T., Tibshirani, R., Friedman, J. (2009) The elements of Statistical Learning. Data Minig, Inference and Prediction. Second Edition.
33. IBM (2015). Metodología Fundamental para la Ciencia de Datos. Disponible en <https://www.ibm.com/downloads/cas/6RZMKDN8>
34. Institute of Economics and Policy (IEP) (2022a). Global Peace Index 2022. Measuring peace in a complex world.
35. Institute of Economics and Policy (IEP) (2022b). Global Terrorism Index 2022. Measuring the impact of Terrorism.
36. Kosse, A., Matteri, I. (2022). Gaining momentum – Results of the 2021 BIS survey on central bank digital currencies
37. MacQueen, J. (1967) Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Statistics, 281--297, University of California Press, Berkeley, Calif., 1967.
<https://projecteuclid.org/euclid.bsmsp/1200512992>.
38. Ng, A.; Piech, C. Stanford University. Disponible en www.stanford.edu/~cpiech/handouts/kmens.html
39. Peter J. Rousseeuw (1987). "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis". *Computational and Applied Mathematics*. 20: 53–65
40. Popper, N, Matsnev, O, Herrero A (2018) “Russia and Venezuela’s Plan to Sidestep Sanctions: Virtual Currencies,” *The New York Times*, Enero 3,

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

2018. (<https://www.nytimes.com/2018/01/03/technology/russia-venezuela-virtual-currencies.html>)

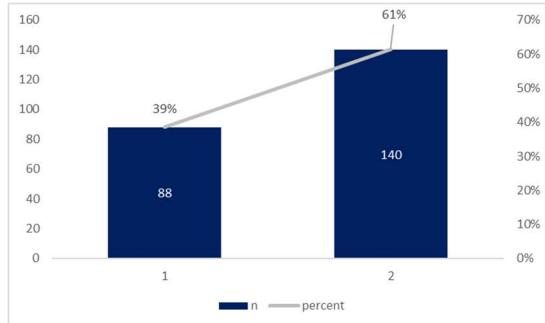
41. Rainerwalt (2021). Supervised vs Unsupervised vs Reinforcement Learning
 - Knowing the differences is a fundamental part of properly understanding machine learning. Disponible en <https://starship-knowledge.com/supervised-vs-unsupervised-vs-reinforcement-learning/>
42. RapidMiner. Cluster distance performance. Disponible en https://docs.rapidminer.com/studio/operators/validation/performance/segregation/cluster_distance_performance.html
43. Rousseeuw, P. (1987) Solhouttes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics. Vol 20 (1987). Pág 53 - 65
44. Sreedhar C.; Kasiviswanath N.; Chenna, P. (2017) Clusterin large datasets using K-Means modified inter and intra cluistering (KM-12C) in Hadoop. Journal of Big Data. 2017.
45. Unidad de Información y Análisis Financiero (UIAF) (2014a). Aplicabilidad de la minería de datos y el análisis de redes sociales en la inteligencia financiera.
46. Unidad de Información y Análisis Financiero (UIAF). /2014b). Técnicas de Minería de Datos para la detección y prevención de lavado de activos y financiación al terrorismo (LA/FT).
47. Ullman S, Poggio T.; Harari, D.; Seibler, D.(2014) Unsupervised Learning. Clusterirng. Massachusstets Instute of Technology – MIT. Center for Brains, Minds & Machines. Fall Semester 2014.

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

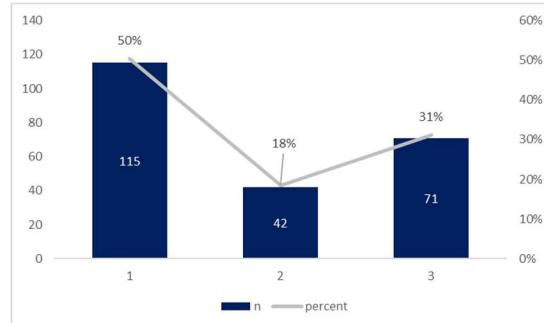
48. US DEPARTMENT OF STATE (2022). 2021 International Narcotics Control Strategy Report. Volumen I: Drug and Chemical Control
49. US DEPARTMENT OF STATE (2022). 2021 International Narcotics Control Strategy Report. Volumen II: Money Laundering
50. US Department of Treasury (2021). Treasury Takes Robust Actions to Counter Ransomware. Disponible en <https://home.treasury.gov/news/press-releases/jy0364>
51. Transparency International (2022). Corruption Perception Index 2021.
52. Wendler, T.; Grottrup, S. Data Mining with SPSS Modeler: theory , exercises and solutions. Springer.
53. Wirth, R., Hipp, J. (2000) CRISP-DM: Towards a Standard Process Model for Data Mining. Proceedings of the 4th International Conference on the practical of knowledge discovery and data mining.

Anexo 1 – Concentraciones de segmentos

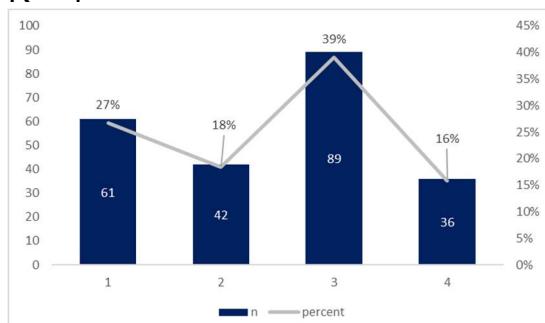
K = 2



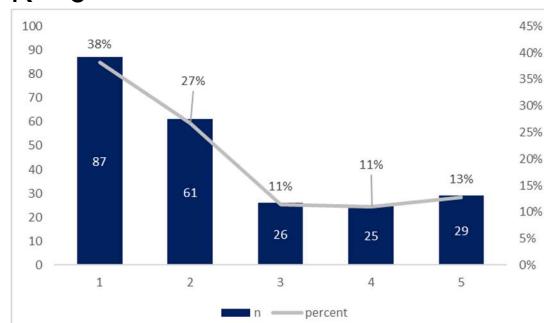
K = 3



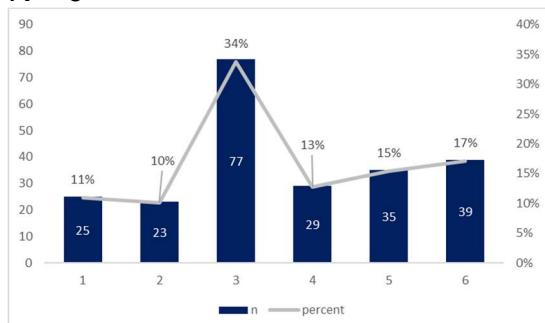
K = 4



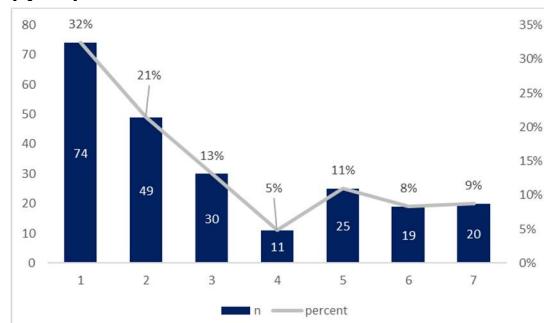
K = 5



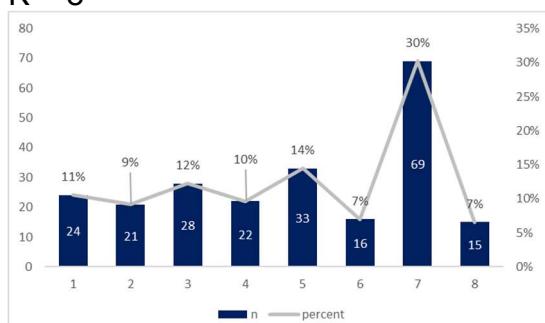
K = 6



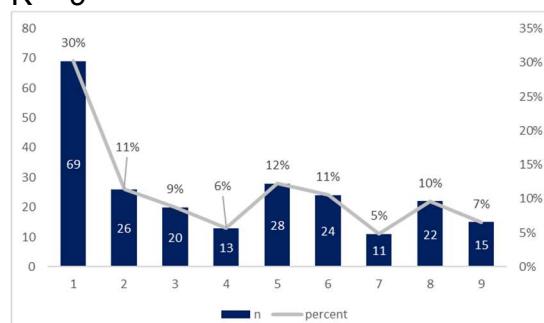
K = 7



K = 8

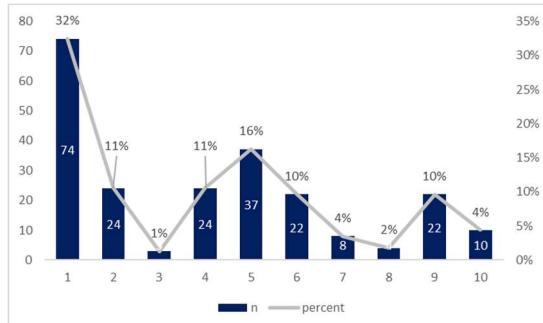


K = 9

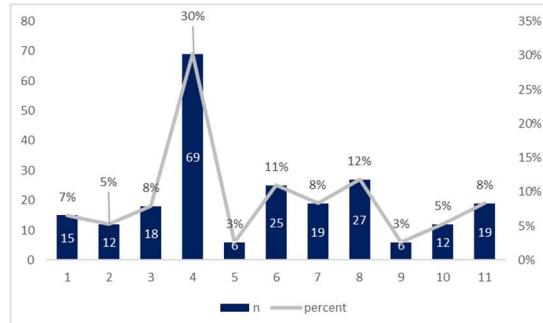


Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

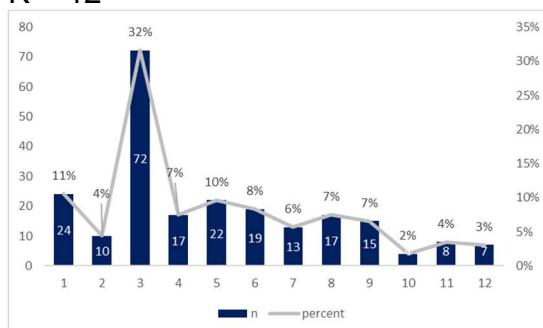
K = 10



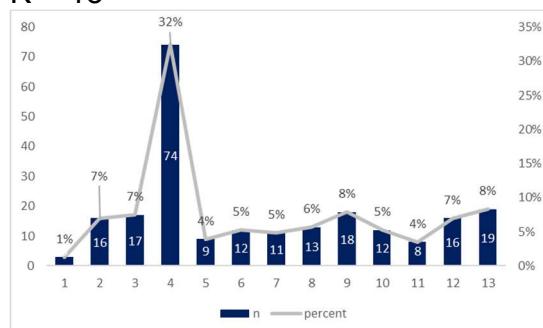
K = 11



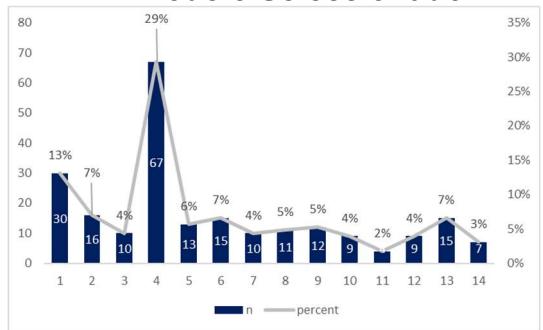
K = 12



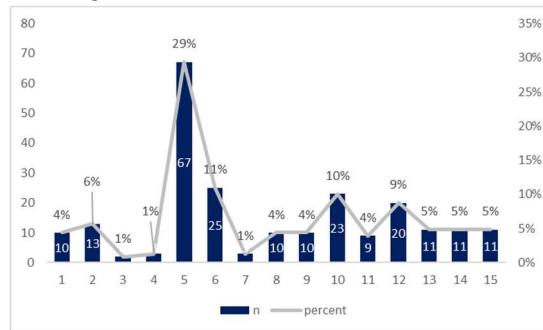
K = 13



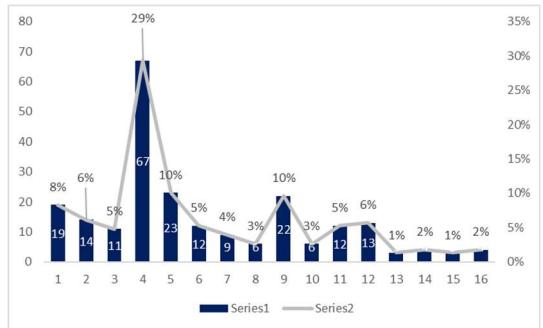
K = 14 – Modelo Seleccionado



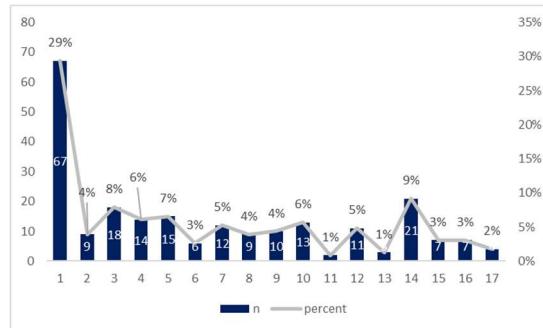
K = 15



K = 16

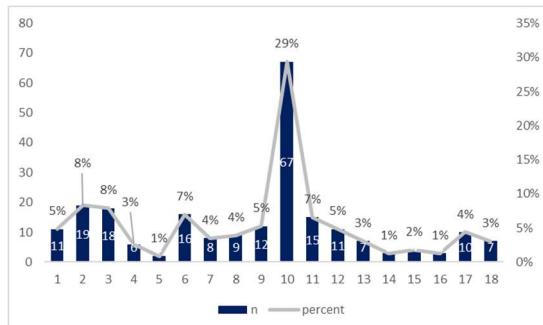


K = 17

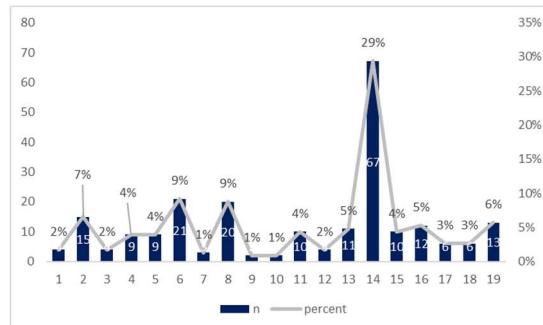


Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

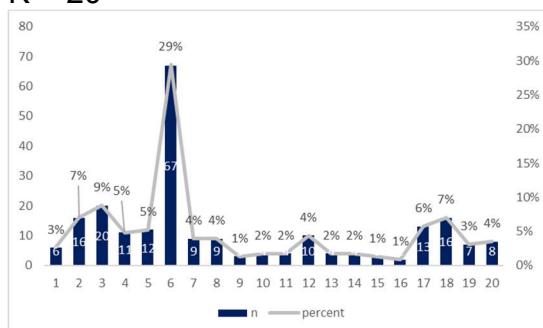
K = 18



K = 19



K = 20



ⁱ Una criptodivisa es “un activo digital privado con su propia unidad de medida, tales como Bitcoin y Ethereum. Las criptodivisas no representan una reclamación ante un banco central y por lo tanto, son diferentes a las DBDC” (Koss, Mattei, 2022)

ⁱⁱ Una CBDC es “una moneda digital emitida por un banco central denominada en la unidad nacional de contabilidad y representa una acreencia para el banco central. Si la CBDC está diseñada para el público en general, se denomina una CBDC de retail o propósito general. De esta forma, ofrece una nueva opción para el público general para almacenar valor y realizar pagos. Una CBDC es diferente a formas actuales de pago sin efectivo para consumidores y negocios, tales como transferencias crédito, débitos directos, pagos con tarjeta e e-money, dado que representa una reclamación directa en el banco central en lugar que una acreencia para una institución financiera. En contraste, una CBDC wholesale se enfoca en un grupo de usuarios finales distinto, las instituciones financieras. Una CBDC wholesale es similar a la actual reserva del banco central”, (Koss, Mattei, 2022)

ⁱⁱⁱ Un *stablecoin* “tal como Tether y USD Coin, son una categoría de criptodivisas que buscan mantener un valor estable al atar su valor a uno o más activos, tales como una moneda fiat, un commodity u otra criptoactivo. FSB (2020). A pesar de compartir diversos diversos riesgos LAFTFPADM con los AV tradicionales, tales como el potencial anonimato de transacciones y su alcance global, ciertas stablecoins

Gestión del riesgo LAFTFPADM de jurisdicciones para AV y PSAV a través de aprendizaje de máquinas no supervisado

pueden tener mayor potencial de adopción masiva, lo cual puede aumentar los riesgos LAFTFPADM, particularmente, la reducción de volatilidad de precio (FATF, 2021B).

^{iv} Una transacción P2P es una “transferencia de AV llevada a cabo sin que se involucre un PSAV u otra entidad obligada, p.e., una transferencia de AV entre dos billeteras no hosteadas cuyos usuarios operan en su propio nombre. Las Transacciones P2P no son sujeto explícito de controles ALA/CFT bajo las Recomendaciones GAIFI”. (FATF, 2021B)

^v La minería de datos es “el estudio de la recolección, limpieza, procesamiento y extracción de conocimiento a partir de diferentes conjuntos de datos” (Gafilat, 2021a). La minería de datos también puede entenderse como “un proceso no trivial de identificación válida, novedosa, potencialmente útil y entendible de patrones comprensibles que se encuentran ocultos en los datos, que a su vez, facilita la toma de decisiones y emplea técnicas de aprendizaje supervisado y no-supervisado” (UIAF, 2014b).