UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN

Reporte De Solución De Caso De Estudio En La Que Presente Objetivo, Alcance, Justificación De La Metodología Y Planeación De Las Etapas Para El Análisis De Datos

MATERIA: Extracción de Conocimiento en Bases de Datos

MAESTR@: Enrique Mascote

ALUMNOS: Kevin Iván Aguirre Silva

Jatzel Israel Cruz Castruita

Jorge Alejandro Hernández Contreras

Carlos Adrián Mata Nevárez

Erick Adrián Sánchez Cervantes

GRUPO: IDGS91N

FECHA:27/09/2025

ÍNDICE

INTRODUCCIÓN

El Objetivo del proyecto será diseñar un sistema de detección de fraude que identifique, en tiempo casi real, posibles transacciones sospechosas en la plataforma de intercambio CryptoSecure.

La finalidad de este proyecto es reducir el riesgo asociado a hackeos, lavado de dinero y operaciones anómalas, garantizando mayor seguridad en el ecosistema de criptomonedas.

En este sentido, el uso de técnicas avanzadas de análisis de grafos, algoritmos de clustering y sistemas de detección de anomalías se convierten en una solución estratégica. Estas herramientas permiten modelar relaciones entre direcciones digitales, reconocer patrones ocultos y generar alertas cuando surgen operaciones que se desvían de lo esperado. Se busca implementar un modelo de análisis de grafos que permita agrupar direcciones relacionadas, detectar patrones irregulares y emitir alertas automáticas cuando una nueva transacción se desvíe significativamente de los comportamientos previamente aprendidos.

El presente trabajo propone un plan de acción basado en la metodología CRISP-DM, ampliamente utilizada en proyectos de minería de datos. Esta metodología ofrece una estructura iterativa y flexible que permite abordar desde la comprensión del problema de negocio hasta el despliegue de un prototipo funcional.

Con ello, se busca demostrar que es posible construir un sistema integral para identificar posibles fraudes en el entorno de las criptomonedas, maximizando la seguridad y anticipándose a amenazas emergentes en un campo donde la innovación tecnológica y la ciberseguridad van de la mano.

ASIGNACIÓN DE CASOS A EQUIPOS

Equipo 8 → Caso 8: Detección de fraude en criptomonedas

Caso 8. Detección de fraude en transacciones de criptomonedas

La plataforma de intercambio CryptoSecure recibe miles de transacciones por minuto entre distintas monedas digitales. Dada la naturaleza pseudónima de blockchain, quieren un sistema que identifique transacciones sospechosas (posibles hackeos, lavado de dinero) en tiempo casi real. Los datos incluyen historiales de direcciones, montos, patrones de red (clusters en la cadena) y reputación de contrapartes. El equipo debe investigar métodos de análisis de grafos (graph embeddings), algoritmos de clustering en batch para agrupar direcciones, y un componente de streaming que alerte cuando nuevas transacciones se desvían de patrones aprendidos. Deben planificar las fases de extracción de datos de nodos de blockchain, transformación, modelado y despliegue de alertas.

OBJETIVO DEL PROYECTO

El objetivo del proyecto es diseñar un sistema de detección de fraude que identifique, en tiempo casi real, transacciones sospechosas en la plataforma de intercambio CryptoSecure. La finalidad es reducir el riesgo asociado a hackeos, lavado de dinero y operaciones anómalas, garantizando mayor seguridad en el ecosistema de criptomonedas.

En concreto, se busca implementar un modelo de análisis de grafos que permita agrupar direcciones relacionadas, detectar patrones irregulares y emitir alertas automáticas cuando una nueva transacción se desvíe significativamente de los comportamientos previamente aprendidos.

ALCANCE

El alcance del proyecto se centra en el diseño de un sistema capaz de detectar fraudes en transacciones de criptomonedas en la plataforma CryptoSecure. Para ello, se trabajará con datos históricos extraídos directamente de la blockchain, incluyendo información de transacciones, montos, frecuencia de movimientos, patrones de red obtenidos mediante análisis de grafos y, en los casos en que sea posible, reputación de las contrapartes involucradas.

El proyecto no abarca cuestiones legales o regulatorias, ni pretende ofrecer una integración completa en un entorno de producción real. Asimismo, se excluye el análisis de criptomonedas emergentes con baja actividad en la red, ya que se priorizarán las blockchains más utilizadas. Tampoco se plantea la erradicación total del fraude, ya que la propuesta se enfoca en un sistema de apoyo para reforzar la seguridad de la plataforma.

Aspecto	*Incluye*	*Excluye*	*Limitaciones*
Fuentes de datos	Historial de transacciones en blockchain, patrones de red (análisis de grafos), reputación de contrapartes.	Criptomonedas emergentes con baja actividad.	Gran volumen y complejidad de los datos que incrementan los costos de cómputo y almacenamiento.
Tipo de análisis	Clustering en batch, detección de anomalías en streaming y modelos de grafos para encontrar patrones irregulares.	Integración legal o normativa de resultados.	Nivel de anonimato en blockchain que impide identificar usuarios reales detrás de direcciones.
Prototipo	Desarrollo de un sistema de detección y generación de alertas en tiempo casi real con métricas de evaluación (precisión, recall, AUC).	Implementación completa en ambientes de producción reales de la plataforma.	Recursos técnicos y de hardware limitados para pruebas en escenarios de alta escala.
Objetivo final	Proporcionar una herramienta de apoyo que mejore la seguridad y reduzca riesgos de fraude en la plataforma *CryptoSecure*.	Prevención total y definitiva del fraude (no es posible garantizarla).	Necesidad de actualización continua del modelo debido a la evolución de tácticas fraudulentas en el ecosistema cripto.

JUSTIFICACIÓN DE LA METODOLOGÍA

Se eligio esta tecnologia por que permite dar un enfoque interativo entre la comprension del negocio y la experimentacion tecnica, se adapta bien a proyectos de deteccion de anomalias donde la preparacion de datos y el modelado requieren ciclos de ajuste, una parte importente de esta tecnologia es que es flexible y ampliamente documentada, lo cual reduce riesgos en la planeación (ibm, 2021).

PLANEACIÓN DE ETAPAS

Fase 1. Comprensión del negocio

Actividades:

- Entrevistar stakeholders de CryptoSecure (equipo de seguridad y operaciones).
- Definir criterios de “transacción sospechosa” junto con expertos.
- Identificar riesgos y casos de uso (hackeo, lavado de dinero, fraude interno).

Entregables: Documento de requisitos y casos de uso de fraude.
Cronograma: Semana 1.

Fase 2. Comprensión de los datos

Actividades:

- Extraer datos históricos de blockchain (direcciones, montos, contrapartes).
- Identificar atributos relevantes: montos, patrones de red, reputación.
- Realizar análisis exploratorio y estadísticas descriptivas.

Entregables: Reporte de análisis exploratorio de datos (EDA) con gráficas y métricas.
Cronograma: Semanas 2–3.

Fase 3. Preparación de los datos

Actividades:

- Construcción de grafos de transacciones.
- Generación de features: embeddings de nodos, métricas de centralidad, ratios de actividad.
- Limpieza y normalización de datos (escalado de montos, manejo de outliers).

Entregables: Dataset transformado y almacenado en Feature Store.
Cronograma: Semanas 4–5.

Fase 4. Modelado

Actividades:

- Entrenar algoritmos de clustering (HDBSCAN, Louvain) para agrupar direcciones.
- Probar embeddings de grafos (Node2Vec, GraphSAGE).
- Implementar un modelo de detección de anomalías en streaming (Isolation Forest / Autoencoder).

Entregables: Prototipo de modelo con métricas de validación.
Cronograma: Semanas 6–8.

Fase 5. Evaluación

Actividades:

- Medir desempeño: precisión, recall, tasa de falsos positivos.
- Validar con casos históricos simulados de fraude.
- Retroalimentación de stakeholders sobre calidad de las alertas.

Entregables: Reporte de evaluación de modelos.
Cronograma: Semana 9.

Fase 6. Despliegue

Actividades:

- Implementar componente de streaming para alertas en tiempo real.
- Configurar dashboard básico con alertas y métricas.
- Documentar el sistema y capacitar a analistas.

Entregables: Sistema de alertas prototipo + manual técnico.
Cronograma: Semanas 10–12.

CONCLUSIÓN

El desarrollo del presente proyecto evidencia la viabilidad de implementar un sistema de detección de fraude en transacciones de criptomonedas mediante el uso de técnicas de análisis de grafos, algoritmos de clustering y mecanismos de detección de anomalías en tiempo casi real.

La aplicación de la metodología CRISP-DM permitió estructurar las fases del trabajo de forma clara, desde la comprensión del negocio hasta la propuesta de despliegue, garantizando un enfoque sistemático y flexible.

Si bien el alcance planteado reconoce limitaciones como la exclusión de aspectos legales y la falta de integración con sistemas en producción, los resultados esperados ofrecen un valor significativo para la plataforma CryptoSecure: mejorar la seguridad, reducir riesgos de hackeos y detectar operaciones irregulares antes de que generen pérdidas mayores.

Este proyecto abre la puerta a futuras ampliaciones que incluyan mayor volumen de datos, integración con sistemas regulatorios, incorporación de criptomonedas emergentes y optimización mediante arquitecturas más avanzadas de aprendizaje automático.

Con ello, se sienta una base sólida para evolucionar hacia un sistema integral de ciberseguridad financiera en el dinámico ecosistema de las criptomonedas.

REFERENCIAS

(17 de Ago de 2021). Obtenido de ibm: https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview