SEON

Guía para el Análisis de Fraude en 2022

SEON Technologies Ltd.
seon.io

info@seon.io
+44 20 8089 2900

Table de Contenidos

Una de las piedras angulares de toda buena estrategia de prevención de fraude es el uso de la analítica de datos.

Al examinar los datos pasados con métodos analíticos, podemos distinguir las características que tienen más probabilidades de ser fraudulentas, y podemos elaborar normas, medidas y procedimientos para proteger a nuestra empresa contra ellas.

La principal ventaja del uso de la analítica sobre el examen de incidentes individuales es que cualquier información obtenida durante el proceso de análisis puede utilizarse para examinar el sistema y descubrir casos potencialmente perjudiciales que podrían haberse pasado por alto inicialmente.

Esta guía repasará los principios del análisis de fraude, permitiéndote descubrir patrones de fraude que de otro modo permanecerían ocultos.

Al invertir en el enfoque analítico, puedes ahorrar dinero a largo plazo al prevenir fraudes que, de otro modo, se habrían convertido en pérdidas directas antes de que se materializaran.

¿Qué es el Análisis de Fraude?

El análisis del fraude es el uso de la analítica de datos en el contexto de la prevención del fraude.

Esto va desde la simple observación de los datos disponibles para encontrar anomalías, ya que las señales fuera de los rangos esperados o normales se consideran sospechosas o de riesgo, hasta las técnicas de big data y el machine learning para detectar y prevenir automáticamente el fraude a escala.

Este último punto es crucial: cualquier servicio o negocio en línea tendrá que apoyarse en la analítica para luchar contra el fraude, ya que el número de transacciones y eventos que pasan por un sistema es sencillamente demasiado grande para que los humanos lo analicen manualmente.

El análisis de fraude es la herramienta de poder metodológico para **descubrir rápidamente áreas de interés,** y el uso de las estadísticas permite establecer reglas para **prevenir eventos indeseables,** así como ayudar al diseño del sistema y del servicio para **minimizar el riesgo.**

El enfoque principal consiste en combinar tanto los datos de inteligencia empresarial como otros datos internos sobre las acciones y transacciones de los clientes e identificar patrones para construir modelos.

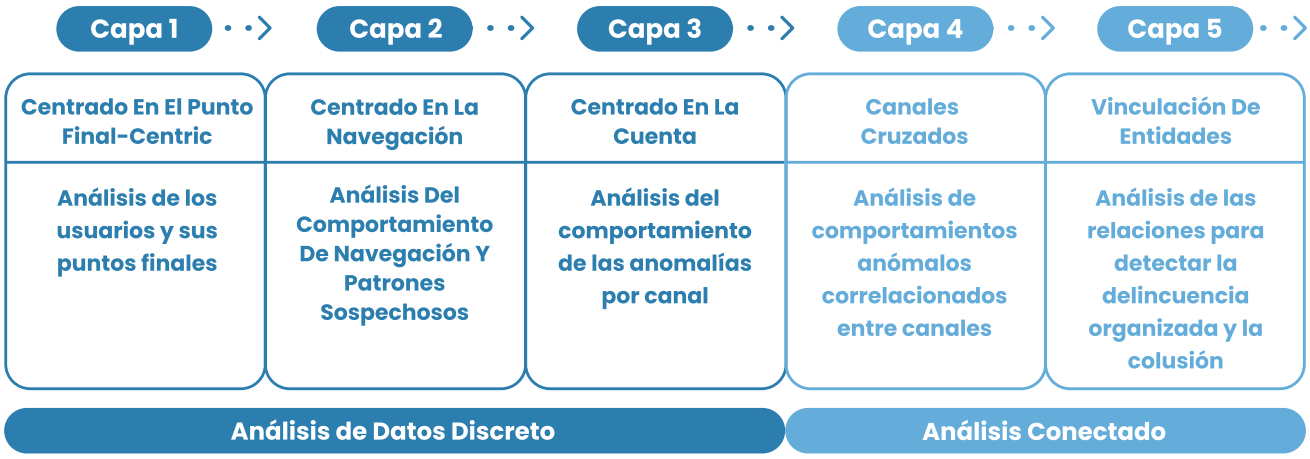
A medida que una empresa evoluciona, dispondrá de más y más datos, a menudo aislados en diferentes sistemas, y el análisis de fraude necesita reunir estos diferentes conjuntos de datos en un solo grupo para establecer conexiones significativas.

Cómo Utilizar el Análisis de Datos para la Detección del Fraude

Los expertos recomiendan utilizar un enfoque por capas para utilizar el análisis de datos para la detección de fraude, basándose en los datos disponibles en diferentes etapas y niveles de agregación.

La idea es que, utilizando el análisis de datos, podemos detectar anomalías en diferentes niveles y con diferente granularidad, lo que hace realmente difícil que incluso los mejores ataques pasen desapercibidos.

Enfoque De Prevención De Fraude Por Capas De Gartner



Estas cinco capas, definidas por Gartner, son:

> **centrado en el punto final:** analizar a los usuarios en sus puntos finales (dispositivos, información de registro, etc.)

> **centrado en la navegación/comportamiento:** examinar los datos de uso del servicio para detectar anomalías/extraños

centrado en la cuenta: buscar comportamientos sospechosos a nivel de canal (por ejemplo, por cohortes de usuarios, mercados)

canales cruzados: examinar cómo se correlacionan los comportamientos anómalos en diferentes canales (por ejemplo, comparando los usuarios orgánicos con los del programa de afiliados)

vinculación de entidades: análisis en profundidad de las conexiones entre diferentes entidades del sistema para descubrir ataques complejos

Esto es útil a la hora de pensar en cómo agrupar diferentes niveles de agregación de datos para realizar el análisis. Piensa en ello como si compararas manzanas con manzanas para detectar una naranja.

Al agregar los datos en estos niveles, podemos establecer fácilmente desencadenantes para la actividad que se encuentra fuera de lo que estadísticamente se espera que ocurra, de una manera que tiene sentido en el contexto dado.

Por ejemplo, si quisiéramos detectar inscripciones fraudulentas o malas,

1. **En primer lugar, observaríamos el aspecto de nuestros buenos usuarios en su fase de registro (en un mercado determinado o en un segmento bien definido).**
2. **A continuación, observaríamos el aspecto de los malos usuarios (los que resultan en transacciones de devolución de cargos, por ejemplo).**
3. **Compararíamos las dos cohortes para encontrar diferencias.**
4. **Tras el análisis, podríamos identificar ciertas diferencias que pueden trasladarse a las cinco capas de agregación de datos, en las que los malos usuarios son valores estadísticos atípicos en comparación con los buenos.**

Utilizando esta información, ahora podemos elaborar reglas y disparadores para alertarnos de posibles riesgos no solo cuando veamos nuevos usuarios que encajen en el modelo del grupo “malo”, sino también para levantar sospechas cuando veamos algo que no está claramente en el grupo “bueno”.

La forma de hacerlo va desde el descubrimiento burdo a través de cuadros de mando de BI construidos a medida hasta la aplicación de modelos de ciencia de datos en grandes conjuntos de datos.

Lo importante es **asegurarse de que los datos son comparables** y tener una buena idea de **cómo se relacionan las diferentes etapas.**

Lo que puede no ser arriesgado en una etapa puede serlo en otra y, a la inversa, lo que parece positivo en un nivel puede ser problemático en otro. Siempre que tu enfoque sea sistemático, deberías estar bien protegido sin ser demasiado estricto.

Técnicas de Análisis de Datos para la Detección de Fraude

Existen diferentes tipos de técnicas de análisis de datos que puedes utilizar para analizar tus datos utilizando estadísticas. No hay un enfoque único que sirva para todo, y deben utilizarse conjuntamente.

Al igual que la filosofía del enfoque por capas, los distintos métodos darán resultados diferentes, cubriendo todos los ángulos posibles para detectar el fraude.

Aprendizaje No Supervisado (Análisis Descriptivo)

Describen métodos estadísticos para descubrir automáticamente valores atípicos en tu conjunto de datos para encontrar fraude o comportamientos de riesgo.

1. **Método estadístico de detección de valores atípicos:**

   Puntuación estándar/puntuación z: el cálculo de la distancia del valor dado con respecto a la media de todo el conjunto de valores puede convertirse en sí mismo en una puntuación de riesgo.

   Punto de ruptura: se observan los cambios repentinos en el comportamiento de la entidad en cuestión.

   Análisis del grupo de pares: cambios repentinos en comparación con los pares.

   Recencia, frecuencia, valor monetario: comparar si una entidad dada difiere del RFM esperado.

   Aprendizaje de reglas de asociación: descubrir asociaciones ocultas entre diferentes variables utilizando reglas “si, entonces”.

2. **Algoritmos de clustering**

Hay muchos tipos de algoritmos de agrupación disponibles, y todos ellos tienen como objetivo agrupar un conjunto de objetos basándose en su similitud entre sí en algún sentido y sus diferencias con otros.

Cuando se trata de la detección de fraude, se dispone de una plétora de puntos de datos para este tipo de agrupación, y esto permite descubrir patrones de fraude que un analista humano podría pasar por alto.

Análisis de Redes Sociales y Gráficos de Fraude

Los pájaros se juntan, como dice el refrán, y esto es doblemente cierto para el fraude.

El análisis de redes sociales significa buscar conexiones de usuarios tanto dentro de tu sistema basándose en puntos de datos (dispositivos, direcciones IP, números de teléfono, correos electrónicos y muchos más), así como utilizar OSINT para trazar conexiones fuera de tu sistema entre tus usuarios.

Las bandas criminales organizadas que tienen una seguridad operativa deficiente probablemente tendrán rastros en línea de sus conexiones en el mundo real.

Análisis Predictivo de Fraude

Se trata del uso de datos históricos y modelos estadísticos para predecir resultados futuros.

Cómo el Análisis Predictivo Ayuda a la Lucha Contra el Fraude

La aplicación de herramientas analíticas a tus datos históricos te permite crear modelos predictivos que pueden señalar con precisión comportamientos sospechosos aplicando

puntuaciones de riesgo y, si la puntuación está por encima de lo aceptado, rechazar automáticamente un registro o una transacción.

Hay un proceso sencillo de cinco pasos a seguir para implementar esto.

1 Etiquetar tu base de datos histórica

Para esto también puedes usar una muestra pero, si puedes, deberías usar todos tus datos disponibles. Se trata simplemente de clasificar las acciones y transacciones históricas como buenas o malas (fraude o no), para que sirvan de entrada a tus modelos de análisis o aprendizaje.

Con SEON, puedes utilizar la API de etiquetas para hacer esto fácilmente, utilizando las transacciones que resultaron ser fraudulentas para entrenar el modelo de machine learning. A continuación, sugerirá automáticamente reglas basadas en los datos proporcionados.

2 Análisis de patrones y variables de fraude

No todas las variables son iguales, y este paso sirve como control de cordura. Descarta la información que se distribuye de forma tan equitativa que no sirve como indicador significativo, ya que solo añadirá ruido a tu modelo.

3 Modelado de datos

Utilizando la metodología proporcionada anteriormente, ahora puedes encontrar valores atípicos o acciones de riesgo que no habías considerado anteriormente. Con estos nuevos conocimientos, deberías ser capaz de señalar cuáles son los indicadores fuertes de fraude, lo que te permitirá establecer reglas contra él.

4 Implementación del modelo

Dentro de SEON, esto se lleva a cabo ya sea estableciendo nuevas reglas basadas en tus hallazgos o utilizando la función de filtrado para descubrir transacciones en tu sistema que se ajusten a tus criterios con el fin de descubrir el fraude.

5 Monitorización y retroalimentación

Con el tiempo, tus modelos implementados pueden quedar obsoletos, mientras que el fraude puede cambiar para escapar a la detección.

Por ello, es necesario supervisar los resultados de vez en cuando y realizar los ajustes pertinentes. De nuevo, la API de etiquetas puede resultar útil para actualizar el sistema de acuerdo con los nuevos descubrimientos, ajustando los falsos positivos y negativos.

Soluciones de Análisis de Datos para la Prevención del Fraude

Hay una gran cantidad de herramientas disponibles para elegir cuando se trata de análisis de datos. La pregunta es: ¿qué objetivos quieres alcanzar?

Enfoque A: Cuadros de mando

El primer enfoque consiste en crear cuadros de mando para tus departamentos que muestren los datos agregados de forma que sea fácil encontrar las anomalías. Para ello, puedes utilizar Tableau, Google Data Studio u otro software de elaboración de informes, que tu equipo de BI puede ayudar a montar.

Lo ideal es que la creación de estos cuadros de mando te permita tener **una visión agregada de las actividades,** lo que facilitará la realización de auditorías y controles rutinarios. Recuerda que el fraude puede pasar desapercibido en un nivel de detección, solo para aparecer como un comportamiento inesperado en un nivel de abstracción diferente.

Este enfoque te protege tanto del problema de perderte en los detalles como de contrarrestar los problemas de silos de datos que tienden a formarse a nivel departamental.

Enfoque B: Impulsado por el ser humano

El segundo enfoque sería confiar en tu equipo de ciencia de datos, que tendrá sus propios entornos y herramientas favorables, como R o Python pandas. Se puede confiar en ellos para que construyan sus propios modelos basados en datos históricos, trabajando conjuntamente con tu equipo de prevención de fraude hacia objetivos comunes.

En este caso, el enriquecimiento de datos de SEON de las direcciones IP, los correos electrónicos y los números de teléfono puede descubrir conexiones ocultas en el conjunto de datos añadiendo más variables que pueden ser significativas, como los círculos de fraude que resultan no tener presencia en las redes sociales, o las cuentas

que están conectadas a través de sus patrones de convención de nombres de correo electrónico, etc.

Enfoque C: impulsado por el machine learning

La tercera forma es dejar que la inteligencia artificial haga el análisis por ti.

El módulo de machine learning whitebox de SEON, si se entrena adecuadamente a través de la API de etiquetas, recomendará reglas automáticamente en función de los patrones que considere de riesgo. El razonamiento es transparente y legible para el ser humano, y viene con un comprobador de reglas para que puedas determinar si es preciso o no. Puedes leer más sobre su funcionamiento en nuestra documentación.

SEON

Para ver cómo SEON puede ayudar a su empresa a prepararse para el futuro, visite seon.io

O programe ahora una llamada de presentación de productos personalizada.

Visite nuestro sitio web

Programe una llamada

SEON Technologies Ltd.
seon.io

info@seon.io
+44 20 8089 2900