

How machine learning works for payment fraud detection and prevention

Get started with Stripe >

Last updated 27 June 2023

Introduction

What is machine learning?

How is machine learning used in fraud prevention and detection?

Machine learning fraud certification

Examples of machine learning for fraud detection

In-person payments

Mobile payments

E-commerce

Other relevant use cases

Get started with Stripe >

Global online payment fraud losses in 2022 reached US\$41billion, a figure expected to balloon to US\$48 billion by the end of 2023. Combating payment fraud – and mitigating its devastating financial and reputational damage – has become a top priority for businesses. Beyond payment fraud's immediate financial losses, businesses also face potential erosion of customer trust and loyalty, as well as increased scrutiny from regulators and law enforcement agencies. To combat this growing threat, organisations are turning to machine learning.

Machine learning, a subfield of artificial intelligence (AI), offers a powerful and adaptive solution to tackle the complex and evolving nature of payment fraud. By mobilising large datasets and advanced algorithms, machine learning can identify patterns and anomalies that indicate fraudulent behaviour, making it possible for businesses to detect and prevent fraud in real time. Ultimately, machine learning can help businesses uphold a secure environment around payments to protect their customers, revenue, and reputation.

We'll cover the benefits of machine learning for fraud prevention and how businesses can use this tool in different payment scenarios.

What's in this article?

- What is machine learning?
- How is machine learning used in fraud prevention and detection?
- Machine learning fraud certification
- Examples of machine learning for fraud detection

What is machine learning?

Machine learning is a subfield of AI that focuses on developing algorithms and models that give computers the ability to learn from data, identify patterns from within the data, and make decisions based on their learnings.

There are three main types of machine learning:

Supervised learning

Supervised learning is a type of machine learning in which a computer is taught to make predictions or decisions based on examples. Think of it like a student learning from a teacher: the teacher provides the student with a set of problems and correct answers to those problems, and the student studies these examples, learning to recognise patterns. When the student faces a new problem, they can use their previous knowledge to find the correct answer.

In supervised learning, the computer algorithm is given a dataset with both the input data (problems) and the correct output (answers). The algorithm studies this dataset and learns the relationship between the input and output. Eventually, the algorithm can make predictions or decisions for new data that it has not seen before.

Unsupervised learning

Unsupervised learning is a type of machine learning in which a computer learns to identify patterns or structures in data without being given any specific examples or correct answers. This is similar to how a detective would try to solve a case without any initial leads – by looking for clues and connections in the available information to uncover hidden patterns or relationships. In unsupervised learning, the computer algorithm is given a dataset with only input data, without any corresponding correct outputs (answers). The algorithm's job is to analyse this data and discover underlying patterns.

Reinforcement learning

Reinforcement learning is a type of machine learning in which a computer learns to make decisions by interacting with an environment and receiving feedback in the form of rewards or penalties. Think of how you might train a dog to perform tricks. When the dog performs the trick correctly, you offer it a treat (a reward), and when the dog doesn't do the trick, you may give it a gentle correction (a penalty). Over time, the dog learns to perform the trick correctly to maximise the number of treats it receives.

In reinforcement learning, the computer algorithm, often called an agent, explores an environment and makes decisions. For each decision that it makes, it receives feedback as either a reward or a penalty. The algorithm's goal is to learn the best strategy, or policy, to make decisions that maximise its cumulative rewards over time. It does this through trial and error, adapting and improving its strategy based on the feedback.

Machine-learning techniques are used in many different scenarios, including natural language processing, image and speech recognition, medical diagnosis, financial analysis, and autonomous vehicles.

How is machine learning used in fraud prevention and detection?

Increasingly, machine learning is being used in fraud prevention and detection due to its ability to analyse large quantities of data, identify patterns, and adapt to new information. Some common applications of machine learning in fraud prevention include:

- Anomaly detection**
Machine-learning algorithms can identify unusual patterns or deviations from normal behaviour in transactional data. By "training" on historical data, the algorithms learn to recognise legitimate transactions and flag suspicious activities that may indicate fraud.
- Risk scoring**
Machine-learning models can assign risk scores to transactions or user accounts based on various factors, such as transaction amount, location, frequency, and past behaviour. Higher risk scores indicate a higher likelihood of fraud, enabling organisations to prioritise their resources and focus on specific transactions or accounts that warrant further investigation.
- Network analysis**
Fraudulent actors often collaborate and form networks to carry out their activities. Machine-learning techniques, like graph analysis, can help uncover these networks by analysing relationships between entities (such as users, accounts, or devices) and identifying unusual connections or clusters.
- Text analysis**
Machine-learning algorithms can analyse unstructured text data, such as emails, social media posts, or customer reviews, to identify patterns or keywords that may indicate fraud or scams.
- Identity verification**
Machine-learning models can analyse and verify user-provided information, such as images of identification documents or facial recognition data, to ensure that an individual is who they claim to be and prevent identity theft.
- Adaptive learning**
One of the key strengths of machine learning is its ability to learn and adapt to new information. As fraudulent actors change their tactics, machine-learning models can be retrained on new data, allowing them to stay up to date and better equipped to detect emerging fraud patterns.

Using machine learning in fraud prevention can be a powerful way for organisations to enhance their detection capabilities, reduce the risk of false positives, and improve overall security and customer experience.

