Machine Learning for Fraud Detection

Shiksha Online

Updated on Feb 23, 2023 12:21 IST

Discover the power of Machine Learning for fraud detection.

Machine Learning for Fraud Detection
Order
Machine Learning Model
Fraud Risk Estimate
Accept or Reject

Machine Learning has long been used to solve real-world problems. It is now widely used in all fields, including medicine, e-commerce, banking, and insurance. In this blog, we will see how Machine Learning can be used for fraud detection.

Table of Content

Introduction

In recent years, fraud has been a major issue in the banking, medical, insurance, and information technology sectors. There has been an increase in online transactions via various payment methods such as credit/debit cards, PhonePe, Gpay, Paytm, and so on. As a result, fraudulent activities have increased. Furthermore, fraudsters have honed their skills in identifying and exploiting loopholes. As a result, developing a secure system for authentication and fraud prevention has become a difficult task. This is where machine learning-powered fraud detection algorithms come in handy.

Internet Fraud Types

Let’s take a look at some of the real-world applications of machine learning in fraud cases. You may have encountered these scams in one form or another.

Email Phishing

This is a fraud case in which the fraudsters trick people into responding to an email with their personal information. They can use the information to hack into your system and steal your money.

Machine Learning’s algorithm distinguishes between legitimate and spam email addresses, preventing fraud. They will examine the subject lines, email content, and sender’s email details before classifying the emails as genuine or fraudulent.

What is a Phishing attack?

A cyber attack is an unauthorized attempt to gain unauthorized access to a computer system in order to size, modify, or steal data. Cybercriminals can launch a cyberattack using a...**read more**

Difference Between Phishing and Vishing

The article talks about phishing and vishing, the difference between phishing and vishing, and ways to deal with both issues.

Identity Theft

Criminals steal your identity and access to your bank accounts in this type of fraud. They will change the IDs or passwords, preventing access to these accounts.

Machine Learning ensures no one can change a password or update an account's identity. You will be notified as soon as someone tries to hack into your account: two-factor authentication and other measures, combined with human-like intelligence, help ensure better fraud prevention.

Credit Card Fraud

Fraudsters can access and misuse your credit card information using methods such as phishing. You'd have to pay for the purchases you didn't make. Using Machine Learning to detect credit card fraud can avoid such trade-offs.

Document Forgery

Fraudsters can purchase ID proof of a person and use it to enter a system, use it, and then exit without consequence. Many organisations are vulnerable to this type of fraud because the fraudsters can gain access to their systems by forging an ID document and cheating them.

Machine Learning models can be used to distinguish between a forged and genuine identity, to create a foolproof system.

Advantages of Machine Learning for Fraud Management

Because machines process large datasets much faster than humans, you get the ability to slice and dice massive amounts of data. That is to say:

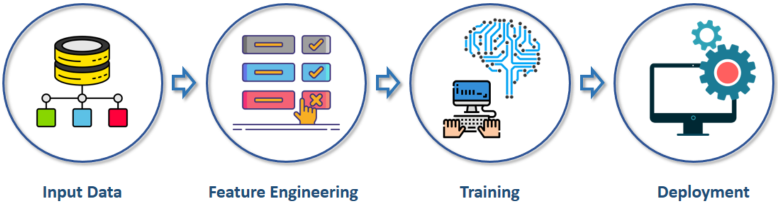
- **Faster and more accurate detection:** The system can quickly identify suspicious patterns and behaviours accurately, which would have taken months for human agents to establish.
- **Reduced manual review time**: Similarly, allowing machines to analyse all data points for you can drastically reduce the amount of time spent manually reviewing information.

- **Larger datasets yield better predictions**: A machine learning engine gets more proficient the more data it is fed. Consequently, while enormous datasets can occasionally make it difficult for people to identify patterns, the situation is exactly the opposite with an AI-driven system.
- **Cost-effective remedy:** You only need one machine-learning system to process all the data you put at it, regardless of volume, as opposed to adding more Risk agents. This is perfect for companies that see seasonal fluctuations in traffic, checkouts, or signups. A machine learning system can help your business grow without significantly raising risk management expenses at the same time.
- **Break-Free availability:** Not to mention, algorithms don't require rest, holidays, or pauses. Even the greatest fraud managers could show up to work on Monday morning with a backlog of manual reviews. Fraud attacks can occur around the clock. By separating situations that are false or acceptable, machines can speed up the process.

How Does A Machine Learning System Work for Fraud Detection?

We can utilise a variety of machine learning techniques; for example, here is a brief explanation of how supervised machine learning works.

Working of a Machine Learning System



Input data

When it comes to detecting fraud, the more data you have, the better.

Note: The data must be labelled for supervised machine learning.

Feature Extraction

User behaviour is described by features, and fraudulent behaviour is identified by fraud signals.

We can group features into main categories, each of which has hundreds or thousands of individual features. For Example:

Identity: The number of digits in the customer's email address, the age of their account, the number of devices on which the customer was seen, and the fraud rate of the customer's IP address are all factors to consider.

Orders: The number of orders they placed in their first week, the number of failed transactions, the average order value, and the contents of risky baskets.

Payment Options: Fraud rate of issuing bank, similarity between customer name and billing name, cards from different countries.

Locations: The shipping address corresponds to the billing address, the shipping country corresponds to the country of the customer's IP address, and the fraud rate at the customer's location.

Networks: Number of emails, phone numbers, or payment methods shared on the network.

Model Training

In machine language, model training is the process of feeding data to an ML algorithm to help identify and learn good values for all attributes involved. There are many different types of machine learning models, the most common of which are supervised and unsupervised learning.

Supervised learning is possible when the training data contains both the input and output values. A supervisory signal is any set of data that contains the inputs and the expected output. When the inputs are fed into the model, the training is done based on the deviation of the processed result from the documented result.

While Unsupervised learning entails identifying patterns in data. Following that, additional data is used to fit patterns or clusters. This, too, is an iterative process that improves accuracy based on correlation to expected patterns or clusters. This method has no reference output dataset.

We can train the algorithm on historical data, which we refer to as a training set. The more fraud in this training set the better so that the machine has many examples to learn from.

The algorithm learns how to make predictions based on customer data described by our features, such as fraud/not a fraud.

Steps to Create Your Own Machine Learning Models

Machine learning involves a series of techniques that allow computer systems to predict, classify, order, make decisions and, in general, extract knowledge from the data without the need to explicitly**...read more**

Anomaly Detection in Machine Learning

In this article we have covered Anomaly detection in machine learning and also suggested how we can recognize it.It also includes the applications of anomaly detection.

Deploy the Model

After training, you will have a model tailored to your company that can detect fraud in milliseconds.

We constantly monitor the model to ensure that it is behaving properly, and we are always looking for ways to improve it. We constantly improve, update, and upload a new model for each client so that the system can detect the most recent fraud techniques.

Is Machine Learning a Replacement for Human Intelligence?

Machine learning does not replace the fraud analyst team, but it does allow them to spend less time on manual reviews and data analysis. This means analysts can focus on the most pressing cases and assess alerts more quickly and accurately, while also reducing the number of genuine customers.

Machines struggle to deal with uncertainty. Some cases are novel, difficult, or unique in some way. Edge cases require more attention and may be difficult to determine – this is where expert human intervention becomes unavoidable.

Conclusion

Companies like PayPal employ an in-house AI engine created using open-source technologies to reveal suspicious behaviour and, more significantly, to identify false alarms from actual fraud.

Machine learning approaches are unquestionably more trustworthy than transaction rules and human inspection. The machine learning systems process a sizable number of transactions in real-time and are dynamic and scalable.

Learn more by taking up machine learning courses.

Contributed by Prerna Singh