Chapter 4:

Fraud Detection: using Machine Learning

The financial services industry and the industries that involve financial transactions are suffering from fraud-related losses and damages. 2016 was a banner year for financial scammers. In the US alone, the number of customers who experienced fraud hit a record 15.4 million people, which is 16 percent higher than 2015. Fraudsters stole about $6 billion from banks in 2016. A shift to the digital space opens new channels for financial services distribution. It also created a rich environment for fraudsters.

If earlier criminals had to counterfeit client IDs, now getting a person's account password may be all that's needed to steal money. Customer loyalty and conversions are affected in both environments, the digital and the physical. According to Javelin Strategy & Research, it takes 40+ days to detect fraud for brick-and-mortar financial institutions. Fraud also impacts banks that provide online payments service. For instance, 20 percent of customers change their banks after experiencing scams.

So, the challenge for industry players is to implement real-time claim assessment and improve the accuracy of fraud detection.

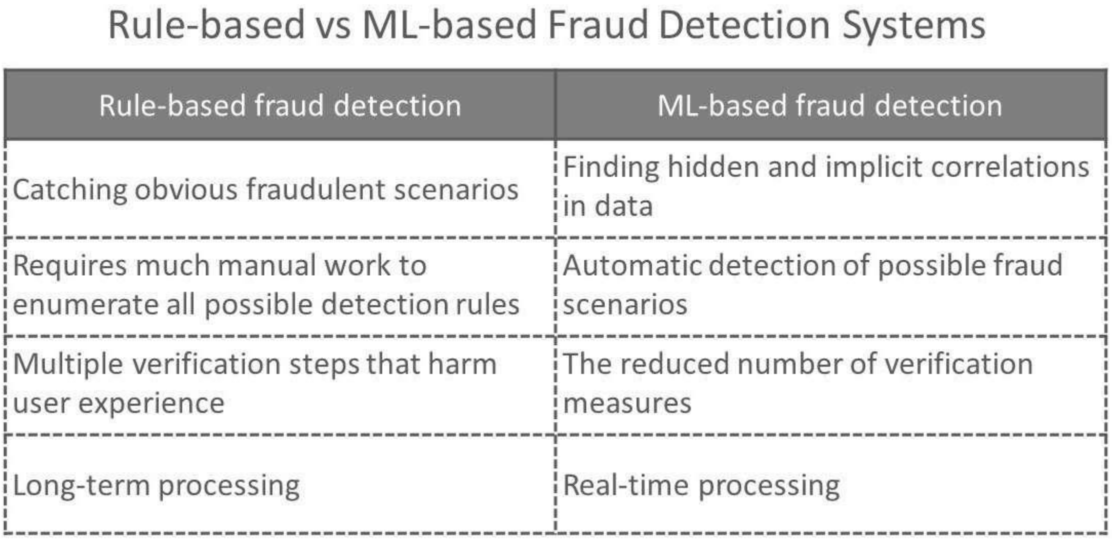
1. Machine learning vs. rule-based systems in fraud detection

The machine learning (ML) approach to fraud detection has received a lot of publicity in recent years and shifted industry interest from rule-based fraud detection systems to ML-based solutions.

What are the differences between machine learning and rule-based approaches?

The rule-based approach. Fraudulent activities in finance can be detected by looking at on-surface and evident signals. Unusually, large transactions or the ones that happen in atypical locations obviously deserve additional verification. Purely rule-based systems entail using algorithms that perform several fraud detection scenarios, manually written by fraud analysts. Today, legacy systems apply about 300 different rules on average to approve a transaction. That's why rule-based systems remain too straightforward. They require adding/adjusting scenarios manually and can hardly detect implicit correlations. On top of that, rule-based systems often use legacy software that can hardly process the real-time data streams that are critical for the digital space.

ML-based fraud detection. However, there are also subtle and hidden events in user behavior that may not be evident but still signal possible fraud. Machine learning allows for creating algorithms that process large datasets with many variables and help find these hidden correlations between user behavior and the likelihood of fraudulent actions. Another strength of machine learning systems compared to rule-based ones is faster data processing and less manual work. For example, smart algorithms fit well with behavior analytics for helping reduce the number of verification steps.



Rule-based vs ML-based Fraud Detection Systems

Rule-based fraud detection	ML-based fraud detection
Catching obvious fraudulent scenarios	Finding hidden and implicit correlations in data
Requires much manual work to enumerate all possible detection rules	Automatic detection of possible fraud scenarios
Multiple verification steps that harm user experience	The reduced number of verification measures
Long-term processing	Real-time processing

Though, while rule-based systems are inferior to ML-driven ones, they still dominate the market.

Leading financial institutions, however, already use the ML technology to combat fraudsters. For instance, MasterCard integrated machine learning and AI to track and process such variables as transaction size, location, time, device, and purchase data. The system assesses account behavior in each operation and provides real-time judgment on whether a transaction is fraudulent. The project aims at reducing the number of false declines in merchant payments. The recent study shows that false declines make merchants lose about $118 billion per year while clients’ loss is about $9 billion per year. It’s the largest area for fraud in financial services. So fraud prevention is a strategic goal for banking and payments industries.

Feedzai, a fintech company, claims that a fine-tuned machine learning solution can detect up to 95 percent of all fraud and minimize the cost of manual reconciliations, which accounts now for 25 percent of fraud expenditures. Capgemini claims that fraud detection systems using machine learning and analytics minimize fraud investigation time by 70 percent and improve detection accuracy by 90 percent.

These facts prove the benefits of using machine learning in anti-fraud systems.

2. Fraud scenarios and their detection

2.1 Insurance claims analysis for fraud detection

Insurance companies spend several days to weeks assessing a claim, but the insurance business is still affected by scams. The most common issues are property damage, car insurance scams, and fake unemployment claims. The ticket to successful detection is a good dataset and carefully selected models.

Fake claims. Semantic analysis is a machine learning task that allows for analyzing both structured, table-type data, and unstructured texts. The feature helps detect fake and falsified claims in the insurance industry. For example, it improves car insurance claims processing. Machine learning algorithms analyze files written by insurance agents, police, and clients, searching for inconsistencies in provided evidence. There are many hidden clues in these textual datasets. The rule-based engines don't catch the suspicious correlations in textual data, and fraud analysts can easily miss important evidence in boring investigation files. That's why analyzing claims is one of the most promising spheres for machine learning applications.

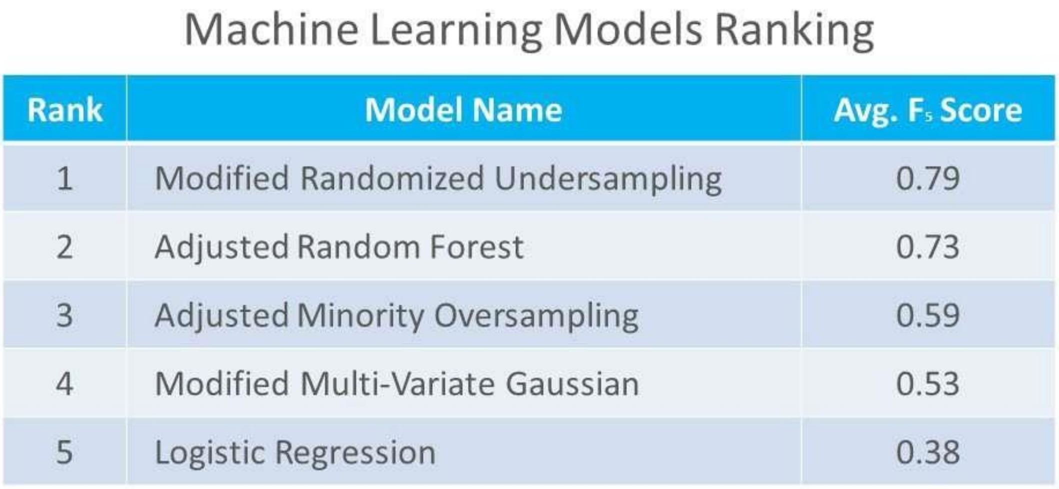
Duplicate claims and overstating repair cost. Smart ML-backed algorithms are also efficient in duplicate claims detection or inconsistencies in car repair cost. Classifying data in repair claims solves the problem by uncovering hidden correlations in claim records or even behaviors of insurance agents, repair services, and clients. For example, the repair service company may provide higher pricing for the customers of a specific agent.

Let's have a look at the results of the AI-based research of insurance vehicles claims conducted by Wipro. The company explored four datasets with such features as a vehicle style, client gender, marital status, license type, injury type, loss date, claim date, police notification date, repair amount, sum insured, market value, etc.

A pre-research analysis disclosed:

- Fraudulent claims are more likely not reported to police.
- Old vehicles are more likely to be involved in fraud.
- Eighty percent of accidents that happen during holidays involve fraud.
- Scams are more likely to involve third parties than legitimate claims.

Then the data was processed using five different machine learning algorithms: Logistic Regression, Modified Multi-Variate Gaussian, Modified Randomized Undersampling, Adjusted Minority Oversampling, and Adjusted Random Forest. Eventually, the best results were achieved by the Modified Randomized Undersampling model that showed 79 percent accuracy.



Machine Learning Models Ranking

Rank	Model Name	Avg. F_5 Score
1	Modified Randomized Undersampling	0.79
2	Adjusted Random Forest	0.73
3	Adjusted Minority Oversampling	0.59
4	Modified Multi-Variate Gaussian	0.53
5	Logistic Regression	0.38

2.2 Anti-fraud solutions for medical claims and healthcare

Healthcare and medical insurance is a rich area for fraud schemes due to a complex and bureaucratic process, which requires many approvals, verifications, and other paperwork. The most common scams are fake claims that use false or invalid social security numbers, claims duplication, billing for medically unnecessary tests, fake diagnosis, etc. Both hospitals and insurance companies are suffering from these issues. Insurance carriers lose money and hospitals take risks being involved in serious crimes, like drug turnover. Multiple data analytics approaches can mitigate such fraud risks.

Upcoding and abuse scams. Upcoding is a specific type of fraud where a healthcare provider or a healthcare worker tries to charge more to a patient or an insurance company. The first step in this case is the implementation of digital analysis based on Benford's Law. It helps reveal unexpected digits in a dataset. The technique detects upcoding of procedures and other abuse attempts.

Machine learning isn't typically required in this case, but it may augment rule-based fraud detection. For example, image recognition techniques can be applied to digitalize paper documentation for further analysis.

Medical receipts and bills. Adding up numerical values helps check sequences of receipts for legal restrictions and find suspicious links between a physician and a patient (or the group of patients) overstating the total limitation for some drugs. Insurance companies can also benefit from this type of analytics by regularly performing reconciliations of bills and thus preventing fake totals.

Personal identity. Image recognition algorithms can also be used for fraud prevention at the personal identification stage. Drug turnover requires a number of verifications, but current supervisory mechanisms aren't ideal in handling this. Machine learning can solve the problem of ID verification by applying face and fingerprint recognition.

2.3 Fraud prevention solutions in eCommerce

The eCommerce scam is closely linked to payments. So, let's discuss the aspects related purely to eCommerce. Typical scams here are identity theft and merchant scams.

Identity theft. This entails a situation in which a scammer breaches a user account, alters personal data, and tries to get money or goods from a retailer using this semi-fake personal information. Nearly all fraud detection tasks here are solved by behavior analytics. Smart algorithms uncover suspicious activities, analyze them, and find inconsistencies in the historical sets of personal data.

Merchant scams. These are related to fraudulent companies or merchants operating through marketplaces. Customers' choices on marketplaces are often driven by reviews. Some fraudsters create fake reviews for their accounts to attract customers. Machine learning algorithms can eliminate the influence of such fraudsters through conducting sentimental and behavior analytics and detecting suspicious activities linked to merchants or their products.

2.4 Fraud detection in banking and credit card payments

Payments are the most digitalized part of the financial industry, which makes them particularly vulnerable to digital fraudulent activities. The rise of mobile payments and the competition for the best customer experience nudge banks to reduce the number of verification stages. This leads to lower efficiency of the rule-based approach. So, banks and payment companies switch to data analytics, machine learning, and AI-driven methods.

Modern fraud detection systems solve a wide range of analytical problems to uncover all scams in the payments streams.

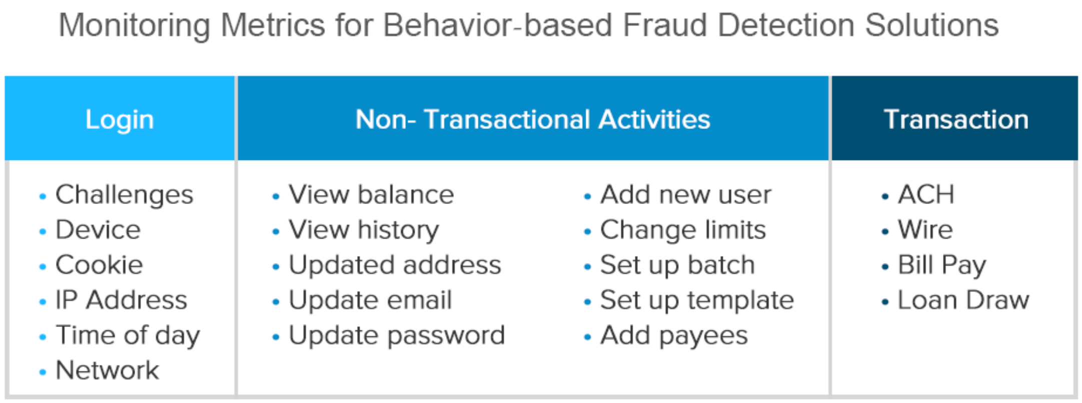
Data credibility assessment. Gap analytics help identify missing values in sequences of transactions. Machine learning algorithms can reconcile paper documents and system data eliminating the human factor. This ensures data credibility by finding gaps in it and verifying personal details via public sources and transactions history.

Duplicate transactions. A common scam method is creating transactions close to original or making a copy of a transaction. For instance, a company tries to charge a counterpart twice with the same invoice by sending it to different branches.

Rule-based systems currently used constantly fail to distinguish between error or unusual transactions from real fraud. For example, a customer can accidentally push a submission button twice or simply decide to buy twice more goods. The system should differentiate suspicious duplicates from human errors.

While duplicate testing can be implemented by conventional methods, machine learning approaches will increase accuracy in distinguishing erroneous duplicates from fraud attempts.

Account theft and unusual transactions. Much of fraud detection in payments is focused on user behavior analysis during transactions.



Monitoring Metrics for Behavior-based Fraud Detection Solutions

Login	Non- Transactional Activities		Transaction
• Challenges	• View balance	• Add new user	• ACH
• Device	• View history	• Change limits	• Wire
• Cookie	• Updated address	• Set up batch	• Bill Pay
• IP Address	• Update email	• Set up template	• Loan Draw
• Time of day	• Update password	• Add payees	
• Network			

For example, a client visits a specific supermarket at 9-10 pm every night. It’s located near the client’s house. The payment sum varies from $10 to $40. Every two days the client also drives to a gas station.

Once a transaction occurs in a different part of town in a bar and the sum is $40, the algorithm will consider this activity suspicious and assign a higher level of fraud likelihood. To check this transaction, the system will send a verification request to a card owner.

Descriptive stats like averages, standard deviations, and high/low values are very useful for analyzing behavior. These metrics allow for comparing separate transactions with personal or intra-group benchmarks. Payments with large standard deviations look suspicious. So, a good practice is to send a request for verification to an account owner if such deviations occur.

2.5 Preventing loan application fraud

Lending is sensitive to scams that abuse personal information. Just a decade ago, it was difficult for fraudsters to get access to IDs, photos, addresses, and mobile phone numbers. Today, nearly all data can be found in social networks or elsewhere on the Internet. This makes the life of financial institutions more difficult. Fraudsters become smarter, and loan applications require more rigorous assessment, while clients want to get money as soon as possible.

Personal details counterfeiting. A common type of fraud is often based on providing false personal information. Scammers provide personal details with a number of misspelling issues or misrepresentations of income or credit qualifications. It makes debt collection more difficult. The problem can be solved in two ways.

The first one entails reviewing the customer relationship history with a bank looking for inconsistencies and quickly verifying record fields via open APIs.

The second path is more sophisticated and requires building a scoring model or calculating fraud probability. Scoring models calculate the fraud possibility of the record and grade it against a standard scale. It helps assess which applications are more likely to be fraudulent. Machine learning and advanced analytics solve the problem of fraud probability assessment by classifying applications into groups.

Such solutions allow for minimizing costs by reducing the need to verify each application and concentrating the efforts on the risky loans only. It also improves general credit scoring – the process of grading customer creditworthiness – by distinguishing fraudsters from bad borrowers. Drawing a distinct line between fraudsters and problematic borrowers also ensures better credit statistics.

An Experian case study with a European bank showed the following benefits of scoring systems:

- Focus on 5 percent of applications allowed for detecting 52 percent of fraud cases;
- Machine learning analytic systems demonstrated 9 times higher detection score than portfolio average.

2.6 Machine learning for anti-money laundering

Regulators, banks, and investment firms are often involved in monitoring possible money laundering activity: They must detect and inform each other about suspicious activities.

For instance, Dr. Miguel Agustín Villalobos and Dr. Eliud Silva describe a case in which a machine learning model was trained on the dataset with transactions conducted by criminals. Such a model combined with the rule-based approach helps discover hidden relations between money movement and criminal activities.

Systems of this type can minimize the workload of small and medium banks involved into monitoring. The described solution revealed 99.6 percent of money laundering transactions and minimized the number of reported transactions from 30 to 1 percent.

3. Common and advanced fraud detection systems

How machine learning applications are built and the common and advanced approaches to creating such fraud detection engines?

3.1 Anomaly detection to reveal suspicious transactions

Anomaly detection is one of the common anti-fraud approaches in data science. It is based on classifying all objects in the available data into two groups: normal distribution and outliers. Outliers, in this case, are the objects (e.g. transactions) that deviate from normal ones and are considered potentially fraudulent.

The variables in data that can be used for fraud detection are numerous. They range from transaction details to images and unstructured texts.

By analyzing these parameters, anomaly detection algorithms can answer the following questions:

1. Do clients access services in an expected way?
2. Are user actions normal?
3. Are transactions typical?
4. Are there any inconsistencies in the information provided by users?

The anomaly detection approach is perhaps the most straightforward as it provides simple binary answers. This may be helpful in some cases. For instance, if the transaction looks suspicious, the system may ask a user to make multiple additional verification steps. Traditional anomaly detection doesn’t allow for revealing fraud, although it may be a good supportive instrument for existing rule-based systems.

But there are more advanced approaches that combine several ML algorithms to reduce uncertainty. They can be implemented using multiple machine learning styles and underlying mathematical models. Let’s have a look at the ones that are most common.

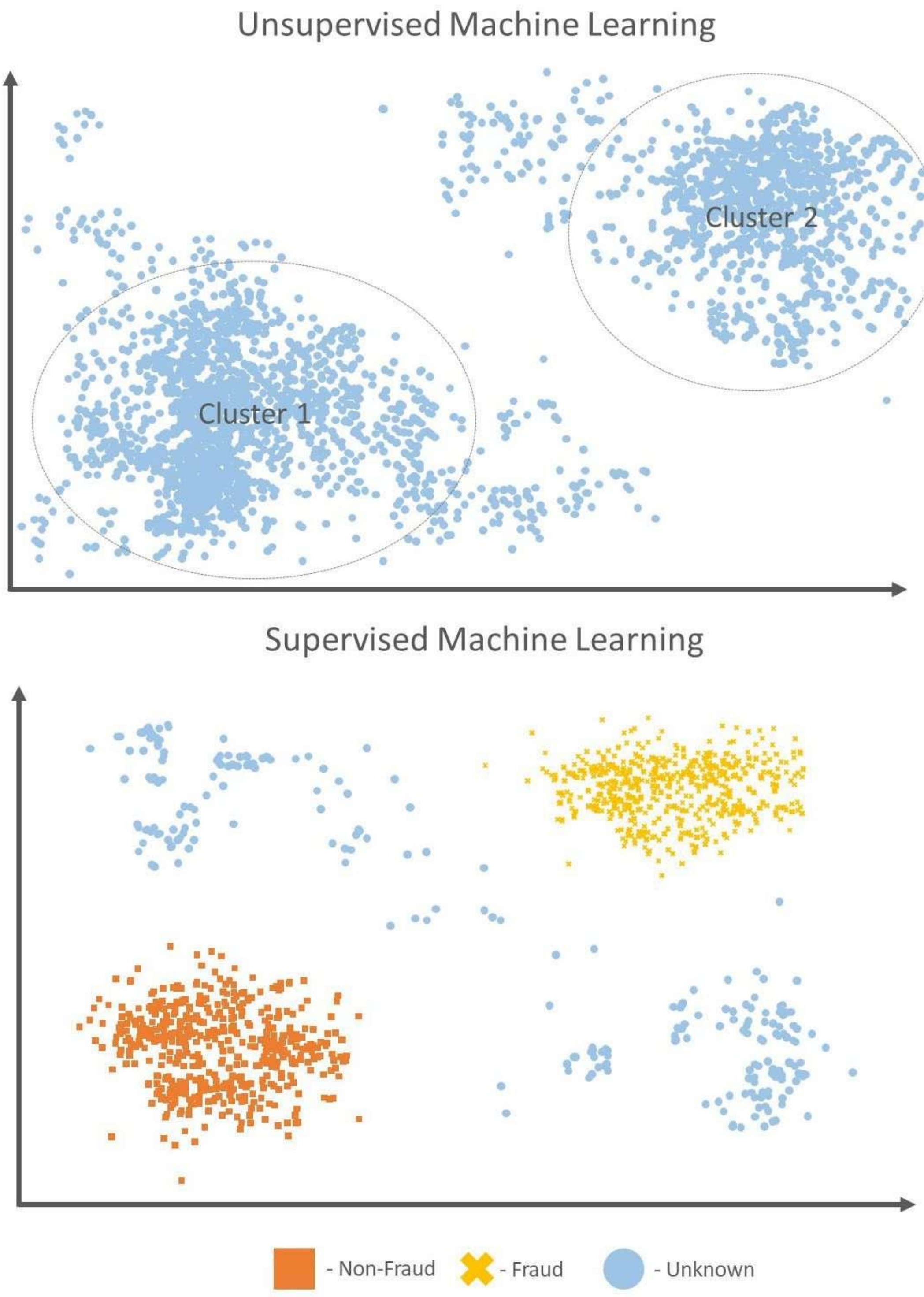
3.2 Advanced fraud detection systems

Advanced systems aren’t limited to finding anomalies but, in many cases, can recognize existing patterns that signal specific fraud scenarios. There are two types of machine learning approaches that are commonly used in anti-fraud systems: *unsupervised* and *supervised machine learning*. They can be used independently or be combined to build more sophisticated anomaly detection algorithms.

Supervised learning entails training an algorithm using labeled historical data. In this case, existing datasets already have target variables marked, and the goal of training is to make the system predict these variables in future data.

Unsupervised learning models process unlabeled data and classify it into different clusters detecting hidden relations between variables in data items.

So, ***how do supervised and unsupervised styles combine to build robust fraud detection systems?***



1. **Labeling data.** While data labeling can be done manually, it's hard for humans to classify new and sophisticated fraud attempts by their implicit similarities. That's why data scientists apply unsupervised learning models to segment data items into clusters accounting for all hidden correlations. This makes data labeling more precise: Not only does the dataset have labeled fraud/non-fraud items, but these labels also nuance different types of fraudulent activities.

2. **Training a supervised model.** Once the data is labeled, the next iteration is to apply this new labeled dataset to train supervised models that will be detecting fraudulent transactions in production use.

3. **Ensembling models.** Ensembling multiple different models is a common approach in data science. While you can make a single model, it will always have its strengths and weaknesses. It will recognize some patterns, but miss the others.

To make predictions more accurate data scientists usually build multiple models using the same method or combine entirely different methods. Thus, all models from the ensemble analyze the same transaction and then "vote" to make a final decision. It allows for leveraging the strengths of multiple different methods and make decision as precise as possible.

4. **Setting an express verification.** Using ensembles requires a great deal of computing power and time to crunch all data. If you were to check all transactions, the time spent on calculations may harm user experience. That's why a good practice is to make verification in two steps. The express verification implies simple anomaly detection or another straightforward ML method to divide all transactions into regular and suspicious ones.

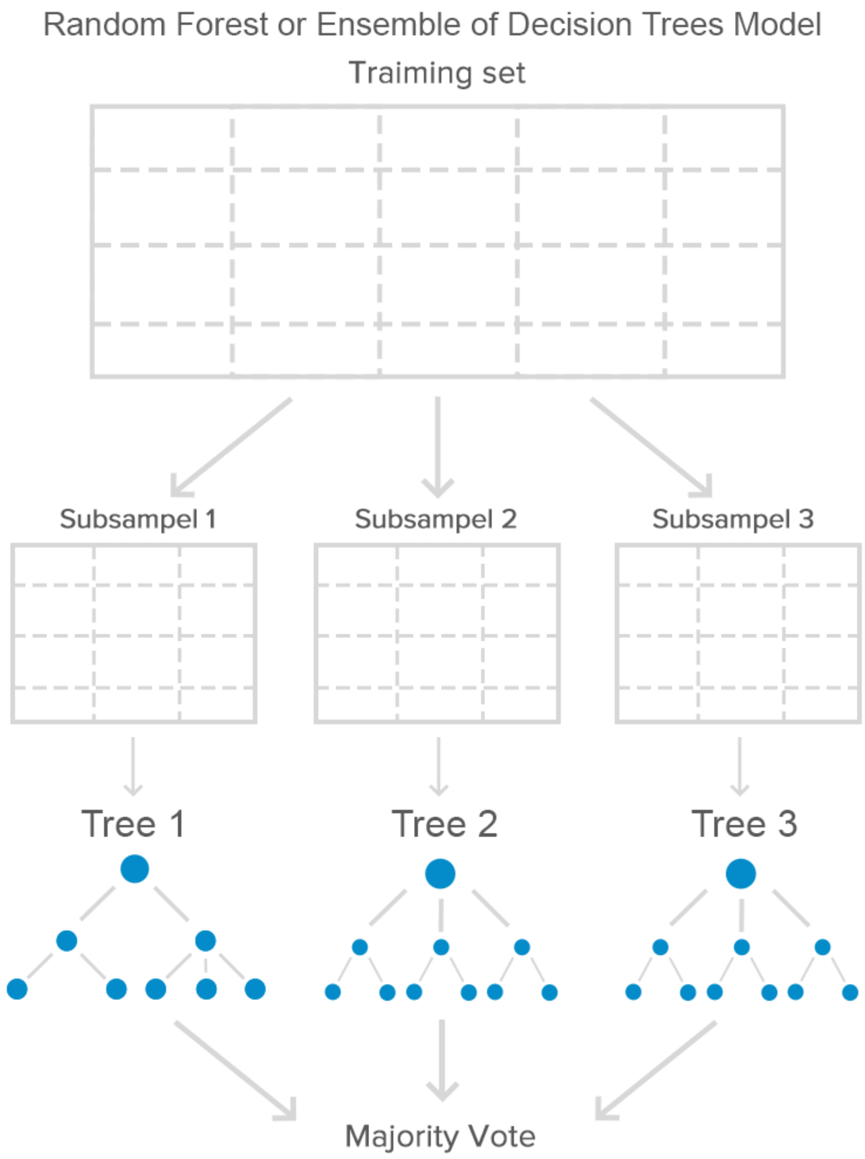
   As regular transactions don't require further verification, the system approves them. Those that look suspicious are sent for advanced verification through a complex ensemble.

4. Supervised fraud detection methods

In terms of actual machine learning methods, there are five commonly used types. We'll cover only supervised learning methods as they can be applied in building complex ensembles.

4.1 Random forests

Random forest (or an ensemble of decision trees) is an algorithm which builds decision trees to classify the data objects. The model selects a variable that enables the best splitting of records and repeats the splitting process multiple times. As a result, if we were to visualize how the algorithm works, the image would look like a tree. To make predictions more precise, data scientists train multiple decision trees on random subsets from a general dataset. To decide whether transaction looks like a fraud, trees vote, and the model provides a consensus judgment.



Random forests are relatively simple systems that you can use to set up fraud detection fast. The most common use case is payments systems.

Pros: Besides their simplicity and speed, random forests can be used with different types of data, including credit card numbers, dates, IP addresses, postal codes, etc. They are considered precise predictors that can work even with datasets that have missing records.

Cons: Sometimes engineers stumble over the problems with overfitting. Overfitting means that the model over-remembers the patterns in the training dataset and fails to make predictions on future data. Another problem is the dataset balance. If a dataset contains mostly normal transactions and just a small fraction of fraudulent ones, the accuracy may decrease.

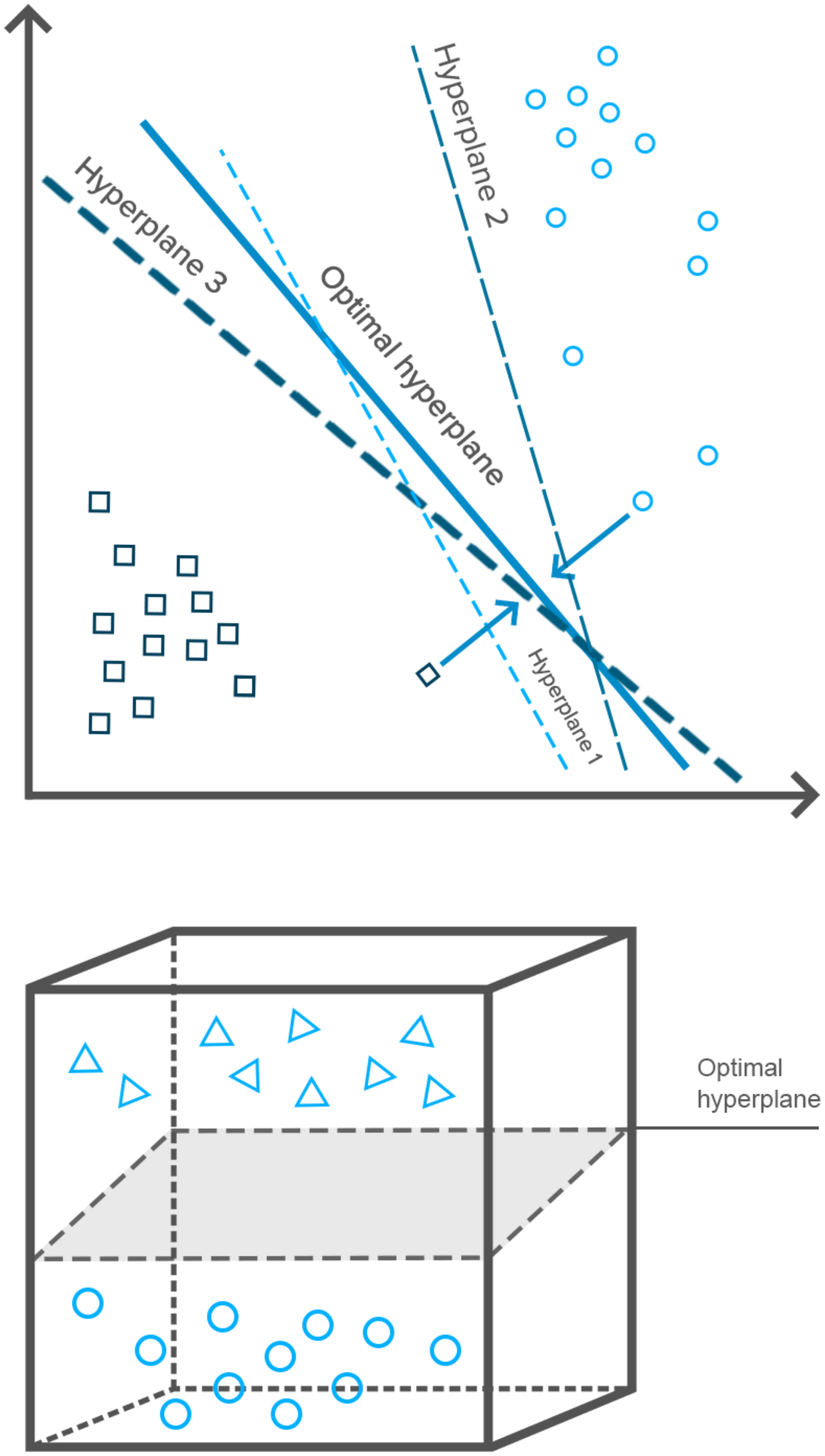
4.2 Support vector machine

A support vector machine (SVM) is a supervised machine learning model that uses a non-probabilistic binary linear classifier to group records in a dataset.

What does it mean?

The algorithm divides data into two categories with a clear gap. The division line is defined by making several hyperplanes in the multidimensional space. Then the algorithm selects the hyperplane which separates records better than the other ones.

Support Vector Machine Model



As some studies show, SVM can be inferior to random forests in credit card transactions with small datasets, but can also approach their accuracy once datasets are large enough.

Pros: Support vector machines are particularly good at working with complex multidimensional systems. They also allow for avoiding the overfitting problem that random forests may experience. Generally, SVM is a very common method in credit card fraud detection. And the abundance of research work makes adjusting SVM-models for credit card fraud detection simpler for a data science team.

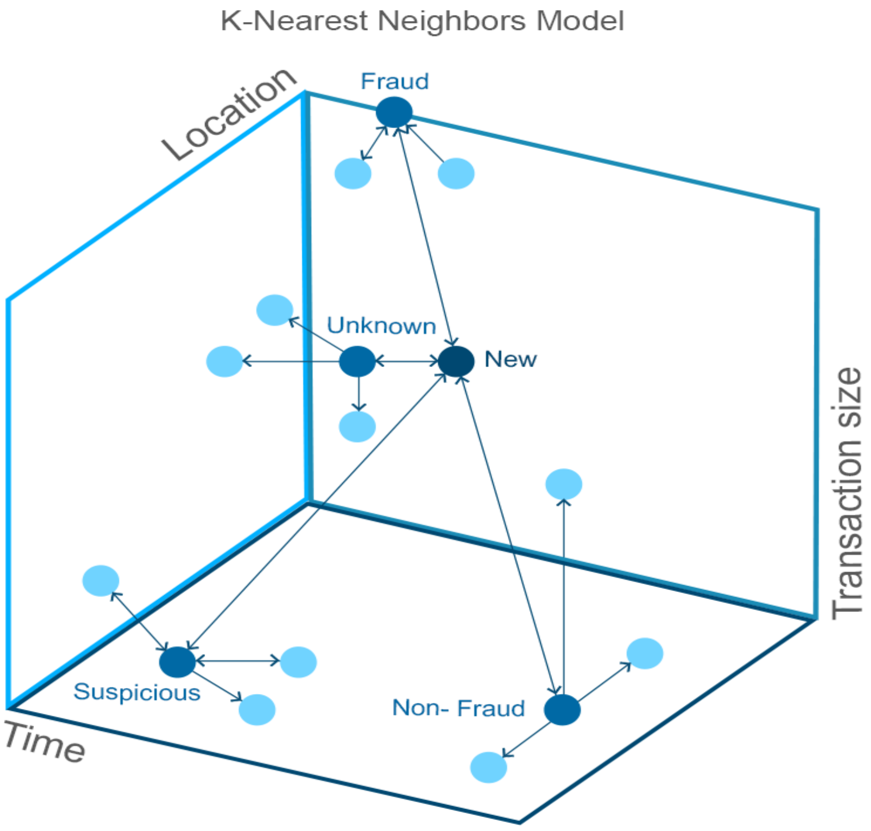
Cons: The complexity of SVM models will require much engineering effort to fine tune the algorithm and achieve high accuracy. Also, SVMs are very slow and computationally heavy. That's why they will require powerful computing architecture.

4.3 K-Nearest Neighbors

K-Nearest Neighbor is an algorithm which classifies records by similarity based on the distance in multidimensional space.

The record is assigned to the class of the nearest neighbors.

The record of each cluster is voting for each new record using the distance parameter.



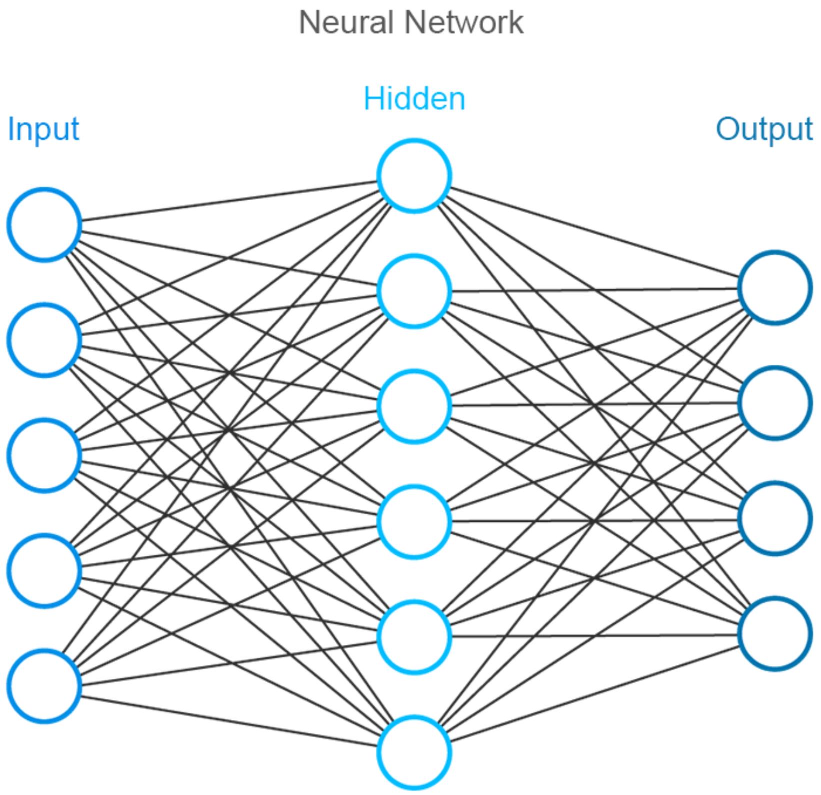
K-nearest neighbors is another common approach used to analyze credit card transactions.

Pros: The method is insensitive to missing and noisy data, which allows for configuring larger datasets with less preparation. It’s also considered highly accurate and doesn’t require much engineering effort to tweak models.

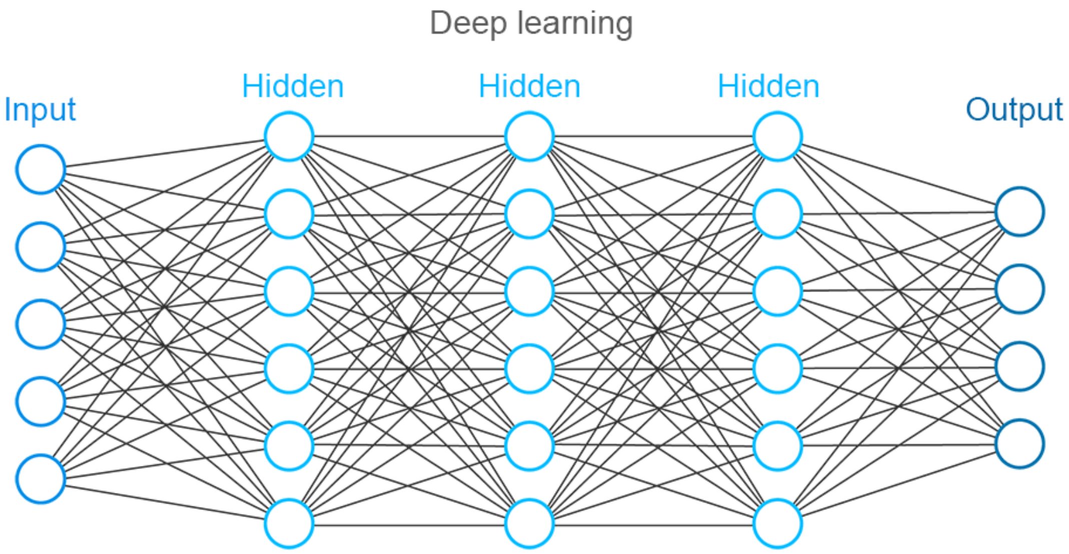
Cons: Like neural networks, k-nearest neighbors require powerful infrastructures and they also lack interpretability.

4.4 Neural networks and Deep Neural Networks

Neural Network is a model that allows for determining non-linear relations between the records. The algorithm structure is built on principles close to those of the human brain neurons. The model is trained on a labeled dataset making input data pass through several layers (i.e. sets of mathematical functions). The models of this type employ 1-2 hidden layers.



Deep Neural Networks works similar to neural networks but employs much more layers than a usual Neural Network. This provides more accurate results, as well as requires more computing power and time for data processing.



Deep learning has created a revolution in data science over the past years. Consequently, it also impacted the financial services industry. Currently, neural networks are being applied both for transactional verification and insurance claims.

Pros: Neural networks and especially deep neural networks are powerful at finding non-linear and very complex relations in large datasets. This works both for transactional data and for text and image analysis, which may be used in insurance cases. They usually provide high accuracy, which makes neural networks a necessary part of a modern fraud detection ensemble.

Cons: Neural networks are state-of-the-art systems that are very difficult to build and tweak to reach efficiency. They require highly skilled professionals and powerful computing architecture. For this reason, we don't recommend using the method for express analysis of all transactions. Another major problem with deep neural networks is the lack of interpretability. While they may be highly accurate, it's nearly impossible to define how specifically the system arrived at one conclusion or the other.

*************************** End of the Module ******************************