The Promise of Machine Learning

A Guide for Financial Institutions
Machine Learning & Fraud

A growing need for real-time fraud identification

Fraud is up – way up – with fraud incidents increasing by over 130% in the past year.¹ Fraud is touching a broader range of companies, too. The percent of organizations that have experienced fraud has increased steadily since 2013.² And if small- and medium-sized organizations think fraud happens only to bigger companies, they should think again. While 80% of organizations with annual revenue of over $1 billion were victims of payments fraud in 2017, almost as many (73%) of companies with annual revenue of less than $1 billion were, too.³ If fraudsters can exploit a weakness, they will do so irrespective of company size.

<table>
<thead>
<tr>
<th>Annual revenue</th>
<th>Organizations that experienced fraud in 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>less than $1 billion</td>
<td>73%</td>
</tr>
<tr>
<td>at least $1 billion</td>
<td>80%</td>
</tr>
</tbody>
</table>

Percent of organizations that experienced attempted and/or actual Payment Fraud³ | 2007-2017

<table>
<thead>
<tr>
<th>Year</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>71%</td>
</tr>
<tr>
<td>2008</td>
<td>71%</td>
</tr>
<tr>
<td>2009</td>
<td>73%</td>
</tr>
<tr>
<td>2010</td>
<td>71%</td>
</tr>
<tr>
<td>2011</td>
<td>68%</td>
</tr>
<tr>
<td>2012</td>
<td>61%</td>
</tr>
<tr>
<td>2013</td>
<td>60%</td>
</tr>
<tr>
<td>2014</td>
<td>62%</td>
</tr>
<tr>
<td>2015</td>
<td>73%</td>
</tr>
<tr>
<td>2016</td>
<td>74%</td>
</tr>
<tr>
<td>2017</td>
<td>78%</td>
</tr>
</tbody>
</table>

2017 ATTEMPTED/ACTUAL PAYMENT FRAUD ACTIVITY³

<table>
<thead>
<tr>
<th>Annual revenue</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>at least $1 billion &amp; fewer than 26 payment accounts</td>
<td>80%</td>
</tr>
<tr>
<td>at least $1 billion &amp; more than 100 payment accounts</td>
<td>74%</td>
</tr>
</tbody>
</table>

Sources:

What explains the growing prevalence of fraud today?

For one thing, fraudsters are leveraging the same tools available to financial institutions. The same payment method diversity, data abundance, and technology that enable companies to improve their business are being used by fraudsters to launch more attacks, faster. As a result, financial institutions must address three key dynamics when designing fraud management strategies:

**Magnitude of attacks are exponentially higher**
- Fraudsters are employing distributed networks, internal knowledge, big data, and even machine learning to detect vulnerabilities and maximize the size of the attacks.

**Weakest links create the most exposure**
- Financial systems are interconnected and consist of a long value chain – a networked ecosystem of multiple entities connecting buyers and sellers. Fraud flows to the least protected components.
- New customer channels (e.g., mobile and social), new products, and new business lines present new risk vectors. In fact, fraud through remote channels is up to seven times more difficult to prevent than in-person fraud.

**Unexpected attacks can be unsettling and disruptive**
- Organizations can go from not having a fraud problem to being devastated in just a few days.

---

**The following are telling statistics about payment fraud attacks on financial institutions:**

- **3 out of 4** financial institutions have incurred fraud losses.
- **96%** of debit card issuers and **77%** of credit card issuers have experienced fraud losses.
- **77%** of check issuers have reported fraud losses.
- **13%** of institutions that offer wire transfers have experienced fraud losses.

**Sources:**


Fraud’s unique characteristics

In addition to navigating the shifting fraud ecosystem, financial institutions must also be prepared to manage the unique characteristics of fraud along with the associated business and technical challenges.

Fraud has long-tail distribution
› Fraudsters are launching micro attacks at all times. There are too many unique cases for humans to pursue each individually.

Fraud is adversarial
› Fraudsters are dedicated to committing fraud. They are using every tool to take advantage of system vulnerabilities within financial institutions to make money. They are also using data abundance and the latest technology to maximize the effectiveness of the attacks.

Fraud patterns change quickly
› Slow-learning countermeasures cannot keep up

Fraud mimics good customer behavior
› Good customers are penalized by overly intrusive countermeasures
The impact of poor fraud detection

Poor fraud detection hurts financial institutions in three key ways:

1. **Financial impact of customer friction:**
   Today’s always-online customers expect fast and seamless experiences. If they encounter friction when accessing their accounts, they are likely to choose different payment cards or close their accounts altogether.

2. **Cost impact of massive fraud attacks:**
   Data breaches are like icebergs. Their hidden costs are a lot larger than their surface costs. And beneath the monetary loss, there’s everything else, which can include: legal penalties, compliance fines, attorney costs, insurance increases, operational disruption, and the loss of public confidence.

3. **Poor fraud detection degrades trust:**
   When organizations cannot tell their legitimate customers from fraudsters, they end up treating everyone like a fraudster by blocking customers or making them do additional steps to verify their identities.

Machine learning for fraud prevention

Organizations that want to defend themselves need to have a superior, fast-learning solution that can evolve constantly. Machine learning achieves this. When applied to fraud management, machine learning can:

- Help to reduce manual review queues through fast, iterating machine models;
- Easily adapt to new business lines using experiential data;
- Help to reduce false positives with behavior analysis;
- Be channel-agnostic;
- Augment human decision-making with increased precision.
Machine Learning Today

Benefits of machine learning

Machine learning as a data science that uncovers patterns and hidden insights is not an entirely new concept; it has been around since the use of neural networks, starting in the 1980s. So why is there so much buzz around machine learning today?

First, as businesses continue to evolve and migrate to the internet, and as modern money is transacted electronically in an ever-growing cashless banking economy, commerce is becoming the business of big data science. This rapidly expanding “dataverse” fuels modern artificial intelligence, making big data an inextricable component of today’s fraud management. Just like IBM’s Deep Blue computer outplayed Garry Kasparov by having learned from millions of chess games, machine learning, in general, requires access to large amounts of data to be able to learn and generalize knowledge. The growth of big data puts additional pressure on fraud solutions to be more sophisticated than fraudsters and to react quickly.

Second, advancements in technology and science have enabled game-changing differences in how machine learning algorithms have evolved and are being applied. Traditionally, human-generated rule sets were the most prevalent approach to fraud management, and they are still used today. But the quantum leap in computing power, combined with the availability of big data over the last few years, has disrupted how data is being used to identify and prevent fraud.

Machine learning uses AI-enabled computer systems to learn, predict, act, and explain without being programmed. Simply put, machine learning eliminates the use of pre-programmed rule sets, no matter how complex.

Machine learning enables:

- **Real-time decisions**
  Advances of in-memory, event streaming technology allows for risk scoring and decision-making in the sub-second range (i.e., ultra-low latency)

- **Increased effectiveness**
  Extremely subtle patterns and variations can be detected and delivered (e.g., precision, recall) better than humans in many tasks

- **Big data processing**
  Advances in distributed data processing allow analyzing more data while still maintaining real-time decisions without trade-offs between data and latency

- **Error-free processing**
  Enormous amounts of data can now be processed without human bias or error

- **Reduced cycle time**
  Learning cycles are continuous, unlike batch learning where models become out-of-date. With machine learning, the same transactions being scored also update and teach the machine learning models

- **Cost efficiencies**
  The many, diverse cases of fraud are far too costly to catch through human intervention
Machine learning engines

Mathematical algorithms power machine learning. But which type of algorithm should companies use? The answer is, "It depends." No one type of algorithm is universally best in all situations. Choosing the best type of algorithm depends on the problem type, size, available resources, and more. Having said that, Random Forests (also known as "Ensemble of Decision Trees") and Deep Learning have been shown to perform very well in a number of scenarios, with Support Vector Machines a close second per the table below.

Random Forests are more robust for a number of real-world problems such as missing data, noise, outliers, and errors. Random Forests also allow multiple types of data (numbers of different scales, text, Booleans, etc.) that can scale very well, parallelizes very easily, is fast to train and score, and requires less effort to achieve the best results. It is no surprise that Random Forests win many machine learning competitions.6

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
</table>
| **Random Forest (a.k.a. Ensemble of Decision Trees)** | • Generalizes patterns well  
• Robust to different input types (i.e., texts, numbers of scales, etc.)  
• Robust to missing data  
• Robust to outliers and errors  
• Fast to train and score  
• Trivially parallel  
• Requires less tuning  
• Probabilistic output (i.e., a score)  
• Can adjust the threshold to trade off between precision and recall  
• Very good predictive power  
• Found to win against many machine learning competitors6 | • Can become complex to interpret as the number of decisions grows (due to the inherent nature of making decisions under increased capacity), but better than all others, especially with Whitebox scoring to demystify decision nodes6  
• Requires labeled data |
| **Deep Learning**                | • Does not require labeled data  
• Reduces feature design tasks  
• Learns multiple levels of representation (e.g., eyes, head, person)  
• Highly parallel  
• Very good predictive power, especially in text and image classification problems | • Very slow to train, but benefits from recent architecture advances (e.g., GPUs, large clusters)  
• Cannot handle different input types  
• Need scaling inputs  
• Needs tuning  
• Does not provide probability estimates  
• Lacks good interpretability  
• Still missing theoretical foundations |
| **Support Vector Machines (SVM)** | • Able to detect non-linear and complex patterns  
• Effective in very high dimensional spaces  
• Very good predictive power | • Requires labeled data  
• Cannot handle different input types  
• Need scaling inputs  
• Cannot handle missing values  
• Not scalable  
• Slow  
• Needs tuning  
• Does not provide probability estimates  
• Lacks interpretability  
• Still missing theoretical foundations |
| **Neutral Networks**             | • Able to represent complex patterns  
• Good predictive power | • Requires labeled data  
• Cannot handle different input types  
• Need scaling inputs  
• Cannot handle missing values  
• Not scalable  
• Slow  
• Needs tuning  
• Lacks interpretability |
| **K-Nearest Neighbors**          | • Robust to missing data  
• Robust to outliers  
• Good predictive power | • Requires labeled data  
• Cannot handle different input types  
• Need scaling inputs  
• Cannot handle missing values  
• Needs tuning  
• Lacks interpretability |

Source: 6 Feedzai
Applying Machine Learning

The application of machine learning has redefined previous strategies and tools in fraud management, delivering benefits that were previously not possible with traditional methods.

Machine learning directly addresses many business challenges that are time-consuming and expensive. For example, manual reviews alone account for almost 25% of the total cost of fraud prevention in the financial services sector.¹

For financial institutions, machine learning can be applied in scenarios where large amounts of data can be used to understand and infer behavior for effective decision making.

- **Account opening**: Enroll only the right customers; respond to loan applications
- **Transaction scoring**: Assess the risk of prepaid, debit, credit, and ATM transactions
- **Account takeover prevention**: Monitor good user accounts for hacking activity

Besides machine learning, models can be used to very efficiently perform analytics and deliver risk scores in real-time, all with greater accuracy by leveraging large amounts of user data.

Behavioral analytics build digital footprints, which can then be used to learn from past data in order to make predictions on future, unseen data patterns. Machine learning algorithms can synthesize this data collected from multiple sources – online and offline – to baseline behavior profiles. User attributes and other data fields used by machine learning algorithms can automatically learn patterns that are used to make predictions.

Machine learning can also be used to automatically derive outcome measurements such as a statistical risk (the measurement of the likelihood of incurring a loss). The effectiveness of the statistical risk score depends on the model’s ability to detect anomalies from known patterns, identify matches to known patterns, and uncover new patterns.

**Explainable logic**

Not all machine learning systems are created equal, especially in terms of how well they explain themselves. It is hard to decipher the machine logic in systems that use neural network algorithms from the 1980s. These systems that make decisions inside black boxes lack transparency around their decisions, which creates two serious problems.

First, there is a control problem, because humans cannot manage and improve a system that they do not understand. Second, there is a validation problem, because an organization cannot audit or validate the decisions that happen inside a black box.

Compare that to a platform that does Whitebox processing to provide clear, human-understandable reasons for its decisions. Certain machine learning platforms predict patterns using Random Forests, which are made of tens of thousands of decision trees. A Whitebox system can take the few top-most factors from this ensemble of decision trees, then weigh and communicate them to the human in a simple way.

For example, a Whitebox explanation might justify a user being declined because there were four recent transactions in distant locations. The idea is to put humans in control by communicating clearly about why certain transactions were approved or declined and arming them with the information they need to review customer account activity more intelligently.

Source:

Hypergranular profiles

A machine learning platform reaches its decisions to create profiles of normal behavior and flags instances of abnormal behavior. But some platforms can zoom in to a higher level of resolution than others.

Does the platform create profiles of loose-fitting cohorts? Or, does it use Segment-of-One profiles to look at individual entities? Segment-of-One profiles do not look at “women between the ages of 30 and 35.” They look at “Susan.” And, they do not look at “devices in this zip code.” They look at “Susan’s device.”

By creating profiles of many different entities (e.g., the branch, the day of the week, the credit card), machine learning platforms can reach new levels of sensitivity, and can help detect fraud with greater power and accuracy.

Machine learning technology enables the creation of Segment-of-One profiles, rather than broad cohorts of many. It allows for the constant updating of these granular profiles based on behavior making detection real-time and at scale.

This is a huge win for financial organizations fighting fraud because fraud just does not happen in batches. Hypergranular profiles complete customer knowledge and fraud knowledge, and can scale as organizations grow and keep up as the payments occur even faster.\(^8\)

---

COHORTS
Come by the Dozen

» Customer Segments (e.g., Baby Boomer, Millennial)
» Account Type (e.g., Gold, Platinum)
» Device Class (iPhone, iPad)

SEGMENTS OF ONE
Come by the Millions

» Each Customer (Susan)
» Each Account Number (123456789)
» Each Device ID (ABCD123456XYZ)

---

Source:

Looking Ahead

Exploding the limitations of machine learning

Though machine learning models are intuitive and can react to cases of potential fraud virtually in real time, there is one main obstacle, which is the steep learning curve.

Expertise in data science, as well as the amount of time and data needed to create models can be beyond the reach of a financial institution. A steep learning curve means data scientists who work on machine learning need to master many different tools such as R, Weka, Python, DBMS, NoSQL data stores, Hadoop jobs, streaming systems, and more.

Without large amounts of data, a machine learning algorithm cannot learn. The existence of efficient algorithms to process this data very quickly opened up the possibility for sophisticated machine-learning algorithms such as spam detection, efficient content recommendations, autonomous vehicles, image recognition, natural language processing, automatic translation, and of course, fraud management.

Respond & React Quickly to Fraud

The promise of machine learning for fraud prevention

Though machine learning models have their limitations, their ability to find interactive relationships in, and to learn from, data allows them to account for evolving fraud behavior. One such model is First Data’s STAR Predictive Fraud Score™, which is based on continuous machine learning that analyzes behavior patterns. This model is updated immediately with new intelligence to automatically discover patterns and has the ability to consume new data sources. As a result, this fraud scoring platform is able to analyze big data in the form of millions of profiles and combinations, and keeps track of every card while looking out for anomalies.

Multiple fraud prevention methodologies in place today have been successful at keeping fraud rates low for typical payment fraud. However, the evolving landscape of financial institutions poses new challenges, necessitating large amounts of computational power to find new solutions that can respond and react quickly to fraud. In this aspect, machine learning is a promising science. It can help to trim manual reviews, adapt to new business lines, and shrink false positives, all while being channel-agnostic, and it enables businesses to defend themselves with a state-of-the-art solution. Machine learning can easily scale to meet the demands of big data with greater flexibility than traditional methods, resulting in enhanced fraud prevention and customer protection.
STAR Predictive Fraud Score

Harness the power of machine learning and Random Forests

STAR Predictive Fraud Score is a first-of-its-kind solution that can give financial institutions better insights into possible fraud and false positives, which in turn helps approve more good payments and prevent fraudulent ones.

Based on machine-learning algorithms, STAR Predictive Fraud Score provides a real-time score for network transactions including ATM, PIN, PINless, and signature debit. STAR’s exclusive, in-house models utilize substantial portions of proprietary, omni-channel data to generate predictive fraud scores.

STAR Predictive Fraud Score is designed to:

› Generate higher approval rates
› Reduce fraud and risk exposure
› Improve customer satisfaction
› Potentially reduce fraud management cost
› Score network transactions during authorization
› Provide human-readable reason codes
› Integrate with First Data fraud products and/or internal risk management tools
› Function on all transaction types: Card present, card-not-present, mobile, chip card and ATM

Every day, STAR® Network helps protect:

$2 billion in scoring volume
30 million transactions total
3,000 transactions per second

Want to learn more?

STAR Network is dedicated to delivering solutions like STAR Predictive Fraud Score that provides a strong defense against fraud.

Visit STAR.com to find out how STAR can help your financial institution build a fraud management program to prevent and detect fraud, increasing the safety and security of every debit transaction.

Source:
9 Statistics referenced from First Data internal information.
Machine learning is a promising science. It can trim manual reviews, adapt to new business lines, and shrink false positives, all while being channel-agnostic, and it enables financial institutions to defend themselves with a state-of-the-art solution. From payment fraud to abuse, machine learning can scale to meet the demands of big data with greater flexibility than traditional methods, resulting in enhanced fraud prevention and customer protection.

About STAR Network
As the largest independent debit network in the U.S., STAR Network offers financial institutions more choice for debit acceptance. Full-service debit capabilities include PIN Debit, PINless, and signature debit functionality for in-person and online commerce. Coupled with innovative network fraud technology, STAR is committed to delivering easier, faster, and safer payments. Visit STAR.com to learn more.

About Feedzai
Feedzai is coding the future of commerce with today’s most advanced risk management platform powered by big data and machine learning. Founded and developed by data scientists and aerospace engineers, Feedzai has one mission: To make banking and commerce safe. The world’s largest banks, processors, and retailers use Feedzai’s fraud prevention and anti-money laundering products to manage risk while improving customer experience.