WHITE PAPER

5 Keys to Successfully Applying Machine Learning and AI in Enterprise Fraud Detection

Like many aspects of our lives, technology is providing consumers with more options when it comes to making financial transactions. From relatively new interactions, such as mobile commerce and peer-to-peer payment apps, to near real-time payment windows, the changes are rapid and expansive. Consumers want immediate payment processing whenever and however they choose to transact. They want a banking relationship that's 100% convenient and 100% secure, and modern financial institutions are doing their best to accommodate these market demands. But there are gaps.

The sense of urgency being applied to satisfying consumer expectations is outpacing the sense of urgency being applied to new fraud detection strategies. This often results in gaps, where fraud defenses lag the introduction of new consumer-focused banking enhancements. Sophisticated, organized fraudsters are equipped to exploit these gaps, the byproducts of payment technology innovation. In order to remain effective, enterprise fraud strategies will require an adaptive and cohesive understanding of individual consumer behavior across all interactions. Channel-specific fraud strategies will no longer cut it.

The good news is that payment fraud is an ideal use case for machine learning (ML) and artificial intelligence (AI). The bad news is that the hype surrounding these topics has made it difficult to distinguish myth from reality. This paper pulls from FICO's experience in using AI to prevent billions of dollars in fraud. It is intended to cut through the noise and explain the factors that are essential for success.

The intended audience for this research includes fraud operations, corporate IT and risk executives within global financial institutions, as well as e-commerce retailers, payment processors and others looking to reduce payment fraud.

A brief history of fraud

Physical theft

Sale of fake and stolen payment cards

Identity theft and application fraud

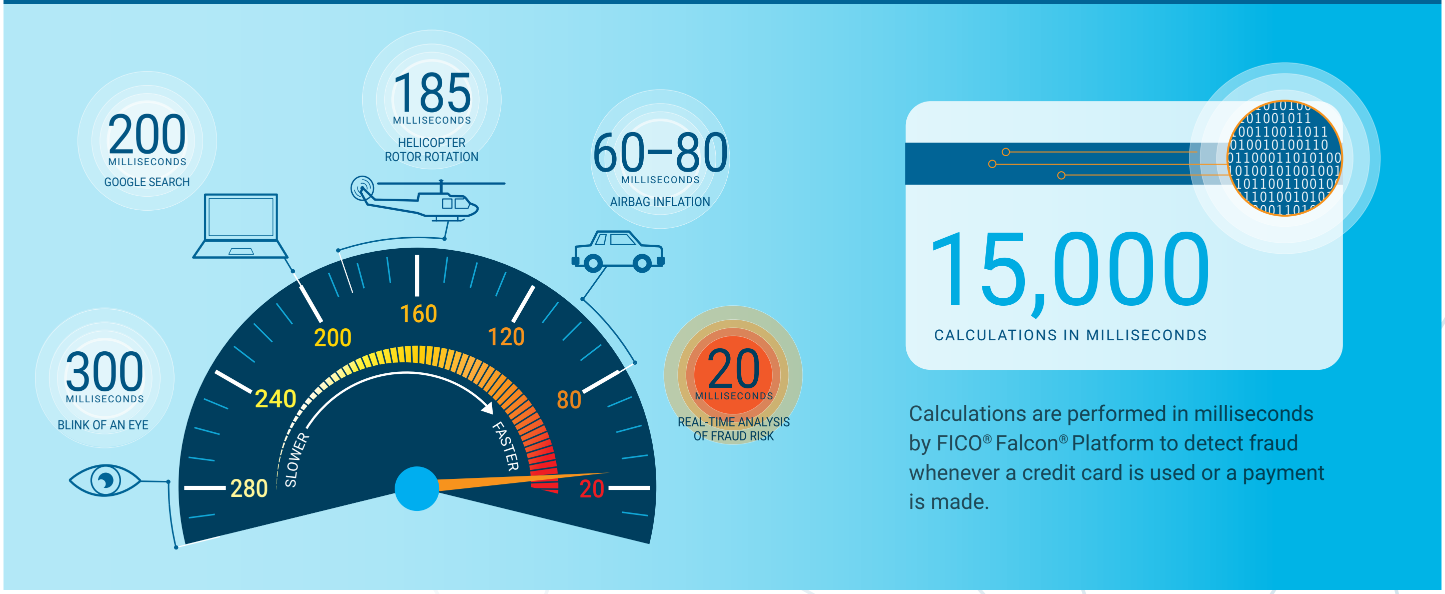
Phishing, social engineering and mass data compromises

Organized fraud rings using advanced analytics and cyber attacks

Engagement of Machine Learning and Artificial Intelligence

Machine learning helps data scientists efficiently determine which transactions are most likely to be fraudulent, while significantly reducing false positives. The techniques are extremely effective in fraud prevention and detection, as they allow for the automated discovery of patterns across large volumes of streaming transactions. If done properly, machine learning can clearly distinguish legitimate and fraudulent behaviors while adapting over time to new, previously unseen fraud tactics. This can become quite complex as there is a need to interpret patterns in the data and apply data science to continually improve the ability to distinguish normal behavior from abnormal behavior. This requires thousands of computations to be accurately performed in milliseconds. Without a proper understanding of the domain, as well as fraud-specific data science techniques, you can easily employ machine learning algorithms that learn the wrong thing, resulting in a costly mistake that is difficult to unwind. Just as people can learn bad habits, so too can a poorly architected machine learning model.

How FAST Is Payment Fraud Detection?



Calculations are performed in milliseconds by FICO® Falcon® Platform to detect fraud whenever a credit card is used or a payment is made.

KEY 1

Integrating supervised and unsupervised AI models in a cohesive strategy

Payment fraud has become extremely sophisticated as banks face dynamic threat vectors devised by organized crime networks. As a result, defense strategies based on any single, one-size-fits-all analytic technique will produce sub-par results. Each use case should be supported by expertly crafted anomaly detection techniques that are optimal for the problem at hand. As a result, both supervised and unsupervised models play important roles in fraud detection and must be woven into comprehensive, next-generation fraud strategies.

A supervised model, the most common form of machine learning across all disciplines, is a model that is trained on a rich set of properly "tagged" transactions. Each transaction is tagged as either fraud or non-fraud. The models are trained by ingesting massive amounts of tagged transaction details in order to learn patterns that best reflect legitimate behaviors. When developing a supervised model, the amount of clean, relevant training data is directly correlated with model accuracy.

While supervised learning is predictive, unsupervised learning is more focused on the automated discovery of discordant patterns. Unsupervised models are designed to spot anomalous behavior in cases where tagged transaction data is relatively thin or non-existent. In these cases, a form of self-learning must be employed to surface patterns in the data that are invisible to other forms of analytics.

Unsupervised models are designed to discover outliers that represent previously unseen forms of fraud. These AI-based techniques detect behavior anomalies by identifying transactions that do not conform to the majority. For accuracy, these discrepancies are evaluated at the individual level as well as through sophisticated peer group comparison.

By choosing an optimal blend of supervised and unsupervised AI techniques you can detect previously unseen forms of suspicious behavior while quickly recognizing the more subtle patterns of fraud that have been previously observed across billions of accounts.

> "The trajectory of fraud attacks on the financial value chain is rising on a number of fronts. Data breaches abound, putting an unprecedented quantity of payment card data, personally identifiable information (PII) and stolen credentials in the hands of organized crime rings."
>
> **Julie Conroy, Research Director at Aite Group**

KEY 2

Applying behavioral profiling analytics

Behavioral analytics use machine learning to understand and anticipate behaviors at a granular level across each aspect of a transaction. The information is tracked in profiles that represent the behaviors of each individual, merchant, account and device. These profiles are updated with each transaction, in real time, in order to compute analytic characteristics that provide informed predictions of future behavior.

Profiles contain details of monetary and non-monetary transactions. Non-monetary may include a change of address, a request for a duplicate card or a recent password reset. Monetary transaction details support the development of patterns that may represent an individual's typical spend velocity, the hours and days when someone tends to transact, and the time period between geographically disperse payment locations, to name a few examples. Profiles are very powerful as they supply an up-to-date view of activity used to avoid transaction abandonment caused by frustrating false positives.

A robust enterprise fraud solution combines a range of analytic models and profiles, which contain the details necessary to understand evolving transaction patterns in real time.

Given the sophistication and speed of organized fraud rings, behavioral profiles must be updated with each transaction. This is a key component of helping financial institutions anticipate individual behaviors and execute fraud detection strategies, at scale, which distinguish both legitimate and illicit behavior changes. A sample of specific profile categories that are critical for effective fraud detection includes:

Transaction profiles	Behavioral Machine Learning profiles for each consumer's financial and non-financial activity. Updated in real-time with each transaction, across all channels.
Collaborative profiles	Improves risk sensitivity by identifying behaviors that differ from typical behaviors within individuals' peer groups.
Behavior sorted lists	Deep learning behavioral analytics that identify and rank recurrent activities that are unique to each individual, such as favorite merchants, ATMs or destination accounts, in order to significantly reduce false positives.
Merchant profiles	Aggregated transactions at merchant-level to form behavioral metrics for a more comprehensive view of risk.
Multi-layered self-calibrating profiles	Detects behavioral outliers in real time even with limited or no data to train the model and automatically adjusts to accommodate new behavioral patterns. Advanced instances use deep learning to further improve pattern recognition.
User-defined profiles	Flexible, custom-defined profiles for entities such as devices or IP addresses.
Global intelligent profiles	Real-time adaptive risk ranking to monitor and respond to the riskiest profiles for improved fraud assessment.

KEY 3

Distinguishing specialized from generic behavior analytics

As in many professional fields (medicine, law, construction, etc.), there are generalists and specialists. The same applies to the development of behavioral analytics; there are general modeling techniques and highly specialized techniques. For example, your communications provider may use a behavioral model to spot indications that you're likely to leave their service in the next 90 days. Or an online gaming site may employ general behavioral analytics to identify users in need of intervention. While human behavior may appear to be a universal concept, the ability to effectively detect anomalous behavior in one domain is not an indication of effectiveness in another.

In fraud detection, artificial intelligence relies on raw data as well as predictive characteristics that serve as inputs to a model that produces a score. These characteristics represent the inferred patterns or relationships within the data that are often discovered with machine learning. Data scientists with strong knowledge of the fraud domain then improve this discovery process by evaluating and refining the weights, portions and combinations of predictive characteristics for optimal model performance.

Many providers of fraud detection analytics choose to ignore the importance of domain knowledge in the model development process. Instead, they rely on generic behavior models that must learn to identify patterns of fraud slowly over time, based on relatively few cases.

Consider this example: A 42-year-old woman from Sacramento, CA, is a frequent domestic traveler. She is attempting to withdraw the equivalent of $300 US from an ATM in Seoul. Your fraud system has less than a second to make a risk determination. Is this anomalous behavior? It may be. That's relatively easy to determine. But is it indicative of fraud? That's a tougher question that only specialized fraud analytics, honed on huge quantities of data, can accurately assess. In order to maintain a positive consumer experience, specialized fraud analytics must be used to assess the "tough" questions. This is where advanced profiling, fraud-specific predictive characteristics, and adaptive capabilities separate themselves from generic behavior analytics.

In a world of real-time payment processing and rapidly changing consumer preferences, generic behavior models are not sufficient for cross-channel, enterprise fraud solutions. After all, when and how someone chooses to transact is not as predictable as his or her likelihood to cancel a fitness club membership. There are significant financial and reputational risks when pointing a generic behavior model at a fraud use case.

KEY 4

Leveraging large datasets in model development and training

> "Card not present (CNP) fraud has gone from 50% of gross fraud losses in 2008 to 70% in 2016."
>
> **FICO Evolution of Card Fraud in Europe 2016**

An intuitive notion, supported by independent research, indicates that depth and breadth of data is more impactful to machine learning model performance than cleverness of the algorithm. It's the computing equivalent of human experience. This suggests that, when possible, you can improve predictive accuracy by expanding the dataset used to craft the predictive characteristic used in a machine learning model.

Think about it: There's a reason why physicians see thousands of patients during their training. It's this amount of experience, or learning, that allows them to accurately diagnose within their area of specialization. In fraud detection, a model will benefit from the experience gained by ingesting millions or billions of examples, consisting of both legitimate and fraudulent transactions. Superior fraud detection is achieved by analyzing an abundance of transactional data in order to effectively understand behavior, and assess risk, at an individual level.

At FICO, we have performed extensive research on different modeling techniques. Clearly, across a variety of use cases, the volume and variety of training data are more critical to prediction than the type of algorithm used. This research, and similar independent research throughout the AI community, indicates that fraud models that are developed and trained using data from thousands of institutions will be more accurate than models that rely on a relatively thin dataset.

KEY 5

Adaptive analytics and self-learning AI in enterprise fraud

> "Recent innovations in self-driving cars and advanced voice recognition highlight the need for domain-specific AI techniques. The same holds true in payment fraud, where even the criminals are analytically sophisticated."
>
> **Scott Zoldi, Chief Analytics Officer at FICO**

Fraudsters ensure that protecting customers' accounts is very complex and dynamic, a challenge where machine learning thrives. For continual performance improvement, fraud detection professionals should consider adaptive technologies designed to sharpen responses, particularly on marginal decisions. These are the transactions that are very close to the investigative triggers, either just above or just below the cutoff. It is on these margins where accuracy is most critical as there is a fine line between a false positive event — a legitimate transaction which has scored too high, and a false negative event — a fraudulent transaction which has scored too low. Adaptive analytics sharpen this distinction with up-to-date knowledge of the threat vectors an institution is facing.

Adaptive technologies improve sensitivity to shifting fraud patterns by automatically adapting to recent confirmed case disposition, resulting in a more precise separation between frauds and non-frauds. When an analyst investigates a transaction, the outcome — whether the transaction is confirmed as legitimate or fraudulent — is fed back into the system to accurately reflect the fraud environment that analysts are facing, including new tactics and subtle fraud patterns that have been dormant for some time. This adaptive modeling technique automatically modifies the weights of predictive features within the underlying fraud models. It is a powerful tool that improves fraud detection performance on the margins and stops new types of fraud attacks.

FICO uses adaptive analytics to balance the benefits of a model developed with data from thousands of institutions with the ability to quickly learn unique fraud patterns from an individual institution and boost fraud performance and responsiveness.

Transform your fraud prevention and detection with machine learning

Banks and payment processors are committed to enabling new customer experiences; these innovations are vital to their business. As a result, there is a need to craft adaptable fraud strategies that perform against evolving fraud tactics, which are increasingly using machine learning as well. This requires fraud models that are highly adaptable and accurate and allow fraud teams to remain invisible to legitimate customers.

The keys to achieving this level of success are quite clear. The most effective fraud models, both supervised and unsupervised, consist of fraud-specific predictive characteristics. These characteristics, which are vetted by domain experts, are often behavior-based machine learning models in their own right. Massive amounts of global transaction data are then analyzed to produce fraud models with the optimal mix of portions and combinations of predictive characteristics. In production, these models leverage behavioral analytics, which interpret the behaviors of each person, merchant and device, in real time. Lastly, adaptive technologies leverage AI to evolve the underlying analytics based on feedback from investigative outcomes.

If properly applied, machine learning and artificial intelligence help provide the foundation for highly effective fraud detection controls. Yet, these techniques by themselves are not sufficient. Proper awareness is necessary to navigate the haze surrounding these topics and develop a clear strategy that will scale and adapt with your needs.

Why FICO

9000+ financial institutions around the globe trust the FICO® Falcon® Platform to identify the individual behaviors that require attention while remaining invisible to consumers during legitimate transactions.

To learn more about how machine learning and artificial intelligence can help you prevent enterprise fraud, visit us at www.fico.com/AI-MachineLearning.

FICO®

FOR MORE INFORMATION
www.fico.com
www.fico.com/blogs

NORTH AMERICA
+1 888 342 6336
info@fico.com

LATIN AMERICA & CARIBBEAN
+55 11 5189 8267
LAC_info@fico.com

EUROPE, MIDDLE EAST & AFRICA
+44 (0) 207 940 8718
emeainfo@fico.com

ASIA PACIFIC
+65 6422 7700
infoasia@fico.com