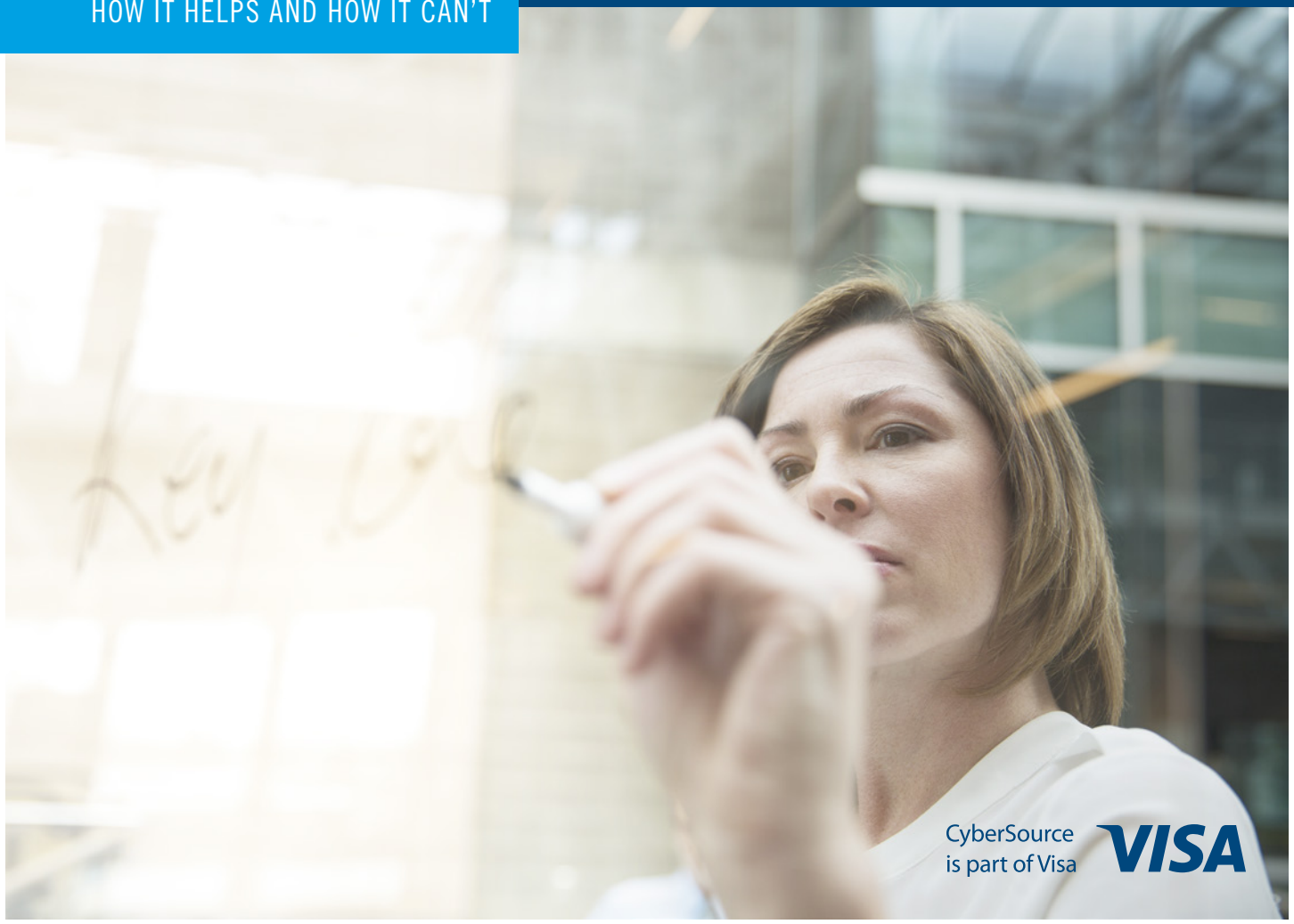
CyberSource®

Payment Management for the Digital Economy

THE ROLE OF MACHINE LEARNING IN FRAUD MANAGEMENT

HOW IT HELPS AND HOW IT CAN'T



PIONEERING MACHINE LEARNING IN PAYMENTS

Everywhere you look today, there are examples cropping up of how machine learning is revolutionizing different industries. In media and entertainment, Spotify® and Netflix® sort through billions of data points to find patterns in music, movies and television that consumers have enjoyed — and then make suggestions based on consumers' tastes. In retail, Amazon® prompts consumers to buy everything from diapers to desk chairs based on shoppers' prior purchases. In finance, machine learning is helping investors anticipate market trends and is powering innovations underlying everything from self-driving cars to voice-assistant applications.

In the payments industry, machine learning is similarly becoming an increasingly important tool to help businesses combat fraud. As new technologies transform the way we pay — originally with credit and debit cards and, more recently, with kiosks, smartphones and other mobile devices — the number of transactions flowing through global payment networks has increased. At the same time, criminals have grown more sophisticated and more adept at using technology — even big data and analytics — to disguise illicit activity. As a result, it is getting harder and harder for businesses to determine which transactions to approve and which ones to reject.

Machine learning methodologies, when deployed as part of automated fraud screening systems, can help businesses make the right call. This white paper explores the analytical capabilities of machine learning and the role it plays in detecting fraud.

WHAT IS MACHINE LEARNING?

Machine learning relies on complex statistical methods and high-octane computing power. At its core, however, is a very simple concept. By identifying the most influential cause-and-effect relationships from the past, a machine can *learn* to make accurate predictions about the future.

Machine learning begins with high-powered computers which are guided by human intelligence to sift through billions of historical data points and identify these cause-and-effect relationships. Then, all this information is fed into a variety of algorithms

to arrive at predictions. As computers get better at identifying these cause-and-effect relationships, they leverage the insights they have gained and use them to refine the algorithms. That is the "learning" that is taking place — at processing speeds far faster than the human mind.

There are a number of strengths that make machine learning such a powerful approach, specifically when it comes to fighting fraud. Machine learning helps:

Facilitate real-time decision-making. Rule-based systems, where people create ad hoc rules to determine which types of orders to accept or reject, require a great deal of time-consuming, manual interaction. Machine learning can help evaluate huge numbers of transactions in real time.

Improve accuracy. As criminals have grown more sophisticated, they have become more adept at disguising fraud. Machine learning can often be more effective than humans at detecting subtle or nonintuitive patterns to help identify fraudulent transactions. It can also help avoid false positives — good orders that are erroneously identified as fraudulent.

Rapidly respond to change. Because fraudsters are always changing their tactics, it's a constant cat-and-mouse game. Machine learning is continuously analyzing and processing new data and then autonomously updating its models to reflect the latest trends.

Lower costs. Significant advances in technology have reduced the costs associated with machine learning solutions and computing systems capable of running them. As machines learning improves accuracy, it also reduces costly false positives and minimizes the time and expense of manual reviews.

AN ENSEMBLE OF MACHINE LEARNING METHODS

Machine learning methods range from basic to very sophisticated. Here are a few that can be useful in fraud management.

REGRESSION ANALYSIS

Regression analysis is a popular, longstanding statistical technique that measures the strength of cause-and-effect relationships in structured data sets. Regression analysis tends to become more sophisticated when applied to fraud detection due to the number of variables and size of the data sets. It can provide value by assessing the predictive power of individual variables or combinations of variables as part of a larger fraud strategy.

ARTIFICIAL NEURAL NETWORKS

Artificial neural network models mimic the way the brain works by building an interconnected network with multiple layers of neurons. Within each layer, the neurons are weighted combinations or functions of inputs. Neural networks typically rely on a transformation layer that converts raw data into more discrete data that the neural network then processes. The model is trained to minimize prediction errors and can be applied supervised or unsupervised. There are many varieties of neural networks, such as deep learning. Although neural networks are well-suited to find subtle interactions between multiple variables, they typically cannot explain why they produce a given score.

DECISION TREES

Decision trees, similar to flowcharts, segment data into meaningful groups. A decision tree is made up of decision points based on different data elements, categorizing the data into smaller and smaller groups. The end of a given tree path provides a score representing the predicted outcome and accuracy level.

Decision trees are able to take in unstructured data with minimal transformation. In addition, the path through a tree provides readily available insight into the logic behind the score. There are numerous decision tree methods, many of which rely on using multiple trees together. Some of the more popular variants are random forests, boosted trees and stochastic gradient boosting trees.

OTHER CLASSIFICATION SYSTEMS AND METHODOLOGIES

There are countless other types of machine learning methods, including customized systems with unique classification algorithms. There are also many methodologies that can improve the training processes of a machine learning system, such as boosting, filtering and time series analysis.

AN EVOLVING APPROACH TO FIGHTING FRAUD

Not long after the first credit card was swiped over a half century ago, businesses began contending with the burgeoning problem of fraudulent digital transactions. At first, businesses literally had to make fraud assessments by hand. They scoured big, thick books listing thousands of "hot credit card numbers" to see if any matched the one presented to them by a customer. In the decades that followed, that process became increasingly automated — first with the help of call center representatives and then with technology programmed to help identify suspicious activity.

In our era of electronic commerce, the moment of truth for every transaction occurs during the checkout process, when a customer uses his or her payment information to complete a mobile or digital payment. In that instant, businesses must decide whether to approve the transaction, completing the payment, or reject it, potentially losing a good sale and not providing a great customer experience.

As computer processing power has increased exponentially and new statistical methods have become more robust, predictive analytics and machine learning in particular has played an increasingly important role in fraud management.

BETTER DATA, BETTER RESULTS

The capabilities of machine learning models are magnified by having more and higher quality data, or both "breadth and depth." Decision Manager relies on the following to help better manage and detect fraud patterns:

Volume of Data: A fraud management solution needs access to a vast store of historical transaction data to help train its models and maximize the likelihood that it will uncover patterns of suspicious activity.

Richness of Data: It is not just the number of past transactions that counts — it is important to get as much information about each transaction as possible. Pulling data from different sources can enhance data quality and fill in missing information gaps.

Relevancy of Data: By collecting data from payment processors, businesses and major payment networks, it is possible to tap into a vast reservoir of "truth information" — data regarding chargebacks, credit backs and manual reviews based on the actual outcome of past transactions. This data is critical to distinguishing between good and bad transactions.

DIFFERENT MODELS CAN HELP MAXIMIZE IMPACT

In general, fraud management solutions rely on two different types of machine learning models to combat payment fraud. At one end of the spectrum are *static models*, which learn how to identify fraud at a point in time by sifting through millions of past transactions. Static models are effective at identifying historical fraud patterns and tend to work well right after they are built. The trouble is there is no way to update or adjust these models as new patterns of fraudulent activity emerge.

At the other end of the spectrum are *self-learning models*, which continuously incorporate data points from new transactions to adapt and recognize ever-evolving fraud patterns. Self-learning models are very effective at identifying the latest fraud tactics. However, the "black box" nature of these models makes it all but impossible for a human to track, control or adjust what the machine learns — which means the model can suddenly cause huge problems if it draws incorrect insights and begins blocking good customers.

Not all machine learning systems rely solely on static models alone or self-learning models alone — there is a middle ground that can capture some of the benefits of each. By merging static and self-learning techniques, it is possible to immediately incorporate newly available information, such as recent chargebacks, while improving the consistency and robustness of scoring.

THE LIMITATIONS OF MACHINE LEARNING

Effective fraud screening strategies should analyze and act on data insights quickly. But machine learning is not a silver bullet. There are still several significant limitations to the approach:

- Machine learning relies on good input data. There must be enough relevant data points to identify legitimate cause-and-effect relationships. Without the appropriate data, the machine may learn the wrong thing — and make erroneous or irrelevant fraud assessments.

- Machine learning is only as good as the human data scientists behind it. Even the most advanced technology cannot replace the expertise and judgment it takes to effectively filter and process data and evaluate the meaning of the risk score.

- Machine learning is often a black box, especially when self-learning techniques are employed. The machine can learn the wrong thing. For example, under normal circumstances, orders with overnight shipping might often be considered fraudulent. While that may hold true for most of the year, it could reject a lot of good customers during the holiday season.

A way to counteract the downsides of machine learning is to combine an automated machine learning system with a rules-based approach. Rules can act as guide rails that give businesses a higher degree of immediate control over fraud decisions.

REAL-TIME FUSION MODELING

Machine learning has served as a core part of CyberSource's Decision Manager platform almost from the start. Today, the CyberSource fraud management solution serves as a testament to this continued wave of development. Not only are there now more than 260 anomaly detectors, but there are 15 region, channel and industry-specific risk models, each optimized to identify fraud in different scenarios. Meanwhile, machine learning has become ingrained in our fraud-fighting capabilities as part of the patented approach we call *real-time fusion modeling.*

Real-time fusion modeling is the centerpiece of the CyberSource Decision Manager approach to fraud scoring. It leverages the proven effectiveness of conventional static models with the more agile data analysis capabilities of today's most advanced self-learning models to help businesses more effectively and efficiently manage and detect fraud.

Importantly, the CyberSource Decision Manager platform does not just rely on one machine learning method to generate its risk assessments. That is because no single model works best in all situations — effectiveness depends on the type and amount of data, the requirements of the user and other factors.

In addition, CyberSource's rules-based engine ensures the business is the ultimate decision-maker. Businesses have the flexibility to set and adjust rules at any time through a user interface. The rules engine provides added transparency, since businesses can easily see what rules were triggered in a decision. CyberSource's real-time fusion model in tandem with a flexible rules engine represents a powerful combination, allowing for swift and accurate responses to unique or emerging fraud trends.

OUR ENTERPRISE FRAUD MANAGEMENT SOLUTION

CyberSource offers a multi-layered fraud management solution — from account monitoring to transaction fraud detection, rules tuning to payer authentication — that can help your business minimize fraud losses, maximize revenue and minimize operational costs. Built on intelligence from more than 68 billion transactions that Visa and CyberSource process annually and proprietary fusion machine learning algorithms, the CyberSource enterprise fraud management solution represents a comprehensive system that helps merchants:

- Protect revenue and operate more efficiently — through more accurate, automated detection, businesses can reduce fraud loss and decreases and streamlines manual review
- Increase sales and maintain customer loyalty — by reducing false positives, increasing order acceptance rates and protecting customer accounts
- Operate with greater flexibility and scalability through confident, real-time control over fraud strategies and the ability to leverage expert resources and infrastructure worldwide

ABOUT US

CyberSource Corporation, a wholly owned subsidiary of Visa Inc., is a payment management company. More than 400,000 businesses worldwide use CyberSource and Authorize.Net brand solutions to process online payments, streamline fraud management and simplify payment security. The company is headquartered in Foster City, California. CyberSource operates in Europe under agreement with Visa Europe. For more information, please visit www.cybersource.com.

CONTACT CYBERSOURCE

NORTH AMERICA
SAN FRANCISCO, CA, UNITED STATES

T. +1 888 330 2300

E. SALES@CYBERSOURCE.COM

LATIN AMERICA & THE CARIBBEAN
MIAMI, FL, UNITED STATES

E. LAC@CYBERSOURCE.COM

ASIA PACIFIC
SINGAPORE

E. AP_ENQUIRIES@CYBERSOURCE.COM

EUROPE, MIDDLE EAST & AFRICA
READING, UNITED KINGDOM

E. EUROPE@CYBERSOURCE.COM

For a complete list of worldwide offices, go to: www.cybersource.com/locations