CONCLUSIONS PAPER

Machine Learning Use Cases in Financial Crimes

Ten practical and achievable ways to put machine learning to work

sas
THE POWER TO KNOW.

Contents

Featuring
Kerem Muezzinoglu, Data Scientist, Fraud Management
Advanced Analytics Team, SAS
Carl Suplee, Director of Product Management for
Fraud and Security Intelligence, SAS
David Stewart, Director, Security Intelligence Solutions for Banking, SAS

Are Your Fraud Systems Keeping Pace?

Fraudsters are crafty, and they have to be. As financial institutions discover and block their usual tactics, they have to change course. They launch more sophisticated and complex schemes, involve more entities to cover their tracks, cruise just under the radar, and move in new directions whenever another path is blocked.

In short, the art and science of fraud are constantly evolving. Are your fraud and financial crimes systems evolving to keep pace?

Most organizations still use rules-based systems as their primary defense. Rules are great for uncovering known patterns, but rules alone are not very good for uncovering unknown schemes, adapting to new fraud patterns, or handling increasingly sophisticated forms of financial crimes.

That's where **machine learning** comes in. Unlike rules-based systems, which are fairly easy for criminals to test and circumvent, machine learning adapts to changing behaviors in a population.

Machine learning systems automatically create analytic models that adapt to what they find in the data. Over time, the algorithm "learns" how to deliver more accurate results, whether the goal is to make smarter credit decisions, retail offers, medical diagnoses or to detect fraud. It's easy to see the value of adaptive modeling to keep pace with emerging fraud tactics.

With advances in machine learning, detection systems can learn, adapt and uncover emerging patterns of financial crime – all the better to catch it before the losses are significant, or to prevent it altogether.

Machine Learning Finds New Application

So machine learning has been around for decades. What's new is that it can now be applied to huge quantities of data – big data – at a realistic price. Cheaper data storage, distributed processing, more powerful computers and new analytical tools have lowered the barrier to entry.

Data scientists can crunch much bigger data sets, much faster, on commodity hardware. They can run hundreds or thousands of iterations of these analytical techniques at blazing speeds on inexpensive computing platforms, so models achieve much better results than previously possible.

However, in an informal webcast poll, only 12 percent of respondents said their firms were currently using machine learning in production. Nearly 40 percent reported that their firms are either piloting a project or plan to do so in the next three to six months – while nearly half of respondents have no plans to adopt its capabilities.

Let's take a look at why the 88 percent who aren't yet using machine learning should accelerate their plans or consider a second look. SAS advanced analytics specialists have worked with major financial firms to develop 10 use cases to:

- **Automate tasks** that formerly required human intervention, such as gathering data for case investigations, and
- **Detect more financial crimes risk** that rules and less sophisticated analytic techniques might miss.

From Checkers to Go – Machine Learning Outsmarts and Outwits

Nearly 60 years ago, computer pioneer Arthur Lee Samuel defined machine learning as "a field of study that gives computers the ability to learn without being explicitly programmed." Much like a Montessori school student, the computer program explores the stimuli set before it (the data) and learns from that exposure rather than being overtly taught (programmed).

Samuel devised a checkers-playing program that appears to be the world's first self-learning program. It wouldn't have beaten a world champion – that honor went to IBM's Deep Blue chess-playing computer nearly 40 years later – but it could beat a respectable amateur at checkers.

In 1979, SAS introduced its first machine learning algorithm for discriminant analysis. In the 1980s, SAS and others wrote algorithms to perform logistical regression and pattern recognition. The 1990s saw advancements beyond classical statistical procedures to include such approaches as decision trees, neural networks and ensemble models (which combine multiple learning algorithms for more accurate predictions).

In the last decade, we've seen IBM's Watson computer use natural language processing and machine learning to beat two of the most successful *Jeopardy!* contestants of all time. In 2016, Google's AlphaGo program used machine learning to repeatedly beat a world champion at Go, a game that was previously thought too complex for computers to solve.

When Amazon and Netflix recommend things you might like, machine learning is behind the scenes. When the Siri and Alexa intelligent personal assistants help you organize your life and make recommendations, or when facial recognition authenticates your online or mobile payments, machine learning is at work.

The concept of machine learning has been around for decades. What's new is that it can now be applied to huge quantities of data – big data – at a realistic price.

Different types of machine learning

The right machine learning approach for detection depends on the input you have for training the machine and what you hope to achieve.

With **supervised learning**, the computer program is presented with sample inputs and their associated outputs, and the goal is to devise a general rule that maps those inputs to outputs. To train a supervised model for fraud detection, you present it with records associated with both fraudulent and legitimate activities, and the model seeks to define a function or instruction set that can predict the presence of fraud when applied to new data.

With **unsupervised learning**, the learning algorithm isn't given any labels (dependent variables in statistics terms). The algorithm is on its own to find structure in the input, such as discovering hidden patterns in the data. In this case, since you don't know which data represents financial crime, you want the model to create a function that describes the structure of the data, flags as an anomaly anything that doesn't fit the norm, and then applies this knowledge to new and unseen data.

Machine Learning for Process Automation

As fraud risk and compliance pressures mount, it's not feasible to just add more people to the problem. Instead, we have to find ways to streamline their work. Machine learning is a natural fit for automating manual or repetitive processes, particularly for uncovering financial crimes, because manual processes raise the cost of compliance to unacceptable levels.

Let's look at six ways to use machine learning to free humans from tasks that computers can do for them much faster.

Automatically enrich alerts and boost SAR conversion rates

Would you like to reduce false positives in your anti-money laundering or compliance groups, while focusing on the risks that are most important to the firm? One way financial institutions are achieving that is through an **AML auto-referral** or **hibernation function**. This function starts with your existing alert generation process and automatically enriches alerts with a broad range of relevant information .

Enrichment adds potentially significant detail about the customers, accounts or beneficiaries associated with the alert, such as:

- Prior cases, suspicious activity reports (SARs) or currency transaction reports (CTRs).
- Existing scoring processes that assess the risk of a transaction, series of transactions, customer or accounts.
- External information such as law enforcement inquiries, subpoenas or negative news.

Fraud detection, online searches, network intrusion detection, real-time credit scoring, email spam filtering – these applications all apply some form of machine learning. They tap into knowledge gained from previous exposures to continuously improve the algorithm and make new and more informed actions.

Guided by machine learning, the automated referral process:

- Brings all those components together.
- Applies a scoring model against all that activity and data.
- Consolidates the resulting score into an entity-level score.
- Runs the entity-level score through risk assessment processes and transaction monitoring.
- Identifies what is truly out-of-threshold and needs to be presented to a human being.

"Unlike suppression, hibernation considers all activity," said Carl Suplee, Director of Product Management for Fraud and Security Intelligence at SAS. "It doesn't remove anything from the equation. This holistic approach can identify risks that would be missed when viewed in isolation. For example, maybe a particular pattern of transactions doesn't by itself point to fraud, but when it occurs in tandem with several other attributes or events, it looks like a big risk to the institution.

"This hibernation approach enables you to gather all the really risky activity together, and then put all that context in front of investigators or analysts to do a full investigation. We've seen this type of automated triage increase SAR conversion rates 30 to 50 percent."

Automate the process of collecting data for case investigation

On average, 60 percent to 70 percent of an investigator's or analyst's time is spent collecting data from disparate systems and entities inside and outside the organization – and that's probably a conservative estimate. It takes a lot of time to simply find and gather everything needed to support an investigation.

"This is a prime opportunity to let technology do what it does best, to automate some of that search and retrieval," said Suplee. "Machine learning can guide systems to automatically search and retrieve data, run queries against databases, or have application programming interfaces (APIs, or 'web calls') collect information from third-party data providers without human intervention."

Data could be pulled from most anywhere, such as your KYC (know your customer) system, account opening system, payment system or wire transfer system. It could be prior cases and SAR narratives, or images, such as checks, statements and more. It could be third-party data, such as from LexisNexis or Google Map checks.

"Suppose you want 18 months of transactional history and you don't necessarily want to load it all into your transaction monitoring or case management system," said Suplee. "You can just have calls go out and pull the data forward. Automation brings a lot of efficiencies to these processes that just take a while.

"This use case is not necessarily about reducing false positives [as with hibernation], but about reducing the man hours spent gathering information. It's a quick win. We're seeing clients reduce the time to case decision by 20 to 30 percent."

Automate the complex process of generating business rules

Virtually all decisions in finance are driven by rules. There may be scores underneath, but the verdict comes from a business rule or rule set. Rules represent the accumulation of domain knowledge, conceived with industry expertise. Traditionally, business users shape these rules based on their needs and observations.

But as noted earlier, fraudsters are good at testing and circumventing rules. To be effective, rules must adapt to reflect what's happening in the world. Imagine the power of having machine learning examine masses of data to help establish rules while keeping them current.

Some financial channels are particularly suitable for governance by rules, even without a risk score. In payments, for instance, activity is characterized by many categorical attributes associated with many different parties – each category signaling different levels of risk.

"A primary tool for handling categorical data and implementing rules with machine learning is the decision tree," said Kerem Muezzinoglu, Principal Staff Scientist on the SAS Fraud Management Advanced Analytics Team. "The history of the decision tree goes back more than three decades, but it's a fundamental building block of many cutting-edge methods today.

"There are many principled methods to build decision trees. A tree generated by a machine, composed without human involvement, is pretty transparent and follows concrete design principles, designed to maximize separation. When you follow a branch on a decision tree, you can see all the data elements that go into that decision.

"The typical branch of seven or eight nodes that points to a fraud-rich bucket in our data is a pretty easy discovery for a machine, but it is a very difficult discovery for a human being. However, when you present such a branch to a human, it is pretty understandable by that human. We owe this clarity to the highly structured and relational characteristics of the transactional data. Decision trees can actually teach us, if we follow what they have discovered."

Automatically generate natural language notes and narratives

In its simplest form, natural language generation is turning data into plain language. For automating financial crimes detection processes, it's about taking data from multiple sources and locations – internal and external – and turning that data into readable text to support an investigation.

A prime opportunity for this technique is in generating notes and narratives for alerts, SARs and CTRs. Natural language processing makes it simple for an analyst to read, understand and augment the content, then transfer it into reports.

A natural language process, coupled with image recognition, can identify document types and apply context-specific analytics based on this classification to deliver about an 85 percent accuracy rate.

"Natural language generation is an easy-to-use form of artificial intelligence that firms definitely need to consider, because it's another quick gain in efficiency," said Suplee.

Create a virtual digital assistant to support the human's process

With the SAS surveillance bot, artificial intelligence meets machine learning. A surveillance bot captures data – clicks, workflow, data movement, data points, external data inputs, manual data entry, etc. – and uses all that information to help predict future steps for the human process.

"Suppose I always go to Screen A, then Screen B, and then do a negative news search for a certain kind of alert," said Suplee. "Instead of having to initiate these repetitive and predictable actions, the system can learn the typical user interaction for a specific type of alert, and automatically offer it."

The system learns from the human, and then uses that knowledge to ease and automate the task for the human. "The system is basically saying, 'Hey, this is the exact same or similar thing you did before. I'll just do it for you instead,'" said Suplee. Looking further, machine learning could also recommend what that route should be, the next step the investigator or analyst should take, based on what has been done in the past.

Automatically recognize and scan document types

If you're an analyst in a typical global trade unit, you'll deal with international letters of credit with packets of information, each trade packet containing 20 or 30 different types of documents. Cognitive computing can automate the tedious process of recognizing the different formats of documents within these packets.

For example, classification analysis determines if a document is a bill of lading or an invoice. "If it is classified as an invoice, then you can start looking for over-invoicing, under-invoicing, terms that may be consistent with dual-use goods, and other risk indicators," said David Stewart, Director of Security Intelligence Solutions for Banking at SAS. "In pilots with some top 10 global banks, we've been able to automate those processes with about an 85 percent accuracy rate, thereby eliminating a lot of manual effort.

"In another pilot, we scanned about 9,000 SWIFT messages looking for things like Palestinian boycott language. For a human to review those messages would require about five to seven minutes per message. During this pilot we found we could do the image recognition and contextual analysis of those messages in less than a second per message."

In that pilot, about 2.5 percent of those 9,000 messages called for some type of further action – and two of those messages were actually reportable based on sanctions laws. Bottom line: An important compliance task that formerly required two person-weeks of time was done in less than a minute.

"For an institution of that size and staffing level, this form of automation was an eight-figure take-out for them in a year," said Stewart. "When you look at the global trade finance problem, the possibilities are definitely intriguing."

Natural language generation

- Uses data in the system or external
- Turns the data into plain language text
- Auto-generates a case or SAR narrative

Machine Learning for Detecting Fraud and Financial Crimes

Detect rare events

Detecting rare events in real time requires tracking the entire history of the many entities involved in a transaction. A high-quality score can only be generated by observing and tracking individuals/entities involved in the transaction – not only their separate history but also their relationships.

"Since new entities and relationships arise and change in nature all the time, one cannot really foresee and hard-code all relationships and possible behaviors during the development of a model," said Muezzinoglu. "However representative the historical data may be, we really need to teach a model, a relevant detection system, how to evaluate a profile, a time series, an entire history of these entities and their relationships.

"So our approach in tracking real-time behavior is very elaborate. We store everything as raw time series, with all the activity – amounts, country codes and many other attributes of the transaction. At every entity, we track these attributes, preserve them, store them as time series, make them available for real-time decisioning in milliseconds, and convert them into predictors on the fly."

There is still a fair amount of processing on the time series, but it's important to preserve the richness of history detail rather than summarized history. At the same time, a multitude of summaries are being calculated and compared against the current event to improve fraud detection.

Detect more digital payments fraud

SAS deployed a digital payment model that has demonstrated rapid success for real-time fraud detection. It alerts on 50 percent of fraud at a cost of only 0.5 percent of the portfolio, with very few false positives.

The model tracks each sender account, beneficiary and online user, watching key indicators such as these:

- **For sending and beneficiary accounts** – Destination country, amount, same day of month payments, tenure of the relationship for senders and beneficiaries.
- **For online users** – Device fingerprint, browser data, payment template details, authentication signals.

"The system stores all of this information, plus text entries, as time series and uses it in real time," said Muezzinoglu. "Then we derive tens of thousands of variables on these time series and provide different summaries. We also use hard-coded risk tables, such as looking at certain countries or static behavior that the model deems risky.

"On top of that, we build an inference engine and try many different machine learning techniques, including neural networks, random forests and boosted decision trees. Through multiple iterations – usually three full development cycles – we find the best-performing technique, the 'champion.' The resulting model is typically a champion over hundreds of candidates."

Uncover more fraud by adding insight from text analytics

In building fraud detection models, no details should be ignored. There are the obvious details, such as amount, geolocation and distance, but certain transactions also come with descriptions and annotations.

Since this text is provided by the user, it can reflect some unique characteristics related to the user. This text can disambiguate certain attributes that numerical data, such as account numbers and transaction amounts, might not. For example, is the beneficiary a business or an individual? Are certain words or phrases being consistently used? Can common synonyms, such as Co, Inc. and LLC, link separate entities as one?

A SAS model has proven to be about 80 percent accurate in differentiating businesses from individuals and has improved detection about 10 percent in some case studies. "In short, no details should be overlooked, however inconsequential they may appear," said Muezzinoglu. "We've seen small notes serve as a very rich data source."

Use unsupervised learning to find out what you don't know

A large global financial institution engaged SAS to mine a large data set and find something they didn't know. Specifically, were there high-risk customers in their customer base that weren't properly classified?

Through unsupervised learning, a model can analyze a large amount of data and find things that are out of norm and might create some type of reputational risk for the institution. "In this case, you didn't necessarily know if people are good or bad; you're just looking for those who represent 'edge cases,' people who are behaving out of norm, relative to their peers," said Stewart.

"We applied some of our newer algorithms – a random forest model with 200 trees – and split out what we deemed to be customers who were behaving as money service businesses but who were not declared as such during the onboarding process, nor were they legally registered as money services businesses. In a population of nearly 2 billion transactions (aggregated in about 10 minutes), we found 416 suspected money service businesses, 89 previously unknown or unregistered MSBs, which, through further triage, resulted in dozens of productive cases."

Through unsupervised learning, a model can analyze a large amount of data and find things that are out of norm and might create some type of reputational risk for the institution.

Five Don'ts for Success

1. Don't be intimidated. Machine learning in its simplest form is automation.

People can get hung up on the term "artificial intelligence," thinking it represents something novel, complex and unattainable. But your organization is probably already using AI in some manner and, as stated earlier, machine learning has been around for a long time.

In its simplest form, artificial intelligence is a mechanism to help automate processes. Embrace it, figure out ways to take advantage. In an informal webinar poll, almost 40 percent of respondents said they would apply AI more for fraud detection, about 10 percent for automating alert scoring, and about 20 percent for robotic process automation. Most large firms are already doing at least one or two of these with machine learning.

2. Don't put all your faith in artificial intelligence.

Machine learning should be seen as an adjunct to human wisdom and experience. In spite of huge advancements in analytical algorithms, machines should not be entirely autonomous in decision making – particularly in financial services. They require proper oversight, for instance, deciding when to update models, how to collect and incorporate domain knowledge in retraining models, and evaluating unexpected inputs. Human intelligence is still indispensable in the process, although at an increasingly higher level.

3. Don't pursue machine learning just for cost take-out.

For the ability to automate processes and decisions that formerly required manual intervention, machine learning can certainly boost efficiency. But don't make your business case on cost-reduction per se. Think more in terms of freeing your data scientists, analysts and investigators from mundane activity that can be automated, so they can apply their time to more valuable things, such as uncovering new risks or more thoroughly studying known risks.

4. Don't sacrifice transparency for efficiency.

Machine learning technology, as cool as it can be, should not overshadow accountability. Decisions must always be transparent and explainable to reviewers. This precept should guide both the development and production life of a machine learning system. Even this transparency costs money, and even if it somewhat compromises efficiency, that long-term accountability is a necessary benefit.

5. Don't pursue AI for the sake of it.

Sooner or later, fraud and risk management professionals will get asked at the water cooler, "Hey, are we doing AI? This other firm is doing it. Why aren't we? We should do AI."

"We get a lot of inquiries I feel may have come from that kind of conversation," said Stewart. "But it's not about pursuing AI for its own sake. There are tangible, proven examples where it can help automate processes, help you be more efficient, or detect behaviors that other traditional approaches could miss. Start with the business objective in mind."

About the Presenters

Kerem Muezzinoglu is a Principal Staff Scientist on the SAS Fraud Management Advanced Analytics Team. He specializes in building predictive analytics solutions for financial fraud.

Carl Suplee is Director of Product Management for Fraud and Security Intelligence at SAS. He is responsible for product direction, strategic development, and domain expertise for SAS financial crimes solutions.

David Stewart is Director of Security Intelligence Solutions for Banking at SAS. David leads strategy development, drives product management, and provides marketing counsel for SAS fraud and compliance solutions worldwide.

Learn More

To learn more about this topic, please visit sas.com/fraud and to learn more about how SAS delivers artificial intelligence, vist sas.com/ai.

To contact your local SAS office, please visit: sas.com/offices