How to use this guide

We have pulled together a range of materials that reflect the reality of using machine learning to tackle and manage the ever-evolving fraud threat. An introduction to the basics of machine learning is followed by a look at the changing role of the analyst in this new reality. Finally we provide a guide to what questions to ask prospective fraud vendors in the form of an Request for Proposal template.
Machine learning and fraud prevention

Machine learning is a form of artificial intelligence that enables computers to learn without being explicitly programmed. It’s especially good at recognising patterns and spotting anomalies in data.

How does Ravelin use machine learning to prevent fraud?

Fraudulent and legitimate transactions have different characteristics. Algorithms are created based on those differences to predict the likelihood of fraud. At Ravelin, we prioritise techniques that allow us to explain how decisions are made. Our customers want to know why and how fraud is happening, so we choose from the following approaches and blend them together to provide a probabilistic fraud score.

Random Forest
This technique uses multiple decision trees to improve classification performance. It allows us to smooth the error which might exist in a single tree, increasing our overall performance and accuracy while maintaining our ability to explain the results to users.

Logistic regression

This statistical technique uses an algorithm to compare a merchant’s good transactions with chargebacks to predict the likelihood of whether a new transaction will be a chargeback. For very large merchants, these models are specific to their customer base.

Neural networks

Neural networks attempt to mimic how the human brain learns, particularly in pattern recognition. They are trained on legitimate patterns and are therefore adept at flagging fraudulent ones. It complements other techniques and improves with exposure to data.

What considerations should be made when using machine learning to detect fraud?

Machine learning is not a silver bullet for fraud prevention. It takes a significant amount of data for machine learning models to become accurate. For some merchants, it is useful to apply a basic set of initial rules and allow the models to ‘warm up’ with more data. At Ravelin, we want to be able to explain the reasons why a customer is suspected of fraud. Our merchants need to be able to confirm fraud and consequently train the system. A further
difficulty surround detecting connections in data. To counter this, we enhance our models with Graph networks.

Why is machine learning the best approach to manage fraud?

While there are limitations to machine learning, machines are much better at dealing with large datasets than we are. They can recognize thousands of features on a user’s purchasing journey, see deep into the data and make concrete predictions. In large datasets, machine learning approaches get even better. This avoids rising costs as merchants increase their customer base, whilst retaining efficiency. The goal is to avoid human intervention unless their input adds insight. Our algorithms should do the heavy lifting. Lastly, machine learning allows our merchants to get results, fast – in microseconds, in fact.

The fraud analyst’s basic role is to spot which transactions are real and which are fraudulent. A fraud analyst can wear multiple hats, partly identifying fraudulent transactions from data, partly doing research to identify trends. They often interact with other areas of business to offer insight and help optimize processes. In the face of cybercrime, the fraud analyst is our super hero... But to produce those heroics, they need support. Ecommerce reached $2 trillion in 2016. Machine learning has an important role to play in managing this scale. But what does the introduction of number-crunching technologies mean for analysts?

The Changing World of Fraud Analysis

The last decade has seen some astonishing changes in the way analysts work. Not least is the coming of automation, which has impacted industries from banking to healthcare.

Fraud analysis is now entering that period of automation. Machine learning-based systems like those at Ravelin are being used to learn trends and spot patterns in transactions which point to fraud. The computer starts off with some initial variables, then as more data comes in the program adjusts the output. It improves over time, giving more accurate results. Humans do this too, of course, and are infinitely more adaptable.

Machines, however, can perform analysis on a huge scale and consume a breadth of data. In essence, what is being automated is the repetitive

Machine learning and the evolving role of the analyst

Fraud analysts are vital. They need to be equipped with the latest technology like machine learning to effectively stop fraud.

According to a Nilson report, global card fraud exceeded $24 billion in 2016. A recent study by Javelin found that financial fraud costs retailers around 7.5% of annual revenue, with digital retailers feeling the greatest impact. The volume, variety and complexity of the fraud attacks are all growing, placing a greater burden on the shoulders of fraud analysts.
– even boring – tasks of scanning data for patterns. For the intelligent minds of fraud analysts, this can be a kind of torture.

Workarounds and efficiencies in scanning data are well sought after. For many years, this resulted in the creation of rules to enhance decision making. Now it is the turn of machines that can use a much greater set of data than rules could ever hope to, and can apply a much broader set of approaches that are suited to varying types of fraud.

Friend or Foe?

One of the worries about techniques like machine learning is that it will ‘take our jobs’. However, we have rarely seen this happen in any other industry. For instance, the digital spreadsheet certainly removed many manual data entry roles in accountancy, but it gave wing to many more value-adding accountancy roles in its wake.

Keeping the machines honest

Machines are far from perfect. They can overweight signals, block good transactions, misinterpret information and wrongly allow fraudulent payments. They can take time to adapt to a new threat.

Any reputable vendor will be completely frank about this. Obviously, these errors should be kept within agreed limits, but the interpretation of these missed and bad calls are the key to improving performance. Working with your vendor, it’s important to establish productive feedback loops. For instance, within Ravelin it is possible to confirm and reject decisions that are made by the machines. These are strong signals to the machine learning models that will enforce or weaken a decision matrix, ensuring that the models continually provide improved precision and recall.

Further to that we recommend weekly meetings with a vendor to discuss trends and suggest areas where performance can be improved. Detailed analysis of customers or transactions that have been wrongly determined are invaluable feedback for a data science team. It helps improve the models to enhance future performance.

Fraud investigation for profit and for service

A somewhat surprising outcome of the adoption of machine learning is the increased relevance of the fraud team within the customer’s experience. Whereas in the past the sheer volume of manual review would often restrict in depth investigations, it’s now possible to set aside time to investigate any customer queries. For instance, a perfectly legitimate customer can be blocked for fraud. It’s possible to rapidly check a fraud decision, and if it appears to be a genuine error, the vendor can overturn it and try to save the customer. However, fraudsters are more than happy to call or email in and try to overturn a decision.

Having the fraudster’s complete history, the reason for the decision, and their link analysis network in a single dashboard means an analyst can confirm
the decision with complete confidence. Machine learning tools are often misconstrued as a ‘black box’ environment that is impossible to query. Ravelin provides rich detailed information, sensibly displayed to allow fraud analysts to confirm decisions one way or the other within seconds.

The role of the fraud analyst will therefore move closer to the customer service team and should be increasingly seen as part of the customer experience than of revenue protection. Recruiting people with the correct qualities to handle this diverse role may be necessary to manage this tweak in responsibilities.

Giving the Fraud Analyst Wings

Machine learning should be seen as a useful tool for the fraud analyst. Rather than replacing the analyst, it works to enhance the role, taking the leg work out of fraud analytics. The analyst team will be much more concerned with strategy, with the accuracy of the results and with improving the decisions. There will be a significant shift in the interpretation of the role towards value-adding intelligence that is critical to revenue generation.

Asking the right questions

The decision to purchase a new or replacement fraud detection system is a difficult one. There are many factors to consider, and many important questions to be asked.

There has been a seismic shift in the technology landscape which has become very evident within fraud prevention. That shift is the emergence of machine learning as an approach to solving the fraud problem for online businesses. Using this transformative technique, digital companies are fighting back. But for the non-expert, it can be difficult to know which questions to ask. What makes one vendor’s approach better than another, and how do you determine which vendor has the best capabilities?

To help we have written a Request for Proposal (RFP) document that will aid you in making that determination. We acknowledge that we are biased, of course – we’re strong believers in the abilities of this technology to reduce online fraud levels. But that bias also means we want clients to ask informed and meaningful questions of their machine learning fraud detection vendor.

We have chosen questions that we think will provide fair and balanced insight into the technical and functional choices of the vendors with which you have selected to engage. In any case, if there are questions you don’t like or sections you think are missing then simply remove or add them. Please get in touch if you have any feedback.
Ravelin’s RFP template

Overall Fraud Functionality
1. Describe your approach to solving fraud for online businesses.
2. What skills or training are required to use your system?
3. Describe the implementation process including how long it takes and what aid is given.
4. What code or scripts are required to provide data to your system?
5. Are you able to build models and rules around our data?

Machine Learning Requirements
1. Which ML techniques do you deploy and why?
2. Are the same models deployed across all of your clients or are they specific to each client?
3. How do you explain decisions to clients?
4. How do your models adapt to changing fraud patterns?
5. How do your models manage incomplete data?
6. How do we train the machines on an ongoing basis?
7. How do we feed chargeback information into the models?
8. How fast are the fraud scores returned to clients?
9. Do we need to add any steps to the purchase process or code to our site/app to work with the ML fraud system?
10. Can we override the models’ decisions? How?

Graph Network Requirements
1. What are the key benefits of using graph networks?
2. How do you visualise your graph networks, and how are they accessed?
3. How accurate have your networks been to date? How have you validated that?
4. How do you manage incomplete data in graph networks?
5. How long does it take to spin up a graph network visualisation?
6. Can I see a network for every customer in the system?
7. What drawbacks do graph networks have and how do you manage them?
8. Do we need to add any steps to the purchase process, or add any code to our site or app to work with the graph network technology?
9. How fast are fraud scores returned to clients and is there any interruption to the customer journey while waiting for response?
10. Can we override the graph network decisions? How?

Support and SLA Requirements
1. What support will we get during the integration process?
2. What is your incident report process?
3. What is the target resolution time for minor and major incidents?
4. How often does your system go down for maintenance? What will be the impact when it does?
5. How do you manage peak traffic?
Pricing and Commercials

1. Do you charge for integration support or any other setup fees?
2. Do you have monthly management or support fees?
3. What is the estimated cost per transaction? How will this change as we grow?
4. Do you guarantee chargebacks? How does that work?
5. Are there any additional fees?

Overview

This report has explored both the merits and setbacks of machine learning technology, explaining why we at Ravelin use it as our primary method of fraud prevention. Although machine learning techniques are changing the role of the fraud analyst, human actors are still integral to the process – especially when interpreting results and communicating with clients. Finally, we have provided an RFP template to help customers ask the right questions when considering a machine learning fraud detection system.