A Geek's Guide to Machine Learning and Risk Analytics and Decisioning

By Jarryd-Lee Mandy

Introduction

Artificial Intelligence (AI), Machine Learning (ML) – whatever you want to call it, these buzzwords are appearing more and more throughout the business and social world. So what are they and what do they mean?

Despite the growing interest, AI/ML isn't new at all. In fact, the models themselves have been around since the 1970s and '80s. In the financial sector, banks have been using ML to mitigate fraud and detect irregular buyer behaviors and patterns for the last decade or more.

Fraud is a growing concern and is costing the financial sector millions of dollars in losses each year. A 2015 research note from Barclays stated that the United States is responsible for 47 percent of the world's card fraud despite accounting for only 24 percent of total worldwide card volume [1]. A 2014 Federal Trade Commission report shows that credit cards and other consumer payment methods produced the greatest losses over other types of fraud [2].

One of the ways in which UK financial firms have tried to reduce fraud is with the implementation of the Chip and Pin system. It was seen as an effective means to prevent and reduce card fraud. But a research paper by Murdoch *et al* (2010) showed how fundamentally flawed Chip and Pin is [3].

As technology evolves, so do the cunning methods for perpetrating a fraudulent crime. Financial firms are now relying on sophisticated artificial intelligence software to evolve, adapt and learn in line with the behavior patterns of fraudsters in order to track, detect and prevent fraud far more quickly than traditional methods. The use of AI has also been implemented in industries outside financial services including insurance, retail and telecommunications.

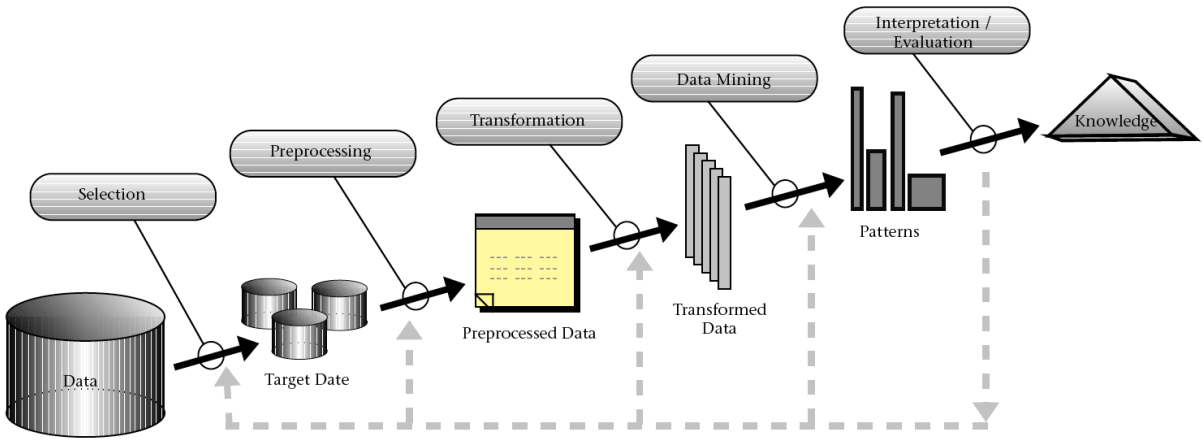
Obviously, it is in the interest of the card issuer or bank to implement strategies to reduce the risk of fraud. Unfortunately, this often requires a compromise between expense and inconvenience to the merchant and the customer. Merchants are at far more risk than the end credit card user as they are ultimately responsible for incurring the cost of a fraudulent purchase and the potential loss of the customer resulting from the bad experience. Other costs to the merchant include direct fraud costs, cost of manual order review, cost of reviewing tools and cost of rejecting orders [4].

This Provenir report describes the use of AI tools in credit card fraud to mitigate risk. We will be looking at various AI detection methods including Artificial Neural Networks (ANN), Fuzzy Neural Networks (FNN), Bayesian Neural Networks (BNN) and Expert Systems.

PROVENIR

An Overview of Fraud Prevention and Detection Techniques

The modern information age is flooded with a rapidly growing and astonishingly huge amount of data. In the U.S alone, the total number of credit card transactions totaled 26.2 billion in 2012 [5]. The processing of these data sets by banks and credit card issuers requires complex statistical algorithms to extract the raw quantitative data.



An overview of the processes that compose Knowledge Discovery in Databases (KDD) (Source: Fayyad *et al*, 1996) [6].

These systems work by comparing the observed and collected data with expected values. Expected values can be calculated in a number of ways. For example, a behavior model would look at the way a customer's bank account has been used in the past, and any deviance from usual purchasing habits would return a *suspicion score*. This method works by flagging a transaction with a typical score, usually between 1 and 999. The higher the score, the more suspicious the transaction is likely to be, or, the more similarities it shares between other fraudulent values.

Typically, the measures taken to combat fraud can be distinguished into two categories – *Prevention* and *Detection*.

- *Fraud Prevention* constitutes the necessary steps to prevent fraud from occurring in the first place, with various preventative methods used to deter fraudsters, such as MasterCard SecureCode and Verified by Visa.
- *Fraud Detection*, the focus of this report, comes into play once fraud prevention fails. Detection consists of identifying and detecting the fraudulent activity as quickly as possible and implementing the necessary methods to block and prevent the card from being used by the perpetrator again [7]. Issues arise when criminals adapt and change their tactics once they are aware that a prevention method is in place, therefore the need for more intelligent and sophisticated technology which 'learns' is essential for the detection of fraud.

The techniques used to detect fraud also fall into two primary classes – *Statistical techniques* (clustering, algorithms) and *Artificial Intelligence* (ANN, FNN, Data Mining) [8]. Both of these methods still involve mining through the available data and highlighting any anomalies (which can be defined by a set of rules) from the purchasing and transaction data of the customer. The difference is that where we used human analysts to manually search useable knowledge in the past, today we make use by machine learning [9] [10].

Artificial Intelligence Models

Artificial Neural Networks

Also known as connectionism, parallel distributed processing, neuro-computing and machine learning algorithms, Artificial Neural Networks (ANNs) were first developed during the late 1980s and have since become a fundamental tool in combating fraud [11]. ANNs work by imitating the way the human brain learns, using complex input, hidden, and output layers.

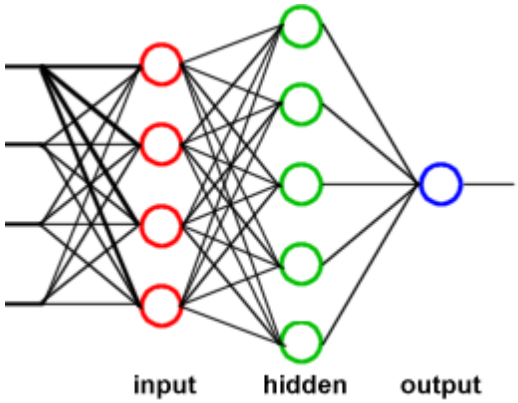


Diagram representing a feed-forward multilayer perceptron (the most common type of ANN). (Source: www.oscarkilo.net)

The input nodes retrieve information from an outside source (for credit card fraud detection, this would be the transactional data of a customer's account) and the output nodes send the results from the neural networks back to the external source. The hidden nodes in-between the input and output nodes have no interaction with the external source and become more complex in their configuration and nature depending on the complexity of the problem at hand [12].

The various nodes in each layer of the neural network are connected by edges where each edge represents a particular weight between two connected nodes. (In the human brain, these are called synapses.) The information that the neural network learns through supervised or unsupervised learning is stored in these weights.
An example of the way neural networks learn is similar to the way children learn to recognize animals. After seeing a dog, the child can then generalize on various other breeds of dogs, categorizing and defining them as 'dogs' without having seen them before.

An important feature of neural networks is that when they learn, they have the option to be *supervised* or *unsupervised*.

- For unsupervised neural network learning, the system makes use of clustering, which groups patterns based on similarity. The two main unsupervised learning methods are *Hebbian* and *Kohonen*. Hebbian learning takes place by association, meaning that if two neurons which are on either side of a synapse are activated simultaneously, the strength of that synapse will be increased. *Kohonen* (also called Self-Organizing Maps) learning takes place by learning the categorization of the input space [13].
- For supervised neural network learning (*back-propagation*), the correct output values for certain input data are determined before starting the algorithm, and the system then learns the function between the paired input and output nodes [14].

A user can train a neural network by running through examples of past data. The learning process occurs when the output data is compared to that of the ANN's predicted output. The weights for each connection are then adjusted based on the exampled data, allowing the system to learn new patterns and behavior and improve accuracy without having to be taught or shown it [15].

Fuzzy Neural Networks

Fuzzy Neural Networks (FNNs) are a branch of hybrid intelligence systems which make use of fuzzy logic together with ANNs to detect fraudulent activity. The idea was first developed and proposed by Zadeh and has since been used and implemented successfully in a variety of industries [16]. The core framework for fuzzy logic is to provide an accurate method for describing human perceptions. Some experts believe that the use of fuzzy rules can provide a more natural estimate as to the amount of deviation from the normal [17].

FNNs, like Expert Systems, make use of IF-THEN-ELSE statements and heuristic rules to handle uncertainty in applications, resulting in better approximate reasoning without the need for analytical precision. The use of traditional IF-THEN-ELSE statements and heuristic rules (see Expert Systems below) has been controversial, and therefore has not been as widely implemented as some of the other AI fraud detection systems [18].

Expert Systems

Expert Systems saw increased usability and growth during the 1980s with the expansion of computer processing power, programming and AI. It was used in credit card fraud detection by using a rule-based system which proved to be fairly popular when no other intelligent systems were around. These systems were used to imitate and replicate the knowledge of an 'expert' person and can be defined into two classes – *facts* and *heuristic* [19] [20].

- Facts are classified as a quantity of information, such as the credit card transaction history or an individual's credit rating.

- Heuristic is where a person of 'expert' knowledge defines a set of rules that they would usually follow by protocol as a result of their 'expert' experience, education, observation and training.

Expert systems work by taking this human knowledge and transferring it into a logical language that a computer can understand and follow in order to solve a problem. A fundamental part of expert systems is their extensive database of stored rules which are defined by a typical IF-THEN-ELSE format. For example, a rule based system using IF-THEN-ELSE may look like the following:
IF the amount of purchase is greater (>) than $1000 and the card acceptance authorization is through 'eBay', THEN raise a suspicion score and require further verification, ELSE approve transaction.

Limitations of Expert Systems however are that they require considerable storage space and rely heavily on extensive programming of expert human knowledge in order to make decisions. Some experts believe that by using a rule-based system with neural networks, the performance of detecting fraudulent activity increases [21].

Bayesian Neural Networks

These types of networks take a slightly different approach to the general guidelines and rules of learning that are commonly seen in ANNs and FNNs. Typically, Bayesian Neural Networks use *Naive Bayesian Classifiers*, a simple method of classification, to classify transaction activity.

Bayesian learning can be trained very efficiently in a supervised learning setting and uses probability to represent uncertainty about relationships that have been learnt as opposed to variations on maximum likelihood estimation [22]. Where neural networks try to find a set of weights for each node (process of learning) to best fit the data inputted, Bayesian learning makes prior predictions by means of probability distribution over the network weights as to what the true relationship might be [23]. One study looked at the comparison of using both ANNs and Bayesian Belief Network algorithms in fraud detection, and found that the use of Bayesian Neural Networks, although slower, were in fact more accurate than the use of ANNs alone [24].

In fact, many believe the use of Bayesian methods to be highly effective in real world data sets as they offer better predictive accuracy [25]. This is supported by research which concluded that the use of Bayesian Neural Networks were far superior and accurate in detecting credit card transactional fraud than Naive Bayesian Classifier [26].

The Data

The following table compares the research findings to highlight which combination of models provides the highest prediction accuracy.

Study	Method/Technique	Application	Prediction Accuracy (%)		
			Fraud	Non-Fraud	Total
Aleskerov et al (1997)	ANNs	Credit Card	n/a	n/a	n/a
Bell and Carcello (2000)	Statistical	Financial Reporting	**55**	**96**	**87**
Brause et al (1999)	Data mining and ANNs	Credit Card	n/a	n/a	n/a
Bolton and Hand (2002)	Clustering	Credit Card	n/a	n/a	n/a
Calderon and Green (1994)	Statistical	Financial Reporting	**20**	**89**	n/a
Dorronsoro et al (1997)	ANNs	Credit Card	n/a	n/a	n/a
Ezawa and Norton (1996)	BNNs	Credit Card	n/a	n/a	n/a
Ghosh and Reilly (1994)	ANNs	Credit Card	n/a	n/a	n/a
Green and Choi (1997)	ANNs	Financial Reporting	**68**	**74**	**72**
Leonard (1995)	Expert System	Credit Card	n/a	n/a	n/a
Lin et al (2003)	FNNs	Financial Reporting	**35**	**86**	**76**
Quash and Sriganesh (2007)	ANNs (Self-Organising Maps)	Credit Card	n/a	n/a	n/a
Zaslavsky and Strizkak (2006)	ANNs	Credit Card	n/a	n/a	n/a

Summary of the most notable investigations into the use of Artificial Intelligence at mitigating fraud.

The greatest challenge when talking about artificial intelligence/machine learning is actually in understanding what data sets we are looking at, and what model/combination of models to apply. Amazon's Machine Learning offering is one example of an automated process which analyses the data and automatically selects the best model to use in the scenario. Other big players who have similar offerings are IBM Watson, Google and Microsoft.

Conclusion

Provenir's clients are continually looking at new and innovative ways to improve their risk decisioning. Traditional banks offering consumer, SME and commercial loans and credit, auto lenders, payment providers and fintech companies are using Provenir technology to help them make faster and better decisions about potential fraud. Integrating artificial intelligence/machine learning capabilities into the risk decisioning process can increase the organization's ability to accurately assess the level of risk in order to detect and prevent fraud.

Provenir provides model integration adaptors for machine learning models, including Amazon Machine Learning (AML) that can automatically listen for and label business-defined events, calculate attributes and update machine learning models. By combining Provenir technology with machine learning, organizations can increase both the efficiency and predictive accuracy of their risk decisioning.

Sources

[1] Barclays' Security in Payments: A Look into Fraud, Fraud Prevention, & the Future, May 22, 2015
[2] https://www.ftc.gov/system/files/documents/reports/consumer-sentinel-network-data-book-january-december-2014/sentinel-cy2014-1.pdf
[3] Murdoch, S. Drimer, S. Anderson, R. Bond, M. (2010). *Chip and Pin is Broken.* Available: http://www.cl.cam.ac.uk/research/security/banking/nopin/oakland10chipbroken.pdf.
[4] Widder, A. Ammon, R. Schaeffer, P. Wolff, C (2008). Combining Discriminant Analysis and Neural networks for Fraud Detection on the Base of Complex Event Processing. *In: GECCO*, march 04-06, Gaithersburg, USA. pp.n/a.
[5] https://www.frbservices.org/files/communications/pdf/research/2013_payments_study_summary.pdf
[6] http://www.infovis-wiki.net/index.php?title=Knowledge_Discovery_in_Databases_(KDD)
[7] Bolton, R. Hand, D. (2002): "Statistical fraud detection: A review.(Unsupervised profiling methods for fraud detection)", *Statistical Science,* vol. 17 (3): 235-255.
[8] G.K. Palshikar, The Hidden Truth – Frauds and Their Control: A Critical Application for Business Intelligence, Intelligent Enterprise, vol. 5, no. 9, 28 May 2002, pp. 46–51.
[9] Trybula WJ (1997). "Data mining and knowledge discovery". Annual Review of information Science and Technology 32: 197-229.
[10] Decker P (1998). "Data mining's hidden dangers". Banking Strategies, 74(2): 6-14.
[11] Tafti, H. A .M. Nikbakht, E. (1993). Neural Networks and Expert Systems: New Horizons in Business Finance Applications. *Information Management & Computer Security*. 1 (1), 22-28.
[12] Jordan, M. I. Bishop, C. M. (1997) 03:1996. *In Tucker, A. B. (Ed.)*. The Computer Science and Engineering Handbook: CRC Press.
[13] Jans, M. Lybaert, N. Vanhoof, K. (2009). A framework for internal fraud risk reduction at IT intergrating business processes: The IFR2 Framework. *The International Journal of Digital Accounting Research*. 9 (1), 1-29.
[14] Widder, A. Ammon, R. Schaeffer, P. Wolff, C (2008). Combining Discriminant Analysis and Neural networks for Fraud Detection on the Base of Complex Event Processing. *In: GECCO*, march 04-06, Gaithersburg, USA. pp.n/a.
[15] Roberts, S. Penny, W. (1997). Neural Networks: Friend or Foes?. *Sensor Review*. 17 (1), 64-70.
[16] Zadeh, L. (1965). "Fuzzy Sets", Information and Control, Vol. 8 No. 3, pp 338-353
[17] Gonzalez, F. Dasgupta, D (2002). D. An immunity-based technique to characterize intrusions in computer networks. IEEE Transactions on Evolutionary Computation, 6(3):281-291
[18] Zimmermann, H. (2001). *Fuzzy Set Theory and its Applications*, Springer.
[19] Griesser, J.W. (1992). "Experts among Us", *Business Horizons,* Vol.35 No. 3, May-June, 77-80.
20 Barr, A. Feigenbaum, E.A. (1982). *The Handbook of Artificial Intelligence,* William Kaufman, Menlo Park, CA, Vol. 2.
[21] *ACI Worldwide* [online]. (2008). Available from: http://www.aciworldwide.com/downloads/EN-NeuralNetworks.pdf.
[22] Neal, R. M. (1996) *Bayesian Learning for Neural Networks*, Lecture Notes in Statistics No. 118
[23] Buntine, W. L. Weigend, A. S. (1991) "Bayesian back-propagation", Complex Systems, 5, 603-643.
[24] Maes, S. Tuyls, K., Vanschoenwinkel, B. Manderick, B. (2002). Credit Card Fraud Detection using Bayesian and Neural Networks, Proc. of the 1st International NAISO Congress on Neuro Fuzzy Technologies.
[25] Phua, C., Alahakoon, D., & V Lee. 2004. 'Minority Report in Fraud Detection: Classification of Skewed Data'. *ACM SIGKDD Explorations: Special Issue on Imbalanced Data Sets*, 6; 50-59.
[26] http://dl.acm.org/citation.cfm?id=1712801

Contact us at:

✉ **marketing@Provenir.com**

@ProvenirGlobal

Provenir Ltd.

PROVENIR