Credit Card Fraud Detection using Bayes Theorem

Juliet Chebet Moso
Dedan Kimathi University of Technology

Jonah Kipcirchir Kenei
University of Nairobi
Email: jonah.kenei [AT] gmail.com

Abstract— **Fraudsters are masters at devising new methods of fabricating transactions thus requiring a consistent development and advancement of techniques for detecting and mitigating the falsifications. Numerous strategies have been proposed and used in the identification and mitigation of fraudulent transactions. Fraud detection and prevention involves analysis of the spending behavior of customers with the main aim being to uncover undesirable behaviour. It is focused on identification of suspicious events in an expeditious manner. Bayesian networks are suitable for circumstances where some data is already known and received data is partially unavailable or uncertain. The objective of utilizing Bayes rule is based on its ability to accurately predict the value of a selected discrete class variable given a set of attributes. The naïve Bayes technique is preferred due to its simplicity in dealing with training data and also its ability to handle missing values.**

Keywords- Bayes rule, fraud detection, Classification, Bayes belief networks

I. INTRODUCTION

There has been a steady growth in the usage of credit and debit cards in the recent past with most customers preferring to use these cards as opposed to using cash. This growth however has led to a significant growth in the cases of fraudulent transactions completed using these cards. The last decade has also witnessed a significant growth in communication technologies which have resulted in an increase in spending power due to the ease of doing business. This has however had a down side in that criminals have become more sophisticated with more ways of committing scams available to them. According to the Concise Oxford Dictionary fraud can be defined as "*criminal deception; the use of false representations to gain an unjust advantage*".

Credit card fraud occurs when an unauthorized person steals a credit card, gains access to credit card information, or Personal Identification Number (PIN), and makes use of this information to purchase items or withdraw money from an Automated Teller Machine (ATM). Fraud detection and prevention involves analysis of the spending behavior of customers with the main aim being to uncover undesirable behaviour or spending patterns. It is mainly focused on identification of suspicious events in an expeditious manner. Millions of dollars are lost annually through fraudulent schemes with the figures increasing yearly due to the innovation of new methods of committing this vice.

According to Allan (2014) 9.4 million dollars was lost to fraudsters by commercial banks in Kenya in the first half of 2014 who working with the bank staff were able to exploit gaps in the institutions' online banking platforms. Furthermore, based on Banking Fraud Investigations Department (BFID) data 525 fraud cases in financial institution resulted in a loss of 8.5 million dollars in the first three months of 2014. However, with the introduction of the chip and PIN debit cards in the second quarter of 2014 and consumer awareness campaigns the amount lost through fraudulent schemes dropped to 907,797 dollars. The data indicates that between January and June 2014, money lost through card fraud was at $112,773.

Based on BFID data 17.52 million dollars was embezzled from customers' bank accounts between April 2012 and April 2013 (Mwaura 2013). The data further shows that 11.2 million dollars was lost in the period between November 2012 and April 2013. The investigation agency BFID identified various schemes and techniques used to commit fraud which included identity theft, electronic funds transfer , credit card fraud , bad cheques, loan fraud, falsification of documents and online fraud (Mwaura 2013; Allan 2014). The rising instances of fraud and cybercrime imply that finance institutions need to urgently spend money on detection and precautionary systems.

Bayesian Networks (BNs) also called belief networks, are probabilistic graphical models, extensively used in knowledge representation and reasoning under uncertainty, where the each node depicts a variable and directed links between the nodes showing the relationship between them. A BN comprises of a directed acyclic graph of 'nodes' and 'connections' that conceptualize a system where the value of each node is defined in terms of different, mutually exclusive, 'states' (McCann, Marcot & Ellis 2006). Dependencies between variables are represented using conditional probability distributions which in turn describe the relationships between nodes. The full specification is as follows (eds. Stuart & Peter 2010):

i. Each node in the belief network represents a discrete or continuous variable.

ii. The parent child relationship is represented by directed links from the parent node to the child node forming a directed acyclic graph.

iii. Each node is assigned a conditional probability distribution value which measures the effect of the parents on the specific node.

The more general case of Bayes' rule for cases where there is more than one variable is given as follows:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

The above equation represents a set of equations, with each equation dealing with specific values of the variables. In a situation where there exists some background evidence the equation may be rewritten as follows:

$$P(Y|X,e) = \frac{P(X|Y,e)P(Y,e)}{P(X,e)}$$

The structure of a belief network stipulates the conditional independence that exists between the various entities forming the network. Given two nodes X and Y, a directed link from X to Y shows a causal relationship with an action performed on X directly affecting Y. After a belief network is constructed it is necessary to compute the conditional probability distribution for each variable within the network. The full joint distribution for all the variables is finally specified (implicitly) while considering both the network topology and the conditional probability distributions of each variable.

BNs are suitable in scenarios where the data being classified contains information which is already known and new instances/data is partially unavailable or contains some level of uncertainty (Sherly 2012; Suvasini *et al*. 2009). The objective of utilizing Bayes rule is based on its ability to accurately predict the value of a selected discrete class variable given a set of attributes (Joseph 2011; Manoel, Xidi & Alair 2008).

II. PROBLEM STATEMENT

The main objective is to demonstrate how Bayes theorem can be used to identify falsified credit card transactions given a set of training data. The objective of utilizing Bayes rule is based on its ability to accurately predict the value of a selected discrete class variable given a set of attributes. Suppose we have two classes C_1, C_2 representing fraudulent and legal transactions respectively. Given an instance $X = (X_1, X_2,\ldots,X_n)$ with each item characterized by an attribute vector $Z= (Z_1, Z_2,\ldots,Z_n)$. Bayes' theorem can be used to compute the maximum probability of each class given the instances (i.e. P(Ci|X)) using the steps outlined below:

1. Given the hypothesis fraud (F) and legal (L), probabilities are computed as follows:

$$P(F|X) = \frac{P(X|F)P(F)}{P(X)}$$
$$P(L|X) = \frac{P(X|L)P(L)}{P(X)}$$

   Since Naïve Bayes assumes independence *P(X)* is dropped since it is constant for all classes, leaving [*P(X|F) P(F)*] and [*P(X|L) P(L)*] as the significant terms to be computed.
2. Next, the computation of the class prior probabilities is done as follows:

   $P(F) = y_i / y$

   Where, y is the total number of training examples and y_i is the total number of fraudulent transactions in the dataset.
3. A basic assumption is made on the independence of attributes as follows.

$$P(X|F) = \prod_{k=1}^{n} P(x_k|F)$$
$$P(X|L) = \prod_{k=1}^{n} P(x_k|L)$$

The probabilities $P(x_1|F)$, $P(x_2|F$,may be computed from the training dataset as follows:

$P(x_k|F)=y_{ik}/y_i$

Where y_i is the number of fraudulent transactions in the training dataset and y_{ik} is the number of training examples for the class with value x_k for Z_k

III. BACKGROUND AND RELATED WORK

Fraud detection is a continuously evolving discipline with criminals adapting new strategies every so often. There is also a continuous entry of new criminals into the field that are not aware of the mechanisms that have been put in place to identify fraudulent activities and as such they commit fraud using methods which are easily detectable. It is therefore prudent to apply earlier detection tools as well as the latest developments. Presently, financial institutions use "if-then" rules for fraud detection which fire based on the set conditions being met with a transaction suspected to be fraudulent being rejected or an alert issued requiring further examination.

Rule based systems perform very well in scenarios where fraud patterns do not change but might not suffice in a dynamic environment which requires other techniques which can be able to detect changing fraud pattern. Artificial intelligence as an ongoing area of research has recently seen its applicability in fraud detection. Nonetheless, there has been a slow growth in the development of methods directly targeting credit card fraud due to privacy issues affecting financial data and lack of public databases (Bolton *et al*. 2002). Even so, the research on credit card fraud detection is growing with artificial intelligence methods being successfully applied with specific emphasis on the use of neural networks (Maes *et al.* 2002), artificial immune systems (Gadi, Wang & do Lago 2008), peer group analysis (Weston *et al*. 2008), association rules (Sánchez et al. 2009), Bayesian learning (Maes *et al*. 2002) and support vector machines (Bhattacharyya et al. 2011).

The Bayesian belief network was formally presented by Cooper and Herskovits (1992). BBNs are great tools for summarizing evidence of causal relationships in form of a network of probabilities. According to Heckerman, Geiger and Chickering (1995), the BN has become a popular representation for encoding uncertain expert knowledge in expert systems. They are able to handle incomplete data sets and represent causal relationships. BBNs are best used in scenarios where information is imprecise, uncertain, incomplete and conflicting.

Maes *et al*. (2002) applied artificial neural networks (ANN) and Bayesian belief networks (BBN) to detect fraud on a real world dataset provided by Europay International. Their approach used STAGE and backpropagation algorithm to identify fraudulent transactions. The best prediction rate was obtained for the experiment in which the features were pre-

processed. It was found that by performing a correlation analysis on the features and removing the feature that was strongly correlated with many of the other features, clear improvements to the results were obtained. Moreover, their results showed that BBNs yielded better fraud detection results with a shorter training period is shorter, though ANN could compute fraud predictions faster in the testing stage.

Ezawa and Norton (1996) argued that neural networks, regression and nearest-neighbour classifiers can be very slow while decision trees may not be able to adequately represent certain discrete variables. In their implementation of Bayesian network models on telecommunications dataset they discovered that the model with the greatest number pf variables and some dependencies gave the best classification performance. Viaene, Derrig and Dedene (2004) applied AdaBoosted naive Bayes scoring where they used weights in their preparation of evidence. The weights facilitated the calculation of comparative prominence of each component and demonstrating the combination of evidence for and contra fraud as a balance of evidence. Their framework exhibited better accuracy and an improvement on the cross entropy and Brier scores when compared to unboosted and boosted naive Bayes.

Chan *et al*. (1999) in their implementation of four classifiers C4.5, CART, naive Bayes and RIPPER on heterogeneous datasets using stacking and pruning of the base classifiers were able to achieve high cost savings and better efficiency on credit card transactions. Phua, Alahakoon and Lee (2004) in their analysis of automobile insurance claims dataset using C4.5, backpropagation neural networks and naive Bayes classifiers with minority oversampling with replacement were able to yield the best cost savings though stacking and bagging.

The ratio of fraudulent transactions in most databases applied in credit card fraud detection systems usually ranges between 0.005% and 0.5% (Gadi, Wang & do Lago 2008; Bhattacharyya et al. 2011). This poses a challenge especially in the training phase of diverse algorithms (Hastie & Tibshirani 2009). To mitigate this under-sampling procedures are mostly adopted with subsets of the databases being designed with higher ratio of fraudulent cases (Hulse & Khoshgoftaar 2007). The distinctiveness of credit card fraud is that erroneously predicting a falsified transaction as authentic can be a very costly affair with great potential of loss of money.

Credit card fraud is perpetrated using different methods and techniques. Ashish & Jagdish (2014) cited various types of credit card fraud, which include: card theft, skimming, card not present, and identity theft involving cards which can take the form of an application fraud or account takeover. Pratiksha and Tarun (2013) further classified different types of fraud based on types of transactions performed as shown in figure 1 below.

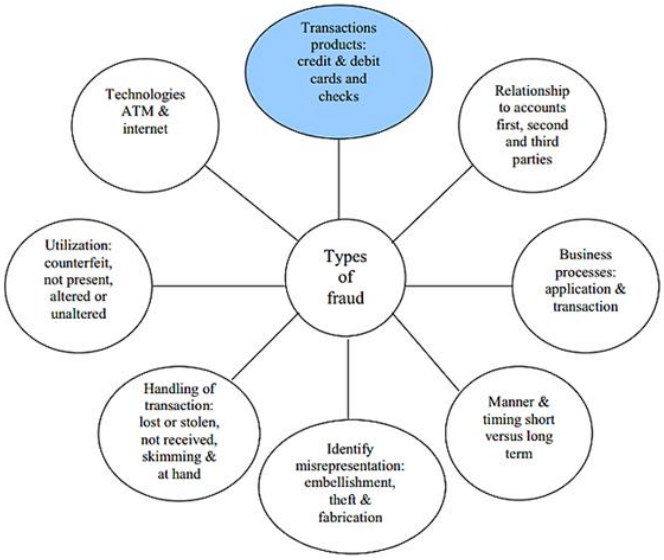

Figure 1: Various transactions and frauds (Pratiksha & Tarun 2013)

BNs can ease the acquisition of knowledge from variables which exhibit causal relationships (Uusitalo, 2007). The topological structure of a BN visibly shows the associations between different system components. This system structure can enable dialogue with people from different disciplines including stakeholder participation (Martín de Santa Olalla et al. 2005). Bayesian Networks can be updated easily using inference mechanisms when new knowledge is introduced (Ticehurst et al. 2008).

BNs are very useful tools when addressing data with uncertainty and also for relating model simulation, observations and expert knowledge (Uusitalo, 2007). BNs are able to acquire knowledge on the structure and parameters of a system from observed data. This knowledge can disclose the causal relationships between variables as well as their dependence and independence. The 'optimal' BN structure can be estimated using the highest probability score for possible candidate structures (Norsys, 2005). Kontkanen et al. (cited in Uusitalo, 2007) proved that given small sample sizes BNs are known to produce good prediction accuracy

Despite the fact that expert knowledge can effectively be represented using Bayesian models, a challenge exists in getting the experts to agree on the structure of the models and the selection criteria of the nodes based on their significance to the problem being solved. In addition, expressing knowledge in form of probability distributions may prove difficult to the experts (Uusitalo, 2007). Knowledge acquisition from experts generally follows an iterative procedure with the nodes, states and interrelationship in the BN being well understood first before the descriptions of distributions and confidence intervals of variables (Pollino, 2008).

Continuous data may require discretization since a number of BN software packages have limited capability for dealing with these kinds of data. Discretization breaks up the continuous variable into discrete interval values which can

effectively be used to offer some control on the size of the network, though the original distribution of the variable may not be fully represented resulting to a lower precision of variable values (Nyberg, Marcot & Sulyma 2006). According to Barton et al. (2008) discretization assumptions can considerably affect the outcome estimates. The acyclic property of BNs is essential in calculation of probability calculus, but the feedback effects cannot be included in the network (Barton *et al*. 2008).

IV. METHODOLOGY

To model fraud detection, two Bayesian networks were used to express the behaviour of users. One Bayesian network is modelled to represent the behaviour of a fraudulent (F) user and the second network represents a legitimate (L) user. In implementing the network the user behaviour x derived from his credit card transaction statistics is used and the probability of the evidence x relative to the two hypotheses (fraud and legal) is obtained. This means, it gives conclusions as to the extent the user behaviour can be considered fraudulent or non-fraudulent

Assuming the likelihood of fraud is P(F) and legal is P(L) = 1- P(F) using Bayes' rule, the probability of fraud, given the evidence x is given by:

$$P(F|X) = \frac{P(X|F)P(F)}{P(X)}$$

Where the denominator p(x) is given by:

$$P(X) = P(F)P(X|F) + P(L)P(X|L)$$

Given two classes, C_1 representing fraud and C_2 representing legal transactions and an instance X = (X_1, X_2,…,X_n) with each item characterized by an attribute vector Z= (Z_1, Z_2,…,Z_n). Bayes' theorem can be used to compute the maximum probability of each class given the instances (i.e. $P(C_i|X)$) using the steps outlined below:

1. Given the hypothesis fraud (F) and legal (L), probabilities are computed as follows:

$$P(F|X) = \frac{P(X|F)P(F)}{P(X)}$$
$$P(L|X) = \frac{P(X|L)P(L)}{P(X)}$$

   Since Naïve Bayes assumes independence *P(X)* is dropped since it is constant for all classes, leaving [*P(X|F) P(F)*] and [*P(X|L) P(L)*] as the significant terms to be computed.
2. Next, the computation of the class prior probabilities is done as follows:

   $P(F) = y_i / y$

   Where, y is the total number of training examples and y_i is the total number of fraudulent transactions in the dataset.
3. A basic assumption is made on the independence of attributes as follows.

$$P(X|F) = \prod_{k=1}^{n} P(x_k|F)$$
$$P(X|L) = \prod_{k=1}^{n} P(x_k|L)$$

The probabilities $P(x_1|F)$, $P(x_2|F$,may be computed from the training dataset as follows:

$P(x_k|F)=y_{ik}/y_i$

Where y_i is the number of fraudulent transactions in the training dataset and y_{ik} is the number of training examples for the class with value x_k for Z_k

V. RESULTS AND DISCUSSION

Real world credit card datasets are not made readily available by financial institution due to customer privacy policies. Due to unavailability of actual credit card transaction data, we used a synthetic data set generated from a random sample which represents a hypothetical situation.

Table 1: Training data set

	Cardholder Name	Gender	Age	Amount	Foreign Purchase	Internet Purchase	Class
1	Robert Mwiti	M	25	50000	1	1	legal
2	Ashley Juma	F	32	20500	0	1	fraud
3	Thomas Wafula	M	40	14700	0	0	legal
4	Susan Njeri	F	35	39600	0	1	legal
5	Christopher Diaz	M	22	12800	0	0	legal
6	Zakayo Sigei	M	36	15000	0	0	legal
7	Samuel Osoo	M	21	60000	1	1	fraud
8	Hezbon Odoyo	M	39	70000	1	0	legal
9	Nick Opamba	M	29	3000	0	1	legal
10	Rose Nyawira	F	33	18200	0	1	legal
11	Kevin Nyamunga	M	42	75649	0	1	fraud
12	Saingodeu Moses	M	42	50609	1	1	legal
13	Jonathan Mutai	M	37	42811	1	0	legal
14	Margaret Miyon	F	32	12558	1	0	fraud
15	Byron Ouma	M	28	4928	1	0	legal
16	Ezekiel Mpating	M	42	31683	1	0	legal
17	Patrick Songa	M	48	41970	0	1	legal
18	Maggie Gwiyo	F	49	43006	1	1	fraud
19	Michael Kioko	M	32	7503	0	0	legal
20	Daniel Onyango	M	27	77634	0	1	fraud
21	Brenda Mwita	F	31	66983	0	1	fraud
22	Mati Nzia	M	23	4657	0	0	legal
23	Azizi Manda	F	34	12723	0	1	legal
24	Jonathan Munene	M	52	42767	1	0	legal
25	Jane Njahira	F	47	24295	1	0	legal
26	Kigen Geoffrey	M	51	24228	1	0	legal
27	David Ndubi	M	36	14303	1	0	legal
28	Emmanuel Tibo	M	27	33010	1	0	legal
29	Nancy Kemoi	F	38	15990	1	0	legal
30	Jasmine Cheluget	F	41	21305	0	1	legal
T	Mary Wairimu	F	50	48613	1	0	?

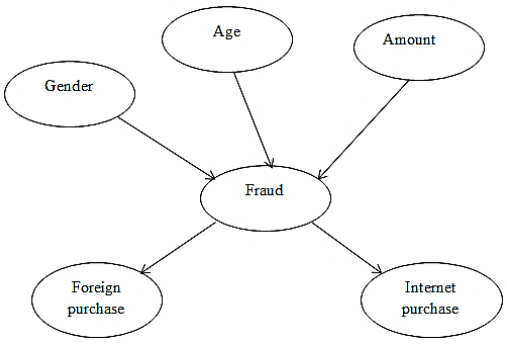

Figure 2: Bayesian network for credit card fraud detection

The sample data set in table 1 contains both continuous and discrete valued attributes. The continuous attributes, age and amount are discretized to fixed interval values to facilitate calculation of probabilities using Bayesian theorem which works best with discrete valued attributes. The values, counts and probabilities of each attribute are shown in table 2 below:

Table 2 Prior probabilities associated with each attribute

Attribute	Value	Count		Probabilities	
		legal	fraud	legal	fraud
Gender	M	17	3	0.74	0.43
	F	6	4	0.26	0.57
Age	21-25	3	1	0.13	0.14
	26-30	3	1	0.13	0.14
	31-35	4	3	0.17	0.43
	36-40	6	0	0.26	0.00
	41-45	3	1	0.13	0.14
	46-50	2	1	0.09	0.14
	51-55	2	0	0.09	0.00
Amount	< 20000	11	1	0.48	0.14
	20000-40000	6	1	0.26	0.14
	40001-60000	5	2	0.22	0.29
	> 60000	1	3	0.04	0.43
Foreign Purchase	0	11	4	0.48	0.57
	1	12	3	0.52	0.43
Internet Purchase	0	15	1	0.65	0.14
	1	8	6	0.35	0.86

From table 2, using Bayesian rule we can calculate the class of each instance as either fraud or legal. The prior probabilities are calculated using the following equations:

P(fraud) = yi / y = 7/30 = 0.23

P(legal) = yi / y =23/30 = 0.77

Using the values in table 2 we can be able to categorize a new transaction. Assuming, we are given a new transaction T = (Mary Wairimu, F, 50), using the probability values for gender and age we can be able to classify this transaction as follows:

P(T |legal) = 0.26 * 0.09 = 0.0234

P(T|fraud) = 0.57 * 0.14 = 0.0798

Thus, the likelihood of this transaction being legitimate = 0.0234 *0.77=0.0180

Likelihood of it being a fraud = 0.0798 *0.23= 0.0184

To compute the probability P(T), we do a summation of the likelihood values as follows:

P(T) = 0.0180 + 0.0184 = 0.0364

The actual probabilities of each event occurring is given by:

P(legal | T) = 0.0180/ 0.0364= 0.495

P(fraud |T) = 0.0184 / 0.0364= 0.505

Consequently, by considering the actual probabilities computed above we can classify the new transaction T as fraud since it has the highest probability value of 0.505.

The naïve Bayes methodology is a simple and user-friendly method requiring a single scan of the training data during classification. It is also able to handle datasets with missing values by omitting the probabilities of the incomplete entries during computations of likelihood of membership in each class. Despite this simplicity, it may not always give the best results due to the fact that attributes may be dependent on each other, which may be solved by ignoring the dependent attributes. This approach also does not handle continuous data, a challenge which is solved through discretization of the values of the dataset which in itself may be tiresome and may influence the accuracy of the results generated.

VI. CONCLUSION

Undoubtedly, credit card fraud is a deed of criminal duplicity. In this research various types of fraud have been discussed such as Card theft, Skimming, Card not present, Application fraud and Account takeover. As a matter of principle it is expected that banks and credit card companies should endeavor to identify all fraudulent cases. However, the economics of implementing a fraud detection system may be highly influenced by the level of skill of the fraudster being targeted.

Every cardholder exhibits a unique spending pattern which can effectively be used to generate their activity profile. Most of the current fraud detection methods consider the activity profiles when generating rules based on the spending patterns. However, should a cardholder acquire new tastes resulting in new spending patterns the detection system may not be effective since the rules are generally static. The ideal detection system should be able learn and adjust accordingly with the change in spending patterns so as to minimize loss and reduce the number of false alarms.

A fraudster may also device new ways of avoiding detection for example by making a few costly purchases or a large number of inexpensive purchases. Hence, it is essential to develop fraud detection systems which can incorporate multiple evidences including patterns of genuine cardholders and those of fraudsters. BBNs are great tools for summarizing evidence of causal relationships in form of a network of probabilities. The BN has become a popular representation for encoding uncertain expert knowledge in expert systems. They are able to handle incomplete data sets and represent causal relationships. BBNs are best used in scenarios where information is imprecise, uncertain, incomplete and conflicting.

REFERENCES

[1] Allan, O 2014, 'Kenya's commercial banks lose $9.4m to fraud in just six months', The EastAfrican 15 November. Available from: <http://www.theeastafrican.co.ke >. [23 April 2018].

[2] Ashish, G & Jagdish, R 2014, 'Fraud Detection in credit Card Transaction Using Hybrid Model', International Journal Of Engineering And Computer Science ISSN:2319-7242, Vol. 3, no. 1, pp. 3730-3735

[3] Banking Fraud and Investigations Department, 2012, Quarterly Banking Fraud statistics, 4.

[4] Barton, DN, Saloranta, T, Moe, SJ, Eggestad , HO & Kuikka, S 2008, 'Bayesian belief networks as a meta-modelling tool in integrated river basin management -- Pros and cons in evaluating nutrient abatement decisions under uncertainty in a Norwegian river basin', Ecological Economics, 66, pp.91-104

[5] Bhattacharyya, S, Jha, S, Tharakunnel, K & Westland, JC 2011, 'Data mining for credit card fraud: A comparative study', Decision Support Systems, vol. 50, no. 3, pp. 602–613.

[6] Bolton, RJ, Hand, D, Provost, F & Breiman, L 2002 'Statistical Fraud Detection: A Review', Statistical Science, vol. 17, no. 3, pp. 235–255.

[7] Chan, P, Fan, W, Prodromidis, A & Stolfo, S 1999, 'Distributed Data Mining in Credit Card Fraud Detection', IEEE Intelligent Systems, vol. 14, pp. 67-74

[8] Cooper, GF & Herskovits, E 1992, 'A Bayesian method for the induction of probabilistic network from data', Machine Learning, pp. 309-347.

[9] Ezawa, K & Norton, S 1996, 'Constructing Bayesian Networks to Predict Uncollectible Telecommunications Accounts', IEEE Expert, pp. 45-51.

[10] Gadi, M, Wang, X & do Lago, A 2008, 'Credit card fraud detection with artificial immune system', Artificial Immune Systems.

[11] Hastie, T & Tibshirani, R 2009, 'The Elements of Statistical Learning: Data Mining, Inference, and Prediction'.

[12] Heckerman, D, Geiger, D & Chickering, DM 1995, 'Learning Bayesian networks: The combination of knowledge and statistical data', Machine Learning, vol.20, no. 3, pp.197-243.

[13] Hulse, JV & Khoshgoftaar, TM 2007, 'Experimental Perspectives on Learning from Imbalanced Data', International Conference on Machine Learning.

[14] Joseph, KP 2011, 'Improving Credit Card Fraud Detection using a Meta-Learning Strategy', A thesis submitted in conformity with the requirements for the degree of Master of Applied Science Graduate Department of Chemical Engineering and Applied Chemistry University of Toronto.

[15] Maes, S, Tuyls, K, Vanschoenwinkel, B & Manderick, B 2002, 'Credit card fraud detection using Bayesian and neural networks', in Proceedings of NF2002,

[16] Manoel, F, Xidi, W & Alair, P 2008, 'Comparison with Parametric Optimization in Credit Card Fraud Detection', pp. 279-285

[17] Martín De Santa Olalla, FJ, Domínguez, A, Artigao , A, Fabeiro, C & Ortega, JF 2005, 'Integrated water resources management of the Hydrogeological Unit "Eastern Mancha" using Bayesian Belief Networks', Agricultural Water Management, vol. 77,pp. 21-36

[18] McCann, RK, Marcot, BG & Ellis, R 2006, 'Bayesian belief networks: applications in ecology and natural resource management', Canadian Journal of Forest Research, vol. 36, pp. 3053-3062

[19] Mwaura, K 2013, 'Gone in 12 months: How fraudsters stole $17m from Kenya's banks', The EastAfrican 18 May. Available from: <http://www.theeastafrican.co.ke >. [23 April 2015].

[20] Nyberg, JB, Marcot, BG & Sulyma, R 2006, 'Using Bayesian belief networks in adaptive management', Canadian Journal of Forest Research, 36, 3104

[21] NORSYS 2005, Netica. www.norsys.com.

[22] Pollino, CA 2008, Application of Bayesian Networks in Natural Resource Management (SRES3035). 11-22 February 2008. Canberra, Australian National University.

[23] Pratiksha, LM & Tarun, Y 2013, 'Credit and ATM Card Fraud Prevention Using Multiple Cryptographic Algorithm', International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, no. 8, pp. 1300-1305

[24] Phua, C, Alahakoon, D & Lee, V 2004, 'Minority Report in Fraud Detection: Classification of Skewed Data', SIGKDD Explorations, vol. 6(1), pp. 50-59

[25] Ticehurst, JL, Letcher, RA & Rissik, D 2008, 'Integration modelling and decision support: the Coastal Lake Assessment and Management (CLAM) tool', Mathematics and Computers in Simulation.

[26] Sánchez, D, Vila, M, Cerda, L & Serrano, J 2009, 'Association rules applied to credit card fraud detection', Expert Systems with Applications, vol. 36, no. 2, pp. 3630–3640.

[27] Sherly, KK 2012, 'A comparative assessment of supervised data mining techniques for fraud prevention', TIST.Int.J.Sci.Tech.Res, Vol.1, pp.1-6.

[28] Stuart, JR & Peter, N (eds.) 2010, 'Artificial Intelligence A Modern Approach',3rd edition, Prentice Hall, Upper Saddle River, New Jersey, pp. 510-558

[29] Suvasini, P, Amlan, K, Shamik, S, Majumdar, AK 2009, 'Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning', Elsevier, Information Fusion, vol. 10, pp. 354–363

[30] Uusitalo, L 2007, 'Advantages and challenges of Bayesian networks in environmental modelling', Ecological Modelling, 203, pp. 312-318.

[31] Viaene, S, Derrig, R & Dedene, G 2004, 'A Case Study of Applying Boosting Naive Bayes to Claim Fraud Diagnosis', IEEE Transactions on Knowledge and Data Engineering, vol. 16(5), pp. 612-620.

[32] Weston, DJ, Hand, D, Adams, NM, Whitrow, C & Juszczak, P 2008, 'Plastic card fraud detection using peer group analysis', Advances in Data Analysis and Classification, vol. 2, no. 1, pp. 45–62.