*IOSR Journal of Computer Engineering (IOSR-JCE)*
*e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 4, Ver. IV (Jul.-Aug. 2016), PP 26-32*
*www.iosrjournals.org*

# A Bayesian Classification Model for Fraud Detection over ATM Platforms

Milgo, Carolyne

**Abstract:** *The banking system relies greatly on the use of Automated Teller Machines, debit and credit cards as a vital element of its payment processing systems. The efficient functioning of payment processing systems allows transactions to be completed safely and on time thereby contributing to operational performance. Customer accounts are equally exposed to risk with respect to fraud. The major security risk identified is identity and access management of customer funds. This research sought to use classification techniques to implement a novel fraud detection solution to establish legitimacy of customer transactions. It also sought to identify the main security issues encountered with use of ATM cards and establish internal control mechanisms which can be implemented to deter possibility of card fraud. The data was obtained from a local bank and was consequently preprocessed and fed into WEKA which was used to develop the training model which was to be used for classification of incoming transactions.*

**Keywords:** *Classification, Bayesian Network, Automated Teller Machines, Data Mining*

## I. Introduction

The architecture of ATM platforms is designed such that systems are linked directly with the bank's servers enabling customers withdraw cash through terminals and selected merchants. The use of these channels should exercise the same controls as a physical cashier or teller in the bank. This implies that it must have the ability to recognize and authenticate the user as a bonafide customer of the bank in order to accept some form of instruction which the customer gives. This recognition is based on the classical secret PIN, which is the four digit secret code given to each customer. The recent wave of fraudulent attacks on customer accounts necessitates the need to implement novel fraud detection measures.

### Classification

Data mining is one area which facilitates the exploration of data to discover previously unknown patterns (Sentil, 2010). The specific data mining approach chosen is classification which entails defining classes such that objects can be classified into pre-defined classes. Classification begins with identification of classes within a data set. The user has to initially determine the classes so as to assign new instances to the predefined classes. Assigning an object or item to a class is based on degree of similarity to previous objects within the classes. A model is first built in a process referred to as training process. Using the training set, different algorithms can be used to find the class or category which the new data or instance points to. Within this research, the goal is to learn a classification approach from the data that can be used to predict new classes of instances or cases. Learning therefore requires a data set, D, a task, T and a performance measure, M. The algorithm should learn from D, to perform a task T. The task performed is measured by M.

### Existing Techniques of Handling Fraud

There various methods that have been advanced to handle fraud using credit cards. One method is the use of support vector machines (SVM). SVM is a discriminative classifier usually defined by a separating plane. It functions by non-linearly projecting the training data in the input space to a feature space of a higher dimension by making use of a kernel function thereby allowing data which is usually nonlinearly separable to be classified in the input space. SVMs are rooted on the concept of decision planes which represent decision boundaries. The decision plane serves as a boundary separating objects belonging to different class memberships. In simple terms, given a set of labeled training data, the algorithms should output an optimal hyper plane which categorizes new instances or objects. The main challenge with the use of SVMs lies in the choice of the kernel (Kulis, 2008). Furthermore, building the model is complex and entails time demanding calculations (Delen, 2014).

The decision tree has also been developed to handle fraud detection. The decision tree is a tree structure which attempts to separate given records into mutually exclusive subgroups. Starting from the root node, each node is split into child nodes in a binary or multi split fashion. The decision tree therefore represents a table of tree shapes connecting lines to nodes. Each node is either a branch node and is followed by more nodes or one node assigned by classification. Basically, the decision tree takes a set of features as well as their associated values of input and uses that to classify new cases by traversing the decision tree. The main disadvantage it

DOI: 10.9790/0661-1804042632

presents is instability since small changes or variations in the data may result in the generation of a completely different tree. Decision trees are also susceptible to the over fitting problem where the classifier generates a classifier which perfectly fits the training data but has lost the capability of generalizing instances not presented during training. The advantage of decision trees is with reference to its white box approach which represents techniques that are easily interpretable. The advantages of employing white box approaches include increasing user confidence in prediction and ease of grasping insight regarding the classification problem (Barros, Carvalho, & Freitas 2015). SVMs and neural networks are examples of techniques which use the black box approach where results are more difficult to interpret (Russell, Meadows & Russell, 2008; Tiwary, 2015; Pun, 2011). The Naïve Bayes is a supervised learning method which uses a training data set with known target classes to predict the classes of future instances. It is a powerful probabilistic method which is superior in its speed of learning. The naïve Bayes assumes that the presence or absence of a particular attribute of a set is not based on the presence or absence of any other attributes in the same set. The method is named 'naïve' because it naïvely assumes independence of the attributes given the class. Classification is done by applying the Bayes rule to compute the probability of the correct class given a set of attributes of a credit card transaction as follows

$$(Fraud\backslash Evidences) = \frac{P(Evidences\backslash Fraud)xP(Fraud)}{P(Evidences)}$$

An advantage of the Naïve Bayes classifier is that it only requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. It is simple to implement and fast to train. Naïve Bayes is considered robust to noisy data as well as missing data.

The K-*Nearest Algorithm is a lazy learning* algorithm that stores available instances and classifies new instances or cases based on a similarity measure. *k*NN is an example of an instance based learner. The nearest neighbor classification technique entails a process where each new instance is compared to existing ones using a distance metric. The closest existing instance is used to assign the class to the new one. The main disadvantage is with respect to the fact that it is computationally intensive, especially when the size of the training set grows. *k*NN is time consuming as the distances between the target instances and all the instances in the training data needs to be calculated. It does not perform well in high dimensional data and because of its lazy learning strategy.

## II. Bayesian Classification Model

The chosen classification technique is Bayesian Network. Bayesian networks are directed cyclic graphs which have nodes representing variables which may be observable parameters or unknown variables. Each node is associated with a probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node. The Bayesian Network provides a compact representation of joint probability distributions. The advantage is drawn from its representation power and generalization. Other predictive classification methods focus on learning only the relationship between the class label and input variables. In contrast, Bayesian Networks represent the joint probability distribution over the class label and input attributes. They also offer good generalization with limited training data (Mittal & Kassim, 2007).

Bayesian networks ease many of the theoretical and computational difficulties of rule based systems by utilizing graphical features for representing and managing probabilistic knowledge. They also have the ability to handle incomplete data sets. Bayesian networks can predict the class label when only partial information about the attribute is available. Bayesian belief networks are very effective for modeling situations where information about the past and/or the current situation is vague, incomplete, conflicting, and uncertain, whereas rule-based models result in ineffective or inaccurate predictions when the data is uncertain or unavailable (Mittal & Kassim, 2007). The process flow is indicated in figure 1.1 details the implementation of the solution

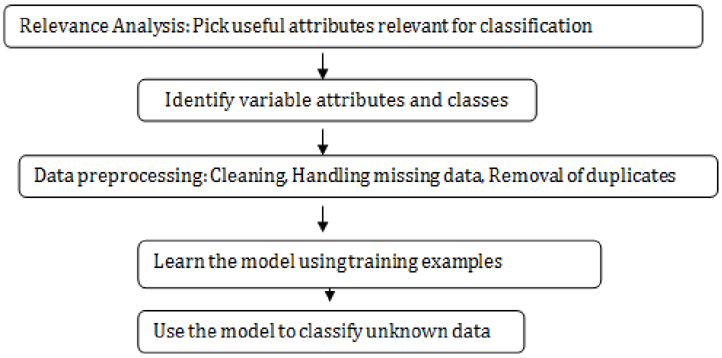

Relevance Analysis: Pick useful attributes relevant for classification
Identify variable attributes and classes
Data preprocessing: Cleaning, Handling missing data, Removal of duplicates
Learn the model using training examples
Use the model to classify unknown data

**Figure 1.1** The classification process

WEKA (Waikato Environment for Knowledge Analysis) is a machine learning toolkit developed by Waikato University, New Zealand. It is equipped with a collection of data preprocessing and modeling techniques. It has the ability to provide the missing value treatment, eliminate noise, split data, sort the sample and reduce dimensions. It equally supports combination of embedded algorithms. WEKA effectively delivers on data visualization, the mining process visualization as well as result visualization hence enhancing comprehension and evaluation.

The classification model was built using WEKA. It should be noted that WEKA offers a black box approach with respect to the computational details of the classifier. The Bayesian Network Classifier was chosen for this task. Ideally, almost every vast block of data contains anomalies is incomplete attribute wise and maybe even holds redundancies. The building of the purified training block for Bayesian Net Classifier incorporates three major parts: data validation attributes selection and handling of missing attributes. The attributes selected were considered of importance in the data mining process and building of the model. The model was developed using results from WEKA and thereafter imported to Java to allow classification of incoming transactions.

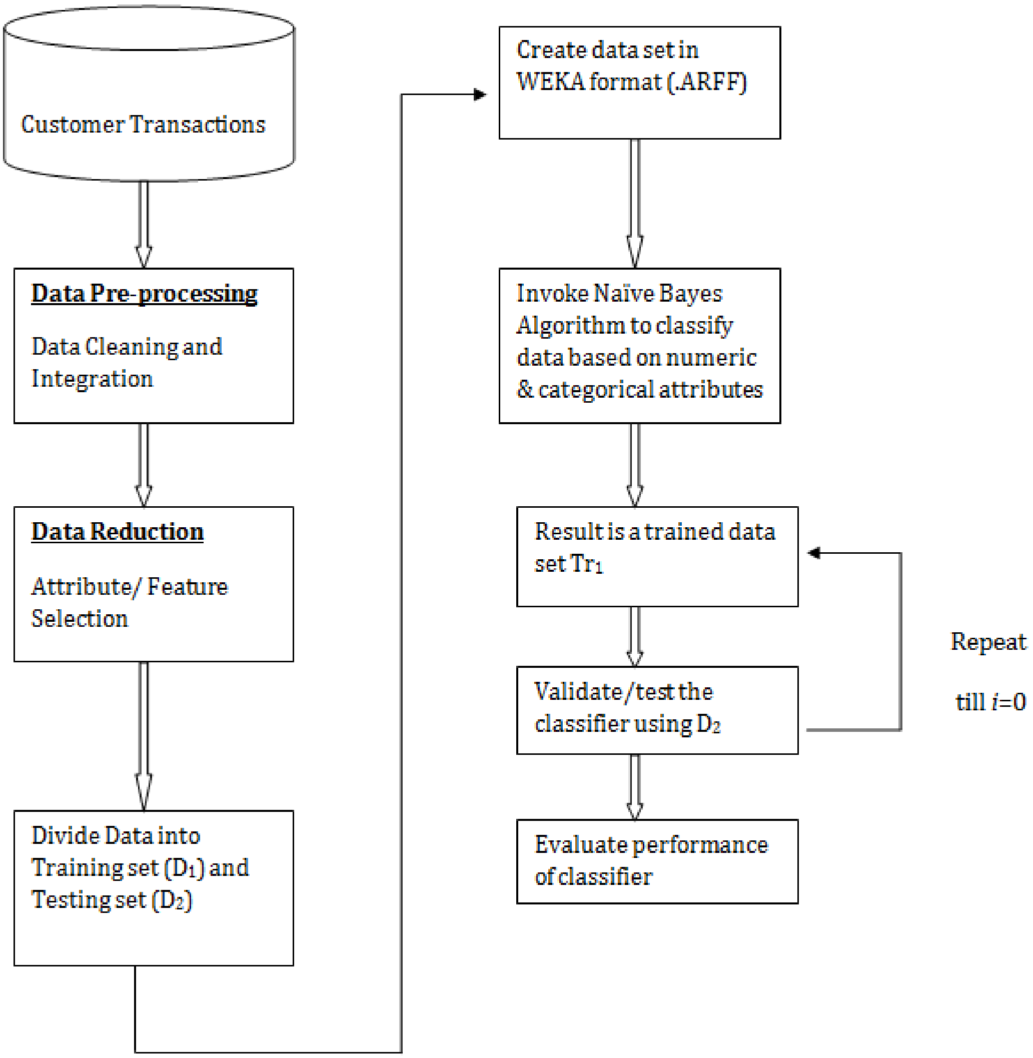

Customer Transactions
Data Pre-processing
Data Cleaning and Integration
Data Reduction
Attribute/ Feature Selection
Divide Data into Training set (D1) and Testing set (D2)
Create data set in WEKA format (.ARFF)
Invoke Naïve Bayes Algorithm to classify data based on numeric & categorical attributes
Result is a trained data set Tr1
Validate/test the classifier using D2
Evaluate performance of classifier
Repeat
till i=0

**Figure 1.2** Arbitrary working of the model

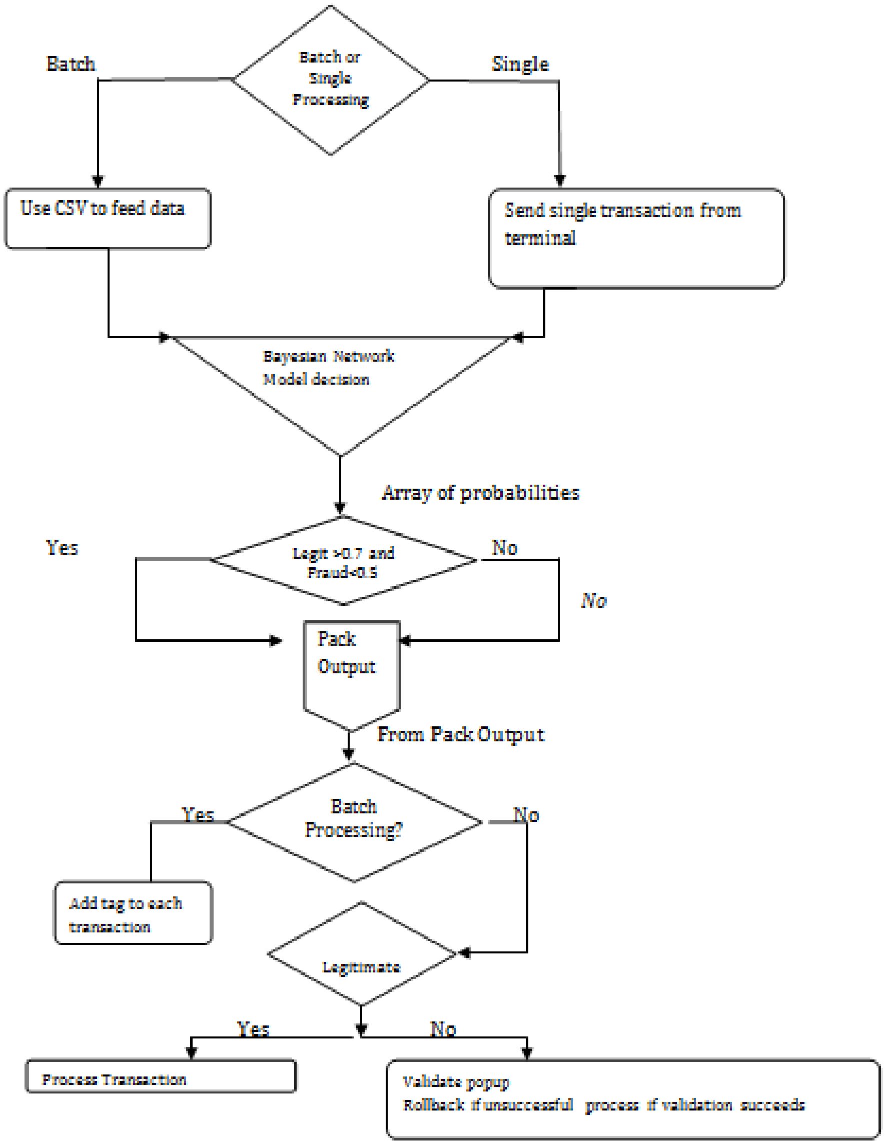


**Figure 1.3** Breakdown of the solution

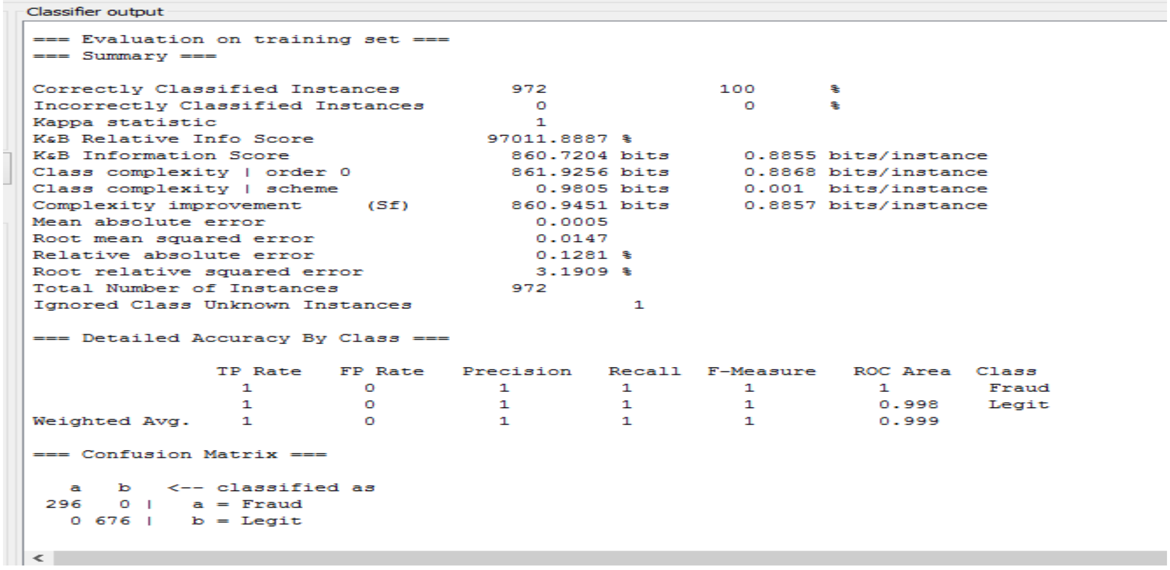
### Description of the Data

The data collected contains attributes relating to transactions. The identified attributes for the classifier include the transaction time, city, country, currency type, amount, transaction details, postings status represented by approved or rejected as well as return codes which indicate the reason for transactions being declined. The data set is labeled and thereby customer transactions can be identified as fraudulent or legitimate. The test data that is used to establish the accuracy of the data is also labeled only that the class labels are removed to measure how well the classifier performs.

The data obtained in .xls (spreadsheet) format was cleaned and coded and converted into .CSV files. The data was grouped according to the card number or target number which was a unique identification for each customer typically represented by a 12 digit code. WEKA permits processing of files in .CSV format. A total number of 972 instances was used to create the model. Based on the data collected, each customer had a value in the attributes identified for building the classifier.

### Evaluating the Classifier

The classifier revealed the following as shown in figure 1.4. The rate of incorrectly classified instances stood at 0% while the correctly classified instances were 100%. The percentage of correctly classified instances is often called the accuracy or the sample accuracy. The confusion matrix indicates 0 as the number of False Negatives and False Positives which means that there were no instances falsely classified into the defined classes. The error rates in WEKA are ideally used for numeric prediction and not the classification task itself.



Classifier output

```
=== Evaluation on training set ===
=== Summary ===

Correctly Classified Instances         972               100      %
Incorrectly Classified Instances         0                 0      %
Kappa statistic                          1
K&B Relative Info Score              97011.8887 %
K&B Information Score                  860.7204 bits      0.8855 bits/instance
Class complexity | order 0             861.9256 bits      0.8868 bits/instance
Class complexity | scheme                0.9805 bits      0.001  bits/instance
Complexity improvement     (Sf)        860.9451 bits      0.8857 bits/instance
Mean absolute error                      0.0005
Root mean squared error                  0.0147
Relative absolute error                  0.1281 %
Root relative squared error              3.1909 %
Total Number of Instances              972
Ignored Class Unknown Instances                  1

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 1         0          1          1         1          1        Fraud
                 1         0          1          1         1          0.998    Legit
Weighted Avg.    1         0          1          1         1          0.999

=== Confusion Matrix ===

   a   b   <-- classified as
 296   0 |   a = Fraud
   0 676 |   b = Legit
```

**Figure 1.4** Evaluation on the training set

### Introduction of the Test Set to Establish Accuracy

A test set in the format of a .CSV file was introduced to the model. The test set was used to evaluate the performance of the model. The test set had known classes (that is fraudulent and legitimate) which were ignored during the test phase to identify the possible classes which the model would classify the test data set. The data set for testing contained 1300 instances which were previously labeled. Upon running the classifier, the number of instances correctly classified amounted to 1103 instances. The remaining instances, 197, were incorrectly classified hence giving the classifier an accuracy of 84.85%. The incorrectly classified instances represent transactions which were legitimate classified as fraudulent or those that were fraudulent incorrectly classified as legitimate.

### Transaction Time Vs. Class

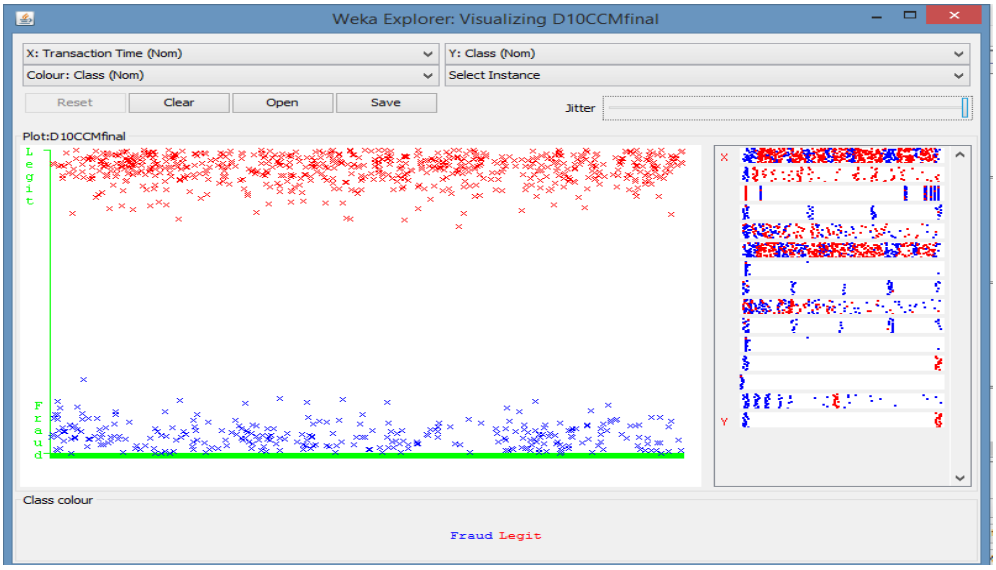
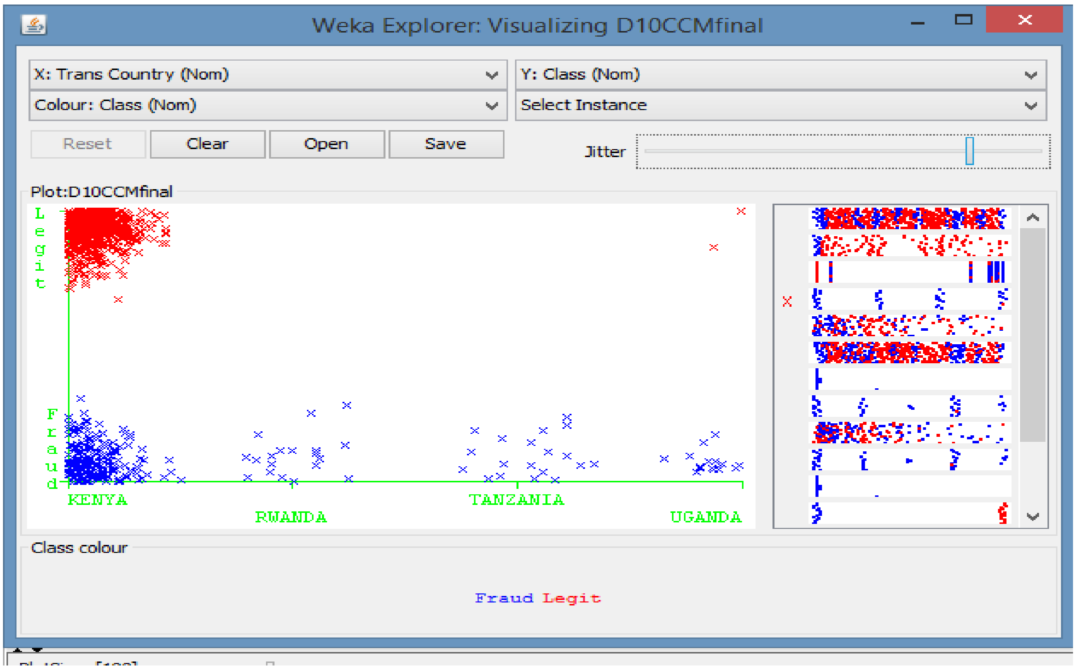


**Figure 1.5** Graph of transaction time Vs. class

The figure 1.5 depicts the fact that legitimate and fraudulent transactions did not have defined transaction times as captured in the data. Fraudulent and legitimate transactions thereby did not have unique times meaning that they had equal distribution irrespective of the class the transactions belonged to.

### Transaction country Vs. Class

Figure 1.6 demonstrates the visualization of the relation between the transaction country and the class indicates a higher level of fraudulent transactions for Kenyan Cardholders were perpetuated in countries outside Kenya notably Rwanda, Tanzania and Uganda. This means that there was a higher probability of a transaction being fraudulent if it was undertaken in the adjacent countries for ATM cards which belonged to Kenyan registered account holders.



Weka Explorer: Visualizing D10CCMfinal

X: Trans Country (Nom)

Y: Class (Nom)

Colour: Class (Nom)

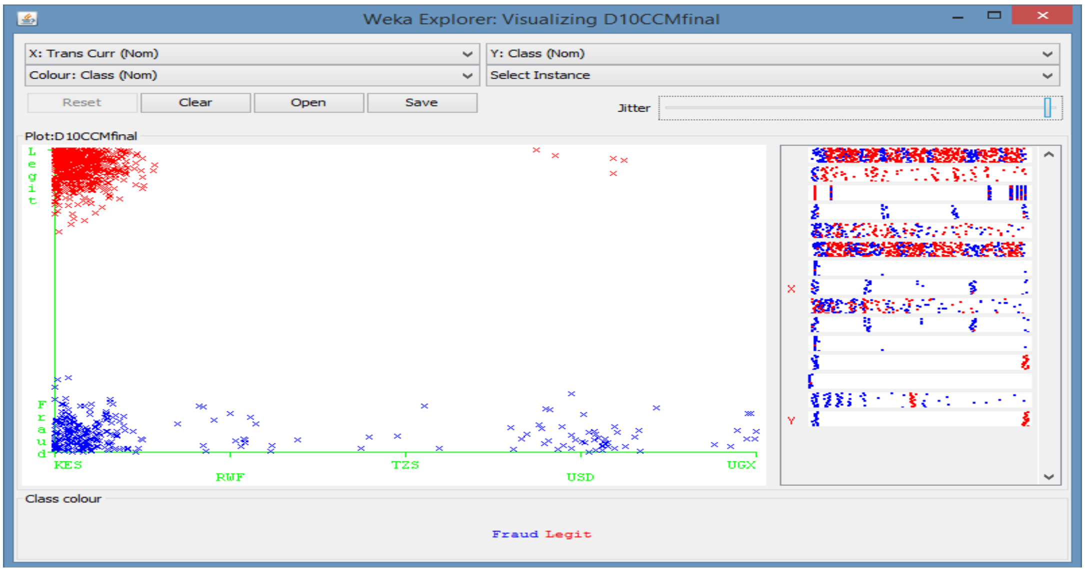
Select Instance

Plot:D10CCMfinal

**Figure 1.6** Graph of transaction country Vs. class

### Transaction Currency Vs. Class

The model generated by WEKA indicates that transactions often settled by the Kenya Shilling currency had chances of being fraudulent or legitimate as indicated in figure 1.7. Transactions settled in Rwandan Franc, Ugandan Shillings and Tanzanian Shillings were likely to be fraudulent as demonstrated in figure 1.7. Transactions involving the use of USD currency registered minimal legitimate transactions and a higher number of fraudulent transactions.



Weka Explorer: Visualizing D10CCMfinal

X: Trans Curr (Nom)

Y: Class (Nom)

Colour: Class (Nom)

Select Instance

Plot:D10CCMfinal

**Figure 1.7** Graph of transaction currency Vs. class

## III. Discussion of the Findings

The main security issues associated with the use of ATM based systems in card fraud entailed the existence of counterfeit cards which were reproduced for legitimate customers for use in unauthorized account access. The cards vulnerable for alteration were the cards not compliant with the EMV technology. Counterfeit card fraud basically entailed material change to a debit or credit card involving cloning and altering and replacing data contained in the card in order to allow illicit transactions concurrently or subsequently. Stolen cards were also a major area of concern as these cards were reported to have been used to perpetuate fraud. Card data compromise was also cited as a security issue which occurred through interception of card information from remote networks by fraudsters often using key-logging software. This gives the fraudster access to digitized card information which is useful in generating counterfeit cards.

The main security challenges identified in card fraud detection in Kenya included the lack of real time fraud detection mechanisms as banks relied heavily on complaints raised by customers and reports by the risk and fraud departments which identify possible high risk terminals and compromised merchants. The cost of implementing the fraud detection systems was also a challenge with respect to implementation. Most of the solutions available were off the shelf and were liable to integration and compatibility issues in the event that there was need to customize them and adopt them within the bank.

The research also sought to establish internal control mechanisms which banks used to deter possible fraud over ATM platforms. The study found that payment processing systems have been interfaced with mobile phone platforms such that each time a transaction is undertaken by a customer, an SMS alert is sent to the cardholder in a matter of seconds. These alerts cover transactions done over ATMs, tellers, point of sale terminals and any other purchases include on e-commerce platforms. In the event that the cardholder receives an unrecognized transaction, the customer is usually required to contact the issuing bank on the numbers provided in the alert or the back of their cards which have help lines.

## IV. Conclusions

The Bayesian Network proves to be a powerful probabilistic method. The accuracy results obtained for this study are reliable indicating that its predictive capability is reliable. The existent control mechanisms were however; this was not entirely beneficial considering the need to authenticate the legitimacy of a transaction before authorization to debit the customers' account actually occurs. Most of the fraud cases reported was only discovered once the transaction had been completed. This model thereby sought to ensure that the fraudulent or legitimate nature of the transaction be established before allowing the transaction to complete. Upon detection of a transaction as being fraudulent, the model provides further analysis to ascertain the true state of the transaction.

This is undertaken by use of defined random questions which are drawn from the customer's account details. This serves as a control mechanism to deter possible fraud as in the event that the customer fails to provide the correct answer to the questions posed, the transaction is declined and marked as fraudulent. Any wrong answer leads to the transaction being rolled back and an alarm can be raised for that particular bank client. If the right answers are given the transaction is carried on.

Recommended areas for further research are in the areas of tokenization and the use of behavioral analytics to authenticate the identity of customers. Despite the use of EMV (Europay, MasterCard and Visa) technology in Kenya, tokenization serves as a complement to EMV. EMV secures the communication between card and POS terminal, EMV does not encrypt the data hence this is where tokenization comes in. Behavioral analytics is another area which can assist merchants and card issuers tackle fraud. It is a fraud mitigation technology which uses tools to detect fraud by monitoring the user session and transactions in a bid to unearth suspicious patterns.

## References

[1]. Barros, C., Carvalho, A. , & Freitas, A. (2015). *Automatic Design of Decision-Tree Induction Algorithms.* New York, NY: Springer.

[2]. Delen, D. (2014). *Real-World Data Mining: Applied Business Analytics and Decision Making*. New Jersey, NJ: . FT Press.

[3]. Kulis, B. (2008). *Scalable Kernel Methods for Machine Learning*. ProQuest.

[4]. Mittal,A. , & Kassim, A. (2007). *Bayesian Network Technologies: Applications and Graphical Models: Applications and Graphical Models*. New Tork, NY: Idea Group Inc.

[5]. Pun, J. (2011). *Improving Credit Card Fraud Detection using a Meta-Learning Strategy. Published Thesis.*Graduate Department of Chemical Engineering and Applied Chemistry. University of Toronto.

[6]. Russell,S., Meadows, A. , & Russell R. (2008). *Microarray Technology in Practic*e. Cambridge, UK: Academic Press.

[7]. Tiwary, C. (2015). *Learning Apache Mahout*. .Packt Publishing.