Neural Network and Classification Approach in Identifying Customer Behaviour in the Banking Sector: A Case Study of an International Bank

Francisca Nonyelum Ogwueleka[1], Sanjay Misra[2], Ricardo Colomo-Palacios[3], Luis Fernandez[4]

[1]Department of Computer Science, Federal University of Technology, Minna
[2]Department of Computer Engineering, Atilim University, Ankara, Turkey
3 Department of Computer Science, *Universidad Carlos III de Madrid, Spain*
4 Universidad de Alcalá, Depto. de Ciencias de la Computación, Alcalá de Henares, 28400 Madrd, Spain

Abstract

The customer relationship focus for banks is in development of main competencies and strategies of building strong profitable customer relationship through considering and managing the customer impression, influence on the culture of the bank, satisfactory treatment; and assessment of valued relationship building. Artificial neural network (ANN) is used after data segmentation and classification, where the designed model register records into two class sets, that is, the training and testing sets. ANN predicts new customer behaviour observations from previous customer observed behaviour after executing the process of learning from existing data. This paper proposes an ANN model, which is developed using a six-step procedure. The backpropagation algorithm is used to train the ANN by adjusting its weights to minimize the difference between the current ANN output and the desired output. An evaluation process is conducted to determine if the ANN has learned how to perform. The training process is halted periodically and its performance is tested until an acceptable result is obtained. The principles underlying detection software are grounded in classical statistical decision theory.

Keywords: Customer relationship management (CRM), artificial neural network (ANN), classification, banks, back propagation algorithm.

1. Introduction

A bank generates profits from transaction fees on financial services and from the interest it charges for lending. Many banks offer ancillary financial services to make additional profits such as selling insurance products, investment products or stock broking. Banks have vast databases and important business information can be extracted from these data stores for decision-making concerning the customer transaction behavioural pattern.

Banks are facing the increased competition due to two different reasons including: the entrance of financial and insurance firms in the traditional banking market, and the wide range of offered products and services to public. As a consequence the banking industry strives to succeed by putting the topic of rapid and changing customers needs to their agenda (Krishnan, Ramaswamy, Meyer & Damien, 1999).

Customer Relationship Management (CRM) first, seeks on how to get closer to the customer by utilizing the hidden data in broaden databases and then transform the company into customer-centric organizations with a greater focus on customer profitability as compared to line profitability. CRM helps banks to improve the productivity of its interactions with customers while at the same time making the interactions seem supportive through individualization. To succeed with CRM, banks need to make their products and operational drive equivalent to prospects and customers, to cautiously manage the customer life cycle.

Objectives of CRM include increased cross-selling possibilities, better lead management, better customer response and improved customer loyalty (Chin, 2000). The potential areas of application of data mining techniques are practically without limit, including data coming from the social networks (García-Crespo et al., 2010). In banking services, a wide variety of data analysis solutions are provided by the banks' computer service personnel's and the bank's consultancy group in the banking section depending on the type of data analysis problems encountered. Examples are observed in comprehensive solution to address the needs of current ever-evolving banking business scenario with different frameworks to enable banks to critically analyze and evaluate various factors affecting their business transactions, thereby optimizing and enhancing the overall reporting capabilities of their composite application suites. By providing a single point of access as one of the data analysis solution, some of these frameworks greatly reduce the costs involved in identifying, gathering, and processing mission critical data. Some other data analysis framework was for calculation of statistical parameters, classification (to find patterns amongst data elements), stratification of numbers (to identify unusual i.e., excessively high or low entries), duplicate testing (to identify duplicate transactions such as payments, claims, etc), gap testing – to identify missing values in sequential data where there should be none, validating entry dates (to identify suspicious or inappropriate times for postings or data entry), etc. SAS software developed by Smith (2003) enables managers to analyze data from virtually any source to develop a deep understanding of customer behaviour, propensities, and profitability. Banks can identify their best customers, implement and measure strategies to retain them, cross-sell and up-sell to them, and make the most effective use of all available assets and channels.

In the area of CRM, the applications are customer segmentation, profitability, prospecting and acquisition, affinity and cross sell, retention and attrition, channel utilization, and risk analyses. The knowledge obtained from CRM enables banking industry to estimate the profitability of individual accounts. Customers are distinguished in terms of their profitability. Banks can also build predictive churn models to retain their best customers by identifying symptoms of dissatisfaction and churning.

Artificial neural networks are massively parallel-distributed processor that has the natural propensity for storing experiential knowledge and making it available for use (Haykin, 1999). The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. Further, the existing literature shows that ANN is applied in several areas however none of them applies ANN for predicting the behaviour of customers in banks to our knowledge. This becomes motivation for us to work in this area.

The objectives of this research are the use of classification and clustering methods on data available in the banking industry to identify abnormal patterns and trends; and the use of the identified patterns and trends relating to abnormal customer transaction activities to provide training to an artificial neural network, which will then learn to detect such abnormal behaviour.

This paper is organised in the following way. The section 2 produces the background of the work, which contains the related study, introduction to ANN, and its relation of CRM in banks. The proposed model and its application are given in section 3. The performance evaluation and the comparison with other methods are in section 4 and 5 respectively. The conclusions drawn are given in section 6.

2. Related Studies

Cardwatch (Aleskerov, Freisleben & Rao, 1997) features neural networks trained with the past data of a particular customer. It makes the network process the current spending patterns to detect possible anomalies. Brause and Langsdorf (1999) proposed the rule based association system combined with the neuro-adaptive approach. Falcon developed by Hecht-Nielsen Neurocomputer Corp (HNC) uses feed-forward artificial neural networks trained on a variant of a back propagation-training algorithm (Hassibi, 2000). A neural multi-layer perceptron (MLP) based classifier is another example using neural networks (Dorronsoro, Ginel, Sanchez & Cruz, 1997). It acts only on the information of the operation itself and of its immediate previous history, but not on historic databases of past cardholder activities. A parallel Granular Neural Network (GNN) method uses fuzzy neural network and rule-based approach (Syeda, Zhang & Pan, 2002). YongSong (2004) studied the effects of variable selection and class distribution on the performance of specific logic regression and ANN implementations in a CRM setting. Kwok, Choy, Lau and Kwok (2007) proposed a strategic customer relationship management system (SCRMs) with hybrid online analytical processing (OLAP) neural approach. Roland, Geier and Kolliker (2004) discussed and compared the best method for customer base between ANN, classification trees and regression but applied it in medicine. Feng & Linwen (2007) discussed improvement of customer satisfaction by the expert system using ANN. In a related work, Rodriguez and Edwards (2009) have presented a work on enterprise risk management (ERM). This paper discusses the application of knowledge in order to control deviations from strategic objectives, shareholders' values and stakeholders' relationships in financial sectors. Alis, Karakurt and Melli (2000) presented a work on data mining techniques for database marketing at Garanti Bank in Turkey. The authors have claimed that the bank has decided to pursue a more aggressive marketing strategy supported with the results of the analysis using data mining applications.

Customer satisfaction is recognized as being highly associated with customer value and with product price; whereas service quality is not generally considered to be dependent upon price. The more satisfied the customers, the more tolerant to price increases they are likely to be, thus resulting in greater profits (Anderson, Fornell, & Lehmann, 1994; Garvin, 1988). Several studies have investigated why individuals choose a specific bank. Important consumer selection factors include: convenience, service facilities, reputation, and interest rates (Kennington, Hill, & Rakowska, 1996; Lariviere & Poel, 2004). According to Delvin (1995) customers have less time to spend on activities such as visiting a bank and therefore want a higher degree of convenience and accessibility. Onut, Erdem, and Hosver (2008), presented a model for banking performance enhancement. This paper discusses the importance of CRM and its potential to acquire new customers, retains existing ones and maximizes their lifetime value.

Changes in customer behaviour are an inevitable aspect of surviving in a competitive and mature market (Zineldin, 1996). Krishnan, Ramaswamy, Meyer & Damien (1999), via a Bayesian analysis, found that satisfaction with product offerings is a primary driver of overall customer satisfaction. Classification has emerged as an important decision-making tool, and has been applied to a variety of problems in marketing, including customer classification (Sharma, 1994). As a conclusion, to predict a customer behavior is an important issue for good CRM in banking sector but not too many studies has been performed on this issue. Further, we have also observed that a very successful method, artificial neural network, which may provide an effective solution for this issue, is not applied yet. These all issues motivate us to apply ANN in predicting the behavior of customers in banking sectors.

2.1 Artificial Neural Networks

The neural network model is configured such that the application of a set of inputs produces the desired set of outputs. The neural network is trained by feeding the patterns and allowing it changes its weights according to some learning rule. The decision making process of NN is inclusive of two phases: learning and reasoning. Before a decision is reached, NN is trained with previous evaluation examples. A back-propagation algorithm is employed. During learning, a set is presented to the network at a time and is propagated forward to determine the calculated output. Error between calculated and target output is determined. The mean square error for all the trained records is then back-propagated through the net in order to adjust the weights with the purpose to reduce the next mean square error. Learning is done through iteratively adjusting the weights until the mean squared error reaches the minimum acceptable error set previously. After learning, the knowledge of making decision is kept implicitly in the weight connections. By then, trained set is ready to resolve a new case.

Zhang, Xinghua, Yuechuang, and Congdong (2010) applied backpropagation neural network to research the models of customer relationship evaluation criteria system based on customer loyalty and satisfaction, which has great advantages and feasibility. Chris, Wang and David (2002) worked on a particular dichotomy existing between neural networks and chi-square automated interaction detection (CHAID) to accomplish the goals of today's customer relationship management philosophy. We employed backpropagation in our study, which resulted to the set presented to the network, propagated forward to determine the calculated output.

Neural networks have broad applicability to real world business problems and are successful when applied in banks. Since neural networks are best at identifying patterns or trends in data, they are well suited for prediction or forecasting needs including customer relationship management, risk management, sales forecasting, data validation, customer research, target marketing, etc. Neural networks use a set of processing elements or nodes analogous to neurons in the brain. These processing elements are interconnected in a network that can then identify patterns in data once it is exposed to the data, that is, the network learns from experience just as people do. This distinguishes neural networks from traditional computing programs that simply follow instructions in a fixed sequential order.

2.2 Reasons for ANN use in CRM

YongSeog (2006) studied the effects of variable selection and class distribution on the performance of specific regression (i.e., a primitive classier system) and artificial neural network (ANN; a relatively more sophisticated classifier system) implementations in a customer relationship management (CRM) setting. The ensemble models are constructed by combining the predictions of multiple classifiers and the paper showed that ANN ensembles with variable selection show the most stable performance over various class distributions. Kwok, Choy, Lau and Kwok (2007) proposed a system aimed at establishing a cost-effective strategic CRM solution for achieving total customer satisfaction, integrating the data warehouse concept with two emerging technologies, Online Analytical Processing (OLAP) and Artificial Neural Networks (ANNs), to support in the customer relation strategy. The system was applied in Ka Shui Manufactory Company Limited to support their customer relationship planning and the result shows that significant improvements were made in customer service efficiency and cost reduction.

Artificial Neural Networks (ANN) is capable of predicting new observations from previous observations results after executing a process called learning from existing data. It is a computer model based on the architecture of the brain. It first detects the pattern from data sets; then it predicts the best classifiers; and finally, it learns from the mistakes. It works best in classification and prediction as well as clustering methods.

A sample solution is designed to prevent, detect, analyze and follow up customers banking behaviour in terms of transaction analysis to observe if the customer is a normal and loyal one [profitable customer that will stay] or not. With this, Intercontinental Bank Plc, where we applied our model, can monitor the activities of customer's behavioural transactions by using a robust and powerful technology based on rules, parameters and indicators. This design allows Intercontinental Bank Plc to determine a customer that will leave and profitable one that will stay.

By adding the intelligence of neural network technology to an already successful classified rule-based system, the bank can increase its capital base from normal, loyal customers. The neural network is adaptive, able to learn from patterns of normal customer behaviour and adapting to the evolving of behaviour of normal transactions. The recall process of the neural networks is extremely fast and can make decisions in real time.

Some advantages of neural network's ability in CRM system are: significantly reduces losses due to the prediction of customers leaving the bank; identify new methods or customer relationship methods to increase bank strategic advantage over their rivals; can work in real time, online or batch modes and will reinforce customer trust; improve operational efficiencies; and also give the bank the flexibility to easily incorporate data from many sources to the neural models.

3. Proposed Methodology

To meet customers' demands and improve their satisfaction, enterprises should have abilities of customer segmentation. Different customers have different characteristics such as preference, values, and profit. Data used was collected from the data accumulated in the bank database, which involves customer's demographic data, transaction data, account balance, etc. The data was categorized into static and dynamic data according to their features. Static data refers to the data that does not often change, that is, personal details that shows the fundamental states of the bank customers such as customer's name, address, age, sex, and income. Dynamic data are the customer data that changes from time to time such as the customer transaction data.

In this research, bank customers are grouped by their purchase history, which is the transaction data, while one attribute is used to substitute clustering result, which was incorporated in the classification process. Feature selection was used to obtain valuable attributes from static attributes. These selected static attributes was combined with the clustering result and taken as the input vector of the neural network.

To cluster the bank customers by their purchase behaviour history, three dynamic attributes selected from the transaction data as the clustering features are the trend attributes, the purchase frequency, and the customer consumption amount. In trend attribute, attributes used in the transaction data are those that describe all aspects of the customer's transaction amount trend of the reflecting long-term trend. The least square method (formula 1) was used to get the trend

$$T = \frac{\sum_{t=1}^{n} t y(t)}{\sum_{t=1}^{n} t^2} \quad \ldots\ldots \quad (1)$$

where t is the time coordinate, n is the number of trend trial time and the $y(t)$ is the customer's monetary expense at time t.

In purchase frequency, it displays customer's purchase habit. Bank customers have different transaction frequency and monetary contribution to the bank. In customer consumption amount, the mean value of customer's purchasing amount is used to reflect the customer's contributions to the overall turnover of the bank. The formula (2) used is

$$\mu = \frac{\sum_{t=1}^{n} y(t)}{n} \quad \ldots\ldots \quad (2)$$

Using the standard variance formula, the variation of customer purchase sequence was obtained. Its computation method is shown in formula (3)

$$\sigma = \sqrt{\frac{\sum_{t=1}^{n} (y(t) - \mu)^2}{n}} \quad \ldots\ldots \quad (3)$$

In this study, all layers use sigmoid function as the transfer function. The scaling output is in the range of x to 1-x linearly making the output range to be between 0 and 1. During the training of the neural network, 0.2 was used in case of possible prediction to output the value greater than 0.8 so that the output will not be smaller than 0.2.

Each bank customer has a transaction sequence describing the customer's behaviours. There are two attributes at any given time, which are the customer consumption and monetary amount of the transaction category. During the analysis, 50 transaction codes out of 133 transaction codes were selected as valuable transaction category to identify different customer's behaviours.

3.1 Classification approach

Classification is the process of finding classifiers/models or functions that map records into one of some discrete class set (Ogwueleka, 2009a). The proposed model construction has two types of data sets, the training and testing set. The training data sets, with prescribed class labels of customer behaviour are fed into the model to enable the model to find parameters or characters that distinguish one class (normal/abnormal) from the other. This step is called the learning process. The testing data sets, without pre-classified labels, are fed into the model. The model automatically assigns the precise behaviour class labels for those testing items. If the results of testing are unsatisfactory, then more training iterations are required. But if the results are satisfactory, the model can be used to predict the classes of target items whose class behaviour labels are unknown.

3.2 ANN Model design

Neural network is an interrelated set of nodes, comprising of neurons in input layer, one or more hidden layers and output layer as shown in Figure 1.

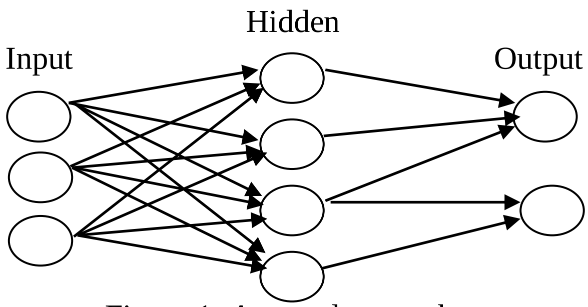

Figure 1: A neural network

The middle layer is at the midst of the input layer and is seen as an internal representation of the input pattern. It is at this layer that patterns are identified from the distilling the characteristics of the input and are then sent to the output layer. The result of the output layer is a judgment on the classification of the input patterns. So the middle layer can be called the distilling character layer. In this study, the input and output variables are first classified. The input variables, which are customer satisfaction factors were entered into the network: geographic position, the variety of goods price, the overall price level, the overall quality of goods, the time of waiting for payment, service attitude and shopping environment. The output variable is only one possible measurement value.

The system has four (4) input layer made up of customer's personal/demographic data [customer's name, address, age, sex, telephone number and email address]; transaction data [withdrawal and deposit]; account details [amount of money per transaction made and account balance]; and purchase history. The hidden layer is two (2) and the output layer is also two [normal loyal customer and abnormal customer]. The weight used was 0.2 and the number of iterations used to produce the best model was 32.

While setting up the ANN model, we follow a six-step procedure. First, the data was defined and presented to the ANN as a pattern of input data with the desired target. Secondly, the data are categorized into training set or validation set as the ANN model used the training set in its learning process to develop the model while the validation set was used to test the model for its predictive ability and to determine when to stop the training of the ANN. Thirdly, the ANN structure was defined by selecting the number of hidden layers to be constructed and the number of neurons for each hidden layer. Fourthly, all ANN parameters are set before the training process started. Next, the training process is started. The training process involves the computation of the output from the input data and the weights. The back propagation algorithm was used to train the ANN by adjusting its weights to minimize the difference between the current ANN output and the desired output. During the training of the neural network, 0.2 was used in case of possible prediction to output the value greater than 0.8 so that the output will not be smaller than 0.2.

Finally, an evaluation process was conducted to determine if the ANN has 'learned' to perform. In the evaluation process, there was periodic pausing of the training process and testing of its performance until an acceptable result was obtained. When the acceptable result was obtained, the ANN model was deemed to be fully trained and ready for use. The model approach employed for clear understanding and easy implementation is depicted in Figure 2.

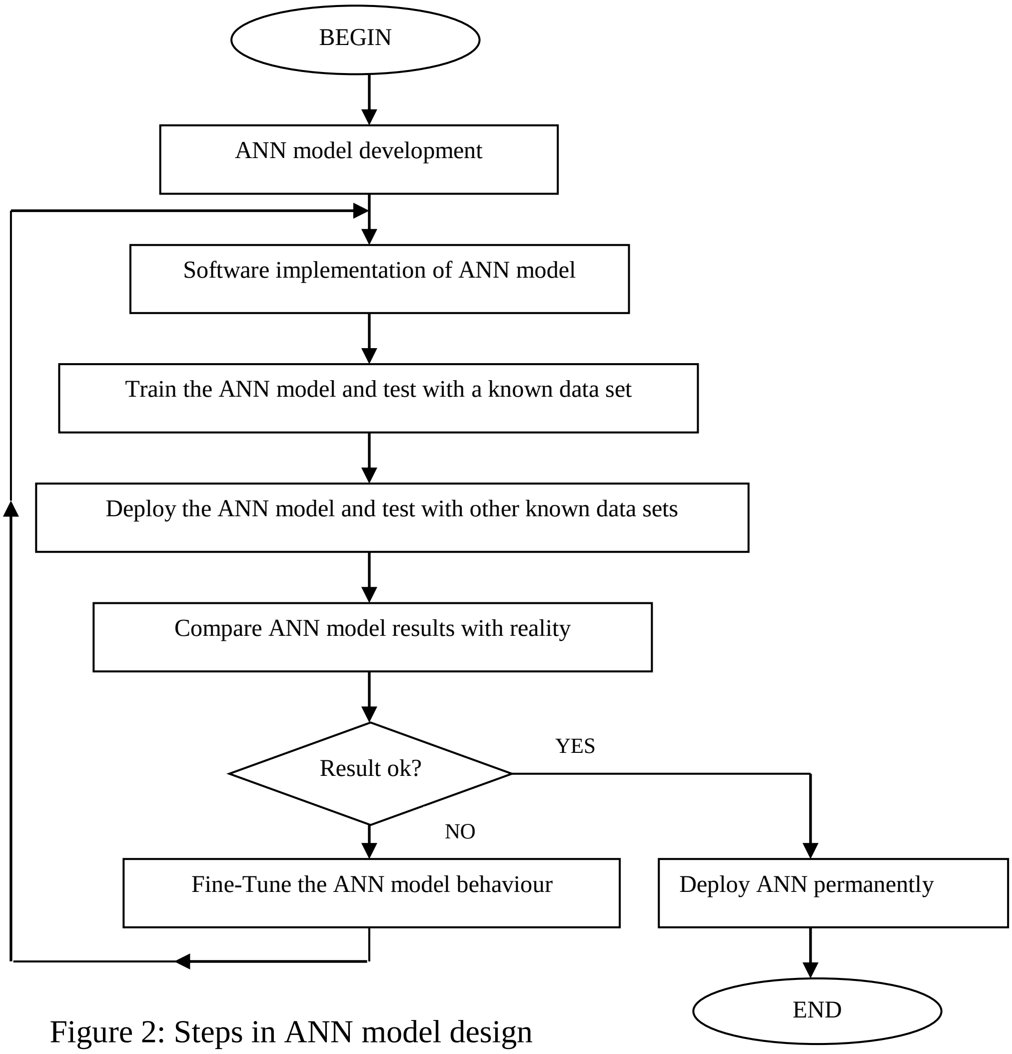

Figure 2: Steps in ANN model design

As there are no fixed rules in determining the ANN structure or its parameter values, we have constructed various ANNs with different structures and parameters before the acceptable model was determined. By performing periodic testing of the ANN on the test set and recording the results of both the training and test data sets, the number of iterations that produces the best model was obtained. All this is done to reset the ANN and train the network up to that number of iterations.

We have applied our ANNN model in a renowned international bank, Intercontinental Bank Plc.

3.3 Organisational Background

Intercontinental Bank transaction banking activities can be characterized as retail banking, dealing directly with individuals and small businesses, and investment banking, relating to activities on the financial markets. It is charged with controlling interest rates and money supply across the whole economy. They act as lender of last resort in event of a crisis.

3.4 Identifying Customer Behaviour

The classification of rule learning based on the given history of a customer banking transactions and customer mode are analyzed and labelled as abnormal (behaviour of customer that will leave] and normal behaviours [profitable customers that will stay).

The set of rules for the account was searched. After rules are selected, a set of profiling monitors was built for the purpose of investigating the sensitivities of accounts to general rules. The construction of profiling monitors consists of two stages, profiling and usage stage. In the profiling stage, a general rule was applied to a portion of a customer's normal transaction usage to evaluate the behaviour's normal activities. In the usage stage, the monitor was applied to the whole part of the account and the result was used to examine the abnormality of the bank transaction monthly use. To improve the confidence of the customer class behaviour, monitors are combined with evidence resulted from the application of monitors to the sample data. There is constant inspection of the network system for the customer behaviour analysis by the bank CRM team.

The network inspection system for the bank customer behaviour (Ogwueleka, 2009b) is shown in Figure 3. Raw data on customer demographic data and transaction data, that is, the behavioural data are extracted and profiled. The data is classified and the rule applied. The behavioural profiled analysis is stored in the bank database.

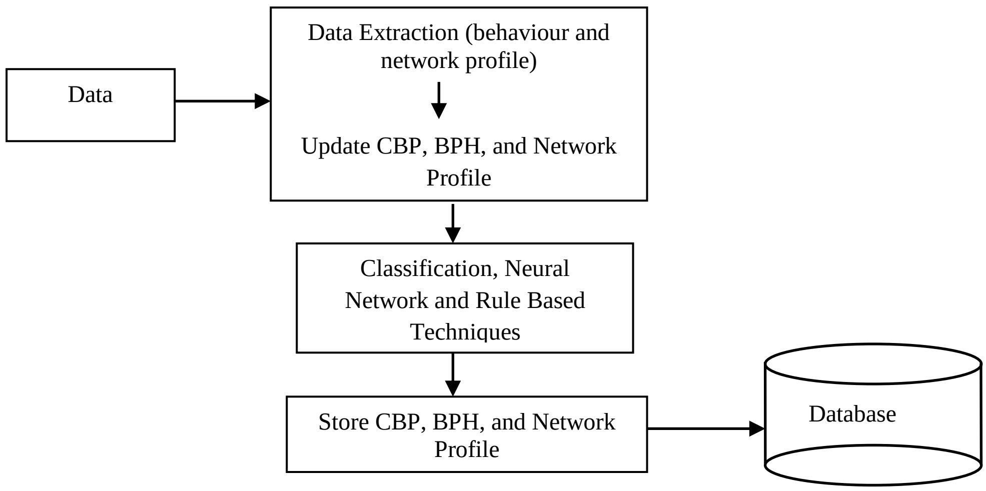

Figure 3: The network inspection system

The features of a customer behaviour profile depend on the nature of the bank service application. The control technique is to maintain histories of customer usage information, relating to the entity, over different time frame. The short term past behaviour is referred to as the Current Behaviour Profile (CBP) and the long term past behaviour as the Behaviour Profile History (BPH).

Performing the differential analysis, the model determines if there is a significant change in the customer behaviour. Performing the absolute analysis, the CBP is observed. If it exceeds predetermined thresholds for bank acceptable network use over the lifetime of the CBP, alerts are raised.

The designed customer behaviour classification task was considered as pattern recognition or classification problem. The set T_n of all customer transaction behaviour in banking operations are divided into two disjoint subsets, namely normal loyal customer behaviour transactions $T^l_n \subseteq T_n$ and abnormal disloyal ones $T^d_n \subseteq T_n$, $T^1_n \cap T^d_n = \emptyset$. If the numerical descriptions, that is, the points in some multidimensional space of abnormal and normal transaction behaviours belong to different areas in the space, then it is possible to make a decision about the image of a new transaction t^{n+1}.

The following two hypotheses are considered as a basis for the classification.

1. Hypothesis H_l : Transaction $t^{n+1} = (t^{n+1}{}_1, \ldots , t^{n+1}{}_m)$ on customer transaction r_k is similar to all previous transactions from the set T_{rk} , which were carried out earlier by the customer. If hypothesis H_l is confirmed for customer transaction behaviour t^{n+1} , then the customer transaction t^{n+1} is classified as normal and included into the set T^1_n .
2. Hypothesis H_d : Transaction $t^{n+1} = (t^{n+1}{}_1, \ldots , t^{n+1}{}_m)$ is similar to earlier executed abnormal transaction behaviour $T^d_n = \{t^i$ – considered abnormal $| \ t^i \in T_n\}$. If hypothesis H_d is confirmed for customer transaction t^{n+1} , then customer transaction behaviour t^{n+1} is classified as abnormal and included into the set T^d_n .

Neural network techniques for clustering and classification was used to check the proposed hypotheses H_l and H_d . This was to create pattern of "normal customer behaviour" and pattern of "abnormal customer behaviour" on the basis of neural network "learning" from the previous customer behaviour transactions T_n executed and to develop "rules" of normal customer behaviour and abnormal customer behaviour. The learning algorithms allow the system to follow the customer behaviour and self-adapt to changes in it. If a transaction does not correspond to the pattern of "normal customer behaviour" or is similar to the "abnormal" pattern it is classified as suspicious for disloyalty.

3.5 The Proposed Algorithm: The Transaction Monitoring Algorithm

The proposed method for transaction analysis, that is, the transaction monitoring algorithm is represented in Figure 4(a), (b), and (c). The process of transaction monitoring consists of three stages: data accumulation, training (building of customer profile) and control of transactions (Ogwueleka, 2009c).

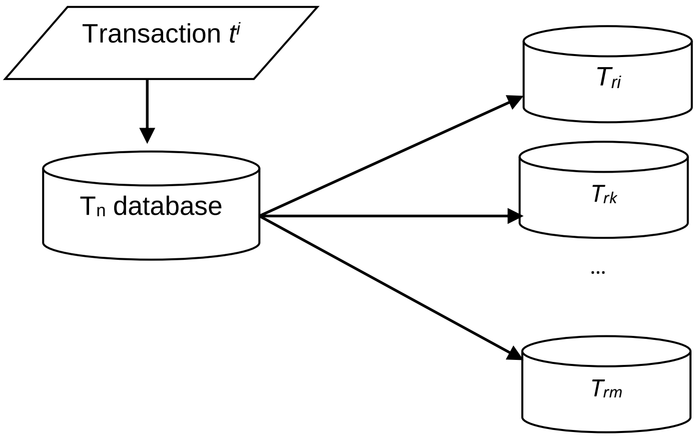

Figure 4(a): Data accumulation stage

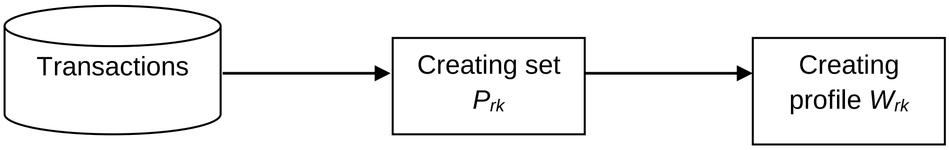

Figure 4(b): Training stage

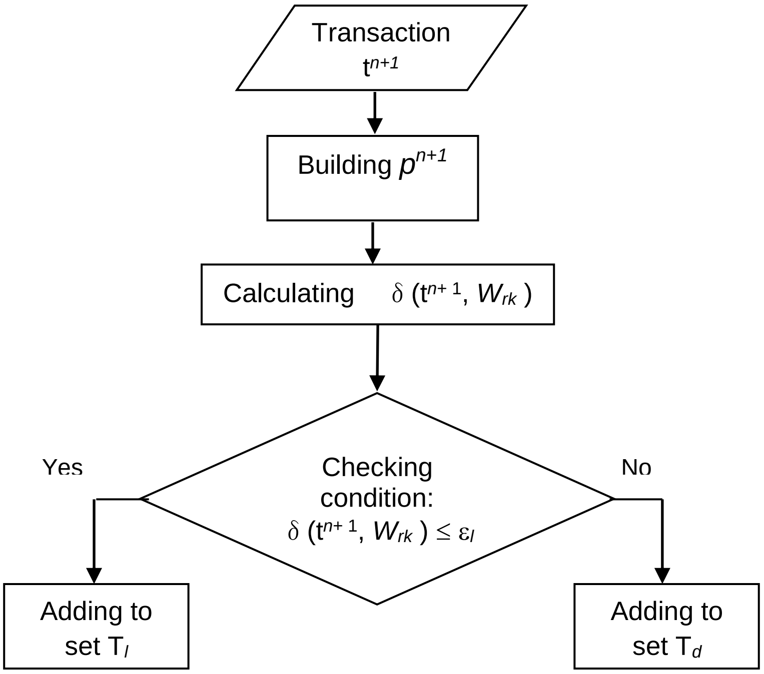

Figure 4(c): Transaction control stage

At the data accumulation stage, the data about the transactions on customer behaviour r_k are collected in the database. If the size of T_{rk} exceeds some predefined level, sufficient to build an adequate profile, then the monitoring process goes to stage two.

At the training stage, which is the stage two of transaction monitoring algorithm (Ogwueleka, 2009c), the customer profile W_{rk} is created building set P_{rk} using the function φ; training the neural network on the basis of set P_{rk}; and building the profile W_{rk} as an outcome of training. (W_{rk} as the total/result of all customer transaction behaviour in banking operations, which was divided into two disjoint subsets of normal loyal customer behaviour transactions and abnormal disloyal ones, obtained as an outcome of the training in building the customer profile and training of the neural network).

After the training stage, comes the stage of customer transaction control process, which involves the building of the vector p^{n+1} and applying the function φ to every new behavioural transaction t^{n+1}; the calculation of the deviation the current behavioural transaction t^{n+1} from the profile W_{rk}; the comparison of the value δ_0 with the threshold ε_l fixed for the profile W_{rk} (where ε_l is a boundary value for the degree of similarity of the transactions on particular customer transaction behaviour c_k to the profile W_{rk}; and finally the consideration of transaction that is normal or abnormal for further analysis. Hence if $\delta_0 \leq \varepsilon_l$ then transaction t^{n+1} is considered normal and the vector t^{n+1} is added to the set $T_l = T_{rk}$;but if $\delta_0 \leq \varepsilon_l$ then transaction t^{n+1} is considered suspicious for abnormal customer and is added to the set T_d for further expert analysis.

3.6 Implementation of the model

The designed system is a classifier that attempts to classify a customer behavioural transactions being normal or abnormal. This is done by generating a score and comparing it with a threshold. The two-stage model created and used in Intercontinental bank has two stage structure of high and low risk in the application of the bank's customer behaviour transaction rules after classification and clustering of the recorded customer transaction data in the database, to distinguish the normal from the abnormal customer transactions. The two-stage model is shown in Figure 5. Figure 5, is the description of how the designed model works using classification method while generating a score along a threshold. Any observed abnormal behaviour is scrutinized by the bank monitoring team and different strategies implemented in order to convert the customer to a normal and loyal one, not one striving to move to another bank.

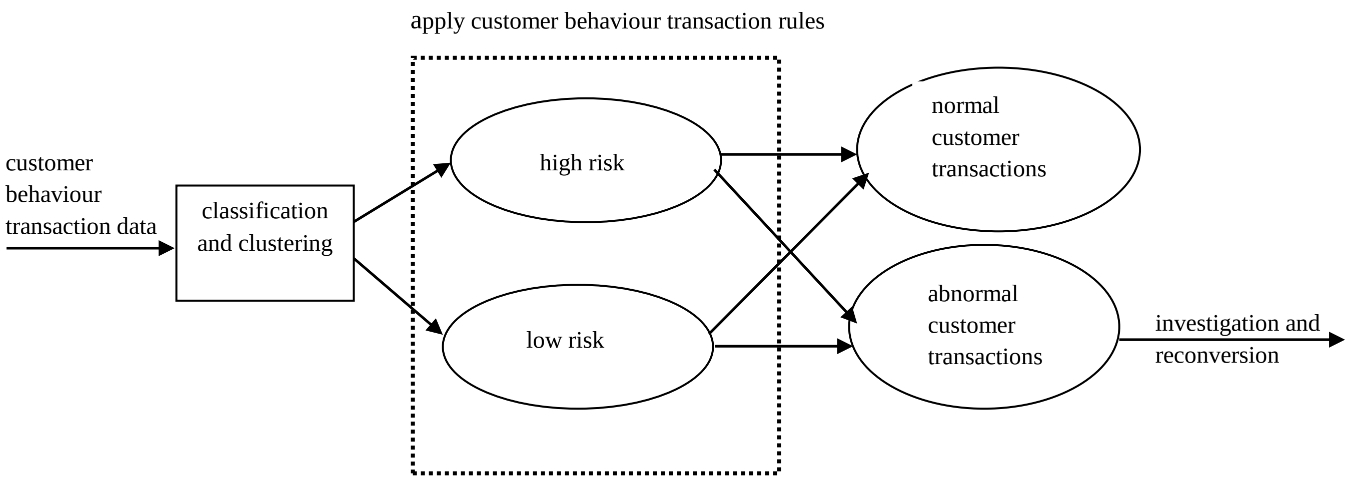

Figure 5: Two-stage model

In this research, the designed system is implemented in credit card section, consisting of two units namely, the withdrawal and deposit unit. Each of the two units is in turn made up of the following subunits: the database interface; the neural network classification, the transaction/business rules; and the visualization. The database interface subunit is tested to ensure that the necessary transaction data is imported and used. In the neural network classification, the available data set is randomly partitioned into a training set and a test set. The training set is further partitioned into subsets used for elimination of the model (i.e. training the algorithm) and subset used for evaluation of the performance of the model (i.e. validation). Transaction/business rules are used to store the knowledge or rules of customer behaviour transactions in an "if-then" format. Transaction parameters are tested with the known rules and thresholds in order to prioritize each customer behavioural transaction into normal and abnormal, as it is very important to test for false and misclassification. In visualization, the graphical user interface (GUI) visualization subunit is also tested to facilitate event driven feature of the designed system, which makes it user friendly. When any of the subunits fails a test, the subunit is redesigned or the program statements rewritten, followed by a retesting, until all the subunits pass the test.

Test data is designed and run on the system with the tested program. The result of the process is compared with a manually prepared result to determine the efficiency and effectiveness of the designed system. The test data for two credit card banking transactions under study, which is withdrawal and deposit, is presented. Test data is designed and run on the system with the tested program. The result of the process is compared with a manually prepared result to determine the efficiency and effectiveness of the designed system. The test data for two credit card banking transactions under study, which is withdrawal and deposit, is presented.

3.7 Actual test result and expected test result for withdrawal customer behaviour transaction

This is where the algorithms are experimented module by module to give an expected result. For actual and expected results; 0 in the corresponding cell refers to normal customer transaction, 1 refers to suspected customer transaction, while 2 refers to an abnormal customer transaction. This notation is used in Table 1 and 2.

Table 1 shows test data for withdrawal transaction used for the program. The expected and the actual results are provided.

Table .1: Withdrawal transaction test data

TransactNo	AccountNo	CreditCardNo	Amount	Last Withdrawal Date	Actual Result	Expected Result
23	SAV23	1000000023	50000	6/8/2008	0	0
24	SAV23	1000000023	100000	6/8/2008	1	1
25	SAV23	1000000023	1500000	7/7/2008	2	1
26	SAV23	1000000023	10000	7/7/2008	0	0
27	SAV24	1000000024	46000	6/8/2008	0	0
28	SAV24	1000000024	20000	7/7/2008	0	0
29	SAV24	1000000024	100000	7/7/2008	0	1
30	SAV24	1000000024	46000	6/8/2008	0	0
31	SAV24	1000000024	2000000	6/8/2008	2	2
32	SAV26	1000000026	100000	7/7/2008	0	0

Test data for deposit customer behaviour transaction

Table 2 shows test data for deposit transaction used for the program. The expected and the actual results are provided.

Table 2: Deposit transaction test data

TransactNo	AccountNo	CreditCardNo	Amount	Last Deposit Date	Actual Result	Expected Result
23	SAV23	1000000023	40000	6/8/2008	0	0
24	SAV23	1000000023	60000	6/8/2008	0	0
25	SAV24	1000000024	100000	6/8/2008	0	0
26	SAV24	1000000024	234000	6/8/2008	0	0
27	SAV24	1000000024	46000	6/8/2008	0	0
28	SAV24	1000000024	20000	7/7/2008	0	0
29	SAV24	1000000024	1500000	7/7/2008	1	1
30	SAV24	1000000024	46000	7/7/2008	0	0
31	SAV24	1000000024	2350000	7/7/2008	2	2
32	SAV26	1000000026	100000	7/7/2008	0	0

4 Performance Evaluation

There are two sources that generate inputs to the detection software: normal (H_0) and abnormal (H_1). The normal source generates loyal customer transactions. The abnormal source generates disloyal customer transactions. In a typical real-life customer behaviour transaction, a large percentage of customer transactions are normal. The skewed nature of the frequency distribution makes detection of abnormal customer transactions difficult. The software observes the customer behaviour transaction but does not know whether it came from a normal or abnormal source. The goal of customer behaviour transaction classification software is to classify each transaction as normal or abnormal. The types of errors can occur in this classification:

(i) Classification of a abnormal transaction as normal (false negative); and

(ii) Classification of a normal transaction as abnormal (false positive)

Probability of classification = P_D = P_r (classify into $H_1 \mid H_1$ is true) or

Probability of false negative = 1 – P_D

Probability of false positive = P_F = P_r (classify into $H_1 \mid H_0$ is true)

Let the numerical values for the normal and abnormal customer behaviour transactions follow exponential distributions with parameters λ_N and λ_F, $\lambda_N > \lambda_F$ respectively.

The probability of classification P_D and probability of false positive P_F as

$$P_D \doteq \int_t^{\infty} \lambda_F e^{-(\lambda_F x)}\, dx = e^{-\lambda_F t}$$

$$P_F = \int_t^{\infty} \lambda_N e^{-(\lambda_N x)}\, dx = e^{-\lambda_N t}.$$

Thus P_D can be expressed as a function of P_F as

$$P_D = P_F^r,$$

where $r = \lambda_F/\lambda_N$ is between 0 and 1.

Trees (2001) stated that the quality profile of most detection software is characterized by a curve that relates its P_D and P_F known as the receiver operating characteristic curve (ROC). ROC can be seen as a function that summarizes the possible performances of a classifier as it facilitates the choice of a decision functions.

The effectiveness of this classification algorithm is measured in terms of the classification errors, which includes system classification rate and false alarm rate. The data used in the application were collected from an Intercontinental bank, which consist of customer transaction data made per day during the observed period, that is, six months. The collection was done according to the two types of bank operations investigated. The aim is to identify abnormality within each of the categories by identifying rechecked suspected customer behaviour transactions using the neural network algorithm. Details of the test datasets are listed in Table 3.

Table 3: Summary of the two data subsets used to test each of the operations of neural network model

Operation	Transaction	Abnormal	Proportion of abnormality
Withdrawal	10,650	5	0.47%
Deposit	8,102	2	0.24%
Total	18,752	7	0.37%

The performance analyses of the respective classification and neural network algorithms are carried out using MATLAB software package and the results compared with the collected data are as shown in Figure 6 and 7.

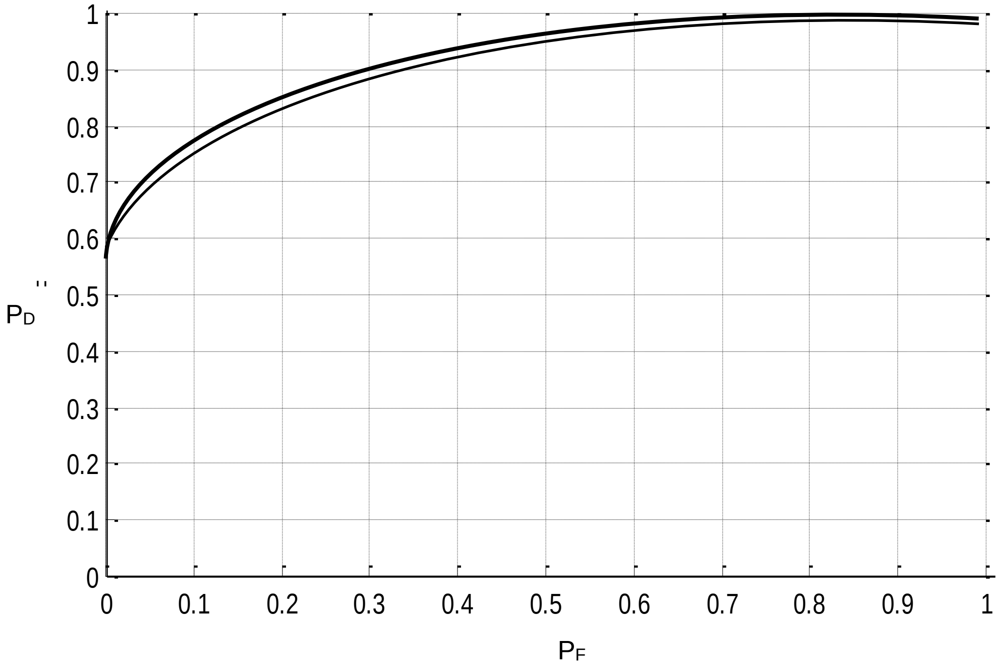

Figure 6: ROC for withdrawal abnormal classification

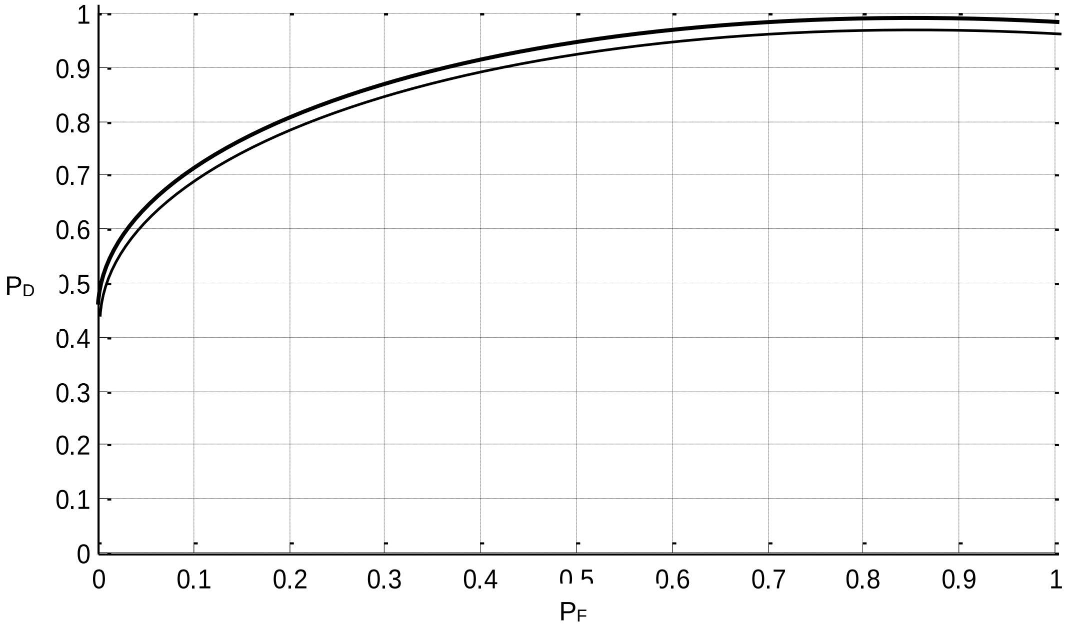

Figure 7: ROC for deposit abnormal classification

Collected data

Model data

From figures 6 and 7, it can be seen that the model results compared satisfactorily well with the collected data results for the two transactions examined in the research.

5 Comparison of customer behavioural transaction performance with other classification detection models

The ROC curve is used for this comparison. Two different commercial products such as quadratic discriminant analysis (QDA) and logistic regression (LOGIT) were selected to test the feasibility of using neural network tools for the purpose customer detection behaviour classification. The performance analysis of the model (deposit transaction) is compared with these two commercial packages and the result is shown in Figure 8.

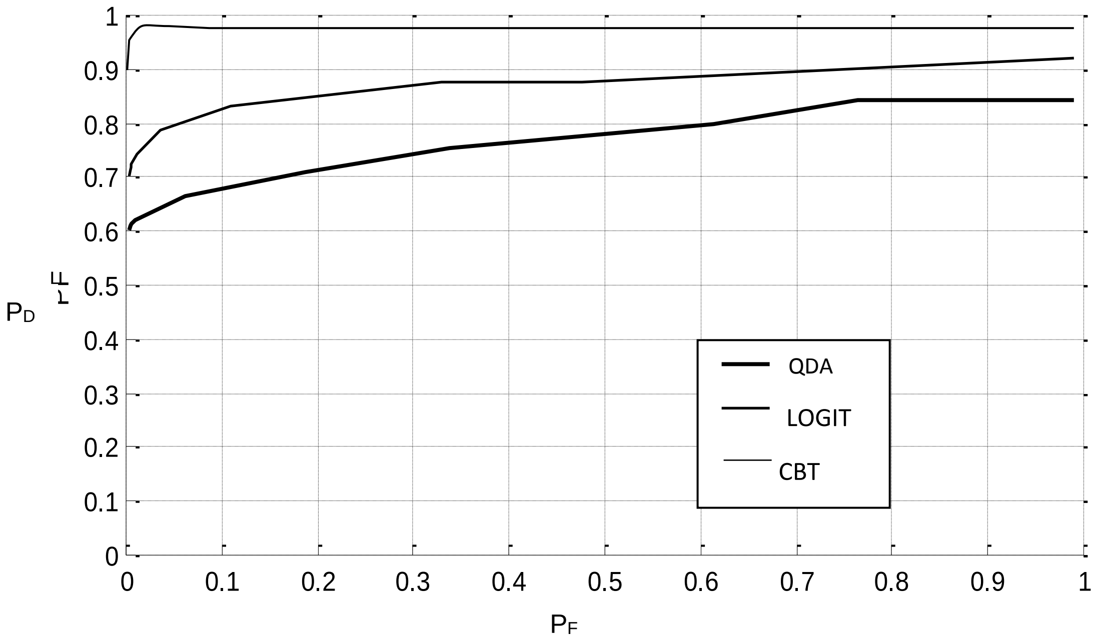

Figure 8: Comparison of customer behaviour transaction (CBT) with other system

From Figure 8, the ROC curve for CBT detects over 94% of classified abnormal cases without causing false alarms. This is followed by the ROC curve for logistic regression with 72% classified customer abnormality with no false alarms. The quadratic discriminant analysis is 61%. This shows that the performance of CBT is in agreement with other algorithms, but performs better. This relative advantage of CBT over others is as a result of combining neural networks with classification transaction rules.

Using two software products enabled this work to illustrate different user interfaces available, alternative to neural networks models, design for decision-making, and the performance metrics of different models. Comparison of performance of neural network model with traditional statistical models increased the confidence in the ability of CBT classification and neural network model in successful modelling of credit card customer abnormality in Intercontinental Bank.

6. Conclusion

Neural networks using unsupervised learning method is applied to the data to generate models as there are no prior sets of normal and abnormal observations. Techniques employed include a combination of profiling, classification, clustering and transaction rules. The baseline model generated that represents normal behaviour and then attempt to detect observations/transactions that show greater departure from this normal transaction helped tremendously in the credit card customer behavioural transaction. Suspected customer transaction behaviours are rechecked and strategies of changing their class taken up immediately, without the knowledge of the customers for the further investigations and subsequent decision-making.

The study created a model, which was used to detect changes in established bank customer behavioural patterns for effective management and maintenance of customer relationship, bank productivity and bank growth for competitive advantage. The system was applied in Intercontinental Bank Plc to support their customer relationship management and planning. The result shows that significant improvements were made in customer service efficiency, customer retention, customer satisfaction and cost reduction; customer balance; customer revival; discovery of new customers; banking operations efficiency; and speedy sale of bank deals.

Cardwatch developed by Aleskerov, Freisleben and Rao (1997) featured neural networks trained with the past data of a particular customer to make the network process the current spending patterns to detect possible anomalies and a neural multi-layer perceptron based classifier designed by Dorronsoro ,Ginel, Sanchez and Cruz (1997), acts only on the information of the operation itself with its immediate previous history, but not on historic databases of past cardholder activities. From our research, we observed that a successful solution to the CRM issues for credit card bank customers has not been applied, which motivated this study and was able to predict the behavior of credit card customers in banking sectors from the performance evaluation.

This study was limited to CRM in physical banking transaction of credit card customers for withdrawal and deposit transactions in an international bank and was not applied to Internet banking or other banking transactions ports. The data collected was basically from the bank's database of physical real time transaction of credit card users

In recommendation therefore, for a total solution in banks CRM implementation, setting up model programs and short-term landmarks should break the bank scheme and design on CRM into controllable portions. All business units or departments in the bank should be integrated in small adjustable method. Thus, for effective bank CRM solution, banks should consider factors such as safety and health support; systems compatibility; contract structure; distribution flexibility; and troubleshooting and problem-solving, etc. If these are put into consideration, an effective and efficient CRM solution will be obtained. Other studies can be done implementing the Internet banking and ATM transactions of the banking industry.

References

Anderson, E., Fornell, C., & Lehmann, D. (1994). Customer Satisfaction, Market Share, and Profitability: Findings from Sweden. Journal of Marketing, 58(3), 53-66.

Alis, O., Karakurt, E., & Melli, P.(2000). Data mining for database marketing at Garanti Bank. In Proceedings of the International conference on data mining, 93-107.

Aleskerov, E., Freisleben, B., & Rao, B. (1997). CARDWATCH: A Neural Network-Based Database Mining System for Credit Card Fraud Detection. Proceedings of the IEEE/IAFE on Computational Intelligence for Financial Engineering, 220-226.

Chris R., Wang J.C., & David C.Y. (2002) Data mining techniques for customer relationship management. Technology in Society, 24(4), 483–502.

Chin J. (2000). It's important to do it well. Straits Times--Computer Times, 14-16.

Delvin, J. (1995). Technology and innovation in retail banking distribution. International Journal of Bank Marketing, 13(4), 19-25.

Dorronsoro, J., Ginel, F., Sanchez, C., & Cruz, C. (1997). Neural Fraud Detection in Credit Card Operations. IEEE Transactions on Neural Networks 8(4), 827-834.

Eduardo R., & Edwards J. S., (2009). Knowledge Management and Enterprise Risk Management Implementation in Financial Services. In proceedings of the 2009 ERM Symposium Society of Acturaries.

Feng Q., & Xu, L. (2007) Improving customer satisfaction by the Expert system using ANN. WSEAS Transactions on Business and Economics, 10(4),147-151.

García-Crespo, A., Colomo-Palacios, R., Gómez-Berbís, J.M., & Ruiz-Mezcua, B. (2010). SEMO: a framework for customer social networks analysis based on semantics. Journal of Information Technology, 25 (2), 178-188.

Garvin, D. (1988). Managing quality: The strategic and competitive edge. The Free Press: New York.

Haykin, S. (1999) Neural Networks: A Comprehensive Foundation, Prentice Hall: New York.

Hassibi K. (2000). Detecting payment card fraud with neural networks. In P.J.G. Lisboa, A.Vellido, B.Edisbury Eds. Business application of Neural Networks. Singopore: World Scientific

Kwok K.C.M., Choy K.L., Lau H.C.W., & Kwok S.K. (2007) A strategic customer relationship management system: a hybrid OLAP-neural approach International Journal of Enterprise Network Management, 1(4), 350-371.

Krishnan, M.S., Ramaswamy, V., Meyer, C.M, & Damien, P. (1999). Customer Satisfaction for Financial Services: the role of products, services, and Information Technology. Management Science, 45(9), 1194-1209.

Kennington, C., Hill, J., & Rakowska, A. (1996). Consumer selection criteria for banks in Poland. International Journal of Bank Marketing, 14(4), 12-21

Langsdorf, T., Brause, R., & Hepp, M. (1999). Credit card fraud detection by adaptive neural data mining. In Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence, 103–106.

Lariviere, B., & Van den Poel, D. (2004). Investigating the role of product features in preventing customer churn, by using survival analysis and choice modelling: the case of financial services. Expert Systems with Applications, 27(2), 277-285.

Ogwueleka, F. N. (2009(a)) Data mining techniques for customer relationship marketing in the banking industry. Advances in Natural and Applied Sciences, 3 (2), 73-78.

Ogwueleka, F. N. (2009(b)). Fraud detection in mobile communications networks using user profiling and classification techniques. Journal of Science and Technology, 29(3), 31- 42.

Onut, S., Erdem, I., & Hosver, B.(2008). Customer Relationship Management in Banking Sector and A Model Design for Banking Performance Enhancement. In A.A. Minai & Y. Bar-Yam (Eds.) Unifying Themes in Complex Systems, Springer.

Roland L., Geier J., & Kolliker M., (2004). Artificial neural networks, classification trees and regression: Which method for which customer base?. Journal of Database Marketing & Customer Strategy Management. 11(4), 344-356.

Sharma, D.D. (1994), Classifying Buyers to Gain Marketing Insight: A Relationships Approach to Professional Services. International Business Review, 3 (1), 15- 30.

Syeda, M, Zhang Y. Q., & Pan Y. (2002). Parallel granular neural networks for fast credit card fraud detection. In Proceedings of the 2002 IEEE International Conference, 1,572–577.

Smith, S.,. (2003). SAS Institute white paper. SAS® Banking Intelligence Solutions for Customer Relationship Management.

Trees, H. V. (2001). Detection, Estimation and Modulation Theory-Part I. John Wiley, New York.

Yong S. K. (2006). Towards a successful CRM: variable selection, sampling, and ensemble. Journal Decision Support Systems. 41(2), 542-553.

Xiang, B., & Jun, Y. (2006) Customer Segmentation based on Neural Network with Clustering Technique, Proceedings of the 5th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Madrid, Spain, February 15-17, 265-268.

Zineldin, M. (1996). Bank strategic positioning and some determinants of bank selection, International Journal of Bank Marketing, 14(6), 12-22.

Zhang, Z., Xinghua, L., Yuechuang, C., & Congdong, L. (2010). Customer Relationship Evaluation System Based on BP Neural Network in Electronic Commerce. In Proceedings of the International Conference on E-Business and E-Government (ICEE), 3098 - 3101.