*i*Journals: International Journal of Software & Hardware Research in Engineering (IJSHRE)
ISSN-2347-4890
Volume 10 Issue 2 February 2022

The Application of Hidden Markov Model in Credit Card Fraud Detection System

[1]Dr. Anusiuba Overcomer Ifeanyi Alex; [2]Okechukwu Ogochukwu Patience; [3]Ekwealor Oluchukwu Uzoamaka; [4]Anusiuba Amarachukwu Angela

[1, 2,3] Department of Computer Science, Faculty of Physical Sciences, Nnamdi Azikiwe University Awka
[4] Anambra State Local Government Commission
[1] oi.anusiuba@unizik.edu.ng [2]op.okechukwu@unizik.edu.ng
[3]uzoamakaekwee@gmail.com [4] amarachukwuanusiuba@gmail.com

ABSTRACT

This work, credit card fraud detection system using Hidden Markov Model is based on card holder's spending habits and can help in eradicating frauds that are associated with credit card transaction; this work thoroughly investigates every credit card transactions to ensure that any falsified transactions are restricted while ensuring that genuine card users are not denied transactions.

The Hidden Markov Model was applied in determining the spending habit and or the profile of credit card holders; more so, with the spending profile established, it becomes possible to determine if an incoming transaction from a card holder is fraudulent or not by comparing any new transaction with the credit card holder's spending history while any variation from the actual spending habit is seen as a probable fraud and will be restricted and further verification is carried out. The methodology adopted for this research is Structured System Analysis and Design Methodology (SSADM). Data were collected and analyzed using PHP-MYSQL programming language for the design and test. The performance evaluation was designed to test the run-time performance of software within the context of an integrated system; this was cautiously carried out in all the testing process including unit and general testing. The performance of the software was justified since it met the aim and objective of the proposed system. Banks and other financial institutions that carry out their transactions with credit cards can adopt this system to detect and prevent all category of credit card fraud; the reliability and potential of this system to ward off credit card fraudsters is unquestionably high.

Keywords: ***Credit Card, Hidden Markov Model (Hmm), Online Shopping, E-Commerce, Credit Card Fraud, Fraud Detection System,***

1.0 Background of the Study

Credit card transactions have become the de facto standard for Internet and Web based e-commerce. The US government estimates that credit cards accounted for approximately US $13 billion in Internet sales during 1998. This figure is expected to grow rapidly each year. Germany and Great Britain have the largest number of on-line shoppers and credit card is the most popular mode of payment (59%). About 350 million transactions per year were reportedly carried out by Barclaycard, the largest credit card company in the UK, towards the end of the last century (Kim and Kim, 2002).

Aleskerov et al (1997), Defines credit card as a small plastic card issued to users as a system of payment. It allows its holder to buy goods and services based on the holder's promise to pay for those goods and services. Most credit cards are issued by banks or credit unions, and are the shape and size specified by the ISO/IEC 7810 ID-1 standard. Credit cards are issued by a credit card issuer, such as a bank or credit union, after an account has been approved by the credit provider, after which cardholders can use it to make purchases at merchants accepting that card.

When a purchase is made, the credit card user agrees to pay the card issuer, The cardholder indicates consent to pay by signing a receipt with a record of the card details and indicating the amount to be paid or by entering a personal identification number PIN Also, many merchants now accept verbal authorizations via telephone and electronic authorization using the internet, known as a card not present transaction (CNP).With rising interest in e-commerce, electronic payment have increased more in number and the most popular way is payment using credit card, probably because of its simplicity and comfort. A user simply enters the credit number, his name, the expiry date of the card; the merchant validates this information and upon approval from credit card Company, ships the goods or provide access to the service. The only thing that needs to be passed between the merchant and buyer is the credit card number.

According to Bhattacharyya, et.al (2011), for extra security, the communication between the user and merchant should be encrypted. The credit cards usage are in two ways; Physical usage this is where an individual does use the credit card to pay for his purchases in any store personally and Virtual/online usage**,** virtual or online usage is where the card owner uses the credit card to pay for purchased items online over the internet by just entering the required credit card details.
A secured credit card is a type of credit card secured by a deposit account owned by the cardholder. Typically, the cardholder must deposit between 100% and 200% of the total amount of credit desired. In some cases, credit card issuers will offer incentives even on their secured card portfolios. In these cases, the deposit required may be significantly less than the required credit limit, and can be as low as 10% of the desired credit limit. This deposit is held in a special savings account. Credit card issuers offer this because they have noticed that delinquencies were notably reduced when the customer perceives something to lose if the balance is not repaid. The cardholder of a secured credit card is still expected to make regular payments, as with a regular credit card, but should they default on a payment, the card issuer has the option of recovering the cost of the purchases paid to the merchants out of the deposit. The advantage of the secured card for an individual with negative or no credit history is that most companies report regularly to the major credit bureaus. This allows for building of positive credit history. Although the deposit is in the hands of the credit card issuer as security in the event of default by the consumer, the deposit will not be debited simply for missing one or two payments. Usually the deposit is only used as an offset when the account is closed, either at the request of the customer or due to severe delinquency (150 to 180 days). This means that an account which is less than 150 days delinquent will continue to accrue interest and fees, and could result in a balance which is much higher than the actual credit limit on the card. In these cases, the total debt may far exceed the original deposit and the cardholder not only forfeits their deposit but is left with an additional debt. Most of these conditions are usually described in a cardholder agreement which the cardholder signs when their account is opened. (Patil, et.al 2010)

Secured credit cards are an option to allow a person with a poor credit history or no credit history to have a credit card which might not otherwise be available. They are often offered as a means of rebuilding one's credit. Fees and service charges for secured credit cards often exceed those charged for ordinary non-secured credit cards, however, for people in certain situations, (for example, after charging off on other credit cards, or people with a long history of delinquency on various forms of debt), secured cards are almost always more expensive than unsecured credit cards.

Credit card fraud cases are increasing every year. In 2008, number of fraudulent activities through credit card had increased by 30 percent because of various ambiguities in issuing and managing credit Cards. Credit card fraudulent is approximately 1.2% of the total transaction amount, although it is not small amount as compared to total transaction amount which is in trillions of dollars in 2007. Card issuers must take more precaution against fraud detection and financial losses. When banks lose money because of credit card fraud, cardholders pay for all of that loss through higher interest rates, higher fees, and reduced benefits. In addition to financial losses, fraud may cause distress, loss of service, and loss of customer confidence (Bhattacharyya, et.al, 2011),

Hence, it is in both the banks' and the cardholders' interest to reduce illegitimate use of credit cards by early fraud detection. Online credit card fraud has become a huge problem on internet environment because this is the beneficial place for fraudsters stealing users' card details. Huge amount of money are stolen through online

transactions and this causes e-consumers significant concerns because of its dramatic speed day-by-day. These issues are challenging the development of online credit card transactions. Consequently, detection of fraudulent transactions and online credit card frauds has become essential to maintain the viability of online transactions and banking systems (Patil et al., 2010).

Over the years, along with the development of fraud detection methods, fraudsters have evolved their fraud methods to avoid detections. Therefore, fraud detection methods need to be developed also. Many algorithms have evolved to detect online credit card frauds based on data mining techniques. This powerful tool helps to extract and analyze various types of data to give decisions to detect and prevent frauds. Some data mining techniques used for credit card fraud detection are neural networks, decision trees, random forests, Hidden Markov Model, Social Network Analysis and logistic regression (Bhattacharyya et al., 2011).

1.1 Statement of the Problem

Presently, due to rapid and wide usage of credit-cards (physical usage and virtual usage), it is difficult to draw out how and for what the credit-cards are used. Therefore, the IP addresses of transactions made online are captured for verification purpose. Thus, when a credit-card fraud is committed, this process needs help from the cyber crime in order to investigate the case (Vatsa 2005). Typically, this process is difficult, time taking and not completely reliable.

Although some systems have been developed to check credit card frauds, it is obvious that we still witness some difficulties and draw backs. This is because the existing systems can only detect fraud after the fraud is done that is, the fraud is detected after the compliant of the card holder and so the card holder faced a lot of trouble before the investigation is concluded. Another major issue with existing systems is that they require labeled data for both genuine as well as fraudulent transactions to train the classifiers and getting real world fraud data is one of biggest problems associated with credit card fraud detection. Also, these approaches cannot detect new kinds of frauds for which labeled data is not available. Also, large data is needed for which a log is maintained that consists of all the transaction.

Therefore it is all this difficulties and irregularities that are associated with the existing system that we wish to address by the introduction of the Hidden Markov Model and the researcher strongly believe that this model will go a long way in address some irregularities and the problems of the existing system.

1.2 Objectives of the Study

The objectives of credit card detection system using Hidden Markov Model include developing a system that should be able to:

i. Identify the spending profile or habit of credit card holders and apply this knowledge in deciding whether a transaction is fraudulent or not.
ii. Determine the spending patterns on every credit card during transaction to figure out any inconsistency with respect to the "usual" spending patterns.

2.0 Analysis of the Existing System

Credit card fraud detection has drawn a lot of research interest and a number of techniques, with special emphasis on data mining and neural networks, have been suggested. Ghosh and Reilly (1994) have proposed credit card fraud detection with a neural network. They have built a detection system which is trained on a large sample of labeled credit card account transactions. These transactions contain example fraud cases due to lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud and NRI (non-received issue) fraud. Recently, Syeda et al (2002) have used parallel granular neural networks (PGNNs) for improving the speed of data mining and knowledge discovery process in credit card fraud detection. A complete system has been implemented for this purpose.

Stolfo et al (1997) suggest a credit card fraud detection system (FDS) using meta-learning techniques to learn

models of fraudulent credit card transactions. Meta-learning is a general strategy that provides a means for combining and integrating a number of separately built classifiers or models. A meta-classifier is thus trained on the correlation of the predictions of the base classifiers. The same group has also worked on a cost-based model for fraud and intrusion detection. They use Java agents for Meta-learning (JAM), which is a distributed data mining system for credit card fraud detection. A number of important performance metrics like TP-FP (True Positive - False Positive) spread and accuracy have been defined by them.

Aleskerov et al (1997) present CARDWATCH, a database mining system used for credit card fraud detection. The system, based on a neural learning module, provides an interface to a variety of commercial databases. Kim and Kim (2002) have identified skewed distribution of data and mix of legitimate and fraudulent transactions as the two main reasons for the complexity of credit card fraud detection. Based on this observation, they use fraud density of real transaction data as a confidence value and generate the weighted fraud score to reduce the number of misdetections. Fan (1999). et. al, suggested the application of distributed data mining in credit fraud detection. Furthermore, Aleskerov et al (1997) developed an approach that involves advanced data mining techniques and neural network algorithms to obtain high fraud coverage.

Chiu and Tsai (2004) proposed web services and data mining technique to establish a collaborative scheme for fraud detection in the banking industry. With this scheme, participating banks share knowledge about the fraud patterns in a heterogeneous and distributed environment. To establish a smooth channel of data exchange, web services techniques such as XML, SOAP and WSDL are used. While Phua et al (2007) did an extensive survey of existing data mining based fraud detection systems and published a comprehensive report. Prodromidis and Stolfo (1999) applied an agent based approach with distributed learning for detecting frauds in credit card transactions. It is based on artificial intelligence and combines inductive learning algorithms and meta-learning methods for achieving higher accuracy.
Moreso, Phua et al (2004) suggest the use of meta classifier similar to in fraud detection problems. They consider naive Bayesian, C4.5 and Back Propagation neural networks as the base classifiers. A meta-classifier is used to determine which classifier should be considered based on skewness in data. Although they do not directly use credit card fraud detection as the target application, their approach is quite generic. More so, Vatsa et al (2005) proposed a game theoretic approach to credit card fraud detection. They model the interaction between an attacker and a fraud detection system as a multi-stage game between two players, each trying to maximize his payoff.

Apart from the above mentioned credit card fraud detection techniques there are some of the recent credit card fraud detection techniques that gained attention. The following is the explanation in brief about each of these techniques:

1. **Fusion of Dempster-Shafer theory and Bayesian learning**
   For credit card fraud detection, this approach is a cross technique that merges the results obtained from present and precedent behavior. Any credit card owner has certain spending pattern for his purchases online that are recorded in his transactions account. This credit card fraud detection system comprises of *mainly* four elements.

i. **Rule-based filter:** the doubt level of every transaction that is made is extracted depending the variations from the normal form of spending patterns

ii. **Dempster-Shafer adder:** in this component, all the transactions that are doubtful obtained by rule-based filter are combined to form a primary belief.

iii. **Transaction History Database:** here all the values formed as a primary belief are combined to form on the whole belief by its theory.

iv. **Bayesian Learner:** here once after any transaction is believed to be suspicious, it is strengthened with fraudulent or weakened with genuine transaction.

This approach has high accuracy and also improves the fraud detection rate when compared with previous credit card detection techniques. The only issue with this mechanism is, it is very expensive and processing speed is less.

2. Blast-Ssahain Credit Card Fraud Detection

Blah-Fraud Detection System algorithm is the improved form of Blast and Ssaha algorithm. These two algorithms are pretty much proficient sequence aligning algorithms in detecting credit card frauds. In the sequence alignment of Blah-Fraud Detection System Algorithm, there are two stages where a profile analyzer obtains the correspondence between the transactions that are incoming in sequence with all the past and sequence of genuine transactions; made, in the past. The abnormal transactions detected by the profile analyzer are then passed into a deviation analyzer for checking with the past fraudulent transactions behavior if present. Thus based on these two analyzers a conclusion is drawn and final decision is taken. The performance of this mechanism in detecting the credit card frauds is good and its accuracy is high. Also processing speed is fast but the problem using this credit card fraud detection approach is that it cannot detect the duplicate transactions or cloned credit card frauds.

3. Fuzzy Darwinian Detection of Credit Card Fraud

This detection method of credit card frauds uses genetic programming in order to develop some fuzzy logic rules that can be helpful in determining the suspicious and non-suspicious classes of transactions. When the information related to a transaction is provided to the Fraud Detection System, the system using the classifiers, will determine the transaction as either safe or suspicious. The absolute system is capable of providing good accuracy rate and less false rate. This is not applicable in the case of online transactions practically as it is highly expensive. Also, its processing speed is very less.

4. Credit Card Fraud Detection Using Bayesian and Neural Networks

This Bayesian and Neural Networks mechanism is an automatic credit card fraud detecting system using a machine learning approach. Both the Bayesian and Neural Networks approaches are suitable for analysis in cases of uncertainty. In general, an artificial neural network will have an inter-connected group of artificial neurons and the pattern classification of frequently used neural networks. This is referred to as feed-forward network and has three layers: input layer, hidden layer and an output layer. Here all the incoming transactions are passed through these three layers known as forward propagation. The artificial neural networks will hold the training information and compares with the inward transactions where the neural networks are originally fed with the common spend pattern and behavior of the card holder. Any suspicious transactions are sent back to the neural networks which classifies the transactions as safe and suspicious. This is where the Bayesian networks use artificial intelligence concept comprising of various methods like data mining and machine learning algorithms to bring out results. The Bayesian belief networks are very efficient in cases where small amount of data is known and incoming information is unstable or not completely available. This helps in data identification and classification and these neural networks need not be re-programmed. These provide good accuracy but require lot of training in achieving high processing speed.

The problem with most of the above-mentioned approaches for credit card fraud detection is that they require labeled data for both genuine as well as fraudulent transactions to train the classifiers. Getting real world fraud data is one of the biggest problems associated with credit card fraud detection. Also, these approaches cannot detect new kind frauds for which labeled data is not available. In contrast, this work wish to explore Credit Card Fraud Detection System using Hidden Markov Model (HMM), which does not require fraud signatures and yet, is able to detect frauds by considering a cardholder's spending habit.

2.1 Flowchart of the Existing System

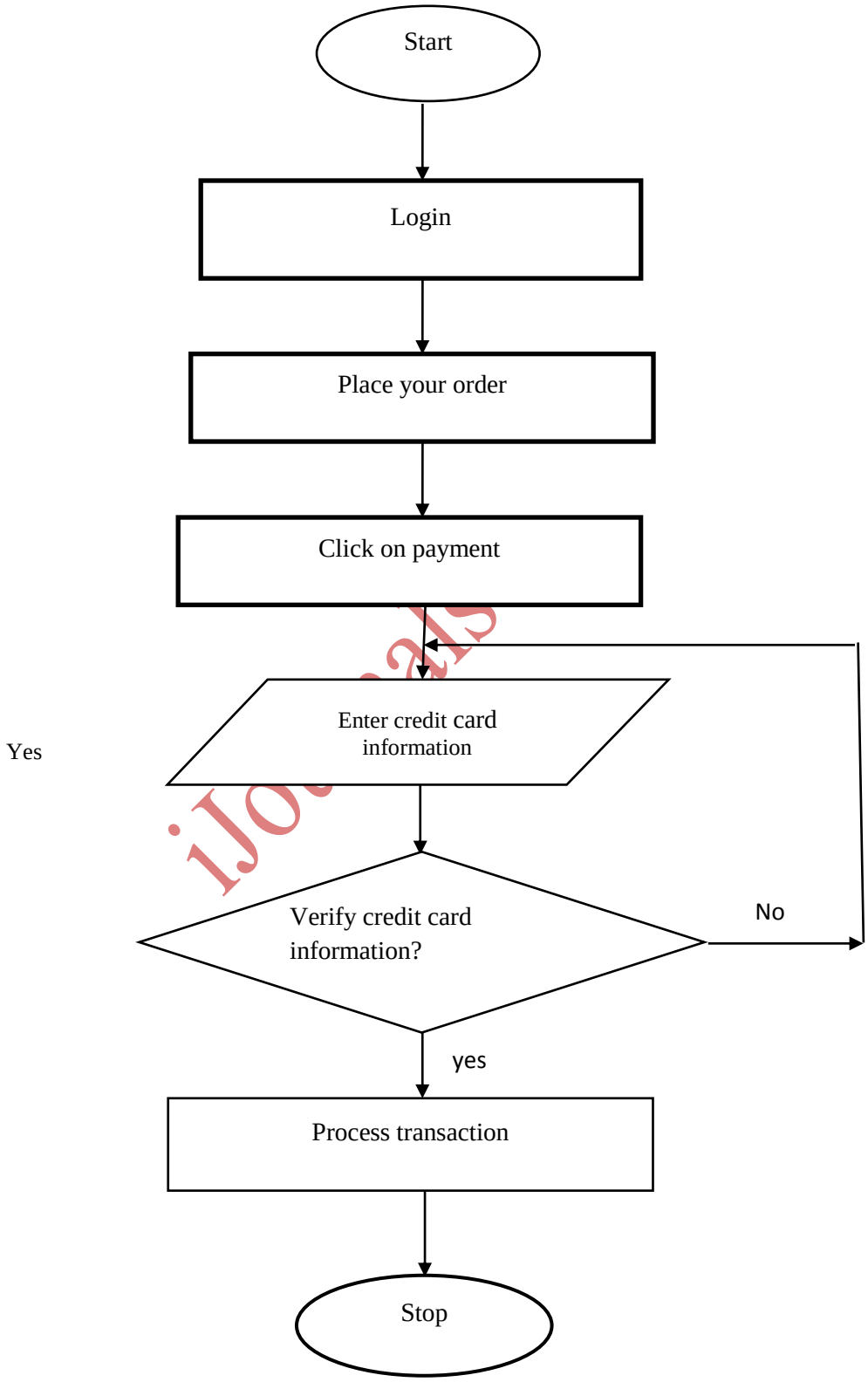


Figure 1: Flowchart of the Existing System

2.2 The Analysis of the New System Using Hidden Markov Model

Having observed the weaknesses and inefficiencies associated with the present fraud detection systems, this work introduced Credit Card Fraud Detection System using Hidden Markov Model (HMM) that does not need any fraud transactions information of the credit-card and still can detect the fraud actions.

This model considers the spending routines of the credit-card holder. Here, transactions of a credit card processing series are modeled by the stochastic procedure. The information related to the purchases with respect to an individual credit card holder is not identified by that particular bank's Fraud Detection System which issued the credit-card. This is the primary factor of Markov chain, which is signified but not noticeable. The credit card transactions can be viewed or known by stochastic process which gives the series of the spending information.

A Hidden Markov Model is a finite set of states; each state is linked with a probability distribution (Rabiner, 1989). Transitions among these states are governed by a set of probabilities called transition probabilities. In a particular state a possible outcome or observation can be generated which is associated symbol of observation of probability distribution. It is only the outcome, not the state that is visible to an external observer and therefore states are "hidden" to the outside; hence the name Hidden Markov Model. Hidden Markov Model is a perfect solution for addressing detection of fraud transaction through credit card. One more important benefit of the HMM-based approach is an extreme decrease in the number of False Positives transactions recognized as malicious by a fraud detection system even though they are really genuine.

In this prediction process, HMM consider mainly three price value ranges such as.

- Low (1),
- Medium (m) and,
- High (h).

First, it will be required to find out transaction amount belonging to a particular category either it will be in low, medium, or high ranges.

It involves two modules which include:

i.Online Shopping

It comprises of many steps, first is to login into a particular site to purchase goods or services, then choose an item and next step is to go to payment mode where credit card information will be required. After filling all these information, the page will be directed to the fraud detection system which will be installed at bank's server or merchant site**.**

ii. Fraud Detection System

All the information about credit card (Like Credit card number, credit card CVV number, credit card Expiry date, name on credit card etc.) will be checked with credit card database. If user entered database is correct then it will ask Personal Identity number (PIN). After matching of Personal Identity number (PIN) with database and account balance of user's credit card is more than the purchase amount, the fraud checking module will be activated. The verification of all data will be checked before the first page load of credit card fraud detection system. If user credit card has less than ten (10) transactions, it will directly ask to provide personal information to do the transaction. Ones database of ten (10) transactions is obtained, fraud detection system will start working. By using this observation, it determines users spending profile.

The purchase amount will be checked with spending profile of user. By transition probabilistic calculation based on HMM, it concludes whether the transaction is real or fraud. If transaction may be concluded as fraudulent transaction then user must enter security information. This information is related with credit card (like account number, security question and answer which are provided at the time of registration). If transaction is not fraudulent then it will direct to give permission for transaction. If the detected transaction is fraudulent then the security information form will arise. It has a set of question where the user has to answer them correctly to do the transaction. These forms consist of information such as personal, professional, address; dates of birth, etc and are available in the database**.** If the information entered by the user matches with

database information, then transaction will be done securely else user transaction will be terminated and transferred to online shopping website.

2.3 Specification of the Hidden Markov Model

Hidden Markov model can be fully specified with five parameters according to (Rabiner, 1989).

1. N, the number of states in the model. We denote the set of all possible states as S= {S_1, S_2…, S_N}, the state at time t as q_t.

2. M, the number of distinct observation symbols per state, i.e., the discrete alphabet size of the output set. We denote the set of all possible output symbols as V = {v_1, v_2, ..., v_m}. the output symbols at time t as O_t. The sequence of observed symbols is denoted as O = 0_1 0_2…..0_T.

3. The state transition probability distribution A={a_{ij}}, where a_{ij} = P[q_{t+1} = Sj|q_t =Si],

   $1 \leq i, j \leq N$.

4. The observation symbol probability distribution in state j, B ={bj (k)}, where bj (k)

   =P[O_t =Vk|qt =Sj], $1 \leq j \leq N$, $1 \leq k \leq M$.

5. The initial state distributions πi = P[q_1 = Si], $1 \leq i \leq N$

A compact representation of the model is λ = (A, B, π), where N, M are implicitly implied by A and B, (Rabiner, 1989).

2.4 Hidden Markov Model for Credit Card Transaction Processing

To map the credit card transaction processing operation in terms of an HMM. we start by first deciding the observation symbols in our model. We quantize the purchase values x into M price ranges VI, V2 ...Vm, forming the observation symbols at the issuing bank. The actual price range for each symbol is configurable based on the spending habit of individual cardholders. Hmm determines these price ranges dynamically by applying a clustering algorithm on the values of each cardholder's transactions.
We use V_k, *k* =1, 2, ...M to represent both the observation symbol as well as the corresponding price range.
In this work, only three price ranges were considered, namely, low (*l*), medium *(m)* and high (*h*). Our set of observation symbols is, therefore, *V* = {*l*, *m*, *h*} making *M*= 3. For example:
Let *l* = (0, ₦200], *m* = (₦200, ₦600] and *h* = (₦600, credit card limit]. If a cardholder performs a transaction of ₦250, then the corresponding observation symbol is *m*.

A credit cardholder makes different kinds of purchases of different amounts over a period of time. One possibility is to consider the sequence of transaction amounts and look for deviations in them. However, the sequence of types of purchase is more stable compared to the sequence of transaction amounts. The reason is that. a cardholder makes purchases depending on his need for procuring different types of items over a period of time. This, in turn, generates a sequence of transaction amounts. Each individual transaction amount usually depends on the corresponding type of purchase. Hence, we consider the transition in the type of purchase as state transition in our model. The type of each purchase is linked to the line of business of the corresponding merchant. This information about the merchant's line of business is not known to the issuing bank running the Fraud Detection System. Thus, the type of purchase of the cardholder is hidden from the Fraud Detection System. The set of all possible types of purchase and equivalently, the set of all possible lines of business of merchants form the set of hidden states of the HMM.

2.5 Dynamic Generation of Observation Symbols

For each cardholder, we train and maintain an HMM. To find the observation symbols corresponding to individual cardholder's transactions dynamically, we run a clustering algorithm on his past transactions. Normally, the transactions that are stored in the issuing bank's database contain many attributes. For this work, only the amounts that the cardholder spent were considered in his transactions. Although there are various clustering techniques. K-means clustering algorithm was used to determine the clusters. K-means is an unsupervised learning algorithm for grouping a given set of data based on the similarity in their attribute values.

Each group formed in the process is called a cluster. The number of clusters *K* is fixed a priori. The grouping is done by assigning each observation to the cluster whose mean is closest to it.

In this work, *K* is the same as the number of observation symbols *M*. Let c_1, c_2, ……,c_m be the centroids of the generated clusters. These centroids or mean values are used to decide the observation symbols when a new transaction comes it.
Let *x* be the amount spent by the cardholder in transaction. Fraud Detection System generates the observation symbol for x (denoted by O_x) as follows:

$$O_x = V_{\underset{i}{arg}\min} | x - ci | \qquad (3.1)$$

Where: V= Observation symbol i.e V = (l, m, h) X= amount spent by card holder in transaction, C_i = the centroid or mean of the generated clusters. ie. cl, cm, ch.

As mentioned before, the number of symbols is 3 in our system. Considering *M*= 3, if we apply X-means algorithm on the transactions in table 1, we get the clusters as shown in table 2 with *cl, cm* and *ch* as the respective centroids. It may be noted that the amounts 6, 8 and 10 have been clustered together as *cl* resulting in a centroid of 8. The percentage (p) of total number of transactions in this cluster is thus 30%. Similarly, amounts 15, 15, 20,20 and 25 have been grouped in the cluster *cm* with centroid 19 while amounts 40 and 80 have been grouped together in cluster *ch* with centroid 60. *cm* and *ch*, thus, contain 50% and 20% of the total number of transactions. When the Fraud Detection System receives a transaction for this cardholder, it measures the distance of the purchase amount x with respect to the means *cl*, *cm* and *ch*, to decide (using Eq. 3.1) the cluster to which the transaction belongs and hence, the corresponding observation symbol.

For example, if x = $10, then from table 3.2 using Eq. 3.1, the observation symbol is
V_1 = l.

Table 1: Sample Transactions with the Amount spent in each Transaction

Transaction number	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Amount (₦)	40	25	15	6	8	20	15	20	10	80

Table 2: Output of K-means Clustering Algorithm

Cluster mean/centroid name	C1	Cm	Ch
Observation symbol	V1=1	V2=m	V3=h
Mean value (Centroid)	8	19	60
Percentage of total transactions (p)	30	50	20

2.6 Spending Profile of Cardholders

The spending profile of a cardholder suggests his normal spending behavior. Cardholders can be broadly categorized into three groups based on their spending habits, namely, high spending (hs) group, medium spending (ms) group and low spending (Is) group.

Cardholders that belong to the high spending group, normally use their credit cards for buying high-priced items. Similar definition applies to the other two categories also. Spending profiles of cardholders are determined at the end of the clustering step.

Let *Y* be the percentage of total number of transactions of the cardholder that belong to cluster with mean *ci.* Then, the spending profile (SP) of the cardholder *g* is determined as follows:

$$SP(g) = V_{\underset{i}{arg}\max}(Y) \qquad (3.2)$$

Where: SP (g) = Spending profile of card holder g, V observation symbol ie. V = (1, m, h), Y = the percentage of total number of transactions of the card holder that belongs to each cluster.

Thus, spending profile denotes the cluster number to which most of the transactions of the cardholder belong. From the example in table 3.2, the spending profile of the cardholder is 2, i.e. *m* and hence the cardholder belongs to the medium spending group.

2.7 Advantages of the Proposed System

The new system will be able to offer the following benefits:

- It can prevent all identifiable frauds after effective tracking.
- Helps to obtain high fraud coverage combined with a low false alarm rate by drastically reducing the number of false positive transactions.
- This project helps in detecting the credit card frauds easily and also with high processing speed.
- The system can produce the most accurate detection.
- There is no loss of revenue due to delivery of goods to fraudsters since the fraud can be detected before it is committed.
- In case of the existing system even the original card holder is also checked for fraud detection. But in this system no need to check the original user as we maintain a log.
- There is no waste of productivity in reviewing all online orders manually.

2.8 The Overall Data Flow Diagram of the Proposed Solution

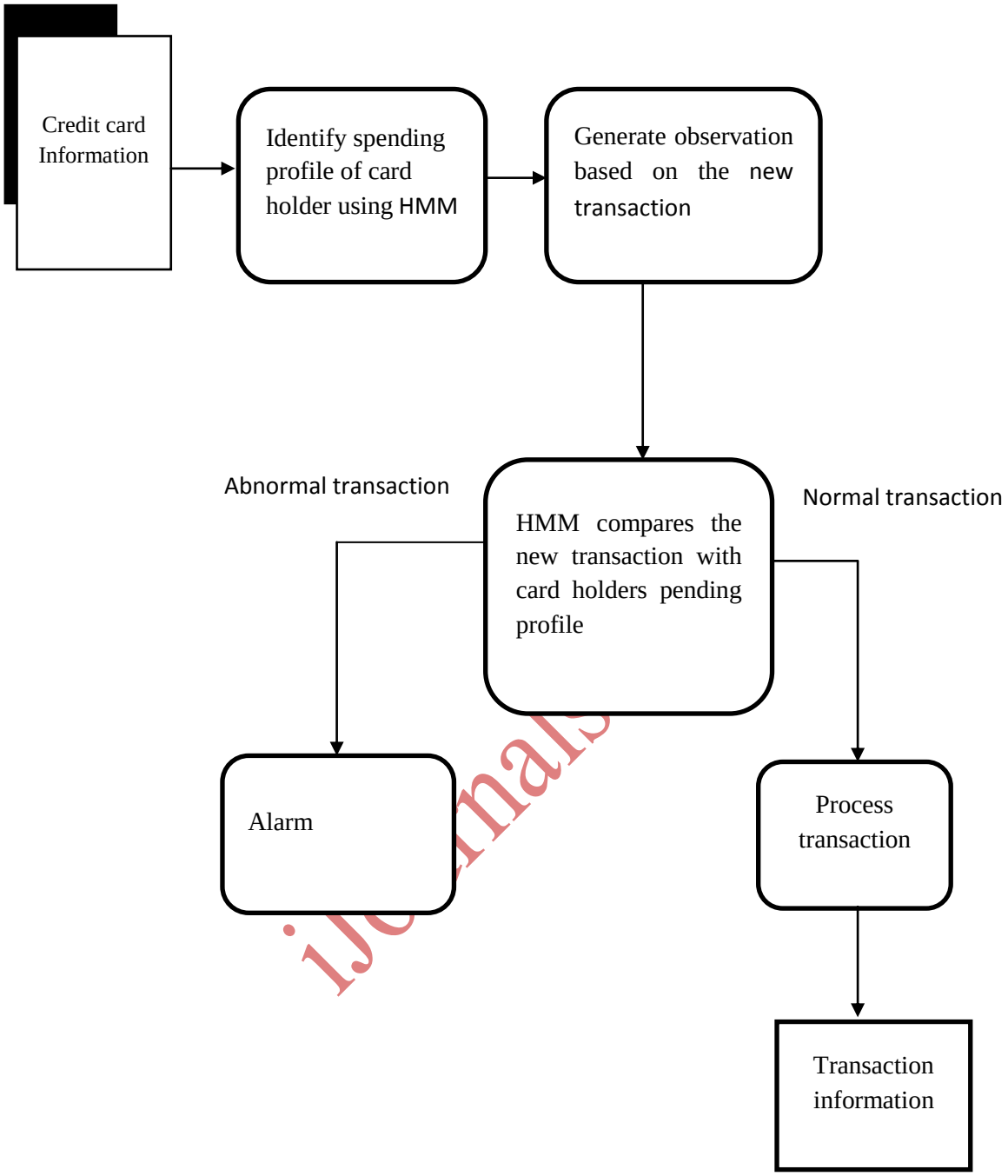


Figure 9: Overall Data Flow Diagram of the Proposed Solution

2.9 Flowchart of the Proposed System

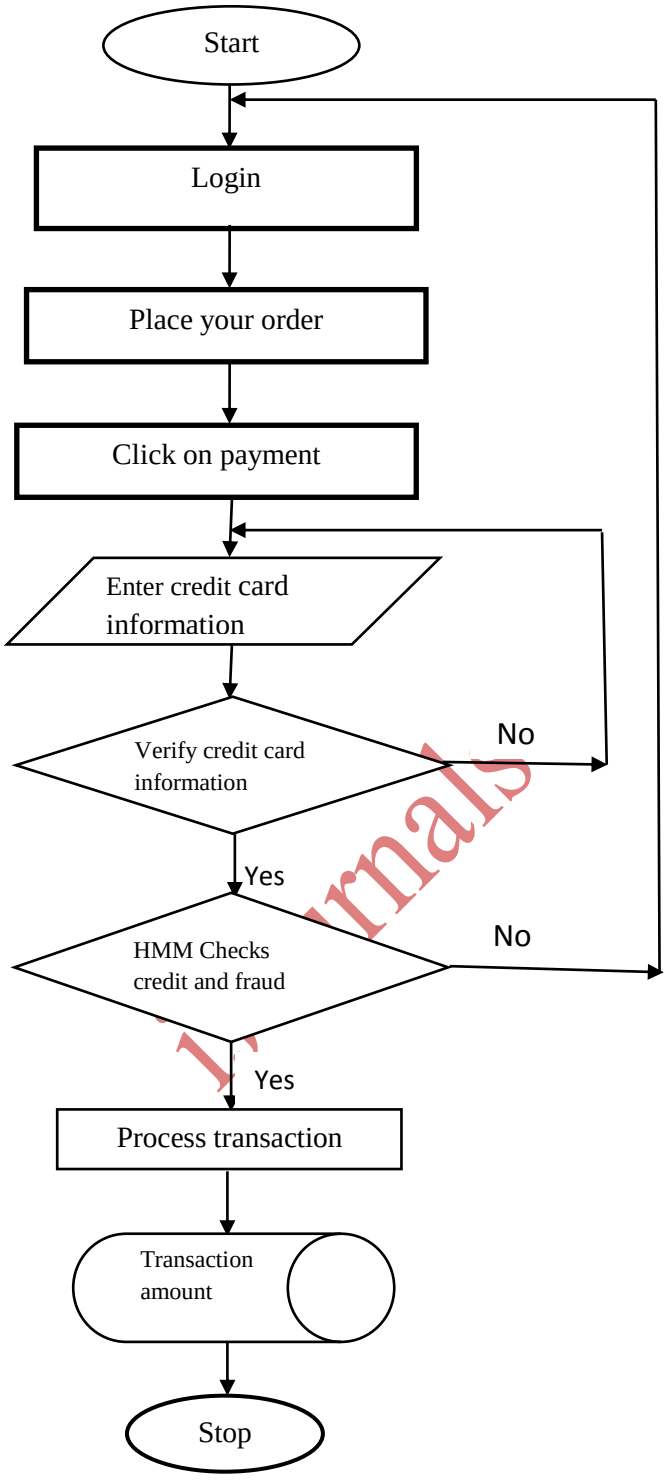


Figure 2: Flowchart of the Proposed System

h Level Model of the Proposed System

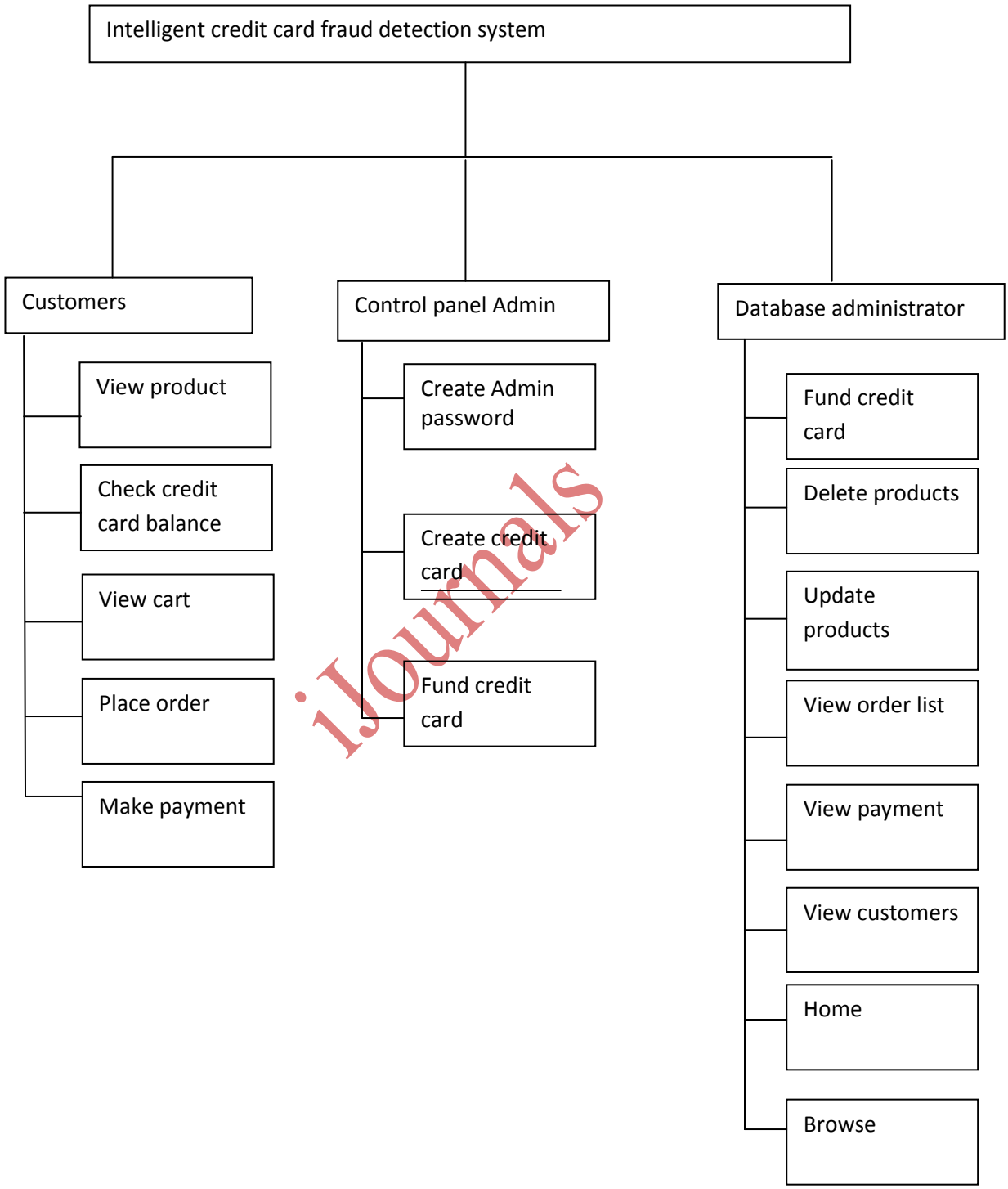


Figure 3: High Level Model of the Proposed System

2.11 Main Menu

The main menu is made up of three sub-menus which include customers' transactions, control panel admin credit cards and database administrator online products.

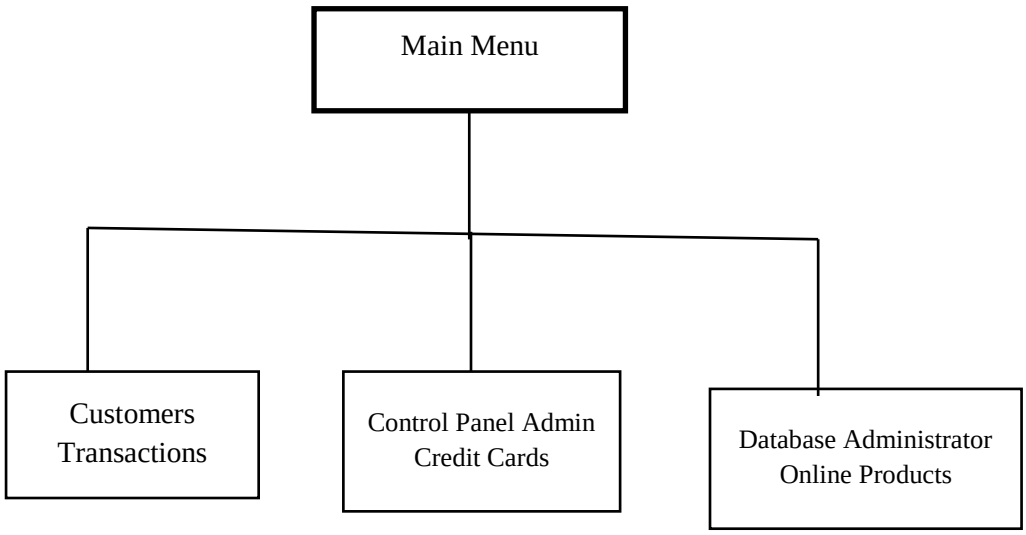


Figure 4: Main Menu

Table 3: Ecard table

FIELD	TYPE	NULL	KEY	DEFAULT
First name	Varchar (20)	YES		(NULL)
Last name	Varchar (20)	YES		(NULL)
Card number	Varchar (40)	NO	PRI	
Signature	Varchar (20)	YES		(NULL)
Expdate	Date	YES		(NULL)
Pin	Int (6)	YES		(NULL)
Amount	Double	YES		(NULL)
Age	Int (6)	YES		(NULL)
Pet name	Varchar (20)	YES		(NULL)

Table 4: Order Table

Field	Type	Null	Key	Default
Sn	Int(4)	NO	PRI	(NULL)
Id	Int(4)	YES		(NULL)
User	Varchar (20)	YES		(NULL)
Dates	Date	YES		(NULL)
Status	Varchar (10)	YES		(NULL)
Price	Double	YES		(NULL)
Product	Varchar (20)	YES		(NULL)

2.12 Input Specification

The new system was designed in such a way that well-structured forms will be used to capture customer's transaction information as well as some administrative updates. Below are some of the input forms designed in the new system.

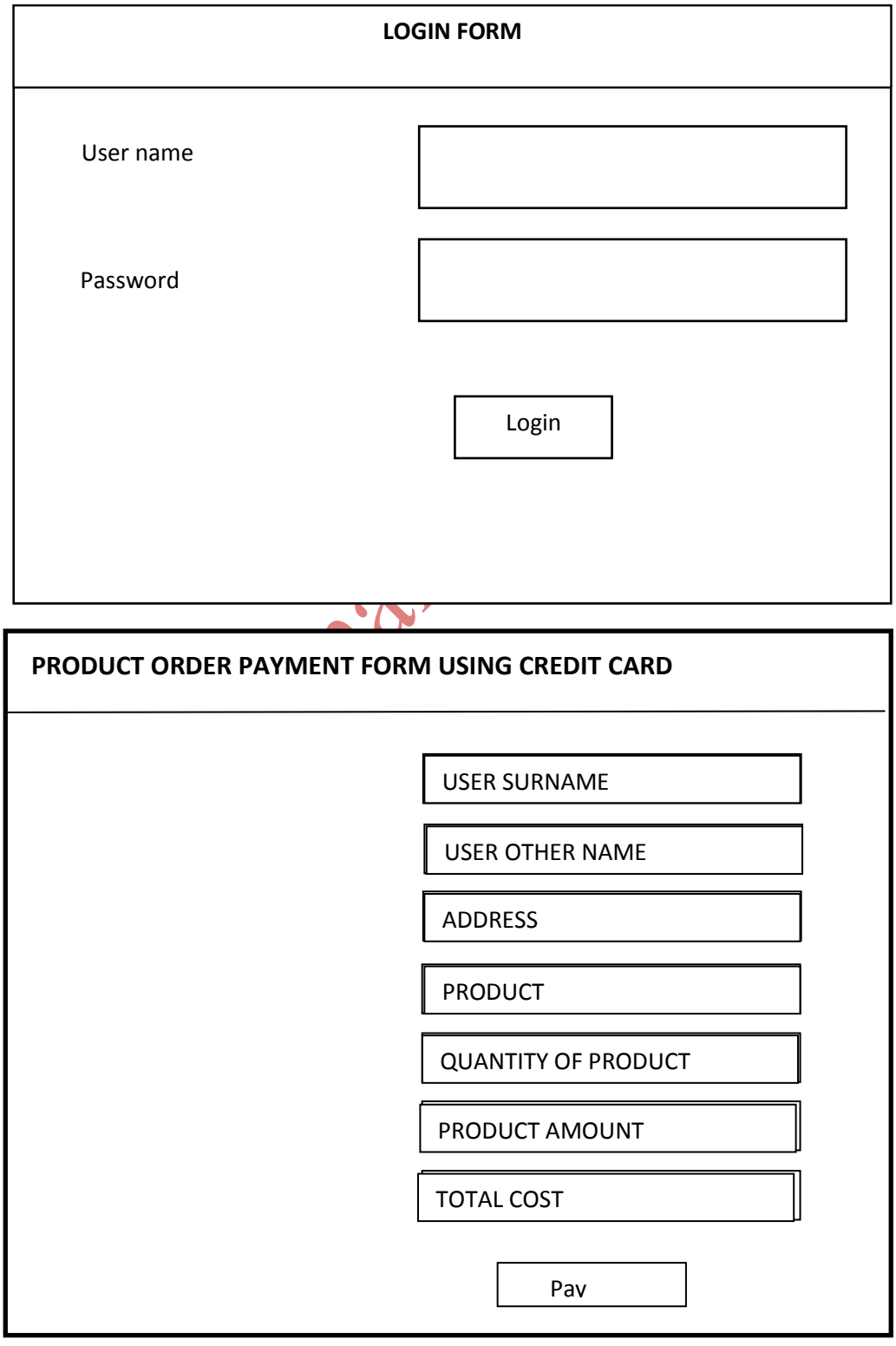
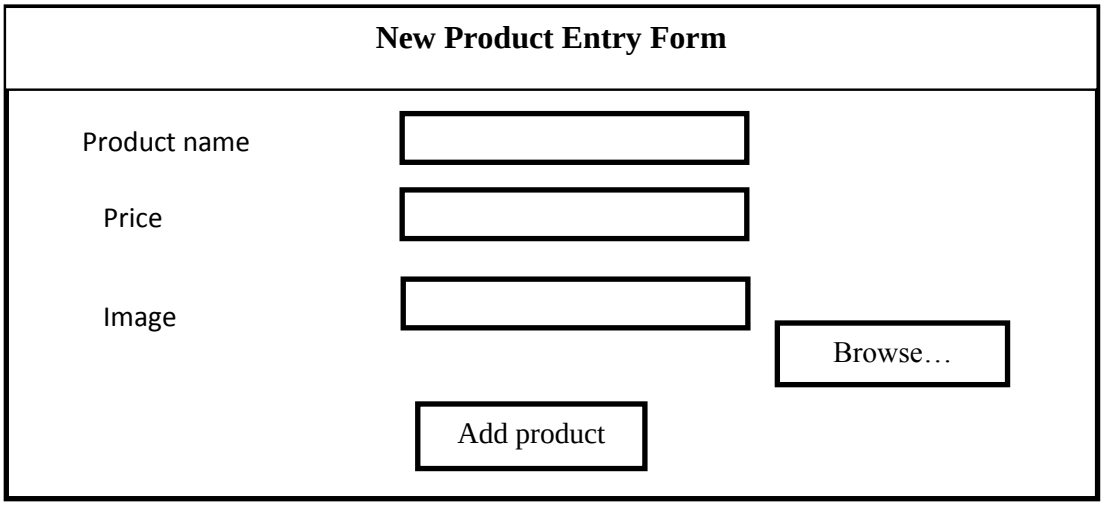
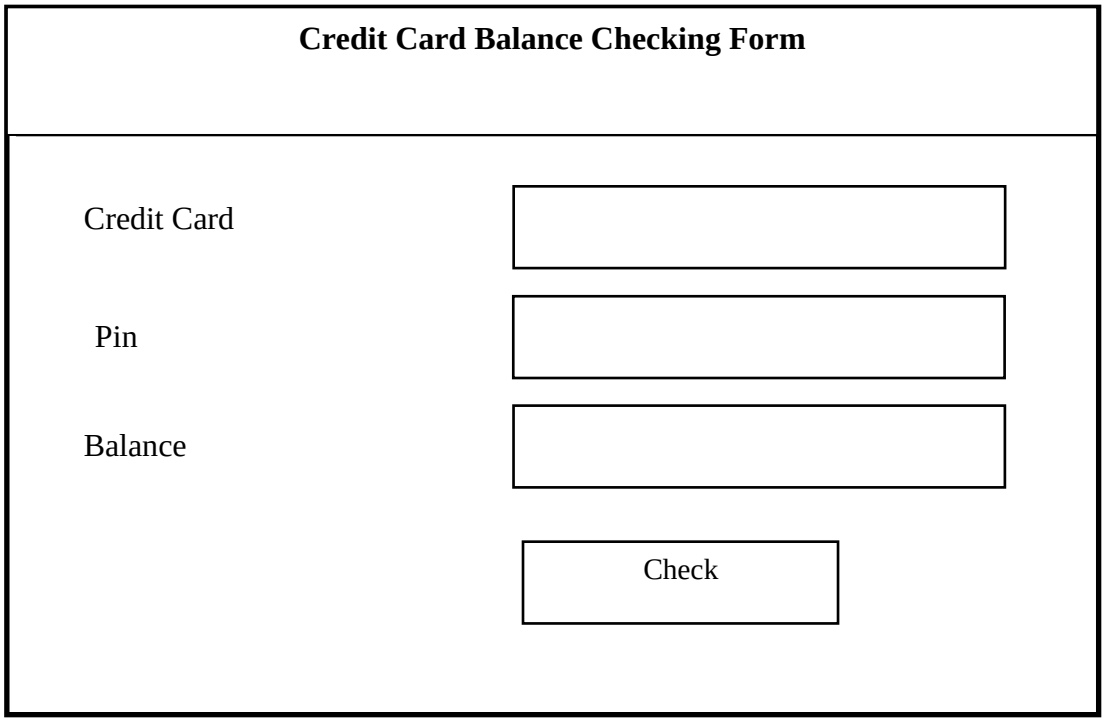


Figure 6: Product Order Payment Form Using Credit Card





2.13 Output Specification

The website was created to enable admin users to retrieve vital information from the site for management use. Reports on credit card balance, product ordering, customer's file and income report can be generated from the system. Below are some of the report formats designed new systems.

Table 5: Product Order Report

User	Product ID	Product name	Price	Status	Date
Chidi	1	Tv	70,000	Pd	2020-08-13
Oge	2	Laptop	82,000	Pd	2020-12-01

Table 6: Income Report

User	Price	Date
Chidi	200,000	2020-04-01

Table 7: Customer Register

Username	First Name	Last Name	Phone	E-mail	Home
Overcomer	Alex	Ifeanyi	08035509616	chi@gmail.com	Awka

3.3 Conclusion

The work, an intelligent based credit card fraud detection system using Hidden Markov Model which is based on card holder's spending habits has been proved to be very effective in eradicating frauds associated with credit card transaction, as it thoroughly investigates every credit card transactions to ensure that any fraudulent transactions are restricted while reducing false positive transaction by ensuring that genuine card users are not denied transactions.

3.4 Areas of Application of this Work

This work will be very helpful for the banks and other organizations that issue credit-cards in detecting the credit card frauds committed. This will equally help the organizations that issue credit-cards provide better services thereby gathering more customers in wide sense.

3.6 Recommendation

To ensure that customer's confidence is not lost through insecurity of credit card transactions, I recommend that banks and other financial institutions dealing on credit cards, should embrace this system.

3.2 Appendix 1: Sample Output

Login Page

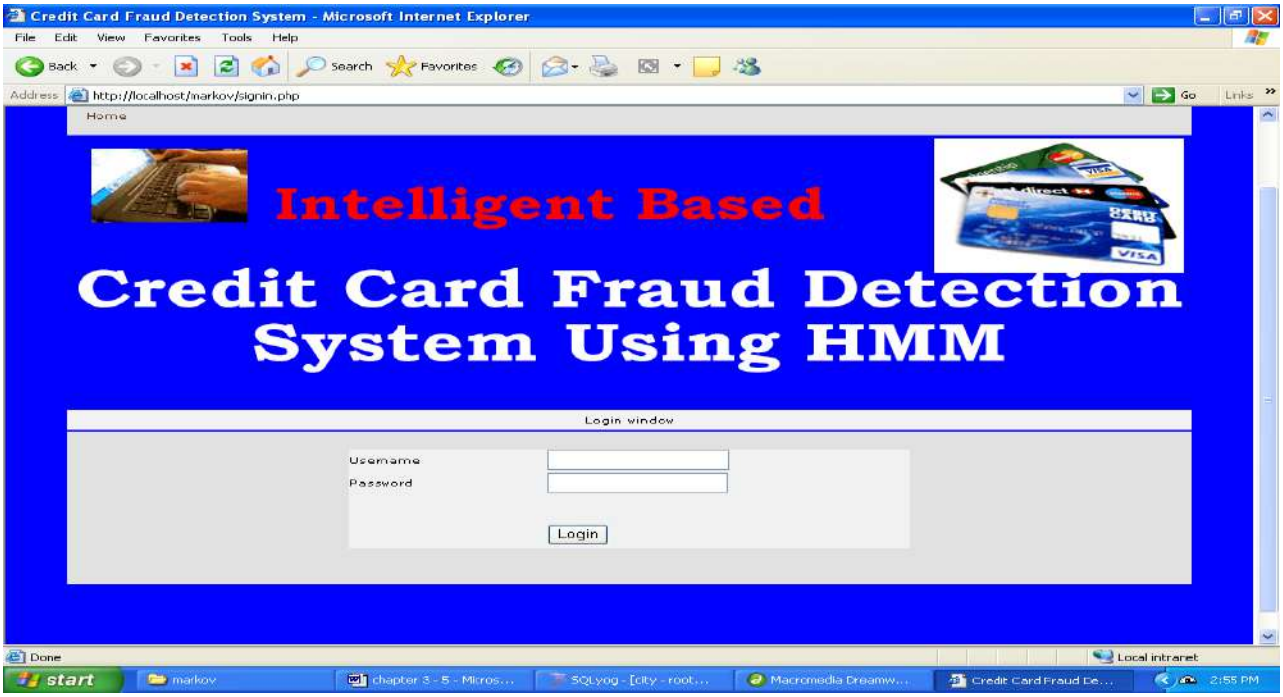


Figure 14: Product Order Payment Form Using Credit Card

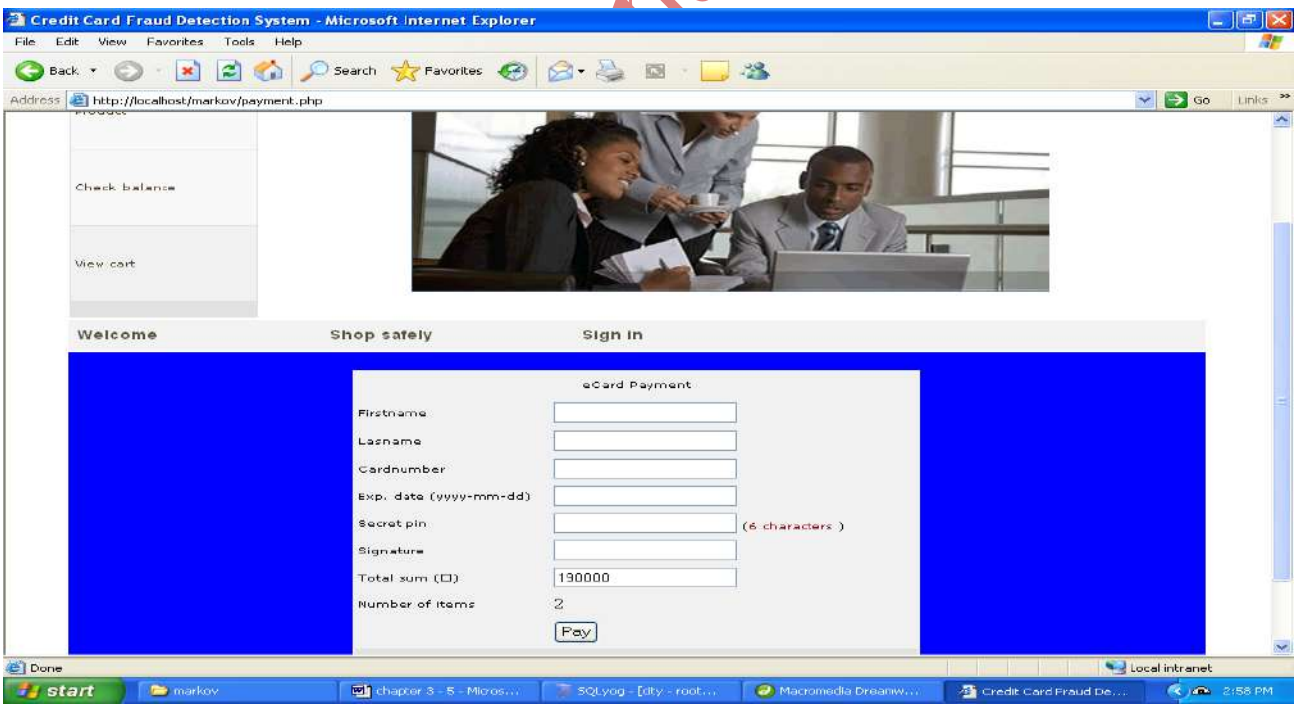


Figure 15: Credit Card Balance Checking Form

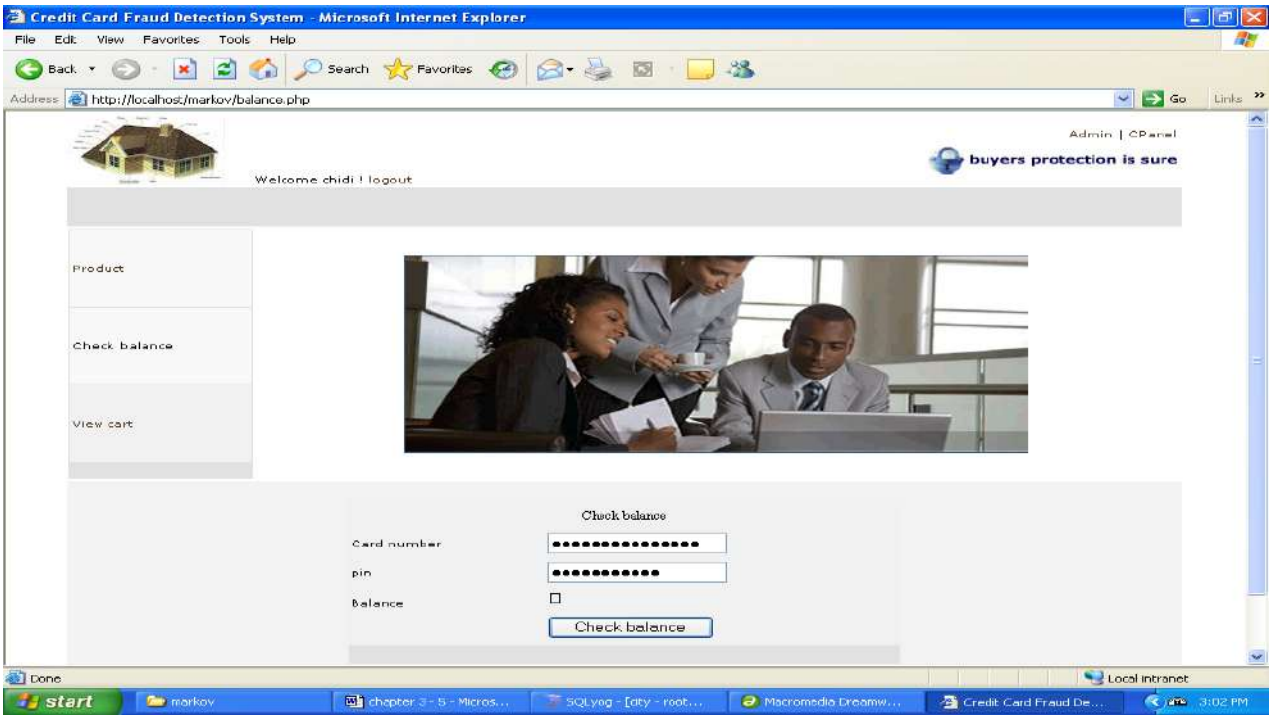


Figure 16: New Product Entry Form

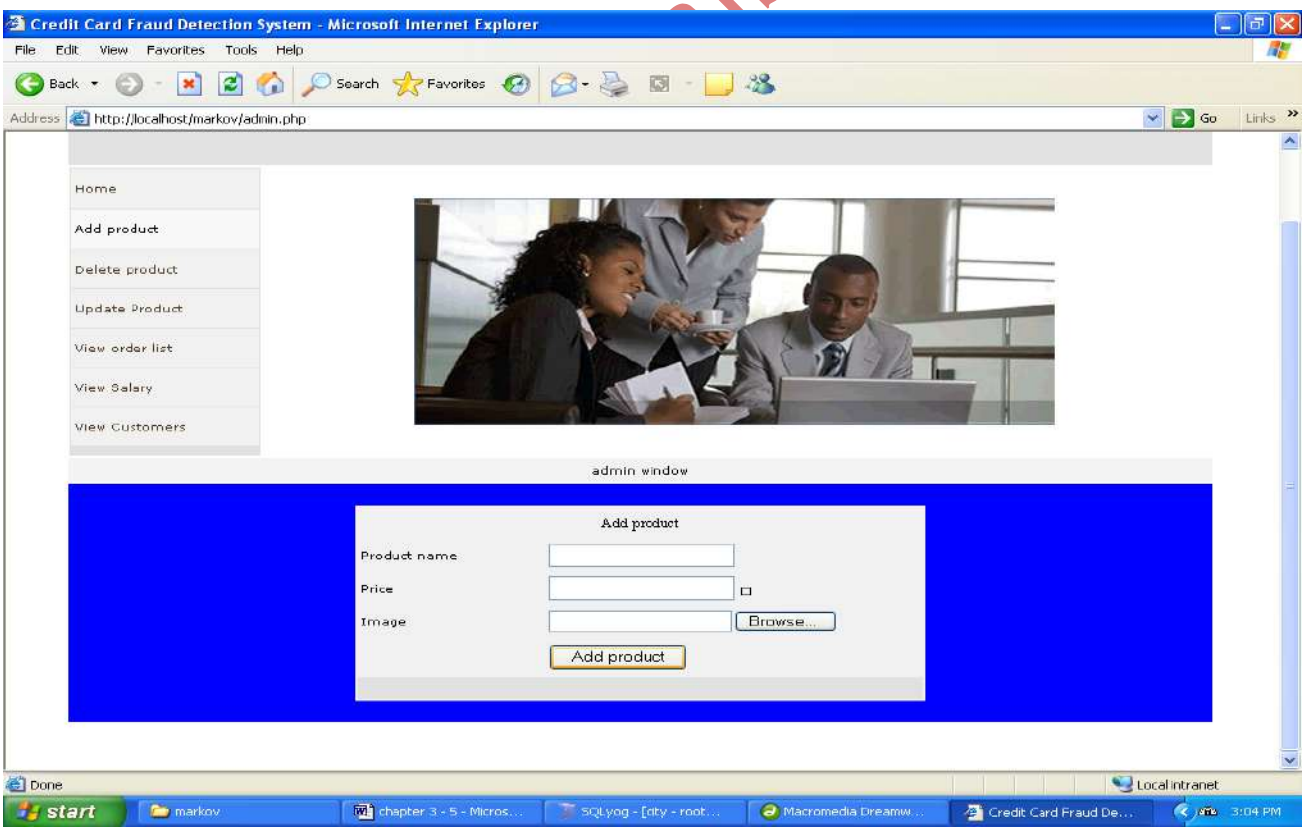


Figure 17: Credit Card Account Opening Form

REFERENCES

Aleskerov, E., Freisleben, B., AndRao. B(1997) Card watch: A Neural Network Based
Database Mining System For Credit Card Fraud Detection, Proceedings of IEEE/Iafe: Computational Intelligence For Financial Eng. Pp. 220-226.

Bhattacharyya, S., Jha, S., TharakunneL K. & Westland. J. C. (2011), Data mining for credit card fraud: A comparative study. *Decision Support Systems* 50, pp. 602-613, DOI: 10.1016/j.dss.2010.08.008.

Chiu C., and Tsai, C., (2004). A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection, Proceedings of IEEE International Conference e-Technology, e-Commerce and e-Service. pp. 177-181.

Fan,W., Prodromidis, A. L., And Stolfo, S. J.. (1999). Distributed DataMiningIn Credit Card Fraud Detection, IEEE Intelligent Systems, Vol. 14. No. 5. Pp 67-74.

Ghosh, S. , Reilly, D.L. (1994). Credit Card Fraud Detection with a Neural- Network, Proc. 27th Hawaii Int'l Conf. System Sciences: Information Systems: Decision Support And Knowledge-Based Systems, Vol. 3, Pp. 621-630.

Kim, M..T. and Kim. T.S. 2002. A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection, *Proc. International Conference on Intelligent Data Engineering and Automated Learning*, Lecture Notes in Computer Science, Springer Verlag, no. 2412, pp. 378-383.

Patil, D. D., Karad, S. M., Wadhai, V.M., Gokhale, J.A. & Halgacaiuir. P. S. (2010). Efficient Scalable Multi-Level Classification; scheme far Credit Card Fraud Detection, International Journal of Computer Science and Network Security. Vol. 10, No. 8, Pp. 17-18.

Phua, C., Lee. V.. Smith. K. and Gayler R (2007) . A comprehensive Survey of Data Mining based Fraud Detection Research, Pp. 190-200.

Phua, C., Alahakooo. D. and Lee, V.(2004). Minority Report in Fraud Detection
Classification of Skewed Data. *ACMSIGKDD Explorations, Newsletter*, vol. 6. no 1, pp. 50-59.

Rabiner. L.R. (1989). A Tutorial on Hidden Markov Models and Selected Applications In Speech Recognition, *Proceedings of The IEEE*, Vol. 77, No. 2, Pp. 257-286.

Vatsa.Y..Sural, S. and Majumdar, A.K.(2005). A Game-theoretic Approach to Credit Card
FraudDetection, *Proc. 1st International Conference on Information Systems Security*, Lecture Notes in Computer Science, Springer Verlag, pp. 263-276.