Prediction of Credit Default Risk in Financial Institutions Using Artificial Neural Network

[1]Anyanwu Onyekachi Julian and [2]Amanze Bethran Chibuike
[1]Federal College of Education (Technical), Umunze Anambra State
and [2]Imo State University, Owerri, Imo State
DOI: 10.56201/ijcsmt.v10.no5.2024.pg123.134

Abstract

This research study, examined the Prediction of Credit Default Risk in Financial Institutions using Artificial Neural Network. Credit default risk has remained one of the important fundamental and critical issues widely studied in the financial institutions in Nigeria. Credit Default risk comes into play when a loan borrower fails to repay his loan within the agreed financial contract. Banks and other financial institutions depend heavily on statistical and machine learning method in predicting loan default to the potential losses of granted loans. These machine learning applications cannot achieve full potential prediction without the semantic context in the data. Neural Network initiate the behavior of the human brain to solve both linear and non-linear statistical problems. The study observe that credit risk is the greatest and leading risk in the banking sector as its effects have crippled several financial institutions and have led to the failure of many. Therefore, the study proposed the adoption of Neural Network in predicting credit default risk to improve the prediction model's accuracy and interoperability. Agile methodology, structured system analysis design methodology (SSADM) was used in the software development to get the total records of credit defaulters. This study designed a system for prediction of Credit Default Risk using Artificial Neural Network. The system is an efficient, accurate and reliable predictive tool which can be employed by financial institutions and lender organizations to solve and manage problem of credit default risk. The programming language PHP Script will be use for the software and My Structured Query Language (MySQL) for database.

Keywords: *Artificial Neural Network, Credit Default Risk, Prediction model, Statistical problems*

INTRODUCTION

The term "Financial institutions" used in this Study covers both banks and non – bank financial institutions in Nigeria. These financial institutions are solely established to render monetary transactions or monetary services such as deposits, loans, mortgage among others to people. Individuals and companies rely on financial institutions for transactions and investments. Examples of financial institutions include; commercial banks, venture capital funds, insurance companies, investment banks, private equity funds, asset management companies, microfinance banks, cooperatives, private lending organizations, among others. Each of these financial institutions has the function either to provide services and accounts for the public or to provide more specialized services to certain customers, all boiling down to rendering money- related services to people. Admittedly, it has been ascertained that financial institutions are an inevitable aspect of every country or economy because of the role they play, and so, form the financial backbone of every nation. Ultimately, any country that is financially stable invariably enjoys economic stability. In the same way, bankruptcies of financial institutions can also create panic in a country. However, the primary aim of any financial firm is to make profit. They make profits through offering loans to individuals, corporations and organizations. Literature have revealed that loans are the major source of income to financial institutions. When loans are paid back by the borrowers in due time, they add profit to the financial institutions but disrupt cash flow in such financial institution when they are withheld by borrowers, whether intentionally or unintentionally. As such, financial institutions can run into trouble of loss in their lending services where no proper lending measures are put in place or proper credit risk evaluation is not conducted on a borrower before a loan is granted. Borrowers default on loans when they fail to make payment on their loans as and when due, for any reason. This default on loans is called credit risk. Credit default risk or credit risk is a kind of financial risk that a money lender suffers when a borrower is unable or fails to make payments on a loan according to the agreed terms and conditions upon which the financial contract is made. It is also described as a measure of a borrower's ability to determine his credit worthiness (Deepanshu, 2019). On this note, a borrower can be classified as either good or bad. A borrower is considered to be good if he repaid or has always repaid correctly his loan and has never been late in paying for thirty (30) days or more while a bad borrower is one who has experienced at least once a delay in the repayment of his loan for 30 days or more (Mohamad & Boujelbene, 2017). Hence, a lender can be a financial institution such as bank, a non – financial institution such as community (microfinance) bank or an individual while a borrower can be an individual, government or a business organization. Although every financial institution's lending process is guided by a structured credit approval process which includes a well- established procedure for comprehensive credit approval, they have continued to record bad debts as a result of the borrowers' inability or failure to pay back their loans within the agreed time frame. This has been a major concern in the financial sector worldwide particularly in Nigeria, and among other types of risks faced by banks, credit risk is rated to be on the lead. Ramazan and Gulden (2019) described credit risk as the greatest risk that affects bank performance. The effects of credit default risk include sudden shutdown of businesses, insolvency, financial crisis, revocation of license by Central Bank of Nigeria, among others. In line with these effects, Afriyie and Akotey (2012) noted that effective credit risk management has become very important in the survival and growth of financial institutions. In order to minimize this type of risk, it is required that the lender carries out a proper credit risk analysis on the borrower before a loan is granted. Credit risk analysis is the procedure by which a lender will determine a potential borrower. This practice enables a financial organization or a lender to ascertain in advance the borrower's ability or inability to pay back a loan and the associated interest on the loan within a specified period of time and to make the right decision based on the outcome of the analysis. Credit simply means money made available by a creditor or lender to a borrower under certain terms and conditions. A borrower in the study can also be referred to as obligator or counterparty. This money provided by the lender is given out with the intention of receiving it back from the borrower within a specified period of time with interest. In this case, the borrower may or may not be able to pay back the credit extended to him either willingly or unwillingly within the agreed time. Now, this probability of the borrower not being able to repay his credit (loan) as and when due is referred to as risk of credit or simply put, credit risk. In a more specific manner, credit risk implies the probability of loss of loan within a financial contract.

Review of Related Works

Series of studies have been carried out on the effect of credit risk on banks and other financial institutions worldwide and the application of neural networks as an effective model in predicting credit risk. Ramazan and Gulden (2019) performed a research to examine the effect of credit risk on bank performance in Turkey. The data used were drawn from 26 commercial banks in Turkey between 2005 – 2017; Return on Asset (ROA) and Return on Equality (ROE) were used as financial indicators while Non-Performing loans (NPLs) served as credit risk indicator. The study revealed that credit risk is the fundamental variable of the bank's insolvency risk, and therefore concluded that credit risk reduces the financial performance of banks and therefore classified credit risk as a major problem for the Turkish banks. Yimka et al., (2015) examined the role of credit risk management in value creation process among commercial banks in Nigeria. They analyzed the impact of four (4) antecedents to credit risk thus; loan and advance loss provision, total loan and advances, non-performing loans and total asset on accounting Return on Equity (ROE) and Return on Asset (ROA) and used panel data of 10 commercial banks listed on Nigeria Stock Exchange (NSE) between 2006 and 2010. Based on the data analysis result, the study concluded that credit risk management has notable effect on financial performance of commercial banks. Taiwo et al. (2017) focused on investigating the impact of non-performing loans on the performance of Nigeria's Deposit Money Banks (DMBs) and Bank lending growth from 1998 – 2014. The research was carried out using secondary data acquired from the CBN statistical bulletin 2014 and World Bank (WDI) 2015 and utilized multiple linear regression model for analysis of time series data. The result unveiled that credit risk management has a trivial effect on the growth of total loan and advances by Nigerian Deposit Money Banks. In a bid to finding a more efficient and effective approach to the analysis of credit risk, many studies have been carried out to prove that artificial neural networks are one of the efficient modern non-linear statistical models for predictions. Bakpo and Kabari (2009) in their study proposed artificial neural networks for the development of credit risk evaluation software that can meet the rising needs and equipment of the loan managing system. Based on the data analysis, the research revealed that neural networks can be used to classify good and bad borrowers with the accuracy of 94.60%. Pacellli and Azzollini (2010) undertook a study to analyze the ability of an artificial neural network developed to forecast the credit risk of a panel of Italian Manufacturing Companies. The paper compared a feed-forward multi-layer neural network developed in their study to predict the credit risk of a panel of Italian Manufacturing Companies to another feed-forwarded neural network developed in another study in 2004 with a similar panel of companies; noting that the two neural networks are similar with a difference in the activation function used. The result obtained from the study showed that the neural networks represent an alternative to traditional method of classification because of their ability to adjust in complex conditions. The study further confirmed that artificial neural networks are particularly suited to analyze and interpret-revealing hidden relationships that govern the data. The study also recorded that the neural network models have some limitations as the risk of inability to exit from local Minima; they need a lot of examples to extract the prototype cases to be included in the training set and the lack of transparency in the identification of parameters most discriminatory. The paper therefore, concluded that the flexibility and objectivity of neural network model can provide strong support in combination with linear methods of analysis for the efficiency of the process of credit risk management of a bank. Ezenkwu et al. (2016) carried out a research on the development of a credit risk evaluation system using multi layer Feed forward neural networks. The datasets used comprised of 1000 records, 6 input columns and 2 outputs. One of neural networks developed in the study was built using 7 input neurons in the input layer, 15 hidden neurons in the hidden layer and 1 output neuron in the output layer for loan applicants' classification and the other network for credit line prediction was built with 8 neurons in the input layer, 15 neurons in the hidden layer and 1 neuron in the output layer. The study disclosed that the proposed neural networks used in the classification of loan applicants and for credit line prediction gave 100% correct classifications in both training and test set. Ghatge and Halkarnikar (2015) focused on determining neural network as a suitable modeling technique for predicting business firm loan satisfactory or not. The study stated that the artificial neural network approach will classify loan applicants as creditworthy or not and also predict the allowance amount more closely aligned with credit default expense acquired during the fiscal period than traditional management approaches would. Also, the study submitted that using back propagation neural network and expert evaluation in credit risk evaluation have the very good consistency. Khemakhem and Boujelbene (2015) aimed at exploring a new practical way based on the neural networks that would help the banker to predict the non-payment risk of companies asking for a loan. The researchers were motivated by the insufficiency of traditional prevision models. 86 Tunisian companies were used as sample and 15 financial ratios were calculated, from the period of 2005 – 2007. The comparison of the result of the two techniques (neural network and discrimination analysis) showed that a neural technique is more accurate in terms of predictability. Ahmad and Ashkan (2011) proposed a hybrid model of support vector machine, neural networks and decision tree which uses ensemble learning for credit granting and credit scoring. Ten classifiers agents were used as the members of ensemble model. These techniques as base classifiers were compared based on their accuracy in classification. The study used a real data set to test the model and classifiers. The result showed that the proposed hybrid ensemble model has better classification accuracy and performance when compared to other credit scoring methods, but among the three classifiers, the Support Vector Machine had the best performance and accuracy. Eriki and Udegbunam (2013) carried out a research to assess the quality of neural network in predicting distress as against the discriminant analysis and its applications in enhancing managers' decision. The study used forty-four films listed on the Nigeria Stock Market between 1987 and 2006. In the study, the performance of Neural Network and Discriminant Analysis statistical technique were compared and the performance of both was further compared with the performance obtained by mere guesswork, the result showed that both neural network and DA performed better that guesswork but neural network performed better than discriminant analysis technique. The study submitted that the performance of the neural network highlights its importance as an invaluable tool in the business decision making process and further suggested that neural network could help managers in decision-making to avert the present down trend of the Nigerian Stock Market. Teles et al. (2020) compared two approaches Bayesian network with artificial neural networks (multiplayer perception MLP) for predicting recovered value in a credit operation. The study utilized a total number of 1,890 records retrieved from a datasets collected from a financial institution. For the missing attributes in the datasets, the global mean was used to replace the missing attributes The Artificial Neural Network. The ANN was based on continuous raw data and coded in PHP Script while Bayesian Naïve input data were converted into Boolean before coded in PHP Script. The study submitted that the two approaches compared in the research work provided reliable outcomes but Artificial Neural Network is more efficient for predicting credit risk with an average score of 82% accuracy. Khashman (2010) worked on credit risk evaluation system that uses supervised neural network models based on the back-propagation learning algorithm. In this Study, three neural networks were trained and implemented to classify a credit application into approved or rejected. The neural networks were trained with real world credit application cases from the German credit approval datasets. The datasets has 1000 cases; where each case has 24 numerical attributes based on which credit application is accepted or rejected. The Study also used nine learning schemes with different training-to-validation ratios and the comparison between their implementation results was provided. The Study concluded that the Credit risk evaluation neural network model performed best when the LS_4 learning scheme was used. The LS4 learning scheme used 400 cases for training and 600 cases for validation i.e. 40%:60%) and suggested that the proposed neural system under LS_4 can efficiently be used in automatic processing of credit applications. Mohammadi and Zangeneh (2016) in their work focused on assessing customer credit risk using artificial neural networks. In the Study, carried a multilayer perceptron neural network model was trained with different Back-propagation (BP) algorithms considered in designing an evaluation model i.e. Levenberg – Marquardt (LM), Gradient Descent, Conjugate gradient, Resilient, BFGS Quasi-newton, and One-step secant of all these algorithms, the results revealed that the LM algorithm converges faster to the network and achieves better performance than the other algorithms. However, in comparing classification performance, the Study recorded that neural network performed better than some classification algorithms such as Logistic Regression and Decision Tree in assessing credit risk of customers.

Methodology Adopted

This Study adopted Agile methodology for its software development. Agile is an iterative and incremental approach to software development that emphasized on the quick delivery of an application. Agile involves breaking the project into more accessible, more feasible units, or iterations, leading to greater quality output across all teams, including improved and elaborate development, collaboration, and testing.

Analysis of the Existing System.

Naïve Bayesian (NB) has been used in modeling credit default risk. Naïve Bayesian (NB) Classifier is a probabilistic classifier, a machine learning model that is used for classification task. This approach consists of supervised statistical classifier on the assumption of conditional independence where probability models are estimated using labeled data i.e., such instance is assigned to a class (Bradley, 1997). This classifier is built on Bayes' Theorem (Siddharth & Hoa, 2020; Gandhi, 2018).

Generally, Bayesian theorem proposes two types of probabilities, namely;

i. Posterior probability of H conditioned on
X: [P(H/X]; and
ii. Prior probability of H regardless of any observation or condition or information

Where

X is data tuple (evidence) and H is the Hypothesis.

Hence, the probability that the hypothesis H holds so long as the evidence or observed data tuple X is given as:

P(H/X) = P(X/H) P(H)/ P(X) (Okesola & Adewale, 2017).

Furthermore, Naïve Bayes (NB) classifier uses a set of training data to estimate the probability terms required for classification (Diad & Hindi, 2017). This performance is measured by the accuracy of estimated required probability terms, which is always a challenge as the training data is not easily available (Wu et al., 2008). A Naïve Bayes classifier presumes that the presence or absence of a particular feature of the class does not have any relation with the presence or absence of any other features given in the same class variable (Neeraj et al., 2015). Siddharth and Hoa (2020) further explained that Naïve Bayes Classifier assumes that each feature is independent and does not interact with each other, such that each feature independently and equally contributes to a probability of a sample to belong to a specific class. Figure 1.0 shows a sample structure of a Naïve Bayes Classifier.

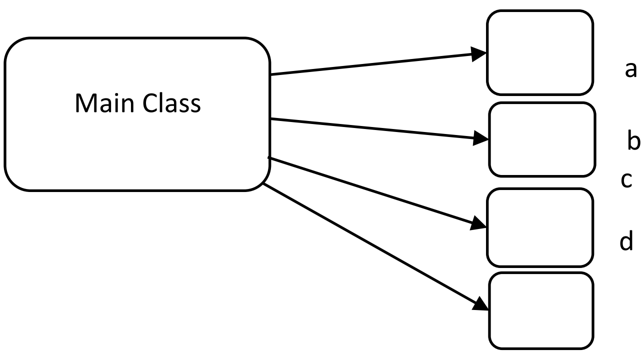

Figure 1: A Simple Structure of NB Classifier

In Figure 1.0 above, the class node is the parent node of all the nodes. One node represents the main class and others such as a, b, c, and d represent other attribute nodes of a particular sample. Also, a Naive-Bayes Classifier has been used as an effective classifier for many years, and has two advantages compared to other classifiers. It is easy to develop, requires no structure learning procedure. Secondly, the classification process is very efficient. Both advantages are due to its assumption that all the features are independent of each other. (Cheng, & Greiner,2013). Naïve Bayes are useful for graphically representing probabilistic relationships between variables. They are important in data modeling especially when data are missing because variables are characterized based on the combination of graphical models and statistical approaches (Hamadi & Aida, 2011). They also have the ability to identify the comparable relationship between variables and help to predict the tendency using probability distribution functions regardless of the nature of the data. They are supervised machine learning algorithms that can learn and represent probabilistic knowledge (John & Langley, 1995). However, the data flow diagram below illustrates the process operation of the existing system. The procedures of applying credit default down to credit approval stage from the bank credit manager. This actually the manual method of bank credit processes, indeed, every financial institutions goes through the below processes before credit can be granted to the customer.

Data Flow Diagram of the Existing System.

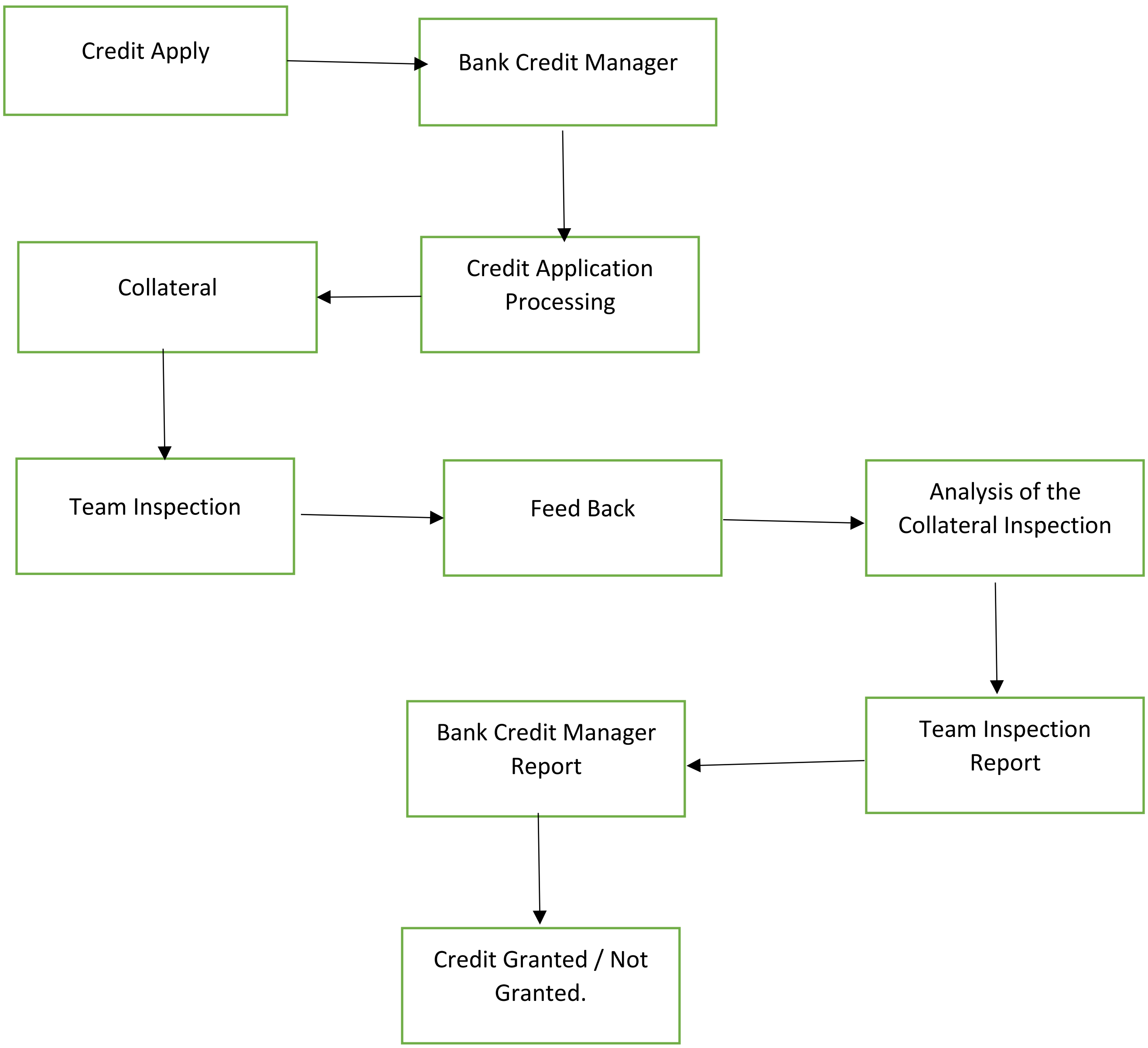

Figure 2: Data Flow Diagram of the Existing System

Analysis of the Proposed System

The below diagram is the data flow diagram of the proposed new system, which shows how the new system will function when implemented. It is typically known for its ability to software complex relationship among data. Unlike the recurrent type, information or data moves in one direction to another for credit default processing. This system is comprises of major processes during the time of credit application from the bank customers; the diagram depict the necessary steps involved before the credit application will be granted to the customers. The function of the neurons in the input system is to receive incoming data requirements, and send the requirements (input) forward to the bank credit manager using a specific function known as the prediction analysis function. The neurons in the input system represent the number of predictor variables and the number of input variables determines the number of neurons to be used in the input system which the diagram illustrated. Moreover, the new system is based on extended gradient -descent based on Delta learning rule which is popularly known as Back propagation rule. The analysis process of customer credit application based on the new system flow which shown on Figure 2 Data Flow Diagram of the Existing System.

Data Flow Diagram for the New System

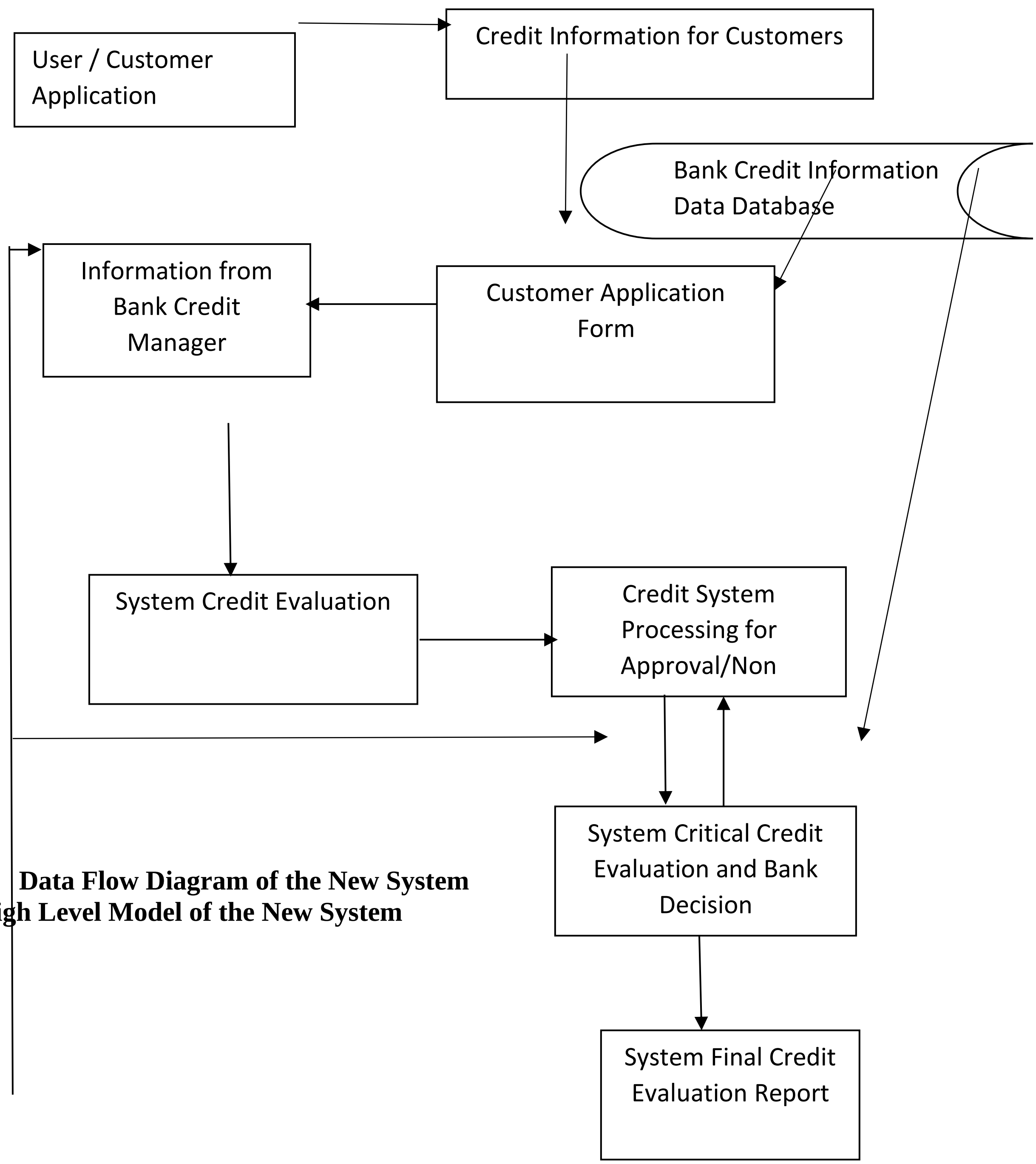

Figure 3: Data Flow Diagram of the New System High Level Model of the New System

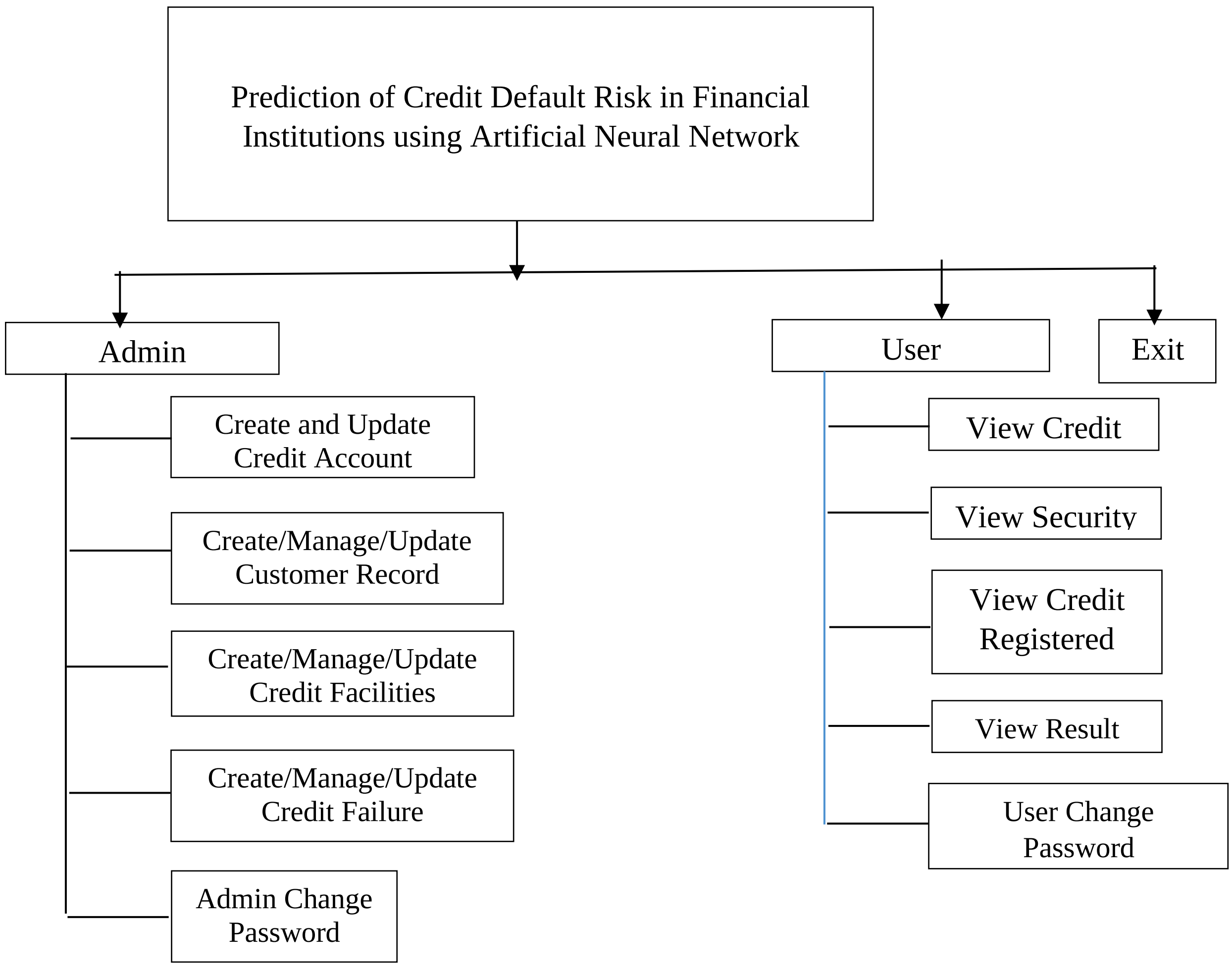

Figure 4: The High Level Model of the new system

CONCLUSION

Neural networks have been found to be powerful tools in predicting credit default risk and problem- solving tools in the modern-day computing. Their ability to learn by example makes them more flexible and useful in a wide range of applications in diverse fields. In this Study, a feed – forward neural network model was developed and trained using PHP script software. Agile methodology was employed in the software development. A total number of 125 valid training samples were retrieved from a credit default risk datasets and were divided into 3 subsets by the PHP software with different percentages i.e., 70% of the datasets for training, 15% for validation, and 15% for testing the network. The model's architecture was built using 10 neurons in the input layer which represent 10 input variables sufficient to predict credit default risk of borrowers, 10 neurons in the hidden layer and 2 output neurons, representing the output variable (dependent variable). The network was trained with Levenberg-Marquardt algorithm backpropagation and sigmoid activation function. The performance of the New System was evaluated using Mean Square Error (MSE) and Regression analysis. The New System gave a mean square error of

0.12933, total target regression value of 0.91523 which shows a strong correlation among the training, validation, and testing samples. Comparing the new system's output with the target output gave an accuracy of 96%. Based on this result, this Study therefore presents neural networks as effective, efficient, accurate and reliable tools which can be employed by financial firms and lender organizations to solve and manage the problem of credit default risk by predicting credit worthiness of borrowers before loans are granted.

References

Afriyie, H.O & Akotey, J.O (2012). Credit risk management and profitability of selected rural banks in Ghana. *African Development Workshop*.https://tinyurl.com/j8b622ec

Ahmad, G., & Ashkan, A. (2011). Application of artificial intelligence techniques for credit risk evaluation. International Journal of Modeling and Optimization, 1(3), 243-249.

Bakpo, F. S., & Kabari, L. G. (2009). Credit risk evaluation system: An Artificial neural. network approach. Nigerian Journal of Technology, *28* (1). 36 – 44. https://tinyurl.com/2p87etua

Bradley, A.P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognition, 30(7), 1145–1159.

Cheng, J. & Greiner, R. (2013). Comparing Bayesian Network Classifiers. Retrieved on 5th August, 2024 from https://arxiv.org/pdf/1301.6684

Deepanshu, B. (2019), A complete guide to Credit risk modelling. Retrieved on 6th August, 2024 from https://www.listendata.com/2019/08/credit-risk-modelling.html.

Diab, D. M., & El Hindi, K. M. (2017). Using differential evolution for fine tuning naïve Bayesian classifiers and its application for text classification. Applied Soft Computing Journal, 54, 183–199. https://doi.org/10.1016/j.asoc.2016.12.043

Eriki, P.O & Udegbunam, R. (2013, October 14). Predicting corporate distress in the Nigerian stock market: Neural network versus multiple discriminant analysis. *African Journal of Business Management*, 7(38), 3856-3863. https://doi.org/ 10.5897/AJBM09.152

Ezenkwu, C.P., Udofia K. M. &. Sihitshuwam, F.D. (2016). Development of a credit risk evaluation system using multilayer feedforward neural networks. *African Journal of Computing & ICT*,8(1). 165 174. https://www.researchgate.net/publication/311650187_Development_of_a_Credit_Risk_Evaluation_System_Using_Multilayer_Feedforward_Neural_Networks

Gandhi, R. (2018). Naive Bayes Classifier. Towards Data Science Blog. https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

Ghatge, A.R., & Halkarnikar. (2015) Estimation of Credit Risk for Business Firms of Nationalized Bank by Neural Network Approach. International Journal of Electronics and Computer Science Engineering, 2(3). 828 – 834. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.434.28&rep=rep1&type=pdf

Hamadi, M. & Aida, A. (2011). Credit-risk evaluation of a Tunisian commercial bank: Logistic regression vs neural network modelling. *International Journal of Accounting & Information Management*, *19*(2), 92 – 119. https://doi.org/ijaim.2011.36619baa.005

John, G.H., & Langley, P. (1995). Estimating continuous distributions in Bayesian classifiers.*Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*,338–345. http://citeseerx.ist.psu.edu/ viewdoc/versions?doi=10.1.1.8.3257.

Khemakhem, S., & Boujelbene, Y. (2017). Artificial Intelligence for credit risk assessment: Artificial Neural Network and Support Vector Machines. *ACRN Oxford Journal of Finance and Risk Perspectives*, 6(2), 1-17. https://tinyurl.com/ycxsbehr

Mohamed, W. T., & Boujelbene, Y. (2017). Bank credit risk: Evidence from Tunisia using Bayesian networks. *Journal of Accounting, Finance and Auditing Studies*, 3/3 (2017) 93-107. https://tinyurl.com/5tdv2nft

Mohammadi, N. & Zangeneh, M. (2016). Customer credit risk assessment using artificial neural networks. *International Journal of Information Technology and Computer Science*, 8(3), 58-66. https://doi.org/10.5815/ijitcs.2016.03.07.

Okesola, O. J., Okokpujie, K. O., Adewale, A. A., John, S. N., & Omoruyi, O. (2017). An improved bank credit scoring model: A naïve bayesian approach. *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*, 228-233.

Pacelli, V. & Azzollini, M. (2011). An artificial neural network approach for credit risk management. *Journal of Intelligent Learning Systems and Applications*, *3*(02), 103-112. http://doi.org/10.4236/jilsa.2011.32012

Ramazan, E., & Gulden, P. (2019). The effect of credit risk on financial performance of deposit banks in Turkey. *Procedia Computer Science*, 158 (2019), 979–987. http://doi.org/10.1016/j.procs.2019.09.139

Siddharth, M.,Hao, L., & Jiabo, H. (2020),*Machine Learning for Subsurface Characterization.* Gulf Professional Publishing. https://doi.org/10.1016/C2018-0-01926-X

Taiwo, J.N., Ucheaga, E.G., Achugamonu, B.U., Adetiloye, K., Okoye, L. &, Agwu, M.E. (2017). Credit risk management: Implications on bank performance and lending growth. *Saudi Journal of Business and Management Studies*,*2*(5B), 584 – 590. https://tinyurl.com/ye23ujxj

Teles, G. Rodrigues, J., Rabelo, R. A. L., & Kozlov, S. (2020). Artificial neural network and Bayesian network models for credit risk prediction. *Journal of Artificial Intelligence and Systems*, 2. 118-132. https://doi.org/10.33969/AIS.2020.21008.

Wu, X., Kumar, V., Quinlan, J.R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, Ng, A., Liu, B., Yu, P.S., Zhou,Z., Steinbach, M., Hand, D.J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*(Vol.14),1-37.https://doi.org/10.1007/s10115-007-0114-2

Yimka, A. T., Abimbola, C., & Adekunle, O. (2015, Feb 2015). Credit risk management and financial performance of selected commercial banks in Nigeria. *Journal of Economic & Financial Studies*, 3(01), 01- 09. https://doi.org/10.18533/jefs.v3i01.73