*American Journal of Applied Mathematics and Statistics, 2018, Vol. 6, No. 6, 266-271*
Available online at http://pubs.sciepub.com/ajams/6/6/8

DOI:10.12691/ajams-6-6-8

# Comparison of Accuracy of Support Vector Machine Model and Logistic Regression Model in Predicting Individual Loan Defaults

**Obare DM*, Muraya MM**

Physical Sciences Department, P.O.Box 109, Chuka University, Chuka, Nairobi, Kenya
*Corresponding author: obaredominic87@gmail.com

*Received October 19, 2018; Revised November 29, 2018; Accepted December 14, 2018*

**Abstract** Prediction of loan defaults is critical to financial institutions in order to minimize losses from loan non-payments. Some of the models that have been used to predict loan default include logistic regression models, linear discriminant analysis models and extreme value theory models. These models are parametric in nature thus they assume that the response being investigated takes a particular functional form. However, there is a possibility that the functional form used to estimate the response is very different from the actual functional form of the response. In such a case, the resulting model will be inaccurate. Support vector machine is non-parametric and does not take any prior assumption of the functional form of the data. The purpose of this study was to compare prediction of individual loan defaults in Kenya using support vector machine and logistic regression models. The data was obtained from equity bank for the period between 2006 and 2016. A sample of 1000 loan applicants whose loans had been approved was used. The variables considered were credit history, purpose of the loan, loan amount, saving account status, employment status, gender, age, security and area of residence. The data was split into training and test data. The train data was used to train the logistic regression and support vector machine models. The study fitted logistic regression and support vector machine models. Logistic regression model showed an accuracy of 0.7727 with the train data and 0.7333 with test data. The logistic regression model showed precision of 0.8440 and 0.8244 with the train and test data. The SVM (linear kernel) model showed an accuracy of 0.8829 and 0.8612 with the train and test respectively. The SVM (linear kernel) showed a precision of 0.8785 with the train data and 0.7831 with the test data. The results showed that support vector machine model performed better than logistic regression model. The study recommended the use of support vector machines in loan default prediction in financial institutions.

***Keywords:*** *loan defaults, prediction model, logistic regression model, support vector machine model*

**Cite This Article:** Obare DM, and Muraya MM, "Comparison of Accuracy of Support Vector Machine Model and Logistic Regression Model in Predicting Individual Loan Defaults." *American Journal of Applied Mathematics and Statistics*, vol. 6, no. 6 (2018): 266-271. doi: 10.12691/ajams-6-6-8.

## 1. Introduction

Loan defaults occurs when the borrowers are not able and, or not willing to repay loans [1]. When loan defaults occur they bring about economic strain since quality borrowers are denied access to credit which they can use to develop the economy. Loan defaults also make financial institutions to incur losses since they lose both the capital and the interest. There has been a growing concern about the constant increase on loans performance in commercial banks in Kenya [2]. This is because commercial banks have been offering loans to customers as long as they can prove their ability to pay. It is not easy to accurately prove this ability and this would later lead to defaults. However, this can be achieved through modelling the probability of loan default.

Credit risk decisions are key determinants for the success of financial institutions because of huge losses that result from wrong decisions [3]. Hence, credit risk evaluation is essential before making any lending decision [4]. Due to the significance of credit risk a number of statistical models have been proposed to predict loan defaults. These models include, Artificial Neural Networks, genetic algorithms, genetic programming, and some hybrid models have been used to evaluate credit risk with promising results in terms of performance accuracy. These models have several drawbacks: (1) lack of explanatory power; (2) reliance on the restrictive assumptions of statistical techniques; and (3) numerous variables, which result in multiple dimensions and complex data [5].

Survival analysis models have also been proposed for credit risk modelling [6]. Stepanova and Thomas [7] and Tong *et al.* [8], showed that the survival analysis models have high performance compared to logistic regression in

terms of precision. Survival analysis models are flexible since they employ parametric and semi-parametric models depending on the choice of the researcher and the underlying nature of data, they also don't make distributional assumption as to the appropriateness of the response variable, lifetime of bank loans in this case.

Machine learning algorithms have also been suggested to predict loan defaults [9]. Galindo & Tamayo [10] tested decision tree algorithms on mortgage-loan data to detect defaults, and also they compared their results to the K-nearest neighbor and probit models. They found that Neural Networks provided the best results with the smallest error, followed by K-Nearest Neighbor algorithm, and probit models performed dismally since it predicted with the largest error. Butaru *et al.* [11] investigated consumer's delinquency using decision trees, logistic regression and random forest with data from six banks. They found out that in terms of prediction random forest and decision tree models perform about the same in terms of accuracy, both consistently outperforming the logistic regression model.

In Kenya loan defaults are on the rise and this is a critical source of economic strain. For this reason, these defaults must be controlled and monitored [12]. The best method would be through modelling loan defaults. In Kenya, several predictive models have been used to predict loan defaults [2]. These models include; linear discriminant analysis, logistic regression models and generalized extreme value regression models. These models are parametric since they assume the response being investigated takes a particular functional form. Logistic regression model has been used to analyze default risk. Martin *et al*., [13] applied logit model as the basis for developing financial ratios and probabilistic prediction of bankruptcy. The results showed that coefficient estimates for this model were efficient in the use of relatively small samples because it overcomes problems arising from linear regression [14]. However, this model is suitable only for qualitative research and the model's effectiveness also depends on the assumption that irrelevant alternatives are independent [15]. Due to this assumption, there is a possibility that the functional form used to estimate the response is very different from the actual form of the response. In such a case, the resulting model will not fit the data well and the estimates from the model will also be poor.

Support vector machine model had not been used to model individual loan defaults in Kenya. It is a new novice algorithm that should be embraced in analyzing data that does not have any prior functional form Zhou *et al.* [16]. This study seeks to investigate if the SVM will produce more accurate results in predicting individual loan defaults compared to Logistic regression model. Support Vector Machine is able to fit complex feature spaces when compared with some of the traditional learning algorithms without the addition of high power features [17]. It is non-parametric method in that it avoids the assumption of a particular functional form of the response. For this reason, it has the potential to accurately fit a wider range of the possible shapes of the response [18]. It is a very flexible model and it can fit many different functional forms of the response. It seeks to estimate the response that gets as close to as possible to the data points without being too rough or wiggly [19].

This study used R-Statistical software [20] to analyze secondary data obtained from Equity bank for a period between 2006-2016. Probabilities of loan defaults were determined by using logistic regression model. Support vector machine model was fitted by machine learning technique. Logistic and SVM model were compared by prediction accuracies and F1 scores.

## 2. Methodology

This study was carried out at Equity bank headquarters in Nairobi using a secondary data for the loan applicants whose loans were approved for the period between 2006-2016. The data was obtained in form of an excel sheet from Equity bank of Kenya headquarters. Over ten thousand (10000) client's information was provided. A mixed method research design was used, it adopts both quantitative and qualitative approaches in a single study [21]. Thirty percent (30%) of the data collected from equity bank of Kenya was used as the sample size [22]. Defaulted loans for individuals were stratified based on number of days past due date for the monthly loan repayments. The independent variables in the data were duration of the loan, the credit history of the applicant, the purpose of the loan, the loan amount, the nature of the saving account, the employment status, the gender of the applicant, the age of the applicant, the security used when acquiring the loan and the area of the applicant The dependent variable was the nature of the loan which was classified as performing or non- performing (loan default). Data analysis was done using logistic regression model and support vector machine in R statistical software [20]. The data was coded for easy analysis using the R- Statistical software. Non-performing loan was coded 1 and a performing loan 0. Equivalent number of dummy variables were created for the purposes of coding and comparing independent variables. In fitting the models by machine learning, the data set was divided into a training and testing set. The training set had a sample of 700 applicants. The machine was trained to divide the sample into seven sub samples. That is, a sample of 100, 200, 300, 400, 500, 600 and 700. Both the Support Vector Machine and Logistic regression models were fitted using each subsample and tests the behavior of the model obtained against the test data in each case by use bias-variance curves. The fitted models were compared using prediction accuracies and F1 score to determine the best model.

## 3. Results and Discussion

### 3.1. Logistic Regression Model

Logistic parameters were interpreted using the odds ratio, all the other covariates were kept constant, it can be deduced that an individual with a current account operating between ksh 0 to ksh 50,000 is 100[exp (-1.32029) -1] = 73.29% less likely to default a loan. For a month increase in the duration of repayment, an individual is 100[1-exp (-0.0291)] = 2.9% less likely to default a loan. A person who has all his credits in the bank paid fully is 100[1-exp (-1.05245)] =65.09% less likely to default a

loan. An individual who borrows a loan to purchase furniture is 100[exp (0.151374) – 1) = 16.3% more likely to default a loan. When the amount borrowed increases by ksh 1,000, the chances of defaulting a loan decreases by 100[1-exp (-0.00014)] = 1.4E-4%. An individual who has been employed for 1- 4 years is 100[1-exp (-0.14987)] = 13.9% less likely to default a loan. When the age of an applicant increases by 1 year, the chances of defaulting a loan increases by 100[exp (0.012752) – 1] = 13.6%. This study revealed that the factors were statistically significant in the prediction of loan default repayment. This is in agreement with the study by Edinam & Agbemava, [14]. The results of study revealed that six factors, i.e., marital status, dependents, type of collateral security, duration and loan type were statistically significant in the prediction of loan default payment with a predicted default rate of 86.67%. This agrees with a study of Ameyaw-Amankwah, [23], which was carried out in Ghana on the effects of client's social and economic factors in relation to likelihood of defaulting a loan. It revealed that a person's gender age and economic status are very significant in assessing the creditworthiness of an individual.

The logistic regression model developed is

$$\begin{aligned} y_i = {} & 4.854957 - 1.79749X_1 - 1.32029X_2 \\ & - 0.69829X_3 - 0.0291X_4 - 1.58278X_5 \\ & - 1.05245X_6 - 0.7383X_7 - 0.93359X_8 \\ & - 0.38994X_9 + 1.03945X_{10} + 0.151374X_{11} \\ & + 0.626872X_{12} - 0.1698X_{13} - 0.38755X_{14} \\ & + 0.337959X_{15} + 0.768652X_{16} - 0.00014X_{17} \\ & - 0.14164X_{18} - 0.45998X_{19} - 0.14987X_{20} \\ & + 0.661989X_{21} + 0.459736X_{22} + 0.245808X_{23} \\ & + 0.38296X_{24} + 0.05287X_{25} + 0.012752X_{26} \\ & + 0.0532008X_{27} + \varepsilon_i \end{aligned} \quad (1)$$

Where $X_i$ = Variables of interest. $i = 1, 2\ldots\ldots$ and $\varepsilon_i$ is the error term.

**Table 1. Logistic Regression Model Summary**

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| (Intercept) | 4.854957 | 1.485136 | 3.269031 | 0.001079 |
| Current account < 0 | -1.79749 | 0.274589 | -6.54612 | 5.90E-11 |
| Current account 0- 50000` | -1.32029 | 0.272092 | -4.85236 | 1.22E-06 |
| Current account > 50000` | -0.69829 | 0.460049 | -1.51786 | 0.129049 |
| Duration | -0.0291 | 0.010556 | -2.75706 | 0.005832 |
| No credits taken | -1.58278 | 0.517312 | -3.05962 | 0.002216 |
| All credits at this bank paid duly | -1.05245 | 0.513208 | -2.05074 | 0.040293 |
| Existing credits paid duly until now | -0.7383 | 0.305429 | -2.41725 | 0.015638 |
| Delay in paying in the past | -0.93359 | 0.391398 | -2.38527 | 0.017067 |
| Purpose new car | -0.38994 | 0.569847 | -0.68428 | 0.493795 |
| Purpose used car | 1.039457 | 0.66132 | 1.571792 | 0.115999 |
| Purpose furniture | 0.151374 | 0.586011 | 0.258313 | 0.796166 |
| Purpose radio/tv` | 0.626872 | 0.577064 | 1.086313 | 0.277341 |
| Purpose domestic appliances` | -0.1698 | 1.05148 | -0.16149 | 0.871711 |
| Purpose education | -0.38755 | 0.701585 | -0.55239 | 0.58068 |
| Purpose business | 0.337959 | 0.621207 | 0.544036 | 0.586417 |
| Purpose others | 0.768652 | 1.055595 | 0.728169 | 0.46651 |
| Amount | -0.00014 | 4.99E-05 | -2.75557 | 0.005859 |
| Unemployed | -0.14164 | 0.432815 | -0.32725 | 0.743482 |
| Employed < 1` | -0.45998 | 0.343289 | -1.33993 | 0.18027 |
| Employed 1 - 4` | -0.14987 | 0.297289 | -0.50411 | 0.614185 |
| Employed 4 - 7` | 0.661989 | 0.363535 | 1.820974 | 0.068611 |
| Sex | 0.459736 | 0.224136 | 2.051152 | 0.040252 |
| Property real estate/farm` | 0.245808 | 0.358482 | 0.685691 | 0.492908 |
| Property savings/insurance` | 0.38296 | 0.351547 | 1.089358 | 0.275996 |
| Property car | 0.05287 | 0.314474 | 0.168121 | 0.866488 |
| Age | 0.012752 | 0.010587 | 1.204483 | 0.228403 |
| Area Residence | 0.532008 | 0.70059 | 0.75937 | 0.447631 |

**Table 2. Accuracy Table for Logistic Regression Model**

| | Train | Test |
|---|---|---|
| Accuracy | 0.7727 | 0.7333 |
| Sensitivity | 0.8145 | 0.7934 |
| Specificity | 0.6854 | 0.5862 |
| Positive Predicted Value | 0.8440 | 0.8244 |
| Negative Predicted Value | 0.6387 | 0.5368 |
| Prevalence | 0.6764 | 0.7100 |
| Detection Rate | 0.5509 | 0.5633 |
| Detection prevalence | 0.6527 | 0.6833 |
| Balanced Accuracy | 0.7500 | 0.6898 |

The performance of the model with both the train and test data was shown using an accuracy table (Table 2). The logistic regression model had an accuracy of 0.7727 with the train data and 0.7333 with the test data. The sensitivity with the train and the test data was 0.8145 and 0.7934, respectively. The precision values of the model were 0.8440 and 0.8244 with the train and test data, respectively. These values showed the percentages of defaults that were correctly predicted by the model on the train and test data.

The performance of the logistic regression model with both the train and the test data was shown using a bias variance curve. (Figure 1).

## 3.2. Fitted Support Vector Machine Model Radial Kernel

The SVM was fitted using the R-Statistical software as a machine learning algorithm. The machine was trained to assume that the separation between the defaulters and the non-defaulters was nonlinear (radial kernel). The performance of the model with the train and test data showed that the sample size needed to be increased to arrive at a better model (Figure 2).

The radial kernel model fitted had an accuracy of 0.7814 with the train data and 0.7800 with the test data (Table 3). The sensitivity values of the model were 0.8932 and 0.8873 with the train data and test data, respectively. The precision values also referred to as the positively predicted values were 0.8116 and 0.8082 with the train data and test data, respectively.

## 3.3. Fitted Support Vector Machine Linear Kernel

The machine was trained to assume a linear boundary between the loan defaults and the non-defaults [24]. Increasing the sample size makes the process to achieve the best model. The best model was achieved after a sample of 400 data sample (Figure 3).

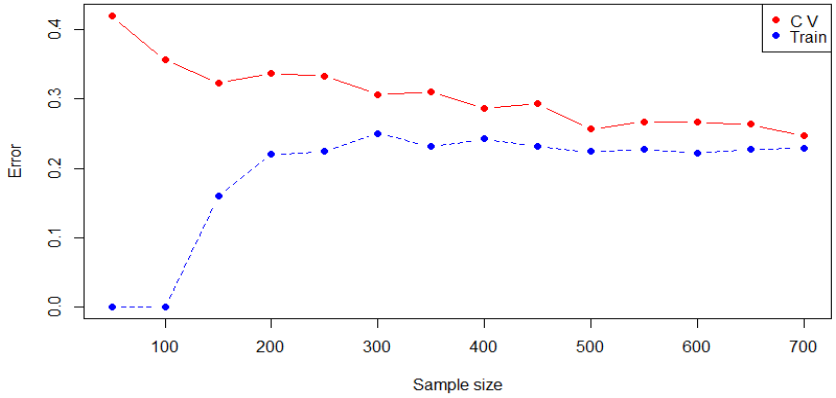


**Figure 1.** Train Errors Vs Test Errors Plot for Logistic Model

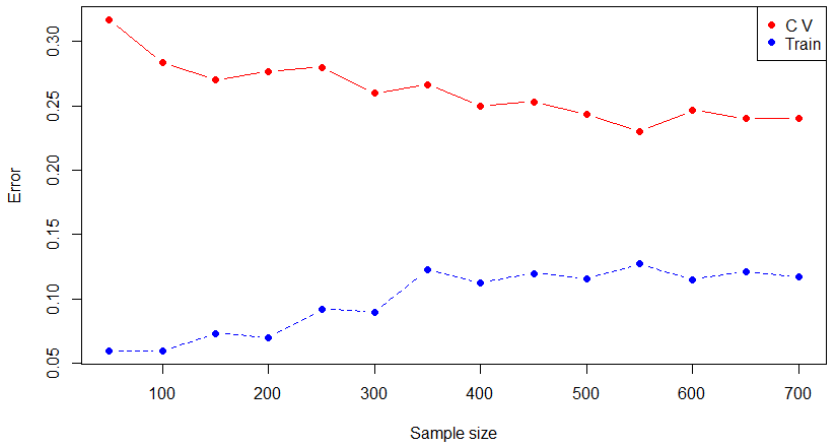


**Figure 2.** Train errors Vs Test Errors Plot for SVM Model (Radial Kernel)

**Table 3. Accuracy Table for SVM Radial Kernel Model**

| | | |
|---|---|---|
| Accuracy | 0.7814 | 0.7800 |
| Sensitivity | 0.8932 | 0.8873 |
| Specificity | 0.5258 | 0.5172 |
| Positive Predicted Value | 0.8116 | 0.8082 |
| Negative Predicted Value | 0.6829 | 0.6522 |
| Prevalence | 0.6957 | 0.7100 |
| Detection Rate | 0.6214 | 0.6300 |
| Detection prevalence | 0.7657 | 0.7700 |
| Balanced Accuracy | 0.7095 | 0.7023 |

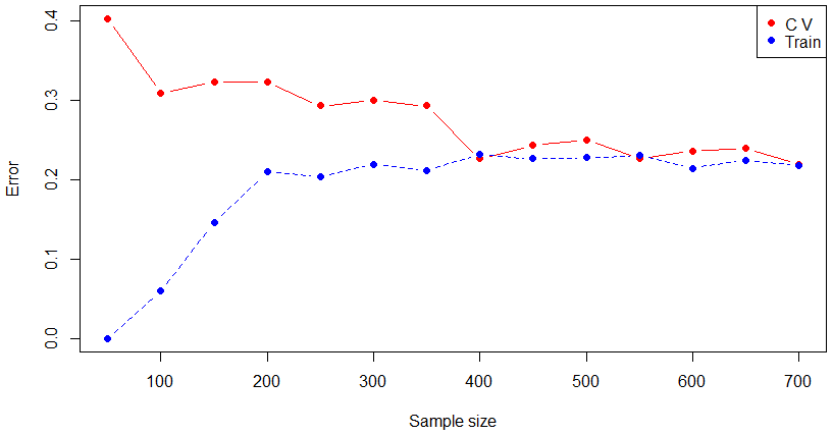


**Figure 3.** Train Errors Vs Test Errors Plot for SVM Model Linear Kernel

**Table 4. Accuracy Table for SVM Linear Kernel Model**

| | Train | Test |
|---|---|---|
| Accuracy | 0.8829 | 0.8612 |
| Sensitivity | 0.9651 | 0.9455 |
| Specificity | 0.6948 | 0.3793 |
| Positive Predicted Value | 0.8785 | 0.7831 |
| Negative Predicted Value | 0.8970 | 0.6471 |
| Prevalence | 0.6957 | 0.7100 |
| Detection Rate | 0.6714 | 0.6500 |
| Detection prevalence | 0.7643 | 0.8300 |
| Balanced Accuracy | 0.8300 | 0.6474 |

**Table 5. Comparison table using train data**

| Model | Accuracy | Precision (Positive predicted value) | Recall (Sensitive) | F1 Score |
|---|---|---|---|---|
| Logistic Regression Model | 0.7727 | 0.8440 | 0.8145 | 0.8290 |
| SVM Radial Kernel | 0.7814 | 0.8116 | 0.8932 | 0.8504 |
| SVM Linear Kernel | 0.8829 | 0.8785 | 0.9651 | 0.9198 |

**Table 6. Comparison table using test data**

| Model | Accuracy | Precision(Positive predicted value) | Recall (Sensitive) | F1 Score |
|---|---|---|---|---|
| Logistic Regression Model | 0.7333 | 0.8244 | 0.7934 | 0.8086 |
| SVM Radial Kernel | 0.7800 | 0.8082 | 0.8873 | 0.8459 |
| SVM Linear Kernel | 0.8612 | 0.7831 | 0.9455 | 0.8567 |

The SVM linear kernel showed an accuracy of 0.8829 and 0.8612 with the train data and test data respectively (Table 4). The recall values of the train data and the test data were 0.9651 and 0.9455 respectively. The model also showed a precision of 0.8785 with the train data and 0.7831 with the test data.

### 3.4. Comparison of the Fitted Models

Comparison of the models was done using the accuracy values and the F1 score values. The F1 score values takes into account the precision and recall values. The general observation was that the support vector machine models performed better than the logistic regression model (Table 5 & Table 6). This agrees with Sebe & Rzvan [17], they showed that SVM performed better than logistic regression in predicting which companies will default on their loans.

The best model in terms of predictive accuracy and F1 score was the SVM linear kernel. This model was followed by the SVM radial kernel and then logistic regression model.

## 4. Conclusion

In order to achieve the objectives of the study, the knowledge of machine learning was utilized and implemented for analysis of the data. The data was obtained from equity bank of Kenya between 2006 - 2016. The data was cleaned and missing values removed through seeding in R., then coded according to the variables for easy analysis. The logistic regression model and support vector machine model were fitted using R-statistical software. First, loan defaults were predicted by using logistic regression model. During this analysis, the data was split into two, train data set and test data set then the probabilities of loan defaults from the train data were developed (Table 1). This helped to tell if an individual is likely to default when compared to the Z-score in relation

to the variables. This was followed by fitting the logistic regression model that can be used to predict individual loan defaults. The logistic regression model was first trained using the train data set, this led to development of the model (equation 1).

Secondly, the data was also fitted using support vector machine model, implemented by machine learning. First the kernels were selected and then applied concurrently using the train data. By training the SVM lead to fitting of two SVM models. The support vector machine model radial kernel was fitted first then followed by the support vector machine linear kernel. The two models were applied again to analyze the test data. This was done in order to ascertain if the models can predict loan defaults accurately as per the train data.

The plots for train errors and test errors were also developed (Figure 1, Figure 2 and Figure 3). This was done in order to determine the effect of increasing the sample size in relation to errors. The three models developed were then compared by use of the model accuracies and the F1 score (Table 5 & Table 6). The train errors and test errors plots for the three models were also used for purposes of comparison.

The fitted models were then compared for their prediction accuracies. The results showed that the Support Vector Machine linear kernel model performs better than the Support Vector Machine radial kernel model and logistic regression model (Table 5 & Table 6). This indicate that the non-parametric behavior of the Support Vector Machine linear kernel model and Support Vector Machine radial kernel models enables them to fit the data better as compared to the parametric models, logistic regression model. The performance of the SVM Model purely depends on the choice of the Kernels. Therefore, the SVM linear kernel should be adopted in predicting loan defaults.

## Acknowledgements

The authors acknowledge Chuka university in particular the faculty of science, engineering and technology for allowing me to undertake this study. Furthermore, the authors thank Equity bank Limited of Kenya for providing the data that was used for this study. The authors are also grateful to all partners who contributed to the success of this study either morally or financially.

## Competing Interests

The authors declare that they have no competing interests.

## References

[1] Hoque, Z. (2005). Linking Environmental Uncertainty to Non-Financial Performance. *The British Accounting Review*. Britain.

[2] Evusa, Z., Mudaki, J. S., & Ojala, D. O. (2015). Evaluation of the Factors Leading to Loan Default at Equity Bank, Kenya. *Journal of Economics and Sustainability. ISSN 2224-607X. pp. Vol.6, No.8.2016.*

[3] Lahsana, A., Anion, R. & Wah, T. (2010). "Credit Scoring Models using soft Computing Methods: a survey". *International Arab Journal of Information Technology*, 7(2), 115-123.

[4] Bekhet, H. & Eletter, S. (2014). "Credit Risk Management for the Jordanian Commercial Banks: Neural Scoring Approach". *Review of Development Finance*, 4, 20-28.

[5] Chen, Y., Cheng, C. (2013). "Hybrid Models based on Rough Set Classifiers for Setting Credit Rating Decision Rules in the Global Banking Industry". Knowledge- Based Systems, 39(1), 224-239.

[6] Banasik, J., Crook, J. N. and Thomas, L.C. (1999). Not If but When Will Borrowers Default. *Journal of the Operational Research Society, 50(12) pp. 1185-1190.*

[7] Stepanova, M., & Thomas, L. (2002). Survival Analysis Methods for Personal Loan Data. *Operations Research*, 50(2), pp.277-289.

[8] Tong, E. N., Mues C. & Thomas, L. (2012). Mixture Cure Models in Credit Scoring: If and When Borrowers Default. *European Journal of Operational Research, 218(1) pp. 132-139.*

[9] Khandani, A.E., Kim, A.J. & Andrew W. Lo. (2010). Consumer Credit-Risk Models via Machine-Learning Algorithms. *Journal of Banking Finance 34:2767-87*.

[10] Galinndo, J., & Pablo T. (2000). Credit Risk Assessment using Statistical and Machine Learning: Basic Methodology and Risk and Risk Modelling Applications. *Computational Economics 15: 107-43.*

[11] Butaru F., Qingqing C., Brian C., Sanmay D., Andrew W. Lo. & Akhtar S. (2016). Risk and risk Management in the Credit Card Industry. *Journal of Banking and Finance 72:218-39.*

[12] Divino, J. A., Lima, E. S., & Orrillo, J. (2013). Interest Rates and Default in Unsecured Loan Markets. *Quantitative Finance, 13*(12), 1925-1934.

[13] Martin, A., Travis L. & Venkatasamy P. (2010). A Framework to Develop Qualitative Bankruptcy Prediction Rules. *St. Joseph's Journal of Humanities and Science* 1:73-83.

[14] Agbemava, E., Nyarko, I. K., Adade, T. C., & Bediako, A. K. (2016). Logistic Regression Analysis of Predictors of Loan Defaults by Customers of Non- Traditional Banks in Ghana. *African Journal of Business Management* 10(2), 33-43.

[15] Calabrese, R. & Osmetti, S. A. (2013). Modelling Small and Medium Enterprise Loan Defaults as Rare Events: The Generalized Extreme Value Regression Model. *Journal of Applied Statistics*, *40*(6), 1172-1188.

[16] Zhou, L., Lai K.K., & Yu. L. (2010). *Least Squares Support Vector Machines Ensemble Models for Credit Scoring*. Expert Systems with Applications 37: 127-133.

[17] Sebe, V., Razvan, A. (2009). Estimating Probabilities of Default using Support Vector Machines. *A master Thesis Presented at centre of Applied Statistics and Economics*. Humbolt University, Berlin.

[18] Huang, C. L., Chen, M. C., & Wang, C. J. (2007). Credit Scoring with a Data Mining Approach based on Support Vector Machines. *Expert systems with Application.* 33 (2007). 847-856.

[19] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning (Vol. 112). New York: Springer.

[20] J. Pinheiro, D., Bates, S DebRoy, D., Sarkar, R., C Team R Package Version 3 (57), 1-89.

[21] Tashakkori, A., & Teddie, C. (2003), The Handbook of Mixed Methods in Social and Behavioural Research, Sage, Thousand Oaks, CA.

[22] Mugenda, A. & Mugenda, O. (1999). Research Methods-Quantitative and Qualitative Approaches, Nairobi. Act Press.

[23] Ameyaw-Amankwah,I. (2011). Causes and Effects of Loan Defaults on the Profitability of Okomfo Anokye Rural Bank. Master Thesis KNUST, Accra, Ghana.

[24] Muller K-R., Mika S., (2001). An Introduction to Kernel-based Learning Algorithms, IEEE Transactions on Neural Networks 12(2), 181-201.