HP
JASD 2019
Hummingbird Publications
Journal of African Sustainable Development
www.hummingpubng.com
© June, 2019
Vol. 17 No. 7

Comparative Analysis of Performance of Different Machine Learning Algorithms for Prediction of Success of Bank Telemarketing

[1]Aliyu Aishatu Mohammed, [2]Morufu Olalere, [3]Abdullahi, Ibrahim Mohammed

[1]Department of Computer Science, Niger State Polytechnic, Zungeru, Niger State, Nigeria [2]Department of Cyber Security Science, Federal University of Technology, Minna Niger State, Nigeria [3]Department of Computer Science, Federal University of Technology, Minna Niger State, Nigeria.

Keyword:
Telemarking,
Random Forest,
Data Mining,
Prediction,
Machine Learning

ABSTRACT

The development in technology has brought revolution in many areas of endeavors across the globe. In recent years, telemarketing has been a popular method of marketing in bank industry. Telemarking is a method of direct marketing in which a salesclerk beseech potential clienteles to buy products or services by means of phone calls. For effective managerial decision, prediction of success of bank telemarketing becomes necessary. Hence, there is need for prediction approach that will predict success of bank telemarketing with high predictive accuracy. As a result, various researchers have proposed different approaches for prediction of success of telemarketing. Machine learning approach is one of the famous approaches used by the previous researchers in this area. Different prediction algorithms have been employed, though not many of these algorithms have been applied in this area. To identify the best machine learning algorithms among the already used and unused becomes impossible. Consequently, this study presents comparative analysis of performance of different machine

learning algorithms for prediction of success of bank telemarketing. To achieve this, a dataset of 45,221 instances with 17 attributes was used to train these algorithms in WEKA environment. The performance of each algorithm was measured in terms of Accuracy, Precision, Recall and F-Measure. Our performance evaluation analysis revealed that Random Forest performed best in terms of accuracy while Voted perceptron has lowest accuracy. In terms of precision rate, SMO perform best while Voted perceptron has lowest performance in terms of precision rate. It is our hope that this study will go a long way in assisting future researchers and bank industry in the selection of predictive algorithms. *Keywords Naive Bayes, Random Forest, Decision Table, Multilayer Perceptron, J48, SMV, LMT, WEKA.*

Introduction

Due to the advance in technology, various devices and means are used in the different fields of industries such as banking, insurance and customer services to carry out business activities which include withdrawal and deposits, telemarketing among others. Telemarketing is a direct marketing used to provide exact solutions for the organizations, including marketing products and services and finding potential clients. As a cost-effective process between the company and its customers, telemarketing is speedily growing all over the world within a different field of industries, including telecommunications, banking, insurance and other financial institutions [1]. In another narration of [2], the telemarketing or telesales is a means of direct marketing in which a salesclerk beseech potential clienteles to buy products or services, through different means such as phone, subsequent face to face, and schedule during web conferencing. Telemarketing can also include recorded sales pitches programmed to be played over the phone via automatic dialing. Similarly, [3] define telemarketing as a collaborating technique of direct marketing that a telemarketer asks prospective customers via the phone to make a sale of merchandise or service. The direct marketing is the marketing discovered to locate prospects for additional services based on the customer data collected in the database known as the database marketing. A database of likely customers

can benefit greatly from the direct marketing such as communication, advertisement and analysis. The most effective telemarketing is to emphases on the value of prospect data, attempting to predict the expected customers that have a higher probability to use the service by using data mining technique.

According to [4] telemarketing business is rapidly growing in some part of the world such as Indonesia and India with increasing number of mobile users and very little communication cost which are basically through phone calls enabling several companies to have more contact list and making the product owners become more absorbed in the business. The portability of telephone has brought a revolution in the field of communication. Telemarketing has developed as a powerful tool for direct marketing due to the rapid growth of mobile telephony which received all the benefits of direct marketing but at a much lower cost [5]. Marketing promotions mainly constitute a technique of outsourcing by organizations with the aim of improving the financial position of their businesses and also having a competitive advantage over their peers. Reaching out to the customer through remote communication to facilitates operational administration of promotions [6].

Several kinds of research has been conducted in the field of telemarketing to understand customer behavior, several banks have embraced the predictive method based on the data mining to predict the customer data for classifying the customers before offering special services [3], identifying the group of customers based on high probability of subscribing to a long-term deposit and customers profiling through prediction [9]. Other compared just four data mining models which are DT, SVM, LR, and NN via a realistic rolling window evaluation and two prediction metrics.

In view of these, this study is aimed at studying and analyzing the effect of various machine learning algorithms in the prediction of success in banking telemarketing. The study employed 15 machine learning algorithms such as Random Forest, PART, Naïve Bayes, Support Vector Machine, Multiple Regression, Simple regression and Bayes Simple. To effectively evaluate these algorithms, accuracy, f-measure, recall and precision are used as the matrix used. A dataset of 45,221 instances with 17 attributes was used on two methods of validation which are 10 cross-valuation and 75% split to compare the effect of each validation method on the instance based on different algorithms.

RELATED WORKS

Sundry studies have been investigated by many researchers and data miners in the phenomenon of Bank Direct Marketing via Telemarketing Promotions.

Decision Support Systems as well other data-driven methodologies were widely explored [6].

According to [5] the mobile telephony has brought a revolution of its kind in the field of communication. Telemarketing has emerged as a powerful tool for direct marketing due to the rapid growth of mobile telephony. The authors conducted a survey which described the state of mobile telephony in India, through personal contacts using structured questionnaire. The findings of the survey have given mixed response leading to the conclusion that marketers have to work harder to make this tool more effective.

The large amounts of data that banks have been collecting for years can have a significant influence on the success of Data Mining (DM) efforts. By using DM to analyze patterns and trends so as bank managers can precisely foresee customer reactions to modifications of interest rates, the likelihood of acceptance of new product offers and which customers will have a higher risk of not paying a loan, thus making customer relationships more lucrative [7].

Many works have been carried out to compare the performances of some prediction algorithms in the success of bank telemarketing. The researchers predicted customer response to bank direct marketing by using four classifiers which are Multilayer Perceptron Neural Network (MLPNN), Decision Tree (C4.5), Logistic Regression and Random Forest (RF) [6]. The findings of their work showed that NN has the better result of 0.80 and ALIFT of 0.67 during the rolling window evaluation. Similarly, the research work of [5] observed the insight and attitude of the customers towards telemarketing, the benefits derived by the enterprise in the awareness of customers and an overall assessment of this marketing tool. Telemarketing has been selected for understanding the role of various social-psychological issues attached to this medium especially in Indian context. The paper describes the state of mobile telephony in India followed by a survey conducted through personal contacts using structured questionnaire. In the study of [4], the authors used ethnography, mystery shopping, in-depth interview and questionnaire as a way of data collecting data with 122 respondents. This study found that there are three types of customer-response toward telemarketing: Permissive, Apathetic and Aggressive.

According to [8]. Compared four DM models: logistic regression, decision trees (DTs), Neural Network (NN) and Support Vector Machine. Using two metrics, area of the receiver operating characteristic curve (AUC) and area of the LIFT cumulative curve (ALIFT), the four models were tested on an evaluation set, using the most recent data (after July 2012) and a rolling window scheme. The NN presented the best results (AUC = 0.8 and ALIFT = 0.7), Their method was

compared with other methods such as data prediction Self Organizing Map and K Means based on criteria Area under Curve. The results indicate that the Area under Curve (AUC used as a benchmark for performance evaluation) in the proposed method is better than other methods.

The study of [9] emphases on helping banks to increase the accuracy of their customer profiling through grouping as well as identifying a group of customers who have a high probability to subscribe to a long-term deposit. In the experiments, three classification algorithms were used, which are Naïve Bayes, Random Forest, and Decision Tree.

The study of [7], developed an Intelligent Bank Market Management System (IBMMS) for bank managers who want to manage capable marketing campaigns. IBMMS is the first system developed by combining the power of data mining with the capabilities of expert systems in this area. Moreover, IBMMS includes important features that enable it to be intelligent: a knowledge base, an inference engine and an advisor. Using this system, a manager can successfully direct marketing campaigns and follow the decision schemas of customers both as an entity and a group, a manager can make decisions that lead to the desired response by customers. Several methods were used by [10] to study analyze and predict numerous algorithm which centered on Naïve Bayes and Decision Tree. The result showed that for a small dataset naïve Bayes performed better while Decision Tree much better in a large dataset. Rapid Miner is the tool used for running the experiment using precision, recall and accuracy.

MACHINE LEARNING ALGORITHMS

The following the machine learning algorithm were used for the proper comparison and evaluation.

The Naïve Bayes is a probabilistic algorithm that is simple and assumes the effect of a variable independently from other variables [11].

Support Vector Machines is seen as the most popular and common algorithms with that is used in learning activities in the area of data mining [12]. In this method, a set of training examples is given with which each example is marked belonging into one of two categories. Then, by using the Support Vector Machines algorithm, a model that can predict whether a new example falls into one category or other is built decision tree structure the leaf node had a decision of expected output.

J48 Classifier this classifier is designed to improve the implementation of the C.4.5 algorithm which is implemented by [13] in 1993. The expected output

based on this classifier is in the form of decision binary trees but with more stability between computation time and accuracy [20] [14]. Regarding the decision tree structure, the leaf node had a decision of expected output.

Multi-layer Perceptron (MLP) Classifier: is one of the most common functions classifiers that prove its effectiveness to deal with several application areas e.g. time series, classification and regression problems [15]. The testing phase can be implemented within a short period of time. On the other hand, the training phase is typically implemented over a long period of time. MLP algorithm can be implemented with various transfer functions e.g. Sigmoid, Linear and Hyperbolic. The number of outputs or expected classes and number of hidden layers is important to design considerations of the MLP algorithm implementations. In the beginning, every node within the neural network had its random weight and bias values, the large weight values present the most effective attributes within a dataset, and on the contrary, the small weight values present the lowest effective attributes within a dataset.

K* algorithm can be defined by [16] as a method of cluster analysis which mainly aims at the partition of „n‟ observation into k‟ clusters in which each observation belongs to the cluster with the nearest mean. We can describe K* algorithm as an instance based learner which uses entropy as a distance measure. The benefits are that it provides a consistent approach to the handling of real-valued attributes, symbolic attributes and missing values. According to [17] K* is a simple, instance-based classifier, like K Nearest Neighbor (K-NN). New data instances, x, are assigned to the class that occurs most frequently amongst the k-nearest data points, yj where j = 1, 2…k.

A Bayes Net learns Bayesian networks made in nominal attributes and no missing values (any such values are replaced globally). Bayes Nets or Bayesian networks are graphical representation for probabilistic relationships among a set of random variables. Given a finite set X= {X1…Xn} of discrete random variables where each variable Xi may take values from a finite set represented by Val (Xi). [18]

IBK (K - Nearest Neighbor): IBK is a k-nearest-neighbor classifier that uses the same distance metric. The number of nearest neighbors can be specified explicitly in the object editor or determined automatically using leave-one-out cross-validation focus to an upper limit given by the specified value. IBK is a k-nearest neighbor classifier. A kind of different search algorithms can be used to speed up the task of finding the nearest neighbors. A linear search is a default but further options include KD-trees, ball trees, and so-called "cover trees". The distance function used is a parameter of the search method. The remaining thing

is the same as for IBL—that is, the Euclidean distance; other options include Chebyshev, Manhattan, and Minkowski distances. [19] Predictions from more than one neighbor can be weighted according to their distance from the test instance and two different formulas are implemented for converting the distance into a weight. [20].

Simple Logistic Regression (SLR) is similar to linear regression, but the nominal variable is dependent and not a measurement. The goal of SLR is if getting the probability certain value of the nominal variable is related with the measurement variable and to forecast the probability of getting a precise value of the nominal variable, given the measurement variable [21].

Decision tables are one type of analysis method that is commonly used by developers for design and specification documents and by others in software engineering and other disciplines such as test engineers, test analysts and test managers. Decision tables are used mainly because of their visibility, clearness, coverage capabilities, low maintenance and automation fitness. Also, decision tables are a useful source of information for model-based testing and work well on rule-based mechanisms or modeling techniques [22] [23].

Voted Perceptron is a perceptron algorithm which takes the advantage of data which are separable linearly with a huge verge. The algorithm is simple, efficient and easy to implement. It can also be used in very high dimensional spaces using kernel functions [24].

According to [25] a Lazy Bayesian Rule (LBR) is an algorithm that constructs lazily a simple Bayesian rule. This is very a high-level accuracy for any kind of prediction for each test instances at a time.

DATASET DESCRIPTION AND ANALYSIS

In other have a better performance of the algorithms, a total number of dataset 45,221 instances was collected from Portuguese retail bank, from May 2008 to June 2017, in a total of 52,944 phone contacts with the aid of UCI Machine Learning Repository. The dataset has 17 attributes of different variable and was converted into attribute related file format (ARFF) (a format compatible for machine learning) supported by the Waikato Environment for Knowledge Analysis (WEKA) tool for the experimentation. Figure 1 shows the experimental process for this study and Table 1 shows the attributes used for the dataset generation in this study.

Figure 1 Process flow for the experimentation of this study

Table 1: Attribute Name and Type

S/N	Attribute Name	Attribute Type
1	Age	**Numeric**
2	Job	**Categorical**
3	Marital status)	**Categorical**
4	Education	**Categorical**
5	Default	**Categorical**
6	Balance	**Numeric**
7	Housing	**Categorical**
8	Loan	**Categorical**
9	Contact	**Categorical**
10	Day	**Numeric**
11	Month	**Categorical**
12	Duration	**Numeric**
13	Campaign	**Numeric**

14	Pdays	**Numeric**
15	Previous	**Numeric**
16	Poutcome	**Categorical**
17	**Categorical**	

To adequately classify the telemarketing dataset, Bayes Net, Naive Bayes, Naive Bayes updateable, Logistic Regression, Multilayer Perceptron, SGD, Simple Logistics, SMO, Voted Perceptron, Lazy K-star, Random Forest J48, Logistic Model Tree, Decision Table and PART were the algorithms used. For proper evaluation, 10 folds cross-validation in this research. The choice of 10 folds was due to results obtained from broad tests on various datasets, with varying learning procedures, that have demonstrated that 10 is about the correct number of folds to get the best gauge error for cross-validation, the data is partitioned arbitrarily into 10 parts in which the class is represented in approximately the same proportions as in the full dataset. Each partition is held out in turn and the learning scheme trained on the remaining nine-tenths; then its error rate is processed on the holdout set. Hence, the learning procedure is carried out 10 times on various training sets. The averages of the 10 error estimates are taken to give an overall error estimate. For Comparative reasons, the dataset was also run using percentage split which allows a certain percentage of the dataset for testing, 75% split was employed for this study work [6]

PERFORMANCE EVALUATION FOR THIS STUDY

The following are the performance measures and their mathematical expressions used in the evaluation of the algorithms which is similar in the research of [26] [26]and [28]

i. True Positive: - This represents instances that are correctly predicted as success.

$$TP = \frac{TP}{TP + FN} \quad (1)$$

ii. **False Positive:** This represents success that are incorrectly predicted as failure.

$$TN = \frac{TN}{TN + FP} \quad (2)$$

iii. **Precision** shows the class agreement of the data labels with the success labels given by the predicted.

$$p = \frac{TP}{TP + FP} \quad (3)$$

iv. **Recall** shows the effectiveness of a predictor to identify the positive labels

$$R = \frac{tp}{tp+fn} \quad (4)$$

v. **F-measure** shows the relation between data's positive labels and those given by a predictor.

$$fmeasure = \frac{2pr}{p+r} \quad (5)$$

vi. **Predictive accuracy**: -this metric is expressed in %, and it shows the number of success predicted in bank telemarketing by each of the predictor.

$$Accuracy = TN = \frac{TP+TN}{TP+TN+FP+FN} = p = \frac{TP+TN}{N} \quad (6)$$

RESULTS AND DISCUSSION

To effectively evaluate the performance of this machine learning algorithms, 45,221 instances were used with 17 attributes. 10 folds cross-validation was used with 75% split for training and 25 for testing. The comparison of performance in terms of Accuracy, Precision, Recall, F Measure, Root Mean Squared Error, Receiver Operator Characteristics Area and Root Relative Squared Error was performance.

Figure 2 shows the comparison between 10 fold cross-validation and 75% percentage split of the performance prediction accuracy of the algorithms. As is shown in the in Figure 2, Rotation Forest has the highest accuracy of 0.950, when 10 cross-validations were applied. Naïve Bayes, Naïve Bayes algorithm has the lowest performance of 0.880 when both validation methods are applied. This implies with a large number of the trained dataset, Rotation Forest performed better when compared with a small number of dataset

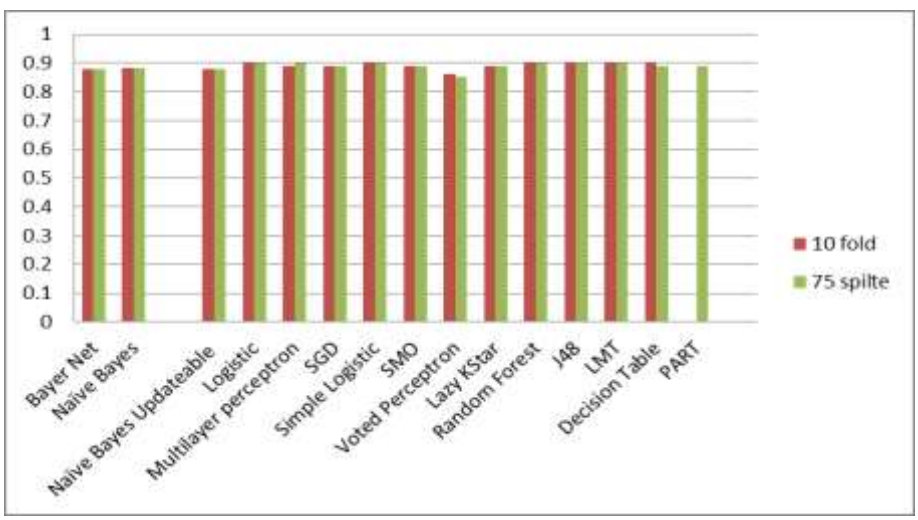

Figure 1: Prediction accuracy

Figure 1. Prediction accuracy of the algorithm with 10 fold cross-validation and 75% split

Figure 3, 4 and 5 indicated the three performance measure which includes precision, recall and f-measure. The Figure 2, 3 and 4 indicated highest precision rate of 0.946 in SMO algorithm when using 10fold cross-validation. In both validation methods, voted perceptron has the lowest performance of 0.821 in all the three performance measures.

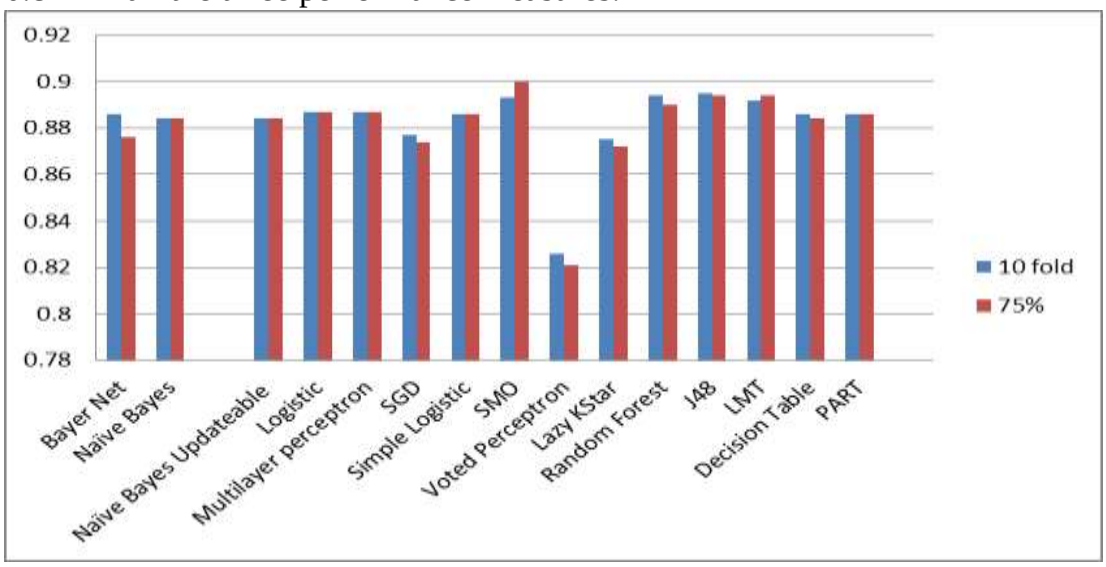

Figure 2: Performance of Precision rate on the algorithms.

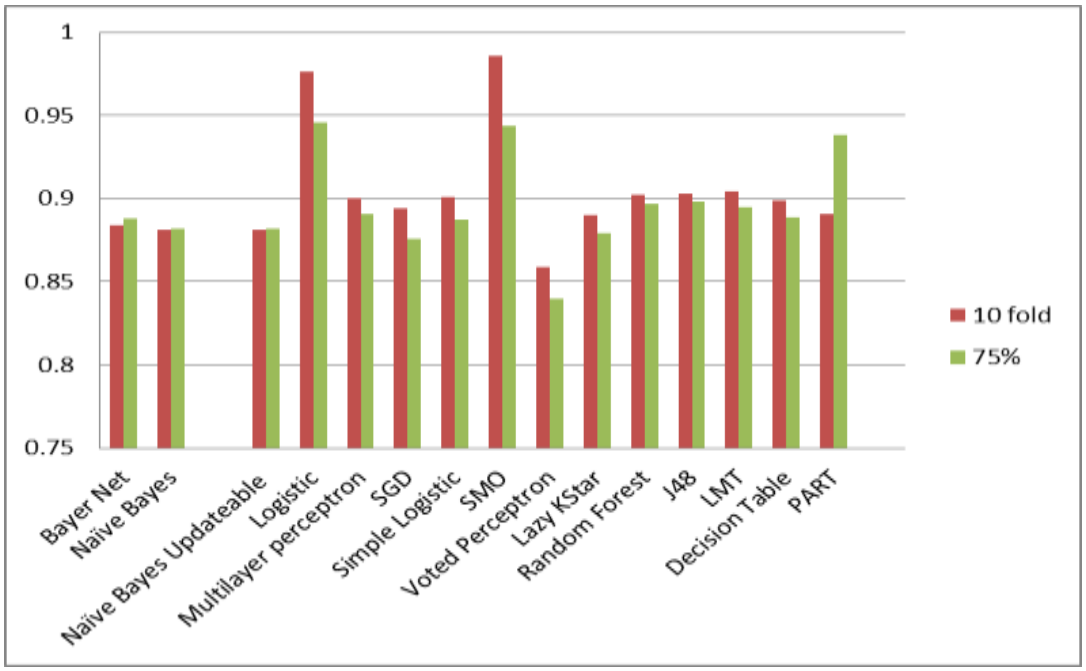

Figure 3: Recall performance on all algorithms

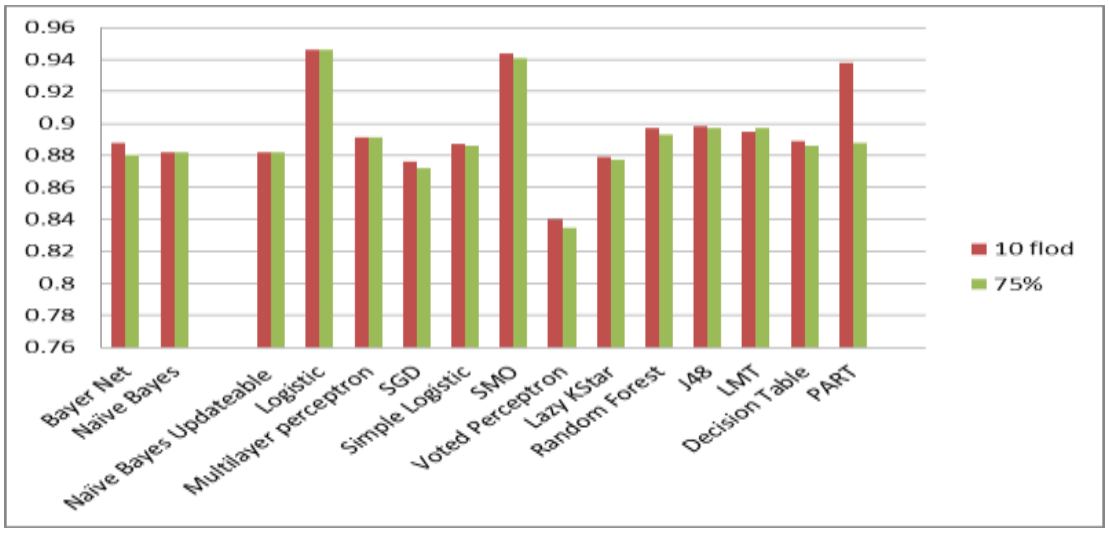

Figure 4: F-measure Performance on all algorithms

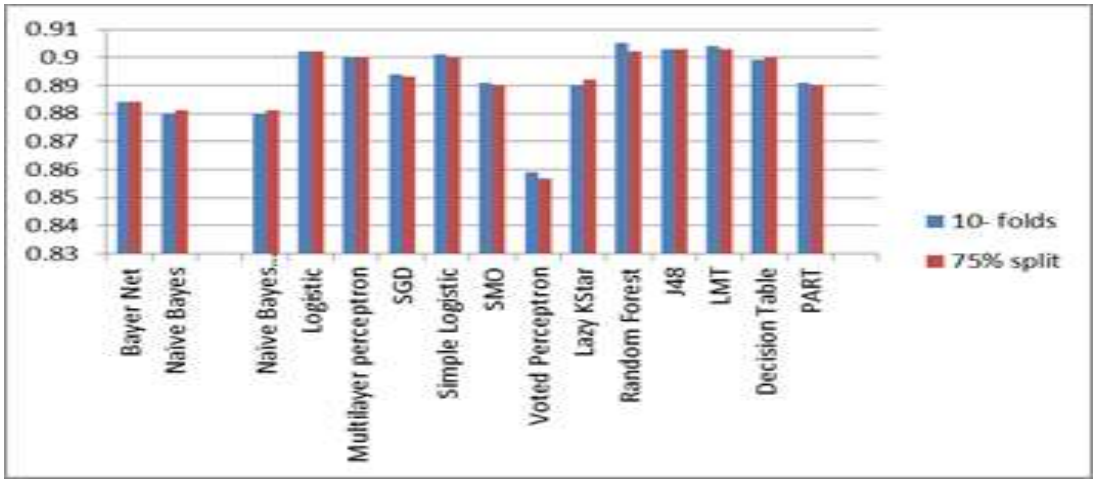

Figure 5 TP Rate performances on all algorithms

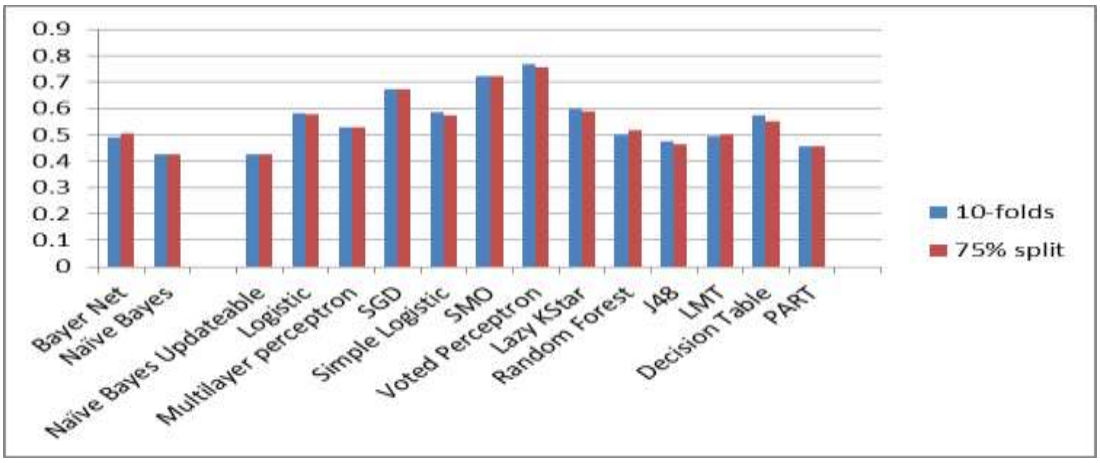

Table.2 Results of Accuracy, Precision, Recall, F-Measure, TP Rate, FP Rate

S/No	Algorithms	TP Rate		FP Rate		Accuracy		Precision		Recall		F-measure	
		10-folds	75% split	10-folds	75% split	10-folds	75% split	10-folds	75% split	10-folds	75% split	10-folds	75% split
1	Bayer Net	0.884	0.884	0.492	0.505	0.88	0.88	0.886	0.876	0.890	0.884	0.888	0.880
2	Naïve Bayes	0.880	0.881	0.426	0.427	0.881	0.881	0.884	0.884	0.880	0.881	0.882	0.882
3	Naïve Bayes Updateable	0.880	0.881	0.426	0.427	0.88	0.88	0.884	0.884	0.880	0.881	0.882	0.882
4	Logistic	0.902	0.902	0.580	0.578	0.90	0.90	0.887	0.887	0.975	0.976	0.946	0.946
5	Multilayer perceptron	0.900	0.900	0.527	0.528	0.89	0.899	0.887	0.887	0.895	0.900	0.891	0.891
6	SGD	0.894	0.893	0.674	0.672	0.89	0.89	0.877	0.874	0.896	0.894	0.876	0.872
7	Simple Logistic	0.901	0.900	0.586	0.575	0.90	0.90	0.886	0.886	0.901	0.901	0.887	0.886
8	SMO	0.891	0.890	0.723	0.722	0.89	0.89	0.893	0.900	0.965	0.986	0.944	0.941
9	Voted Perceptron	0.859	0.857	0.766	0.755	0.86	0.85	0.826	0.821	0.865	0.859	0.840	0.835
10	Lazy KStar	0.890	0.892	0.599	0.590	0.89	0.89	0.875	0.872	0.892	0.890	0.879	0.877
11	Random Forest	0.905	0.902	0.503	0.515	0.950	0.90	0.894	0.890	0.905	0.902	0.897	0.893
12	J48	0.903	0.903	0.475	0.463	0.90	0.90	0.895	0.894	0.903	0.903	0.898	0.897
13	LMT	0.904	0.903	0.493	0.503	0.90	0.90	0.892	0.894	0.903	0.904	0.895	0.897
14	Decision Table	0.899	0.900	0.572	0.551	0.90	0.89	0.886	0.884	0.900	0.899	0.889	0.886
15	PART	0.891	0.890	0.458	0.458	0. 89	0.89	0.886	0.886	0.945	0.891	0.938	0.888

CONCLUSION

The experiment of this study was conducted in WEKA using a direct marketing deposit dataset. This study applied and compare the prediction performances of several machine learning algorithms in terms of prediction ability. This provides an improvement over the study undertaken by in previous research works where

they compared the performances of MLPNN, Decision Tree and Logistic Regression with the classifier coming first. These insights and revelations would go a long way to improve the effectiveness of bank direct marketing as well as telemarketing to banks and other financial institutions. From the results obtained 14 different prediction algorithms (including commonly used algorithms) using two test options established that, some uncommon algorithms perform relatively well on the telemarketing dataset. Random Forest emerged as the best with an accuracy of 0.950 even better than some commonly used prediction algorithms including J48 which records 0.90 accuracy, Naïve Bayes with 0.880and Multilayer Perceptron with 0.860. The results are promising for further applications in telemarketing problems.

References

L., Salim. A. Two-Step System for Direct Bank Telemarketing Outcome Classification. Research article Intell Sys Acc Fin Mgmt. 2017;24:49–55. 2017. Copyright © 2017 John Wiley & Sons, Ltd.

S., Moro, P., Cortez, and P. Rita, A data-driven approach to predict the success of bank telemarketing. Decision Support Systems, 62, 22–31. 2014 .https://doi.org/10.1016/j.dss.2014.03.001

C. Vajiramedhin and A. SuebsingFeature Selection with Data Balancing

E., Amalia, T., Maulana and N., Khairunnisa Permissive, Aggressive or Apathetic? Indonesian

Telemarketing Customer, Procedia - Social and Behavioural Sciences 169) 69 – 74. 2015.

G. T., Deepesh, and P. Gupta A Study of Indian Consumers' Perception on Telemarketing. International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 2, No. 2. 2012 for Prediction of Bank Telemarketing. A**pplied** Mathematical Sciences, Vol. 8, 2014, no. 114, 5667 – 5672. 2014. HIKARI Ltd, www.m-hikari.com http://dx.doi.org/10.12988/ams.2014.47222

A., Justice and J. Manoj. Predicting Customer Response to Bank Direct Telemarketing Campaign. International Conference on Engineering Technology and Technopreneurship (ICE2T) 2017.

A., Keles, and A. Keles, IBMMS Decision Support Tool for Management of Bank Telemarketing Campaigns. International Journal of Database Management Systems 7(5), 1–15. 2015.

S., Moro, P., Cortez, and P. Rita, A data-driven approach to predict the success of bank telemarketing. Decision Support Systems, 62, 22–31. https://doi.org/10.1016/j.dss.2014.03.001 2014.

P., Shamala, M., Aida., F. M., Cik and A., Rodziah. Customer Profiling using Classification Approach for Bank Telemarketing . INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION. 1. 2017 NO 4 - 2 . Retrived from file:///C:/Users/AHARAMBIF/Downloads/68-212-1-PB.pdf and accessed on 2oth april, 2018.

P., Thendral, M., V., Viswanatham, S. Ashmeet Comparison of Performance of Various Data Classification Algorithms with Ensemble Methods Using RAPIDMINER. *International Journal of Advanced Research in Computer Science and Software Engineering*. Volume 6. 2016. ISSN: 2277 128X www.ijarcsse.com

T. Mitchell," Machine Learning," McGraw Hill, New York, 1997.

V. Vapnik, "The Nature of Statitical Learning Theory," Springer, Heidelberg, 1995.

J.R. Quilan, "C4.5: Programs for machine learning," in Elsevier, 2014.

M.S.Bhuller and A. Kair, "Use of data mining in educstion sector," in proceedings of the world congresson engineering and computer science, vol. 1, 2012.

S.K. Pal andS. Mitra, "Multilayer perception, fuzzy sets and classification," in IEEE transactions on neural networks, vol. 3, 1992.

S. Vijayarani and M. Muthulakshmi, "Comparative Analysis of Bayes and Lazy Classification Algorithms," in International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 8, August 2013 Copyright to IJARCCE www.ijarcce.com 3118

T. C. Sharma and M. Jain, "WEKA Approach for Comparative Study of Classification Algorithm"

K. H. Raviya and B. Gajjar, "Performance Evaluation of Different Data Mining Classification Algorithm Using WEKA"

S. Ghosh, S. Roy and S. K. Bandyopadhyay, "A tutorial review on Text Mining Algorithms"..

S.Vijayarani and S.Sudha,"Comparative Analysis of Classification Function Techniques for Heart Disease Prediction"

J. H. McDonald, "Handbook of Biological Statistics," Baltimore, Maryland, (3rd ed.). Sparky House Publishing, 2014, pp. 238-246

A. Linetzki, “Analyzing decision tables A simple process for learning, analyzing and producing test cases from decision tables,” in Testing Experience Magazine, March 2012.

IBM Websphere ILOG Rules

Voted Perceptron, "Voted Perceptron," 29 0ctober 2011. [Online]. Available: http://curtis.ml.cmu.edu/w/courses/index.php/Voted_Perceptron. [Accessed 10 May 2018

Z. Xie, "LBR-Meta: An Efficient Algorithm for Lazy Bayesian Rules," in *220 Handan Road, Shanghai 200433, PR.* , China, 2008.

M. Olalere, A. E., Raji, O. A., Joseph, I, Ismaila, R J. Gbenga. A Naïve Bayes Based Pattern Recognition Model for Detection and Categorization of Structured Query Language Injection Attack. *Internation Journal of Cyber Security and Digital Forensics (IJCSDF)* 7(2): 189-199

M. Naghmeh, A Pattern Recognition Neural Network Model for Detection and Classification of SQL Injection Attacks 2015. Retrieved from https://www.researchgate.net/publication/271072307 and accessed on 20th April 2018

M. Olalere, M. T., Abdullah, R. Mahmod and A. Abdullah. Identification and Evaluation of Discriminative Lexical Features of Malware URL for Real-Time Classification. International Conference on Computer & Communication Engineering. 2016.