Predicting the Success of Bank Telemarketing for Selling Long-term Deposits: An Application of Machine Learning Algorithms

Premkumar Borugadda*
Doctoral Scholar, Department of Computer Science, Pondicherry University, Karaikal Campus, Karaikal - 609605, India.
E-mail: premkumar.jones@gmail.com

*Corresponding Author

Dr. Prabhakar Nandru
Guest faculty, Department of Commerce, School of Management, Pondicherry University, Karaikal Campus, Karaikal - 609605, India.
E-mail: prabhakarnandru2020@gmail.com

Dr. Chendragiri Madhavaiah
Assistant Professor, Department of Management, School of Management, Pondicherry University, Karaikal Campus, Karaikal - 609605, India.
E-mail: drcmadhavaiah@gmail.com

Received: 13/ 04/2021 Revised: 23/06/2021 Accepted: 24/06/2021

Abstract

This study attempts to investigate the demand for the adoption of telemarketing practices for promoting long-term bank deposits to potential bank customers. The study explored the demand for long-term bank deposits by employing various machine learning algorithms like Random Forest (RF), Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), Decision Tree (DT), and Logistic Regression (LR). The dataset related to direct marketing campaigns (phone calls) of a Portuguese banking institution is considered for analysis. The results confirm that the LR model provides 92.48% accuracy, which is the best model for predicting the potential customers who have an interest in long-term deposits through telemarketing. The results of the study also provide insightful information to banks for making telemarketing policy decisions in the success of bank deposits to their existing and prospective bank customers.

Keywords: Long-term deposit; Machine learning algorithms; Telemarketing

1. Introduction

Globalization of economies has impacted all businesses irrespective of scale of operation. Financial services sector is not an exception to dramatic changes in terms of technology, strategy and customer service due to globalization. These changes have become source of competitive advantage for the business sectors such as healthcare, education, banking, and the insurance industry (Raghunandan et al., 2018). The banking sector has significantly influenced by globalization as an offspring which enabled banking sector to experience sea change in providing customer service, state-of-the-art technology, improved productivity and increased profitability, across the globe (Goldberg, 2009; Kamath et al., 2003; Kumbhakar & Sarkar, 2003; Maji & Dey, 2006; Roy & Sinha, 2014; Sayed & Sayed, 2013). For instance, financial institutions like banks and insurance companies started promoting their products and services through innovative marketing strategies that has enhanced the earning capacity of firms (Kishor & Nagamani, 2015). The upcoming technologies, especially disruptive information and communication technologies have posed opportunities and challenges as well to banking system across the globe (Krishan, 2015). It implies that information technology has enabled unprecedented marketing opportunities to banking companies to offer set of quality products and services to customers (Bapat & Mazumdar, 2015). Increasing competition in banking sector to adopt emerging technologies subsequently gain competition, banking companies started adopting an innovative technology-enabled marketing practice called telemarketing (Farooqi & Iqbal, 2019; Noronha & D'Cruz, 2007).

Telemarketing is a direct marking strategy method in which a salesperson acquires the prospective customers' willingness to purchase products or services over phone calls (Palaniappan et al., 2017). Notably, many financial services providers adopted telemarketing strategy to reach out to the new customers, providing better services to existing customers and meeting their specific needs (Moro et al., 2014).

In view of this, current study is aimed to predict the accuracy and preciseness of telemarketing practice of banks for selling long-term deposits. The dataset for this study is accessed from an authenticated source which is accessible publicly and that data is related to direct marketing campaigns of a Portuguese banking institution which often offers long-term deposits to prospective customers. The marketing campaigns of the banking institution are based on phone calls to the prospective customers. The novelty of this study is based on the application of machine learning algorithms viz., Logistic Regression, Decision Tree, Random Forest, Gaussian Naive Bayes, and Support Vector Machine to predict the subscription of bank long-term deposits through telemarketing services.

2. Review of Literature

Few studies examined the predicting the success of bank telemarketing for selling long-term deposits through the application of various machine learning techniques. The prior studies were shown below.

Asare-Frempong and Jayabalan (2017) anticipated a model using four machine learning algorithms: multilayer perceptron neural network, decision tree (C4.5), Logistic regression, and random forest. The dataset is taken from (UCI) Machine Learning Repository, containing 45147 instances with 17 attributes. The random forest gives better accuracy is 86.8%.

Palaniappan et al. (2017) suggested a model using data mining approaches. The dataset was taken from the UCI Machine Learning repository, with 41,188 instances 21 attributes. Three algorithms had been applied, which are Naïve Bayes, Random Forest, and Decision Tree. The experiments measured the accuracy percentage, precision, and recall rates. They found that the decision tree algorithm gives the best accuracy.

Another study by Jiang (2018) explored predicting the success of bank telemarketing using data mining approaches. The dataset was taken from the (UCI) Machine Learning Repository. There are 4119 instances and 21 attributes in this data set. They used the support vector machine, logistic regression, Naïve Bayes, neural network, and decision tree. Among these five algorithms, they got the best accuracy from the logistic regression. The accuracy of the logistic regression model was 92.03%. According to Ilham et al. (2019) proposed a model using machine learning approaches. They used different techniques: Logistic Regression, Naïve Bayes, Random Forest, K-Nearest Neighbor, Support Vector Machine, Neural Network, and, Decision Tree. Preprocessing was not done to the dataset features; it directly uses a ready dataset from the UCI repository. These models' evaluation metrics verify that the most accurate they are founded that 91.07% by using the SVM.

3. Materials and methodology

3.1 Materials

This study used the Jupyter Notebook as a tool and Python 3.7 as a programming language.

3.2 Data acquisition

The data about direct marketing campaigns of a Portuguese banking institution. The data set was collected from the UCI Machine Learning Repository. The data set has 4119 instances and 21 attributes. Among the 21 features, the target variable is 'y.' Descriptions of all 21 features are presented in Table 1

Figure 1. A framework of the machine-learning model

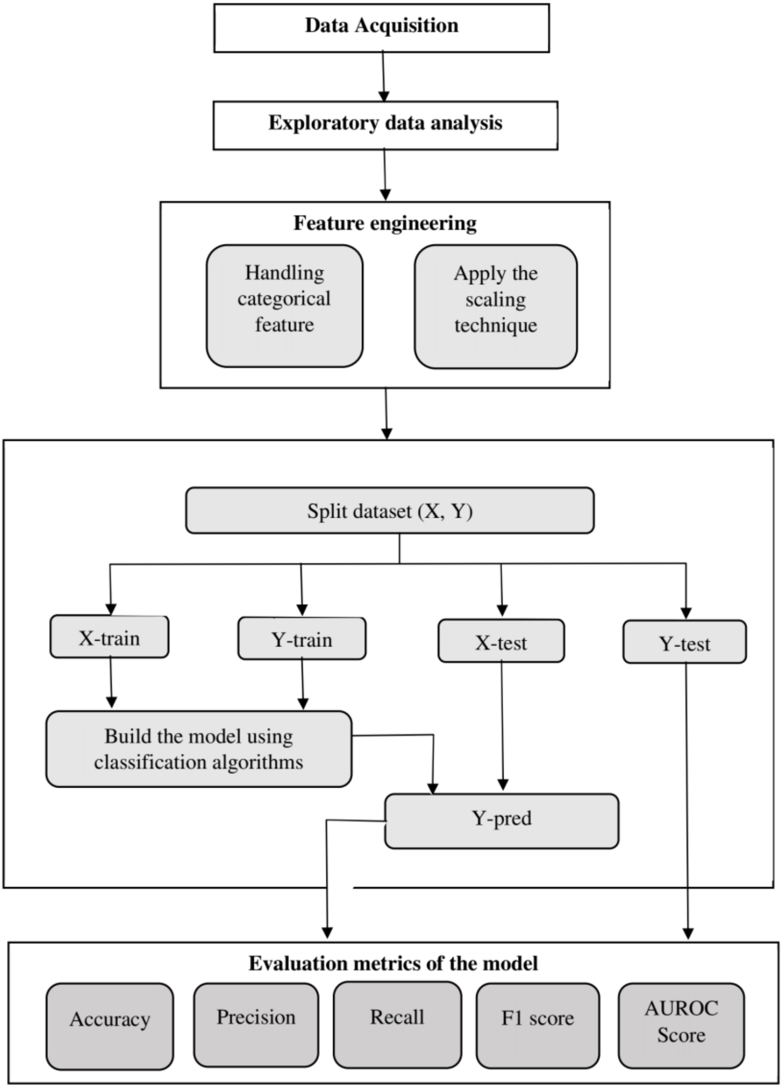


Table 1. Features of the dataset

S. No	Features	Description of Features
1	Age	Age of the client
2	Job	Type of client's job
3	Marital	Client's marital status
4	Education	Highest education of the client
5	Default	Client credit
6	Housing	Housing loan
7	Loan	Personal loan
8	Contact	Type of contact communication with the client
9	Month	Last month of the year contracting to the client
10	Day of week	Last day of the week contracting to the client
11	Duration	Duration of client contact
12	Campaign	Number of contacts performed during the campaign and for this client
13	Plays	Number of days elapsed after the client's last visit
14	Previous	Number of contacts performed before this campaign and for this client
15	Poutcome	Outcome of the previous marketing campaign
16	Emp.var.rate	Employment variation rate
17	Cos.price.IDX	Consumer price index
18	Cons.conf.IDX	Consumer confidence index
19	Euribor3m	Euribor 3-month rate
20	Nr. employed	Number of employees
21	Label	Client subscription

Source: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing

3.3 Exploratory Data Analysis

After importing the dataset, analyze the dataset to check the data's shape, check the unique labels of the target variable (Y), contain the data type of features, and the data's descriptive statistics.

3.4 Feature Engineering

After analyzing the dataset, convert the categorical variable into numeric because it works with only numeric features.
Before the train, the models apply the standard scaler technique because all independent features should have the same order of magnitude for the algorithm to work efficiently.

3.5 Build Model

Before making the model, split the data set into the train data set with 90% and the test data set with 10%. Train the model on 90 % of the training dataset, test the model on 10% of the test dataset, and then predict the target variable using the trained model.
Evaluation Metrics of the model
After building the models, check the classification models' performance like accuracy, precision, recall, F1_score, ROC curve, and AUROC score.

3.6 Classification algorithms
3.6.1 Logistic Regression

Logistic regression is one of the classification algorithms. (Asare-Frempong & Jayabalan, 2017) Logistic regression is the statistical model for classifying the classes. Logistic regression classifies binary labels like pass\fail. Logistic regression is used to predict the categories based on the threshold value.

3.6.2 Decision Tree

A decision tree is a tree structure model; it handles high dimensionality, and it is a supervised algorithm. (Keles & Keles, 2015) It is work on regression and classification models. The top of the tree is called the root node, based on the labels, and the source set splitting into edges (leaf), the end of the branch that doesn't break anymore; that is the model's decision. In the decision tree, the two techniques work on significant entropy and information gain.

3.6.3 Random Forest

Random forest is the combination of hundreds or thousands of decision tree trains each one on separately. The final predictions of the random forest are made by averaging the predictions of each tree (Breiman, 1996)

3.6.4 Support Vector Machine

A support vector machine is applied for both classification and regression problems (Zhen & Wenjuan, 2016). The main goal of SVM is to divide the datasets into classes to find a maximum marginal hyperplane.

3.6.5 Gaussian Naive Bayes

Gaussian naive Bayes classifier is one of the naïve Bayes algorithms (Vedala & Kumar, 2012). All naïve Bayes algorithms are working based on the Bayes theorem.

4. Results and discussion

In this experiment, the dataset was split into a training dataset and a testing dataset. The ratio of the training dataset and the testing dataset was 90% and 10%, respectively. In this experiment, Standard scaling and 10-fold cross-validation were applied to build the machine learning model. With this specification, the performance of Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Gaussian Naïve Bayes (GNB), and Support Vector Machine (SVM) are analyzed and compared.

4.1 Performance Measure

Evaluate the performance of classification models through a confusion matrix.

4.2 Confusion Matrix

The performance of a classification model was applied to a test data set. The confusion matrix is observed in figure 2. The tasting part of the data measures performance of classification models.

Figure 2. Confusion Matrix

	Predicted Negative (N=0)	Predicted Positive (P=1)
Actual Negative (N=0)	TRUE NEGATIVE (TN)	FALSE Positive (FP)
Actual Positive (P=1)	FALSE Positive (FP)	TRUE Positive (TP)

Definition of the Terms of Confusion matrix:

- Positive (P): Actual observation is positive
- Negative (N): Actual observation is negative
- TP (True Positive): Actual observation is positive (P), and it is predicted as positive (P)
- TN (True Negative): Actual observation is Negative (N), and it is predicted as Negative (N)
- FP (False Positive) (Type I error): Actual Observation is Negative(N), but it is predicted as positive (P)
- FN: False Negative (Type II error): Actual Observation is positive (P), but it is predicted as negative(N).

Accuracy: Accuracy is the ratio of correct predicted to total predictions.

$$\text{Accuracy}=\frac{TP+TN}{TP+TP+FP+FN}$$

Precision: Out of the total number of positive instances by the model, what percentage of an actual positive instance

$$\text{Precision}=\frac{TP}{TP+FP}$$

Recall *(Sensitivity):* Ability of a classification model to find all positive instances.

$$\text{Recall}=\frac{TP}{TP+FN}$$

F1-Score: F1 score is used to measuring accuracy. The F1 score range is between 0 and 1, and the score is nearby one means the model is excellent and accurate.

$$\text{F1-Score}=\frac{2*Precision*Recall}{(Precision+Recall)}$$

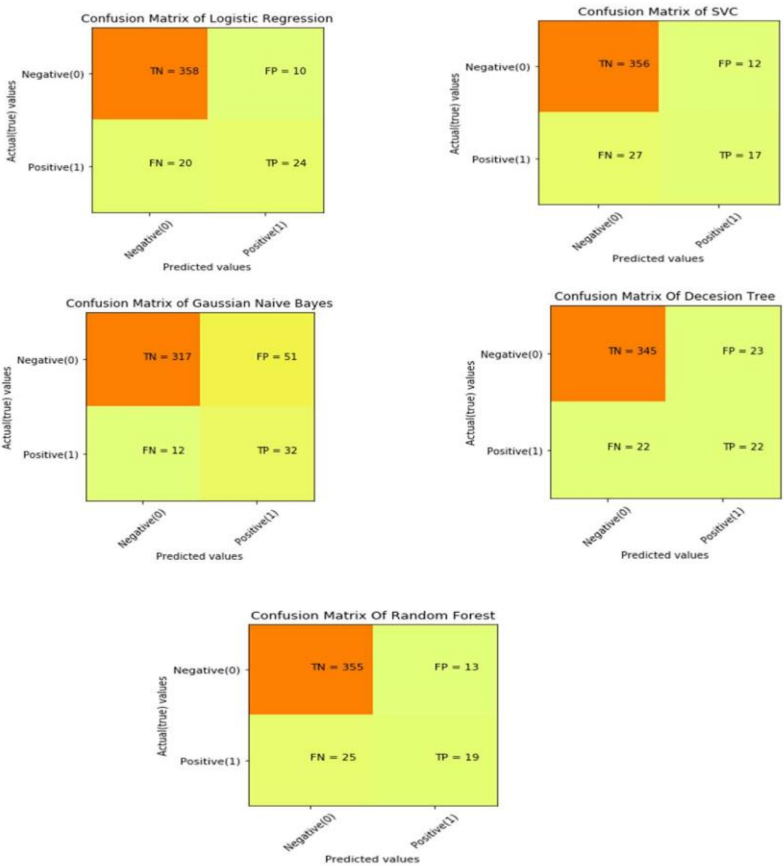
ROC (Receiver Operating Characteristic) Curve: It visualizes the performance of a binary classifier. It is a very suitable method to measure the accuracy of a classification model. The ROC curve is plotted with TPR against the FPR, where TPR is on the y-axis, and FPR is on the x-axis at different classification thresholds.

4.3 Analysis of Results

Confusion matrices of five classification algorithms are shown in figure 3, in which the confusion matrix of logistic regression has fewer wrong predictions as FP=10 and FN=20. The performance comparison of five classification machine learning models is shown in table 2, and classification reports of five ML Algorithms shown in figure 4.ROC curves of the five ML models are shown in figure 5. Among the five models, LR gives AUROC score is 93.62%.

4.3.1 Confusion Matrices

Figure 3 Confusion Matrices of classification Algorithms



The results of five classification models are displayed with five evaluation metrics in Table 2. The logistic regression model gave better results with 92.72 accuracies and 93.62 AUROC Score among five models compared to other models. So, choose a logistic regression model for deployment in real-time applications.

Table 2. The performance comparison of five classification machine learning models

S. No.	Classification Models	Evolution Metrics of Classification Models				
		Accuracy	Precision	Recall	F1-Score	AUROC -score
1.	LR	92.72	70.58	54.54	61.53	93.62
2.	GNB	85.0	38.55	72.72	50.39	84.76
3.	RF	91	59.37	43.18	50.0	93.53
4.	SVM	90.53	58.62	38.63	46.57	91.4
5.	DT	91.0	56.75	47.72	51.85	71.69

Source: Jupyter notebook output

4.3.2 Classification Reports

In a report, observe different evaluation metrics like precision, recall, f1-score, support is called classification report. While using a classification problem, we need to use metrics to check efficient machine learning models. Various classification reports are displayed in below Figure 4. Among five models, the classification report of logistic regression given the best results.

4.3.3 Receiver Operating Characteristic (ROC) Curves

A metric like ROC can give a graphical representation of the performance of the classification model. An excellent model has AUC near to the 1, which means it has a good measure of separability of two classes. A poor model has an AUC near 0, which means it has the worst measure of separability. ROC curve of various machine learning models observed in Figure 5. AUC Score of logistic regression is 93.62 highest score is observed in figure 5a, compared to the rest of the models.

Figure 4. Classification Reports of five Classification Models

Classification Report of Logistic Regression:

	precision	recall	f1-score	support
0	0.9471	0.9728	0.9598	368
1	0.7059	0.5455	0.6154	44
accuracy			0.9272	412
macro avg	0.8265	0.7591	0.7876	412
weighted avg	0.9213	0.9272	0.9230	412

Classification Report of Gaussian Naive Bayes:

	precision	recall	f1-score	support
0	0.9635	0.8614	0.9096	368
1	0.3855	0.7273	0.5039	44
accuracy			0.8471	412
macro avg	0.6745	0.7943	0.7068	412
weighted avg	0.9018	0.8471	0.8663	412

Classification Report of SVC:

	precision	recall	f1-score	support
0	0.9295	0.9674	0.9481	368
1	0.5862	0.3864	0.4658	44
accuracy			0.9053	412
macro avg	0.7579	0.6769	0.7069	412
weighted avg	0.8928	0.9053	0.8966	412

Classification Report of DT:

	precision	recall	f1-score	support
0	0.9387	0.9565	0.9475	368
1	0.5676	0.4773	0.5185	44
accuracy			0.9053	412
macro avg	0.7531	0.7169	0.7330	412
weighted avg	0.8990	0.9053	0.9017	412

Classification report of Random Forest Classifier:

	precision	recall	f1-score	support
0	0.9342	0.9647	0.9492	368
1	0.5938	0.4318	0.5000	44
accuracy			0.9078	412
macro avg	0.7640	0.6982	0.7246	412
weighted avg	0.8979	0.9078	0.9012	412

Figure 5. ROC Curves of five Classification Models

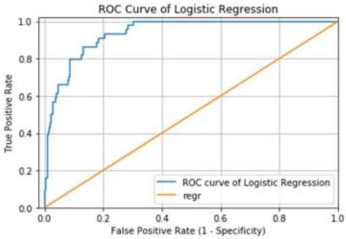


Figure 5a

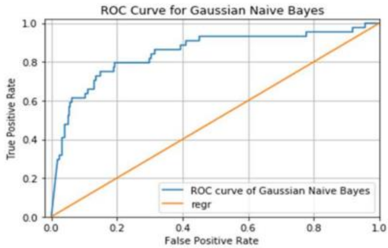


Figure 5b

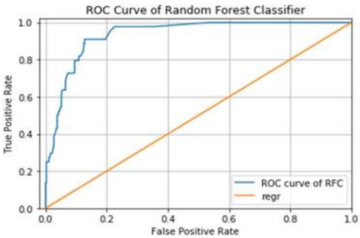


Figure 5c

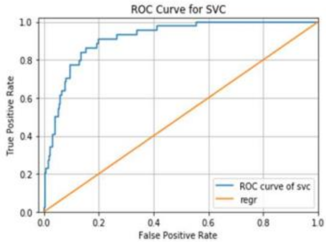


Figure 5d

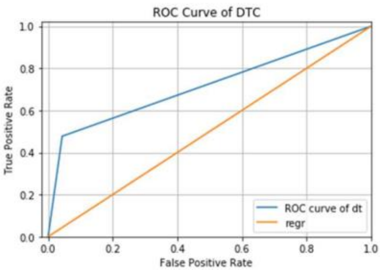


Figure 5e

5. Conclusion

In this study, five algorithms were applied viz., Logistic Regression, Decision Tree, Random Forest, Gaussian Naïve Bayes, and Support Vector Machine. Among these five algorithms, the Logistic Regression Classifier has improved accuracy, i.e., 92.48%, and AUROC scores are 93.62%. Therefore, discerning knowledge to accelerate the telemarking campaign opportunity is emerged to sustain a competitive advantage over the existing and prospective service providers in the banking industry. The results may help service providers to frame strategies pertaining to telemarking marketing campaign that will help increase banks' earning capacity. In order to get success in a telemarketing campaign, the service providers may use employee python (Flask framework) as back-end work for GUI (Graphically User Interface). The LR method was identified as the best method among the five models and this method can be used in the predict section of the web application where service providers can insert input data that would generate a response by predicting client subscribed to a term deposit or not.

The study results also provide insightful information to future researchers to obtain the best results when they apply the features selection techniques and balancing techniques like Oversampling and under-sampling, which enables the development of the machine learning models and offers a substantial contribution to the existing literature.

References

Asare-Frempong, J., & Jayabalan, M. (2017). Predicting customer response to bank direct telemarketing campaign. In 2017 *International Conference on Engineering Technology and Technopreneurship* (ICE2T), (1-4). IEEE.

Bapat, D., & Mazumdar, D. (2015). Assessment of business strategy: Implication for Indian banks. *Journal of Strategy and Management*, *8*(4), 306-325.

Breiman, L. (1996). Bagging predictors. Machine learning, *24*(2), 123-140.

Farooqi, R., & Iqbal, N. (2019). Performance evaluation for competency of bank telemarketing prediction using data mining techniques. *International Journal of Recent Technology and Engineering*, *8*(2), 5666-5674.

Goldberg, L. S. (2009). Understanding banking sector globalization. *IMF Staff Papers*, *56*(1), 171-197.

Ilham, A., Khikmah, L., & Iswara, I. B. A. I. (2019). Long-term deposits prediction: a comparative framework of classification model for predict the success of bank telemarketing. *In Journal of Physics: Conference Series*, *1175* (1), 1-6.

Jiang, Y. (2018). Using Logistic Regression Model to Predict the Success of Bank Telemarketing. *International Journal on Data Science and Technology*, *4*(1), 35.

Kamath, K. V., Kohli, S. S., Shenoy, P. S., Kumar, R., Nayak, R. M., Kuppuswamy, P. T., & Ravichandran, N. (2003). Indian banking sector: Challenges and opportunities. *Vikalpa*, *28*(3), 83-100.

Kaur, N. (2013). Customer Relationship Management in Indian Banking Sector. *BVIMR Management Edge*, *6*(1), 33-43.

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

Kishor, N. R., & Nagamani, K. (2015). Customer relationship management in Indian banking sector. *ACADEMICIA: An International Multidisciplinary Research Journal*, *5*(7), 74-82.

Krishna, V. R. (2015). Impact Of Information and Communication Technology (ICT) On Indian Banking Sector. *Voice of Research*, *3*(4), 54-58.

Kumbhakar, S. C., & Sarkar, S. (2003). Deregulation, ownership and efficiency in Indian banking. *Arthaniti-Journal of Economic Theory and Practice, 2*(1-2), 1-26.

Maji, S. G., & Dey, S. (2006). Productivity and Profitability of Select Public Sector and Private Sector Banks in India: An Empirical Analysis. *The IUP Journal of Bank Management*, (4), 59-67.

Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, *62*(62), 22-31.

Noronha, E., & D'Cruz, P. (2007). Reconciling dichotomous demands: Telemarketing agents in Bangalore and Mumbai, India. *Qualitative Report*, *12*(2), 255-280.

Palaniappan, S., Mustapha, A., Foozy, C. F. M., & Atan, R. (2017). Customer profiling using classification approach for bank telemarketing. *International Journal on Informatics Visualization*, *1*(4-2), 214-217.

Raghunandan, G., Lavina, G. S., & Jose, S. (2018). Electronic Customer Relationship Management an Effective tool in the Banking Sector. *Asian Journal of Management*, 9(2), 913-919.

Roy, S., & Sinha, I. (2014). Determinants of customers' acceptance of electronic payment system in Indian banking sector–A study. *International Journal of Scientific and Engineering Research*, *5*(1), 177-187.

Sayed, G. J., & Sayed, N. S. (2013). Comparative analysis of four private sector banks as per CAMEL rating. *Business Perspectives and Research*, *1*(2), 31-46.

Vedala, R., & Kumar, B. R. (2012). An application of naive bayes classification for credit scoring in e-lending platform. In 2012 *International Conference on Data Science & Engineering (ICDSE)*, 81-84. IEEE.

Zhen, W., & Wenjuan, S. (2016). Commercial bank credit risk assessment method based on improved svm. In 2016 *International Conference on Intelligent Transportation, Big Data & Smart City* (ICITBS), 353-356. IEEE.