



Parameter Tuned Machine Learning based Decision Support System for Bank Telemarketing

Noura Metawa¹, Amany Ahmed Elshimy²

¹ Faculty of Commerce, Mansoura University, Egypt

² Faculty of Commerce, Damietta University, Egypt

Emails: N_metawe@mans.edu.eg; amany.elshimy2020@du.edu.eg

Abstract

In banking sectors, telemarketing is the major support of selling the products or services. Banking advertisement and marketing are mainly depending upon the comprehensive knowledge of objective data regarding the market and the actual client requirements for the bank gainful way. Decision Support Systems (DSS) play a vital part in telemarketing sector, which determines a specific class of automatized facts to assist the company to make decisions. Machine learning (ML) is commonly used in the DSS which integrates the data and computer application for precise prediction of results. This paper presents an effective parameter tuned ML based DSS (PTML-DSS) for bank telemarketing sector. The proposed PTML-DSS technique follows a three-level process namely preprocessing, classification, and parameter optimization. Initially, the marketing data is preprocessed to get rid of unwanted information. In addition, gradient boosting decision tree (GBDT) based classifier model is used to classify the data. Besides, firefly algorithm (FFA) is applied for tuning the parameters involved in the GBDT model. In order to verify the improved performance of the PTML-DSS technique, a series of simulations were performed, and the results are inspected under varying aspects. The resultant values reported the improved performance of the PTML-DSS technique over the other techniques.

Keywords: Telemarketing, Banking sector, Machine learning, Decision Support System, Loan approval.

Introduction

Marketing is method of revealing the targeted users to products through appropriate channels and systems [1, 2]. This eventually facilitates the manner for buying the products or services and help in defining the needs of the products and motivate customer to purchase that product [3]. The goal is to rise the trades of products and services for financial, enterprises, and business institutions. Also, it aids in maintaining the reputation of the business.

The latest developments in digital technology and the accelerated development of the global market are completely changing consumer spending and life patterns. Consumers' preference for contactless and remote interaction channels has grown, and they have become accustomed to using mobile technology to obtain the services and information they need almost anytime and anywhere. In order to cope with this situation and obtain a competitive economic advantage,

while avoiding potential negative business results, companies are striving to provide services that adapt to the digital age, while increasing the convenience of contactless channels and the proportion of direct sales. Therefore, as an important means of implementing direct sales strategy, telemarketing is highlighted. The focus of telemarketing is shifting from passive incoming calls to outgoing calls. This is an active and profitable marketing method. In the inbound method, customers are encouraged to subscribe to products or services when they call the call center. In contrast, in the outbound approach, telemarketers call customers and invite them to subscribe to products or services. Therefore, it is very important to develop technology that can accurately select potential customers who may purchase the product. [27]

Telemarketing is form of direct marketing in which salesperson approaches the customer either face to face or phone call and persuades him to buy the product [4, 5]. Telemarketing attains most popularity in 20th century and still gaining it. Nowadays, telephone (fixed-line or mobile) has been broadly used. It is cost effective and keeps the customers up to date. In banking sector, marketing is the backbone to sell its product or service. Banking advertising and marketing are mostly based on an intensive knowledge of objective information about the market and the actual client needs for the bank's profitable manner. Making right decisions in organizational operations is sometimes proved a great challenge where the quality of decision really matters [6-8].

Decision Support Systems (DSS) are classified as a particular class of computerized facts and figures that helps the organization or administration in their decision making actions. The concept of DSS originates from a balance which lies between the data generated by computers and the judgment of humans [9].

The objective of decision support systems is to enhance the effectiveness of the decisions. This is a great tool which can analyze the sales data and provide further predictions. The purposes which can be established from the DSS are such as, analysis, optimization, forecasting, and simulation. The research subjects who use DSS for decision making, come-up with more effective decisions than those who did not use it [10]. Nowadays, DSS is contributing a meaningful role in many fields such as medical diagnosis, business and management, investment portfolios, command and control of military units, and statistics. DSS uses statistical data to overcome the deficiencies and helps the decision makers to take the right decision. Data mining (DM) plays vital role to support the Decision support systems which are based on the data obtained from the data mining models: rules, patterns, and relationships. Mining Information is the procedure of discovering, modeling, and selecting higher amount of information to clarify and find unidentified patterns. The objective of data mining in decision support systems is to suggest a tool which is easily accessible for business users to analyze the data mining models [11-13].

Machine learning can be divided into two main categories, namely supervised learning and unsupervised learning. In supervised learning, the output of the algorithm is known and we use the input data to predict the output. Examples of supervised learning are regression and classification. In contrast, in unsupervised learning, we only have input data and the corresponding output variables are not selected. An example of unsupervised learning is clustering. Feature selection is the process of selecting a subset of relevant variables from the model. Identify the most important attributes that help predict the output. By using this technique, we can reduce the curse of dimensionality, avoid model over fitting, and shorten training time. In this way, a minimalist model can be made with the least amount of parameters and good explanatory and predictive capabilities [14]

In [14], familiar approaches of SVM, DT, RF, and ANN classification were implemented. For reducing the dimension, FS optimal subsets LR, LASSO, RF methods are used. The aim is to verify the predictive performance and accuracy of this method afterward FS. Sun et al. [15] describe a mining statistics method to extract beneficial data from a current Portuguese bank telemarketing operation attained from the UCIKDD ML Dataset. For validating predictive performance, the presented method is related to common classifier model that includes NB, SVM, and DT.

Che et al. [16] aim at the complicated higher dimension non-linear features of the factor affecting the attainment rates of telemarketing, a t-SNE FE model, and later takes the extracted lower dimension feature as input, utilize non-linear SVM for training and predictions. The result shows that the bank phones depending upon t-SNESVM presented in these studies. The marketing predictive models have generalization and learning capability that could give specific decision making references for the bank and another business for achieving accuracy marketing. Fig. 1 illustrates the different applications of ML techniques.

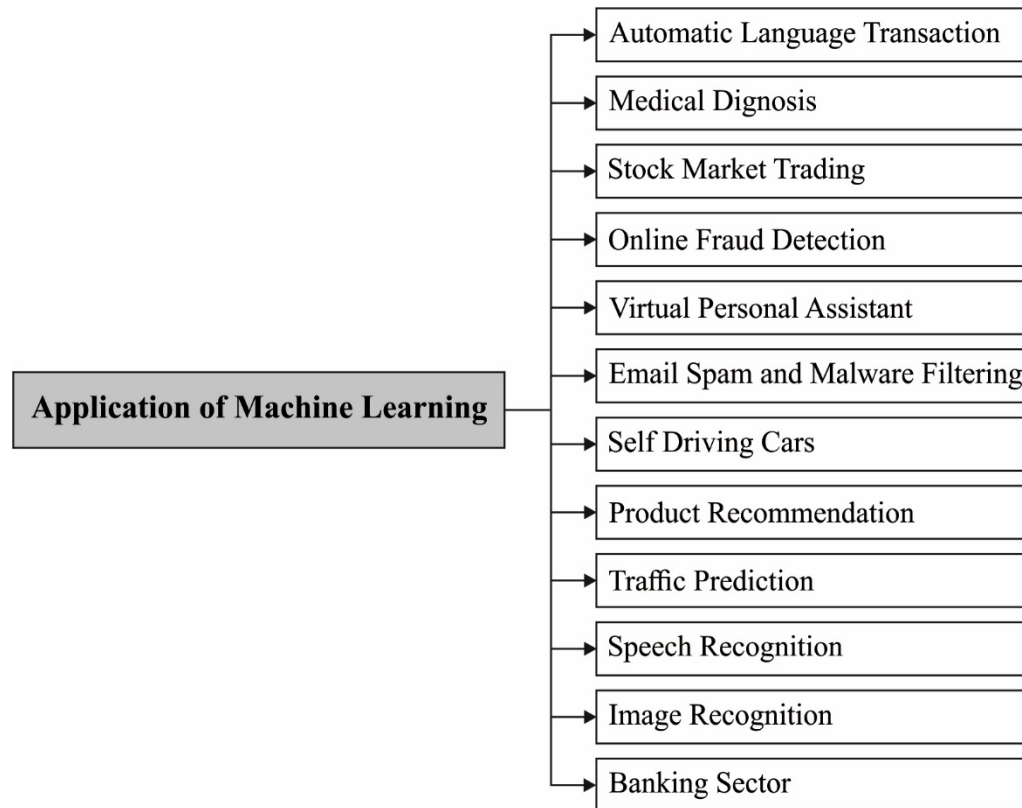


Fig. 1. Different applications of ML models

Kim et al. [17] proposed the DCNN framework which predicts either a provided user is appropriate for banks telemarketing/not. Various layers, learning rate, first values of node, and another variable which must be fixed for constructing DCNN are proposed and analyzed. Ghatashah et al. [18] aimed at enhancing the efficiency of forecasting the will of bank clients for applying to term deposits in extremely imbalanced dataset. It proposes improved ANN methods (viz., cost sensitive) for mitigating the drastic effect of extremely imbalanced information, with no distortion of novel information samples. The produced model is validated, evaluated, and subsequently associated with distinct ML methods. Realtime telemarketing datasets from Portuguese banks are employed in each experiment.

In [19], DL approaches (LSTM, GRU, and SimpleRNN) are employed for predicting the opportunity of contributing to deposits afterward the customers are named within the possibility of the bank's telemarketing operation. Executed methods were verified using the datasets and experiment results are interpreted as well as compared. For improving the attained performance level distinct methods are used to the datasets. Due to the unbalanced structures of the employed datasets, SMOTE method has been employed for reaching high precise results.

This paper presents an effective parameter tuned ML based DSS (PTML-DSS) for bank telemarketing sector. The proposed PTML-DSS technique follows a three-level process namely preprocessing, classification, and parameter optimization. Initially, the marketing data is preprocessed to get rid of unwanted information. In addition, gradient boosting decision tree (GBDT) based classifier model is used to classify the data. Besides, firefly algorithm (FFA) is applied for tuning the parameters involved in the GBDT model. In order to verify the improved performance of the PTML-DSS technique, a series of simulations were performed and the results are inspected under varying aspects.

2 Literature Review

This section explains the previous research work which has been already done in classification using ML techniques.

In [23], the aim of this study is to find a model that can improve the success rate of bank telemarketing. The statistical data mining techniques used in his study include support vector machines (SVM), decision trees (DT), and naive Bayes. Use receiver operating characteristic (ROC) curves to check the performance of these models.

Of all these statistical techniques, SVM provides the most effective results. Regarding attributes, call duration is the most relevant feature, indicating that the longer the call duration, the higher the success rate. After the contact month, the attributes are the number of contacts, the number of days since the last contact, the result of the last contact, and the duration of the first contact.

In [24], objective of the study was to predict the success of bank telemarketing. The data set they used in the study consisted of 150 attributes and was a complete data set from 2008 to 2013. They compared 4 data mining models, namely Logistic Regression (LR), Decision Trees and Data Mining Machines. Support Vectors and Neural Networks (NN). The neural network obtained the best results, and the decision tree showed that the probability of a successful call was higher. Statistical learning algorithms have been successfully used in many classification research problems.

For example, [25] conducted a research to find out the fault diagnosis system for reciprocating compressors. Reciprocating compressors are extensively used in petroleum industry. Data was taken from Oil Corporation (5 years operational data) and uses the Support Vector Machine to analyze it. They come up with the results that SVM accurately predicts the 80% right classification to find the potential faults in compressor.

Kim et al. [26] studied a deep convolution neural network (DCNN) designed to predict the success of bank telemarketing. They analyzed 16 attributes related to finance. The eight numerical attributes include age, balance, duration of last contact, number of contacts, number of days since last contact, number of contacts before a specific activity, and date and month of last contact, while the eight nominal attributes include employment, Marital status, education, loan default status, housing, loan amount and communication method (mobile phone or telephone). Taking into account factors such as the number of layers, the learning rate, the initial value of the node and other parameters, the DCNN-based model was checked in several structural experiments. Compared with other traditional ML models, their proposed model shows higher performance.

In [12], the purpose of this research is to investigate the use of telemarketing practices to promote the long-term bank deposit needs of potential bank customers. The research explored the demand for long-term bank deposits using various machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), Gaussian Naive Bayes (GNB), Decision Tree (DT) and Logistic Regression (LR). Data sets related to direct marketing activities (telephones) of Portuguese banking institutions are considered for analysis. The results confirm that the LR model provides an accuracy of 92.48%, which is the best model for predicting potential customers interested in long-term deposits through telemarketing. The research results also provide banks with valuable information so that they can make telemarketing policy decisions to their current and potential bank customers on the success of bank deposits.

The Proposed DSS for Bank Telemarketing

The proposed PTML-DSS technique follows a three-level process namely preprocessing, GBDT based classification, and FFA based parameter optimization. The detailed working of these subprocesses are offered in the following:

Design of GBDT Model

The GBDT is going to combined techniques on the fundamental of different weak classifiers. Therefore, the last results are biased by weighting technique. The weak classifiers are implemented Regression DT in enormous iterations. All the weak classifiers endure training based on former weak classifier error where GBDT has attained a classifier target by reducing the error in trained procedure. Essentially, the GBDT has a kind of enhancing ML technique model. Therefore, the inclusion technique and forward stage-wise techniques are implemented. The provided equation issues the m th weak classifier, where M refers the amount of classifiers and θ_m depicts a variable of particular classifications.

$$f_M(x) = \sum_{m=1}^M T(x, \theta_m) \quad (1)$$

The forward stagewise technique means iterative approach in front to back, learning 1 weak classification with parameter. The learning procedure of recent weak classifier is dependent upon the weak classifications which are trained in the beginning. Therefore, m th step of boosting classifier was demonstrated as provided below.

$$f_m(x) = f_{m-1}(x) + T(x, \theta_m) \quad (2)$$

Loss Function:

$$L(f_m(x), y) = L(f_{m-1}(x) + T(x, \theta_m), y) \quad (3)$$

To trained model, train a classification $T(x, \theta_m)$ from every iteration that makes sure Loss Function $L(f_m(x), y)$ that exists restricted.

Generally, the GBDT is dependent upon Negative binomial logarithm likelihood $\log(1 + e^{-2yF})$, $y \in (-1, +1)$. It can be easy to optimized loss function, but it can be complicated for optimizing a general function utilizing GD technique [20]. Afterward, Freidman devised a technology which is implemented negative gradient of loss function for fitting the Classification and Regression Tree. The specific approach was demonstrated as:

Assume that N instances sets and initiate weak classifications:

$$f_0(x) = \arg \min_c \sum_{i=1}^N L(y_i, c) \quad (4)$$

The t th round i th sample loss's Negative GD is showcased as:

$$r_{ti} = \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{t-1}(x)} \quad (5)$$

This can be employed (x_i, r_{it}) $i = 1, 2, \dots, m$ to fitting t th Decision RT, the equivalent leaf nodes is R_{tj} , $j = 1, 2, \dots, J$. J denotes an amount of leaves node. The instance of nodes seek to result values c_{tj} for limiting the loss function.

$$c_{tj} = \arg \min \sum_{x_i \in R_{tj}} L(y_i, f_{t-1}(x_i) + c) \quad (6)$$

Recreate round and achieve a powerful classification method.

$$f_t(x) = f_{t-1}(x) + \sum_{j=1}^J c_{tj} I(x \in R_{tj}) \quad (7)$$

Lastly, final robust classifiers are given by,

$$f(x) = f_M(x) = f_0(x) + \sum_{t=1}^M \sum_{j=1}^J c_{tj} I(x \in R_{tj}) \quad (8)$$

In case of classification issue, the Negative binomial logarithm likelihood $L(y,F) = \log(1 + e^{-2yF})$, $y \in (-1, +1)$ has been applied. Thus, negative gradient error is illustrated as;

$$r_{ti} = \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{t-1}(x)} \quad (9)$$

$$= \frac{y_i}{1 + \exp(y_i f(x_i))}$$

Thus, Best Residual Error values for all nodes are given below:

$$c_{tj} = \arg \min \sum_{x_i \in R_{tj}} \log(1 + \exp(-y_i(f_{t-1}(x_i) + c))) \quad (10)$$

Followed by, majorization that applies approximate measures as a substitute.

$$c_{tj} = \frac{\sum_{x_i \in R_{tj}} r_{ti}}{\sum_{x_i \in R_{tj}} |r_{ti}|(1 - |r_{ti}|)} \quad (11)$$

Design of FFA for Parameter Optimization

To finely tune the parameters involved in the GBDT model, the FFA [21] is employed and raises the classification outcomes. In FA, a Meta heuristic method has been proposed by Yang in 2008 to solve optimization problems. Indeed, all fireflies equivalent to all optimum solutions would have their brightness equivalent to the FF of the optimum solution. The firefly actions using dark brightness would search and attain to another firefly produce high brightness levels are related with the recently made solution depending on older solution using an optimal FF. As a result, in FA, all older solutions could be recently made each time based on the assessment of its brightness with another one. Consequently, individual solutions of all older solutions are saved according to the assessment of FF. Assume that all the solutions $i(X_i)$ are a location of firefly i in the present iterations. While the FF of solution i is high compared to other solutions j , the distance among the fireflies $i&j$ is attained as follows.

$$r_{ij} = \sqrt{(X_i - X_j)^2}. \quad (12)$$

Later, the upgraded distance is used to be replaced to other (13) for calculating a novel attraction. Afterward, the novel location for i th considers fireflies could be defined as equivalent to the formation of novel i th solutions. The process of making novel solutions is executed by (14).

$$\beta = \beta_0 e^{-\gamma r_{ij}^2}, \quad (13)$$

$$X_{ijnew} = X_i + \beta \cdot rand \cdot \Delta X_{ij} + rand, \quad (14)$$

Whereas $rand$ denotes an arbitrary value of solutions i and β_0 indicates the attraction at zero distance and is generally fixed to one. X_j represents a solution containing low FF compared to X_i ; and ΔX_{ij} indicates an upgraded step size calculated as the succeeding method.

$$\Delta X_{ij} = (X_j - X_i). \quad (15)$$

Eqs (12), (13), & (14) continue to be defined for the i th solutions till there are no other solutions which contain low FF.

Briefly, for all solutions i , they could contain whether one or high than one new solution or no new solution depending on the fitness comparison between solution i and other solutions amongst the present populations. The declaration could be determined based on the succeeding equation.

$$X_i^{new} = \begin{cases} X_i, & \text{if } X_i \text{ is } X_{Gbest}, \\ X_{iGbest}^{new}, & \text{if } X_j \text{ is } X_{Gbest}, \\ X_{ij}^{new} \text{ with } FT_{best}, & \text{otherwise.} \end{cases} \quad (16)$$

For the primary terms in (16), when the deliberated solutions i is to a globally optimal solution, there would be no other novel solutions made. Next, there would be only individual novel solutions, X_{iGbest}^{new} is made when the deliberated solutions are the succeeding optimal solutions, and X_j indicates the global optimal solutions X_{Gbest} amongst the populations. For other instance, X_i denotes the 3rd optimal solutions or worsen compared to 3rd optimal solutions and it is the worst solutions, there would be from 2 novel solutions to $(N_{pop} - 1)$ novel solutions X_{ij}^{new} . Regarding this, the group of novel solution i would be calculated using FF value assessment and the optimal ones using the lowermost fitness (FT_{Gbest}) is maintained when other is dismissed.

Performance Validation

The performance of the PTML-DSS technique is investigated using a benchmark dataset from Kaggle repository [22]. The dataset includes 11162 instances with 16 attributes. The dataset holds 2 classes with 5289 instances under class 1 and 5873 instances under class 2.

Table 1 and Figs. 2-3 investigates the classification performance of the PTML-DSS technique on the applied benchmark dataset.

The precision analysis of the PTML-DSS technique with existing techniques reported that the DT model has showcased worse outcomes with the lower precision of 0.8530. At the same time, the NB-Tree technique has attained

somewhat increased precision of 0.8580. Next to that, the RF technique has accomplished moderately reasonable performance with a precision of 0.8600. However, the PTML-DSS technique has accomplished superior performance with a precision of 0.9045.

Table 1 Performance Evaluation of Different Classifiers on Applied Dataset

Methods	Precision	Recall	Accuracy	F-Score	Kappa
Proposed Method	0.9045	0.8990	0.8854	0.8705	0.8043
Random-Forest	0.8600	0.8580	0.8576	0.8580	0.7154
NB-Tree	0.8580	0.8560	0.8563	0.8560	0.7127
Decision Tree	0.8530	0.8500	0.8499	0.8500	0.7003

The recall analysis of the PTML-DSS approach with recent algorithms stated that the DT manner has showcased worst outcome with the lesser recall of 0.8500. Likewise, the NB-Tree approach has gained somewhat improved recall of 0.8560. In addition, the RF method has accomplished moderately reasonable performance with the recall of 0.8580. But, the PTML-DSS technique has accomplished higher efficiency with the recall of 0.8990.

The accuracy analysis of the PTML-DSS approach with existing methods described that the DT technique has exhibited worse outcomes with the lower accuracy of 0.8499. At the same time, the NB-Tree technique has achieved slightly improved accuracy of 0.8563. Next to that, the RF technique has accomplished moderately reasonable efficiency with an accuracy of 0.8576. Eventually, the PTML-DSS methodology has accomplished maximum performance with an accuracy of 0.8854.

The F-score analysis of the PTML-DSS algorithm with state-of-art techniques reported that the DT model has outperformed poor outcomes with the lower F-score of 0.8500. Concurrently, the NB-Tree system has achieved somewhat increased F-score of 0.8560. Similarly, the RF algorithm has accomplished moderately reasonable performance with an F-score of 0.8580. Finally, the PTML-DSS manner has accomplished increased effectiveness with an F-score of 0.8705.

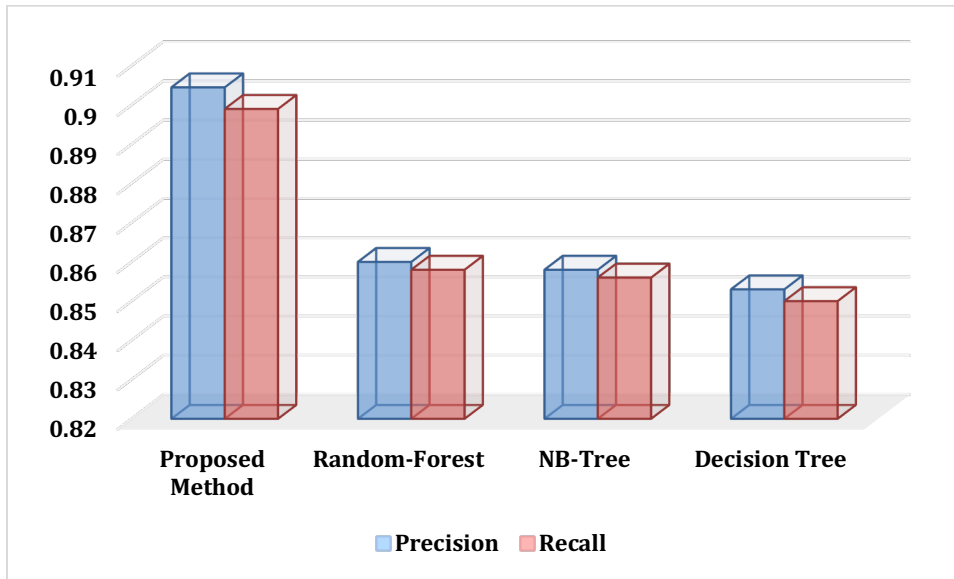


Fig. 2. Precision and recall analysis of PTML-DSS model

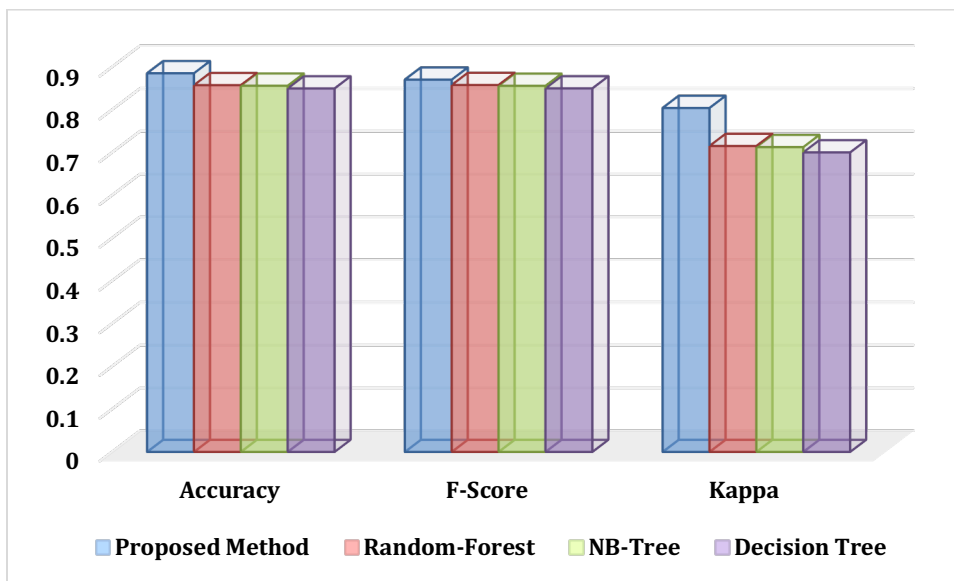


Fig. 3. Comparative analysis of PTML-DSS model with existing techniques

The kappa analysis of the PTML-DSS method with recent algorithms stated that the DT technique has portrayed least result with the minimum kappa of 0.7003. Simultaneously, the NB-Tree manner has attained somewhat increased kappa of 0.7127. Followed by, the RF approach has accomplished moderately reasonable performance with the kappa of 0.7154. Lastly, the PTML-DSS algorithm has accomplished higher performance with the kappa of 0.8043.

Conclusion

This paper has presented a new PTML-DSS technique for bank telemarketing sector. The proposed PTML-DSS technique follows a three-level process namely preprocessing, classification, and parameter optimization. Initially, the marketing data is preprocessed to get rid of unwanted information. In addition, GBDT based classifier model is used to classify the data. Besides, FFA is applied for tuning the parameters involved in the GBDT model. In order to verify the improved performance of the PTML-DSS technique, a series of simulations were performed and the results are inspected under varying aspects. The resultant values reported the improved performance of the PTML-DSS technique over the other techniques. As a part of future extension, the classification performance of the PTML-DSS technique can be improvised by the use of clustering approaches.

References

- [1] Mohammed, A.A., Olalere, M. and Mohammed, A.I., 2019. Comparative Analysis of Performance of Different Machine Learning Algorithms for Prediction of Success of Bank Telemarketing.
- [2] Borugadda, P., Nandru, P. and Madhavaiah, C., 2021. Predicting the Success of Bank Telemarketing for Selling Long-term Deposits: An Application of Machine Learning Algorithms. *St. Theresa Journal of Humanities and Social Sciences*, 7(1), pp.91-108.
- [3] Desai, R., 2021. Performance Enhancement of Hybrid Algorithm for Bank Telemarketing. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(8), pp.2703-2711.
- [4] Khan, W.A., Chung, S.H., Awan, M.U. and Wen, X., 2019. Machine learning facilitated business intelligence (Part I): Neural networks learning algorithms and applications. *Industrial Management & Data Systems*.
- [5] Hein, L., White, L., Miles, A. and Roberts, P., 2018. Analysing the impacts of air quality policies on ecosystem services; a case study for Telemark, Norway. *Journal of environmental management*, 206, pp.650-663.
- [6] Hosseini, S., 2021. A decision support system based on machined learned Bayesian network for predicting successful direct sales marketing. *Journal of Management Analytics*, 8(2), pp.295-315.
- [7] Serrano-Silva, Y.O., Villuendas-Rey, Y. and Yáñez-Márquez, C., 2018. Automatic feature weighting for improving financial Decision Support Systems. *Decision Support Systems*, 107, pp.78-87.
- [8] Asare-Frempong, J. and Jayabalan, M., 2017, September. Predicting customer response to bank direct telemarketing campaign. In *2017 International Conference on Engineering Technology and Technopreneurship (ICE2T)* (pp. 1-4). IEEE.
- [9] Moro, S., Cortez, P. and Rita, P., 2018. A divide-and-conquer strategy using feature relevance and expert knowledge for enhancing a data mining approach to bank telemarketing. *Expert Systems*, 35(3), p.e12253.
- [10] Farooqi, R. and Iqbal, N., 2019. Performance evaluation for competency of bank telemarketing prediction using data mining techniques. *International Journal of Recent Technology and Engineering*, 8(2), pp.5666-5674.
- [11] Jin, W. and He, Y., 2019, April. Three data mining models to predict bank telemarketing. In *IOP Conference Series: Materials Science and Engineering* (Vol. 490, No. 6, p. 062075). IOP Publishing.
- [12] Borugadda, P., Nandru, P. and Madhavaiah, C., 2021. Predicting the Success of Bank Telemarketing for Selling Long-term Deposits: An Application of Machine Learning Algorithms. *St. Theresa Journal of Humanities and Social Sciences*, 7(1), pp.91-108.
- [13] Ghatasheh, N., Faris, H., AlTaharwa, I., Harb, Y. and Harb, A., 2020. Business analytics in telemarketing: cost-sensitive analysis of bank campaigns using artificial neural networks. *Applied Sciences*, 10(7), p.2581.
- [14] Asif, M., 2018. Predicting the Success of Bank Telemarketing using various Classification Algorithms.
- [15] Sun, Y., Cheng, X., Li, H., Lin, Y., Sung, J. and Wang, A., 2019. Prediction of Bank Customer Telemarketing Success Using Data Mining Models. *ADRRI Journal of Engineering and Technology*, 4(5), pp.18-32.
- [16] Che, J., Zhao, S., Li, Y. and Li, K., 2020. Bank Telemarketing Forecasting Model Based on t-SNE-SVM. *Journal of Service Science and Management*, 13(3), pp.435-448.

- [17] Kim, K.H., Lee, C.S., Jo, S.M. and Cho, S.B., 2015, November. Predicting the success of bank telemarketing using deep convolutional neural network. In 2015 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR) (pp. 314-317). IEEE.
- [18] Ghatasheh, N., Faris, H., AlTaharwa, I., Harb, Y. and Harb, A., 2020. Business analytics in telemarketing: cost-sensitive analysis of bank campaigns using artificial neural networks. *Applied Sciences*, 10(7), p.2581.
- [19] Turkmen, E., 2021, February. Deep Learning Based Methods for Processing Data in Telemarketing-Success Prediction. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1161-1166). IEEE.
- [20] Rao, H., Shi, X., Rodrigue, A.K., Feng, J., Xia, Y., Elhoseny, M., Yuan, X. and Gu, L., 2019. Feature selection based on artificial bee colony and gradient boosting decision tree. *Applied Soft Computing*, 74, pp.634-642.
- [21] Yang, X.S. and He, X.S., 2018. Why the firefly algorithm works?. In *Nature-Inspired Algorithms and Applied Optimization* (pp. 245-259). Springer, Cham.
- [22] <https://www.kaggle.com/janiobachmann/bank-marketing-dataset>.
- [23] Moro, S., Laureano, R. & Cortez, P. 2011. Using data mining for bank direct marketing: An application of the crisp-dm methodology, in 'Proceedings of European Simulation and Modelling Conference-ESM'2011', Eurosis, (pp. 117–121).
- [24] Moro, S., Cortez, P. & Rita, P. 2014. 'A data-driven approach to predict the success of bank telemarketing', *Decision Support Systems* 62, 22–31.
- [25] Qi, G., Zhu, Z., Erqinhu, K., Chen, Y., Chai, Y. & Sun, J. 2018. 'Fault-diagnosis for reciprocating compressors using big data and machine learning', *Simulation Modelling Practice and Theory* 80, 104–127.
- [26] Kim, K.H.; Lee, C.S.; Jo, S.M.; Cho, S.B., 2015. Predicting the success of bank telemarketing using deep convolution neural network. In *Proceedings of the 2015 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, Fukuoka, pp. 314–317.
- [27] Gu, J.; Na, J.; Park, J.; Kim, H., 2021. Predicting Success of Outbound Telemarketing in Insurance Policy Loans Using an Explainable Multiple-Filter Convolutional Neural Network. *Appl. Sci.*, 11, 7147.