The Research on Dynamics in Phone Sales Marketing Campaign Based on Machine Learning

Zhuoye Lai*

Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

*Corresponding author: z25lai@uwaterloo.ca

Abstract. Phone market campaign is one of the selling methods that banks use to attract new term deposits. Identifying attributes of customers is essential to increase the successful rate of a phone marketing campaign. This paper uses a dataset from a Portuguese bank where various features and attributes of customers are summarized. Four methods are used to analyze this issue, firstly decision tree method, then random forest method, K nearest neighbor and lastly logistic regression. After evaluating and reviewing their confusion matrices and generated scores, the decision has been made to choose the model of random forest as it possesses the highest mean of all metrics of classification. In conclusion, the duration of contact, age, how many days have passed since last contact, the month of the contact, and the number of contacts on one customer are deemed as more important than other attributes under a random forest classifier. The findings of this paper implicates that the Portuguese bank needs to focus more of these attributes to obtain a sustainable development.

Keywords: Machine Learning, Phone Sales Marketing, Portuguese bank

1. Introduction

Phone sales are an essential component of promoting depositing sales in many industries. Although random customers are found mainly at the very beginning of the phone calling promotion sales, specific clients are mostly preferred to be called afterwards, which require data analysis to reveal attributes of clients. Successful identification, analysis, and confirmation of attributes of clients usually play an important role in increasing the response rate of clients.

Compared to other ways of promoting sales such as internet sales and in-person sales, phone is also the easiest way that has the lowest cost way of promoting businesses. Nowadays, technology has connected individuals in various ways, but phone calls are still the most efficient ways to reach business agreements. According to Cardone (2015) from the website Entrepreneur, phones are used effectively as tools to approach potential customers, follow up with customers that are already on a sale, and finally complete sales [1]. Despite there are other methods like social medias, many require a lot of advertisement and premium fees. Thus, it is worth investigating in how to obtain more potential customers through phone calls.

Moreover, there are advantages on customers' sides to use phone consultations. For instance, some customers are afraid of information leak, given the fact that many websites ask for personal information input, and usually manually typed information are hard to analyze due to different levels of willingness of customers to give out their information. According to Li (2021), it is also noteworthy that many clients feel lazy to consult on sales when they see they will have to type personal information to proceed further [2]. Li (2021) also pointed out that there are multiple important customers features that are necessary to give credit when calling to them, which are duration of calls, callers' backgrounds, etc. The analysis on callers' features is vital to optimize in increasing sales and conducting marketing activities [2].

Although lots of industries have tried to find approaches to analyze phone calling customers' features, banks, especially commercial banks are one of the oldest business entities who specializes in improving the efficiencies of phone calling services. As technology develops further, most fields, including banking industry, have been using machine learning skills to analyze features of a large scale of customers via phone calling. Machine learning consists of a series of built-in algorithms to reveal features and characteristics of datasets in a lot of industries. Several main steps of machine

learning include data prepossessing, data recognition, data analysis, model construction, and future data prediction based on existing built data. The levels of the complication of machine learning are dependent on how deeply analysts want on these datasets. Furthermore, during the process of machine learning, interactions with human-made decision will be frequently required and consulted. Any improvement during the process can be made and then applied to the process [3].

Previous studies have used a dataset from a Portuguese bank and have praised the merits of this dataset of acting prosperously to learn behaviors of customers who may be interested in applying term deposits in the future. To be more specific, both inbound and outbound analysis for phone calls is performed jointly to determine whether a customer is more likely to be attracted on phone sales [4]. The analysis has been useful to increase sales of banks. Consequently, this paper focuses the same dataset by applying the machine learning method. There are a total of four methods that will be used to analyze this dataset which are firstly decision tree method, then random forest method, K nearest neighbor and lastly logistic regression. During the research process, comparisons are made among these four methods and the single best model will also be determined. The model that stands out will be preferred and recommended in the future for sales promotion.

The remaining of this paper is constructed as the following: Section 2 is a combination of multiple data prepossessing procedures including data recognition, data cleaning, and model building; Section 3 is the discussion of the four methods; Section 4 is the displays of simulation results after running models; Section 5 is the discussion based on the simulation results from last section; and finally, section 6 is the conclusion summarizing the whole paper.

2. Data Analysis

2.1 Data Collection

Analysis is performed on 411,89 customers of the Portuguese dataset to find out which attributes are mostly significant in responding phone call sales of its bank term deposit. In this dataset, if the customer agreed to take the term deposit, the dependent variable will be marked as a "yes", and "no" otherwise. The dataset comes from Kaggle: Kaggle.com//datasets/yufengsui/Portuguese-bank-marketing-data-set [5].

2.2 Data Recognition

This paper overviews the original dataset in Table 1. In the dataset, there are categorial values such as "job" and "marital", etc; and there are numerical values such as "duration" and "contact-num", etc; and other specific values such as "date" and "yrmth" because they are not categorical values but are stored with numbers [5].

To be more specific, "default" indicates whether the customer has defaulted on any kind of credit; "contact-num" is the number of times that this specific customer has been contacted; "p-days" represents the days that passed after last call on the customer and "999" means that this customer has never been called before; "p-outcome" shows the outcome of the previous market campaign from "success", "failure" and "nonexistent"; "date" tells the last contact date'; "yrmth" is a six-digit string showing the year and month in "YYYY-MM" form; "y" is the dependent variable marking whether the customer has subscribed a term deposit [5].

Table 1. Data Type Overview

Explanatory Variable	Data Type
age	float64
job	object
marital	object
education	object
default	object
mortgage	object
loan	object
contact_type	object
duration	int64
contact_num	int64
p_days	int64
p_outcome	object
y	object
duration_mins	float64
date	object
yrmth	int64

2.3 Data Cleaning

Data cleaning is extremely important in data prepossessing to improve the efficiency of data analysis, where it can wipe out some useless data such as "N/A" type. Moreover, before any kind of data cleaning, it makes sense that duplicated values must be removed by counting the unique values of customer identification number. In other words, there is one duplicated record from the data (Note that “date” has been transferred to a “datetime64” type of data).

From Table 2, there are a lot of missing values on certain types of data, and these kinds of data types will largely impact data efficiency. Data cleaning is legitimately a three-step process which are screening, diagnosis, and editing [6]. Inappropriate and unavailable missing values sometimes crash the modelling and will output inaccurate results from "wrong" models, which consequently generate negative reports; and that is why data cleaning is important.

Table 2. Data Missing Numbers

Explanatory Variable	Data Type	Missing Count
age	float64	500
job	object	329
marital	object	80
education	object	1725
default	object	8578
mortgage	object	985
loan	object	985
contact_type	object	0
duration	int64	0
contact_num	int64	0
p_days	int64	0
p_outcome	object	0
y	object	0
duration_mins	float64	0
date	object	0
yrmth	int64	0

2.3.1 Numerical Values Cleaning

Table 3 indicates that the "age" column from the top one clearly showed that there are missing values; and from Table 4, all missing values from "age" have been filled with the imputation of the median value and colored red to make sure that these rows of records are complete and readable. Since "age" is the only explanatory variable that needs data imputation, the median strategy is reasonable because other strategies such as mean strategy is inappropriate because the attribute of age does not support fractions. Hence, all missing values under ages are filled with the median of the age group.

Table 3. Numerical Missing

	age	duration	contact_num	p_days
45	NaN	165	4	999
66	NaN	5	1	999
258	NaN	1307	1	999
337	NaN	97	2	999
549	NaN	59	1	999

Table 4. Numerical Filling

	age	duration	contact_num	p_days
45	38	165	4	999
66	38	5	1	999
258	38	1307	1	999
337	38	97	2	999
549	38	59	1	999

2.3.2 Categorical Values Cleaning

After conducting data imputation on numerical data, this research cleans the categorical data. As listed in Table 5, which displays parts of the dataset, there are many different types of missing date, that is, different types of data own different categories of themselves. Behind this thought, the method of most frequent imputer is used. As its name suggests, under each categorical explanatory variable, missing values are imputed by the most frequent ones of themselves.

Table 5. Categorical Missing

	job	marital	education	mortgage
13	services	married	high.school	NaN
17	unemployed	married	NaN	yes
34	blue-collar	married	NaN	no
63	NaN	married	NaN	no
71	student	divorced	NaN	no

Table 6 is what is displayed after the method of most frequent method. Items colored red are the most frequently categories under each categorical variable. For example, under “education”, the most frequent type of education is “university. degree”.

Table 6. Categorical Filling

	job	marital	education	mortgage
13	services	married	high.school	no
17	unemployed	married	university.degree	yes
34	blue-collar	married	university.degree	no
63	admin	married	university.degree	no
71	student	divorced	university.degree	no

In addition, to be considered valid in the later machine learning process, all categorical data types require indicator modification before entering machine learning. At this point, the method of one-hot encoding is to be proposed. One-Hot encoding applies when there are no specific orders within categorical values, and then change all categories into binary data types, that is either one or zero. The choices of determining which categories to be one or zero are purely dependent on the exact needs from data exploration and visualization of that categorical data type. Thus, it is decided between binaries one and zero. When there are more categories within one categorical variable, multiple categories can be jointly marked as one. Categories with higher successful response rate from clients are marked as one, with the remaining to be marked as zero [7]. To be more specific, take the categorical explanatory variable "yrmth" as an example, where single month is extracted as another column to explore which months exactly have a higher response rate. Those months with a higher response rate will be marked as zero.

Figure 1 shows the response rate of each month. March, December, September, and October are the four months with relative larger response rates compared to the remaining six months. Therefore, a listed containing March, December, September, and October will be created to help make up a new column consisting of only binary values, if the client was contacted among March, December, September and October, the value in this column is one.

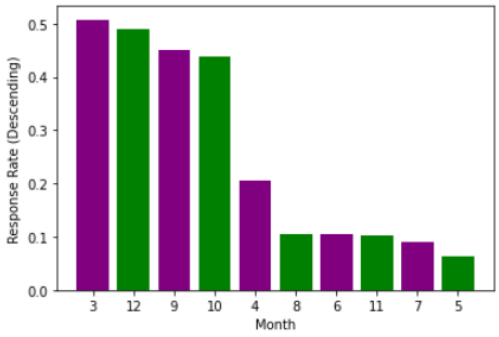

Fig 1. Response Rate Regarding Month

Similar analysis of One-Hot encoding will be performed on other categorical values to ensure that categories are mostly close to the increase response rate [7]. However, note that despite there are some categorical data types such as "loan" having almost the same response rate for all its categories, which could be seen from the Figure 2, it is still vital not to omit this categorical value. The rationale behind this thought is for the completeness of the analysis.

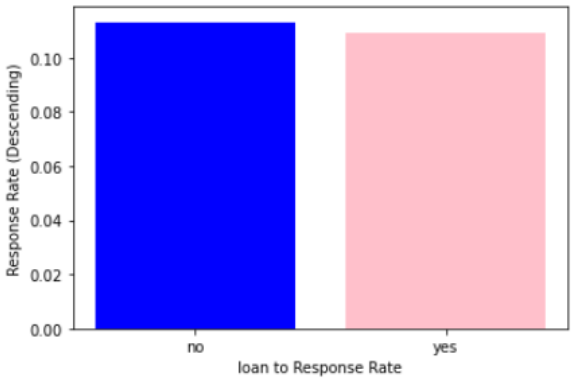

Fig 2. Response Rate Regarding Loan

Furthermore, for the simplicity of this report, after a clear thought for the purpose, it is suggested that two explanatory variables "date" and "yrmth" be omitted. Hence, that leaves two main data types which are numerical and categorical. The numerical data types consist of "age", "duration", "contact-

num", "p-days" and "duration-mins"; whereas categorical data types contain "job", "marital", "education", "default", "mortgage", "loan", "contact-type", "p-outcome" and "y".

2.4 Train and Test Split

To summarize, after a wide range of data cleaning in both numerical and categorical values, the dataset has been split into training sets and testing sets, with the test size to be 0.25 and the random state to be 666.

Explanatory variables (x) contain variables age, duration, number of contacts, days before last contact, month when in contact, whether the customer is: a student, retired, illiterate single; whether the customer has mortgages or loans; and whether the customer responded from last marketing campaign. Dependent variable (y) is whether the customer responded this marketing campaign.

3. Methods

3.1 Feature Selection

Feature importance analysis is one of the most popular data analysis approaches to extract main features from datasets. Feature importance analysis, or feature selection method, is extremely helpful when the dataset is complex and hard to compile. There are, indeed, cases where research based on multiple subsets of the original dataset is needed, when, however, the original dataset is not worthwhile to conduct a complete and over-arching statistical analysis. At these times, the feature selection method is instrumental to data analysts by reducing the amount of work in computation so that they are more focused on what have been filtered afterwards [8]. Moreover, after the feature selection method, the generated output possesses great qualities and reliabilities because most pronounced features are shown in the results of the subset, which means that the subset after the feature selection method will prevent problems such as data overfitting [9].

3.1.1 Decision tree

A decision tree is a simple representation of feature selection methods. By definition, a decision tree maps a known model or other input to their output with desired attributes or features. There are basically two types of trees: One is information node called the leaf node; and the other is called a test node which extends to other nodes. The goal is to reach a leaf node that contains predictive features of the input within a random tree, with the start as a root node to be the input [10].

3.1.2 Random forest

Random forest combines a randomized list of random decision trees which predict, and then return with feature importance analysis of explanatory variables. Random forest serves as an extraordinary technique in classifying and regressing huge datasets and problems. Generally speaking, random forest aggregates the results generated from multiple decision trees to provide a better and accurate overview of the feature importance of datasets [7].

3.2 K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is one of the supervised machine learning algorithms that is used mainly in finding the classification of a dataset. The rationale behind this algorithm is to find the most appropriate group of a data point after identifying the features of the data point and its associated groups (i.e., nearest neighbors) [11]. The value of K is dependent on the most similar number of groups, meaning that K is not best when it is smaller, nor is it greatest when it is bigger, and only most suitable K is needed to complete the process.

3.3 Logistic Regression

The logistic regression model is another well-known model used in machine learning, it does not regress on simple explanatory variables; instead, the logistic regression incorporates a special

technique called "odds", which is the ratio of the probability of the case occurring divided by the probability of the case not occurring. In fact, the natural logarithm is taken for regressors. According to LaValley & Michael (2008) [12], the formula for a simple logistics regression is

$$\ln(\text{odds}(Y)) = \beta_0 + \beta_1 X \quad (1)$$

4. Simulation results

4.1 Scoring Standards

In this section, four methods of building models based upon the clean dataset will be discussed. After featuring analysis for each method, a confusion matrix will be provided to assist in visualizing the results from test sets. There are four areas within a confusion matrix: Top-Left (True Positive/TP): amount of projected correct values that have actually occurred by the test; Bottom-Right (True Negative/TN): amount of projected incorrect values that have not actually occurred by the test; Bottom-Left (False Positive/FP): amount of projected incorrect values that have actually occurred by the test; Top-Right (False Negative/FN): amount of projected correct values that have not actually occurred by the test [9].

Under each method, several scores based on the confusion matrix will be computed. The accuracy score of a confusion matrix is the ratio of the true predictions (TP + TN) divided by the total number of observations (TP+TN+FP+FN) [9]. Moreover, the recall (sensitivity) score is defined as the ratio of true positives (TP) divided by all positive results (TP+FN) [13]; the ratio of true positves (TP) divided by all positive predictions (TP+FP) defines the precision score [13]; the specificity socre is defined as the ratio of true negatives (TN) divided by all negative results (TN+FP) [13]; the F1 score is the one of the balance scores defined as the result of 2*(precision score * recall score) divided by the summation of the recall socre and the precision score, which is used to adjust the both recall and precision score when the dataset is single and individual [13]; balanced score is the arithmatic average between sensitivity and specificity, which is used to deal with imbalanced classifiers [13]; the score of area under the curve (AUC) is interpreted as the probability of how much a classifier values a random positive result higher than a random negative result [14].

4.2 Results Based on Decision Tree

Figure 3 shows the feature importance under the supervision of a decision tree classifier. It indicates that several interested and relevant variables such as duration and age are analyzed regarding the response rate of answering phones by measuring relative importance of each explanatory variable. Moreover, Figure 3 also shows that the duration of contact, age, how many days have passed since last contact, the month of the contact, and the number of contacts on one customer are deemed as more important than other attributes or explanatory variables under a random forest classifier.

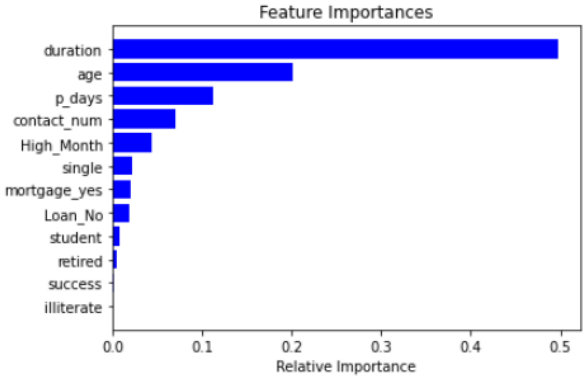

Fig 3. Feature Analysis under Decision Tree

The confusion matrix is provided in Figure 4. Hence, amount of projected incorrect values that have not actually occurred by the test is 514; amount of projected correct values that have actually occurred by the test is 8422; amount of projected incorrect values that have actually occurred by the test is 643; amount of projected correct values that have actually not occurred by the test is 693.

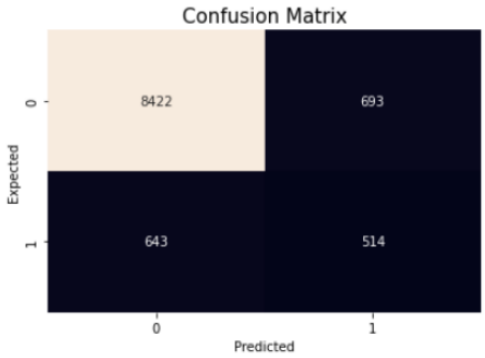

Fig 4. Decision Tree Confusion Matrix

Table 7 shows the evaluation scores for decision tree method generated from the confusion matrix. The train accuracy is 100%, while the test result is 87.0% accurate compared to the training result, which means that under this training and testing split, the decision tree model is slightly overfitting. Scores other than test AUC are all fair and satisfactory. The train AUC returns a score of 100%, while the test AUC is only 68.4%, again meaning that the decision tree classifier is overfitting.

Table 7. Evaluation Scores for Decision Tree Method

Train	Test	Recall (Sensitivity)	Specificity	Precision	Balanced	F1	Train AUC	Test AUC
1.000	0.870	0.924	0.457	0.929	0.691	0.926	1.000	0.684

4.3 Results Based on Random Forest

Figure 5 is the feature importance under the supervision of a random forest classifier [15]. It is clear to see that the duration of contact, age, how many days have passed since last contact, the month of the contact, whether the last contact was successful and the number of contacts on one customer are deemed as more important than other attributes or explanatory variables under a random forest classifier.

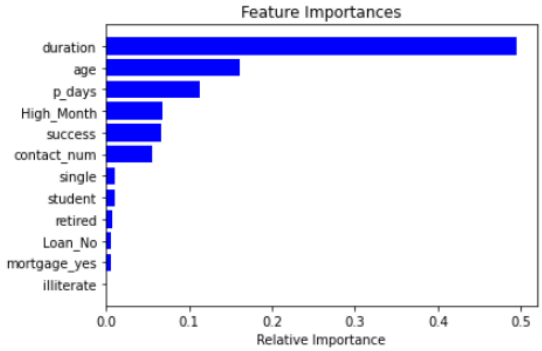

Fig 5. Feature Importance of Random Forest

The confusion matrix is provided in Figure 6. Hence, amount of projected incorrect values that have not actually occurred by the test is 494; amount of projected correct values that have actually occurred by the test is 8803; amount of projected incorrect values that have actually occurred by the test is 671; amount of projected correct values that have actually not occurred by the test is 304.

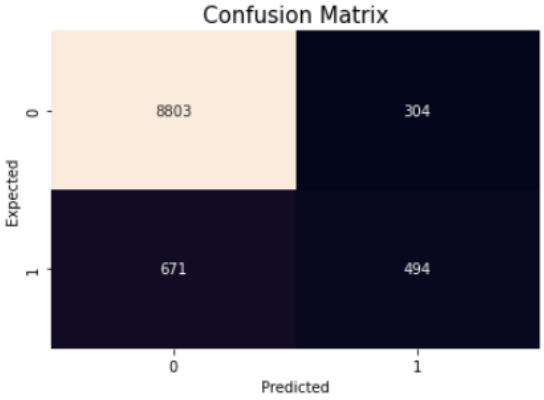

Fig 6. Confusion Matrix of Random Forest

Table 8 shows the evaluation scores for random forest method generated from the confusion matrix. It is a good sign that both the test accuracy score and test AUC score are near 90%, which is acceptable for a single model. It is also worth noting that the F1 score is near 95%, meaning that both recall (sensitivity) and precision scores are great indications of the model. The balanced accuracy score of random forest is like the decision tree method.

Table 8. Evaluation Scores for Random Forest Method

Train	Test	Recall (Sensitivity)	Specificity	Precision	Balanced	F1	Train AUC	Test AUC
0.952	0.905	0.967	0.424	0.929	0.696	0.948	0.964	0.898

4.4 Results Based on K-Nearest Neighbor

The confusion matrix for K-Nearest Neighbor is provided in Figure 7. Hence, amount of projected incorrect values that have not actually occurred by the test is 449; amount of projected correct values that have actually occurred by the test is 8881; amount of projected incorrect values that have actually occurred by the test is 708; amount of projected correct values that have actually not occurred by the test is 234.

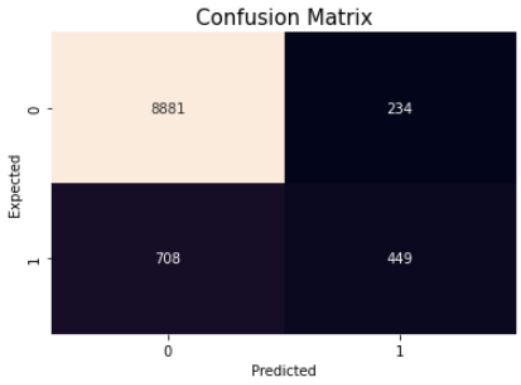

Fig 7. Confusion Matrix of K-Nearest Neighbor

Table 9 shows the evaluation scores for K-Nearest Neighbor method generated from the confusion matrix. Similarly with the random forest method, the F1 score is near 95%, meaning that both recall (sensitivity) and precision scores are great indications of the model. Compared to the random forest classifier, although the test accuracy score is also above 90%, the test AUC score is not as good as the random forest classifier, but the overall performance of the K-Nearest Neighbor classifier is better than the decision tree method. The balanced accuracy score of K-Nearest Neighbor, however, is fewer than both the decision tree method and the random forest method.

Table 9. Evaluation Scores for K-Nearest Neighbor Method

Train	Test	Recall (Sensitivity)	Specificity	Precision	Balanced	F1	Train AUC	Test AUC
0.907	0.908	0.974	0.388	0.926	0.681	0.949	0.890	0.874

4.5 Results Based on Logistic Regression

The confusion matrix for K-Nearest Neighbor is provided in Figure 8. Hence, amount of projected incorrect values that have not actually occurred by the test is 385; amount of projected correct values that have actually occurred by the test is 8897; amount of projected incorrect values that have actually occurred by the test is 772; amount of projected correct values that have actually not occurred by the test is 218.

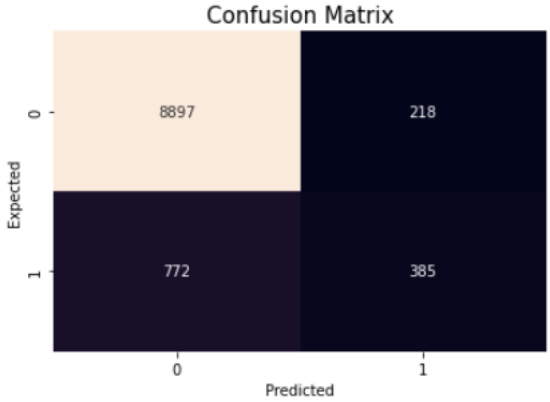

Fig 8. Confusion Matrix of Logistic Regression

Table 10 represents the scores generated from the confusion matrix are placed for K-Nearest Neighbor for further evaluation. Similarly with random forest and K nearest neighbor, the F1 score is near 95%. The test accuracy score is also above 90%, the test AUC score is not as good as the random forest classifier, but the overall performance of the logistic regression classifier is better than the decision tree method. The balanced accuracy score of logistic regression classifier, however, is fewer than all three other methods.

Table 10. Evaluation Scores for Logistic Regression Method

Train	Test	Recall (Sensitivity)	Specificity	Precision	Balanced	F1	Train AUC	Test AUC
0.904	0.904	0.976	0.333	0.920	0.655	0.947	0.896	0.896

5. Discussion

Four methods are being used to analyze the dataset, which are firstly decision tree method, then random forest method, next K nearest neighbor and lastly logistic regression The joint evaluation scores for the four methods are summarized in Table 11. All of the four models are terrific models to evaluate the attributes of a client that banks need to focus when initiating a phone marketing campaign [16]. The higher the scores are, the better the classifier is.

Table 11 summrizes all of the metrics for four classifiers such as precision and recall. Firstly, note that the goal is to find the best model with the highest accuracy scores such that if further investigation is required, the performance of the chosen model would still be satisfactory, which leads to higher testing and testing AUC accuracy scores [14]. By looking up the sum of test and test AUC columns of each method, it is found that random forest possesses the highest scores of accuracy. Therefore, based on the primary criterion, random forest is the best [13]. Furthermore, it is still important to evaluate other metrics in the table. For example, evidently the F1 score of other three methods

outperform that of decision tree and K nearest neighbor has the highest F1 score, but F1 score of random forest and logistic regression is still very close. In addition, random forest has the highest balanced accuracy score among the four methods [13].

In summary, the method of random forest is to be proposed, because it has the highest accuracy scores as well as balanced accuracy scores, in the meantime it F1 socre is not bad.

Table 11. Combination of Scores among four Methods

Train	Test	Recall (Sensitivity)	Specificity	Precision	Balanced	F1	Train AUC	Test AUC	Mean
1.000	0.870	0.924	0.457	0.929	0.691	0.926	1.000	0.684	0.719
0.952	0.905	0.967	0.424	0.929	0.696	0.948	0.964	0.898	0.770
0.907	0.908	0.974	0.388	0.926	0.681	0.949	0.890	0.874	0.752
0.904	0.904	0.976	0.333	0.920	0.655	0.947	0.896	0.896	0.738

6. Conclusion

This paper uses four methods which are firstly decision tree method, then random forest method, K nearest neighbor and lastly logistic regression to research on revealing main features of customers who are more likely to be attracted to term deposits of a Portuguese bank through machine learning.

Since the mean of random forest is the highest, the best model among decision tree, random forest, K nearest neighbor and logistic regression is random forest. In conclusion, the duration of contact, age, how many days have passed since last contact, the month of the contact, and the number of contacts on one customer are deemed as more important than other attributes under a random forest classifier [17]. The Portuguese bank may develop algorithms to better situate customers based on the research results, namely, certain customers within the importance of variables such as the duration of contact and age may require more attention to improve the successful rate of response.

However, as is mentioned above, this paper only uses four methods to determine attributes of individuals, which means that there are other methods that will necessarily produce different results. In addition, results will also vary based on different selection of training and testing data split. In this case, there may be multiple results if training and testing split ratio are different [4]. Furthermore, results from this paper may also be affected by different adoptions of data prepossessing methods. In particular, this paper uses median imputation method for numerical variables and most frequent imputers method for categorical variables. Any methods other than those two methods may yield different results [4]. Finally, it is recommended that individuals choose their methods of selection wisely according to different protocols when analyzing this dataset, as this dataset can provide other research purposes if necessary.

References

[1] Cardone, G. When It Comes to Sales, the Phone Is Your Most Powerful Tool. Entrepreneur. Entrepreneur, April 24, 2015. https://www.entrepreneur.com/article/245434.

[2] Li, C. Why Phone Calls Are Still The Most Effective Lead Generators, March 18, 2021. https://retreaver.com/blog/phone-calls-effective-lead-generation-strategy.

[3] El Naqa, I., Murphy, M.J. (2015). What Is Machine Learning?. In: El Naqa, I., Li, R., Murphy, M. (eds) Machine Learning in Radiation Oncology. Springer, Cham. https://doi.org/10.1007/978-3-319-18305-3_1

[4] Oluwabusola, Oluwaseun Esther. Applying Business Analytics in Practice to a Bank Telemarketing Dataset. University of Strathclyde, September 2015. https://local.cis.strath.ac.uk/wp/extras/msctheses/papers/strath_cis_publication_2714.pdf.

[5] Sui, A. Portuguese Bank Marketing Data Set. Kaggle, April 9, 2019. https://www.kaggle.com/datasets/yufengsui/portuguese-bank-marketing-data-set?resource=download.

[6] Broeck, Jan Van den, Solveig Argeseanu Cunningham, Roger Eeckels, and Kobus Herbst. Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities. PLOS Medicine. Public Library of Science. Accessed August 1, 2022. https://journals.plos.org/plosmedicine/article?id=10.1371%2Fjournal.pmed.0020267.

[7] Brownlee, J. Why One-Hot Encode Data in Machine Learning? Machine Learning Mastery, June 30, 2020. https://machinelearningmastery.com/why-one-hot-encode-data-in-machine-learning/.

[8] Jović A, K Brkić, and N Bogunović. A Review of Feature Selection Methods with Applications. ResearchGate, May 2015. https://www.researchgate.net/profile/Alan-Jovic/publication/308871570_A_review_of_feature_selection_methods_with_applications/links/62792a8db1ad9f66c8ae5cf6/A-review-of-feature-selection-methods-with-applications.pdf.

[9] Asif, M. Predicting the Success of Bank Telemarketing Using Various Classification Algorithms. Örebro University School of Business. Accessed August 1, 2022. https://www.diva-portal.org/smash/get/diva2:675139/FULLTEXT01.pdf.

[10] Quinlan, J. R. Learning Decision Tree Classifiers. ACM Computing Surveys 28, no. 1 (1996): 71–72. https://doi.org/10.1145/234313.234346.

[11] Song, Yang, Jian Huang, Ding Zhou, Hongyuan Zha, and C. Lee Giles. IKNN: Informative K-Nearest Neighbor Pattern Classification." Knowledge Discovery in Databases: PKDD 2007, n.d., 248–64. https://doi.org/10.1007/978-3-540-74976-9_25.

[12] LaValley, Michael P. Logistic Regression. Logistic Regression. Circulation, May 6, 2008. https://www.ahajournals.org/doi/full/10.1161/CIRCULATIONAHA.106.682658.

[13] Olugbenga, M. Balanced Accuracy: When Should You Use It? neptune.ai. neptuneblog, July 22, 2022. https://neptune.ai/blog/balanced-accuracy.

[14] Classification: Roc Curve and AUC. Machine Learning. Accessed August 27, 2022. https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc.

[15] Biau, G, and Erwan S. A Random Forest Guided Tour. TEST 25, no. 2 (2016): 197–227. https://doi.org/10.1007/s11749-016-0481-7.

[16] Wlodarczyk, K, and Kingsley I. Data Analysis of a Portuguese Marketing Campaign Using Bank Marketing ... Data Analysis of a Portuguese Marketing Campaign using Bank Marketing data Set. ResearchGate, December 5, 2019. https://www.researchgate.net/publication 339988208_Data_Analysis_of_a_Portuguese_Marketing_Campaign_using_Bank_Marketing_data_Set.

[17] Ghosh, S. How to Compare Machine Learning Models and Algorithms. neptune.ai, July 22, 2022. https://neptune.ai/blog/how-to-compare-machine-learning-models-and-algorithms.