# Term Deposit Subscription: A Machine Learning Approach

Hong Shi

December 10 2021

## 1 Introduction

A Term Deposit is a deposit that a bank or a financial institution offers at a fixed rate (often better than just opening a deposit account) and the money will be returned back at a specific maturity time. Since the money is locked during the term deposit, the bank is more flexible to invest the money in other financial products or lend the money to borrowers, yielding higher returns. Instead of simply opening deposit accounts, a bank would like to encourage their clients to subscribe for term deposits to increase its profitability.

With increasing economic pressure, financial institutions are in high demand of accurately identifying their potential clients to maintain their profitabilities. This paper analyzes a term deposit marketing campaign of a Portuguese bank. Based on client personal data, deposit campaign information, and social and economics indicators, I first explore current patterns of term deposit subscription decision. And then I use five supervised Machine Learning models (Logistic Regression, Naive Bayes, Decision Tree, Random Forest, Gradient Boosted Tree) to predict client subscription decision. Even though Random Forest algorithm performs the best in terms of its classification accuracy, Naive Bayes algorithm outperforms other algorithms in terms of its ability to capture potential clients.

The rest of this paper is organized as follows: In Data section (Section 2), I would introduce the Portuguese bank marketing dataset and preprocess the data. Next, in Exploratory Data Analysis section (Section 3), I use visualizations to explore current patterns of client subscription decision. In Machine Learning Model section (Section 4), I first introduce methods of handling dataset imbalance, build machine learning models and then evaluate model performances. Finally, in Result and Discussion section (Section 5), I conclude and suggest future directions of this study.[1]

## 2 Data

The data is about direct term deposit marketing campaigns of a Portuguese banking institution collected from May 2008 to November 2010. The data source is from Moro et al. (Moro, Cortez, and Rita 2014) and it is public available at UCI Machine Learning Repository.[2]

The data is analyzed by R (R Core Team 2020), and its packages `tidyverse` (Wickham et al. 2019), `here` (Müller 2020), `rattle` (Williams 2011). I used `bookdown` (Xie 2016) and `kableExtra` (Zhu 2020) to format the document.

### 2.1 Dataset Description

There are 41188 observations in the dataset with 20 features and 1 label column. These features belong to four major categories: (1) client personal data, (2) data related with last contact of current term deposit campaign, (3) data related with current and previous marketing campaigns, (4) data related with social and economics context indicators. Detailed description of dataset features is shown below (Table 1):

[1] Codes and data are available at the GitHub repo: https://github.com/honn-ishinn/bank_marketing.

[2] Link of the dataset: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing.

Table 1: Feature Information of Term Deposit Campaign Dataset

| | **Name** | **Datatype** | **Description** |
|---|---|---|---|
| **Client Data** | age | Numeric | Age |
| | job | Categorical | Type of Job (admin, blue-collar, entrepreneur, housemaid, management, retired, self-employed, services, student, technician, unemployed, unknown) |
| | marital | Categorical | Marital status (divorced, married, single, unknown) |
| | education | Categorical | Education level (basic.4y, basic.6y, basic.9y, high.school, illiterate, professional.course, university degree, unknown) |
| | default | Categorical | Has credit in default? (no, yes, unknown) |
| | housing | Categorical | Has housing loan? (no, yes, unknown) |
| | loan | Categorical | Has personal loan? (no, yes, unknown) |
| **Related to Last Contact of Current Campaign** | contact | Categorical | Contact communication type (cellular, telephone) |
| | month | Categorical | Last contact month of year (jan, feb, ... , nov, dec) |
| | day_of_week | Categorical | Last contact day of week (mon, tue, wed, thu, fri) |
| | duration | Numeric | Last contact duration, in seconds |
| **Related to Current and Previous Campaign** | campaign | Numeric | Total contacts performed during current campaign with the client, including last contact |
| | pdays | Numeric | Number of days after the client was last contacted from a previous campaign |
| | previous | Numeric | Number of contacts performed before current campaign with the client |
| | poutcome | Categorical | Outcome of previous campaign (failure, success, nonexistent) |
| **Social and Economoic Context Indicators** | emp.var.rate | Numeric | Quarterly indicator of employment variation rate |
| | cons.price.idx | Numeric | Monthly indicator of consumer price index |
| | cons.conf.idx | Numeric | Monthly indicator of consumer confidence index |
| | euribor3m | Numeric | Daily indicator of euribor 3 month rate |
| | nr.employed | Numeric | Quarterly indicator of number of employees |
| **Target Label** | y | Categorical | Has the client subscribed |

## 2.2 Data Preprocessing

This dataset contains no missing value so handling missing value is not needed during data preprocessing. However, some features need to be dropped or recoded as follows:

*duration*: As suggested by Mora (Moro, Cortez, and Rita 2014), feature *duration* highly affects the target label *y* that a client will certainly not subscribe to the term deposit if that last contact duration is 0 second. Besides, the contact duration of a call is not known before a call is actually performed, and client subscription is not quickly available after the call. In this paper, I would like to build a realistic binary classification model for client subscription, so feature *duration* will be dropped in advance before further analysis.

*pdays* and *previous*: Among 41188 observations, there are 39673 observations coded as 999 in *pdays* feature, meaning that around 96.3% percent of clients are the first time to receive marketing campaign from this bank. And for those clients who received previous campaign, feature *previous* suggests that around 81.1% clients have only received 1 contact before current campaign. Taking the somehow arbitrary numeric coding of *pdays* and existing high correlation between *pdays* and *previous*[3] into consideration, I decide to combine these two features into a new feature, namely *pcampaign*, in further analysis. *pcampaign* is a categorical variable that indicates whether the client received previous campaign or not (yes, no).

# 3 Exploratory Data Analysis

After preprocessing the data, data visualizations are used to explore current patterns of client subscription decision based on different existing features. Features with relative significant subscription patterns are shown as follows, while other visualizations are attached in the Appendix (Appendix A).

## 3.1 Client Data

Either younger or elder clients are more likely to subscribe to term deposit, while middle-aged are less likely to subscribe to term deposit (Figure 1).

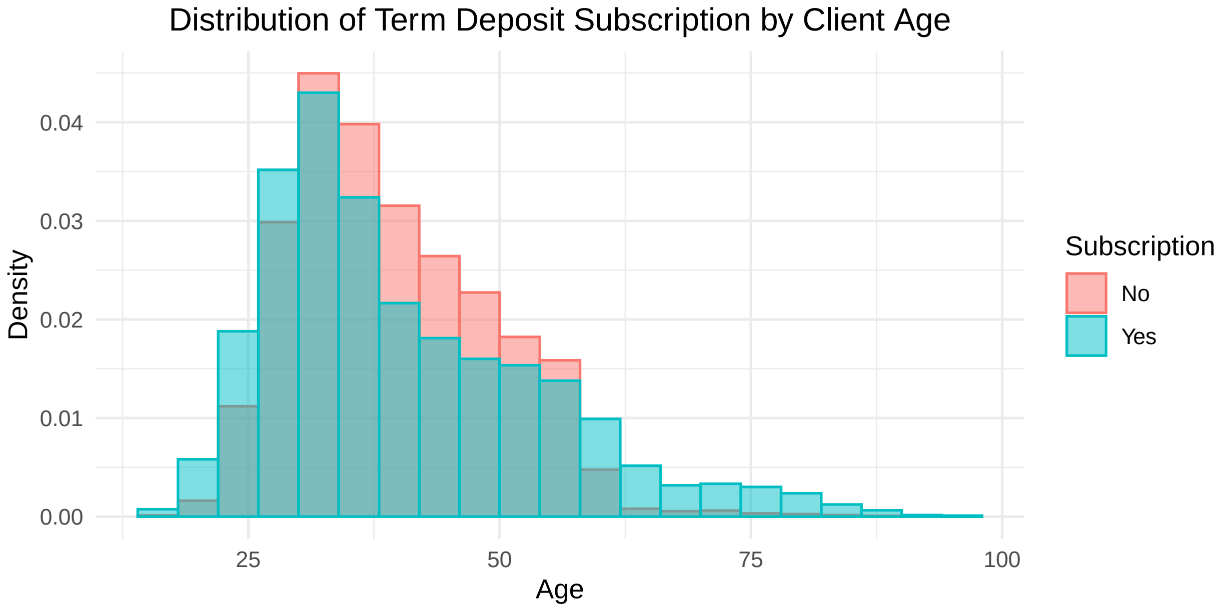


Figure 1: Distribution of Term Deposit Subscription by Client Age

[3] Pearson Correlation is -0.59, the negative sign mainly ascribes to the arbitrary coding of 999 by feature pdays

Clients who are student and retired people are more likely to subscribe to term deposit, while clients who are blue-collar and services people are less likely to subscribe to term deposit (Figure 2).

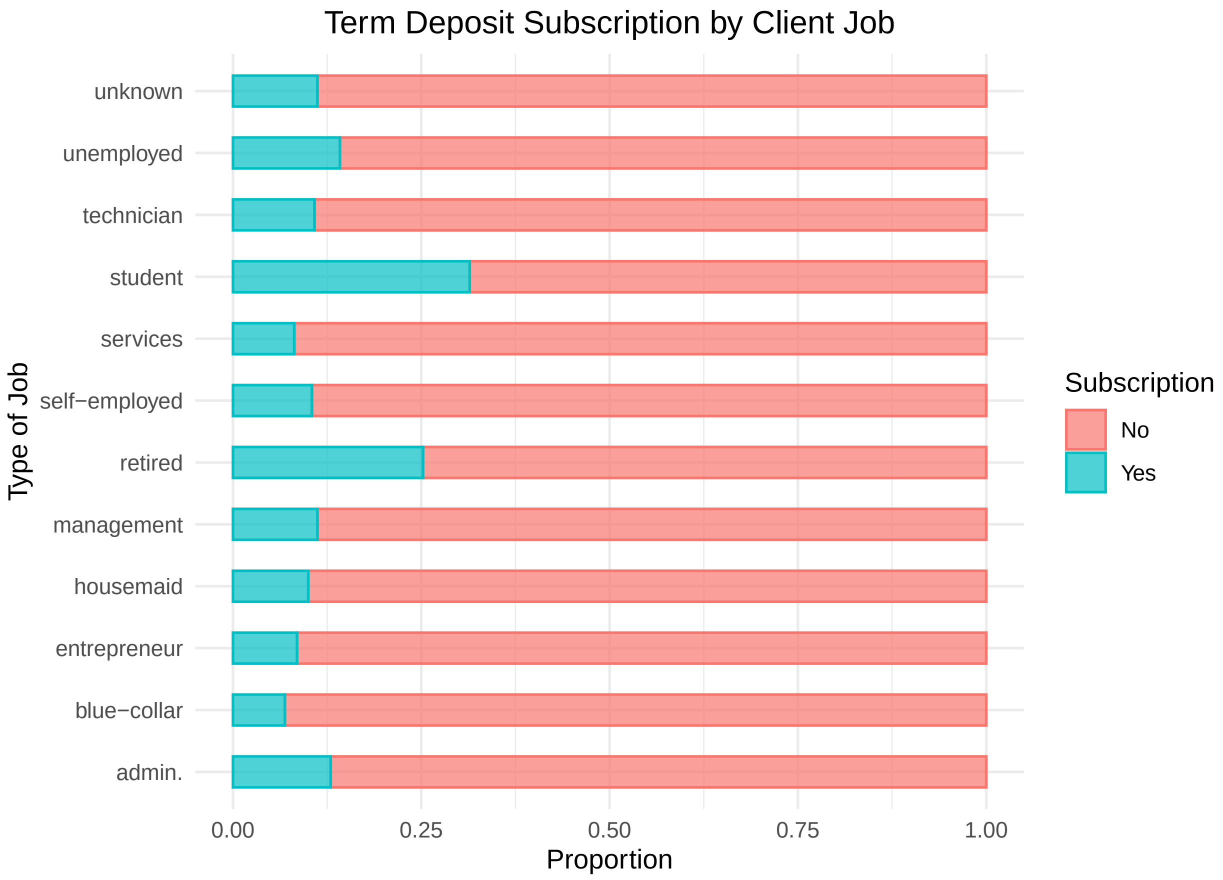


Figure 2: Term Deposit Subscription by Client Job

Illiterate clients are more likely to subscribe to term deposit (Figure 3).

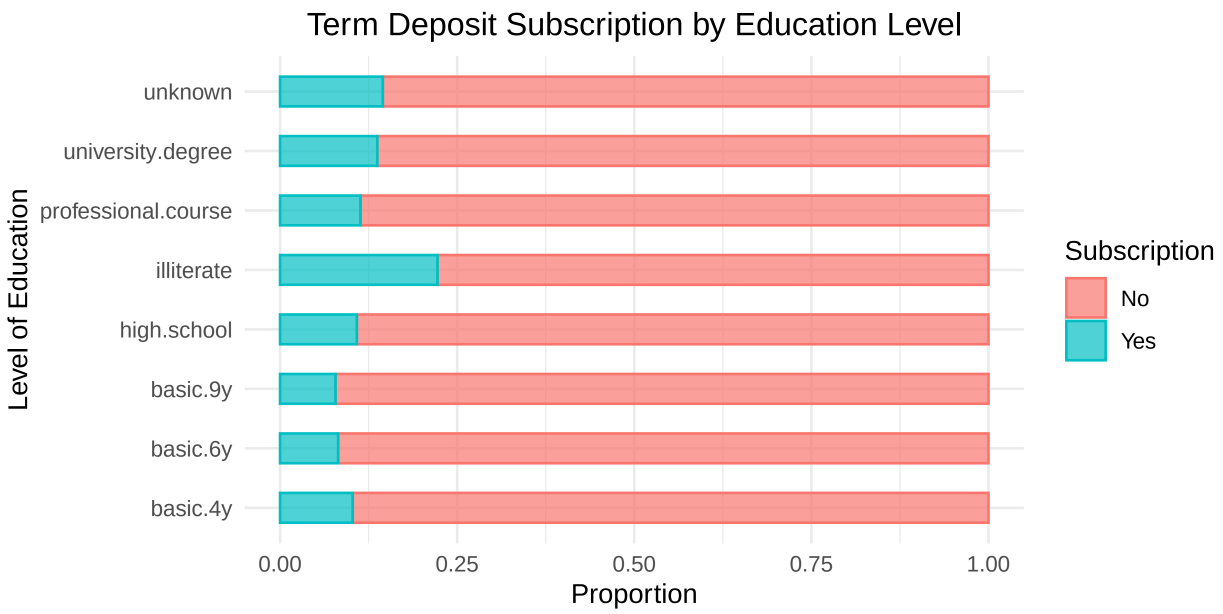


Figure 3: Term Deposit Subscription by Education Level

### 3.2 Related to Last Contact of Current Campaign

Clients last contacted through cellular devices are more likely to subscribe to term deposit (Figure 4).

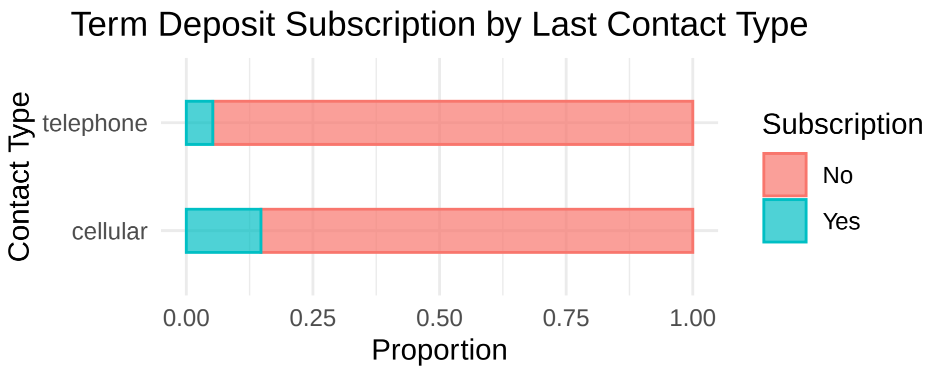


Figure 4: Term Deposit Subscription by Last Contact Type

No contacts are performed in January and February. Many last campaign contacts are from April to August, especially in May, while they do not lead to a high subscription rate. By contrast, there are less last campaign contacts within March, September, October and December, while these last contacts lead to high subscription rates. (Figure 5).

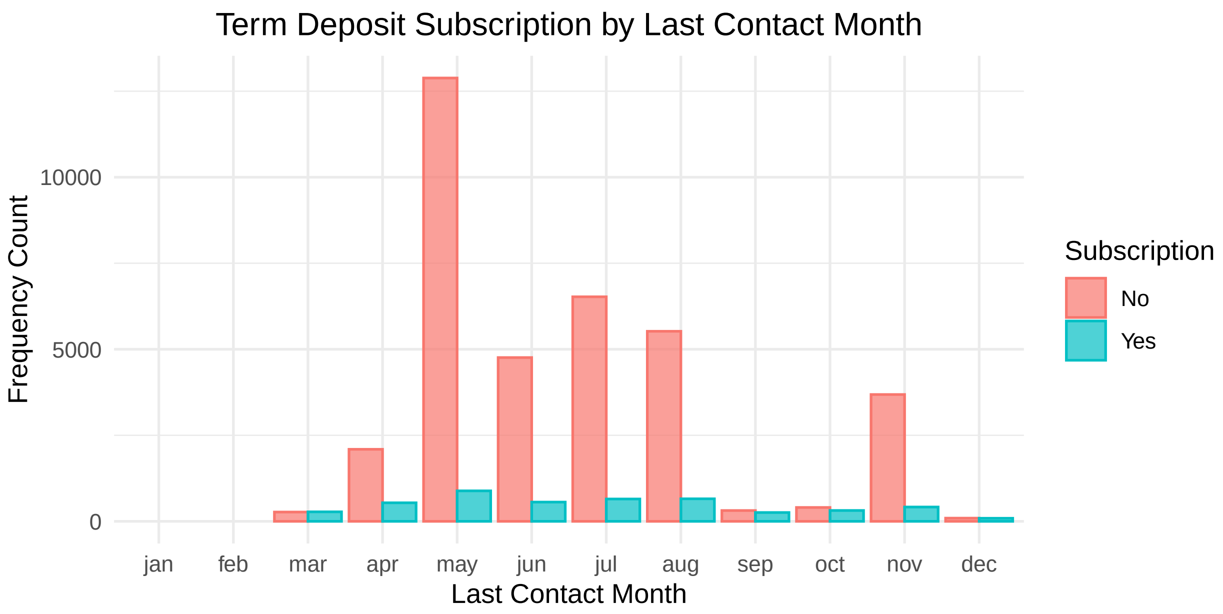


Figure 5: Term Deposit Subscription by Last Contact Month

### 3.3 Related to Current and Previous Campaign

Clients who are more previously contacted by other campaigns of this bank are more likely to subscribe to term deposit (Figure 6). For clients who received previous campaigns, clients accepted previous campaign are more likely to subscribe to term deposit (Figure 7)

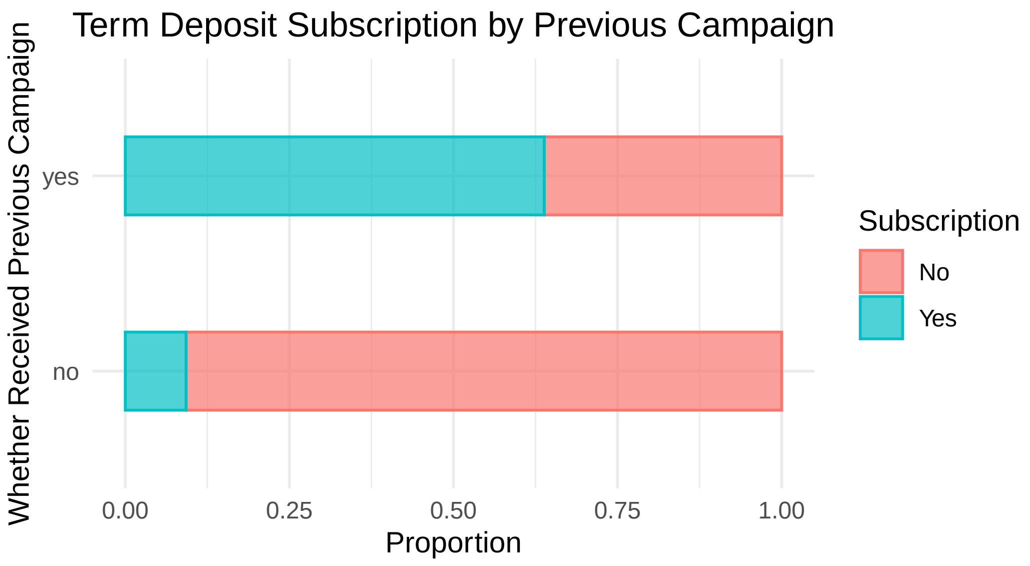


Figure 6: Term Deposit Subscription by Previous Campaign

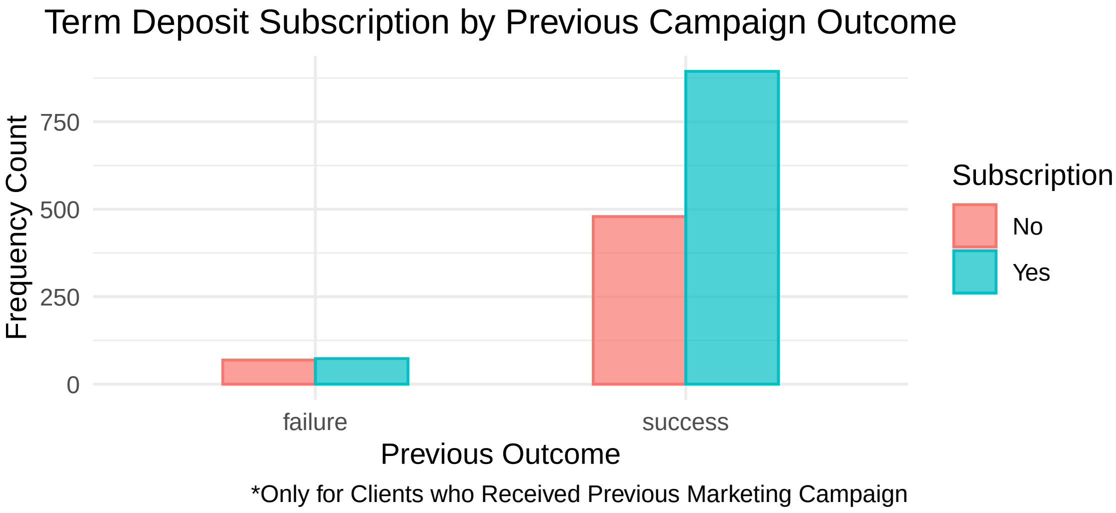


Figure 7: Term Deposit Subscription by Previous Campaign

### 3.4 Social and Economic Context Indicators

Clients are less likely to subscribe to term deposit when employment variation rate[4] is positive, and more likely to subscribe to term deposit when employment variation rate is negative (Figure 8).

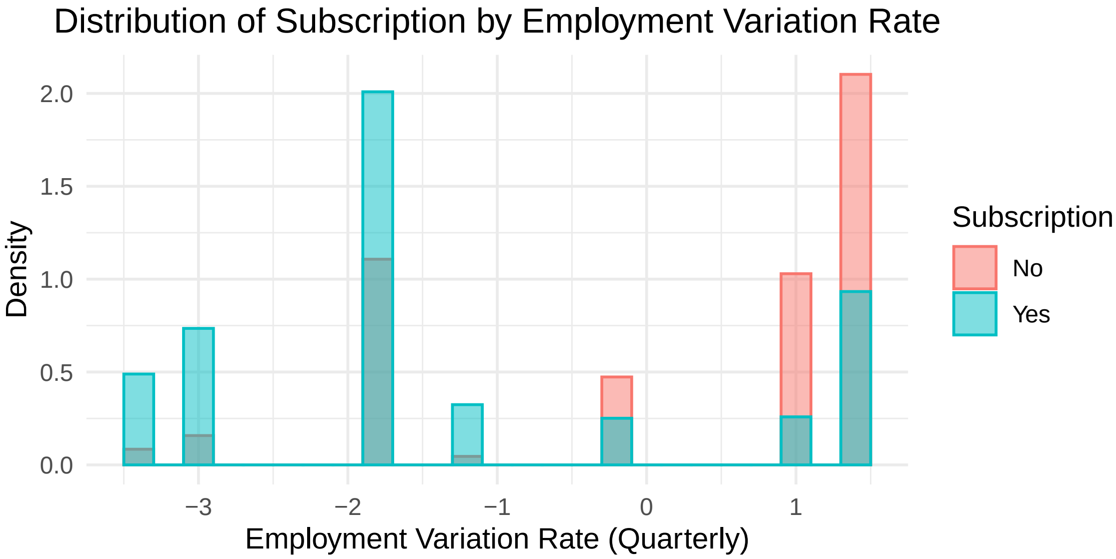


Figure 8: Term Deposit Subscription by Previous Campaign

Clients are less likely to subscribe to term deposit when Euribor 3 month rate[5] is relatively high, and more likely to subscribe to term deposit when Euribor 3 month rate is low (Figure 9).

[4] Refers to cyclical employment variation to indicate how many people are being hired or fired due to shifts in economy conditions

[5] Refers to Euro Interbank Offer Rate, a reference rate that is constructed from the average interest rate at which eurozone banks offer unsecured short-term lending on the inter-bank market

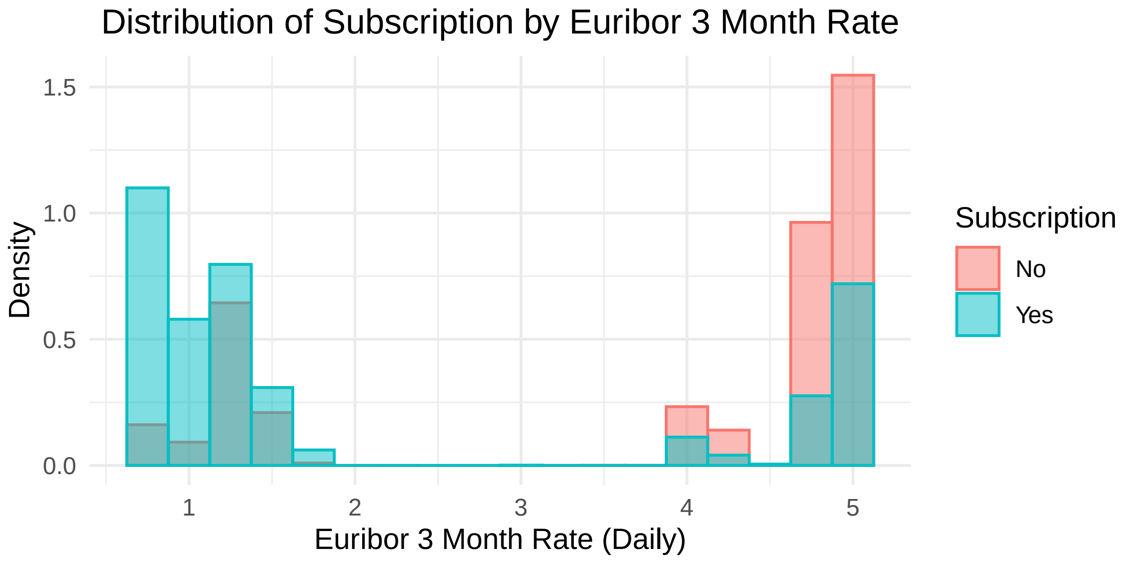


Figure 9: Term Deposit Subscription by Previous Campaign

Clients are less likely to subscribe to term deposit when number of employees[6] are relatively high, and more likely to subscribe to term deposit when number of employees are low (Figure 10).

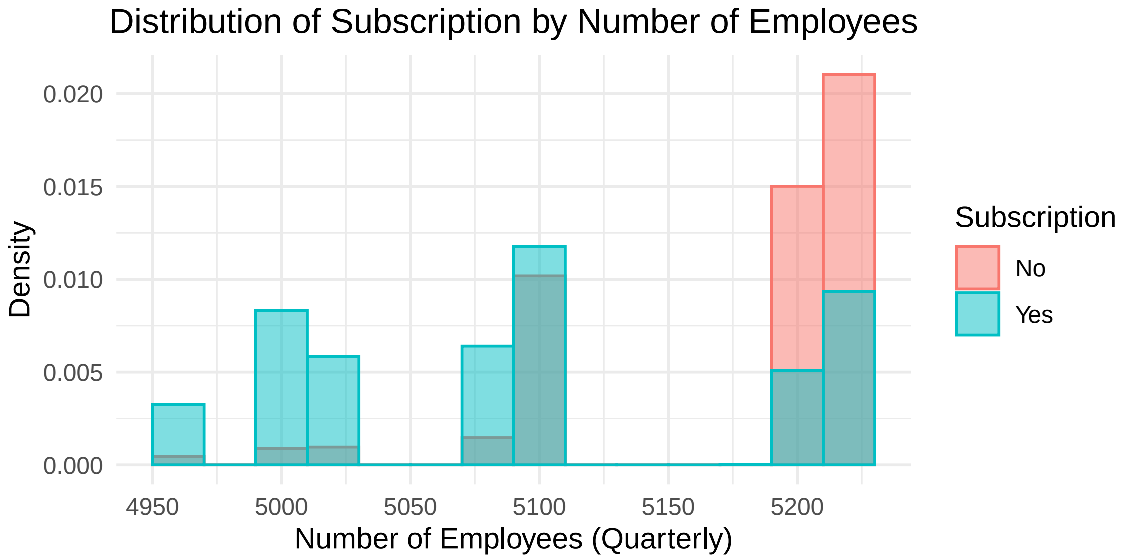


Figure 10: Term Deposit Subscription by Previous Campaign

[6] The dataset unfortunately does not specify unit of number of employees (e.g., in thousands)

# 4 Machine Learning Model

As the bank marketing campaign outcome for a client is to either subscribe the term deposit or not, it is a binary response. Therefore, binary classification algorithms should be used to predict client subscription decision.

Five binary classification algorithms are used to build the client subscription prediction models: Logistic Regression, Naive Bayes, Decision Tree, Random Forest and Gradient Boosted Tree.

Package `naivebayes` (Majka 2019), `rpart` (Therneau and Atkinson 2019), `randomForest` (Liaw and Wiener 2002), `gbm` (Greenwell et al. 2020) are used to build these models. Models are evaluated by `cvms` (Olsen and Zachariae 2021) and `MLmetrics` (Yan 2016) to return classification accuracy, F1 score[7] and confusion matrix.

## 4.1 Resample Imbalance Data

In this dataset, client subscription decisions vary that around 88.7% clients do not subscribe to term deposit while only around 11.3 clients subscribed (Table 2). The imbalance nature of this dataset could cause significant drawback on classifier performance (Sun, Wong, and Kamel 2009).

To solve or at least mitigate the imbalance data issue, I use package `ROSE` (Lunardon, Menardi, and Torelli 2014) to resample the training data. `ROSE` (Random Over-Sampling Examples) is a bootstrap-based technique to balance the rare class. After randomly splitting the data into 80%/20% training/testing set (32950 vs 8238 observations), I randomly over-sample the minority class `yes` and under-sample the majority class `no`. The training set is then resampled to 50000 observations that the ratio of class `yes` and class `no` is around 1:1 (25000 vs 25000 observations). And then I run binary classification algorithms on resampled training set and examine model performances through original testing set of the random split.

Table 2: Term Deposit Campaign Outcome

| Subscription Decision | Count | Proportion |
|---|---|---|
| no | 36548 | 88.73% |
| yes | 4640 | 11.27% |

## 4.2 Logistic Regression

Logistic Regression uses a logistic function to model a binary class and it is widely used in binary classification algorithm.

Following the resample procedure as suggested in Section 4.1, logistic regression yields the classification accuracy of 82.4% and the F1 score of 0.895. By using 5-fold Cross Validation, the robust estimates on testing set are 82.9% for accuracy and 0.899 for F1 score. The close measures between the model and its robust estimates suggest a valid model.

The confusion matrix of logistic regression (Figure 11) shows that the model yields 348 False Negatives, suggesting that 348 clients who actually subscribed to the term deposit campaign are incorrectly predicted not to subscribe to the term deposit. And the false negative rate[8] of the logistic regression model is 36.5%.

[7] Harmonic mean of precision and recall, which precision is (True Positive)/(True Positive + False Positive) and recall is (True Positives)/(True Positive + False Negative)

[8] Also refers to miss rate, which is (False Negative)/(False Negative + True Positive)

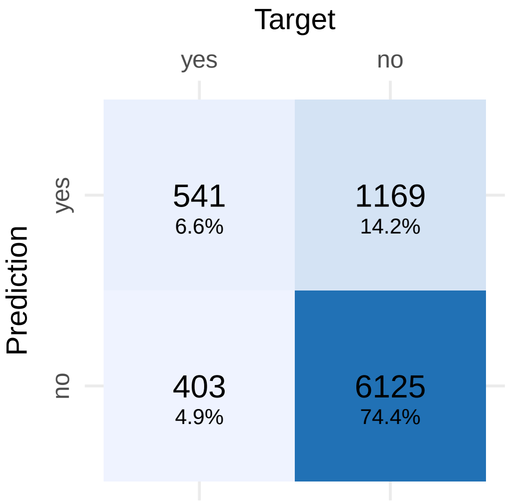


Figure 11: Confusion Matrix of Logistic Regression

## 4.3 Naive Bayes

Naive Bayes is a probabilistic classifier based on Bayes theorem and kernel density estimation could be applied on Bayesian models to achieve better performance (Pérez, Larrañaga, and Inza 2009).

Following the resample procedure as suggested in Section 4.1, Naive Bayes classifier using kernel density yields the classification accuracy of 76.3% and the F1 score of 0.852. By using 5-fold Cross Validation, the robust estimates on testing set are 77.9% for accuracy and 0.864 for F1 score. The close measures between the model and its robust estimates suggest a valid model.

The confusion matrix of Naive Bayes (Figure 12) shows that the model yields 283 False Negatives, suggesting that 283 clients who actually subscribed to the term deposit campaign are incorrectly predicted not to subscribe to the term deposit. And the false negative rate of the Naive Bayes model is 29.7%.

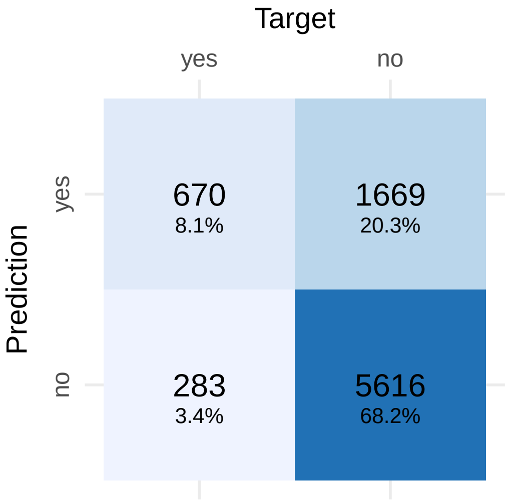


Figure 12: Confusion Matrix of Naive Bayes

### 4.4 Decision Tree

Decision Tree is a non-parametric classifier that uses greedy recursive binary splitting to predict the class label. Due to the greedy modeling approach, the decision tree could grow large and redundant, while a large decision tree could be pruned through cost complexity pruning that optimizes the tree complexity and classification accuracy.

Following the resample procedure as suggested in Section 4.1, the pruned decision tree yields the classification accuracy of 83.5% and the F1 score of 0.903. By using 5-fold Cross Validation, the robust estimates on testing set are 83.54% for accuracy and 0.903 for F1 score. The close measures between the model and its robust estimates suggest a valid model.

The nature of recursive binary splitting provides interpretability of decision tree model. The following pruned decision tree plot (Figure 13) suggests that if the quarterly indicator of number of employees is greater or equal to 5088, and the last campaign contact is in May, June, July, August, November or December, the client will be predicted not to subscribe to the term deposit. If the quarterly indicator of number of employees is less than 5088, the client will be predicted to subscribe to the term deposit.

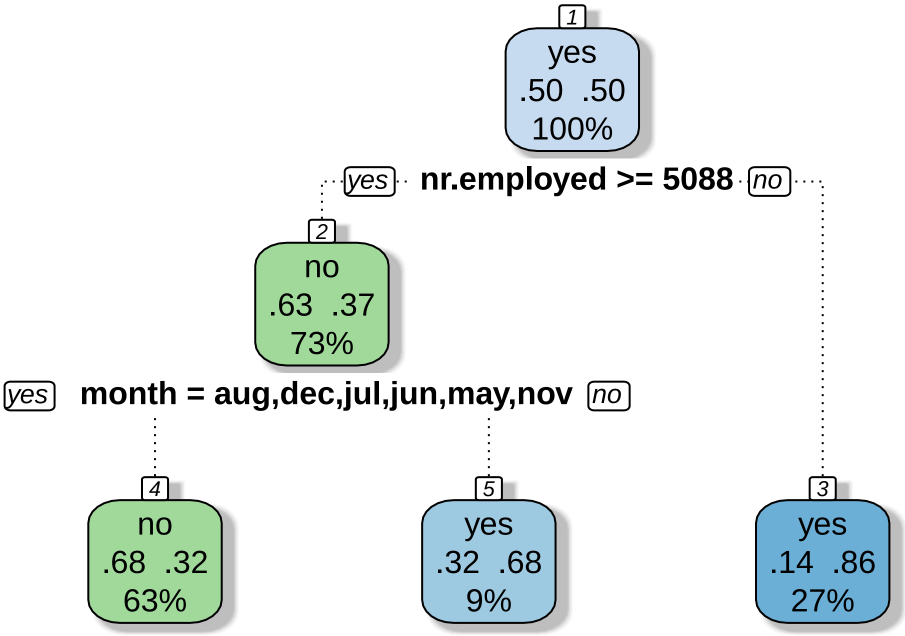


Figure 13: Tree Plot of Decision Tree

The confusion matrix of Decision Tree (Figure 14) shows that the model yields 369 False Negatives, suggesting that 369 clients who actually subscribed to the term deposit campaign are incorrectly predicted not to subscribe to the term deposit. And the false negative rate of the Decision Tree model is 38.7%.

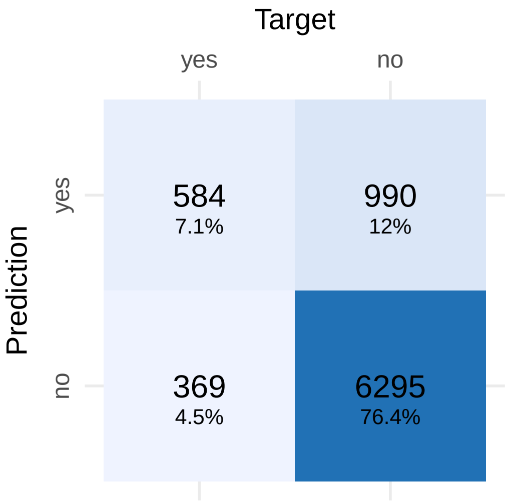


Figure 14: Confusion Matrix of Decision Tree

### 4.4.1 Random Forest

Random Forest is a non-parametric classifier that construct a multitude of decision trees on bootstrapped training samples. When build such tree, a random sample of m features is chosen as split candidates from the full set of p features for each split in a tree that $m{\approx}\sqrt{p}$. After preprocessing, there are 18 features left in the marketing campaign dataset, so a random sample of $m = 4$ feature will be chosen for each split.

Following the resample procedure as suggested in Section 4.1, the random forest yields the classification accuracy of 86.7% and the F1 score of 0.924. By using 5-fold Cross Validation, the robust estimates on testing set are 86.5% for accuracy and 0.923 for F1 score. The close measures between the model and its robust estimates suggest a valid model.

The confusion matrix of Random Forest (Figure 15) shows that the model yields 437 False Negatives, suggesting that 437 clients who actually subscribed to the term deposit campaign are incorrectly predicted not to subscribe to the term deposit. And the false negative rate of the Random Forest model is 45.9%.

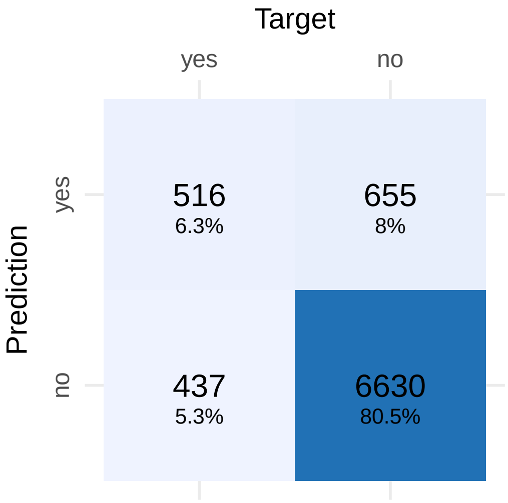


Figure 15: Confusion Matrix of Random Forest

The random forest model allows to identify how important a feature is in classifying the data. Mean Decrease Accuracy and Mean Decrease Gini Coefficient are commonly measures to examine feature importance (Han, Guo, and Yu 2016). Mean Decrease Accuracy expresses how much accuracy the model losses by excluding such feature. Higher classification accuracy loss suggests higher feature importance. Mean Decrease Gini Coefficient expresses how each feature contributes to the homogeneity of nodes and leaves in resulting random forest. Higher Mean Decrease Gini Coefficient suggests higher feature importance.

According to the Mean Decrease Accuracy plot (Figure 16), Top 5 features that contribute most to this random forest model is *job*, *education*, *campaign*, *day_of_week* and *age*, which mainly belongs to client personal data. For Mean Decrease Gini plot (Figure 17), Top 5 features that contribute most to this random forest model is *euribor3m*, *age*, *job*, *nr.employed*, *education*, which mainly belongs to client personal data and social & economic indicators.

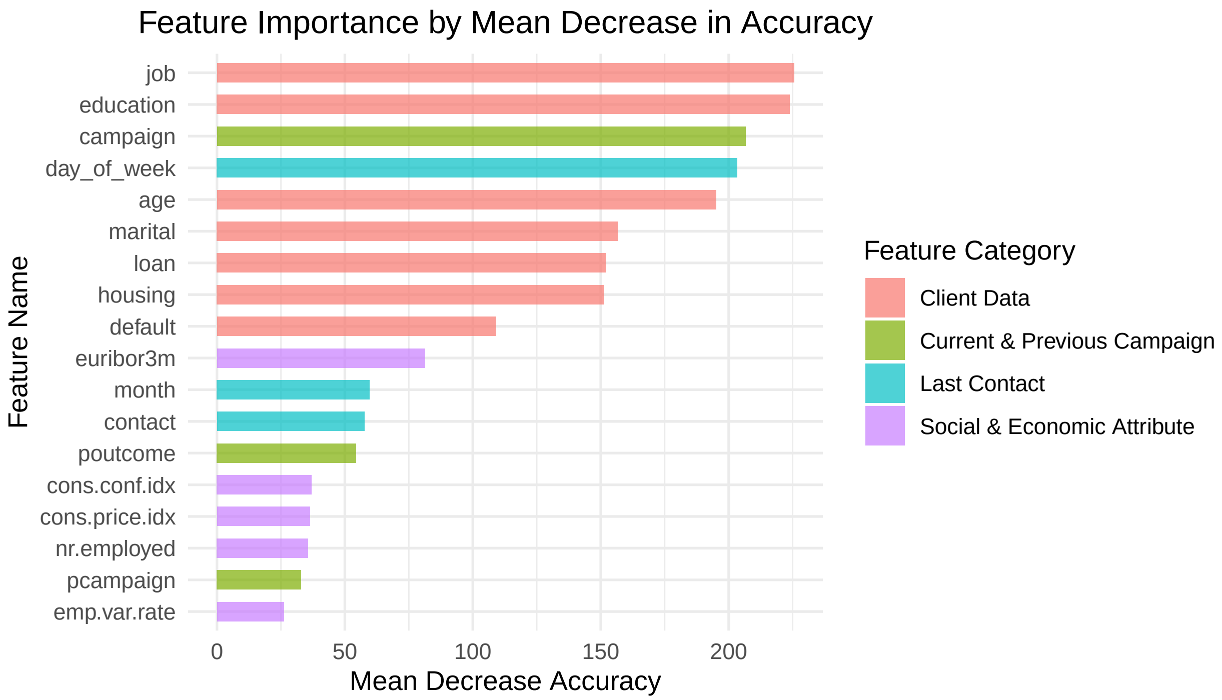


Figure 16: Feature Importance by Mean Decrease in Accuracy

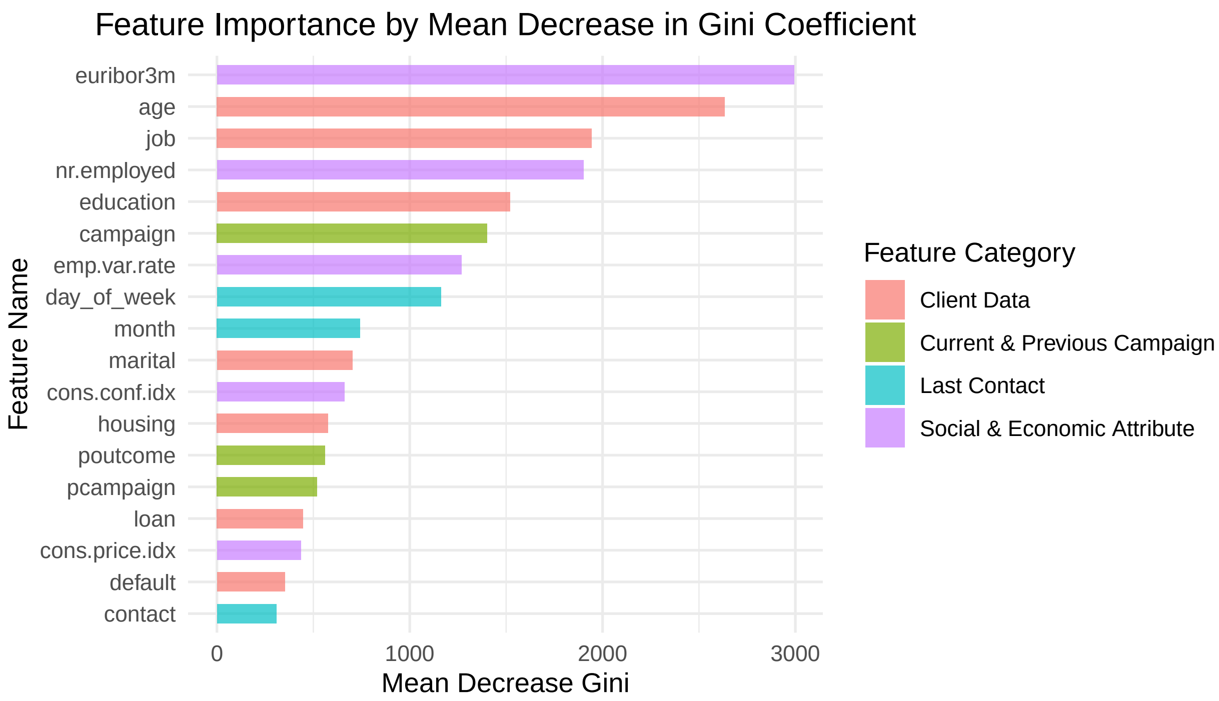


Figure 17: Feature Importance by Mean Decrease in Gini Coefficient

### 4.5 Gradient Boosted Tree

Gradient Boosted Tree use is another non-parametric algorithm based on decision tree. Each tree fits on the original training data and the tree is grown sequentially using information of previously built trees. The model is then boosted by combining a large number of decision trees to improve its performance.

The model performance of gradient boosted tree could be affected by following hyperparameters: (1) total number of $n$ trees to fit, (2) interaction depth $d$ that control level of interaction of features within the model and (3) shrinkage or learning rate $\lambda$ that control the tree expansion. Hyperparameter tuning using 5-fold Cross Validation is performed to find a robust hyperparameter combination[9] that returns the best model performance on the dataset based on F1 score. The tuning result suggests that when $n = 250$, $d = 1$ and $\lambda = 0.001$, the model reaches its best performance on F1 score.

Following the resample procedure as suggested in Section 4.1, the random forest yields the classification accuracy of 87.4% and the F1 score of 0.929. By using 5-fold Cross Validation, the robust estimates on testing set are 87.3% for accuracy and 0.928 for F1 score. The close measures between the model and its robust estimates suggest a valid model.

The confusion matrix of Gradient Boosted Tree (Figure 18) shows that the model yields 506 False Negatives, suggesting that 506 clients who actually subscribed to the term deposit campaign are incorrectly predicted not to subscribe to the term deposit. And the false negative rate of the Gradient Boosted Tree model is 53.1%.

[9] n = 100,250,500 ; d = 1,5,10,15 ; $\lambda$ = 0.001, 0.002, 0.01, 0.02

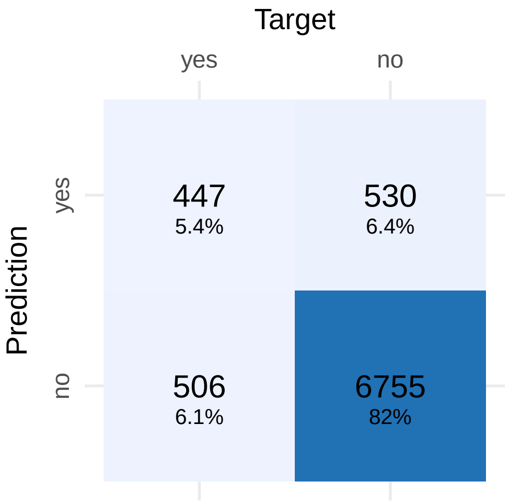


Figure 18: Confusion Matrix of Gradient Boosted Tree

# 5 Result and Discussion

## 5.1 Model Evaluation

The following table (Table 3) summarizes evaluation metrics of all previous models. Among these five algorithms, Random Forest has the highest classification accuracy of 86.7% and F1 score of 0.924 so that it could be considered as a powerful model to predict client term deposit subscription decision.

Table 3: Summary of ML Model Evaluation Metrics

| **Algorithm** | **Accuracy** | **F1 Score** | **False Negatives** | **FN Rate** |
|---|---|---|---|---|
| Logistic Regression | 82.4% | 0.895 | 348 | 36.5% |
| Naive Bayes | 76.3% | 0.852 | 283 | 29.7% |
| Decision Tree | 83.5% | 0.903 | 369 | 38.7% |
| Random Forest | 86.7% | 0.924 | 437 | 45.9% |
| Gradient Boosted Tree | 86.5% | 0.923 | 506 | 53.1% |

However, if a bank wants to implement a realistic predictive model on subscription decision, a Random Forest model may not be a optimal choice. Notice that 437 false negatives of random forest indicate that 437 clients who actually subscribed to the term deposit campaign are incorrectly predicted not to subscribe to the term deposit. The bank might not run the marketing campaign on clients predicted not to be the primary marketing targets, although they do subscribe. Such missed targeting on potential clients could affect the profitability of the bank. Therefore, model with low false negatives and false negative rates shall be considered to capture more potentially valuable clients. In this situation, Naive Bayes outperforms other algorithms with its lowest false negatives of 283 and false negative rate of 29.7%.

## 5.2 Limitation

During the analysis, I identified several limitations that may affect the evaluation result:

1. Imbalance data: As mentioned in Section 4.1, the imbalance nature of the data could affect the classification result. Resampling the training data may not be a perfect solution to this issue.
2. Missing time series feature: Since the data is collected from May 2008 to November 2010, some deposit campaigns were run under 2008 financial crisis. Previous studies have suggested that financial crisis could affect the deposit decision (Han and Melecky 2013), while a time feature that records the actual time of contacting clients is missing from the dataset. Even though some social and economic indicator features in this dataset may reflect certain economic conditions, relationships between these indicators and economic conditions still requires careful examinations.

## 5.3 Discussion and Future Steps

In conclusion, even though a relatively poor performance on classification accuracy, I would suggest using Naive Bayes as a predictive model to determine client subscriptions due to its ability to capture potential clients. And based on previous analysis result, the bank shall take several actions to improve their marketing campaign success rate:

1. Adjust target demographics: According to exploratory data analysis (Figure 1 and Figure 2) and feature importance plot of random forest (Figure 16), there are several client personal features such as age and type of job affecting the subscription rate of the term deposit. The bank may wisely target clients based on demographic information to increase the subscription rate and reduce the labor cost.
2. Choose proper time to run campaigns: According to exploratory data analysis (Figure 5 and tree plot of decision tree (Figure 13, running the marketing campaign at a proper time period could lead to an increase of subscription rate.

There are potentials for improving the analysis result in future studies, including but not limited to: implementing other classification algorithms such as Support Vector Machine, exploring additional features available to improve model performance, etc.

# A Remaining Exploratory Visualizations

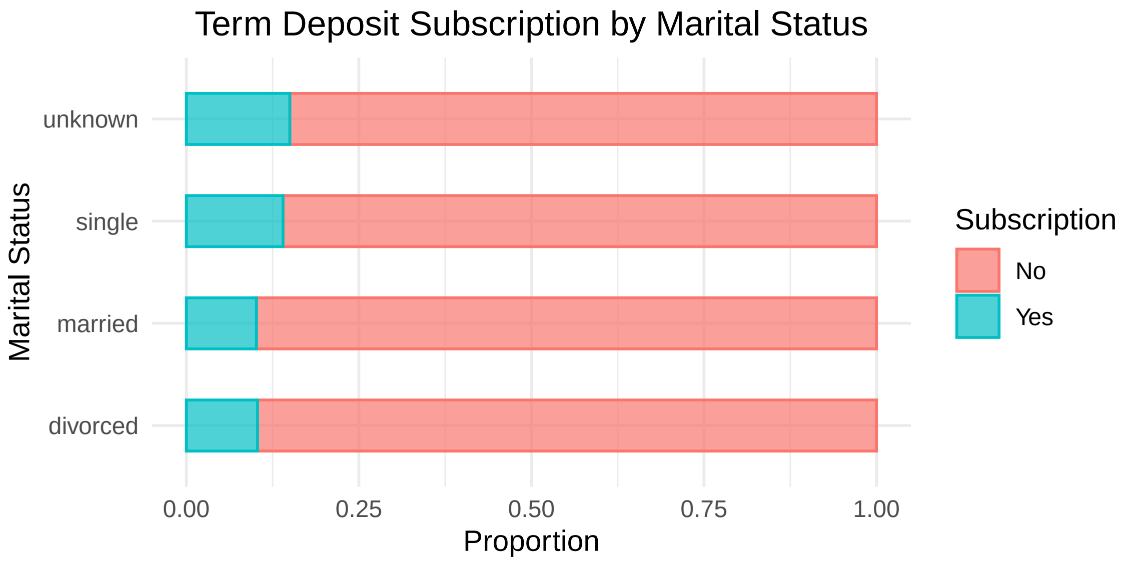
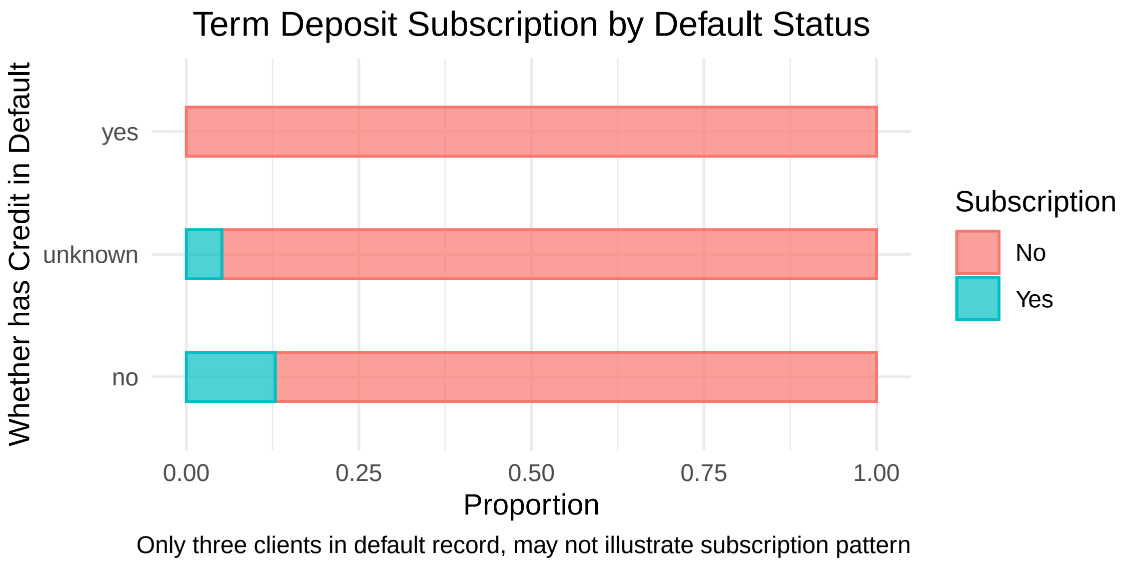


Figure 19: Term Deposit Subscription by Marital Status



Only three clients in default record, may not illustrate subscription pattern

Figure 20: Term Deposit Subscription by Default Status

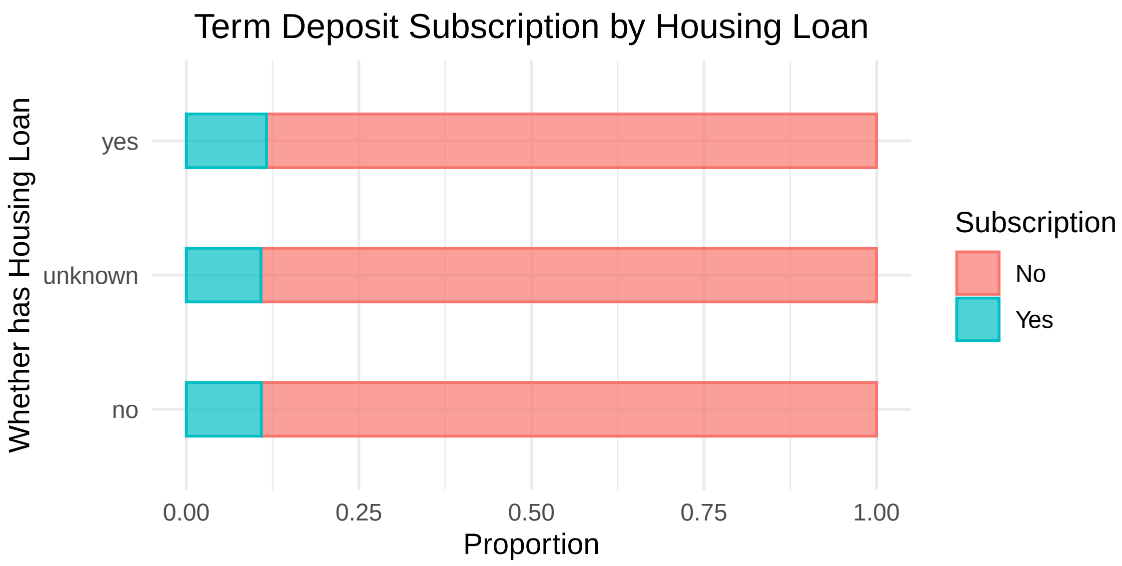


Figure 21: Term Deposit Subscription by Housing Loan

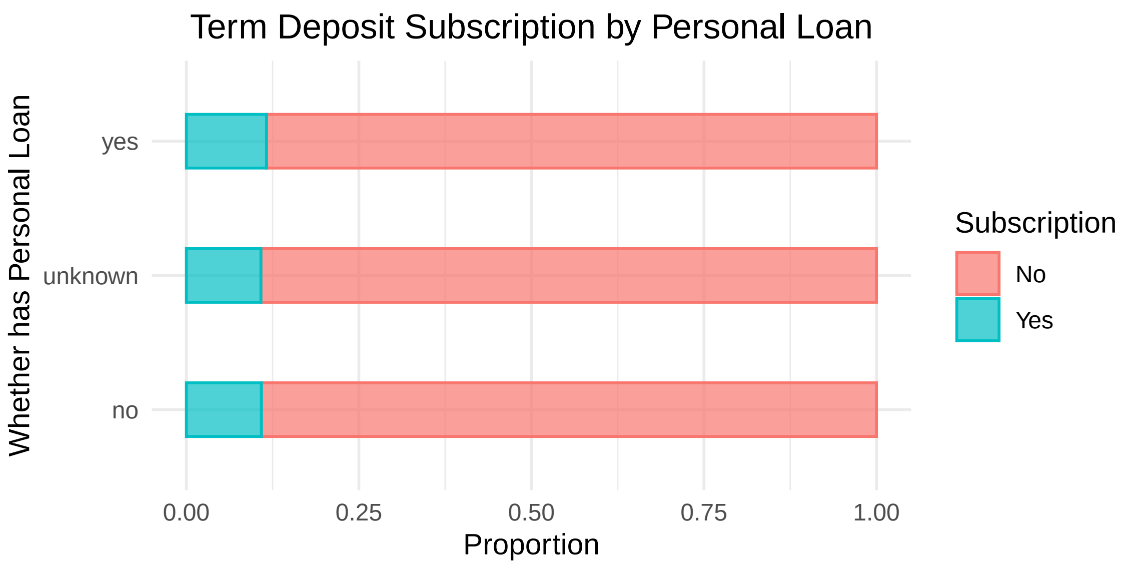


Figure 22: Term Deposit Subscription by Personal Loan

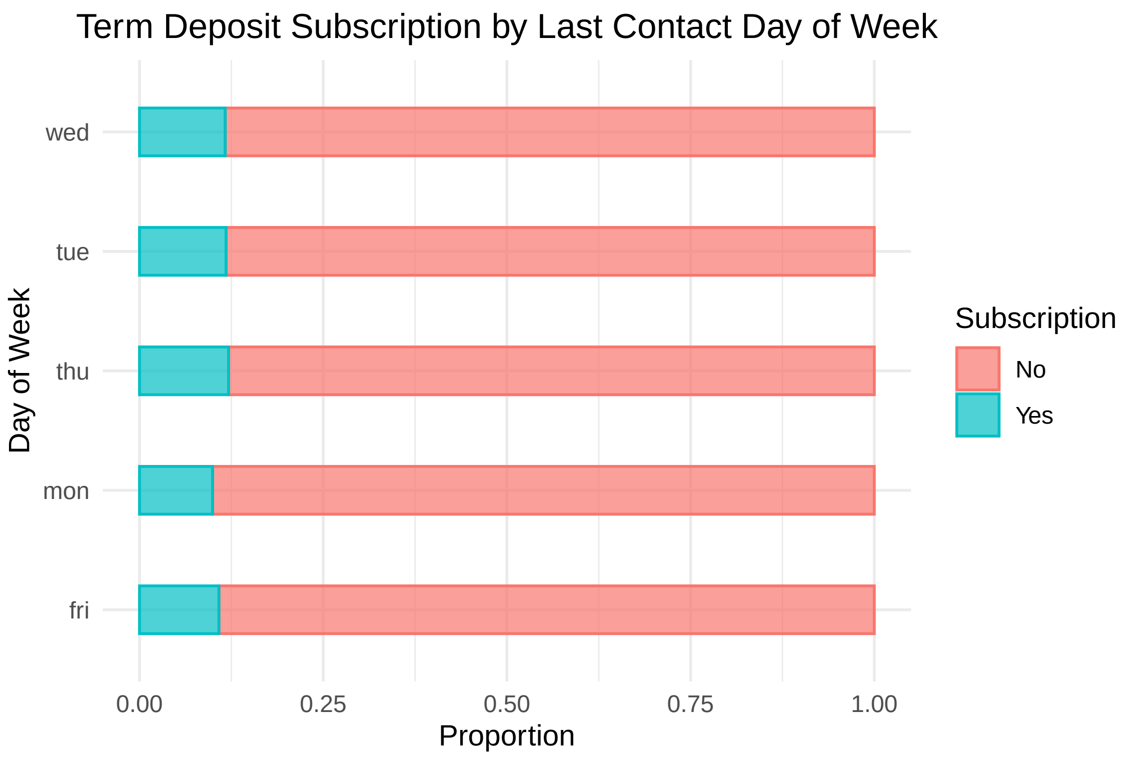


Figure 23: Term Deposit Subscription by Last Contact Day of Week

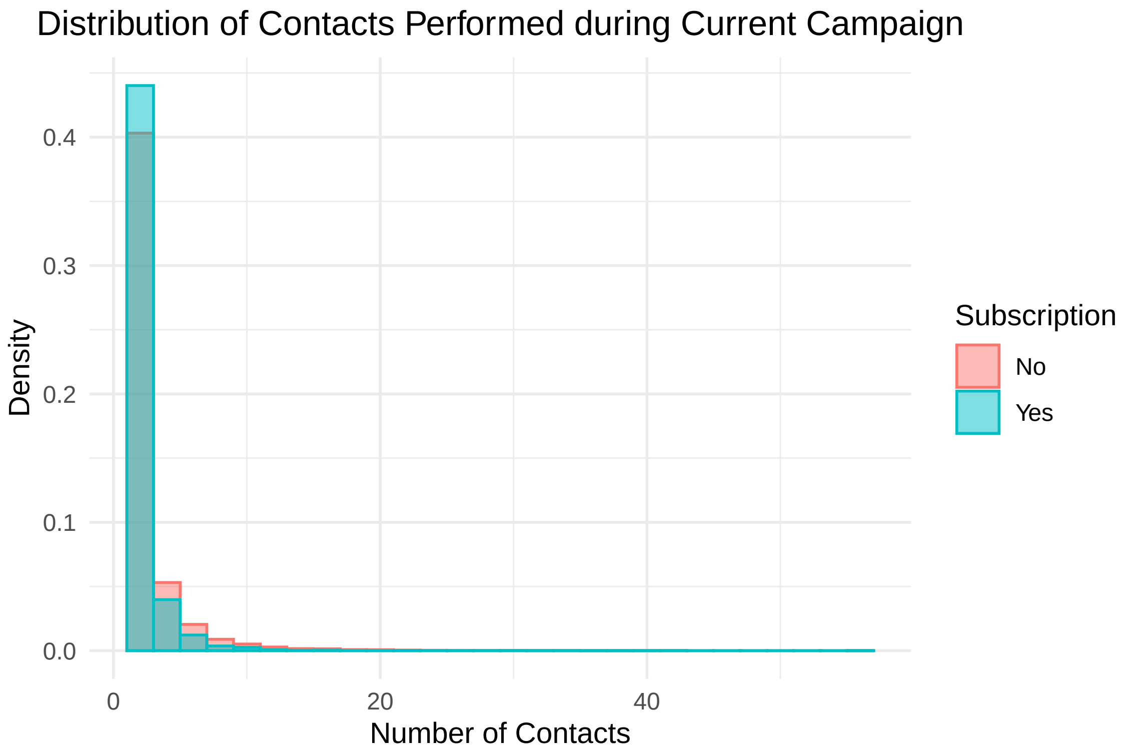


Figure 24: Density Distribution of Contacts Performed during Current Campaign

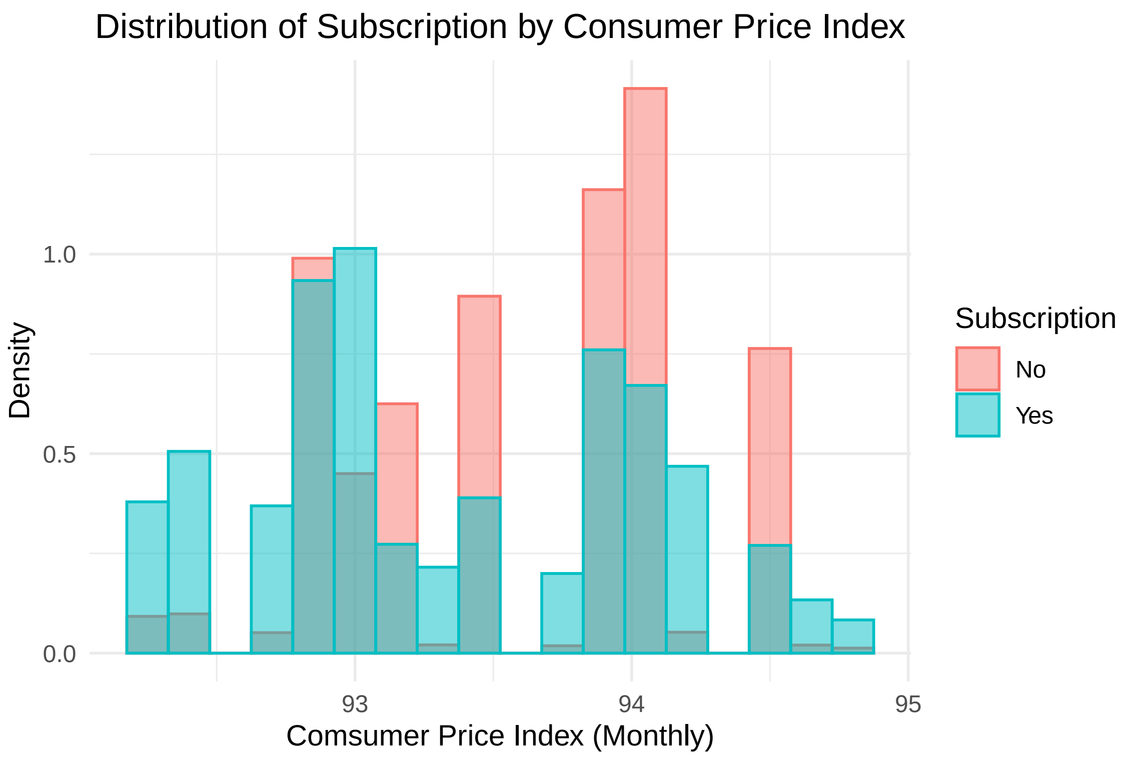


Figure 25: Distribution of Subscription by Consumer Price Index

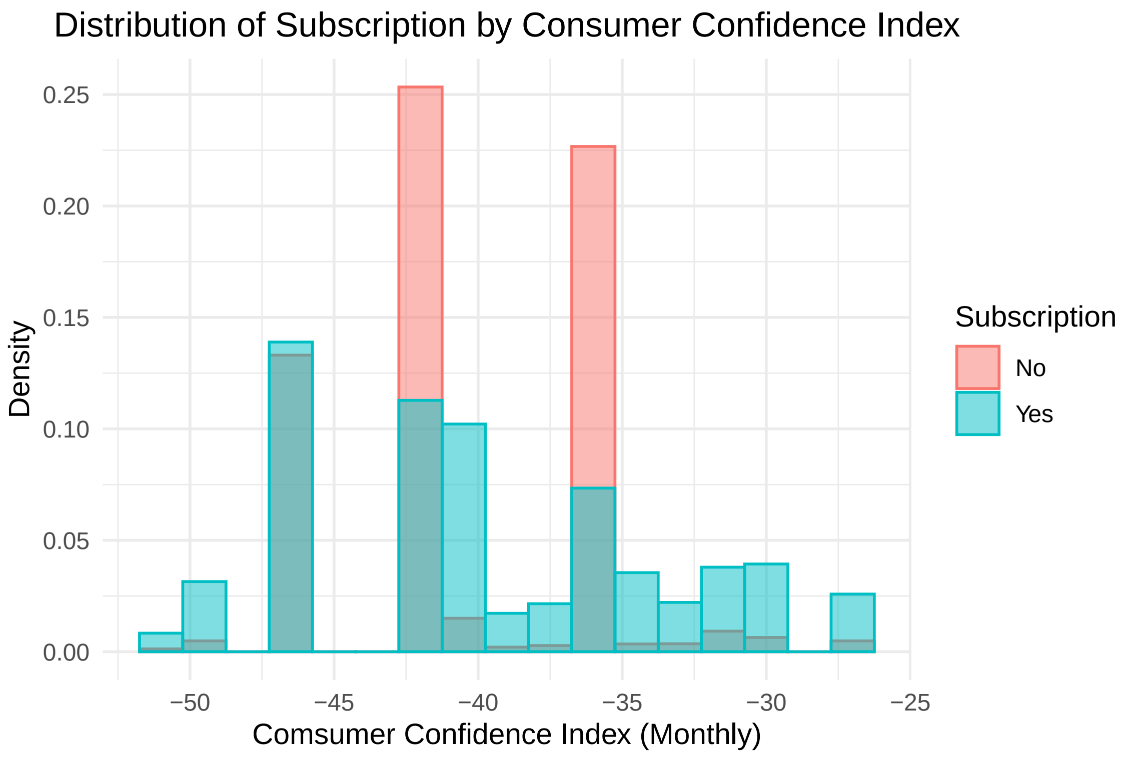


Figure 26: Distribution of Subscription by Consumer Confidence Index

# References

Greenwell, Brandon, Bradley Boehmke, Jay Cunningham, and GBM Developers. 2020. *Gbm: Generalized Boosted Regression Models*. https://CRAN.R-project.org/package=gbm.

Han, Hong, Xiaoling Guo, and Hua Yu. 2016. "Variable Selection Using Mean Decrease Accuracy and Mean Decrease Gini Based on Random Forest." In *2016 7th Ieee International Conference on Software Engineering and Service Science (Icsess)*, 219–24. IEEE.

Han, Rui, and Martin Melecky. 2013. "Financial Inclusion for Stability: Access to Bank Deposits and the Deposit Growth During the Global Financial Crisis."

Liaw, Andy, and Matthew Wiener. 2002. "Classification and Regression by randomForest." *R News* 2 (3): 18–22. https://CRAN.R-project.org/doc/Rnews/.

Lunardon, Nicola, Giovanna Menardi, and Nicola Torelli. 2014. "ROSE: A Package for Binary Imbalanced Learning." *R Journal* 6 (1): 82–92.

Majka, Michal. 2019. *Naivebayes: High Performance Implementation of the Naive Bayes Algorithm in R*. https://CRAN.R-project.org/package=naivebayes.

Moro, Sérgio, Paulo Cortez, and Paulo Rita. 2014. "A Data-Driven Approach to Predict the Success of Bank Telemarketing." *Decision Support Systems* 62: 22–31.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Olsen, Ludvig Renbo, and Hugh Benjamin Zachariae. 2021. *Cvms: Cross-Validation for Model Selection*. https://CRAN.R-project.org/package=cvms.

Pérez, Aritz, Pedro Larrañaga, and Iñaki Inza. 2009. "Bayesian Classifiers Based on Kernel Density Estimation: Flexible Classifiers." *International Journal of Approximate Reasoning* 50 (2): 341–62.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sun, Yanmin, Andrew KC Wong, and Mohamed S Kamel. 2009. "Classification of Imbalanced Data: A Review." *International Journal of Pattern Recognition and Artificial Intelligence* 23 (04): 687–719.

Therneau, Terry, and Beth Atkinson. 2019. *Rpart: Recursive Partitioning and Regression Trees*. https://CRAN.R-project.org/package=rpart.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Williams, Graham J. 2011. *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*. Use R! Springer. http://www.amazon.com/gp/product/1441998896/ref=as_li_qf_sp_asin_tl?ie=UTF8&tag=togaware-20&linkCode=as2&camp=217145&creative=399373&creativeASIN=1441998896.

Xie, Yihui. 2016. *Bookdown: Authoring Books and Technical Documents with R Markdown*. https://github.com/rstudio/bookdown.

Yan, Yachen. 2016. *MLmetrics: Machine Learning Evaluation Metrics*. https://CRAN.R-project.org/package=MLmetrics.

Zhu, Hao. 2020. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.