Norah Sulaiman Saleh Alogla

Bank Marketing

Predict Term Deposit Subscriptions in Bank

In this project I will try to perform exploratory data analysis and build a machine learning model to Predict if a client will subscribe (yes/no) to a term deposit in the bank. And this is defined as a classification problem.

The dataset we found it on kaggle.com and it gives us information about a marketing campaign of a financial institution in which you will have to analyze in order to find ways to look for future strategies in order to improve future marketing campaigns for the bank. Dataset contained 17 different features and 11162 clients. Features were both categorical and numerical. Target variable was binary ("yes" or "no").

The most important column in the dataset is y, which is the output variable (desired target): this will tell us if the client subscribed to a term deposit (binary: 'yes', 'no').

After that I clean the data to make sure everything is ready on my data analysis (Feature Engineering, drop unwanted Features, Handle Missing Values, Handle Categorical Features, Handle Feature Scaling, Remove Outliers). Then we split dataset into training set and testing set.

For this classification project, we used 3 models which are KNN, Random Forest classifier, and XGB Classifier. From the results I found the XGB Classifier has the best performance on the data and achieve the best accuracy. And I used many different methods

to be sure that our model performs well on the data (Accuracy matrix, Confusion Matrix, and the ROC Curve).

Finally, we found the number of the clients who not deposit is more than the clients who made deposits. I suggest making more campaigns to encourage clients to make more deposits in our bank to increase the income.

Dataset link:
https://www.kaggle.com/janiobachmann/bank-marketing-dataset