



Analyzing Bank Marketing Campaigns

Miguel Santana

Flatiron School

Data Science | FT Cohort

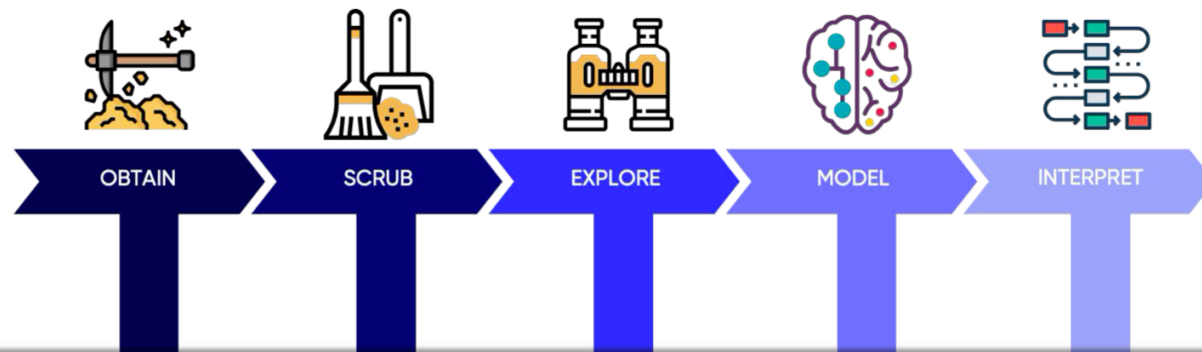
Introduction

- Methodology
- OSEMN Framework
- The Data | Cleaning
- Exploratory Data Analysis
- Modeling
- Interpret | Analyze
 - Features: Age & Number of Employees
 - Features: Campaign Contact & Days Since Last Contact
- Business Recommendations
- Conclusions & Limitations
- Thank you

Methodology

- A Portuguese financial institution provided telemarketing data concerning marketing campaigns with the goal of predicting subscriber term deposits.
- Dataset included client, campaign, social and economic data
- Leveraged OSEMN framework to analyze the dataset and make key business recommendations

Data Science Process



O
Gather data from
relevant sources

S
Clean data to formats
that machine
understands

E
Find significant patterns
and trends using
statistical methods

M
Construct models to
predict and forecast

N
Put the results into
good use

Originally by Hilary Mason and Chris Wiggins

OSEMN Framework

The Data | Cleaning

Available via UCI's Machine Learning Archive and Kaggle

- <http://archive.ics.uci.edu/ml/datasets/Bank+Marketing#>
- <https://www.kaggle.com/henriqueyamahata/bank-marketing>



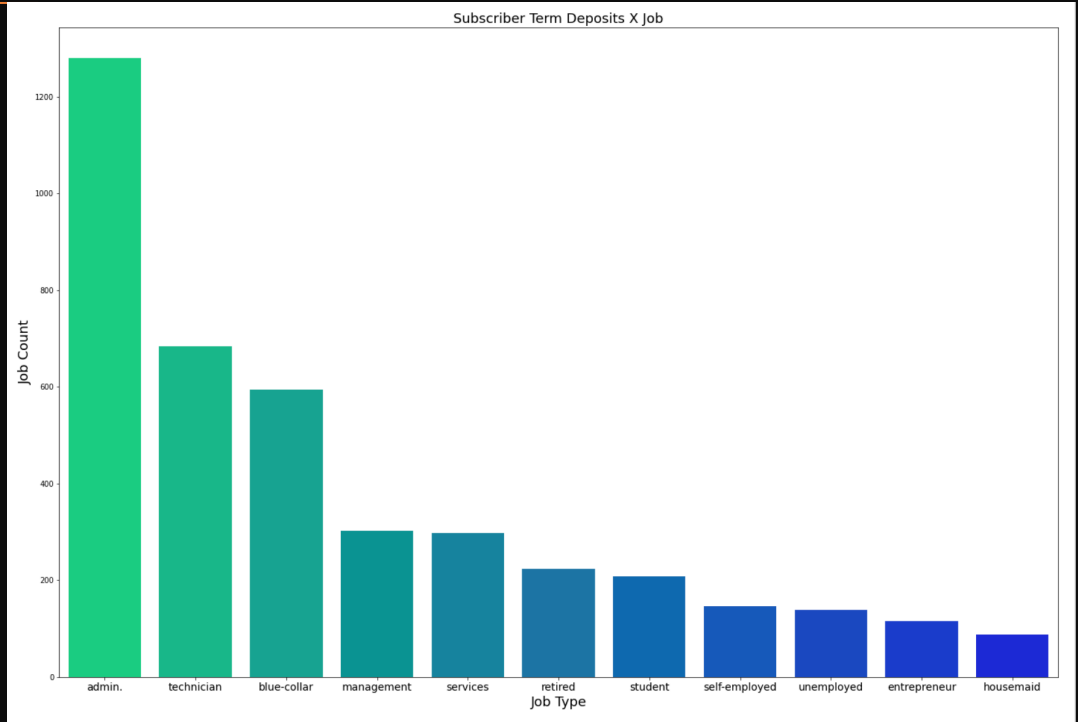
Cleaning the data

- Addressing "unknown" values
- Focused on consumers who provided customer profile answers
- Removed Outliers
 - Age
- Feature Selection

EXPLORATORY DATA ANALYSIS

Most Common Client Features

- Job Type: admin
- Age: early thirties
- Education Level: university degree
(see appendix for education)



Modeling

Machine Learning Algorithms

- Gradient Boosting Classifier
- AdaBoost Classifier
- Logistic Regression
- Support Vector Machine
- K-Nearest Neighbors
- Random Forest
- Gaussian Naïve Bayes
- Decision Tree

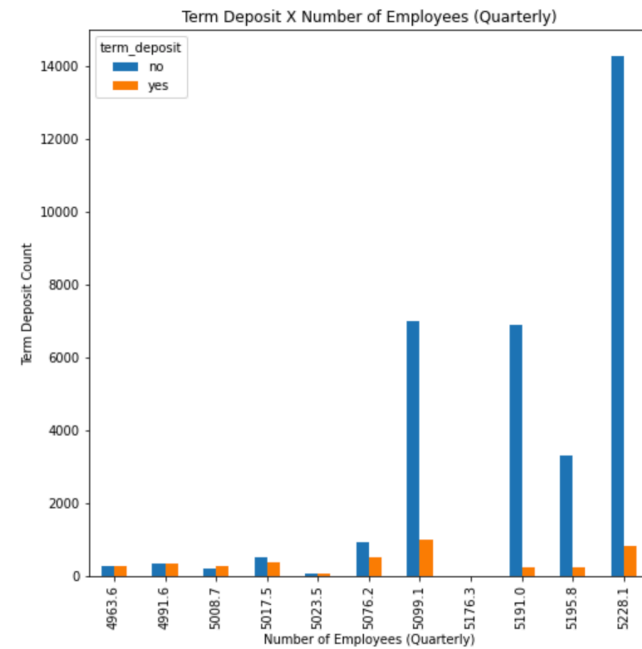
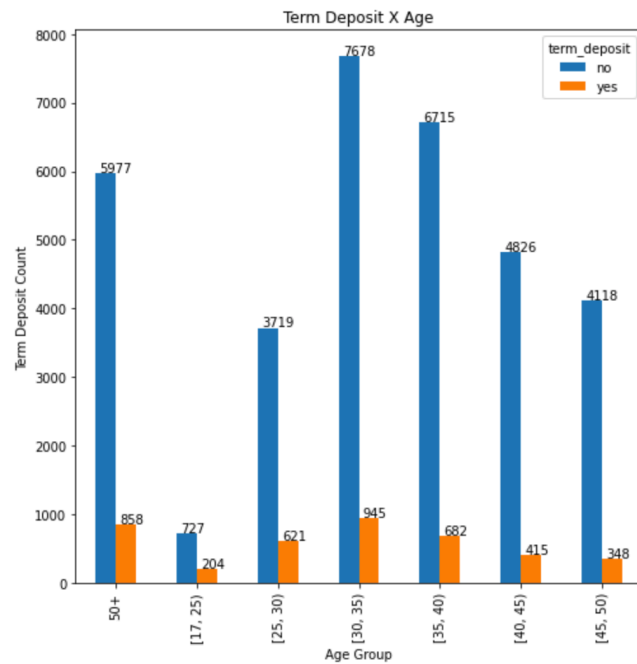
Accuracy Score

- 90.92%
- 90.70%
- 90.58%
- 90.30%
- 89.92%
- 89.73%
- 88.19%
- 84.47%

The top models selected for feature evaluation were Gradient Boosting Classifier and Adaboost Classifier

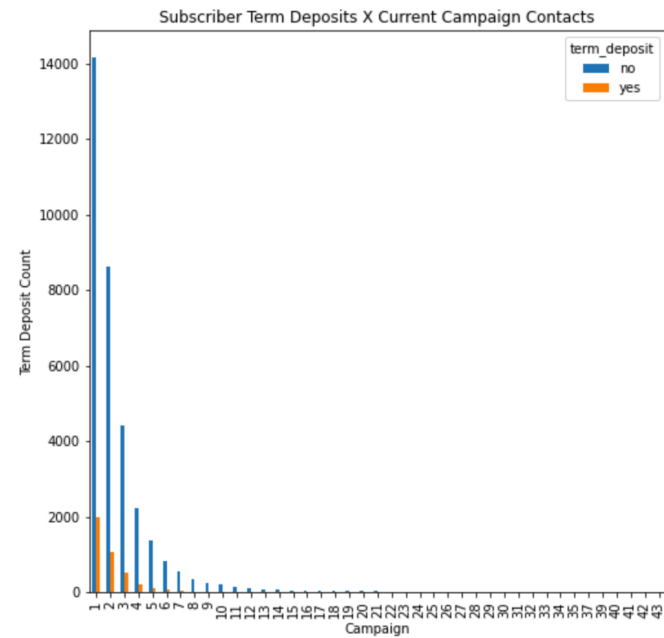
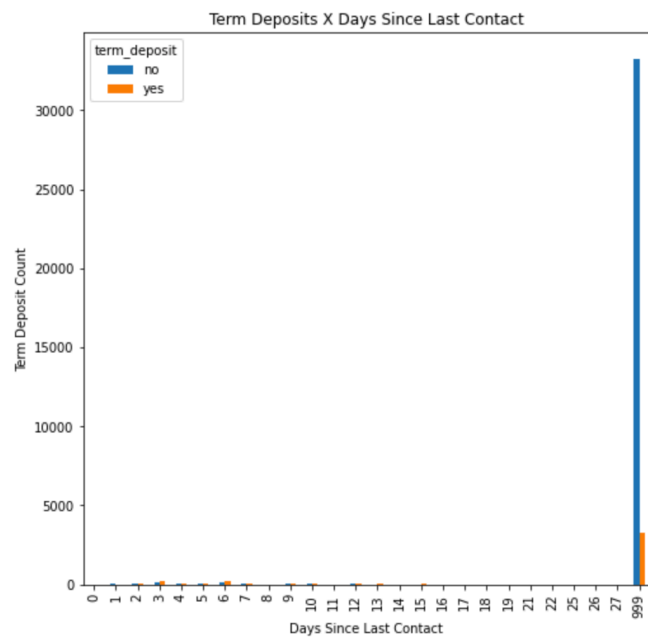
Interpret | Analyze

- Generated top 25 features (per model)
- Observed overlap | selected features
- Selected Top Features
 - Age
 - Number of Employees
 - Number of times contacted during the current campaign
 - Number of days since last contact (999 = no contact)



Features: Age & Number of Employees

Features: Campaign Contact & Days Since Last Contact



Business Recommendations

- 1) Observe current marketing on consumers age 17-25. Use similarly structured marketing on consumers age 30-40.
- 2) Maintain an average employee level of 5076 or lower. Focus on productivity when employee levels go over 5076.
- 3) Deploy an A team to handle first 3 telemarketing contacts per customer.
- 4) Deploy B team to handle new callers and first-time bank clients.

Conclusions & Limitations

Conclusions

- Feature trends illustrate high subscriber counts but higher non-subscriber counts as volume increases.
- Employee effectiveness decreases as additional employees are hired (over 5076 quarterly) and volume increases.

Limited to consumers that:

- Cite job type, marital status and education level
- Between the ages of 17 and 69

THANK YOU!

Questions? Miguel Santana | contact: msantana269@gmail.com

Github Repo: <https://github.com/msantana269/Module-3-Project>



Appendix

Appendix: Educational Trends

