

Module 3 Project: Machine Learning Techniques on Bank Marketing Data

Graham Wilson

4/21/20

- Introduction
- Methodology
- Models
- Data Visualization and Analysis
- Conclusion and Recommendations

Project –

- This project uses a Bank Marketing Dataset to find the best machine learning strategies to improve on the next marketing campaign.

Dataset –

- Bank Marketing Dataset acquired through UCI's data archives:
 - <https://archive.ics.uci.edu/ml/datasets/bank+marketing>

Goal –

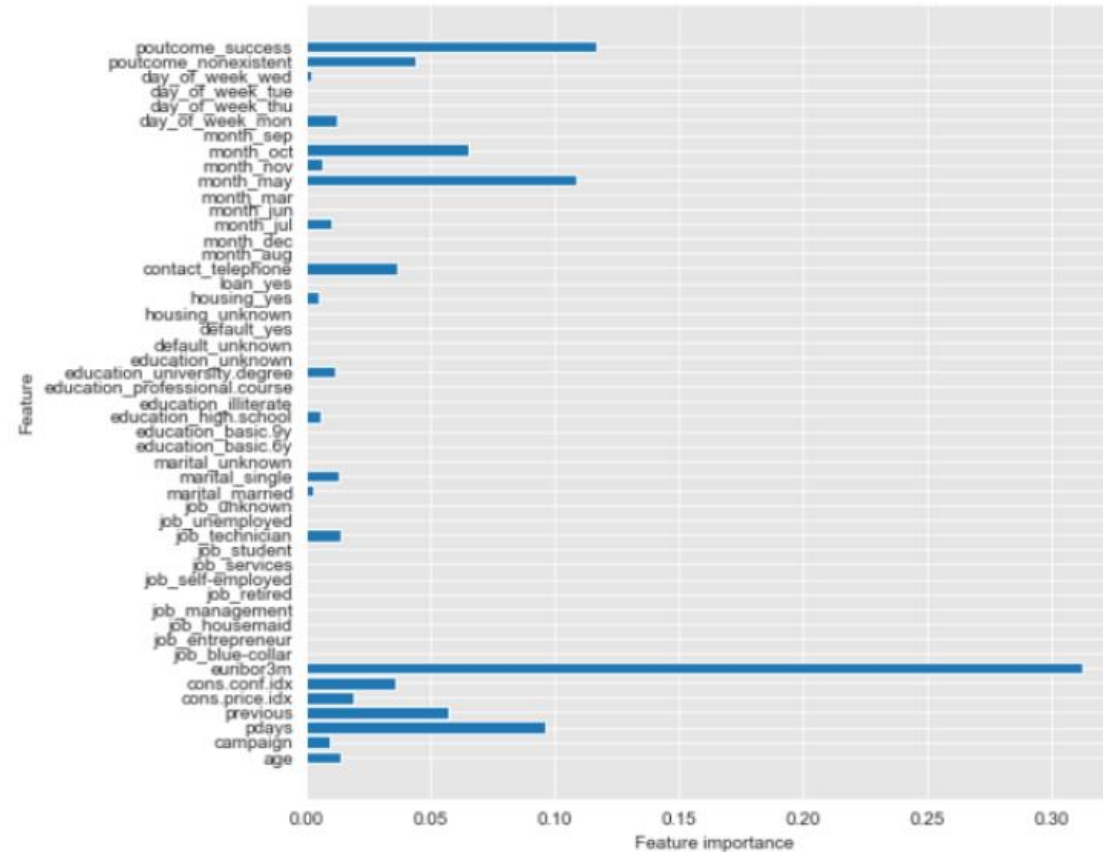
- Provide an accurate, predictive model that determines the outcome of a term deposit marketing campaign for future investors.
- Deliver strategies for increasing acceptance rate of term deposits through marketing optimization.

- Using Python:
 - Import data into Jupyter Notebook
 - Scrub
 - Convert data types
 - Manage extraneous values
 - Characterize and transform relevant variables
 - Explore and Analyze
 - Remove correlative variables
 - Model
 - Logistic regression
 - Decision trees
 - Random forest
 - XG boost
 - Interpret models meaningful
 - Generate visualizations

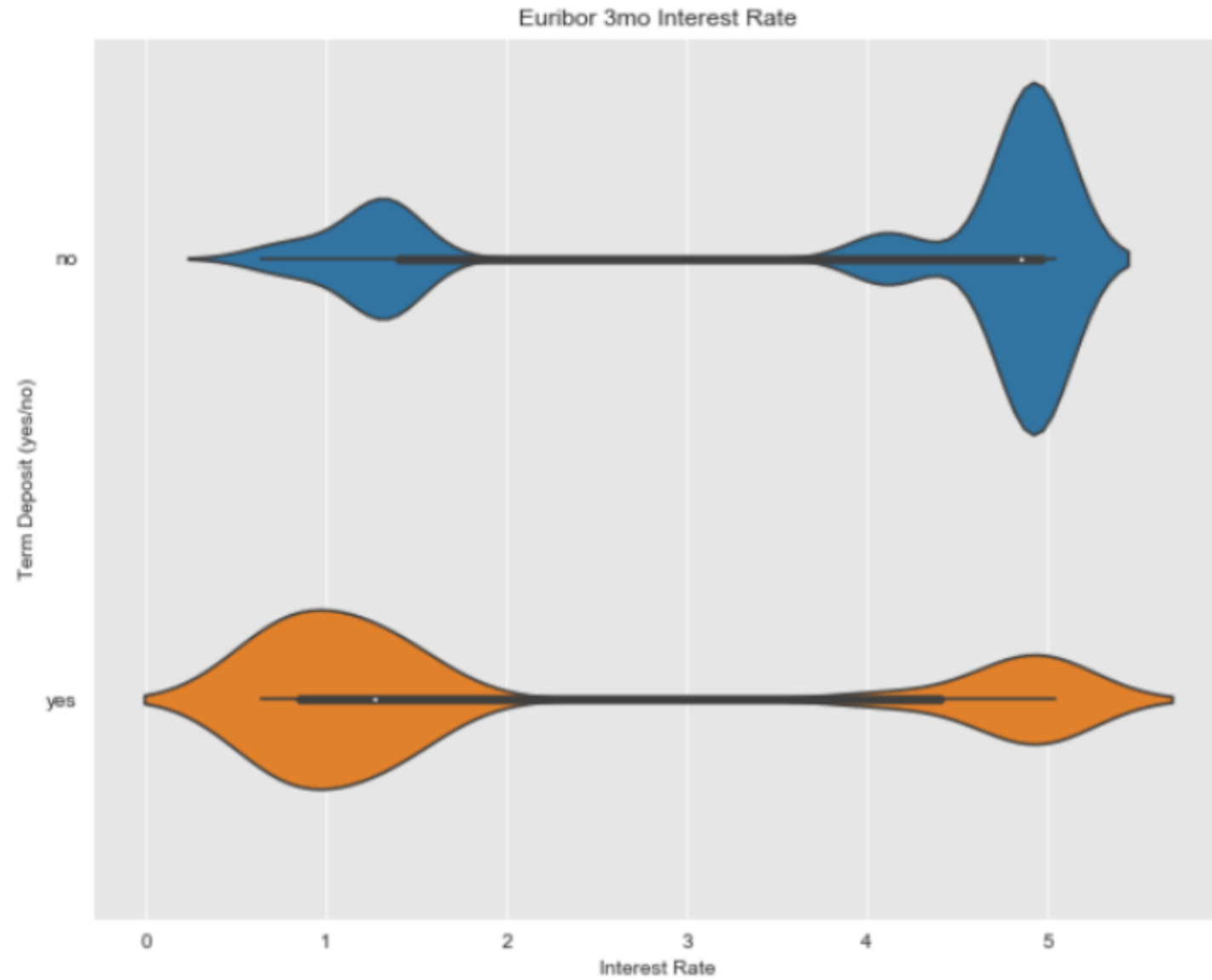
Model Results

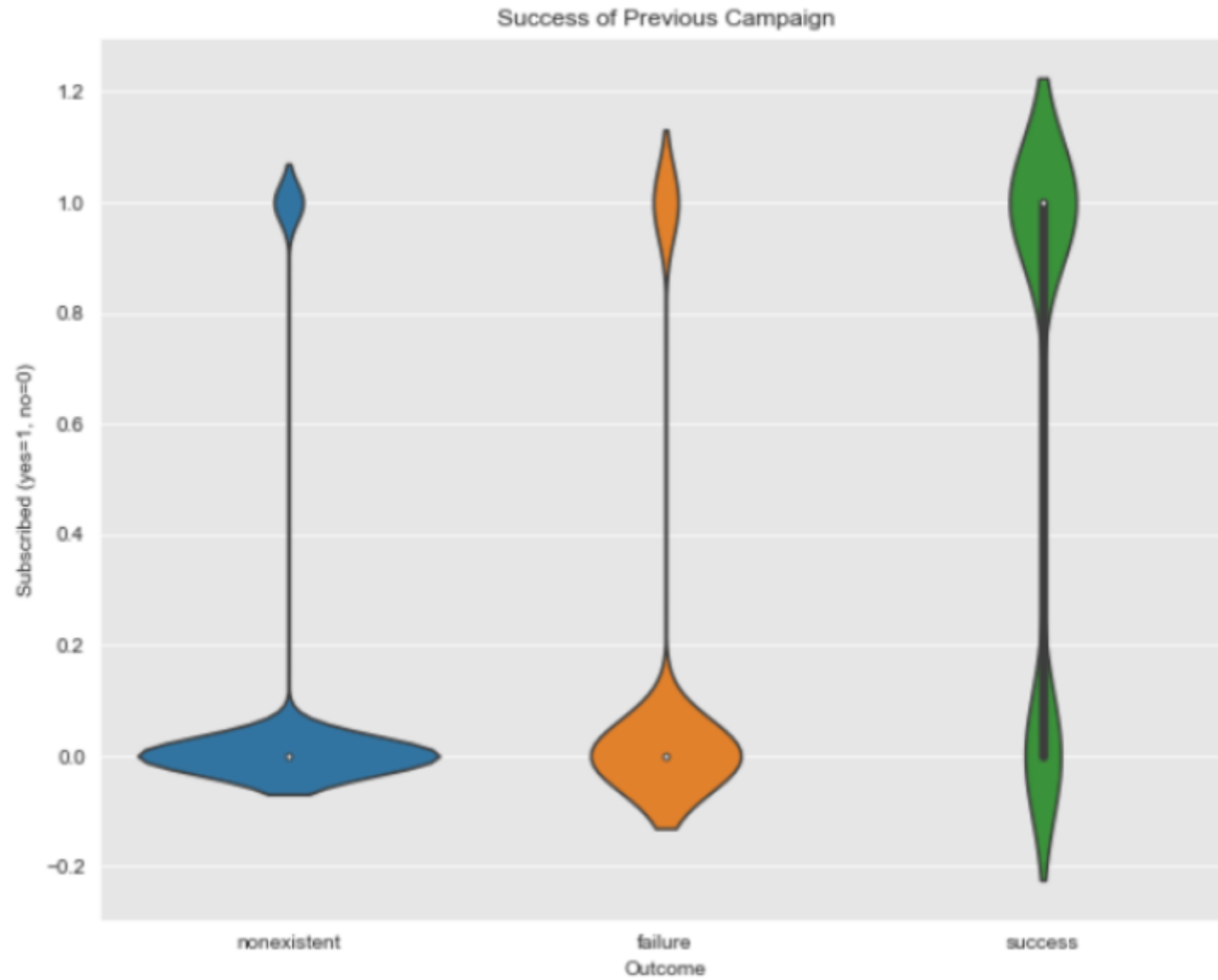
Model	Precision	Recall	F1-Score	Accuracy	Top 3 Important Features
Logistic Regression	0.99	0.8	0.88	0.81	
Logistic Regression (SMOTE)	1	0.72	0.84	0.75	duration, euribor3m, poutcome success
Decision Tree	0.94	0.96	0.95	0.91	euribor3m, poutcome success, month_may
Decision Tree (drop duration)	0.91	0.98	0.95	0.90	euribor3m, poutcome success, pdays
Random Forest	0.91	0.99	0.94	0.90	euribor3m, pdays, cons.conf.idx
Random Forest (Grid Search)	0.91	0.98	0.95	0.90	euribor3m, cons.conf.idx, pdays
XG Boost	0.91	0.98	0.95	0.90	euribor3m, poutcome_success, month_may

XG Boost Feature Importance:

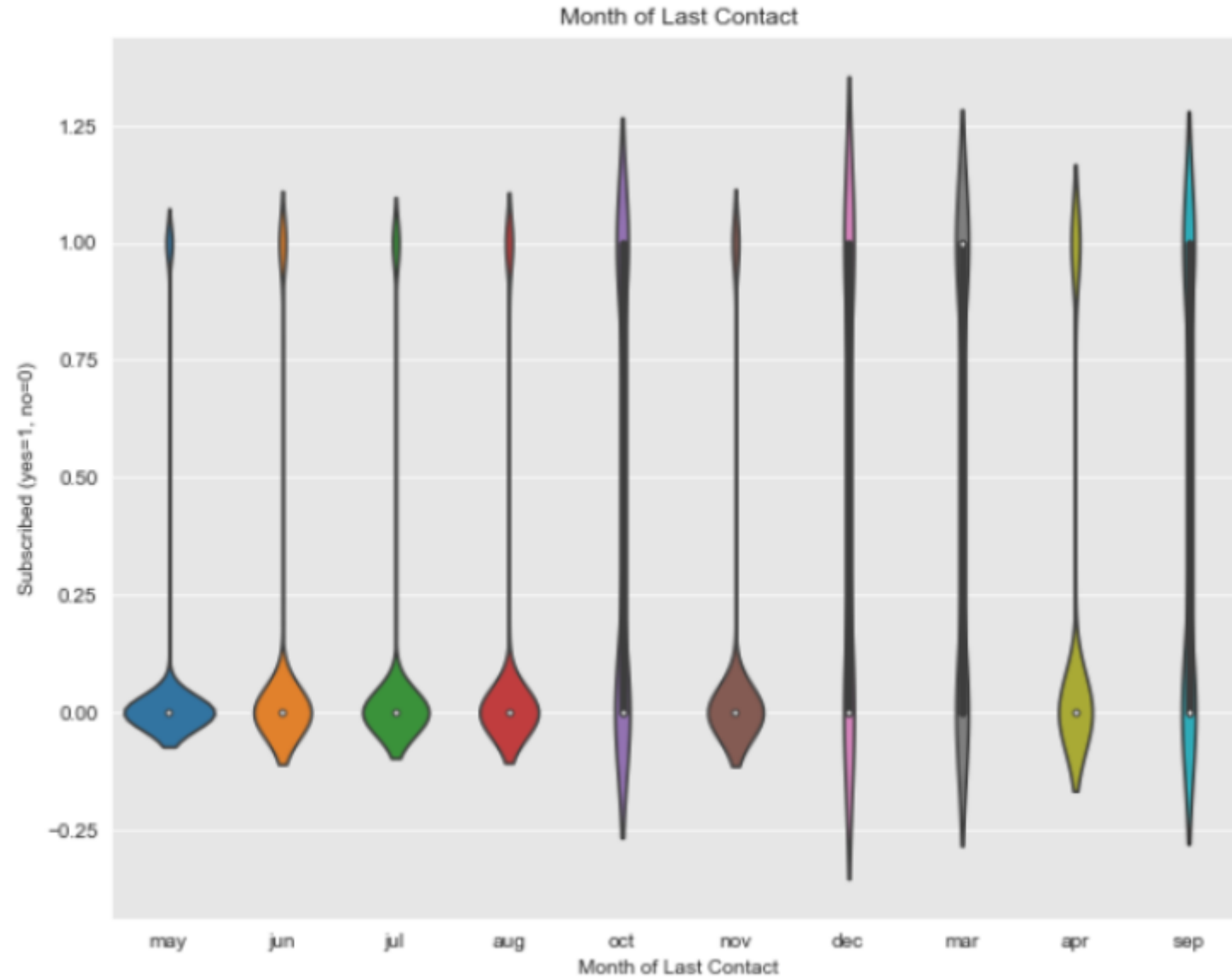


Feature Analysis





Feature Analysis



Models –

- Decision tree, random forest, XG boost all performed similarly
 - Precision – 0.91-0.94
 - Recall – 0.98-0.99
 - F1-Score – 0.94-0.95
 - Accuracy – 0.90

Feature Importance –

- Euribor 3mo interest rates was the most important model variable

Recommendation –

- After investigating 3 of the top important features, it's recommended that banks focus in on contacting previous subscribers closer to the end of the calendar year especially while interest rates are low.

Future Work –

- Investigate other top important features such as income, consumer price index, education, etc. to help with optimizing future bank marketing campaigns.

Thank You