



Rock Central Case

Data-Driven Approach to Predict the Success of Bank Telemarketing

Michael Harder

July 23, 2021



Agenda

Introduction

Data

Modeling

Next Steps

Appendix



Introduction



Background:

- Portuguese retail bank's telemarketing campaign data collected from 2008 - 2013
- Goal of the campaign calls were to sell bank long-term deposits
- By leveraging this data to predict the success of a campaign call, we can better understand which customers we should be targeting with telemarketing campaigns



Objective:

- Create a model to predict if a client will subscribe to term deposit (classification problem)
- Evaluation metric: accuracy



Source: [Moro et al., 2014] S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. Decision Support Systems, Elsevier, 62:22-31, June 2014



Agenda

Introduction

Data

Modeling

Next Steps

Appendix

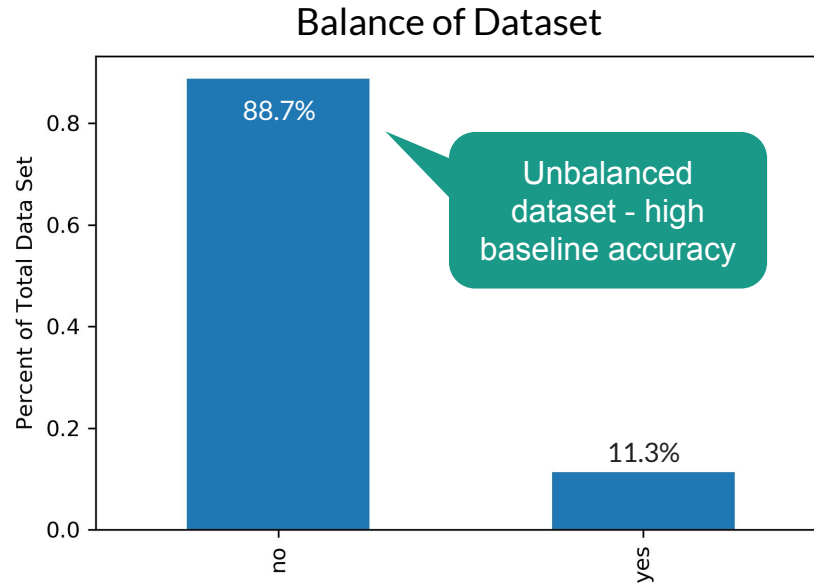


Data

- Made public on the [UCI Machine Learning Repository](#)
- Direct marketing campaigns data from a Portuguese bank
- Roughly **41,000 observations**
- **20 features** which can be used for modeling
 - Call duration provided in dataset but cannot be used for prediction
- Data includes:
 - Bank client data
 - Last contact information of the current campaign
 - Social and economic contextual features
 - Additional client campaign attributes

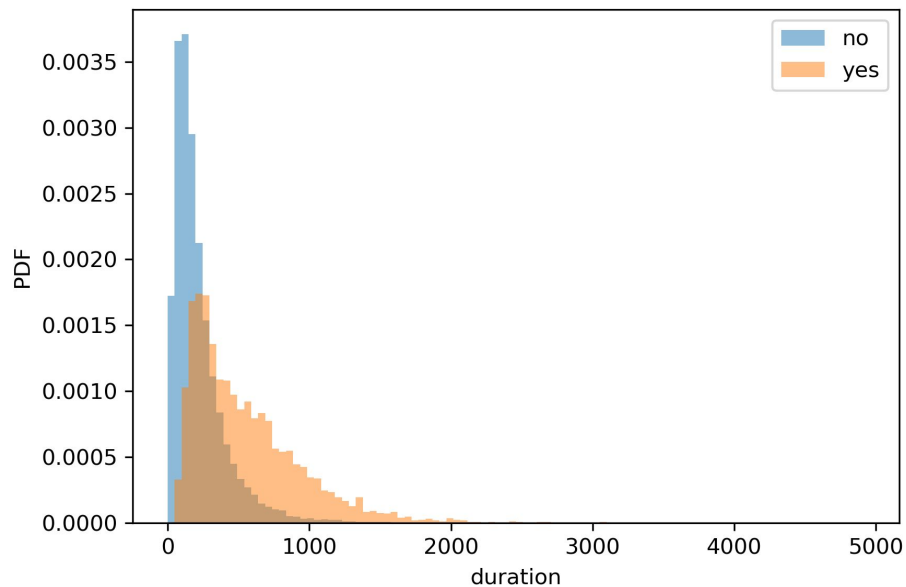


Data Label



Data Duration

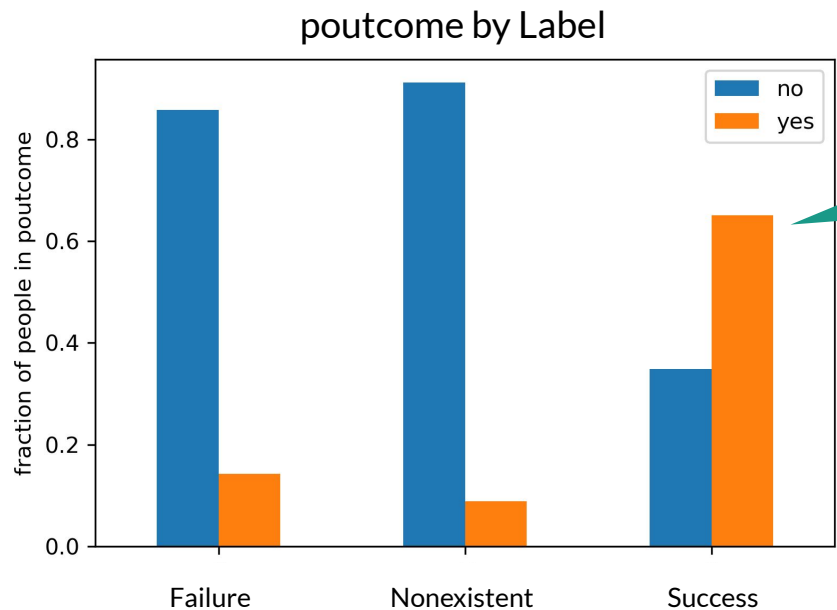
Duration Probability Distribution
Function by Label



'Duration' seems to separate the data well; however, we cannot know call duration before the campaign call.

Data

Previous marketing campaign outcome



'Success' only
accounts for 3.3%
of observations





Agenda

Introduction

Data

Modeling

Next Steps



Modeling

Models tested using stratified k-fold cross validation optimized on accuracy

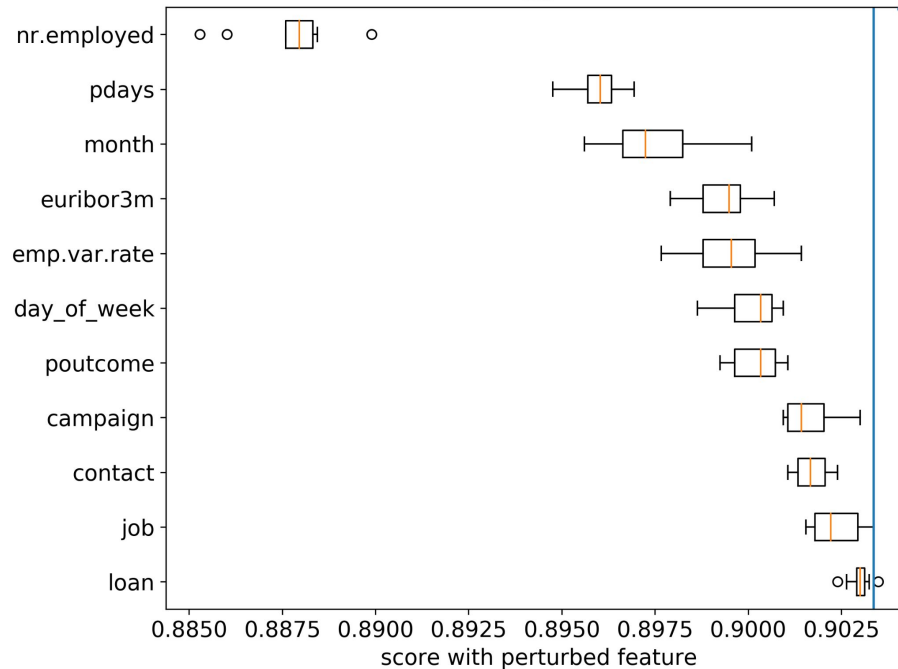
Model	Accuracy	Precision	Recall	F Score
KNeighbors	89.96%	67.0%	21.4%	0.325
SVC	90.13%	66.4%	25.1%	0.364
XGBoost	90.18%	68.1%	24.1%	0.356
MLP	90.34%	68.9%	26.0%	0.377



Modeling

MLP - Feature importance

Features dropped from modeling informed by permutations test: age, previous, default, cons.price.indx, education, marital, cons.conf.inx, housing





Agenda

Introduction

Data

Modeling

Next Steps

Appendix





Next Steps

- **Seek further model improvements**
 - Feature engineering and deploying bayesian optimization over grid search
- **Explore additional features**
 - Some of the most important variables for the model were economic and social context features; similar features could help improve the model accuracy
- **Identify ideal customer profiles**
 - Optimize outreach efforts of telemarketing campaigns
- **Understand client behavior**
 - Model does slightly better than a naive guess; if model improvements are not possible, drive additional business impact by increasing understanding of what makes a client subscribe to term deposit
 - E.g. understand the directional impact of *pdays* - a potentially actionable feature





Agenda

Introduction

Data

Modeling

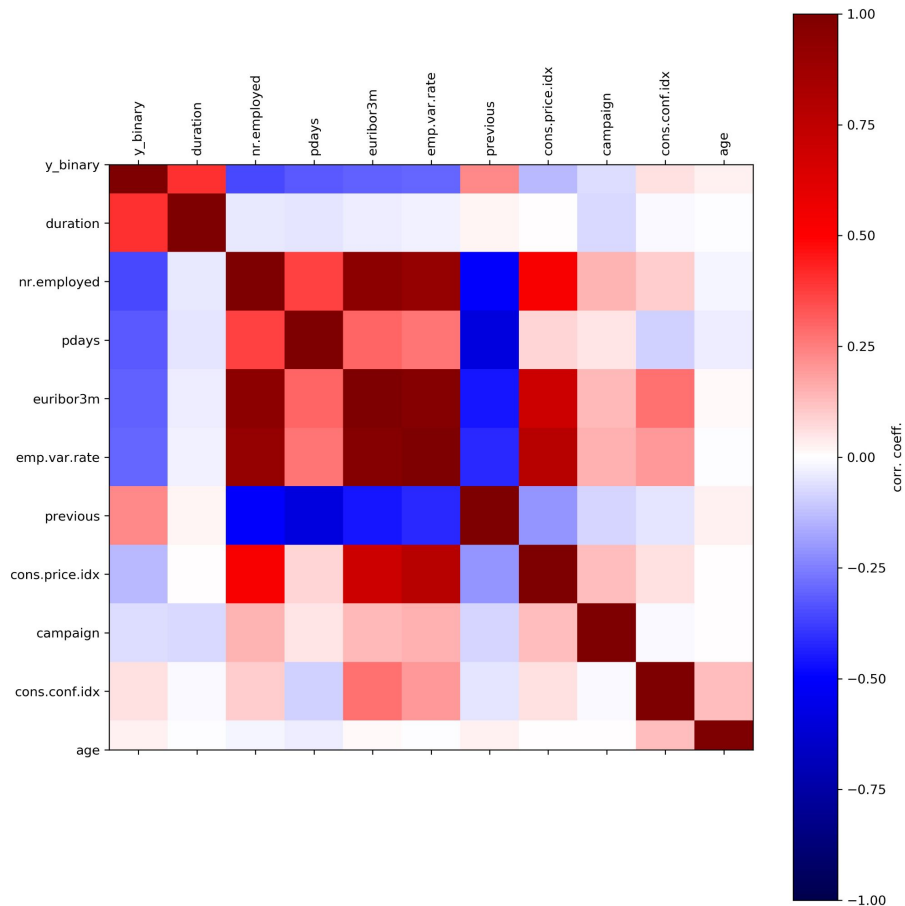
Next Steps

Appendix



Appendix

EDA - Correlation matrix



Appendix

MLP - Confusion matrix

