

The background features a dark blue grid. Overlaid on the grid is a white line chart with approximately 15 data points connected by thin white lines. The chart shows a fluctuating upward trend, starting from the left and ending on the right side of the frame.

# Predictive modeling of perspective customers opening a bank account

Yeonjoo Smith


# Data and Motivation

## □ Features

- Personal information
- Financial information
- Information about campaigns
- Economic situation at the time of the campaign

## □ Outcome: Success for a term deposit subscription





Can we build a model that predicts  
if a client is going to subscribe a  
term deposit using the information  
collected?



If we can build such a model, can  
we have the insight of which  
features are more influential to  
decide the outcome?

# Modeling Process

5

## Preprocessing data


- ✓ Loading and cleaning data
- ✓ Pick features that are relevant to the outcome and not strongly correlated to each other
- ✓ Divide the data set into the train data set and the test data set

## Machine Learning

- ✓ Test different models
- ✓ Feed the training data set of features and outcomes into each models to "train"
- ✓ Pick the optimal model and tune it for the best result for prediction

## Prediction & Evaluating Performance

- ✓ Feed a test feature data into the model
- ✓ Compare the prediction with the corresponding output
- ✓ Calculate the accuracy and other metrics to evaluate the performance



Can we predict if a client is going to  
subscribe a term deposit using the  
information collected?



Yes! For the test data, the prediction  
achieved over 90% accuracy.



If we can build such a model, can  
we have the insight of which  
features are more influential to  
decide the outcome?

# Feature importance

## Economical/ social situations

- euribor3m: 3 mon. maturity Euribor interest rate
- cons\_conf\_idx: consumer confidence index
- cons\_price\_idx: consumer price index

## Previous experience

- poutcome\_success
- previous
- campaign

## Communication type

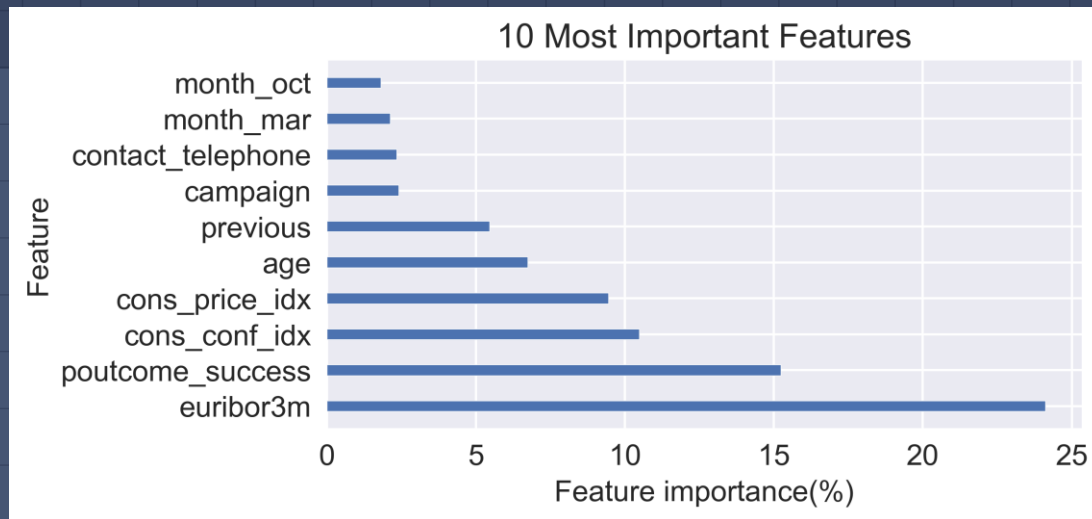
- contact\_telephone

## Time of the campaign

- month\_mar
- month\_oct



## Personal information

- age





## Limitations and Future Direction

- ❑ The model predictions and feature importance should be used as clues not as the whole picture 
- ❑ Combine business ideas and feature importance to develop a plan to increase the success rate 
  - Methods and frequency of contacting clients
  - Maintaining the connection with current/perspective clients
  - Campaign targeting timing and specific age groups

Thank you!