BANK TELEMARKETING STRATEGY ANALYSIS

Sally Cao
Sudarshan Kumar
Sunil Mathur
Drew McKnight
Amanda Morley

MEET OUR TEAM

Sally Cao

Business Analyst, Credit Card Marketing Growth

Sally works in the governances and monitoring team. She creates control management and key risk indicator reports within Centralized Pricing in the credit card space.

Sudarshan Kumar

Software Engineer, Home Lending Advice

Sudarshan is a software engineer that designs and develop software for the loan organization application and the modernization process.

Sunil Mathur

Architect, Data and Engineering

Sunil is part of the CCB Architecture, Data and Engineering team. He is responsible for the Architecture for Digital Behavioral Analytics Platform.

Drew McKnight

Software Engineer, Home Lending Technology

Drew is a software engineer within CCB Lending Technology. He creates Web applications using Java and spring boot.

Amanda Morley

Financial Analysis, Chase Auto Finance

Amanda is responsible for Planning and Analysis within the Chase Auto line of business. She is responsible for monthly reporting, budgeting , forecasting, and quarterly earnings reporting.

Machine Learning Decision Making Via The AI Canvas

What task/decision are you examining?

The task being analyzed is determine whether or not a customer will subscribe to a term deposit using a marketing campaign data set targeted to existing customers from a Portuguese banking institution: link to data https://www.openml.org/d/1461

Prediction

Predict whether an existing customer would subscribe to a term Deposit Product based on a phone Marketing campaign.

Judgment

Look at the Customer demographics and previous campaign data and determine the following payoff from the prediction.
Right - the bank will have a customer enrolled into the term deposit with another product from bank
Wrong – customer does not want the deposit and/or move to new company after annoyed with reaching out about campaign

Action

Provide a campaign that fits the customers needs with the bank.

Customer opens or declines a Deposit account with the Bank

Outcome

Determine the subscribed acceptance rate of the campaign

Training

Historical marketing campaign data matched with the historical outcome data. This data is used to train and calibrate the AI model before it is deployed

Input

Historical banking data on existing customers to get a better understanding on which ones to target for campaign

Feedback

Historical data from similar marketing campaigns in the past is compared with the outcome of the current campaign. The feedback is then incorporated to update the model, improving the AI

How will this AI impact on the overall workflow?

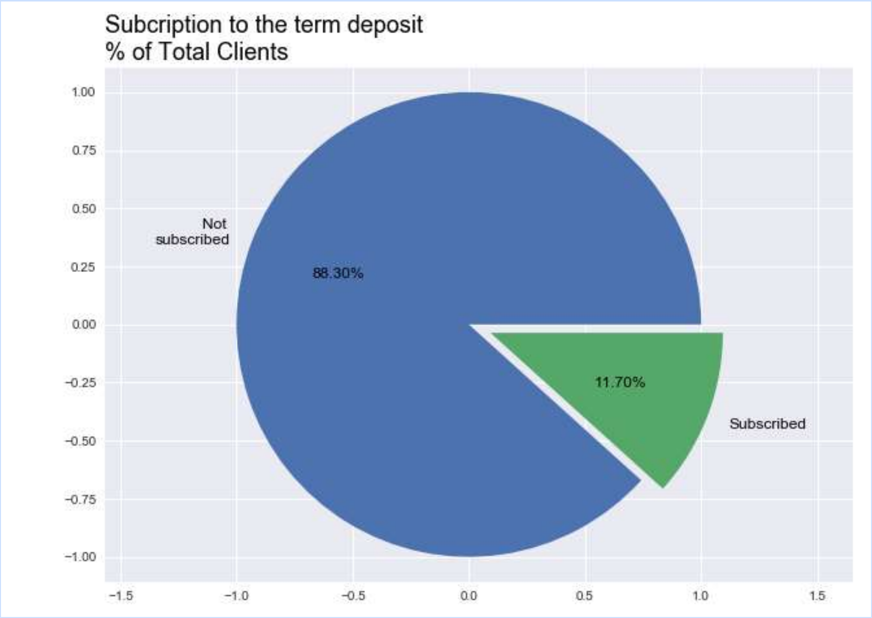
Will give Bank the opportunity to understand Customer preferences based of the feedback received from the calls in the campaign and determine if any new or better product is better suited for the particular Customer demographics.

EDA

Dataset

Key Findings:

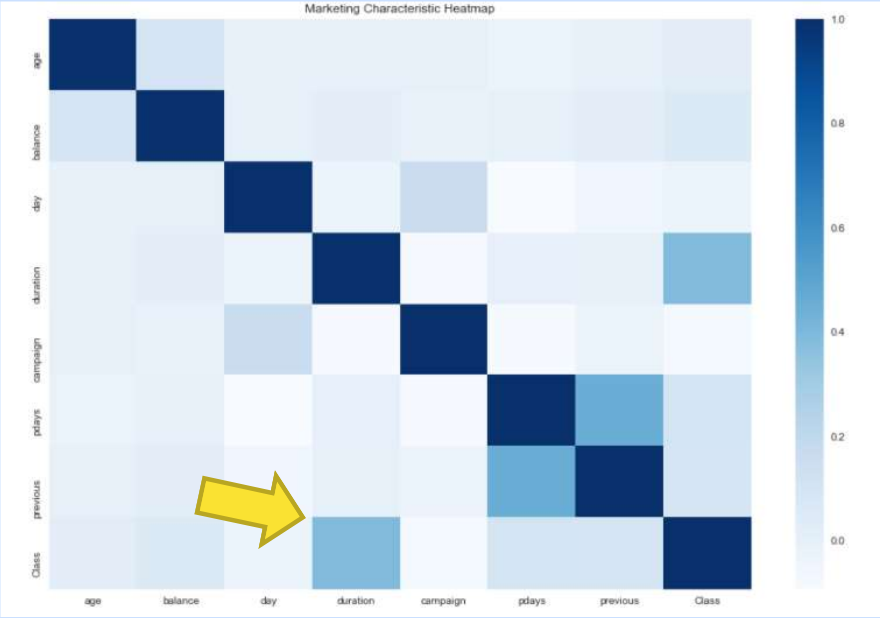
- Relatively clean dataset: 45,211 rows and 17 columns, and no missing values
- Most instances did ***not*** subscribe



Initial Correlation Testing

Key Findings:

- Initial correlation characteristic heatmap on numerical categories
- Duration was noted to be a highly correlated feature in our initial EDA



Tuning & Baseline

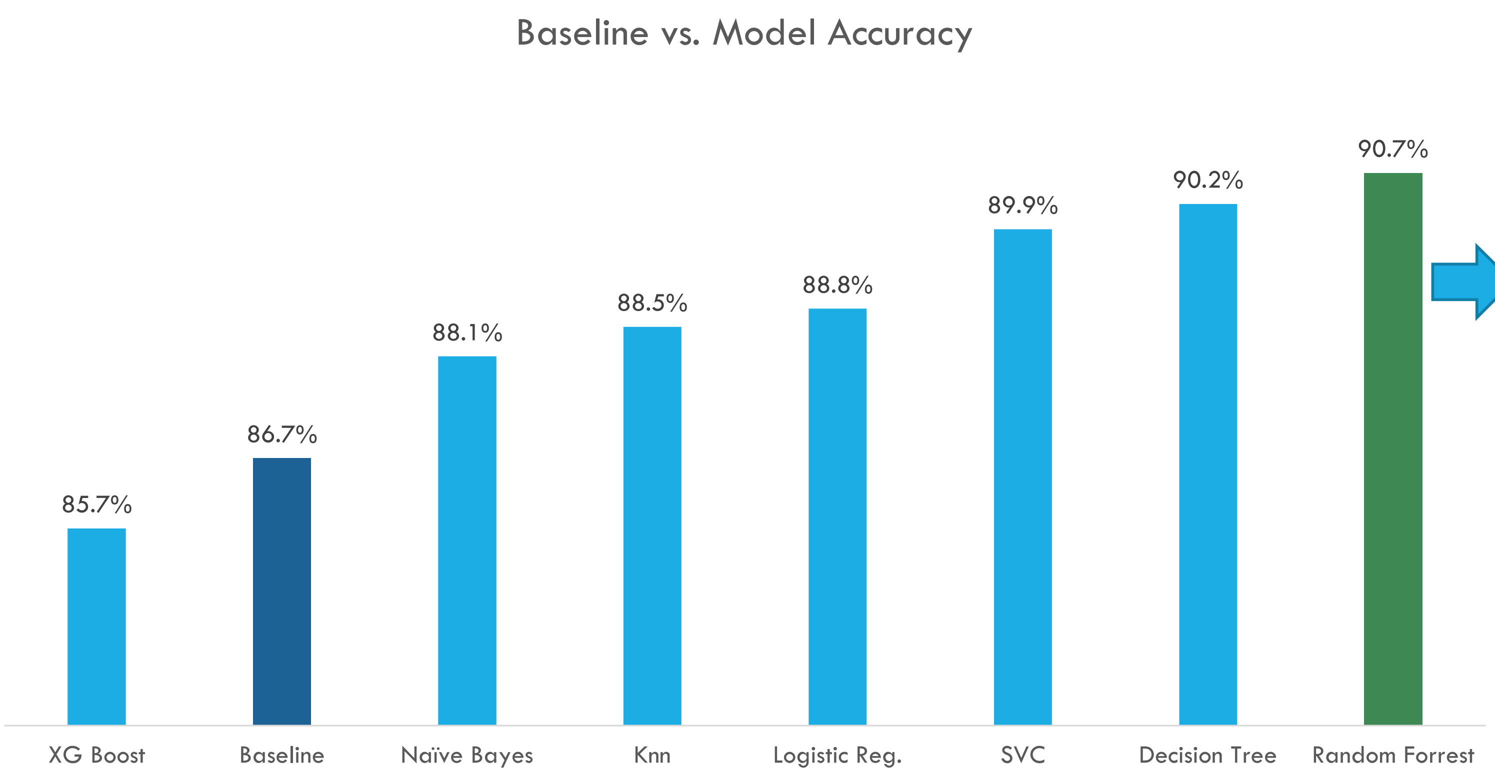
Model	Tuning	Score
Baseline (Non modeled)	Poutcome (Previous Outcome Success)	0.8830
Decision tree		0.9022616677726166
Random Forest	Gini, entropy	0.9064919265649193
SVC		0.8987502764875027
Xgboost (In progress)	Best parameters: {'model__max_depth': 2, 'model__n_estimators': 100, 'pca__n_components': 20}	0.857387
Logistic regression (In progress)		

Model Selection:- "Random Forest" with Best Score "90.63%" is selected as the final model.

MODEL SELECTION

<u>Models:</u>

- We tested 7 different types of models and have chosen Random Forest as the most accurate



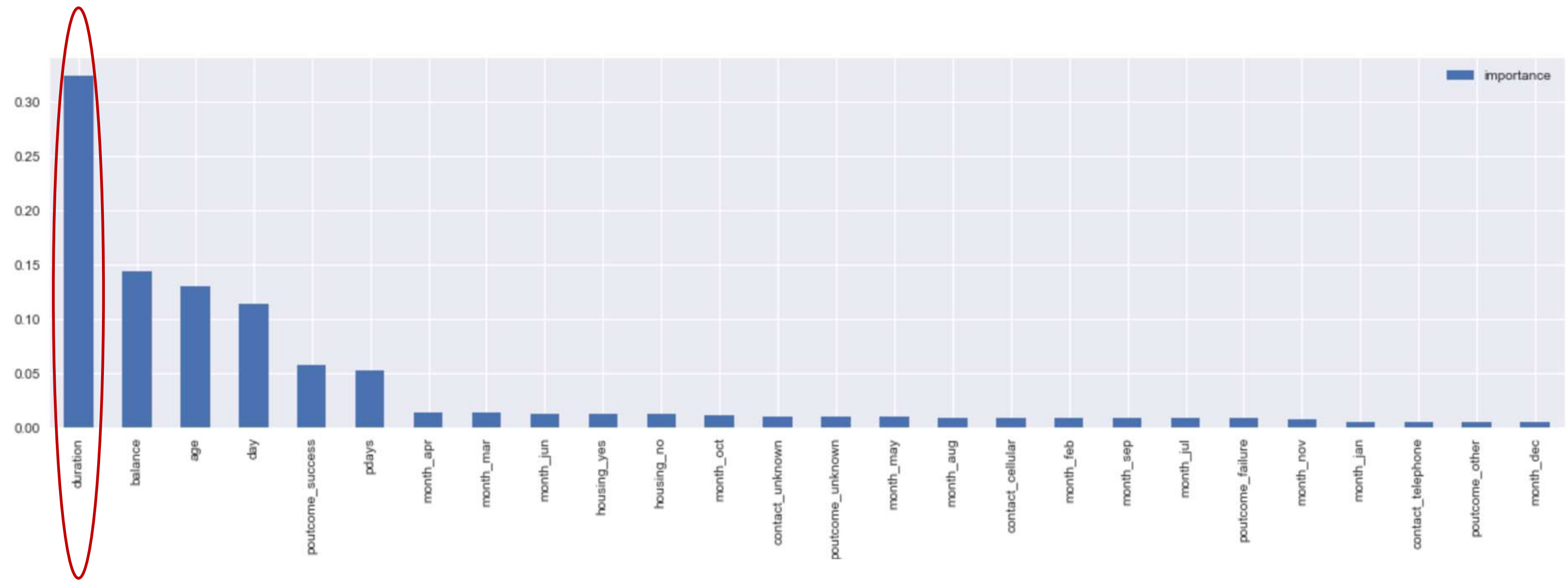
Model features selected in this model:

1. Age
2. Balance
3. Day
4. Duration
5. Pdays
6. Housing
7. Month
8. Outcome
9. Contact

FEATURE IMPORTANCE

Random Forest:

- For our chose model we graphed the feature importance and similar to our early EDA duration is the most important



KEY INSIGHTS

<u>Duration:</u>

- We have noted that duration seems to be the key feature for prediction. Our suggestion is that the bankers are trained to engage the potential clients in meaningful conversation as they will be more likely to subscribe.

<u>Balance:</u>

- Balance was also an important feature that could be researched in advance by the bankers. This could help target better who to market to.

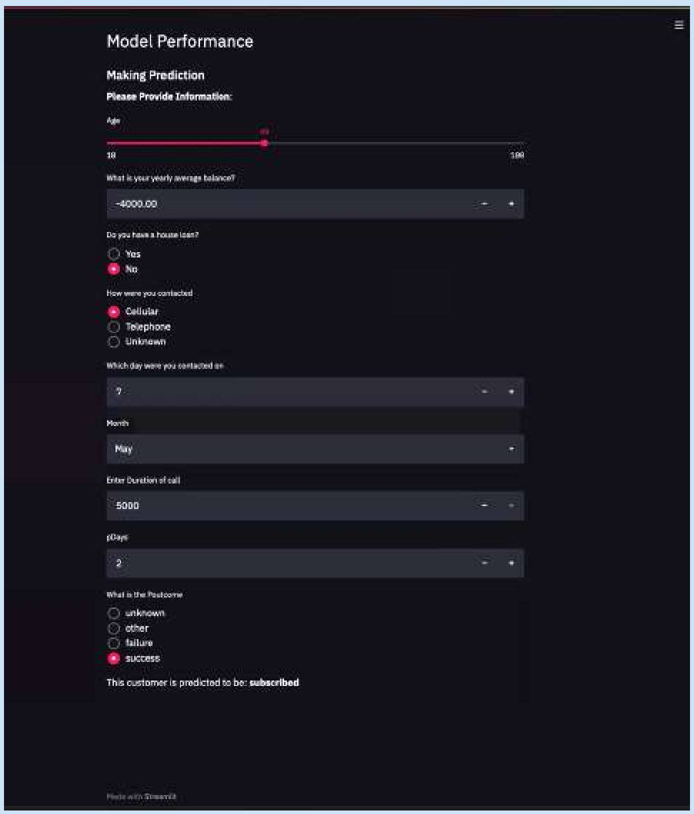
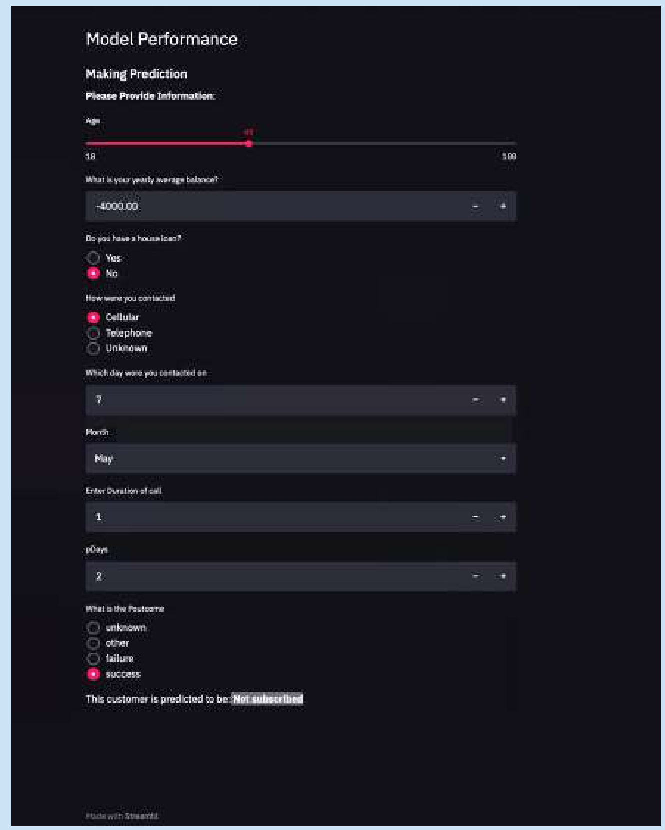
<u>Age:</u>

- Age also appeared to be a considerable feature. This will help to target demographics that are more likely to subscribe.

<u>Day:</u>

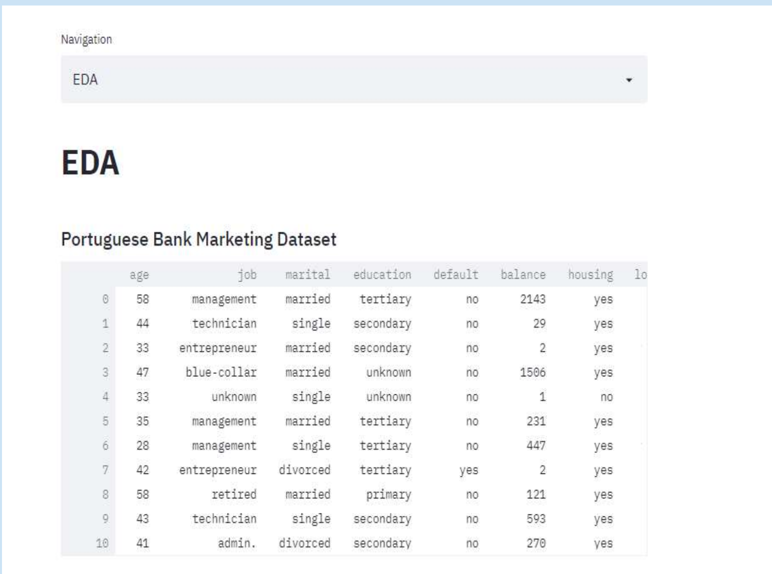
- Day of the week seemed to also be an important feature. This could be driven by subscribers being more relaxed on specific days and could change the way that the bankers are staffed.

ML WEB APP DEMO

Subscribed Demo	Did not Subscribe Demo
	

ADDITIONAL STREAMLIT VISUALIZATIONS

[Streamlit (team5ud.herokuapp.com)]



Navigation

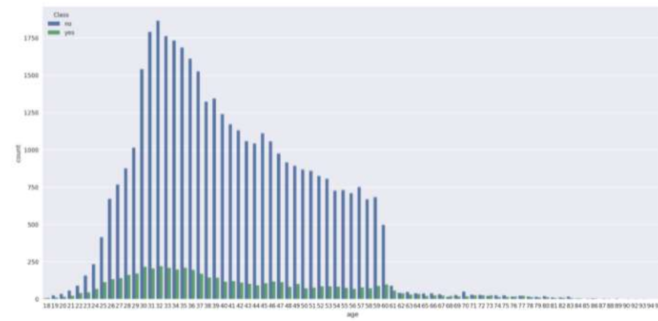
EDA

EDA

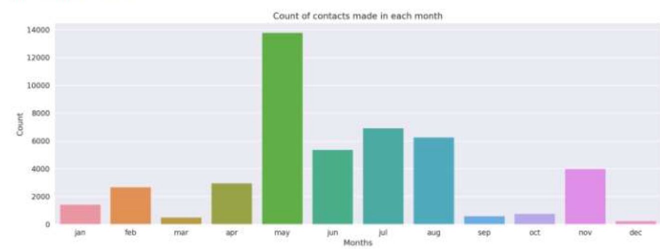
Portuguese Bank Marketing Dataset

	age	job	marital	education	default	balance	housing	lo
0	58	management	married	tertiary	no	2143	yes	
1	44	technician	single	secondary	no	29	yes	
2	33	entrepreneur	married	secondary	no	2	yes	
3	47	blue-collar	married	unknown	no	1506	yes	
4	33	unknown	single	unknown	no	1	no	
5	35	management	married	tertiary	no	231	yes	
6	28	management	single	tertiary	no	447	yes	
7	42	entrepreneur	divorced	tertiary	yes	2	yes	
8	58	retired	married	primary	no	121	yes	
9	43	technician	single	secondary	no	593	yes	
10	41	admin.	divorced	secondary	no	270	yes	

Age Feature



Month Feature



Based off the chart above, May has the most contact made to the customer (13766) but the least amount of subscription rate (7 %) compared to March which has the second least contact