

Group name: crispy chicken sandwich

Group members:

Loh Wan Teng, michellelohwt0799@gmail.com, Malaysia, Universiti Sains Malaysia

Zhechen Zhu, zhechenz@seas.upenn.edu, China, University of Pennsylvania

Zhihui(Angela) Chen, zhihuichen085@gmail.com, China, Brandeis University, Data Science

Problem description: The ABC bank aims to launch a new product, before they do that, they want to develop a model to help them understand what kind of customers would buy the product. In other words, based on the model built on different features of customers, they want to figure out the features that make the most difference to the outcome.

Business understanding: Based on the machine learning model, we hope to work out with the most efficient marketing strategy. The machine learning model would tell which feature matters most, meanwhile visualization results could also tell the clusters in each feature. For example, if it turns out that the job matters most, and people in the type of management are most likely to purchase for the product, then the main target of the marketing would be the people in management category with specific frequency.

Project lifecycle with ddl:

Deadline	Project Lifecycle
19 August 2022 (Week 7)	<ul style="list-style-type: none">• Problem description• Business understanding• Project lifecycle with deadline• Data Intake Report
26 August 2022 (Week 8)	<ul style="list-style-type: none">• Problem description• Data understanding• Data analysis<ul style="list-style-type: none">◦ NA values, outliers, skewed data analysis◦ Data processing and description
2 September 2022 (Week 9)	Data Cleansing and Transformation <ul style="list-style-type: none">• Data cleaning with 2 techniques• Team code review

9 September 2022 (Week 10)	<ul style="list-style-type: none"> • Problem description • EDA • Final Recommendation • EDA submission
16 September 2022 (Week 11)	<ul style="list-style-type: none"> • EDA Presentation • Modeling Technique Proposal
23 September 2022 (Week 12)	<ul style="list-style-type: none"> • Model Selection • Model Building
30 September 2022 (Week 13)	<ul style="list-style-type: none"> • Final Project Submission • Final Project Presentation

Github link: <https://github.com/AZHChen/ds-marketing-ml-project.git>

Data intake report

The dataset used for analysis is bank-full, accessed from UCI database.

Tabular data details: bank-additional-full

Total number of observations	41188 rows
Total number of files	1 file
Total number of features	21 columns
Base format of the file	.csv
Size of data	6.6+ MB