Personalized Product Recommendation

(Midterm Report)

Rishabh Gupta

2nd October 2023

Outline

- Introduction
- Data
- EDA with Data Visualization
- Data Preparation
- Production View
- Project Timeline

Introduction

- Project background and context

The project's objective is to predict the probability of customers adopting Ecoshare—a product that allows both new and existing NRG Energy customers to reduce their environmental impact by investing in carbon offset purchases.

This predictive analysis will then be leveraged to anticipate customer interest and enable seamless integration with NRG Energy's array of products. This integration improves the customer experience by delivering personalized product recommendations aligning with their preferences and requirements.

Data

Data Source : NRG Team

Time Period : Jan’ 2017 – Jul’ 2023

Features : 33

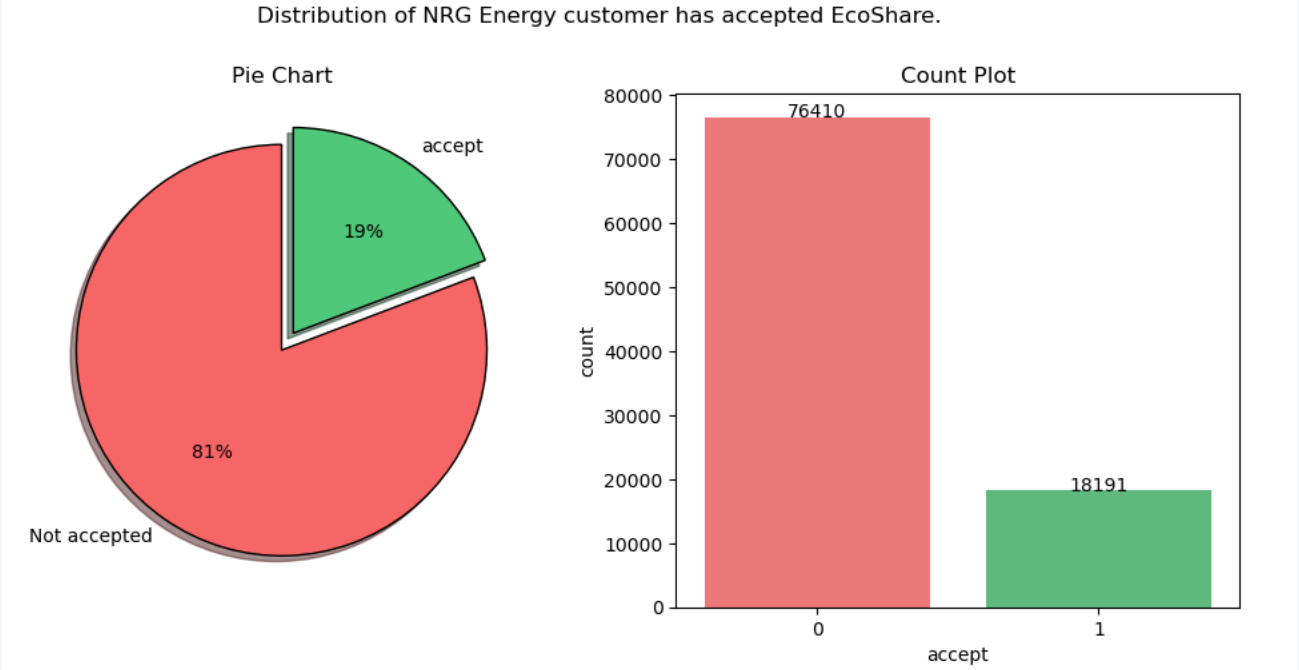
Target Variable : Binary

Unique Customer ~ 55k

Problem Type: Binary Classification (Ranking problem)

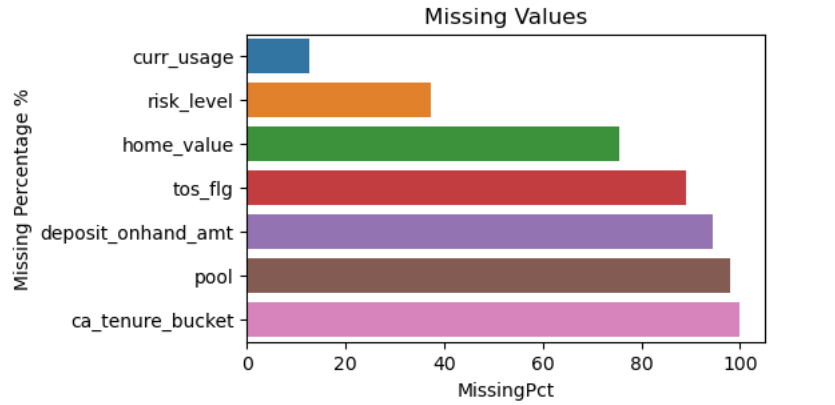
Evaluation Metric: Log loss

Target Variable Distribution

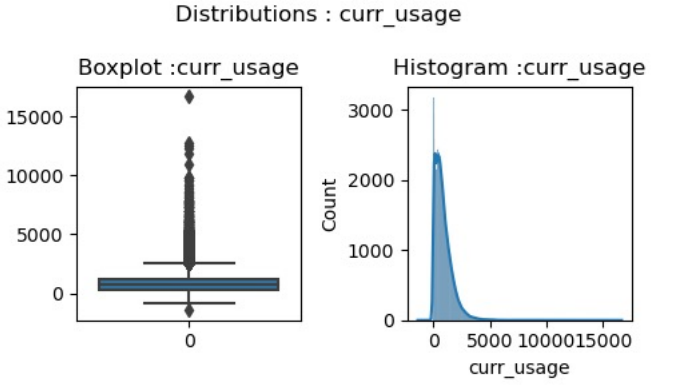


EDA with Data Visualization

- Explored the data by visualizing the relationship between target variable 'accept' and features.
- 5 features have more than 60% null values.
- `ca_tenure_bucket` feature have 100% missing data.



- `curr_usage' feature have negative values probably due to some people might be using solar panels.

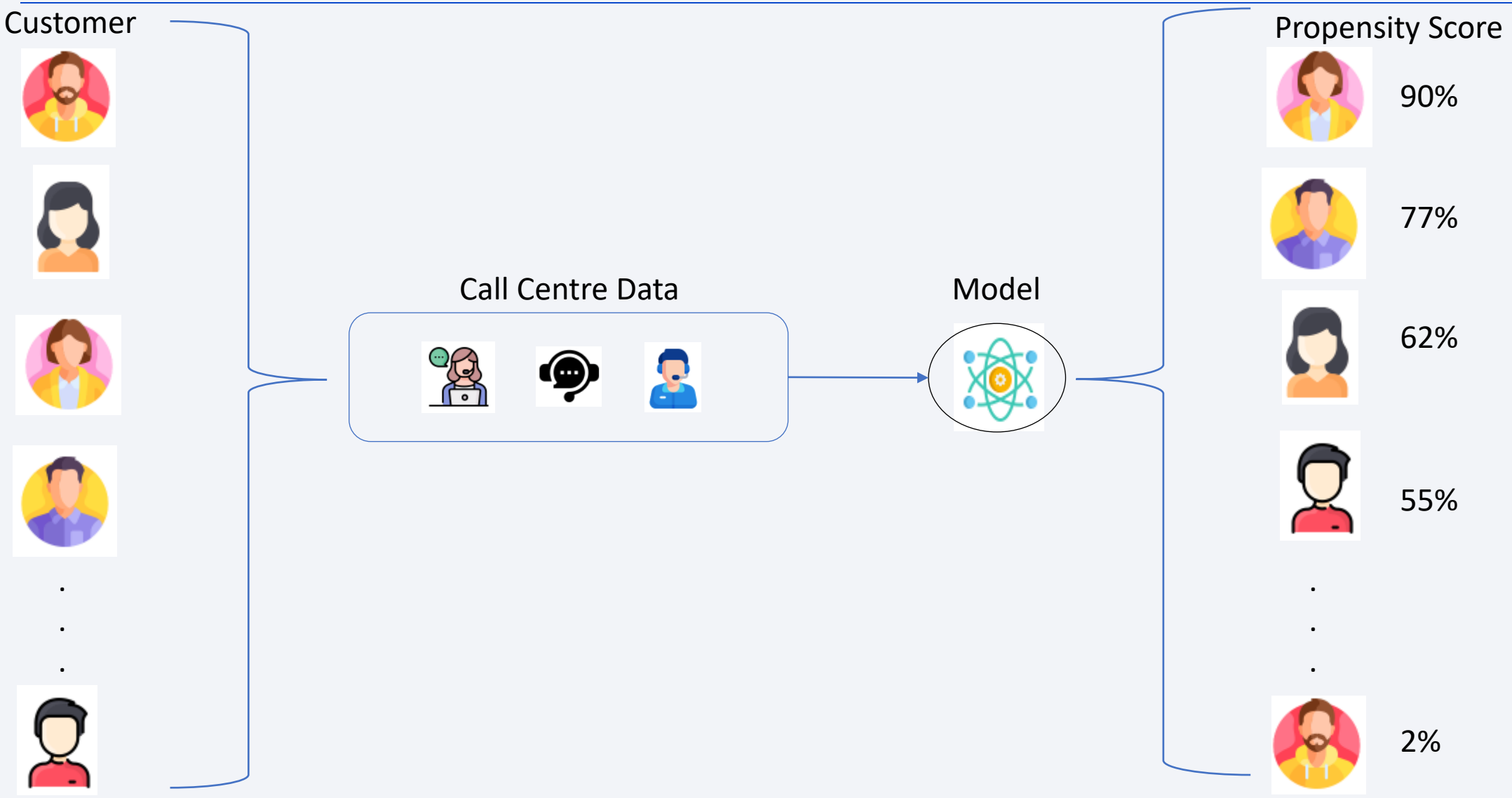


- A single customer may possess multiple residences, resulting in multiple contracts with NRG Energy.

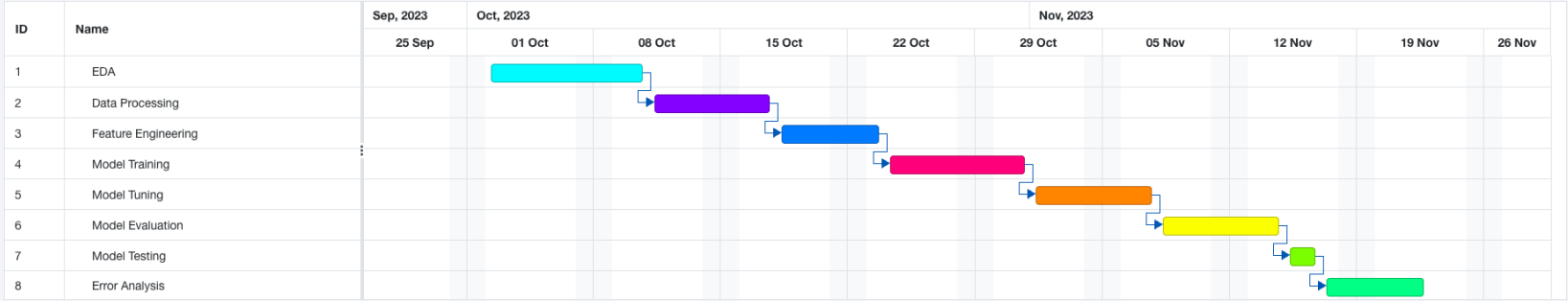
Data Preparation

- Following the discussions during office hours, customers will be presented with the EcoShare product based on their propensity score.
- In preparation for modeling, we will aggregate the data to the customer level through a process of data rollup. This involves consolidating and summarizing the relevant information at the individual customer level.
- We will remove a few features if they do not contribute significantly to predicting the output variable, aiming to streamline and improve the model's performance.
- We will attempt to impute the missing values with appropriate logical data.
- Perhaps incorporating external data, such as weather information, may have an impact.

Production View



Project Timeline




Thank you!