UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

DATA CONSULTING

DATA603-24X
DATA CONSULTING

| SHAHRON SHAJI |
| PRETTY MATHEW |
| MONISHA HENRY |

| JAMES WILLIAMS |
| ZAC TODD |
| KAREN SLABBERT |
| RICKI JONES |

PREDICTING POTENTIAL GYM MEMBERS

WITH SOCIO ECONOMIC PROFILES

OVERVIEW

Using socio-economic profiles like birthplace index, education index, income brackets (lower end), and distance to the gym, the conversion rate of users of Azira Data to gym members of the client was predicted using a gradient boosting regressor.

METHODOLOGY

The process begins with data cleaning and feature extraction to ensure accuracy and relevance. Next, indexes are developed, and conversion rates are calculated to derive meaningful insights. Finally, hypothesis testing and model development are carried out to support data-driven decision-making.

DATA CLEANING
↓
FEATURE EXTRACTION
↓
DEVELOPMENT OF INDEXES
↓
CONVERSION RATE CALCULATION
↓
HYPOTHESIS TESTING
↓
MODEL DEVELOPMENT

NULL HYPOTHESIS

The socio-economic factors (Common_Income_Lower_End, BPI, Education_Index) and the distance variable (Distance_In_Mi) have no significant effect on the Conversion Rate.

Prob(F-statistics) [1.11e-40] < Common Significance Level [0.05]

SOCIO-ECONOMIC FACTORS

Income Bracket

$$Income\ Bracket = Median\ of\ income\ range$$

The lower end of the most common income bracket to which most individuals in the SA1 belongs to.

Birth Place Index

$$BPI = \frac{Number\ of\ People\ Born\ in\ Australia}{Number\ of\ People\ Born\ in\ Australia\ and\ Overseas}$$

It is the probability of being born in Australia of a particular SA1

Education Index

$$Education\ Index = \frac{Total\ Educated}{Total\ People}$$

It denotes the educated individuals relative to population in a SA1

MODEL 1 – CEL Conversion Rate

RMSE - 0.028
MAE - 0.008
R-SQUARED - 0.836

MODEL 2 – First Party Conversion Rate

RMSE - 0.007
MAE - 0.004
R-SQUARED - 0.881

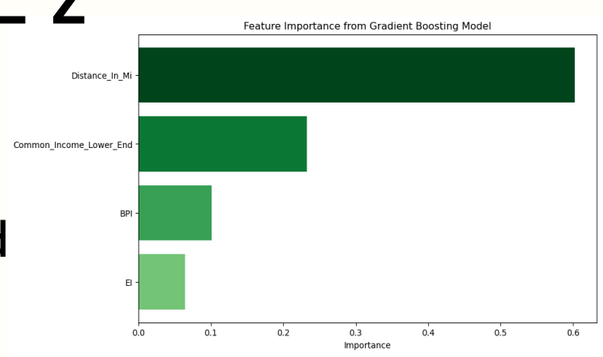
MODEL 3 – CEL to First Party Conversion Rate

RMSE - 0.119
MAE - 0.054
R-SQUARED - 0.763

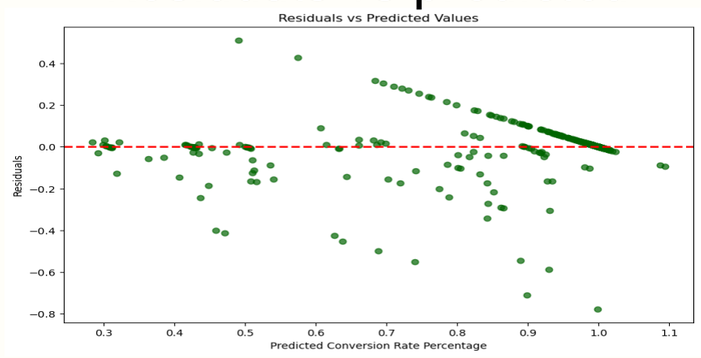
MODEL 2

Working model as per tested

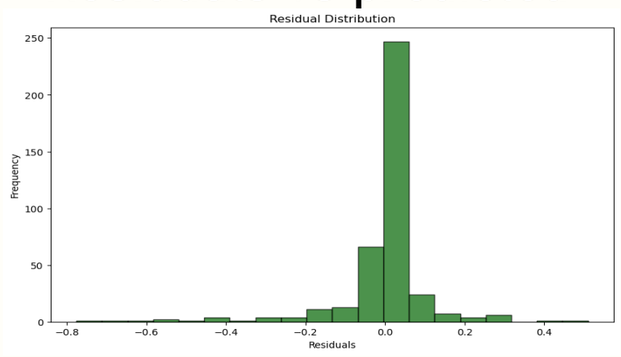
Feature importance



Residuals vs predicted



Residuals vs predicted



CONCLUSION

SUCCESSFULLY CREATED A MODEL THAT PREDICTS MEMBER CONVERSION RATE WITH SOCIO-ECONOMIC DATA AT 80% CONFIDENCE

MODEL PROVES EFFECTIVE AND RELIABLE WITHIN THE CONFIDENCE LEVEL WITH THE DATA PROVIDED

TESTED THE PREDICTIVE MODEL WITH A DIFFERENT AZIRA DATASET FOR A DIFFERENT LOCATION AND FOUND TO BE WITHIN THE CONFIDENCE RANGE