The Analytics Team

- Sprocket Central Pty Ltd

Data Analytics Approach

Raghul V

Contents

Introduction

- Sprocket Central Pty Ltd has reached us for an analytical approach to leverage their existing business. Supporting to that they have also provided sufficient data to interact with.
- However every data is not perfect or accurate, so to carry out further analytical process we have few steps to go through.
- This includes some agenda like data exploration, model development and data interpretation. On the account of best results it is must to carry out these steps carefully.
- This presentation briefs us about the above mentioned agendas.

Data Exploration

- Understanding the characteristics of the given data in means of all aspects like variable distribution, data validity, etc.
- For example, a training dataset may be highly skewed towards the younger age bracket so how will this affect the process of analysing other remaining dataset
- There some limitations in the data provided such as the missing values of a particular deposit and different data types corresponding to the values assigned.
- Transformation of data is required in order to produce the most sought deliverable.
- Merging the cleansed data with the existing ABS dataset.
- Improvising the given dataset if provided additional period/ time by the firm.

Model development

STEP 1: Determining an hypothesis to address the problem statement with the help of existing dataset.

STEP 2: Performing statistical analysis to validate the hypothesis.

STEP 3: Creating calculated fields in reference with the existing data.

- ***Example*** :- Converting date of birth (D.O.B) into age brackets.

STEP 4: Testing the performance of the model,

- Residual deviance
- Akaike Information Criteria
- Receiver Operating Characteristic Curve
- r squared interpretation

(Accordingly the model performance, assumption, and limitations.)

Data Interpretation

- Visualizing the processed data will give the best deliverable to interpret. Hence the deliverable is to be presented to provide the solution for the business problem.

- This process may include some interpretations like finding and iterating the significant values and coefficients from a business perspective.

- Hence the whole process may help us in ideating the best solution for the business problem and also will support the solution with a quantitative and qualitative observations noted during the whole data exploration, model development and visualization processes.

Appendix

Credits to,

- Sprocket Central Pty Ltd
- Australian Bureau of Statistics
- Python open source forum