

### Logistic Regression

#### Q1: Logit Model for Credit Grant Outcome use case

A group of 20 customers possess portfolio ranging from 0.5 to 6.5 million USD in one of the largest Financial Services majors in South America. As an analyst you are tasked to find out how the portfolio amount affect the probability of a customer getting a credit grant? Please refer to the dataset provided in *GitHub - CreditGrantOutcome.csv*.

Portfolio Value is given in million USD = X; And Credit Grant Decision = y which is either 0 or 1, i.e. 1 for getting a grant.

- Find out the Odds ratio for every customer data that is captured here. What is the Odds ratio when Portfolio Value X = 2 m USD?
- Find out optimum values of coefficients beta\_0 and beta\_1?

#### Solution:

Please refer to the Jupyter Notebook "AAI\_LR\_Assignment\_CreditGrantOutcome.ipynb" for details.

So, we can generate odds ratio for every customer data. For X=2, odds ratio is approx. 34.37%

Secondly, beta\_0 = -4.077 and beta\_1 = 1.504

#### Q2: Confusion Matrix Scenario

You are working on a binary classification model to predict which devices are going to fail or not fail in the future.

What is the accuracy of your model given the confusion matrix below? (Rounded to 1 place after decimal)

Actual Value	Predicted (No fail)	Predicted (Fail)
No fail	371	84
Fail	65	122

#### Solution:

Accuracy =  $(TP+TN) / (TP+FN+FP+TN)$

$$= (371+122) / (371+84+65+122)$$

$$= 493 / 642 = 76.79\% \sim 76.8\%$$

### Q3: Logit/GLM Model for Insurance Cross Sell use case

The insurance major is interested to understand the cross-sell opportunities by analysing the information it has. As a Data Scientist, you have access to the Insurance Cross Sell data. (Please refer to *GitHub - InsuranceCrossSell.csv*).

The description of variables are as follows (for your reference):

- *Response* is the binary outcome which indicates whether customer has taken the insurance or not. 1 if insurance taken Else 0.
- Other predictor variables are as follows: Gender, Age, Driving\_License, Region\_Code etc.
  - Gender of customer
  - Age of customer
  - Driving\_License (DL) – 1 if customer has DL, else 0
  - Region\_Code – location code to which customer belongs
  - Previously\_Insured – 1 if customer has prior insurance
  - Vehicle\_Age – 3 categories: <1 year, 1-2 year, >2 years
  - Vehicle\_Damage – whether there has been damage to the vehicle owned by the customer or not. 1 if there is damage in the past.
  - Annual\_Premium – amount paid by customer
  - Policy\_Sales\_Channel – code of the channel through which the policy is being sold
  - Vintage – number of days customer is associated with the insurance

Fit a model (using Logit or GLM) and explain significance of predictors on the "Response" decision.

#### Solution:

Please refer to the Jupyter Notebook "AAI\_LR\_Assignment\_InsuranceCrossSell.ipynb" for details.