# Task 2. Logistic Regression

**Due: Mon Apr 5, 2021 23:55**

**Unlimited Attempts Allowed**

Available until Apr 9, 2021 23:55

**Details**

Gowdafone is a telecommunications company that specialises in consumer and business mobile telephone services. Recently it has experienced a modest sales increase of the services they  offer to their customers. The VP for sales has tapped your expertise to help them predict if a customer will churn or not. 'Churning' is the turnover of customers leaving to going to competitive providers. Using the dataset provided, help Gowdafone to predict customer churn by performing the following tasks:

1. Perform data exploration and visualisation to gain insight about the dataset "train_churn.csv ".
2. Perform pre-processing if needed.
3. Split the train_churn dataset for training and testing (validation).
4. Systematically select features that will be used for your model.
5. Create a model that will predict if a customer will churn 1-YES, 0-NO
6. The model must be trained and tested (validated) using the split churn-dataset.
7. Make churn predictions using the separate test set named  "predict_churn.csv "

The following table describes the features on both datasets except for X20 which is omitted in "predict_churn.csv".

Dataset feature description

| Feature Names | Description |
|---|---|
| X1 | state where the subscriber resides |
| X2 | maintained the account, shown in number of months |
| X3 | area code of the subscriber |
| X4 | tells whether the subscriber has international calling plan |
| X5 | tells whether subscriber has a voice mail plan |
| X6 | number of voicemail messages |
| X7 | total minutes called during the day |
| X8 | total number of day calls |
| X9 | total amount charged for day calls |
| X10 | total minutes called during the evening |
| X11 | total number of evening calls |
| X12 | total amount of charged for evening calls |
| X13 | total minutes called during the night |
| X14 | total number of night calls |
| X15 | total amount charge for night calls made |
| X16 | total minutes of international calls |
| X17 | total international calls made |
| X18 | total amount charged for international calls made |
| X19 | number of times customer service was called |
| X20 | indicates whether customer churned or not |

Files to be submitted in Canvas

1. Jupyter notebook containing the codes demonstrating tasks 1 - 7. Make sure that your codes are error free as I will not debug it when and if errors occur during runtime.
2. CSV file containing the predictions of your model. The submitted predictions must be the same as the predictions produced when I actually run your code.
3. The tasks above  must be completed using either Jupyter Notebook or Google Colab, running Python 3.

Filename conventions :

**xxxx**_predict_churning.ipynb

**xxxx**_predictions.csv

Replace **xxxx** with your student number, for example:

**4321**_predict_churning.ipynb

**4321**_predict.csv

Help Me