DataScience Exercise

# Introduction

This exercise tests your ability to machine learning skills to build models for elegance. You should exploit the latest solution available in the market to your advantage. The less code you write to achieve a defined functionality, the more your score. Use proper packages to solve technical challenges.

## Exercise #1: Identify and build a predictive model for any dataset

You can use any dataset from [Kaggle](https://www.kaggle.com/) (read the following instructions and choose your dataset to cover the steps in the exercise) and then create the following for the chosen dataset. Make sure there are proper comments available in the code region for a better understanding of the dataset and logic.

Please use the Jupyter notebook to complete the exercise. Preferably python. Organize your code in sections as per the following details of the exercise. You are encouraged to learn classical Machine learning techniques and experience their behavior. Notebooks should be uploaded with output for evaluation.

**1. Import needed Libraries and Datasets from Kaggle**

a. Download the dataset

b. Import the required libraries

**2. Data Visualization and Exploration**

a. Print at least five rows for a sanity check to identify all the features present in the dataset and

if the target matches with them.

b. Print the description and shape of the dataset.

c. Provide appropriate visualization to get an insight into the dataset.

d. Try exploring the data and see what insights can be drawn from the dataset.

**3. Data Pre-processing and cleaning**

a. Do the appropriate preprocessing of the data, like identifying NULL or Missing Values, if

any, handling outliers in the dataset, skewed data, etc. Apply appropriate

feature engineering techniques for them.

b. Apply the feature transformation techniques like Standardization, Normalization, etc. You

are free to apply the appropriate transformations depending on the structure and the

complexity of your dataset.

c. Do the correlational analysis on the dataset. Provide a visualization for the same.

**4. Data Preparation**

a. Do the final feature selection and extract them into Column X and the class label into

Column into Y.

b. Split the dataset into training and test sets.

**5. Model Building**

a. Develop ANN Model

b. Train the model and print the training accuracy and loss values.

**6. Performance Evaluation**

a. Print the confusion matrix. Provide appropriate analysis for the same.

b. Predict the test data and display the results for the inference.

## Exercise #2: Deploy the model

Download the created model and deploy and consume the model and expose it as an API. Use postman to call the API and share the postman result in the screenshot.