****

**PROJECT WRITE-UP – APP RATING PREDICTION**

**Submission Deadline : 18-12-2022**

**Submitted on : 14-12-2022**

**Submitted by : Mani Shankar**

**PROJECT WRITE-UP**

**INTRODUCTION :**

The project that I chose to work upon is creating an App Rating Prediction model using Python programming and the coding environment I used is Jupyter labs. The purpose of choosing this topic is that it sounds a bit challenging and I would love to explore it and get better insights from the dataset.

In this write-up, I will elaborate on my approach / Procedural steps that I followed while creating the Prediction model.

**BUSINESS REQUIREMENT:**

The following are the project tasks :

* Formatting the dataset columns
* Converting the datatypes
* Removing unnecessary data (Both rows and columns).
* Identifying dependent and Independent variables for analysis.
* Performing Sanity checks
* Plotting visualisations and interpretations of the visualised data.
* Finding Outliers and correcting it.
* Reducing the skew for some variables
* Adding dummy columns
* Apply train, test and split to build a model
* Build a model using Linear regression technique and report R2

**OBJECTIVES :**

\*To derive the data and visualize it so as to make data more readable.

\*To get better insights on relationship between the variables and its impact.

\*To transform the data for ML models and present the data visualizations and app predictions the most efficient way possible.

\*To analyze the difference between actual and predicted Ratings for the Apps.

\*To use the model to train similar datasets in future.

\*To help businesses to analyze the future trends in terms of App Ratings.

\*To support the business to make important decisions based on predictions.

**MY APPROACH :**

\*Get the data

\*Perform EDA on data

\*Perform pre-processing

\*Have X and y

\*Split the data into train and test set.

\*Train the ML model ---> X\_train and y\_train

\*Predict the results

\*Evaluate the model ----> On train data and test data.