ENDTERM PROJECT:

TOPIC PROPOSAL

By

"DATA DOME" – TEAM 1

SNEHITHA TADAPANENI

SAI RACHANA KANDIKATTU

AMRUTHA JAYACHANDRADHARA

WILONA NGUYEN

PRAMOD KRISHNACHARI

Introduction to Data Science: DS 6101

Under the guidance of,

Prof. Divya Pandove

(TA: Aishwarya Maddula)

RESEARCH TOPIC:

ANALYSIS ON GOOGLE PLAY STORE APPS

Our research focuses on performing Modelling on Google Play Store apps to uncover patterns, trends, and insights regarding app characteristics, user behavior, and installation patterns. We are trying to see how app popularity, defined as the number of installs, is impacted by the top five categories, last updated, app sizes, version, and other factors.

SMART QUESTION

"Which are the top 5 app categories, as identified by classification models (logistic regression, SVM, XGBoost, KNN, and random forest), that significantly influenced app success (measured by installs) based on app data from 2010 to 2018, and how accurately can these models predict success trends within this time period??"

SOURCE OF THE DATASET (NO. OF OBS)

The dataset is taken from Kaggle (Link). There are 10841 observations of 13 features in the taken dataset.

TEAM's GIT REPOSITORY LINK: github.com