**Capstone Project Submission**

| **Team Member’s Name, Email :** |
| --- |
| 1. Shafaq Khan 2. Sumit ojha 3. Ayush Mishra   Email;   * ayushmishra0031@gmail.com * [shafaqkhan1711@gmail.com](mailto:shafquekhan1711@gmail.com) * [sumitojha1309@gmail.com](mailto:sumitojha1309@gmail.com) |
| **Please paste the GitHub Repo link.** |
| Github Link:- https://github.com/Ayushmishra1503/Google-play-store-analysis |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| Google play store Analysis project was done by group of 3 members – Shafaq Khan, Sumit Ojha and Ayush Mishra. In this project we had got 2 different csv files as an input. The csv files were play store data and user reviews. The objective of this project is to deliver insights to understand customer demands better and thus help developers to popularize the product. It is of 10k Play Store apps for analyzing the Android market. This dataset contains details of different applications and reviews from different users. Application distribution platform, for example, Play Store gets overwhelmed with millions of new applications being launched on the platform regularly. Lots of designers and developers work on it to make an app successful on the Play Store. In this highly competitive world, it is an immense challenge for a developer to know whether they are focusing on the right path to make their app successful on the platform.  To launch themselves successfully and create an identity for themselves in this oversaturated market, they need to ensure that majority of the essential factors are incorporated while designing and developing an app that would play an important role in customer's decision-making process. However, the lack of a clear understanding of the inner working and dynamic of popular app markets impacts both the developers and users.  The main objective of this exploratory data analysis project is to understand customer demands better and thus help developers to popularize their product on the Play Store. In this project we analyzed the data of Google play store apps like how many number of apps are available in dataset, what is there status, there ratings, genres, reviews percentage of free and paid apps, etc. For that we used data visualization tools such as pandas ,numpy , seaborn , matplotlib. And with the help of seaborn and matpolitb we plotted the graphs and shown the distribution of apps in different age group ,percentage of free and paid apps,reviews in the apps ratings etc. The apps within the Google App store could also be biased and overrated because higher ratings given by users potentially attract several new users disproportionately. This study therefore investigated the utilization of ensemble classifiers to predict numeric ratings for Google Play store apps supported the user reviews for those apps. Several ensemble classifiers were investigated to gauge their performance on the reviews scraped from the Google App store |