**Topic: Android App Market on Goole Play**

**Concept of the project:**

In Today’s world, there are millions of mobile apps which are getting deployed into the google play store, which are easy to create and are profitable. It can be for business’s purposes or sometimes done for testing, while few are done for malicious purposes to gain money by scamming the users, when they install the app into their phones. By this Data Classification Analysis, we can gather information about:

1.App Categories.

2.Understand different attributes of the dataset.

3.Analysis on the user’s reviews.

**Objective of the Project:**

Retrieve Sentimental analysis of the given app through descriptive analysis from the dataset.

* The analysis will provide the reviews of the users.
* Prediction of the user’s demand can be done.
* Lucrative and favourable app can be built.

**Data Sources used:**

<https://www.kaggle.com/datasets/kamganganthony/the-android-app-market-on-google-play>

**Data Analysis Software used:**

Python and Jupyter Notebook Libraries used:

* Pandas – For cleaning and manipulating the dataset
* Matplotlib – To create data visualization
* Plotly – To create interactive graphs
* Seaborn – To create data visualization
* Warnings – To alert about the condition in the program

**Data sets probable visualizations:**

Bar graphs, Histogram, Scatter plots will be used for the visualization of the given parameters.

**Methodology:**

In this project, the machine learning algorithm is constructed by cleaning the data by removing null values and unwanted characters, correcting the data types, analysing the features which can be worked upon, exploring the app categories to understand the active apps which are dominating the market and about the categories which are least used. After which, the ratings, price and size of the apps are analysed to understand the user’s demands and preferences.

There are various apps which are deployed with ulterior motives of scam and to develop their app development skills, which can be filtered out by checking the price feature. We can even check the difference of popularity between paid and free apps. For the sentimental analysis is done by mining the user review data and deriving the three parameters: Positive, Negative and Neutral.

**Data Visualization:**

The findings from the Project will be the different outcomes from the three features of price, size, and ratings of the apps in the google play store and the sentimental analysis of the user’s reviews.

These can be presented by using the visualization libraries of python: Plotly, Matplotlib and Seaborn.

**Conclusion:**

By the end of the project, we can derive all the information from the description analysis and through that we can construct apps which are lucrative and favourable for the business.