FANTASTIC
5+1

MOBILE APP SUCCESS PREDICTION

IS YOUR APP HERO OR ZERO ???

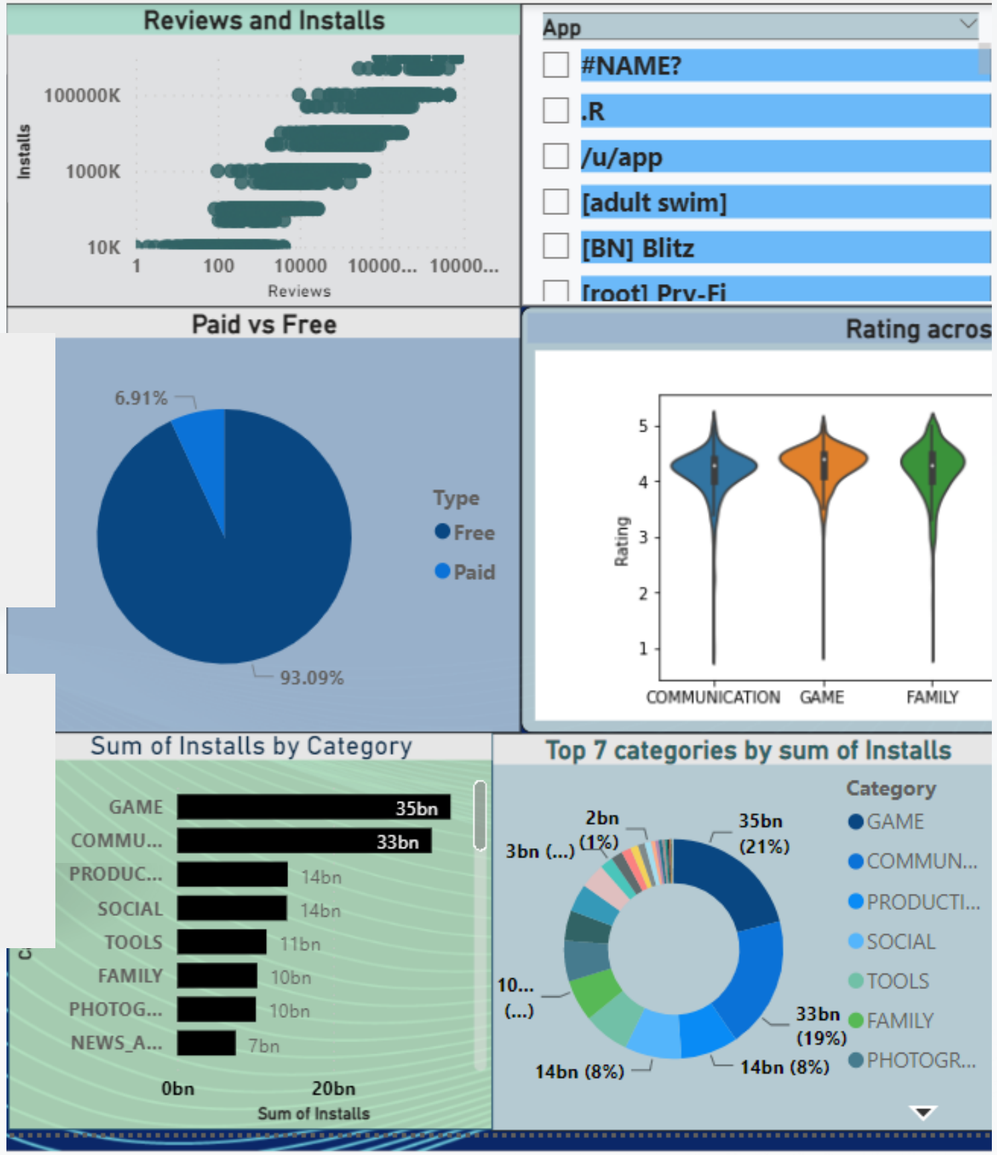
WHAT'S HAPPENNING

Input:

Rating, Reviews, Installs, Type, Price($), Content Rating, Size(MB)

Output:

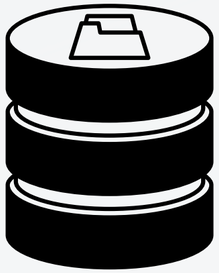
Successful or Unsuccessful



TARGET CUSTOMERS

DATA TALK

Dataset

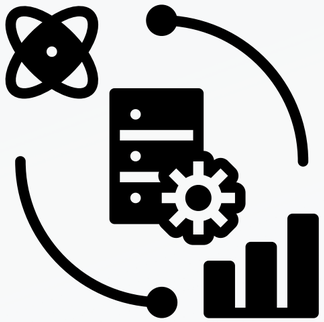


Source: Kaggle

Dataset: Google play store

- Rows: 10841
- Columns: 9

Preprocessing



- Converting into numerical values.
- Cleared the null values and then renamed certain columns for normalization.
- Implemented k means clustering to acquire the target variable.

Technologies Used

- NUMPY
- PANDAS
- MATPLOTLIB
- SEABORN
- PLOTLY
- POWER BI
- CHATGPT

Improved

- Knowledge in machine learning algorithms
- Better handling of data
- Dashboard implementation

BUSINESS DRIVER

- Investing resources in the right mobile apps
- Can increase their chances of generating revenue

- It avoids loss in your business
- If it is not performing well, upgrade with the required features

MATH-STAT TALK

K Means Clustering

In the Euclidean plane, let point p have Cartesian coordinates (p1,q1) and let point q have coordinates (q1,q2). Then the distance between p and q is given by

$$d(p,q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

$$\mathrm{WCSS}(k) = \sum_{j=1}^{k} \sum_{\mathbf{x}_i \in \text{cluster } j} \|\mathbf{x}_i - \bar{\mathbf{x}}_j\|^2,$$

where $\bar{\mathbf{x}}_j$ is the sample mean in cluster j

Decision tree

$$Gini = 1 - \sum_{i=1}^{n} p^2(c_i)$$

$$Entropy = \sum_{i=1}^{n} -p(c_i)log_2(p(c_i))$$

LIVE DEMO

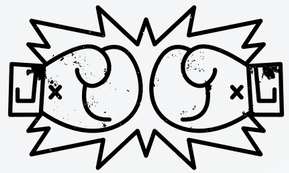
```
In [100]: cmod.predict([[4.1,176000000,43,5000000000,0,0,1]])   # Data of Whatsapp
Out[100]: array([1])

 In [99]: cmod.predict([[4,11000000,17,5000000000,0,0,1]])  # Data of Hotstar
 Out[99]: array([1])

 In [97]: cmod.predict([[3.7,1600,18,10000,0,0,2]]) #Alo Social Random chat
 Out[97]: array([0])

In [101]: cmod.predict([[2.4,2000000,80,1000000,0,0,1]]) # Kik messaging & chat app
Out[101]: array([1])
```

JOURNEY



- Setting the objective at first
- Choosing the right algorithm
- Time Management

CHALLENGES

- Clarity of goals
- Effective Planning
- Continous Improvement

CONQUERED

- More relatable to business world
- 3D graph
- User-Interactive Dashboard

UNIQUENESS

01

WHAT WE LEARNT?

- Combine two algorithms
- Teamwork
- Settling disagreements in case of conflicts of interest

02

COUNTERING DISAGREEMENTS

- Set clear goals and expectations
- Respect different perspective
- Follow up and evaluate

03

NEXT

- Sentiment Analysis on user reviews for better performance and analysis of data.

04

EUREKA MOMENT

- Transition from one topic to another
- Identifying the correct algorithm

THANK YOU

Irene
Anwesh Roy
Joel Bastin
Romel TK
Gladys
Lajiya