

App Name : AnalyzeMyApp

-by Ana Singh

Title of the paper :

An app for developers to predict their ideated app's success rate

Problem statement :

The mobile app development industry is highly competitive, and app developers often struggle with the challenge of determining whether their app idea is likely to be successful or not. Without a reliable method of evaluation, developers risk investing significant time and resources into a product that may not generate the desired return on investment. Factors such as market competition, potential user demand, and effective advertising strategies all play a role in determining an app's success. Therefore, there is a need for an app that provides developers with an accurate and comprehensive evaluation of their app idea, considering various elements like ads, competitions in the market, user demand, and other relevant factors, to help them make informed decisions about whether to pursue their idea or not.

Market/Customer/Business need assessment:

Market Assessment: The mobile app industry is a highly competitive and dynamic market, with new apps constantly entering the market. As per latest Google Play stats, you will be shocked to know that there are 3.48 million apps currently at the Google Play Store. This number of apps on Google Play is on a rise as 3,739 apps are added to the Play Store every single day. Also, As per latest Google Play stats, there are 3.48 million apps currently at the Google Play Store. This number of apps on Google Play is on a rise as **3,739 apps are added to the Play Store every single day.**

Therefore, it is essential for app developers to have a reliable and accurate method of evaluating their app ideas before investing significant resources into their development. The market demand for an app that can evaluate the potential success of an idea is high, as it can help developers make informed decisions about which ideas to pursue.

Customer Assessment: App developers and entrepreneurs are the primary customers for an app that can evaluate the potential success of their app ideas. These customers are typically looking for a tool that can help them determine whether their idea has the potential to generate a return on investment. They also seek guidance on market competition, potential user demand, and effective advertising strategies to make informed decisions.

Business Assessment: An app that evaluates the potential success of an app idea can be a valuable business tool for both small and large app development companies. The app can help developers reduce the risks and costs associated with app development by allowing them to focus their resources on ideas that are more likely to succeed. Such an app can also help businesses improve their return on investment by ensuring that their app ideas have a higher chance of generating revenue.

In conclusion, there is a clear market gap and customer need for an app that evaluates the potential success of an app idea. This tool can be an effective way for developers and businesses to make informed decisions about which ideas to pursue, and reduce the risks and costs associated with app development

Target Specifications and characterization :

The primary customers for an app that evaluates the potential success of an app idea are app developers and entrepreneurs. Here are some characteristics of this customer group:

Tech-Savvy: App developers and entrepreneurs are usually tech-savvy individuals who are familiar with mobile app development, programming languages, and app store optimization.

Creative: App developers and entrepreneurs are often creative individuals who are constantly coming up with new and innovative app ideas.

Business-Oriented: These customers are typically business-oriented and focused on generating revenue from their app ideas. They are looking for tools that can help them determine the viability of their ideas and maximize their return on investment.

Risk-Averse: App development is a risky business, and developers and entrepreneurs are aware of the potential risks associated with app development. Therefore, they are often risk-averse and seek tools that can help them reduce the risks associated with app development.

Time-Constrained: App developers and entrepreneurs are usually working on multiple projects simultaneously and have strict deadlines to meet. Therefore, they are looking for tools that can help them evaluate their app ideas quickly and efficiently.

Market-Aware: These customers are aware of the competitive nature of the app market and understand the importance of conducting market research to determine the viability of their app ideas.

In conclusion, the primary customers for an app that evaluates the potential success of an app idea are tech-savvy, creative, business-oriented, risk-averse, time-constrained, and market-aware individuals. Understanding these customer characteristics is essential for developing an app that meets their needs and provides them with a reliable and effective tool for evaluating their app ideas.

External search :

I checked on the following resources:

<https://www.owox.com/blog/articles/guide-to-mobile-app-analytics/> :

The Ultimate Guide to Mobile App Analytics by OWOX provided an in-depth knowledge of mobile app analytics, explaining what it is, how it works, and why it is essential for app developers and businesses.

<https://www.businessofapps.com/insights/strategies-and-tools-for-successful-testing-in-mobile-app-development/> : Strategies and Tools for Successful Testing in Mobile App Development by Business of Apps

provides an in-depth overview of testing strategies and tools for mobile app development and importance of it.

Bench-marking alternate products :

I looked for comparable applications or tools that offer comparable functionality to find prospective competitors for the app :

Tools for App Store Optimisation : A number of ASO tools are available to assist app developers in improving the visibility and ranking of their apps in app stores as well as the optimisation of their app store listings. Some Examples :

Sensor Tower: A comprehensive ASO tool that offers app store optimization, mobile app intelligence, and market insights.

App Annie: A popular ASO tool that provides market data, app store optimization, and competitive analysis for app developers and marketers.

Mobile Action: A tool that offers ASO, app intelligence, and advertising insights for app developers and marketers.

Tools for Market Research : A range of tools for market research are accessible and offer insights into the app market, including user demographics, rivalry, and market trends. Some of them are :

Google Trends: A free tool that allows users to explore search interest and trends over time for specific keywords or topics.

SEMrush: A marketing tool that offers keyword research, competitive analysis, and market insights to help businesses improve their online presence and marketing strategy.

SimilarWeb: A market intelligence tool that provides insights into web traffic, user engagement, and app performance for businesses and app developers.

Tools for gathering user feedback and analytics : A number of tools for gathering user feedback and analytics are available to assist app developers in comprehending user preferences and behaviour so they can tailor their app properly. Some examples are:

Google Analytics: A free tool that provides website and app analytics, tracking user behavior, traffic sources, and conversion rates.

Mixpanel: A tool that offers real-time analytics, user behavior tracking, and user engagement metrics for web and mobile apps.

Appsee: A tool that offers user behavior analytics, session replays, touch heatmaps, and conversion funnels for mobile apps.

Applicable regulations :

Privacy and data protection laws like (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011 under the Information Technology Act, 2000 , EU's General Data Protection Regulation (GDPR) or the California Consumer Privacy Act (CCPA) when collecting, processing, or storing user data.

Applicable constraints of space , time , budget and expertise :

Space: The amount of storage space required by the app needs to be optimized to minimize the amount of storage space it takes up on users' devices. The storage of data can be done on cloud for low weight storage and processing.

Expertise: Developing an app that can accurately assess the potential success of an app idea requires expertise in various areas, including market research, data analysis, and user experience design , front-end and back-end. Since the storage of data will be on cloud , therefore, developers with necessary expertise about cloud computing will be needed as well.

Business Idea / Monetization model :

Free trial: Customers will be offered a free trial of the app that allows users to assess the value of the service before committing to a subscription.

Subscription options: Customers will be provided with different subscription options, such as monthly or yearly subscriptions. Users can choose the subscription according to their needs and budget.

Pricing tiers: Customers will be offered different pricing tiers with varying levels of access to the app's features. For example, a basic subscription may provide access to the app's core features, while a premium subscription may include additional features, such as personalized recommendations , access to forums or deeper analysis of app data.

Concept generation : Being a computer science student ,I have myself and my friends have faced a common problem that many app developers face - how to determine whether their app idea is likely to be successful in the market. Based on the experience and knowledge of the app development industry, I got the idea to solve the problem of conducting market research, gathering insights into the challenges that developers face when trying to assess the viability of their app ideas.

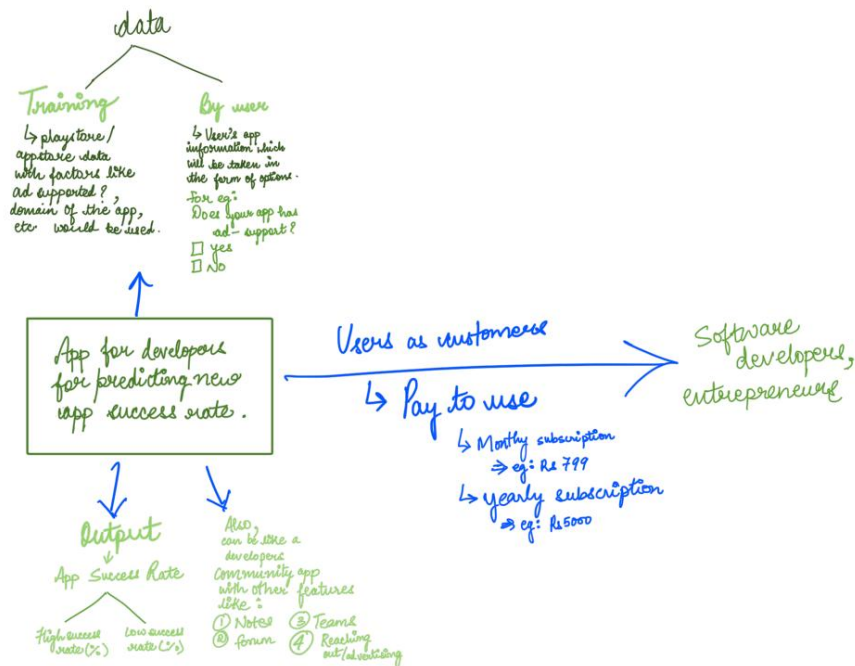
Concept development :

Analyzed the apps from the Google Play Store dataset provided valuable insights into the trends of app success. By examining factors such as app ratings, downloads, and user reviews, identified patterns and

correlations that can inform the development of the app . Then Naive Bayes Algorithm was used to identify the most significant factors that contribute to app success, such as app category, app size, app price, and user engagement metrics. The algorithm then be used to guide the app assessment process and provide recommendations for improving the app idea's potential for success.

Final Product prototype :

The proposed app is a comprehensive tool for developers to assess the potential success of their app ideas. It incorporates various factors such as competition in the market, user feedback, and advertising to provide developers with an accurate and reliable measure of their app's potential success. The app will have a user-friendly interface with intuitive features that enable developers to input information about their app idea and receive a detailed analysis of its viability. The app will provide various charts, graphs, and metrics to make it easy for developers to understand the strengths and weaknesses of their app idea. The app will have a subscription-based business model, with monthly and yearly payment options. Overall, the app will help developers make informed decisions about their app development and marketing strategies, potentially saving them time, money, and effort in the long run.



Product details :

How does it work?

The app for developers to check whether their app idea is going to be successful or not will work by incorporating various elements to analyze an app idea's potential success. The user will input information about their app idea, including its niche, features, target audience, and competitors. The app will then analyze this data using various algorithms and techniques, including machine learning models, to provide an accurate assessment of the app's potential success.

The app will analyze the competition in the market and provide insights into what features and functionalities other apps are offering in the same niche. It will analyze the market size and target audience to determine the potential reach of the app. The app will also analyze user feedback and provide insights into what users like or dislike about similar apps.

Additionally, the app will analyze advertising and marketing strategies for similar apps and suggest appropriate marketing strategies that can be used for the new app. The app will also provide a detailed report and

score that indicates the viability of the app idea and provides recommendations for improving the app's chances of success.

The user will have to subscribe to the app's service, and they will have the option of choosing a monthly or yearly subscription. The app will provide a user-friendly interface to make it easy for developers to input their app idea's details and receive an accurate analysis of its potential success. Overall, the app will provide developers with valuable insights into their app idea's potential, which can save them time, money, and effort in the long run.



Data sources :

The data about the apps was taken from kaggle - google playstore.

Link to the dataset :

<https://www.kaggle.com/datasets/gauthamp10/google-playstore-apps>

Algorithm applied :

A Naive Bayes algorithm was applied to the app idea to predict the success of a new app based on various factors such as competition, user feedback, app features, and marketing strategies.

The Naive Bayes algorithm works by calculating the probability of the success of a new app given a set of features or variables. To apply Naive Bayes to the app idea, data was collected on various factors that impact app success, such as user ratings, downloads, competition, and app features. They would then train the algorithm on this data to calculate the probability of success based on these factors. Once the model is trained, it was used to predict the success of new app ideas and make informed decisions about features, marketing, and target audience. While Naive Bayes is a simple algorithm, it can still provide useful insights into the probability of app success.

Code implementation

4/26/23, 4:11 PM

Untitled0.ipynb - Colaboratory

```
import pandas as pd
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import train_test_split

# Load the dataset
data = pd.read_csv('/content/Google-Playstore.csv')
```

<ipython-input-1-6a8ea388e982>:7: DtypeWarning: Columns (21,22) have mixed types. Specify dtype option on import or set low_memory=
data = pd.read_csv('/content/Google-Playstore.csv')

data

	App Name	App Id	Category	Rating	Rating Count	Installs	
0	Gakondo	com.ishakwe.gakondo	Adventure	0.0	0.0	10+	
1	Ampere Battery Info	com.webserveis.batteryinfo	Tools	4.4	64.0	5,000+	
2	Vibook	com.doantiepvien.crm	Productivity	0.0	0.0	50+	
3	Smart City Trichy Public Service Vehicles 17UC...	cst.stJoseph.ug17ucs548	Communication	5.0	5.0	10+	
		com.horodyski.grower	Tools	0.0	0.0	100+	
		
42986	Black Sparrow Radio	com.radio.m8aa7034c8	Entertainment	4.8	11.0	100+	
42987	Intergalactic Multiples Of Ten	intergalactic.multiplesoften.minigame	Educational	0.0	0.0	10+	

Saving...

data.columns

Index(['App Name', 'App Id', 'Category', 'Rating', 'Rating Count', 'Installs', 'Minimum Installs', 'Maximum Installs', 'Free', 'Price', 'Currency', 'Size', 'Minimum Android', 'Developer Id', 'Developer Website', 'Developer Email', 'Released', 'Last Updated', 'Content Rating', 'Privacy Policy', 'Ad Supported', 'In App Purchases', 'Editors Choice', 'Scraped Time'], dtype='object')

data.info

https://colab.research.google.com/drive/1zr7Gros6LmhQkk1lhCwFPW_vhGDrJoud#scrollTo=NJwR0qraS2Nv&printMode=true

1/8

4/26/23, 4:11 PM

Untitled0.ipynb - Colaboratory

```

Developer Website \
0 ... https://beniyizibyose.tk/#/
1 ... https://webserveis.netlify.app/
2 ... NaN
3 ... http://www.climatesmarttech.com/
4 ... http://www.horodyski.com.pl
...
42986 ... https://www.blacksparrowmusicparlor.com
42987 ... https://spaceworld.us/
42988 ... http://www.autantledire.com/
42989 ... https://digitalsparta.com/
42990 ... NaN

Developer Email Released Last Updated \
0 jean21101999@gmail.com Feb 26, 2020 Feb 26, 2020
1 webserveis@gmail.com May 21, 2020 May 06, 2021
2 vnaerewit@gmail.com Aug 9, 2019 Aug 19, 2019
3 climatesmarttech2@gmail.com Sep 10, 2018 Oct 13, 2018
4 milekhorodyski@gmail.com Feb 21, 2020 Nov 12, 2018
...
42986 blacksparrowmusicparlor@gmail.com Apr 7, 2020 Apr 07, 2020
42987 justin@intergalacticeducation.com Jul 31, 2020 Mar 10, 2021
42988 radiohonam@gmail.com Jul 10, 2018 Oct 27, 2020
42989 support@digitalsparta.com Sep 12, 2019 Sep 18, 2019
42990 mongolcontent@gmail.com Nov 10, 2020 Apr 06, 2021

Content Rating Privacy Policy \
0 Everyone https://beniyizibyose.tk/projects/
1 Everyone https://dev4phones.wordpress.com/licencia-de-uso/
2 Everyone https://www.vietnamairlines.com/vn/en/terms-an...
3 Everyone NaN
4 Everyone http://www.horodyski.com.pl
...
42986 Mature 17+ https://www.blacksparrowmusicparlor.com/privacy
42987 Everyone https://www.intergalacticeducation.com/Privacy...
42988 Everyone http://android.fastcast4u.com/policy_universal...
42989 Teen http://konferenza.com/privacy.pdf
42990 Everyone NaN

Ad Supported In App Purchases Editors Choice Scraped Time
0 False False False 2021-06-15 20:19:35
1 True False False 2021-06-15 20:19:35
...
42986 False False False 2021-06-15 21:02:13
42987 False True False 2021-06-15 21:02:13
42988 True False False 2021-06-15 21:02:13
42989 False False False 2021-06-15 21:02:13
42990 False NaN NaN NaN

[42991 rows x 24 columns]>

```

data_desc.txt

```
data.describe()
```

	Rating	Rating Count	Minimum Installs	Maximum Installs	Price
count	42581.000000	4.258100e+04	4.299000e+04	4.299100e+04	42991.000000
mean	2.210833	2.385493e+03	1.622876e+05	2.644589e+05	0.102779
std	2.109226	9.644367e+04	7.466771e+06	1.115605e+07	2.512067
min	0.000000	0.000000e+00	0.000000e+00	0.000000e+00	0.000000
25%	0.000000	0.000000e+00	5.000000e+01	8.500000e+01	0.000000
50%	3.000000	6.000000e+00	5.000000e+02	6.960000e+02	0.000000
75%	4.300000	4.300000e+01	5.000000e+03	7.426500e+03	0.000000
max	5.000000	1.429116e+07	1.000000e+09	1.616141e+09	399.990000

```
missing_values = data.isna().sum().sort_values() #finding the missing values in the dataset
missing_values
```

```
App Name      0
App Id        0
Category      0
Ad Supported   0
Content Rating 0
Last Updated  0
Maximum Installs 0
Free          0
Price         0
Developer Id   0
In App Purchases 1
Developer Email 1
```

https://colab.research.google.com/drive/1zr7Gros6LmhQkk1lhCwFPW_vhGDkJOud#scrollTo=NJwR0qraS2Nv&printMode=true

2/8

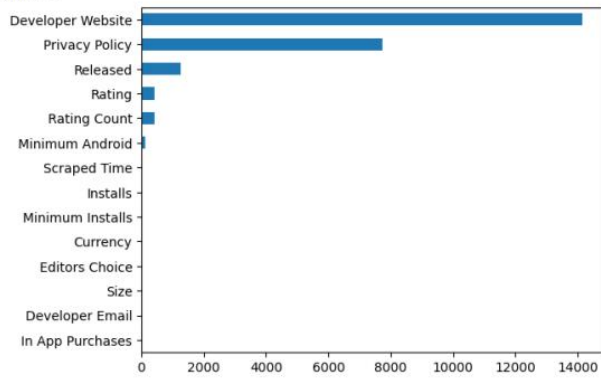
4/26/23, 4:11 PM

Untitled0.ipynb - Colaboratory

```
Size          1
Editors Choice 1
Currency      1
Minimum Installs 1
Installs      1
Scraped Time  1
Minimum Android 130
Rating Count  410
Rating        410
Released      1253
Privacy Policy 7737
Developer Website 14158
dtype: int64
```

```
missing_values[missing_values !=0].plot(kind='barh', )
```

<Axes: >



Saving...



```
Privacy Policy'],axis=1,inplace =True)
```

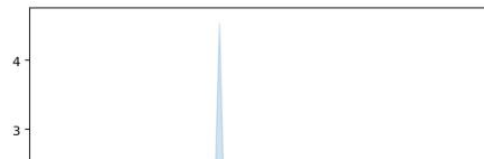
```
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
number_of_price_data = data.Price.unique()
```

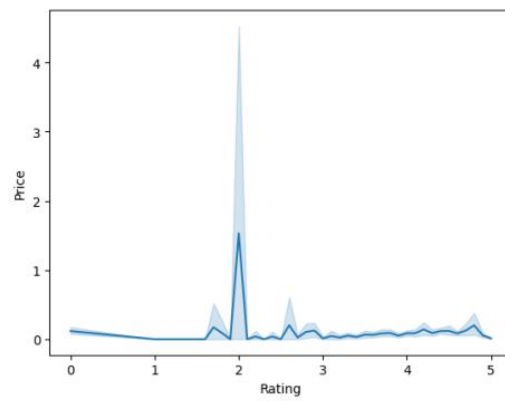
```
len(number_of_price_data)
```

89

```
sns.lineplot(x="Rating", y="Price", data=data)  
plt.show()
```



```
sns.lineplot(x="Rating", y="Price", data=data)
plt.show()
```



https://colab.research.google.com/drive/1zr7Gros6LmhQkk1lhCwFPW_vhGDrJOud#scrollTo=NJwR0qraS2Nv&printMode=true

3/8

4/26/23, 4:11 PM

Untitled0.ipynb - Colaboratory

```
apps_by_Category = data.Category.value_counts()
Top_categories = apps_by_Category.head(10)
Top_categories
```

Education	4465
Music & Audio	2921
Tools	2646
Business	2644
Entertainment	2485
Lifestyle	2199
Books & Reference	2149
Personalization	1703
Health & Fitness	1536
Productivity	1515

Name: Category, dtype: int64

```
data.sort_values(by = 'App Name', ignore_index = True, na_position = 'first')
```

	App Name	App Id	Category	Rating	Rating Count	Install:
0	"Drum Roll" Meme	com.onesoundapps.drumrollmeme	Entertainment	0.0	0.0	100
1	"Look At This Dude" Meme	com.onesoundapps.lookatthisdudememe	Entertainment	0.0	0.0	500
2	"No God please no" Meme	com.onesoundapps.nogodpleasenomeme	Entertainment	4.8	6.0	1,000
3	# storm Area 51	com.hyperiment.sa51	Action	4.2	12.0	1,000
4	Laugh. Date. Experiment	com.davenporttheatrical.dateme	Dating	3.4	7.0	1,000
...
42986	العباب تأوين الفراشات للطفال: ظرين جنوت fairy	com.FairyPages.books2.coloring	Educational	0.0	0.0	500
42987	Superhero Skinpack Justice League mod MCPE	com.BlackStarQQ.Hero	Entertainment	0.0	0.0	100
42988	Mermaid Princess Makeup - Girl Fashion Salon	com.kiwigo.mermaidprincessmakeup.free	Casual	4.0	8080.0	1,000,000
42989	Smart - Brain fitness	com.doggoapps.smart	Education	4.6	8.0	500
42990	Mod Freddy for Minecraft PE	com.EmpireMods.FnaflMod	Entertainment	4.0	691.0	50,000

42991 rows × 7 columns

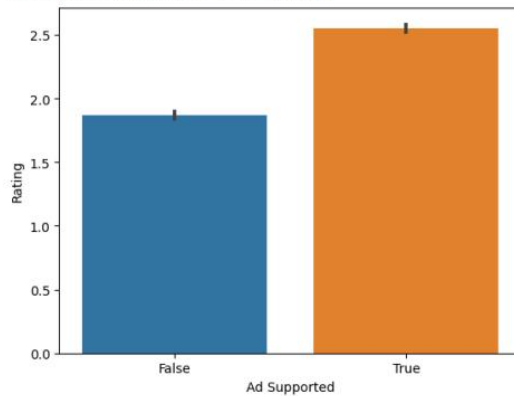


4/26/23, 4:11 PM

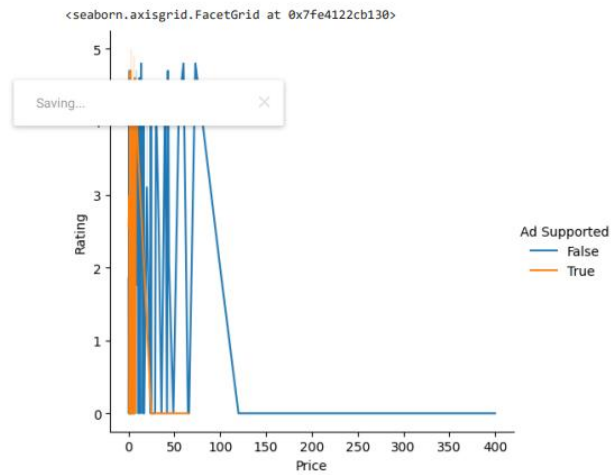
Untitled.ipynb - Colaboratory

```
sns.barplot(data=data, x= 'Ad Supported', y='Rating')
```

<Axes: xlabel='Ad Supported', ylabel='Rating'>



*Expected : * The apps with app would have bad ratings because of interruptions .
Found : The apps with ads have better ratings.



We find that:

If the price is negligible, the rating remains high, and the factor of ads does not hamper it as much, though without ads the rating is slightly higher. As the price increases: 1.1. If ads are supported by the app, then Ratings decrease and can go as low as 0. 1.2. If ads are not supported by the app, then Ratings can vary from high to low.

```
# Select the relevant columns for classification
data = data[['Category', 'Rating', 'Free', 'Installs', 'Ad Supported']]
```

```
data
```



```

      Category Rating Free Installs Ad Supported
0      Adventure 0.0  True    10+      False
1         Tools  4.4  True   5,000+     True
2    Productivity 0.0  True    50+     False
3  Communication 5.0  True    10+     True
4         Tools  0.0  True   100+     False
...
14325 Entertainment 3.2  True  10,000+    True
14326 Education  4.7  True  10,000+    False

type('Installs')

str

# Drop any rows with missing values
data = data.dropna()

# Convert the 'Installs' column to numeric
data['Installs'] = data['Installs'].astype(str)
data['Installs'] = data['Installs'].str.rstrip('+')
data = pd.get_dummies(data, columns = ['Category', 'Free', 'Ad Supported'])
print(one_hot_encoded_data)

```

data

	Rating	Installs	Category_Action	Category_Adventure	Category_Arcade	Category_Art & Design
0			0	1	0	0
1	4.4	5,000	0	0	0	0
2	0.0	50	0	0	0	0
3	5.0	10	0	0	0	0
4	0.0	100	0	0	0	0
...
14325	3.2	10,000	1	0	0	0
14326	4.7	10,000	0	0	0	0

```

data['Installs'] = data['Installs'].str.replace(",","")
data['Installs'] = data['Installs'].astype(int)

```

```
data2 = data.drop(['Rating'],axis=1)
```

data2

```

      Installs Category_Action Category_Adventure Category_Arcade Category_Art & Design Category_
0          10              0              1              0              0
1         5000              0              0              0              0
2          50              0              0              0              0
3          10              0              0              0              0
4          100              0              0              0              0
...
14325        10              1              0              0              0
14326        10              0              0              0              0

```

data

	Rating	Installs	Category_Action	Category_Adventure	Category_Arcade	Category_Casual
0	0.0	10	0	1	0	0
1	4.4	5000	0	0	0	0
2	0.0	50	0	0	0	0
3	5.0	10	0	0	0	0
4	0.0	100	0	0	0	0
...
14324	0.0	10	0	0	0	0
14325	3.2	10000	0	0	0	0
14326	4.7	10000	0	0	0	0
14327	0.0	10	0	0	0	0
14328	3.7	50000	0	0	0	0

Saving...

```

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(data2, data['Rating'].astype('int'), test_size=0.2, random_state=42)

# Train the Naive Bayes classifier
clf = MultinomialNB()
clf.fit(X_train, y_train)

# Predict the ratings of the test set
y_pred = clf.predict(X_test)

# Print the accuracy of the classifier
accuracy = clf.score(X_test, y_test)
print('Accuracy:', accuracy)

Accuracy: 0.5267605633802817

```

Link to the colab file :

https://colab.research.google.com/drive/1zr7Gros6LmhQkk1lhCwFPW_vhGDrJOud?usp=sharing

Accuracy can be improved with more adding more relevant columns in the dataset.

Conclusion :

The app success evaluator aims to develop an app that evaluates the potential success of an app in the market using machine learning techniques. The project used the Google Play Store dataset to analyze app trends and developed a model that predicts the success of an app based on various parameters like category, rating, reviews, ads, and price.

The project used Naive Bayes Machine learning algorithm, to develop a predictive model. Naive Bayes showed an accuracy of 0.5267605633802817 , which can be improved by adding more factors to the dataset, and by training more intensively.

The app works by taking user input for the various parameters and giving a score between 0 to 1.

The project had some limitations, like the dataset being limited to the Google Play Store and not including data from other app stores like Apple's App Store.

Overall, the app success evaluator project provides a useful tool for app developers to evaluate their app's potential for success in the market and make informed decisions for improvement.

-----Thank you-----