Internship Report:

Analyzing User Sentiment Towards Apps on the Google Play Store

Introduction-

The Google Play Store is a vast digital marketplace where users can discover and download millions of apps. User reviews play a crucial role in influencing app popularity and user satisfaction. This internship project aims to analyze user sentiment towards gaming apps on the Google Play Store and identify trends in user reviews.

Background-

The Google Play Store launched in 2008, providing a platform for developers to distribute their apps to Android users worldwide. User reviews serve as valuable feedback for developers, allowing them to improve their apps and address user concerns.

Learning Objectives-

- Gain proficiency in Python programming and data analysis libraries.

- Master the art of visualizing and interpreting data effectively.

- Comprehend the significance of data cleaning and preprocessing.

Activities and Tasks

Data Cleaning and Preprocessing:

1. Handling Missing Values: addressed missing values by replacing them with 0. However, consider exploring other imputation techniques like using the mean or median for numerical features, or mode for categorical features. This might provide more accurate insights.
2. Data Transformation: converted data types and removed special characters. Consider further transformations like scaling numerical features (e.g., using StandardScaler) or applying logarithmic transformations for skewed distributions. This can improve model performance.

Feature Engineering:

3. App Age: calculated app age, which is a valuable feature. Consider exploring other time-based features like days since last update, or creating categorical features based on app age ranges (e.g., new, mature, old).
4. Text Features: performed basic text analysis using word clouds. Consider exploring more advanced techniques like sentiment analysis on app descriptions or user reviews. This can provide insights into user perceptions and preferences.
5. Category Interactions: Explore interactions between categories and other features. For example, create features like "Category_x_Type" to capture the combined effect of category and app type.

Analysis and Visualization:

6. Interactive Visualizations: Consider using Plotly or other interactive visualization libraries to create more engaging and informative charts. This allows users to explore the data in more detail.
7. Multivariate Analysis: performed PCA. Consider exploring other dimensionality reduction techniques like t-SNE or UMAP. This can help visualize relationships between multiple features.
8. Statistical Modeling: Consider building statistical models to predict app ratings, installs, or revenue. This can provide insights into factors that drive app success.

Skills and Competencies

- **Python programming:** Proficient in using Python for data manipulation, analysis, and visualization.
- **Data analysis:** Skilled in data cleaning, preprocessing, and exploration.
- **Data visualization:** Ability to create informative and visually appealing charts and graphs.

Challenges and Solutions

- **Data sparsity:** Encountered challenges in analyzing less popular apps due to limited review data.
- **Solution:** Focused on analyzing more popular apps with a larger volume of reviews.

Outcomes and Impact

- There is a positive correlation between app ratings, reviews, and installs.
- The majority of apps in the dataset are free.
- Category Performance: Analyze the performance of different app categories based on metrics like installs, ratings, and revenue. Identify top-performing categories and potential areas for improvement.
- User Sentiment: Analyze user reviews to understand sentiment towards different app categories or features. This can help identify areas for improvement in app design or functionality.
- Pricing Strategies: Analyze the relationship between price and installs for paid apps. Identify optimal pricing strategies for different app categories.
- Marketing and Promotion: Based on your analysis, provide recommendations for marketing and promotion strategies to increase app visibility and downloads.

Conclusion

- This project performed an in-depth analysis of app and user review data from the Google Play Store.
- The analysis involved data cleaning, exploration, correlation analysis to identify key insights into app features, user sentiment, and market trends.
- The project also identified popular app categories and the relationship between app ratings, reviews, and installs.
- Interactive visualizations were created, including a scatter plot showing revenue versus installs for paid apps and an interactive choropleth map displaying global installs by country.
- The analysis also examined user reviews and applied sentiment analysis techniques to gain insights into user satisfaction and feedback.