Anime User Behavior Analysis: Final Report

1. Introduction and Problem Identification:

We focus on understanding user behaviors for anime consumption. By analyzing user preferences and ratings, we can develop a recommender system to aid streaming platforms in curating a better viewing experience for their users.

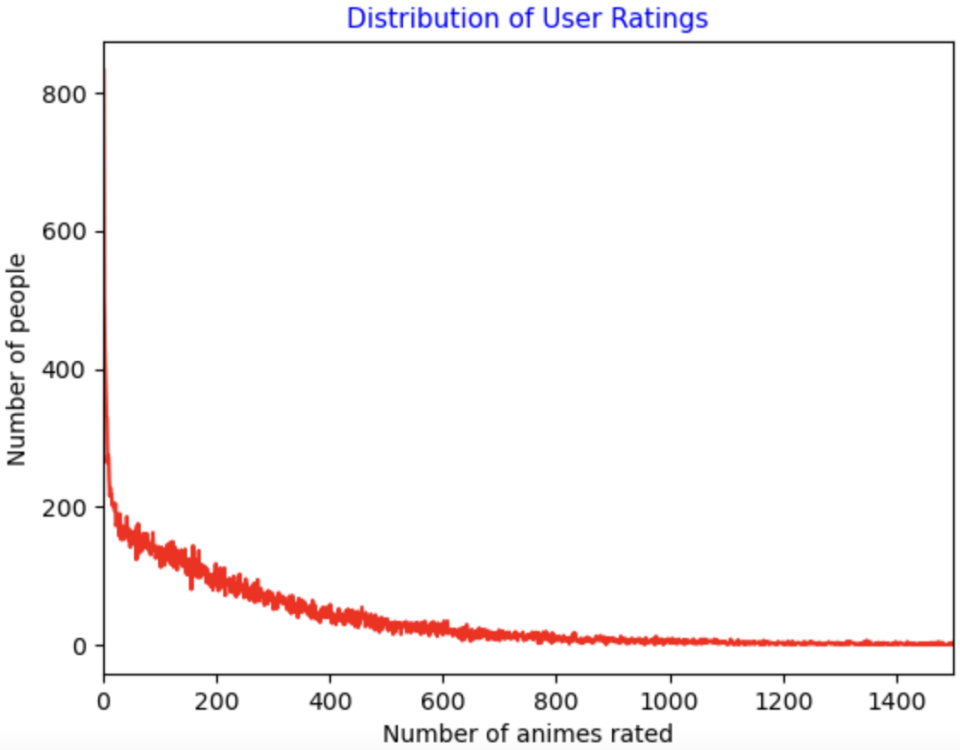
2. Data Collection and Wrangling:

Datasets used:
- `https://www.kaggle.com/datasets/azathoth42/myanimelist?select=UserList.csv
    - ☐ Detailed user data.
- https://www.kaggle.com/datasets/azathoth42/myanimelist?select=AnimeList.csv
    - ☐ Comprehensive list of anime with metadata.
- https://www.kaggle.com/datasets/azathoth42/myanimelist?select=UserAnimeList.csv
    - ☐ Reflects user interactions and ratings with different anime.

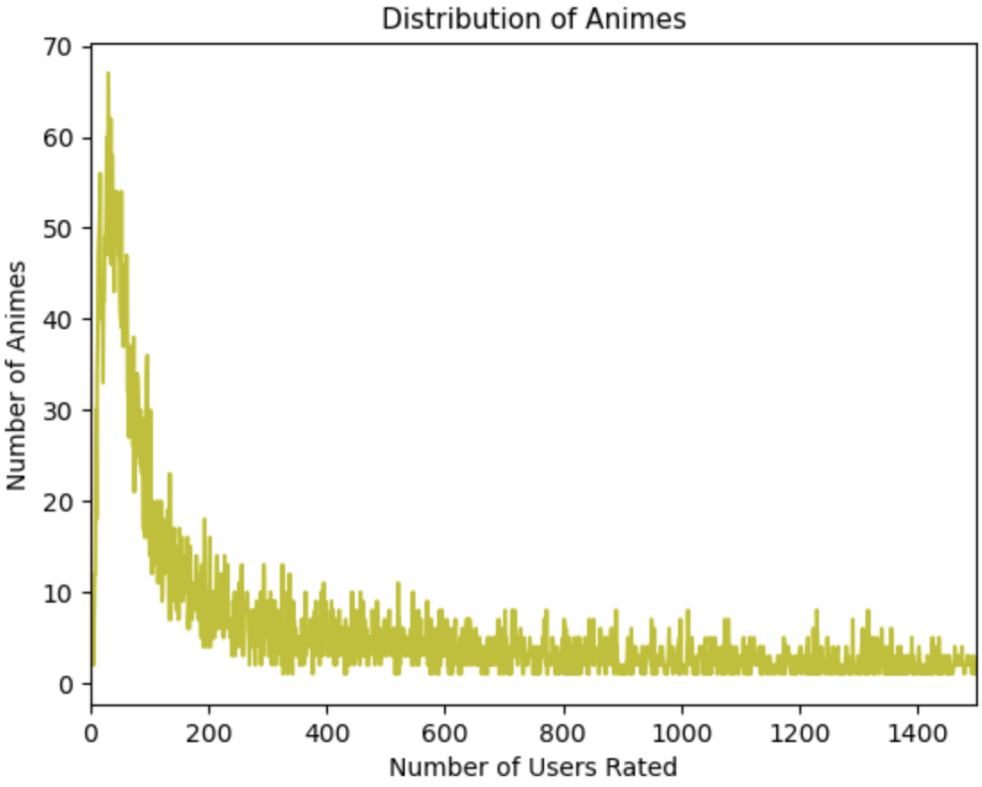
After loading the data, essential columns were filtered, and dataframes were merged to create a holistic dataset for analysis.

3. Exploratory Data Analysis (EDA):

- Distribution of anime ratings by users.



- Number of anime rated by users.



4. Data Pre-processing and Addressing Challenges:

- Filtering users based on the number of anime they've rated to address the cold start problem.
- After merging the datasets, the data was reduced to a manageable size due to initial large dimensions.

5. Model Building, Training, and Evaluation:

Metrics:
Precision@k and Recall@k were chosen due to their significance in evaluating recommender systems.

Models:
Multiple models were tested, including SVD, SlopeOne, NMF, KNN variants, and BaselineOnly. After evaluation:

- `**BaselineOnly**` emerged as the best model in terms of RMSE and Precision@10.
- GridSearchCV was used for hyperparameter tuning. However, default parameters yielded the best RMSE.
- A final recommendation system was built using `**BaselineOnly**`, predicting potential ratings for user-anime combinations. A function was then developed to recommend top anime for any given user based on these predictions.

6. Conclusion and Recommendations:

- The `**BaselineOnly**` recommender system can be utilized by streaming platforms to provide top anime recommendations for their users.
- To enhance user engagement, platforms can use this system to personalize user interfaces, focusing on showcasing top-rated anime.

7. Future Work:

- Addressing the cold start problem for new users or anime.
- Introducing content-based filtering in conjunction with collaborative filtering to enhance recommendation accuracy.
- Integration of real-time user feedback to update the recommendation list dynamically.