

Movie Recommendation System

using
Collaborative Filtering

By: Abdullah Khaled



Content

▶▶▶	Introduction	01
▶▶▶	Methodology	02
▶▶▶	Exploratory Data Analysis (EDA)	03
▶▶▶	Model Training and Evaluation	04
▶▶▶	Recommendations	05
▶▶▶	Conclusion	06
▶▶▶	References	07
▶▶▶	Questions and Discussion	08

Introduction

Project Overview

- **Objective:** To develop a recommendation system using the MovieLens 100k dataset.
- **Dataset:** Contains 100,000 ratings (1-5) from 943 users on 1682 movies.

Importance

- **User Experience:** Enhance user experience by recommending movies tailored to individual preferences.
- **Business Value:** Increase user engagement and retention by providing personalized content.

Methodology

Steps Involved

1. **Data Loading:** Load the MovieLens 100k dataset.
2. **Exploratory Data Analysis (EDA):** Understand the dataset through statistical analysis and visualizations.
3. **Hyperparameter Tuning:** Perform grid search to find the best hyperparameters for the SVD model.
4. **Model Training:** Train the SVD model using the best hyperparameters.
5. **Predictions:** Predict ratings for unseen movies for a given user.
6. **Recommendations:** Recommend the top-rated movies.
7. **Visualization:** Visualize the top recommended movies with their predicted ratings.

Exploratory Data Analysis (EDA)

Dataset Information

943

Users

1682

Movies

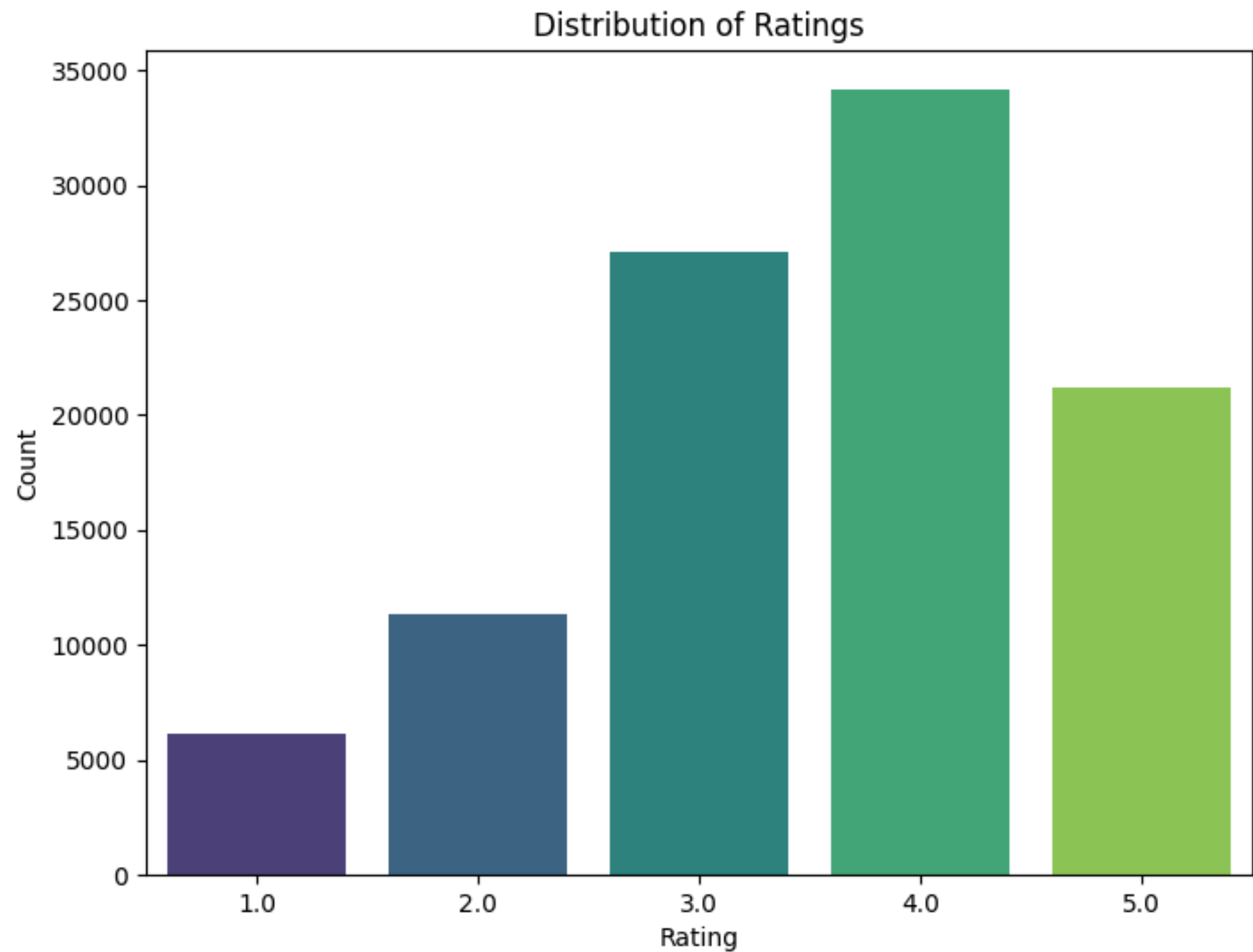
100,000

Ratings

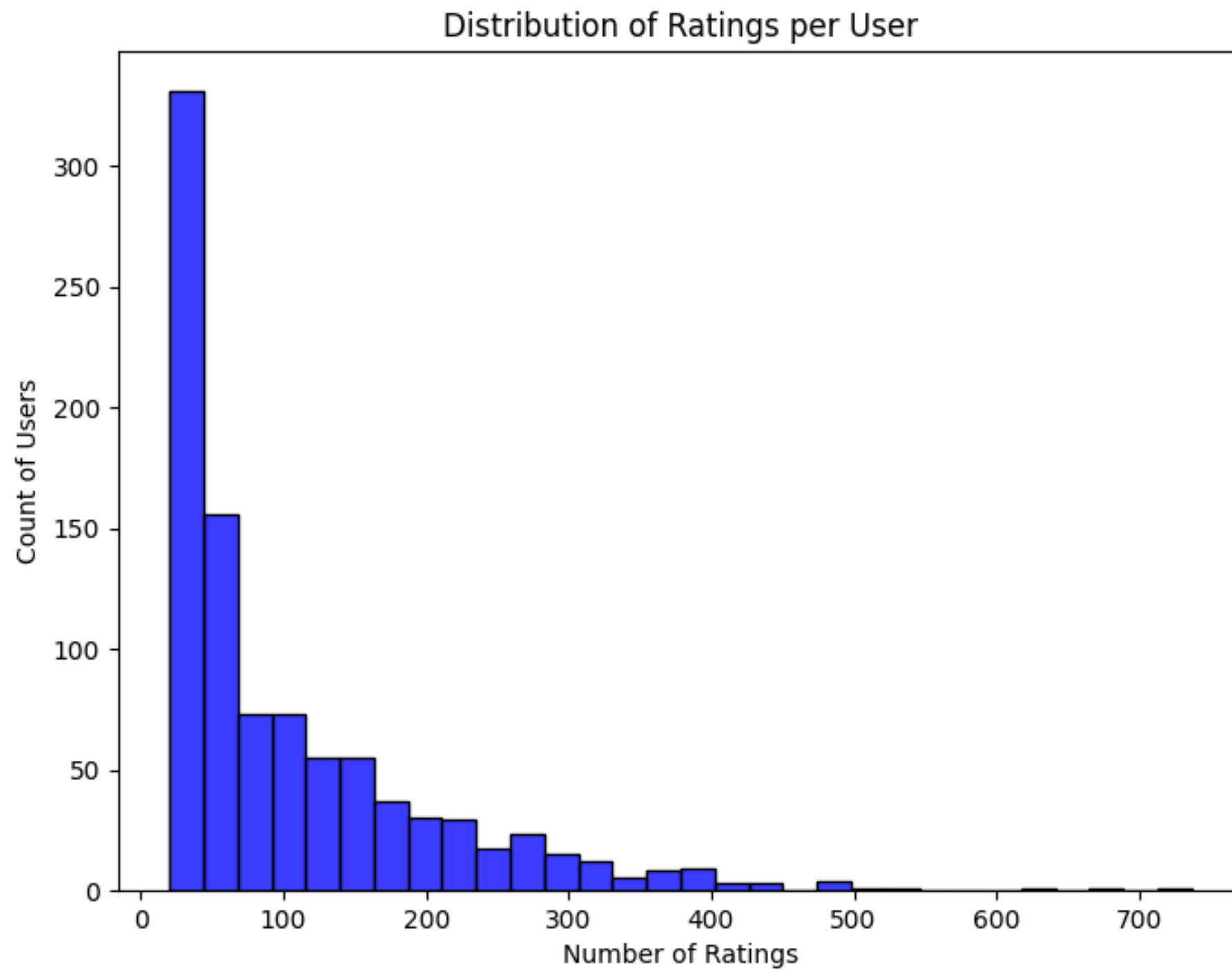
1-5

Rating Scale

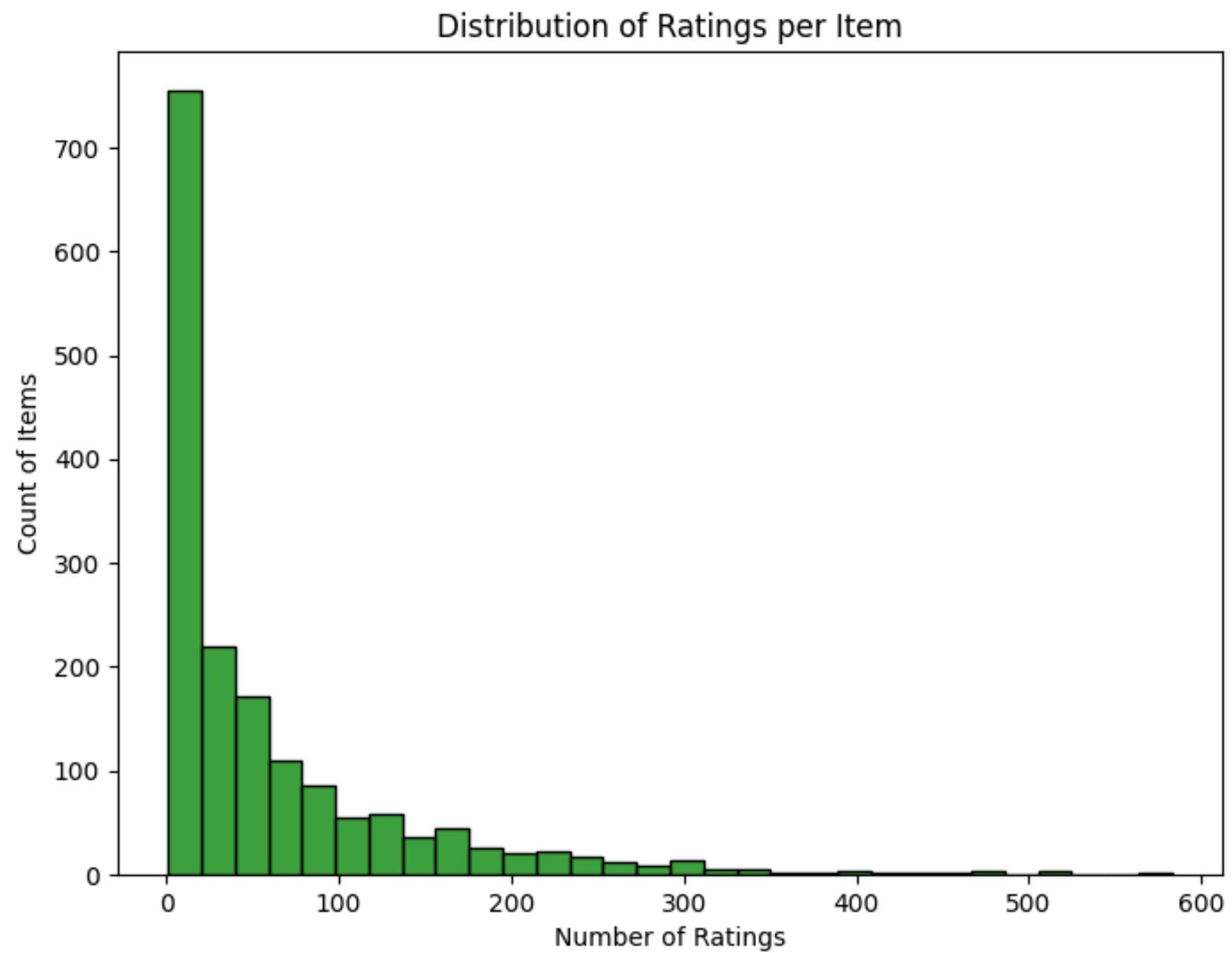
Exploratory Data Analysis (EDA)



Exploratory Data Analysis (EDA)



Exploratory Data Analysis (EDA)



Model Training and Evaluation

Hyperparameter Tuning

Algorithm: Singular Value Decomposition (SVD)

Parameter Grid:

- n_epochs: **[5, 10]**
- lr_all: **[0.002, 0.005]**
- reg_all: **[0.4, 0.6]**

Best Parameters:

- n_epochs: **10**
- lr_all: **0.005**
- reg_all: **0.4**

Model Training and Evaluation

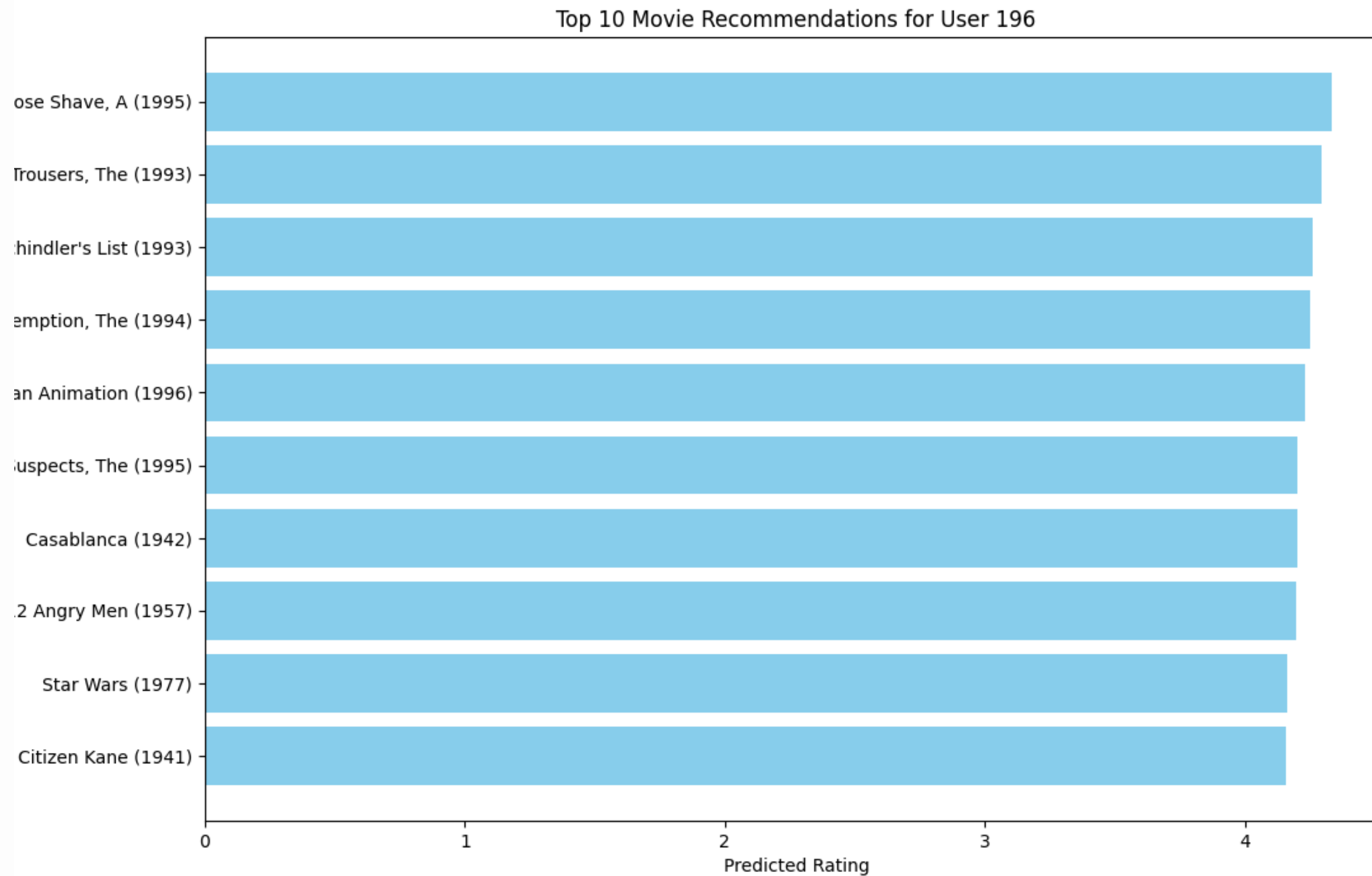
- **Train the Model:** Using the best parameters on the entire dataset.
- **Evaluation Metrics:** RMSE (Root Mean Squared Error), MAE (Mean Absolute Error)
- **Best RMSE Score: 0.9412**

Recommendations

Predicting Ratings:

- **User ID: 196**
- **Predict Ratings:** For all movies that the user has not rated.

Recommendations



Conclusion

Summary

- **Effective Recommendation System:** Built using the MovieLens 100k dataset.
- **Personalized Recommendations:** Enhances user experience by providing tailored movie recommendations.
- **Visualization:** Helps in understanding the recommendations better.

Future Work

- **Incorporate More Features:** Use additional features like movie genres, user demographics for better recommendations.
- **Scalability:** Apply the model to larger datasets like MovieLens 1M or 20M.
- **Real-time Recommendations:** Implement real-time recommendation systems for dynamic user interactions.



References

Notebook: [click here](#)

Resources:

- Real Python: [click here](#)
- Google developers: [click here](#)

The image features a solid blue background. Scattered across this background are several light blue, semi-transparent triangles of various sizes and orientations. Some triangles point towards the top-left, others towards the top-right, and some towards the bottom-left. In the center of the image is a white rectangular box with a thin white border. Inside this box, the word "QUESTIONS?" is written in a large, bold, white, sans-serif font.

QUESTIONS?

THANK YOU!



Abdullah Khaled

Data Analyst | Machine Learning Engineer

 +20 1557504902

 dev.abdullah.khaled@gmail.com

 Beni Suef, Egypt