Movie Recommender System

Collaborative filtering

Description

- The movie recommender system aims to provide personalized movie recommendations to users based on ratings. By using data on movie genres and ratings, the system will predict and suggest movies that users are likely to enjoy.
- The system will use collaborative filtering (user - based) to make suggestions for each user.

Business Problem

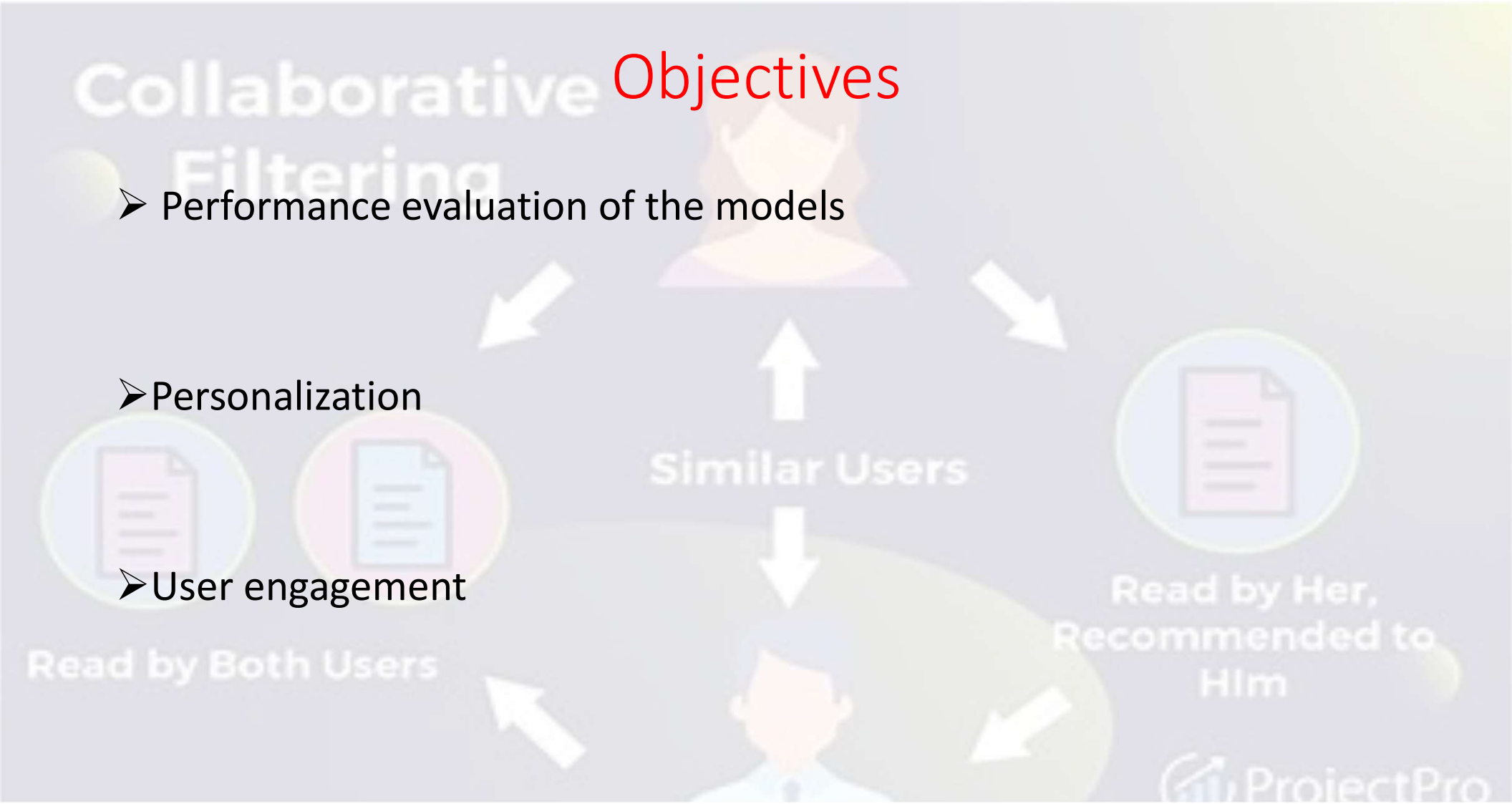
Users finds it difficult to choose content that matches their preferences. Lack of personalized recommendations leads to users dissatisfaction.

The recommender system will:

- Help users to discover new content that they may enjoy based on their unique preferences.
- Enhance user experience by providing personalized movie recommendations.
- Ensure viewers are exposed to a variety of relevant movies

Objectives

- Performance evaluation of the models
- Personalization
- User engagement



Shareholders

- End users - these are the people who consume the content and are relying on recommendations.
- Content creators - these are people who are producing the content
- marketing team - these team will rely on the system to come up with a marketing strategy
- product team`- these are for integrating the system into the platforms

Data Understanding

- Data used was the grouplens.org/datasets/movielens/latest/ sourced from the Group Lens research lab at the University of Minnesota.

- The small dataset containing 100,000 user ratings and a subset of the data i.e. movies and ratings was used.
- The data contains columns:

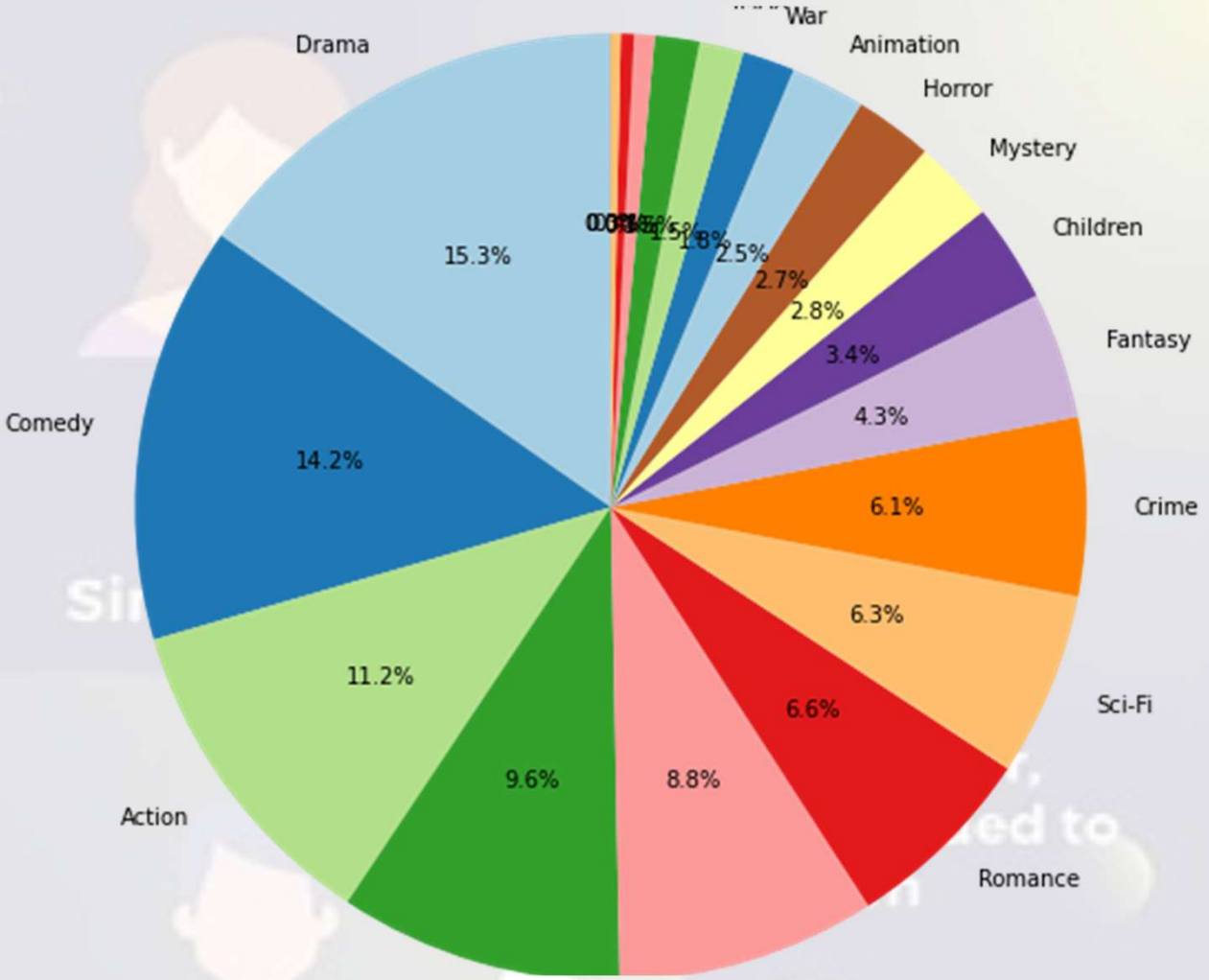
- movie ID - uniquely identifies a movie
- Title - name of a movie

Data Understanding

- Genres - genre to which a movie belongs
- user ID - uniquely identifying each user
- rating - a rating given to movies by users
- timestamp - time the movies were released

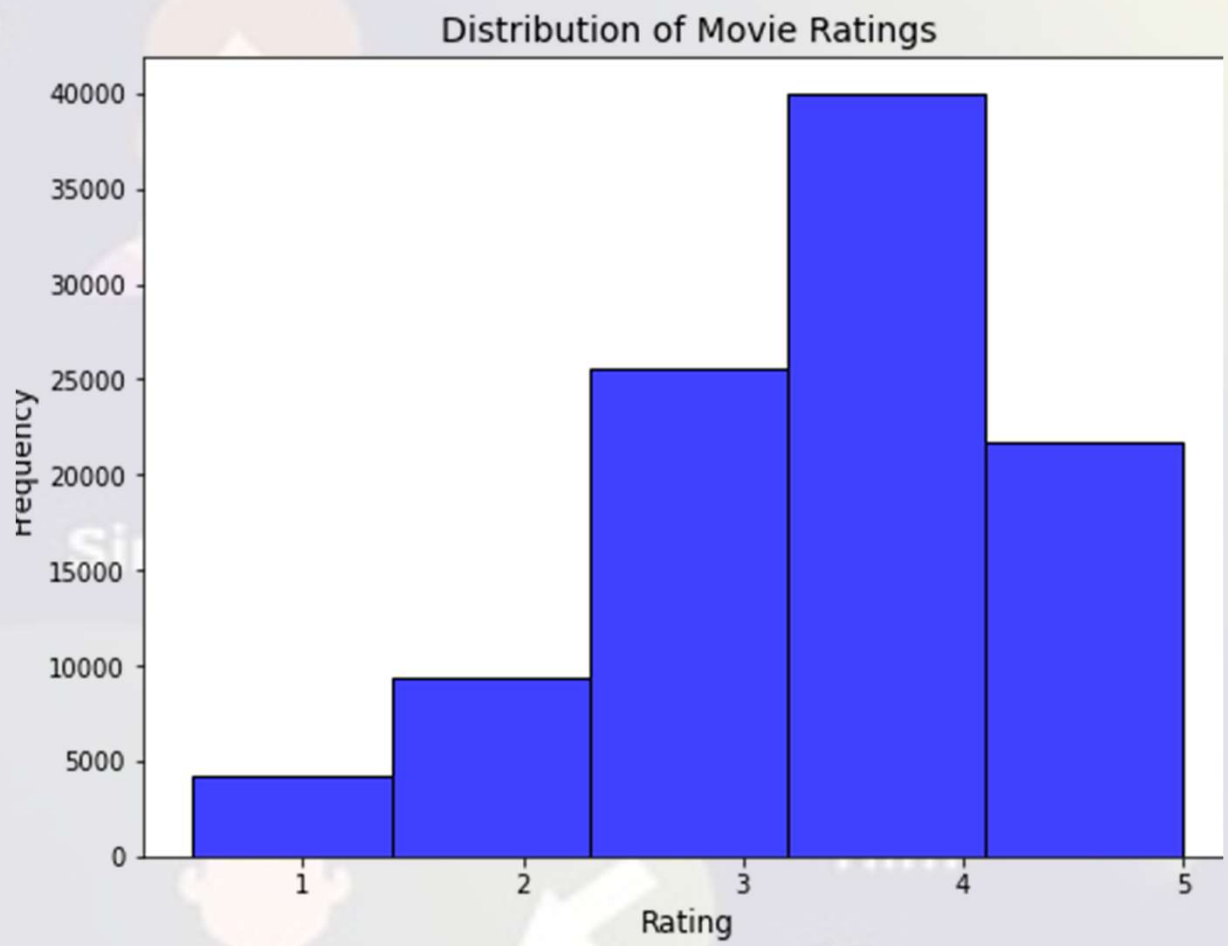
Genre popularity visualizations

- Genres Drama, Comedy and Action are the most popular ones

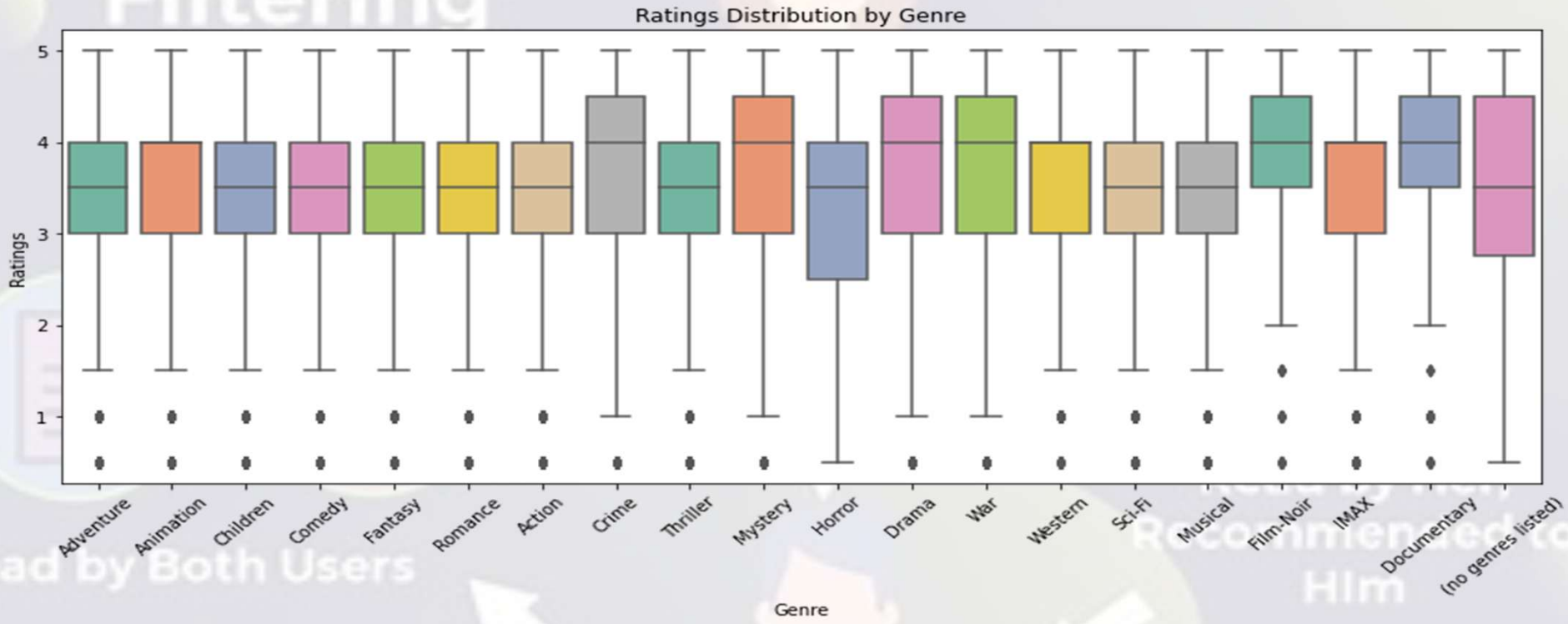


Ratings Distribution

Rating 4 has the highest frequency followed by 3,5,2 and 1



Rating Distribution by Genre

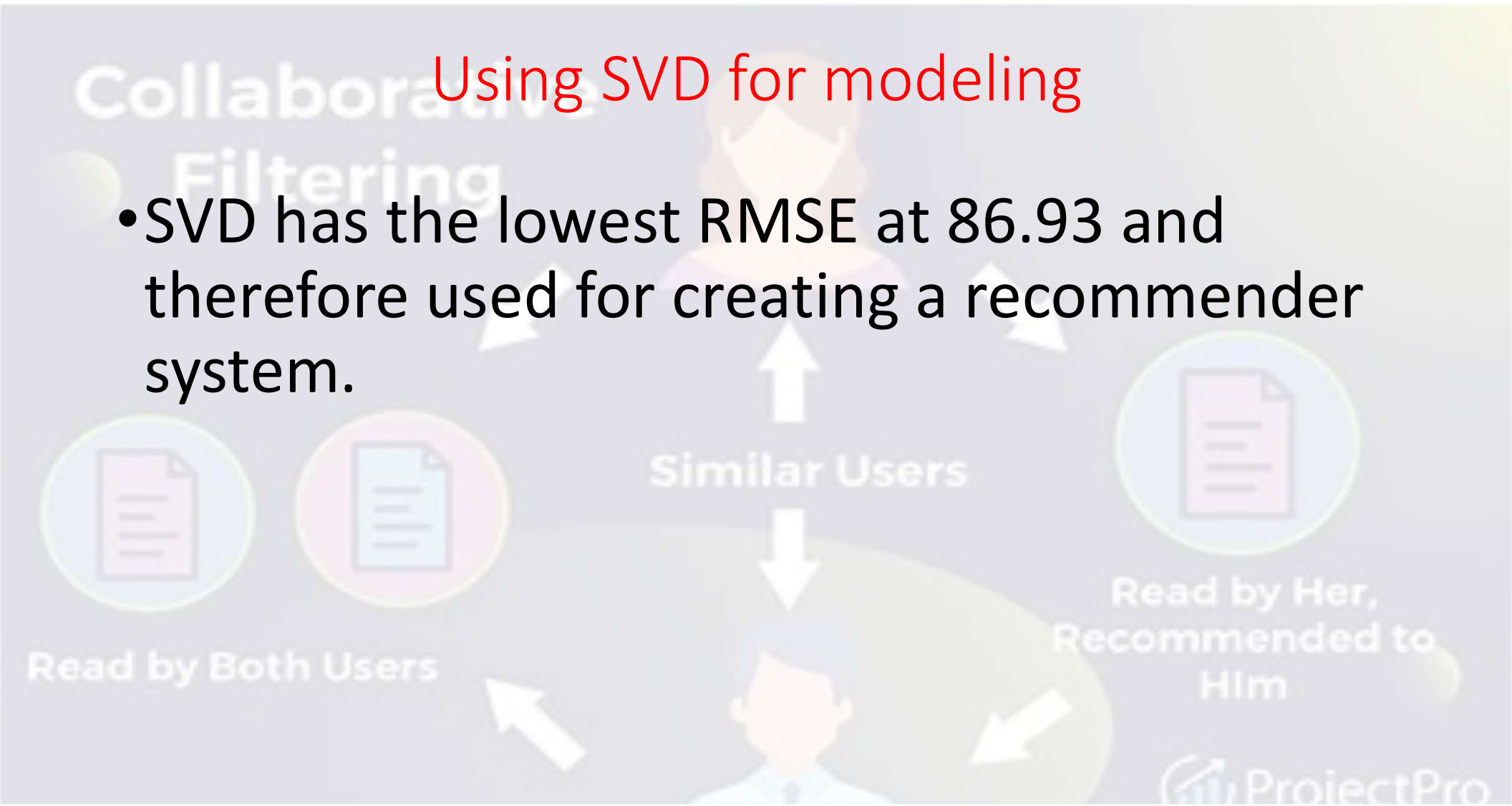


Using different Models

1. Modeling with SVD using gridsearch – Test RMSE = 86.93%
2. Cross validation with KNNBasic – Test RMSE = 97.36%
3. Cross validation with KNNBaseline – Test RMSE = 87.67%
4. Cross validation with KNNWithMeans – Test RMSE = 89.74%

Using SVD for modeling

- SVD has the lowest RMSE at 86.93 and therefore used for creating a recommender system.



Making Recommendations

- Recommendation 1: “Dr. Strangelove".
- Recommendation 2: "The Shawshank Redemption (1994)".
- Recommendation 3: “Lawrence of Arabia (1962)".
- Recommendation 4: “Fight club (1999)".
- Recommendation 5: “Godfather: part II, The (1974)”.

Collaborative
Filtering
Similar Users
Read by Both Users
Read by Her,
Recommended to
Him
ProjectPro

THANK YOU!