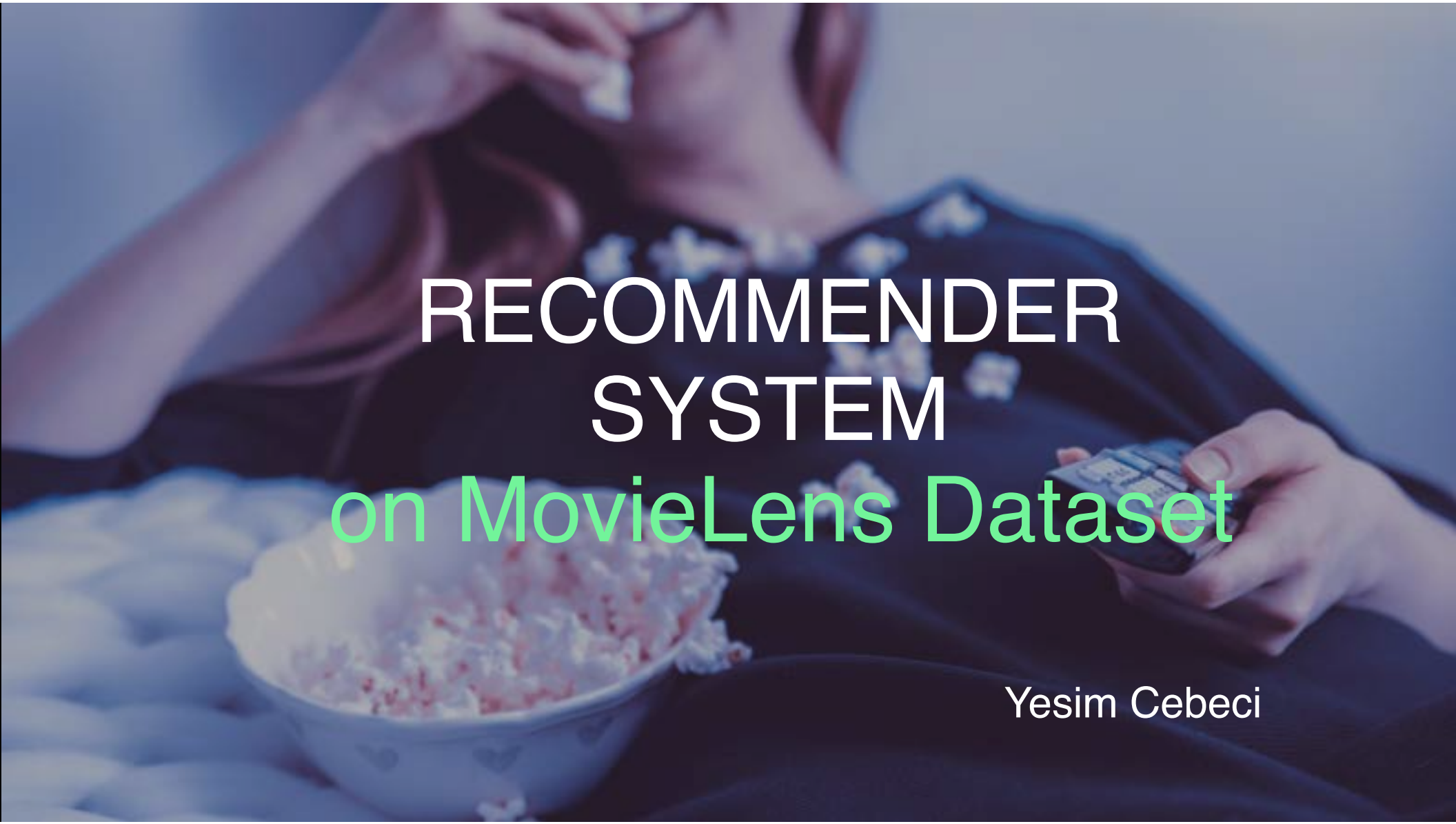
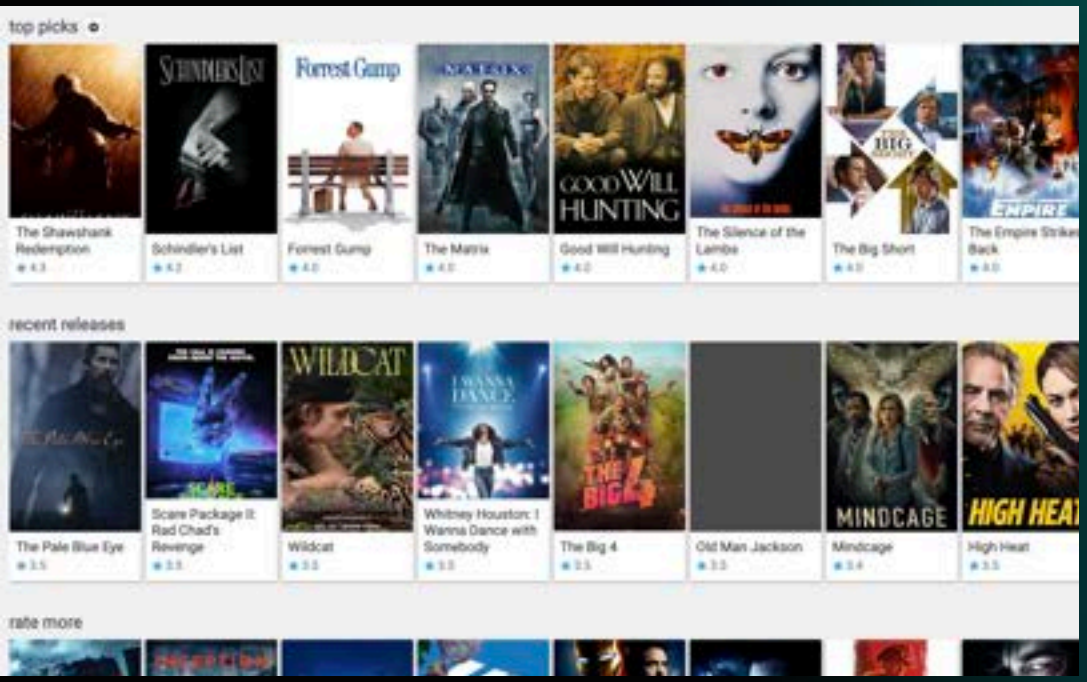


RECOMMENDER SYSTEM
on MovieLens Dataset

Yesim Cebeci

TABLE OF CONTENTS



ABOUT THE MOVIELENS

MovieLens is a website that helps users find movies that they will like.

It uses ratings given by the user to build a custom taste profile of that particular user and then utilizes that information to recommend other movies for the user to watch.

01

Business Understanding

MovieLens wants to improve it’s movie recommendation system that is located on users’ homepages.**The goal is to use users movie ratings and recommend other movies.**

This may save the user time when deciding which movie they would like to watch

02

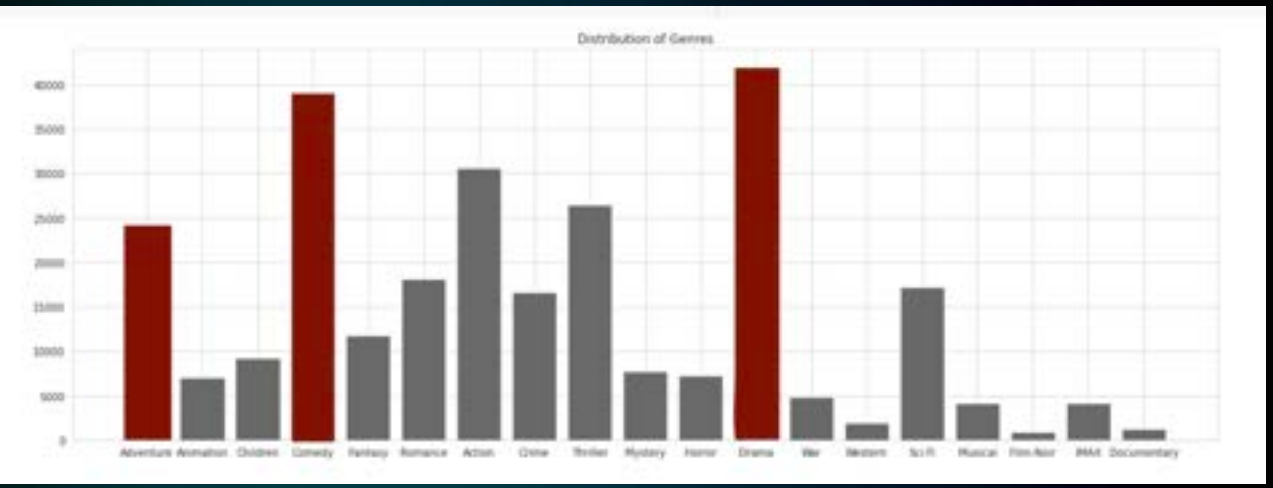
Data Understanding

This dataset provide by **MovieLens website**

- It contains **100836** ratings,
- **9742** movies
- And data were created by **610** users

03

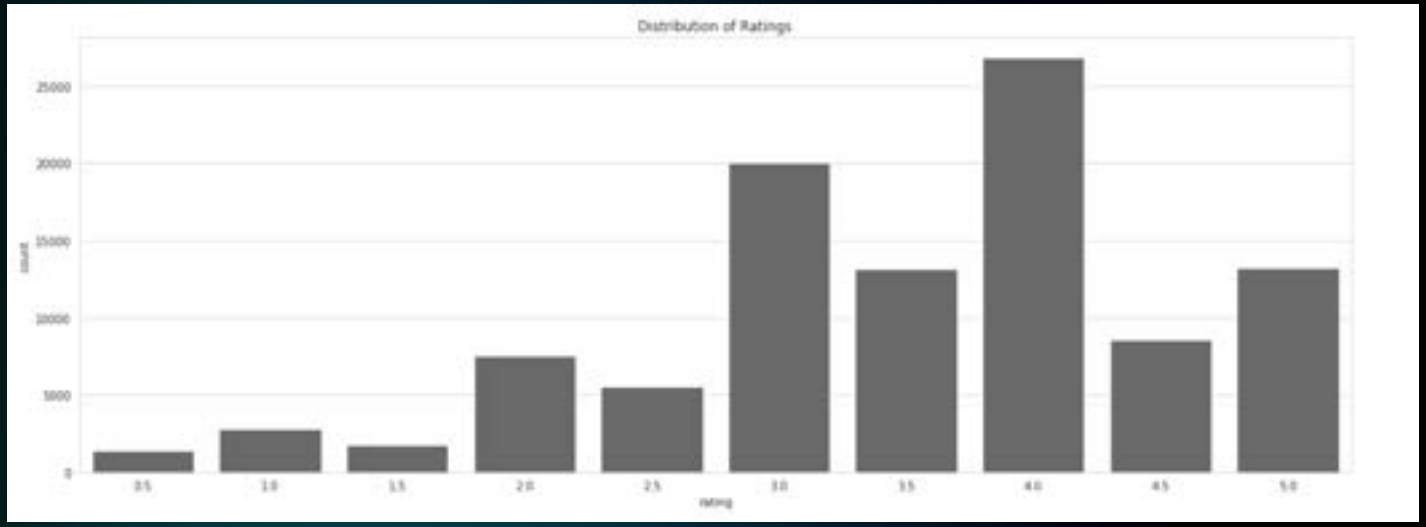
Data Analysis



- Distribution of genres is not normal distributed
- Most of the movies belong to movie genre **Drama** , **Comedy Action, then** Thriller and Adventure

03

Data Analysis



- The **highest** rating is **4** and then followed by rating **3** and then rating **5**.
- The summary stats also show the ratings mean are high

04

Recommendations based on Popularity to a **New User**

Movie	Rating
Shawshank Redemption	4.42
Fight Club	4.27
The Usual Suspects	4.23
Star Wars: A New Hope	4.23
Schindler's List	4.22

- These are the most popular movies which can be recommended to **a new user**.
- Movies are sorted by both the number of ratings and their average rating in our dataset.

04

Recommendations based on Genres to a **New User**

- **Get recommendations based on genre : ‘Pulp Fiction’**

Movie	Genres
Fargo	Comedy,Crime,Drama,Thriller
Freeway	Comedy,Crime,Drama,Thriller
Man Bites Dog	Comedy,Crime,Drama,Thriller
Beautiful Creatures	Comedy,Crime,Drama,Thriller
Confessions of a Dangerous Mind	Comedy,Crime,Drama,Thriller

- These are the movies which can be recommended based on given movie titles genres to a **new user**.
- Movies are sorted by the similarity scores

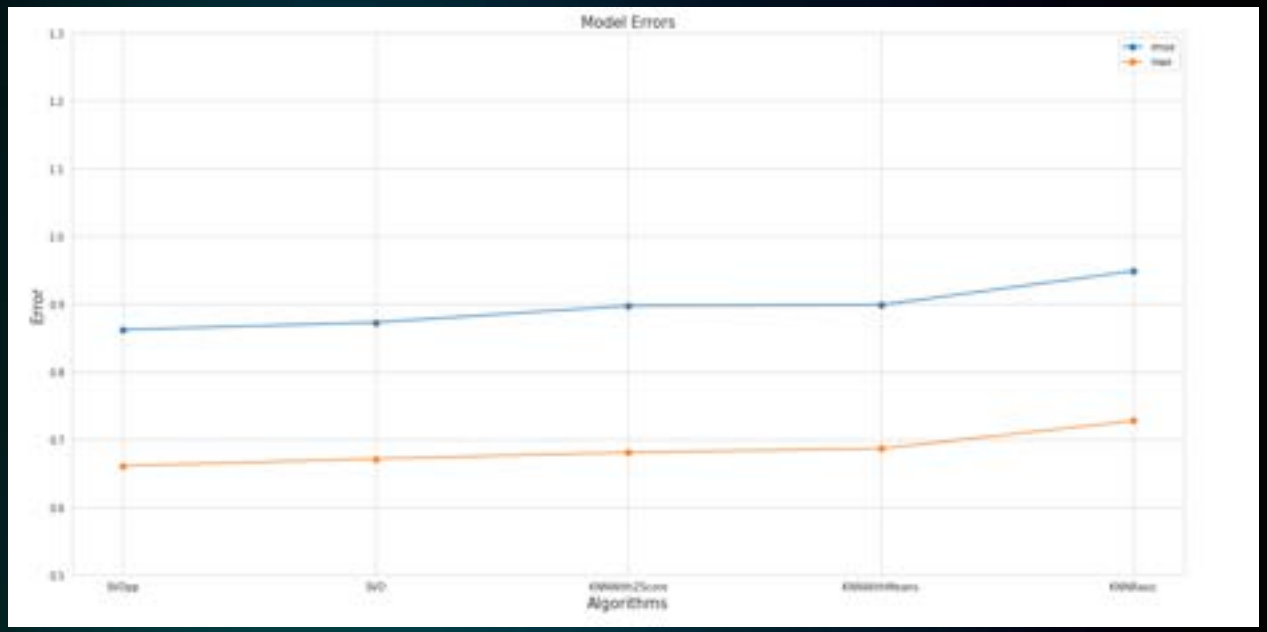
Item-Based Collaborative Filtering to **Existing User**

- **Get recommendations basen on movie similarity: 'Edward Scissorhands'**

Movie	Genres	Similarity
Hamlet	Crime,Drama, Romance	0.88
Man of the House	Comedy	0.81
Billy Elliot	Drama	0.80
Bend It Like Beckham	Crime,Drama, Romance	0.80
Dogma	Adventure,Comedy,Fantasy	0.80

- The main idea behind item based collaborative filtering is to use other users' preferences and taste to recommend new items to a user.
- **similarity** is a numerical score that quantifies the similarity between two movies.

Collaborative Filtering Model Based on User Ratings



- In this section, I chose **RMSE** (Root Mean Squared Error) as our evaluation metrics . We got the best **RMSE** result with **SVDpp** algorithms.

- **RMSE** computes the mean value of all the differences squared between the true and the predicted ratings and then proceeds to calculate the square root out of the result.

Collaborative Filtering Model Based on User Ratings

- **Get movie recommendation for userid = 25**

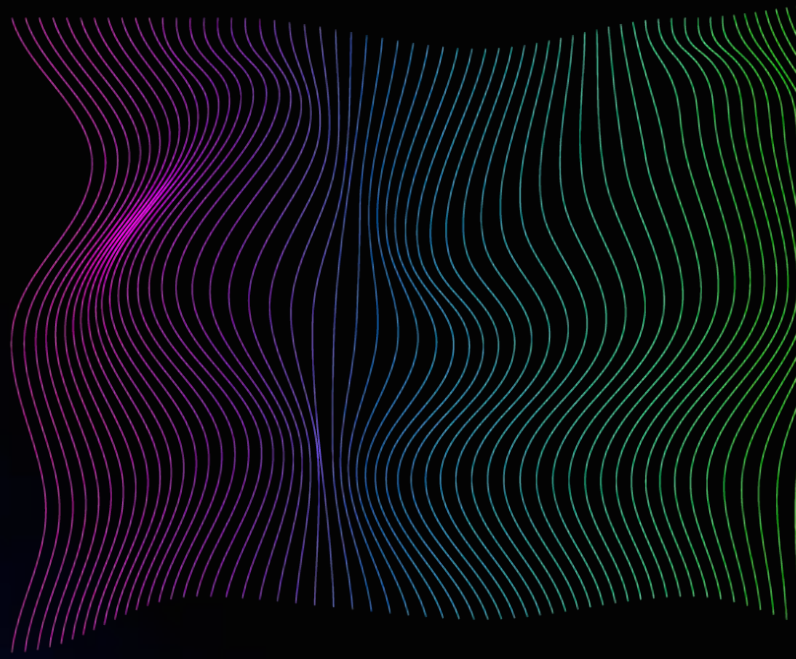
Movie	Genres	Expected Rating
The Shawshank Redemption	Crime,Drama	0.5
Pulp Fiction	Comedy,Crime, Drama	0.49
Silence of the Lambs	Crime,Horror,Thriller	0.49
Star Wars: Episode V	Action,Adventure,Sci-Fi	0.49
Memento	Mystery, Thriller	0.49

- This model recommends those movies that haven't been watched by this user and were loved by other users with similar taste.

05

Conclusion

- First I provided general recommendations based on movie popularity and genre
- Then progressed some collaborative filtering based on movie similarity
- And last,I built an machine learning model and predict ratings on movies that havent been watched then recommended 10 movies to user upon predicted ratings.



05

Recommendations

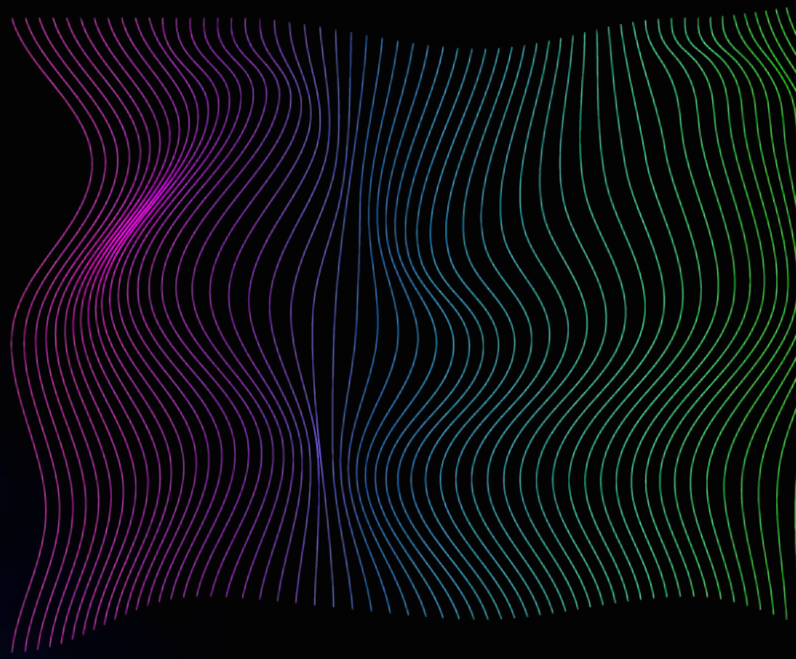
- Most popular genres will be a relevant aspect to take into account when building the content based recommender
- We can use hybrid model of our naive recommendation engine and the model based recommendation engine to solve cold start problem
- Optimizing SVDpp can be more efficient since SVDpp deals with both explicit feedback an implicit feedback

06

Future Work

- Deep learning based recommender system can be built to enhance the performance and provide better recommendations to user.

- And also I can scrape more data from internet and make use of directors or cast dataset to recommend movies to new users upon actors or directors



THANKS

Do you have any questions?

ysm.cebeci@gmail.com

https://www.linkedin.com/in/yesim-cebeci/