

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green color. They are positioned diagonally, with the blue one in front of the green one.

Movie Recommendation System

Htoo Aung Latt

Why add a recommendation system?

Amount of movie on the platform increases every month.

The tremendous growth of the movie catalogue on our platform has an adverse effect of making browsing movies tiring.

- Increase customer satisfaction
- Increase ease of use
- Direct effect on business profit

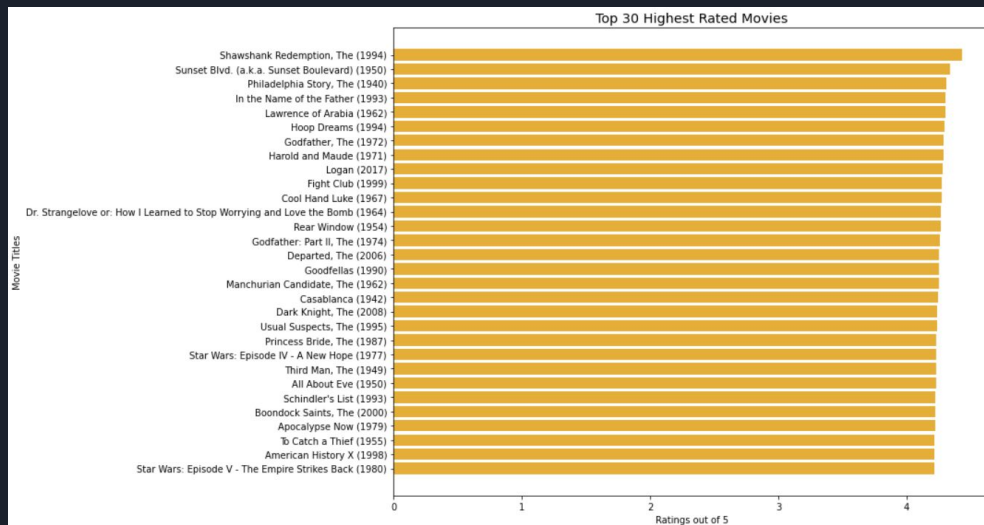


Unpersonalized Recommendation

A list of the top 30 average highest rated movies on our platform.

A good idea to implement in the platform so that users can take a look at the most popular movies.

Fails to recommend for a more specific taste. For that, we need a personalized recommendation system.



movieId	rating	title	genres
318	4.429022	Shawshank Redemption, The (1994)	Crime Drama
922	4.333333	Sunset Blvd. (a.k.a. Sunset Boulevard) (1950)	Drama Film-Noir Romance
898	4.310345	Philadelphia Story, The (1940)	Comedy Drama Romance
475	4.300000	In the Name of the Father (1993)	Drama
1204	4.300000	Lawrence of Arabia (1962)	Adventure Drama War
246	4.293103	Hoop Dreams (1994)	Documentary
858	4.289062	Godfather, The (1972)	Crime Drama
1235	4.288462	Harold and Maude (1971)	Comedy Drama Romance
168252	4.280000	Logan (2017)	Action Sci-Fi
2959	4.272936	Fight Club (1999)	Action Crime Drama Thriller
1276	4.271930	Cool Hand Luke (1967)	Drama
750	4.268041	Dr. Strangelove or: How I Learned to Stop Worr...	Comedy War
904	4.261905	Rear Window (1954)	Mystery Thriller
1221	4.259690	Godfather. Part II, The (1974)	Crime Drama
48516	4.252336	Departed, The (2006)	Crime Drama Thriller
1213	4.250000	Goodfellas (1990)	Crime Drama
1267	4.250000	Manchurian Candidate, The (1962)	Crime Thriller War
912	4.240000	Casablanca (1942)	Drama Romance
58559	4.238255	Dark Knight, The (2008)	Action Crime Drama IMAX
50	4.237745	Usual Suspects, The (1995)	Crime Mystery Thriller
1197	4.232394	Princess Bride, The (1987)	Action Adventure Comedy Fantasy Romance
260	4.231076	Star Wars: Episode IV - A New Hope (1977)	Action Adventure Sci-Fi
1212	4.229167	Third Man, The (1949)	Film-Noir Mystery Thriller
926	4.229167	All About Eve (1950)	Drama



Personalized Recommendations

Content-Based vs Collaborative Filtering

Content Based

- Make use of tags to find similar items
- Requires manpower and time to tag the items

Collaborative Filtering

The main idea is that similar users share the same interests and users tend to like items that are similar.

- KNN Neighbour / Memory-Based
 - User-based - users are compared
 - Item-based - items are compared. Tagging is not needed, instead common users between items are compared.
- Model-Based/ Matrix Factorization
 - Scales well with large data
 - Complete input data is not required

K-Nearest Neighbour Item Based Recommender

`similar_movies("Toy Story")`

	Title	Distance
1	Back to the Future (1985)	0.469619
2	Groundhog Day (1993)	0.463985
3	Mission: Impossible (1996)	0.461087
4	Star Wars: Episode VI - Return of the Jedi (1983)	0.458911
5	Lion King, The (1994)	0.458855
6	Forrest Gump (1994)	0.451747
7	Star Wars: Episode IV - A New Hope (1977)	0.442612
8	Independence Day (a.k.a. ID4) (1996)	0.433992
9	Jurassic Park (1993)	0.433073
10	Toy Story 2 (1999)	0.427399

`similar_movies("Mean Girls")`

	Title	Distance
1	Friends with Benefits (2011)	0.528908
2	Harold and Kumar Go to White Castle (2004)	0.522654
3	Knocked Up (2007)	0.517965
4	Girl Next Door, The (2004)	0.515201
5	Bridesmaids (2011)	0.505240
6	Bring It On (2000)	0.500826
7	Napoleon Dynamite (2004)	0.495866
8	How to Lose a Guy in 10 Days (2003)	0.488317
9	Elf (2003)	0.466852
10	Legally Blonde (2001)	0.440918

Matrix Factorization/ Model-Based

Latent Variable Model

- explain complex relationships between variables using simple relationships between variables and their unobservable, underlying “latent” variables.
- preferences of users can be determined by hidden factors called embeddings that can be gotten using matrix factorization.

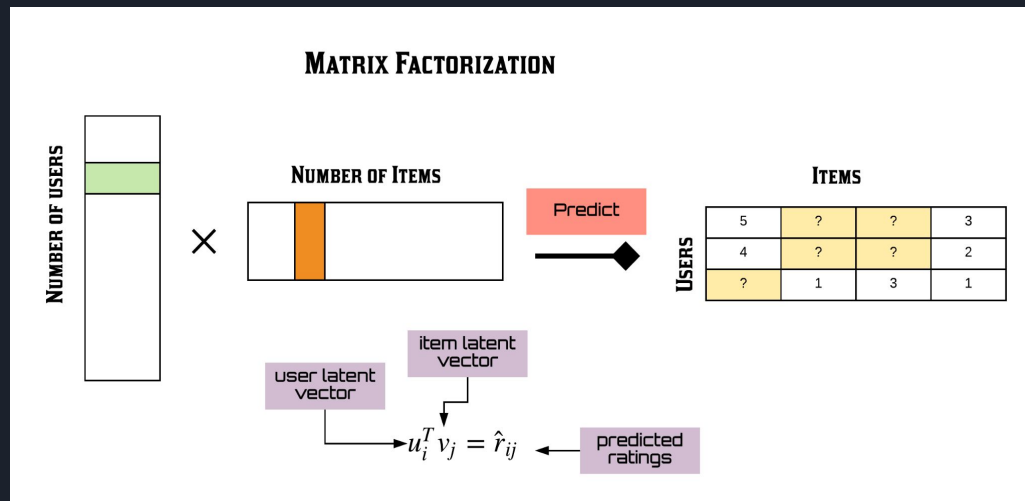
Examples of Hidden Embeddings

Movie - How recently the movie is released.

- How action-driven. How dialogue-heavy.

User - How much does the user like romantic movies.

- How much does the user like old movies.

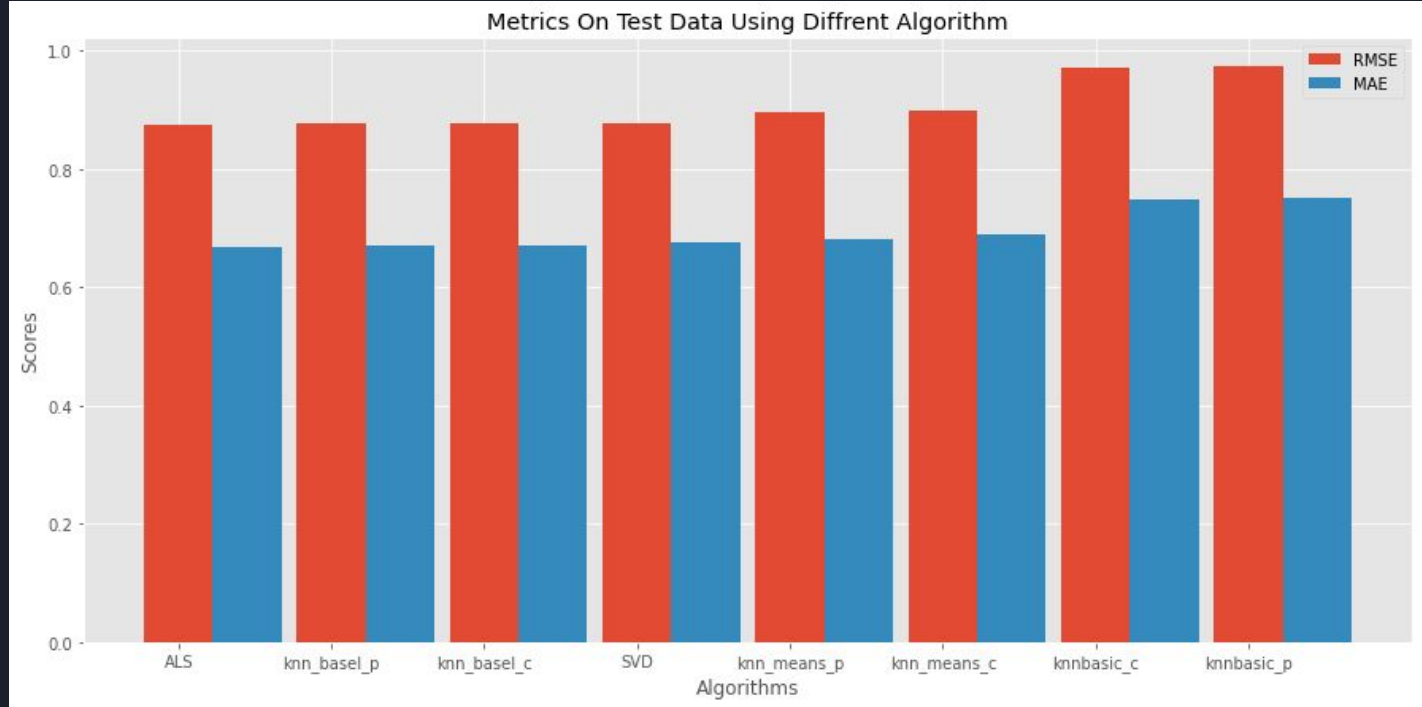


Results of different methods

The scores of the top two memory-based models and the matrix factorization models are very close.

Matrix Factorization models scales better with large dataset.

ALS shines when dealing with enormous and sparse datasets.



The Final Model - Hybrid Recommender

To the right: user with ID 200 gives new ratings to movies in the “War” genre. “n” is entered for movies the user haven’t seen

Below: The new ratings are added to the database before predicted ratings are calculated and returned.

```
recommender(200, genre='War', new_ratings=new_ratings_df)
```

	Ranking	Rating	title	genres
movieId				
32892	3.0	5.342887	Ivan's Childhood (a.k.a. My Name is Ivan) (Iva...	Drama War
26073	17.0	5.122548	Human Condition III, The (Ningen no joken III)...	Drama War
6818	124.0	4.805233	Come and See (Idi i smotri) (1985)	Drama War
170705	126.0	4.797941	Band of Brothers (2001)	Action Drama War
98154	129.0	4.794590	Lincoln (2012)	Drama War

```
new_ratings_df = movie_rater(200, genre="War")
```

```
movieId      title      genres
6671    57528  Rambo (Rambo 4) (2008)  Action|Drama|Thriller|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
3
movieId      title      genres
9497    170705  Band of Brothers (2001)  Action|Drama|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
5
movieId      title      genres
909      1208  Apocalypse Now (1979)  Action|Drama|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
5
movieId      title      genres
4077      5817  Ararat (2002)  Drama|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
n
movieId      title      genres
4877      7307  Flesh & Blood (1985)  Action|Adventure|Drama|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
n
movieId      title      genres
7344      78142  Baaria (2009)  Comedy|Drama|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
4
movieId      title      genres
933      1233  Boot, Das (Boat, The) (1981)  Action|Drama|War
How do you rate this movie on a scale of 1-5, press n if you have not seen :
4
```




Thank You