Recommender Systems

What should I watch next?

Evangelos Tsakeridis Katerina Maxouri Fivos Tzavellos

Mean Rating

- Formula: $\mathrm{R} = \frac{\sum x}{N}$
- Simple, Quick and Easy
- Skewed and Unreliable

Name	Mean
Lost: Season 1	4.661100
The Simpsons: Season 6	4.588155
Family Guy: Freakin' Sweet Collection	4.523465
Inu-Yasha	4.462002
Six Feet Under: Season 4	4.461650
The Best of Friends: Vol. 4	4.445550
Lord of the Rings: The Fellowship of the Ring	4.433316
Gilmore Girls: Season 3	4.432909
The West Wing: Season 3	4.428331
Stargate SG-1: Season 8	4.425880

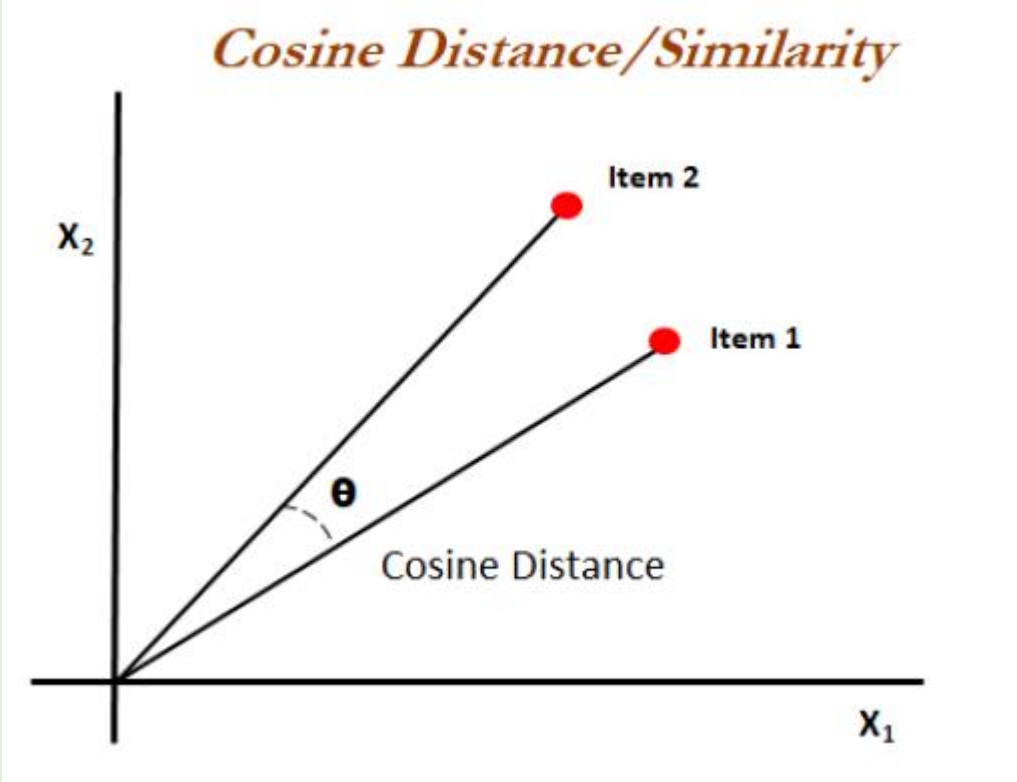
Weighted Mean Rating

- Formula: $W = \frac{Rv+Cm}{v+c}$
  Where:
  - W = weighted rating
  - R = average for the movie
  - v = number of votes for the movie
  - m = minimum votes
- Simple, Quick and Easy

weighted_rating	Name
4.193170	Lord of the Rings: The Fellowship of the Ring
4.148703	Finding Nemo (Widescreen)
4.104791	The Sixth Sense
4.084403	The Silence of the Lambs
4.051759	Braveheart
4.034834	The Godfather
3.995032	Pirates of the Caribbean: The Curse of the Bla...
3.948087	Shrek 2
3.911413	Ray
3.911147	The Wizard of Oz: Collector's Edition

Cosine Similarity

- Formula: $\cos \vartheta = \frac{AB}{\|A\|\|B\|}$
- Item-Based and User-Based



Cosine Similarity

Item-Based

Look for movies with similar features

```
recommend_movie("Chip and Potato")
```

executed in 42ms, finished 23:12:25 2020-01-21

```
Pac-Man and the Ghostly Adventures
The Minimighty Kids
Super Monsters
Hatchimals | Adventures in Hatchtopia
JingleKids
Super Monsters Monster Party
My Little Pony: Friendship Is Magic
Edgar Rice Burroughs' Tarzan and Jane
The Deep
ChuChu TV Kids Songs, Learning Videos & Bedtime Stories
```

User-Based

Look for user that like the same movies

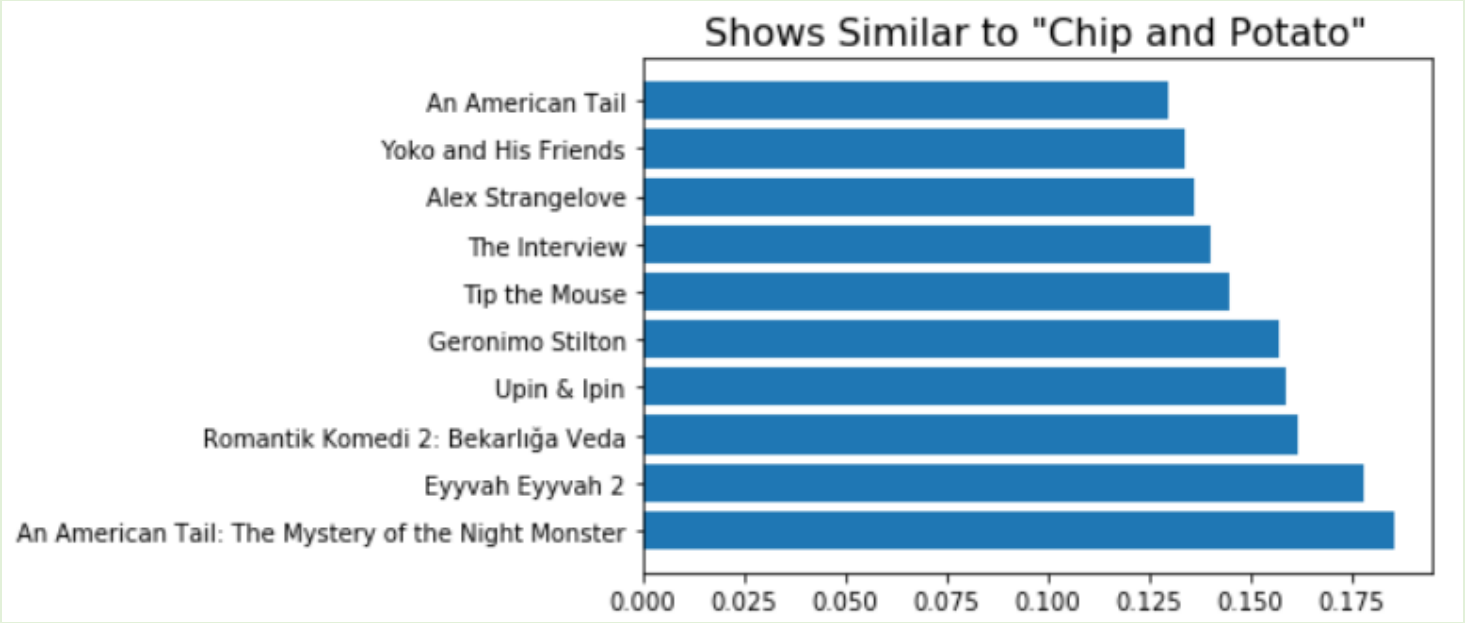
Ranking Of Top 10 Recommended Movies For A User Based On Similarity

Movie_Id	0	Year	Name
4496	3.58	1993.0	Farewell My Concubine
1395	3.58	1963.0	Charade
1561	3.58	2003.0	American Wedding
1558	3.58	1990.0	Rocky V
1553	3.58	1995.0	Persuasion
1547	3.58	1939.0	The Women
1543	3.58	1995.0	Nick of Time
1542	3.58	1993.0	Sleepless in Seattle
1532	3.58	1992.0	Mo' Money
1530	3.58	2002.0	Rose Red

Cosine TFIDF Description Similarity

TF - IDF (Term Frequency - Inverse Document Frequency)

- Text -> Real-Valued Vector
- Apply Cosine Similarity
- Profit(?)



Collaborative Filtering - SVD

- Singular Value Decomposition
- Formula: $X = USV^T$
  Where:
  - U: user features
  - S: diagonal matrix of singular values
  - V^T: movie features
- Independent of the contents of recommended items
- It can be closely integrated with social networks
- Good accuracy in terms of recommendations

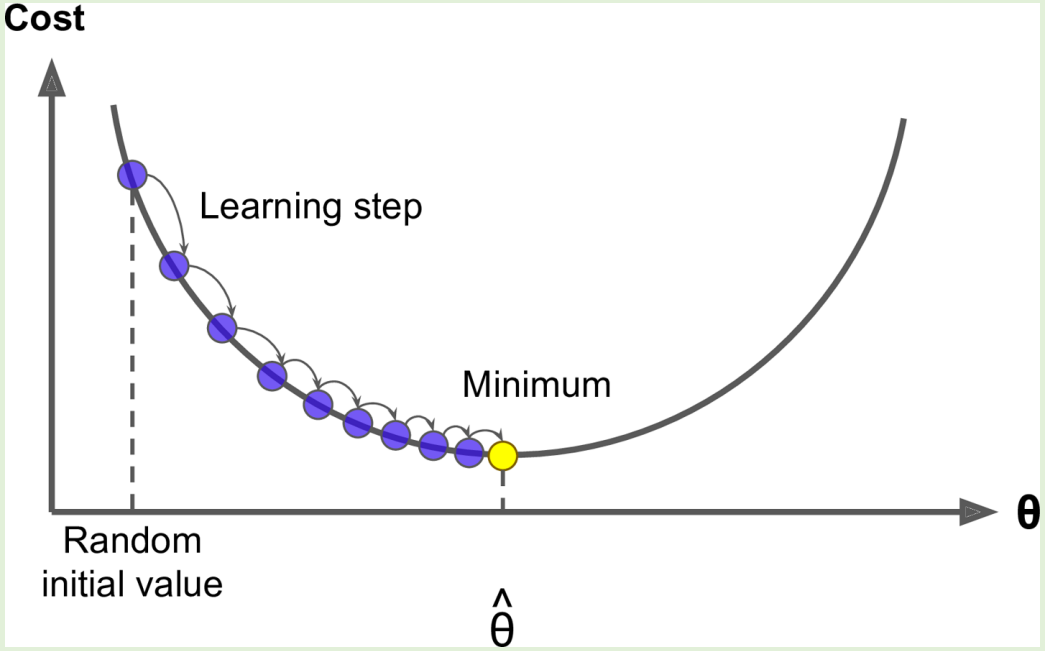
```
recommend_movies(268, df_title)
```

executed in 927ms, finished 23:30:42 2020-01-21

```
The Family Man
Training Day
The Perfect Storm
Taking Lives
Love Actually
Eternal Sunshine of the Spotless Mind
The Others
Lock
Ever After: A Cinderella Story
Sixteen Candles
```

Matrix Factorization With Keras And Gradient Descent

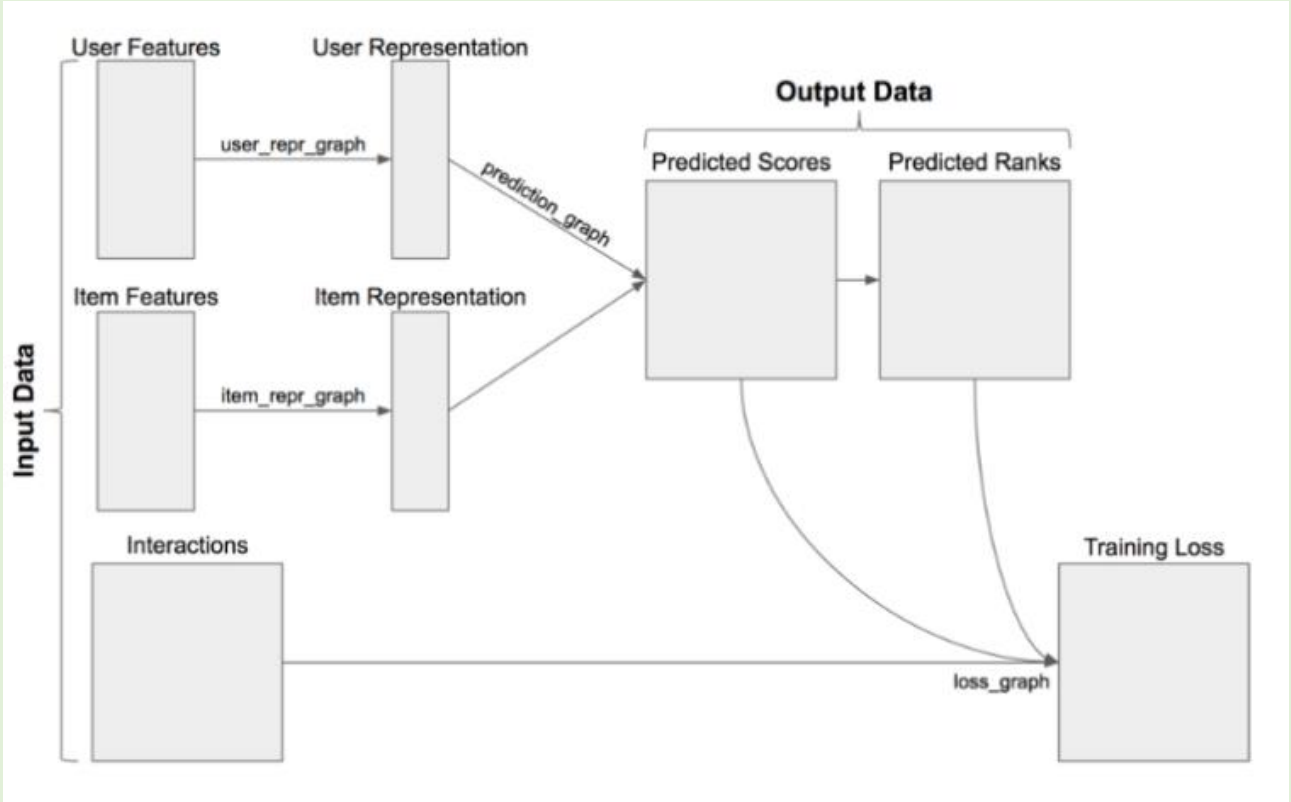
- Formula: $X = U * M + b_U + b_M + b_G \approx \hat{X}$
  Where:
  - X: user-movie ratings matrix (number of users x number of movies)
  - U: users (number of movies x k)
  - M: movies (number of users x k)
  - k: latent variable that encodes a movie or user
  - b_U, b_M, b_G: user, movie and global biases
  - $\hat{X}$: approximation of X




```
Train on 12482969 samples, validate on 1386997 samples
12482969/12482969 [==============================] - 396s 32us/sample - loss: 2.0394 - val_loss: 0.9527
```

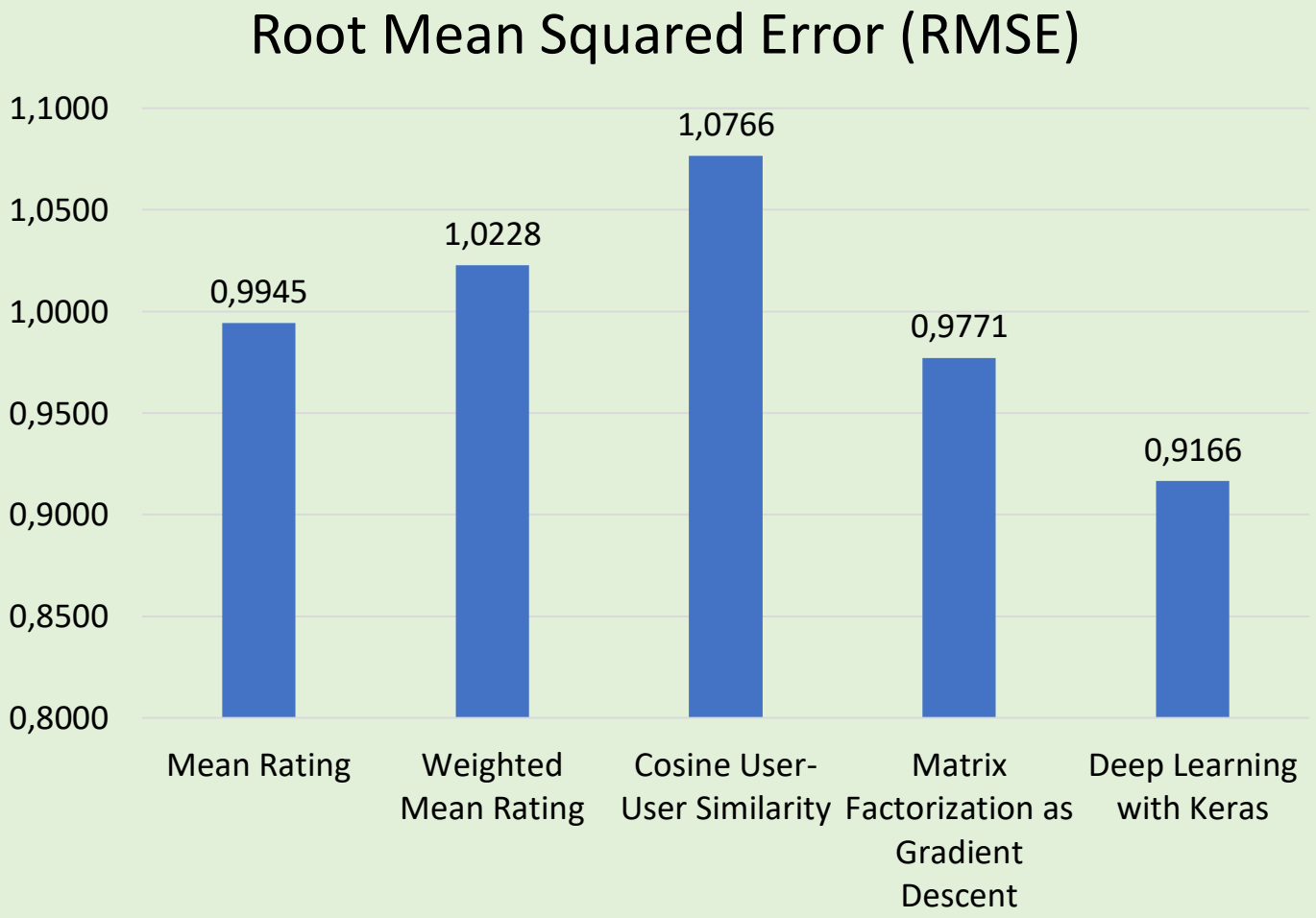
Deep Learning - Keras

Let an algorithm figure out what features are important and how much they matter.




```
Train on 12482969 samples, validate on 1386997 samples
12482969/12482969 [==============================] - 822s 66us/sample - loss: 0.8720 - val_loss: 0.8387
```

Results



Other important factors:

- Diversity
- Recommender persistence
- Privacy
- User demographics
- Robustness
- Serendipity
- Trust
- Labelling

Thank you for your time.

Questions?