

What should we watch tonight? A Movie Recommender System

Top Picks for BENJAMIN



Benjamin W. Sturm
June 28, 2018

Recommender System Goals

Take control over the movies I watch

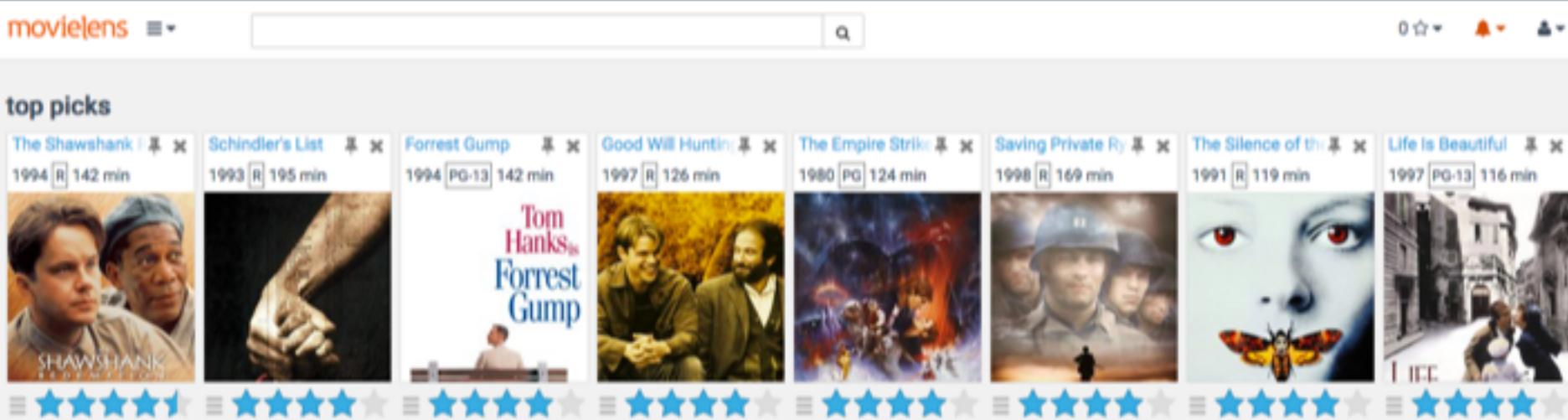
Provide top recommended movies and a predicted score

Be able to make recommendations for couples



Movie Recommender Data

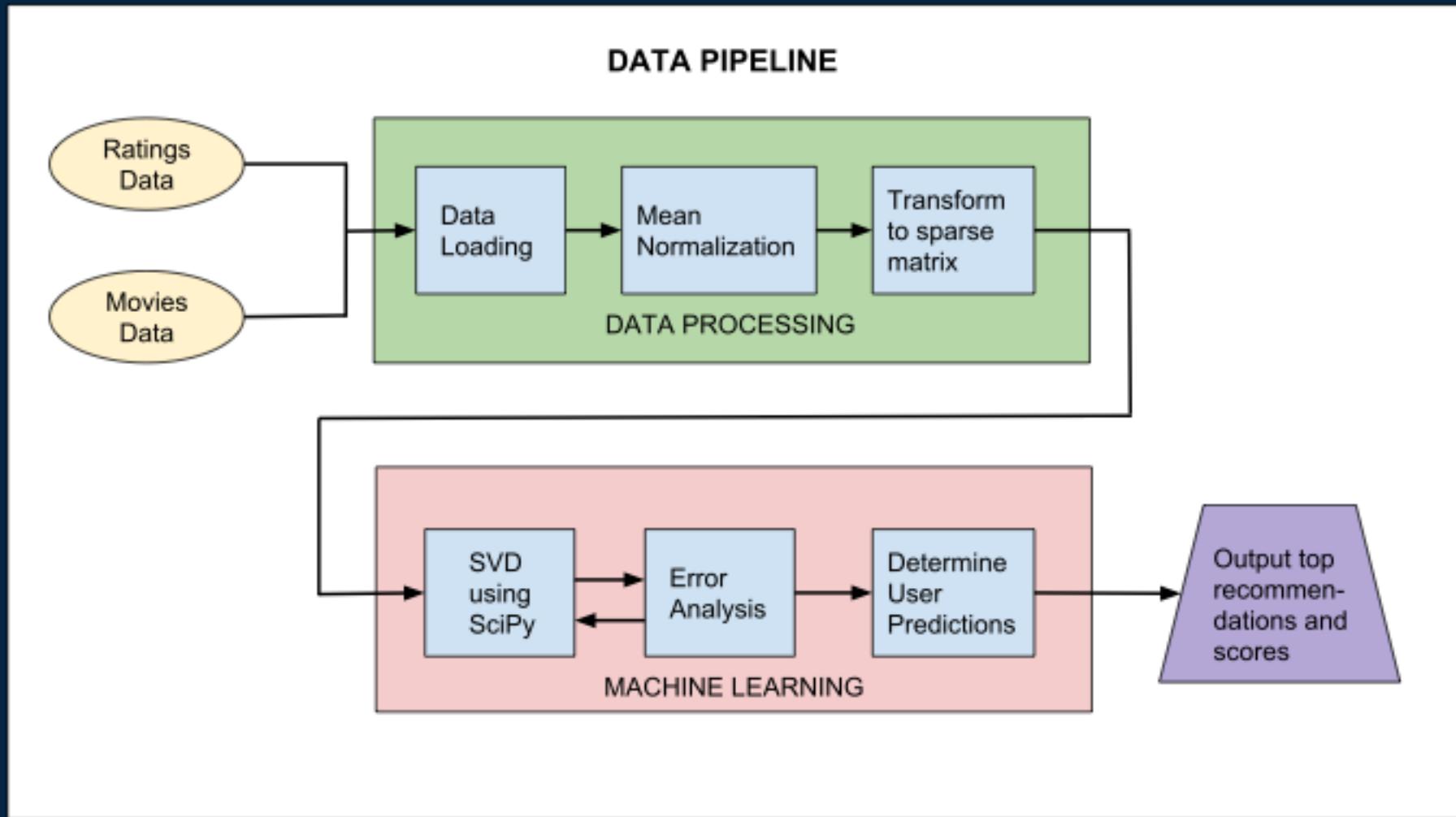
- Employed the MovieLens 20 Million Ratings Dataset



- Total number of users: 138,000
- Total number of movies: 27,000
- Scores range from 0.5 – 5.0

Modeling Approach & Data Pipeline

Implemented a collaborative filtering based recommender



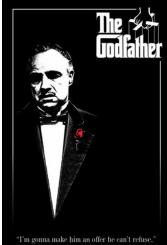
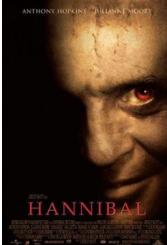
Model Tuning

Tuned my model by adjusting the number of components k

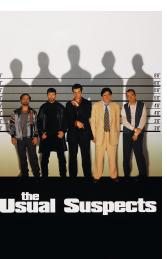
k	Mean Absolute Error	Root Mean Squared Error
100	0.744	0.958
50	0.742	0.958
15	0.737	0.946

Recommendations for a user that enjoys Thrillers

User Ratings

Movie	 The Godfather "I'm gonna make him an offer he can't refuse."	 HANNIBAL	 FIGHT CLUB MISCHIEF. MAYHEM. SOAP. BRAD PITT EDWARD NORTON KELLA DURRUM CARTER	 Austin Powers The Spy Who Shagged Me	 Ernest saves Christmas
Rating	5.0	5.0	5.0	3.0	1.0

Top Recommendations

Movie	 SHAWSHANK REDEMPTION Two Thunders Up!	 The Silence of the Lambs	 A Quad City Double Feature PULP FICTION A Quentin Tarantino Film JOHN TRAVOLTA SAMUEL L. JACKSON TIMOTHY O'ROURKE RALPH MCQUARIE DAVID BOWIE MARS VALERIA MARINA DE NICOLESI VITO RAVANELLO ROBERT DUVALL ROSEANNE ASSANTE CHRISTOPHER WALKEN DIR. QUENTIN TARANTINO	 BRAVEHEART	 the Usual Suspects
Predicted Rating	5.0	5.0	4.9	4.9	4.7

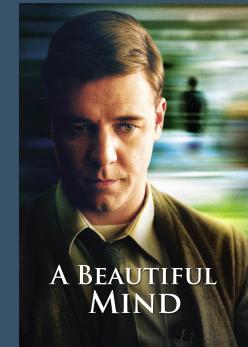
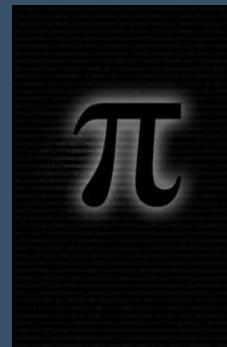
Recommendations for Couples

This is a challenging problem due to unique individual preferences

My solution:

- Create separate recommendations for each user
- Merge the two lists using a group rating function

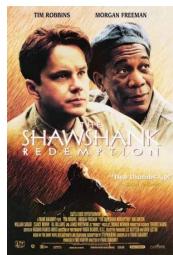
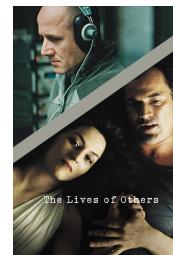
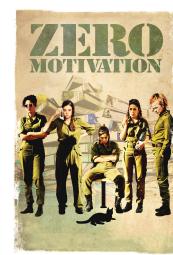
$$R_G = \left(\frac{R_1 + R_2}{2} \right) - \left(\frac{|R_1 - R_2|}{5} \right)$$



RUser1	5.0	4.2
RUser2	3.0	3.8
Ravg	4.0	4.0
RG	3.6	3.92

Recommendations for my better half and I

Top Picks or Us

Movie						
Rating	4.5	4.3	4.3	4.5	4.3	4.3



Outlook for Future Work

Do topic modeling on movie tags dataset

Incorporate topic data into a hybrid model

Build a web app

Thank you!



Benjamin W. Sturm

- Ph.D. in Nuclear Engineering

email: bwsturm@gmail.com

LinkedIn: bwsturm

Github: bwsturm