

A Content-based Movie Recommender System

Solving Cold-start Problem

Yanjie He

December 6, 2018

Example: The Recommender System of Netflix

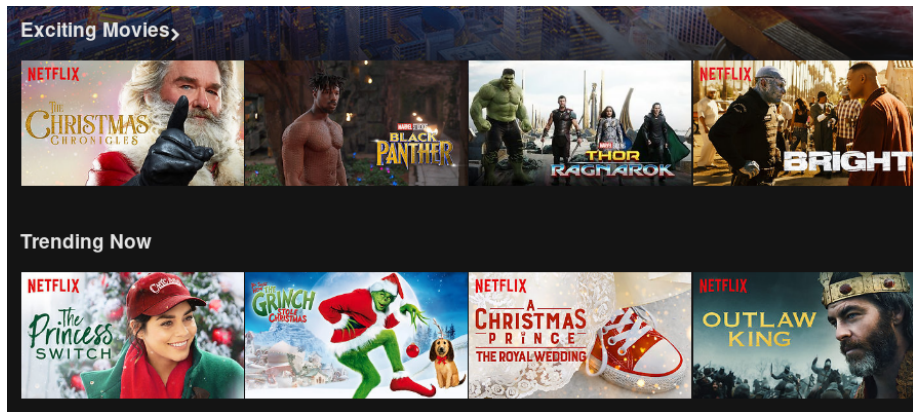


Figure: Netflix Recommender System

Rating Matrix

						
	4	3			5	
	5		4		4	
	4		5	3	4	
		3				5
		4				4
			2	4		5

Figure: Rating Matrix

Matrix Factorization

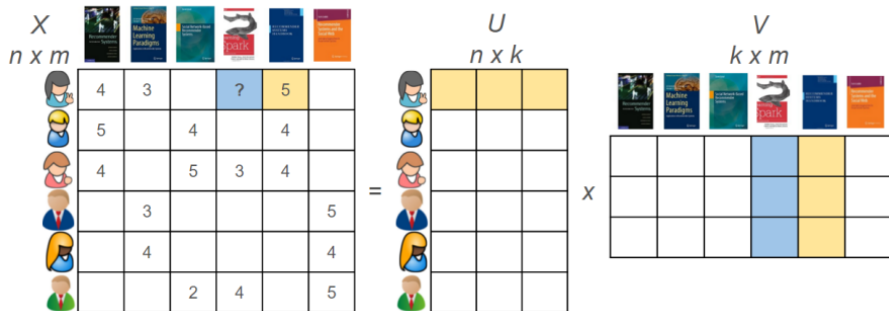


Figure: Matrix Factorization

Cold-start Problem



Figure: Cold-start Problem

Genres

The movielens dataset contains the genres of each movie, the genres are listed as followings: Action, Adventure, Animation, Children's, Comedy, Crime, Documentary, Drama, Fantasy, Film-Noir, Horror, Musical, Mystery, Romance, Sci-Fi, Thriller, War, and Western.

One-hot Encoding

movieId	Action	Adventure	Animation	Children	Comedy	Crime	Documentary
1	0	1	1	1	1	0	0
2	0	1	0	1	0	0	0
3	0	0	0	0	1	0	0
4	0	0	0	0	1	0	0
6	1	0	0	0	0	1	0
7	0	0	0	0	1	0	0
8	0	1	0	1	0	0	0
9	1	0	0	0	0	0	0
10	1	1	0	0	0	0	0

Table: One-hot Encoding for Genres

Information Retrieval

- Movie trailers (images and audio)
- Movie network

Image Processing

- Since every movie trailer is a video, and a video is consist of frames.
- For a video which fps (frame per second) is 24, every second in this movie has 24 frames.
- One frame for each second. A three-minute movie trailer can give us $60 * 3 = 180$ frames.

RGB

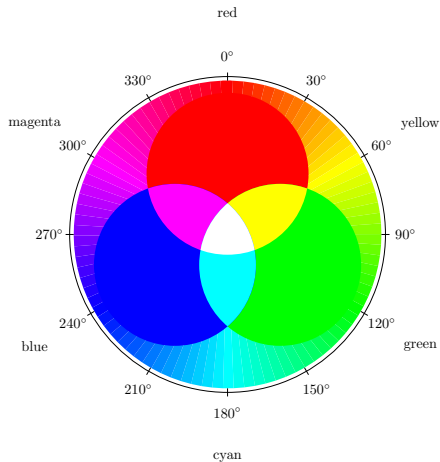


Figure: RGB Color

HSV

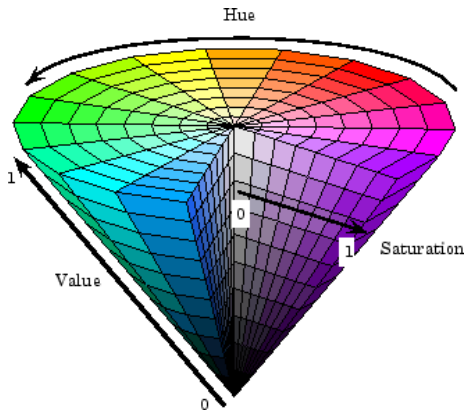


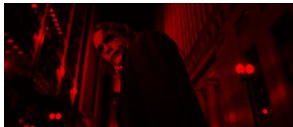
Figure: HSV Color

Image Processing

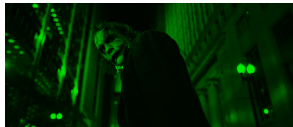


Figure: One frame of a movie trailer

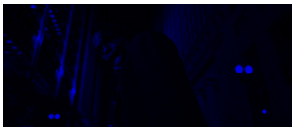
Image Processing (RGB Colorspace)



(a) Red



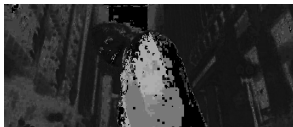
(b) Green



(c) Blue

Figure: One frame of a movie trailer in RGB (Red, Green, Blue)

Image Processing (HSV Colorspace)



(a) Hue



(b) Saturation



(c) Value

Figure: One frame of a movie trailer in HSV (Hue, Saturation, Value)

HSV Color Histogram

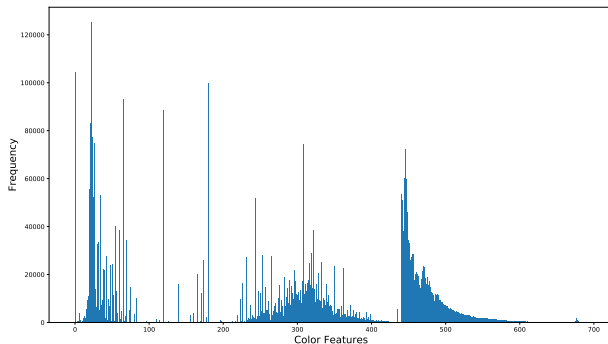


Figure: HSV histogram

Audio Processing (Wave)

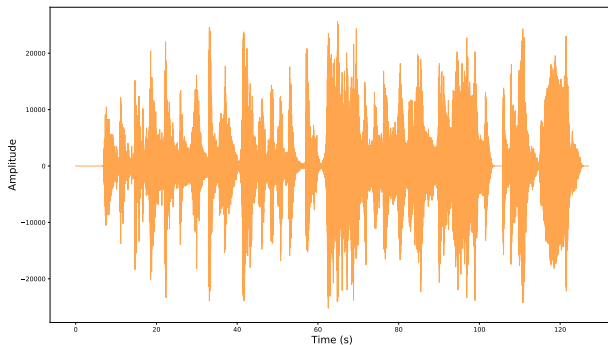


Figure: Amplitude of a Movie Trailer

Audio Processing (Chromatogram)

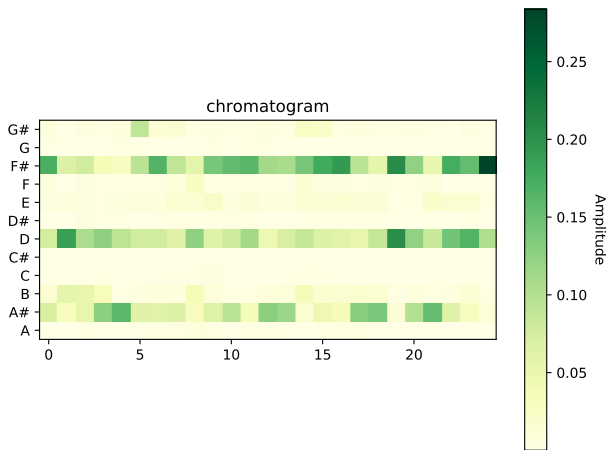


Figure: Chromatogram of a Movie Trailer

Toy Story (1995): Memory-based Recommendations

title	distance
Emperor's New Groove, The (2000)	0.0
Antz (1998)	0.0
Adventures of Rocky and Bullwinkle, The (2000)	0.0
Land Before Time, The (1988)	1.0
Over the Hedge (2006)	1.0
Ice Age 2: The Meltdown (2006)	1.0
Bug's Life, A (1998)	1.0
Wallace & Gromit in The Curse of the Were-Rabb...	1.0
Ice Age (2002)	1.0
Madagascar (2005)	1.0

Table: Recommendations Based on Genres

Toy Story (1995): Memory-based Recommendations

title	distance
Adventures of Rocky and Bullwinkle, The (2000)	0.185827
Antz (1998)	0.240271
Emperor's New Groove, The (2000)	0.357686
Space Jam (1996)	1.014048
Toy Story 3 (2010)	1.016265
Lord of the Rings, The (1978)	1.017530
Horton Hears a Who! (2008)	1.018153
Borrowers, The (1997)	1.018853
Over the Hedge (2006)	1.021717
Madagascar (2005)	1.022015

Table: Recommendations Based on both Genres and Trailers' Features

Network Definition

	Interstellar (2014)	The Dark Knight Rises (2012)
Director:	Christopher Nolan	Christopher Nolan
Writers:	Christopher Nolan	Jonathan Nolan Christopher Nolan David S. Goyer
Stars:	Leonardo DiCaprio Joseph Gordon-Levitt Ellen Page Tom Hardy Ken Watanabe Marion Cotillard	Christian Bale Gary Oldman Tom Hardy Joseph Gordon-Levitt Anne Hathaway Marion Cotillard

Social Network Analysis

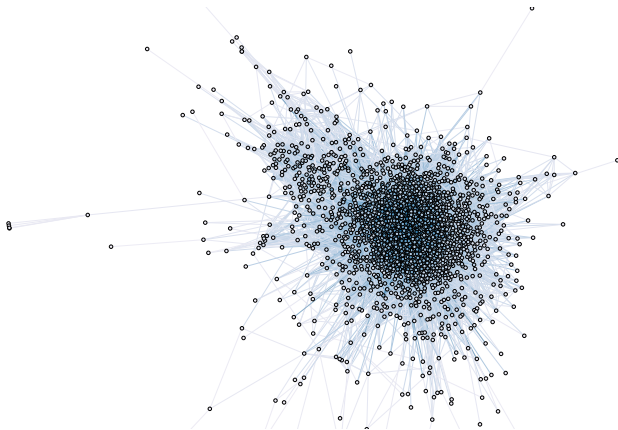


Figure: Movie Network Graph

Thank you for your time!

Any questions?