

# RECOMMENDER SYSTEMS

By: Mike Rhinehart



# WHAT'S ON THE AGENDA?

- Introduction
- Data Preparation
- Exploratory Data Analysis
- Types of Recommender Systems
- Memory-Based Collaborative Recommender System
- Model-Based Collaborative Recommender System
- Content-Based Recommender System



# INTRODUCTION

- The MovieLens datasets are full of data describing how people rate movies. As it turns out, these datasets have been useful to lots of folks, from recommender systems researchers to the readers of popular-press programming books. It was collected by GroupLens researchers over various periods of time and by far the most popular when it comes to implementing a recommender system.



	userId	movieId	rating	timestamp
0	1	31	2.5	1260759144
1	1	1029	3.0	1260759179
2	1	1061	3.0	1260759182
3	1	1129	2.0	1260759185
4	1	1172	4.0	1260759205

	movieId	title	genres
0	1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy
1	2	Jumanji (1995)	Adventure Children Fantasy
2	3	Grumpier Old Men (1995)	Comedy Romance
3	4	Waiting to Exhale (1995)	Comedy Drama Romance
4	5	Father of the Bride Part II (1995)	Comedy

	userId	movieId	rating	title	genres
0	1	31	2.5	Dangerous Minds (1995)	Drama
1	7	31	3.0	Dangerous Minds (1995)	Drama
2	31	31	4.0	Dangerous Minds (1995)	Drama
3	32	31	4.0	Dangerous Minds (1995)	Drama
4	36	31	3.0	Dangerous Minds (1995)	Drama

# DATA PREPARATION

- Read in the movies and ratings datasets
- Merge both datasets on the MovieId
- Removed irrelevant fields
- 100000 rows and 5 columns



# EXPLORATORY DATA ANALYSIS

- Check for missing values
- Which user had the highest number of ratings?
  - User 547 had the highest number ratings with a value of 2391.
- Which movie title had the highest number of ratings?
  - Forest Gump (1994) had the highest number of ratings with a value of 341.
- Notice the highest rating is 5 and the lowest rating is 0.5.
  - The average for the entire dataset is 3.5.





# TYPES OF RECOMMENDER SYSTEMS

- There are two types of recommendation systems, Content-Based and Collaborative Filtering.
  - Collaborative recommendation systems focuses on the similarity attribute of the users, that is, it finds people with similar tastes based on a similarity measure from the large group of users.
  - Content-Based recommendation systems focus on the attributes of the items and give you recommendations based on the similarity between them.





# MEMORY-BASED COLLABORATIVE RECOMMENDER

- In Memory-Based Collaborative Filtering, this recommendation system Memory-based algorithms approach the collaborative filtering problem by using the entire database.
  - **User-User:** It identifies other people with similar tastes to a target user and combines their ratings to make recommendations for that user.
  - **Item-Item:** It identifies global product associations from user ratings, but uses these product associations to provide personalized recommendations based on a user's own product ratings.





# MODEL-BASED COLLABORATIVE RECOMMENDER

- In Model-Based Collaborative Recommender Systems involve building a model based on the dataset of ratings. In other words, we extract some information from the dataset, and use that as a "model" to make recommendations without having to use the complete dataset every time.



# CONTENT-BASED RECOMMENDER

- Content-based recommendation engine works with existing profiles of users. A profile has information about a user and their taste.
  - The engine compares the items that are already positively rated by the user with the items he/she didn't rate and looks for similarities between them.



# QUESTIONS

- Go check out my GitHub for a more in-depth look at my recommender systems I implemented using Python.
  - <https://github.com/mvrhine/What-kind-of-movies-you-like-to-watch->
- You can contact me via LinkedIn for questions about this project.
  - [www.linkedin.com/in/michael-rhinehart](http://www.linkedin.com/in/michael-rhinehart)

