Personalized Movie Recommendation System

Using Content-Based and Collaborative Filter Methods to Provide Movie Recommendations

By Melvin Garcia
Sept. 9, 2021
Phase 4
Flatiron Cohort - onl01-dtsc-pt-011121

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- MovieTime’s current recommendation system suffers from an apparent cold-start problem
- Users are found spending a significant amount of time search for a new movie compared to other streaming platforms
- Using MovieTime’s new user-generated tag feature
- Generate hybrid recommendation system incorporating new tag feature data

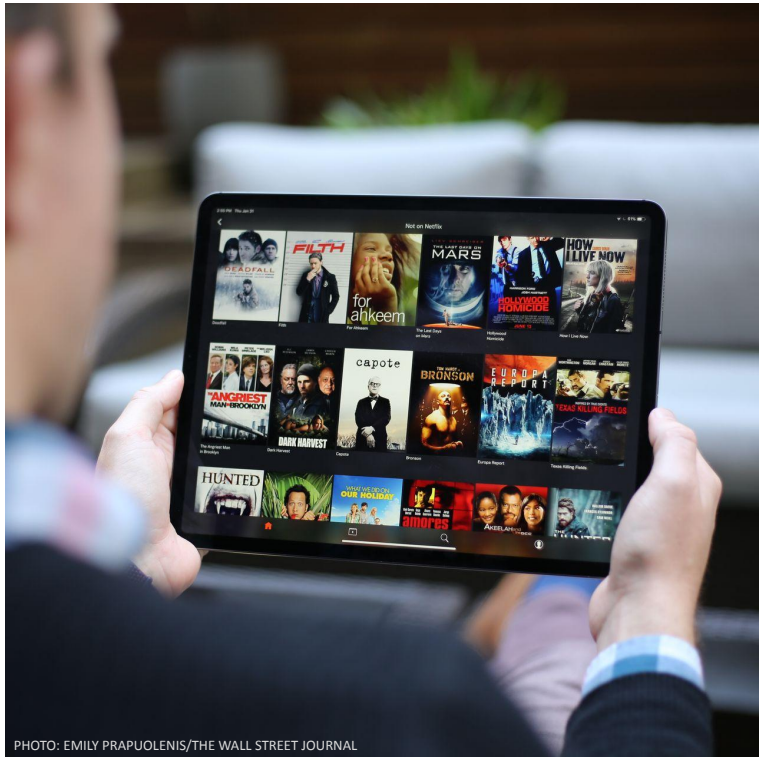
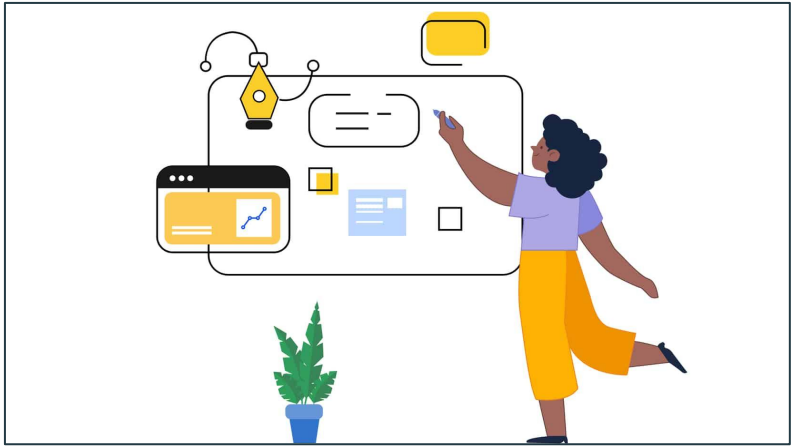
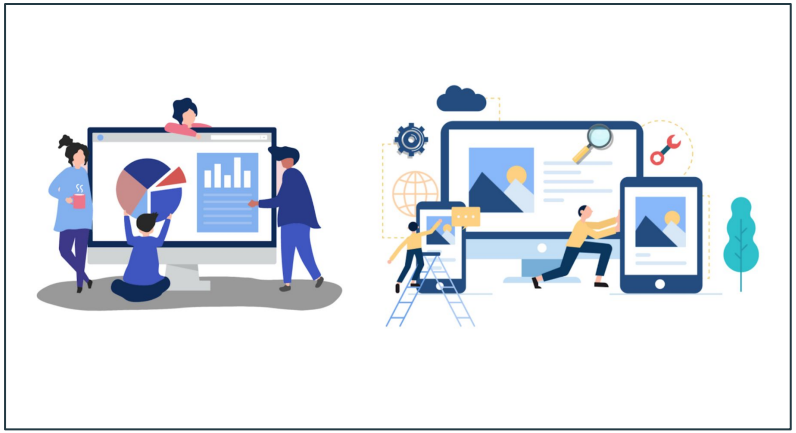


PHOTO: EMILY PRAPUOLENIS/THE WALL STREET JOURNAL

Data

- MovieTime data consists of
  - User-movie ratings
  - User-generated tags for each movie
  - Tags containing -- Genres, Studio, Cast etc
- 25M movie ratings --
  - 620K movies
  - 162K users
  - 1M tags
- Working sample of 100K ratings --
  - 9K movies
  - 600 users
  - 3.6K tags





https://www.divami.com/blog/best-practices-for-user-interaction-and-user-engagement/

Methods

Prepare & Explore Data

1. Explore and confirm main tables to use
2. Set up key to join disparate data sources
3. Examine movie rating distribution, and tags

Model Exploration & Iteration

1. Replicate MovieTime's current user-user collaborative filter (CF) model
2. Compare different versions of the CF model
3. Move forward with best CF model to incorporate into end-hybrid model

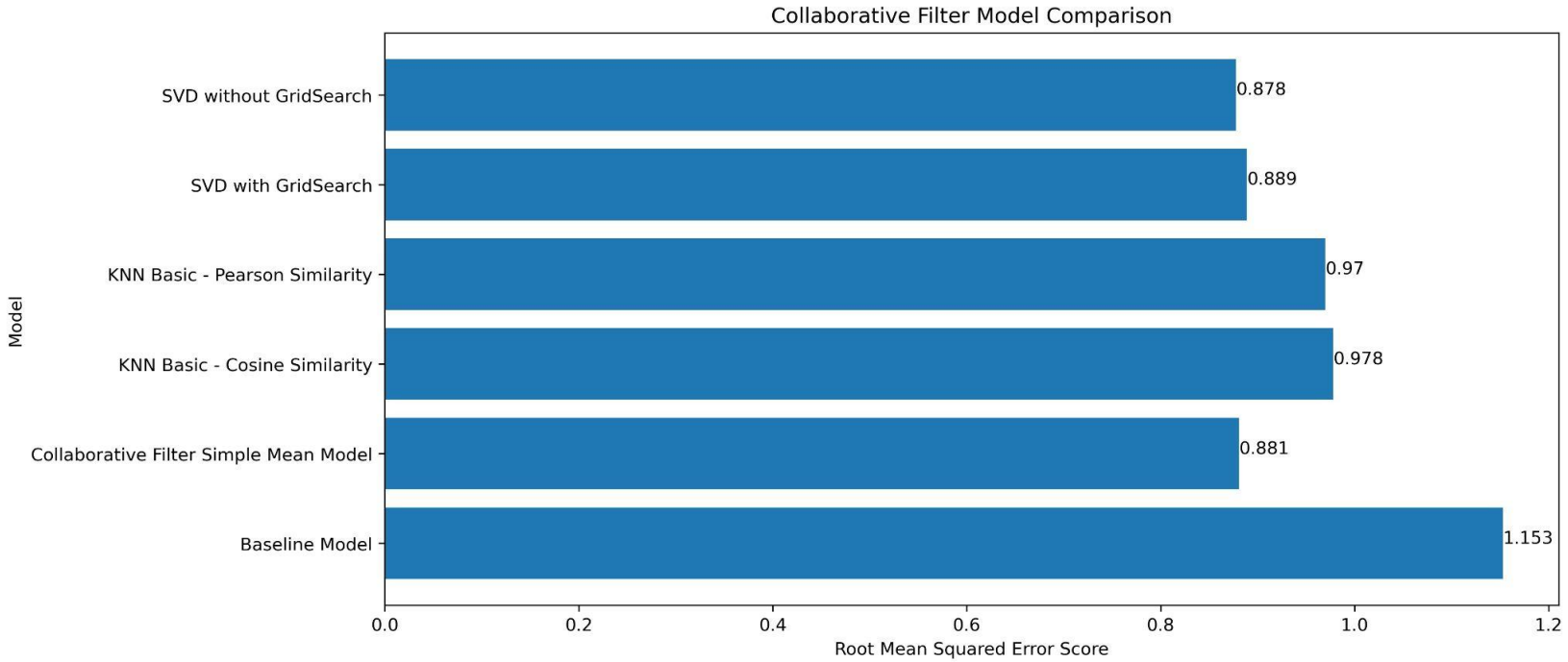
Using Tags for Content-Based Model

1. Extract genres and tags per movie
2. Generate similarity metric to output top N most similar movies to a given movie input (e.g. *Toy Story*)

Creating a Hybrid Model

1. Combine the CF and content-based model
2. Given a user ID and movie title, output a sorted list of movies by predicted rating that are most similar to the given movie title

Results - Collaborative Filter Model Comparison



Model Exploration & Iteration

Results - Toy Story Content Model Example

Movies Most Similar to ***Toy Story***

Movie Title	
A Bug's Life	Disney
Toy Story 2	Disney
Up	Disney
Sintel	
Guardians of the Galaxy 2	Disney
The Cat Returns	
Kiki's Delivery Service	
Alice in Wonderland	Disney
Sinbad: Legend of the Seven Seas	Disney
Who Framed Roger Rabbit?	

Using Tags for Content-Based Model

Results - Toy Story Content Model Example

Movies Most Similar to ***Toy Story***

Using Tags for Content-Based Model

Movie Title		
A Bug's Life	Disney Pixar	Animated
Toy Story 2	Disney Pixar	
Up	Disney Pixar	
Sintel		
Guardians of the Galaxy 2	Disney	
The Cat Returns		
Kiki's Delivery Service		
Alice in Wonderland	Disney	
Sinbad: Legend of the Seven Seas	Disney	
Who Framed Roger Rabbit?		

Results - User-to-User Hybrid Model Comparison

Predicted Ratings for **User ID 1**

Title	Predicted Rating
Up	4.468
Laputa: Castle in the Sky	4.465
My Neighbor Totoro	4.120
Toy Story 2	4.358
Kiki’s Delivery Service	4.352

Results - User-to-User Hybrid Model Comparison

Predicted Ratings for **User ID 1**

Title	Predicted Rating
Up	4.468
Laputa: Castle in the Sky	4.465
My Neighbor Totoro	4.120
Toy Story 2	4.358
Kiki's Delivery Service	4.352

Predicted Ratings for **User ID 448**

Title	Predicted Rating
Up	3.015
Laputa: Castle in the Sky	3.021
My Neighbor Totoro	2.960
Toy Story 2	2.905
Kiki's Delivery Service	2.881

Creating a Hybrid Model

Average Rating Difference = **~1.5**

Conclusions

- In our content-based model using the user generated tags -- it was solid with determining most similar styles of movies (e.g. Disney, Pixar, and Animated films)
- The hybrid model was able to combine the predicted ratings from the CF model and the find most similar movies for a given movie title input
- Distinguish between two users movie tastes over the *Toy Story* movie

Next Steps

- Expand and train our models on a larger dataset
- Enrich our content based model by adding plot descriptions, cast, and crew to our dataset
- Prove (or disprove) higher user engagement metrics using the new hybrid recommender system through a set of designed A/B tests

Thank You!

Email: garciamelvin4@gmail.com
GitHub: @melvyg
LinkedIn: linkedin.com/in/melvinmgarcia/