Recommendation Systems

For Movies

Omar Mustaf

Wondershare PDFelement

Wondershare PDFelement

Agenda

1. The business problem to solve
2. The Dataset
3. Recommendation models overview
4. Evaluating the models

Wondershare PDFelement

Wondershare PDFelement

Introduction

In today's digital landscape, user engagement and satisfaction are pivotal for any platform's success.Recommender systems play a crucial role in achieving these goals.

Wondershare PDFelement

This presentation delves into our recommendation models, aiming to optimize user experiences and elevate engagement..

Wondershare PDFelement

Project Goals

Engagement

70% of time watched on YouTube driven by recommender system (CNet)

Retention

Recommendations saves Netflix more than $1B per year in subscriber churn (Netflix)

Revenue

35% of the purchases on Amazon are the result of their recommender system (McKinsey)

What Movies will the users be interested in?

Wondershare PDFelement

Dataset used

Dataset

We will use MovieLens dataset, created and maintained by GroupLens research group out of University of Minnesota: https://grouplens.org/datasets/movielens/

Wondershare PDFelement

Wondershare PDFelement

Wondershare PDFelement

Dataset info

Offers multiple versions, we will use the one targeting education: ml-latest.zip (5MB)
Dataset contains user ratings with timestamps, user-applied tags, movie genres.

100k
Ratings

1000
Users

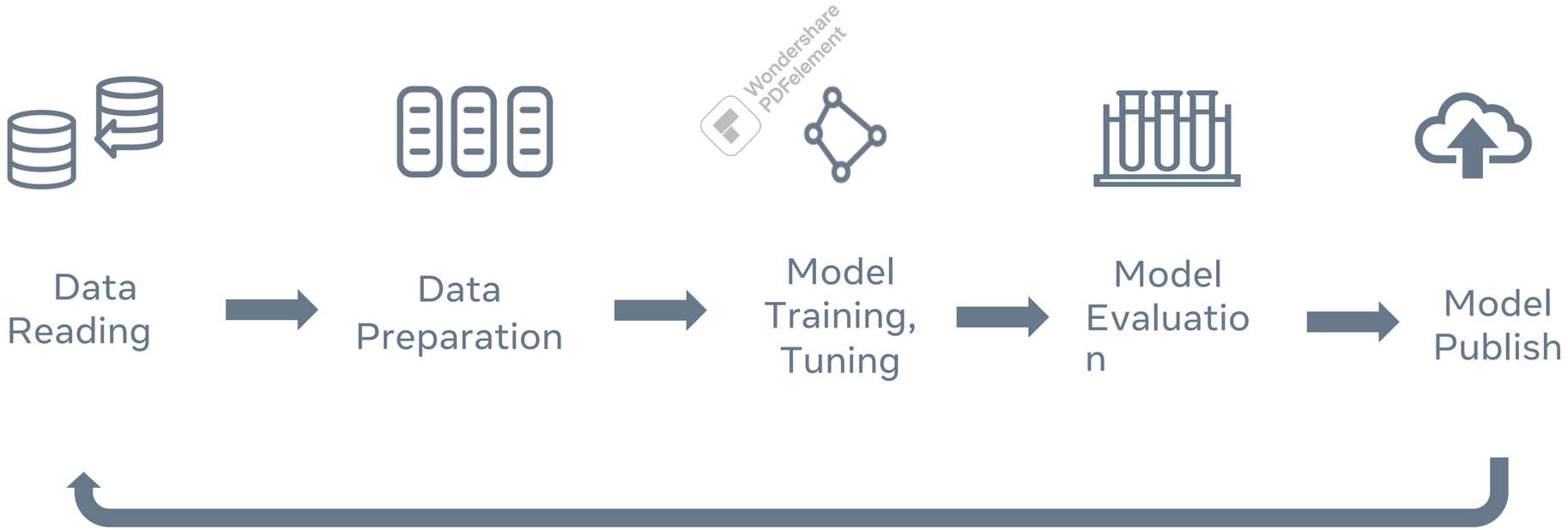
1700
Movies

Wondershare PDFelement

Sparsity: 0.998 (!)

Wondershare PDFelement

Project Steps



Wondershare PDFelement

2 Recommendation Models Overview

Wondershare PDFelement

Wondershare PDFelement

Baseline model: Popularity

Before we start with complex modeling, it is important to set a baseline!

- We will set a baseline for a simple popularity-based model:
  - This model recommend top movies to all the users regardless of preference

Wondershare PDFelement

- Model Pros and Cons:
  - This model is simple to make and it appeals to a broad audience
  - The downside is it lacks personalization and user preferences

Can we do better?

Cosine similarity

How it Works

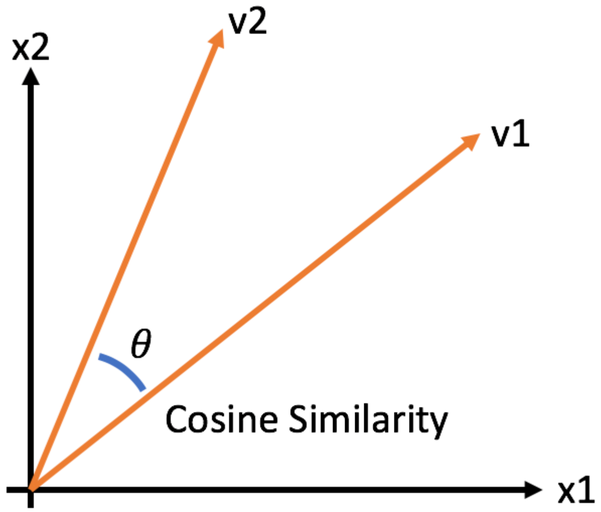
- This model works by getting the cosine of the angle to determine item similarity

Advantage

. Effective for capturing item relationship based on content

Usage

I used this model in my website to recommend similar movies to the input

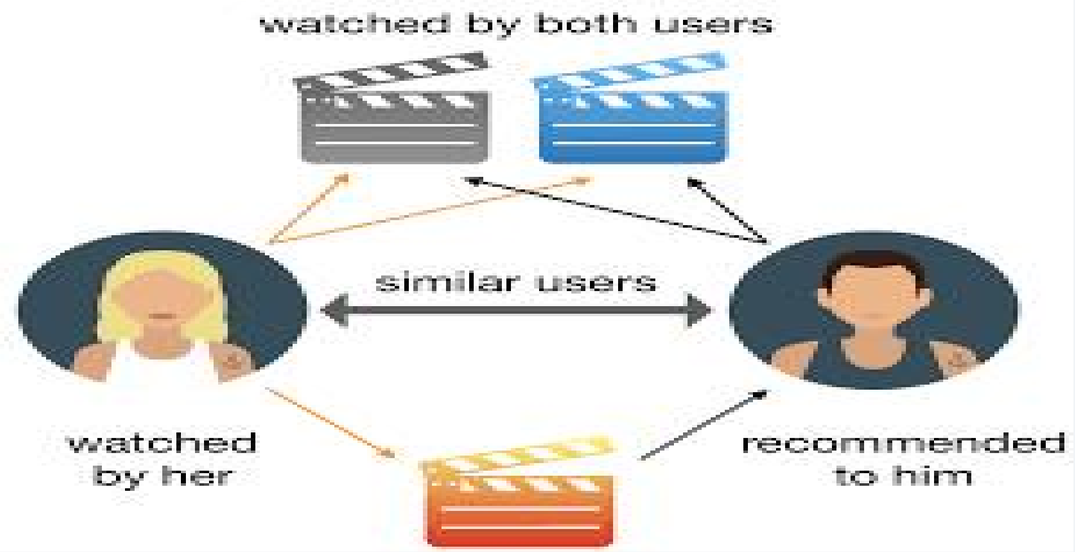


Wondershare PDFelement

Final:AlS Model

How It Works

- This model works by using user-item interactions to make personalized recommendations
- ALS model is a collaborative model which means similiar users are recommended similiar movies



Advantages

This model performs well in making personalized user recommendation

Wondershare PDFelement

Business Problem

Evaluation Overview

- How Th

Wondershare PDFelement

Evalution

Metrics:

- Data: 100K samples (70% training, 30% test)
- Metrics used:
  - RMSE ,RME
  - Novelty ,Personalization
  - Coverage and Recall -

Findings

Results: RMSE 0.
88 and
RME 0.78

The model seems to be doing good on the test data with a good RMSE and RME

Wondershare PDFelement

Wondershare PDFelement

Personalization and Novelty

Novelty

Novelty evaluates how unique recommendations are.

- High novelty means recommendation introduces fresh,less explored items
- Our ALS is 3 times more novel than the baseline model

Wondershare PDFelement

Personalization

Personalization measures how tailored recommendation are to individual users

High personalization means recommendations are unique while low signifies recommendations are similiar.

Findings

+3x vs PopModel with 2.70

0.05 personalization

Coverage

Increasing model capacity enables the model to:

- Coverage assesses the diversity and breadth of recommendation made by the system

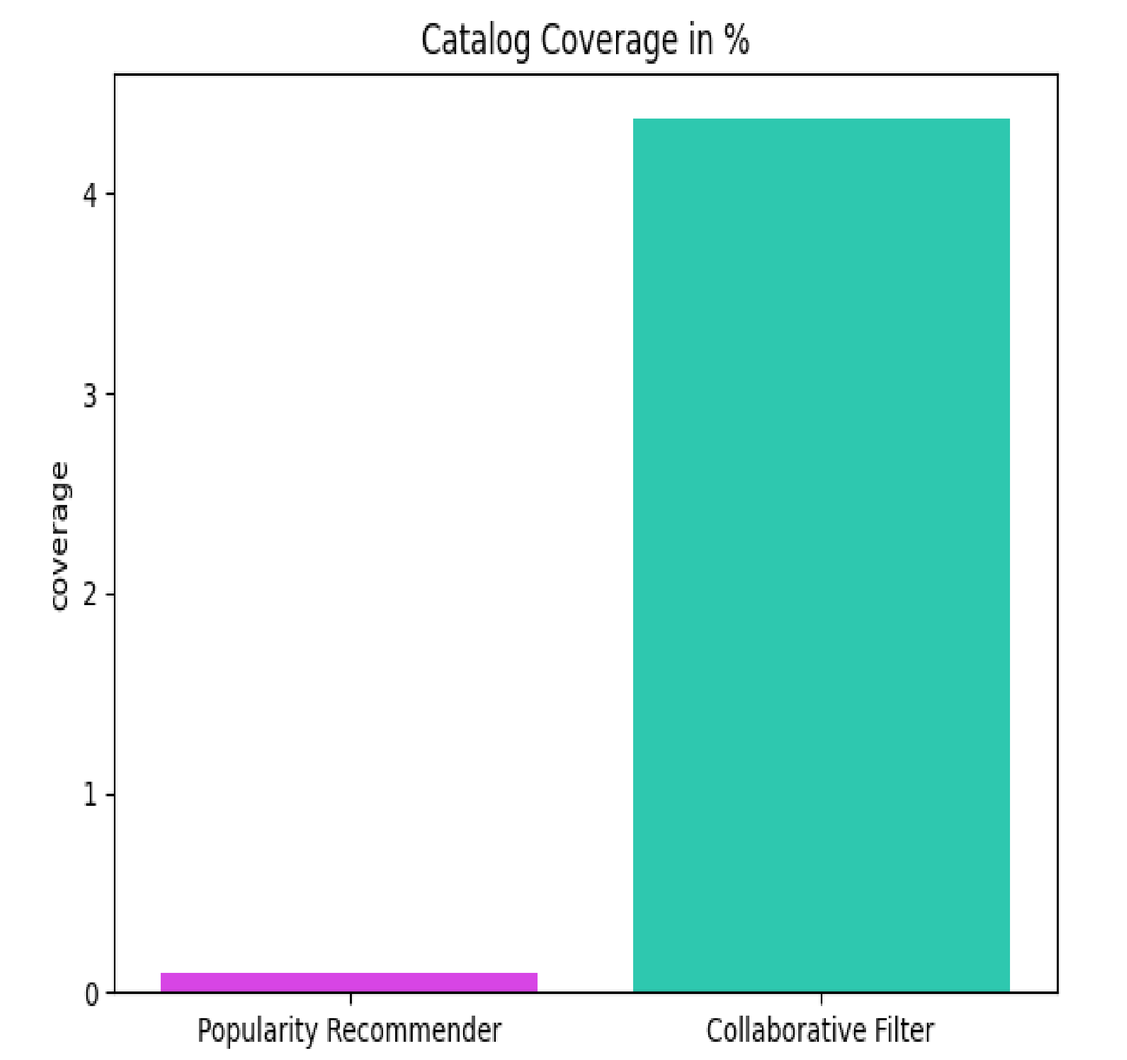
Why it matters:

- High coverage ensures that a wide variety of items are recommended to users
- Low coverage may lead to users seeing a limited subset of available items

Use Cases:

- E-commerce platforms aim for high coverage to present a range of products
- Movie Platforms strive for high coverage leading to users finding good movies

+40% vs PopModel



“Bigger is better ”

Wondershare PDFelement

3 EVALUATING THE MODEL

MAR@K

MAR@K:

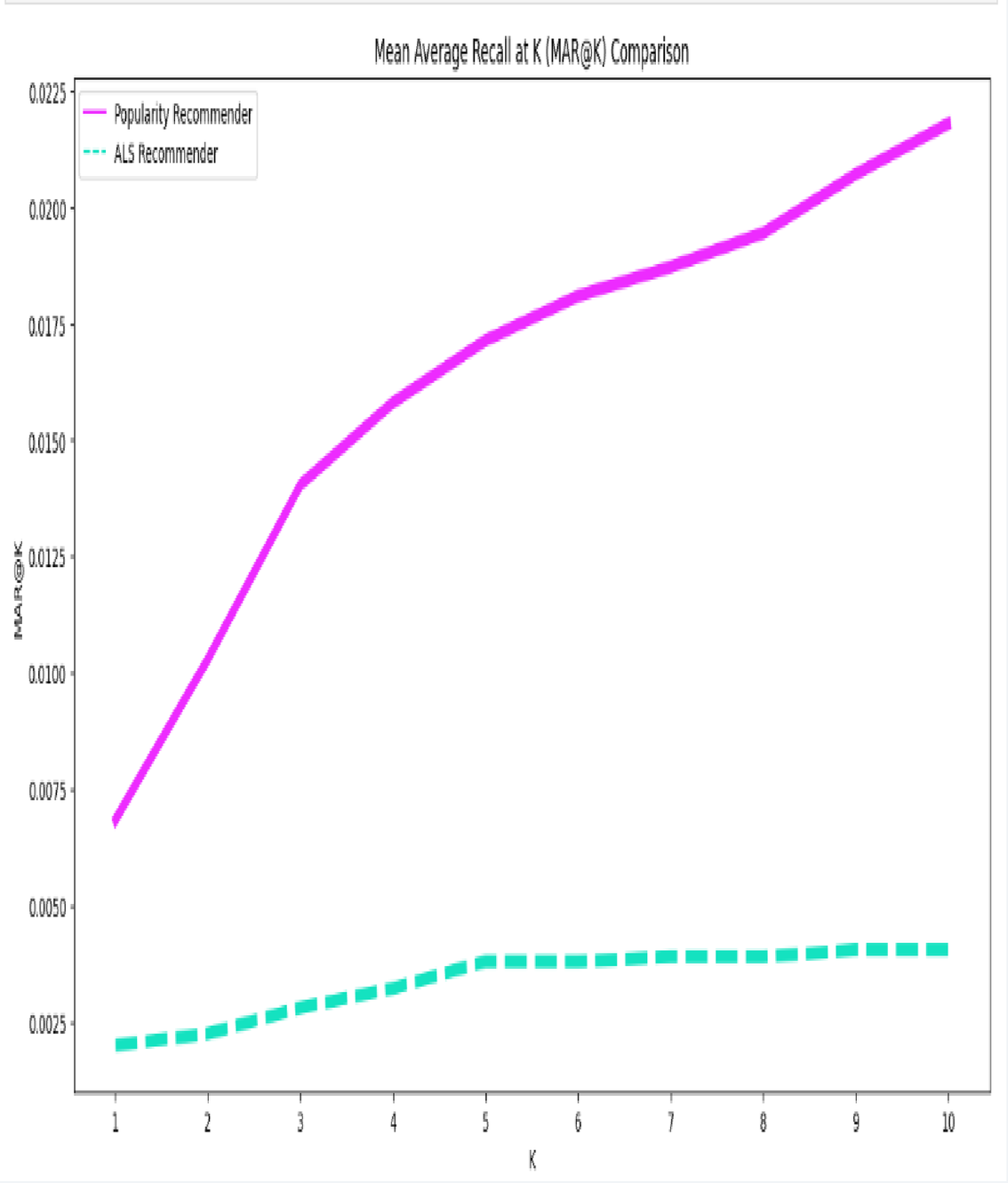
- This metric assesses the quality of ranked recommendation by measuring how well relevant items are placed within top positions
- ALS model has lower MAR@K reflecting personalized recommendation while the popular has high MAR@K meaning it excels in promoting popular items
- This metric suffers from popularity bias

Use Case:

- This metric is used in movies and music recommendation system to ensure relevant items are prominently featured

Wondershare PDFelement

Results:+16.7% vs



Wondershare PDFelement

further improvement

Data

- More data
- Better data

...

Model

- Hybrid Model
- Deep Learning Model

...

Other

- Better user interface
- Feedback from users

...

Wondershare PDFelement

Conclusion

Key takeaways:

- High coverage and personalization from our model results in user retention and engagement

- Use more data for better personalization

- Retrain the model with ratings from users

Wondershare PDFelement

Wondershare
PDFelement

Thank You!

Wondershare
PDFelement