Movie Recommendation Systems

Mod 4 Project by Jahir Miah and James Moulds

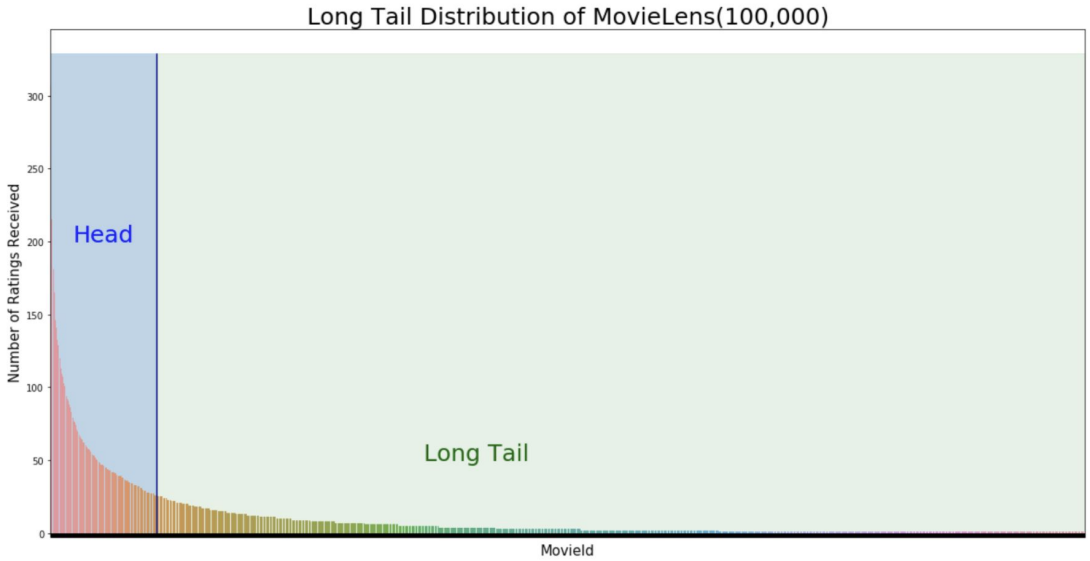
Business Problem

We've been asked to provide a recommendation system to a new player in the movie streaming space. They would like to be able to recommend movies in order to retain their customer base and attract new users. We have split the problem into three streams.

1- Use the existing service content information to improve recommendations to new or potential clients.

2. Use existing service clientele information to improve recommendations to other, similar, existing clients,

3. Use external data to understand the kind of movies that the service should consider licensing.



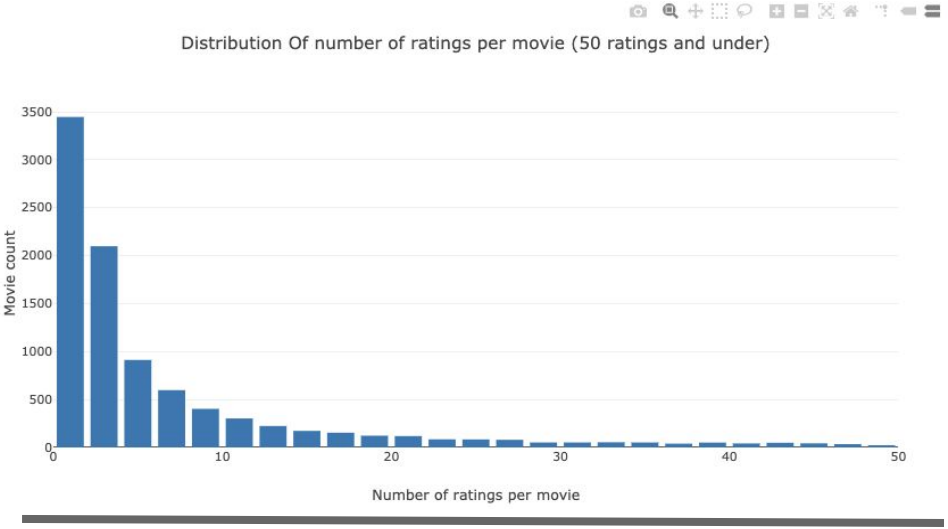
Data Insight

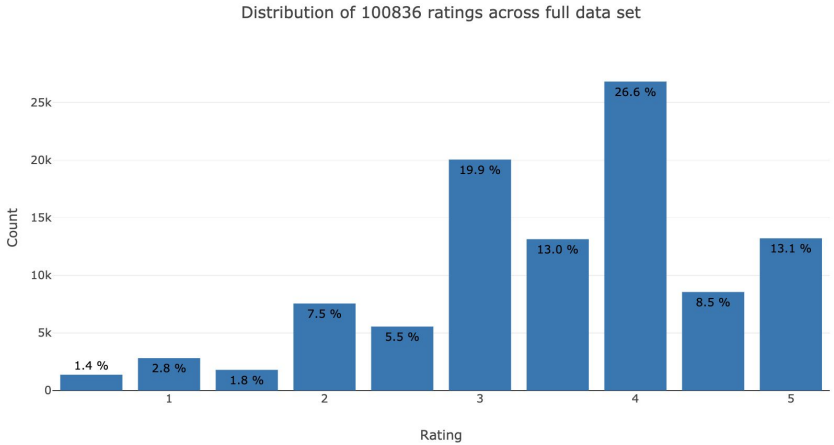
In our data set 7455 out of 9724 titles have under 10 ratings!

That's 80% of the titles being responsible for around 20% of the ratings.

Conversely the other 20% of the titles hold the 80% of the ratings. .

When recommending we have attempted to take into account this bias to ensure a mix of high and low popularity titles are selected from all sections of the dataset.





Alternative methods

Our current models have used the full data set with no filters to get a baseline product into production.

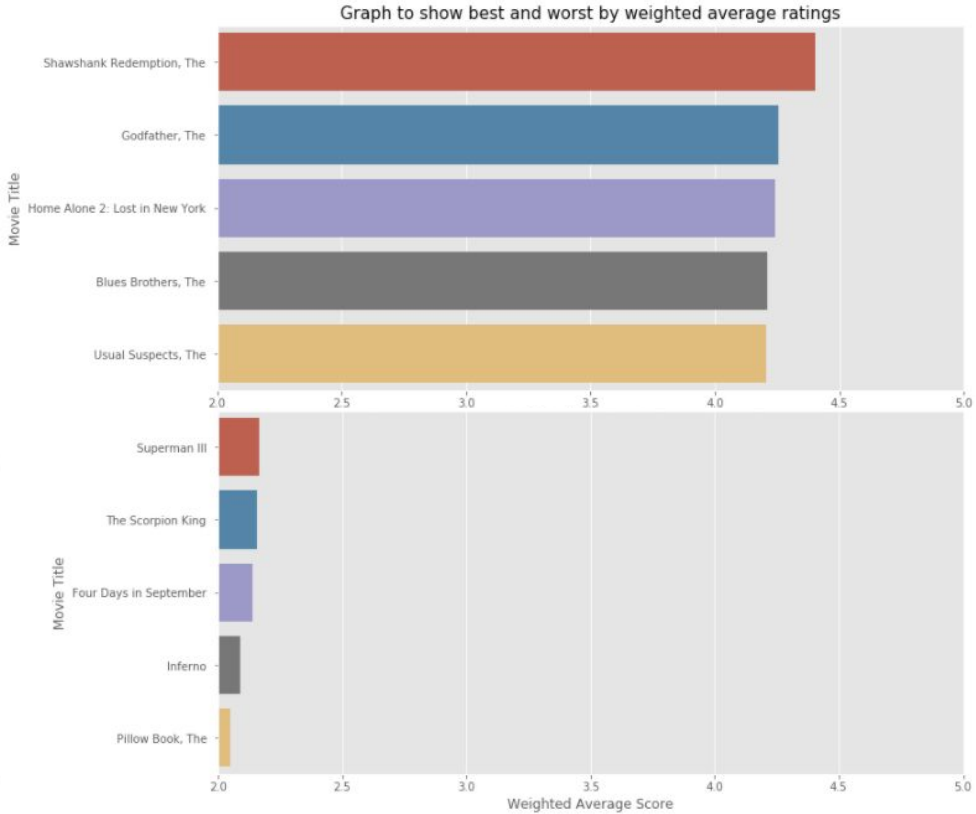
In future we would like to investigate weighted movie scores (as used by IMDB) to get a better hold on popularity and also import full movie synopses to inform our content recommenders.

$$W = \frac{Rv + Cm}{v + m}$$

where:

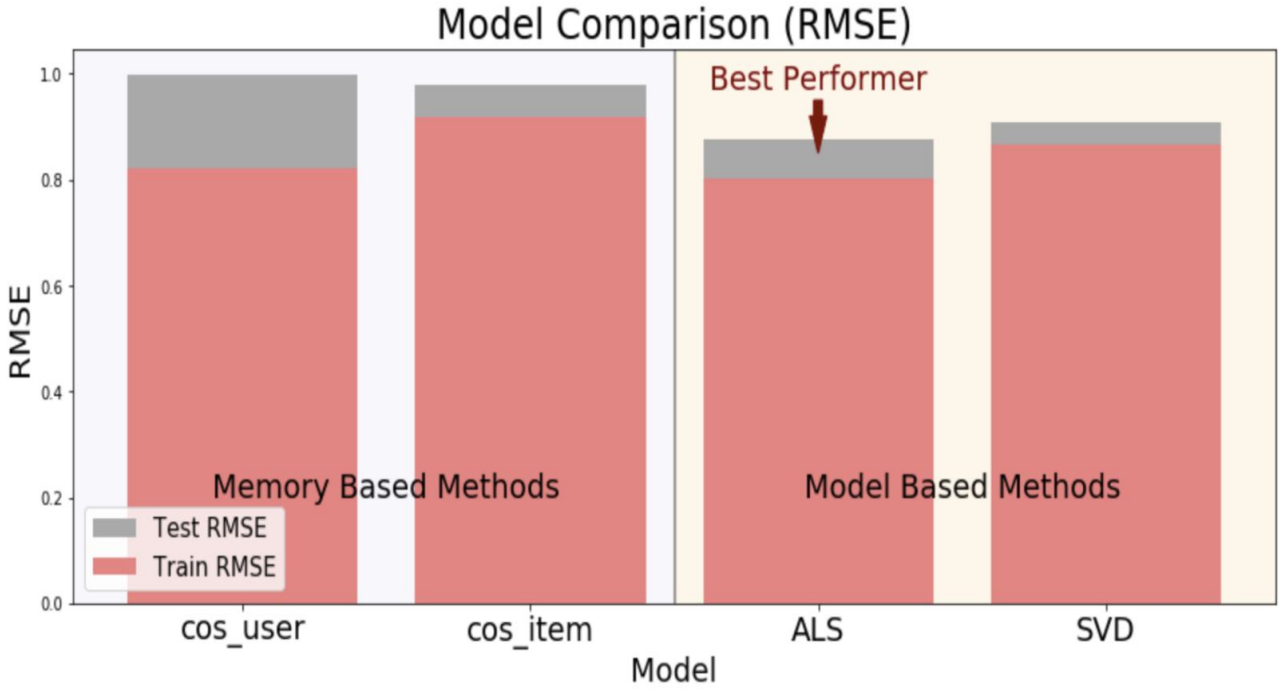
W = Weighted Rating
R = average for the movie as a number from 0 to 10 (mean) = (Rating)
v = number of votes for the movie = (votes)
m = minimum votes required to be listed in the Top 250 (currently 3000)
C = the mean vote across the whole report (currently 6.9)

The current rating system if IMDb.



Our System

When we tested our preferred model we found that our predictions are currently within 1 rating score against known within ratings.



Demonstration

A Guided walk through of the current recommendation product