Why is popularity important in recommendation engines?

What is popularity of Recommendation Engines

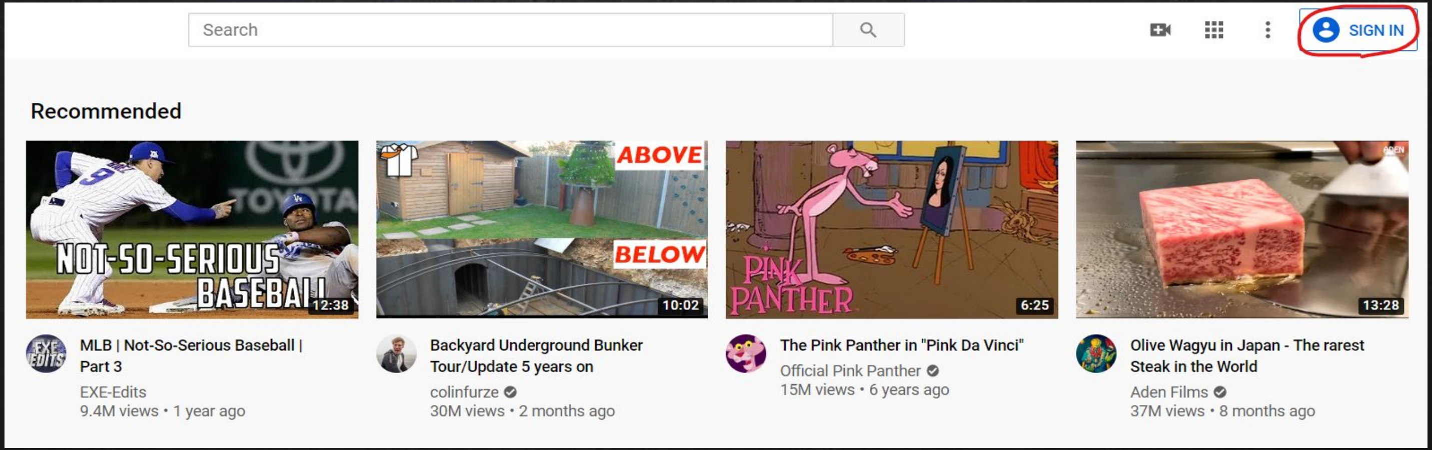
As our mind, popularity in recommendation system means that the rank of each features of products in the dataset when we analysis a new customer behavior based on our dataset. According to the rank, we could recommend corresponding product to our customer, so popularity is important in recommendation engines.

Recommendations by popularity

- Pros :
  - Show trend by time, location, culture…etc (Explicit data).
  - Easy to apply when you have no detail data (Implicit data).
  - High probability to recommend good item since it is liked by most of people.

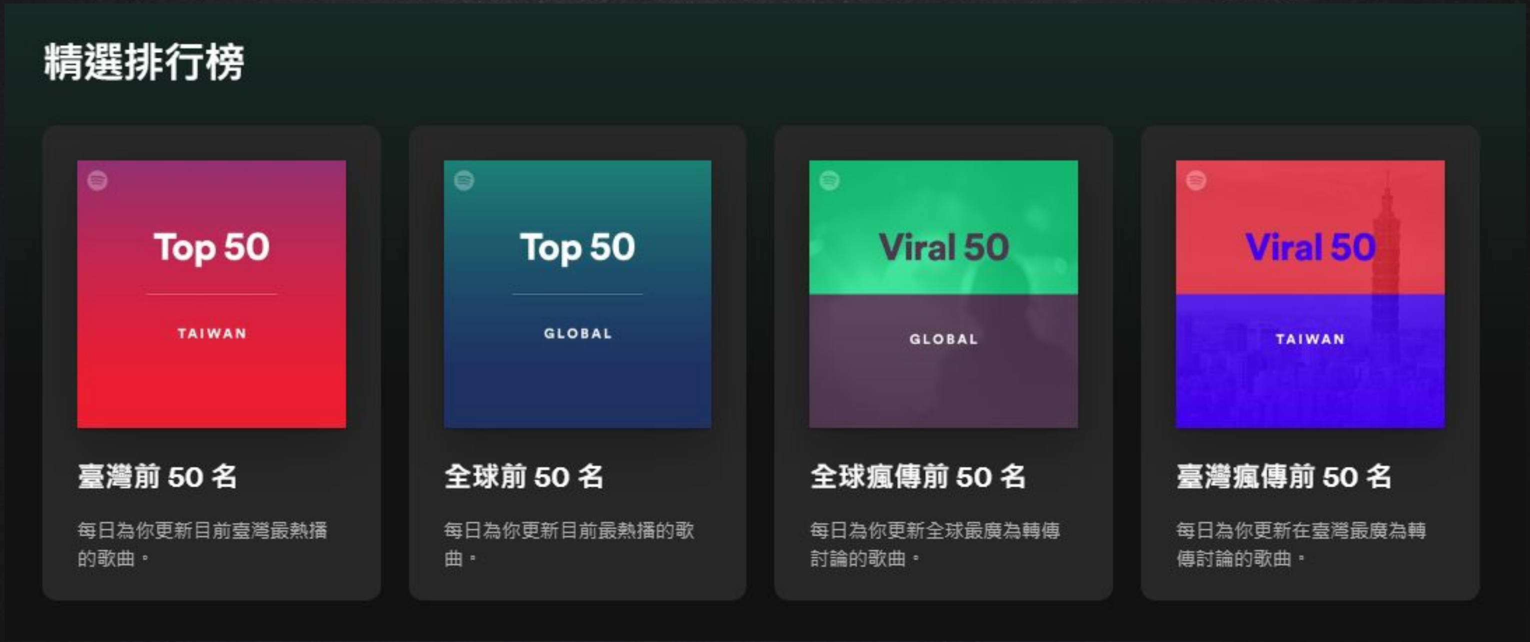
- Cons
  - Recommendations might be same or similar for all users.

Visitor Cold Start



- When a new user added in, and we have no historical data of new user. It's called **Visitor Cold Start.**
- Recommend the most popular items first (could be recently overall or regionally) . After users show their preference, following recommendation will be easier.
- A good first step!

Hot playlist



- ✘ Sometimes user/service provider want to know the trend
- ✘ For people who have no particular preference
- ✘ Trend observing

Evaluation Metrics for Recommender Systems

Ranking Metrics

Precision

$$\text{Precision} = \frac{tp}{tp + fp}$$

tp **represents the number of items recommended to a user that he/she likes**

tp+fp **represents the total items that a user like**

Recall

$$\text{Recall} = \frac{tp}{tp + fn}$$

tp+fn **represents the total items recommended to a user**

Ranking Metrics

Mean Averaged Precision @ K

- ✘ Averaged Precision @ K from 1 to K

$$MAP_i = \frac{1}{|R_i|} \sum_{k=1}^{|R_i|} P(R_i[k])$$

Mean Averaged Recall @ K

- ✘ Averaged Recall @ K from 1 to K

Ranking metrics

Mean reciprocal rank (MRR)

$$\mathrm{MRR} = \frac{1}{\mathrm{n}} \cdot \sum_{i=1}^{\mathrm{n}} \frac{1}{\mathrm{r}(\mathrm{Q}_i)}$$

Larger the mean reciprocal rank, better the recommendations

Ranking Metrics

Normalized Discounted Cumulative Gain(NDCG @ k)

DCG: $\sum_{j=0}^{n-1} \frac{rel_{D_i}(R_i(j))}{\ln(j+1)}$ $\quad rel_D(r) = \begin{cases} 1 & \text{if } r \in D, \\ 0 & \text{otherwise.} \end{cases}$

n: min(max(|R|, |D|), k)
|R|: length of recommended items
|D|: length of ground truth

Ideal DCG: $\sum_{j=0}^{\min(|D|,k)-1} \frac{1}{\ln(j+1)}$

NDCG : DCG/Ideal DCG, for each user

Prediction accuracy metrics

RMSE

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$