EXPLOITING THE EVALUATION FREQUENCY OF THE ITEMS TO ENHANCE THE RECOMMENDATION ACCURACY

Roberto Saia (roberto.saia@unica.it)
Ludovico Boratto (ludovico.boratto@unica.it)
Salvatore Carta (salvatore@unica.it)

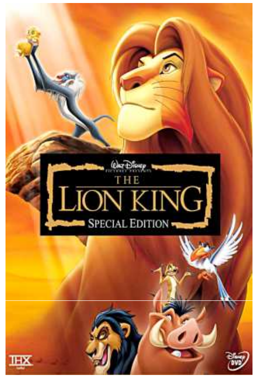
International *Conference* on Computer Applications & Technology
Rome, Italy – November 5-7, 2015

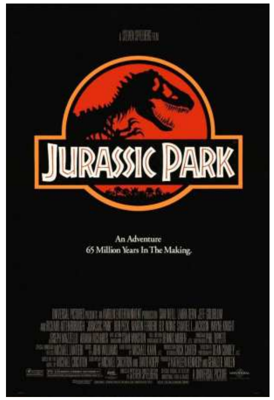
Recommender Systems

- Objective: suggest items that might be interesting for a user
- Analyze the items a user previously evaluated or considered

Popularity in Recommender Systems

- Items' popularity is an aspect widely studied in the literature
- Good to identify items of potential interest to the users
- …but too trivial to produce interesting recommendations
- E.g., the *non-personalized model* [Cremonesi et al.] recommends a fixed list of popular items to all the users





Semantics-aware Recommender Systems

- Semantic Web is generating much interest nowadays
- Employment of ontologies and semantic analysis tools in recommender systems too
  - Semantics-aware recommendation
  - Consider the semantics behind the item descriptions
- <u>…but they are not the most effective recommendation approaches</u>

Review: “I love Nike Air!”

Effective disambiguation between the brand and the goddess

NIKE

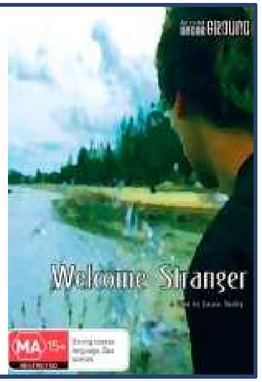
Intuition

- Popularity and semantics can provide useful information
  - Not in the filtering/prediction process
  - But in the decision making process
- Idea:
  - Use more advanced and state-of-the-art approaches to predict
  - Re-rank the results by considering item popularity and semantic correlation with the user profile (i.e., in the following example, by swapping the movies)

1

Welcome Stranger

Predicted rating: 4.9/5



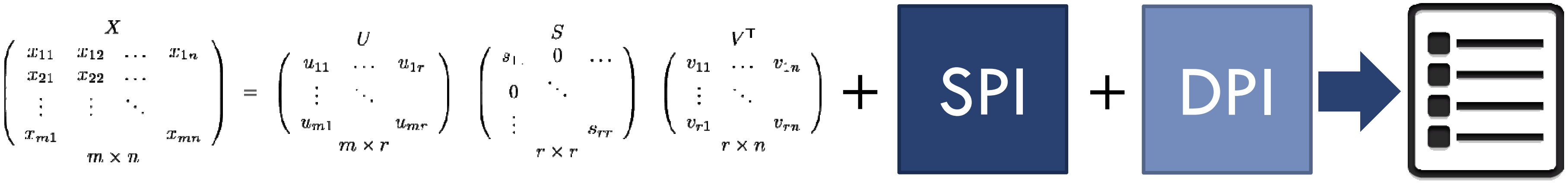
2

Titanic

Predicted rating: 4.9/5

Our proposal

- Novel recommendation approach
  1. Build the predictions with the state of the art in recommender systems (SVD++)
  2. Calculate the semantic similarity between the item and the user profile (*Semantic Popularity Index*, SPI)
  3. Calculate the frequency of evaluation of an item (*Domain Popularity Index*, DPI)
  4. Recommend considering the predicted rating, SPI, and DPI



Scientific contributions

- The concept of semantic similarity of an item w.r.t. to a user profile is novel
- The employment of the popularity as a form of post-processing is novel
- Novel recommendation approach (*Popularity-Based SVD*++, PBSVD++)
- Evaluation on 3 real-world datasets
  - Our proposal outperforms the state-of-the-art recommendation approach (SDV++)

Our proposal

Semantic Popularity Index (SPI)

- For each item *i* we generate a semantic description sd_i
  - sd_i is a set of semantic objects (synsets) in the Wordnet ontology
  - Each synset contains a semantic disambiguation of the word in the description
- For each user *u* we consider the union of the sd_i of her/his items ($sd_{I,u}$)

$$SPI = \frac{|sd_i \cap sd_{I,u}|}{|sd_i|}$$

Domain Popularity Index (DPI)

- $np_{i,U}$ is the number of positive preferences expressed by all users *U* for an item *i*
- DPI(i,U) represents the frequency of evaluation of the item w.r.t. the others

$$DPI(i,U) = \frac{np_{i,U}}{\sum_{\forall j \in I} np_{j,U}}$$

PBSVD++ Algorithm

- Let $rating_{i,u}$ be a SVD++ prediction and R_u the list of recommended items
- We normalize it in a [0,1] range:

$$STD(i,u) = \frac{rating_{i,u}}{\sum_{\forall j \in R_u} rating_{j,u}}$$

- The final rating is:

$$\rho_{i,u} = STD(i,u) + \left(\frac{SPI(i,u)}{\sum_{\forall j \in R_u} SPI(j,u)} \cdot \frac{DPI(i,U)}{\sum_{\forall j \in R_u} DPI(j,U)} \right)$$

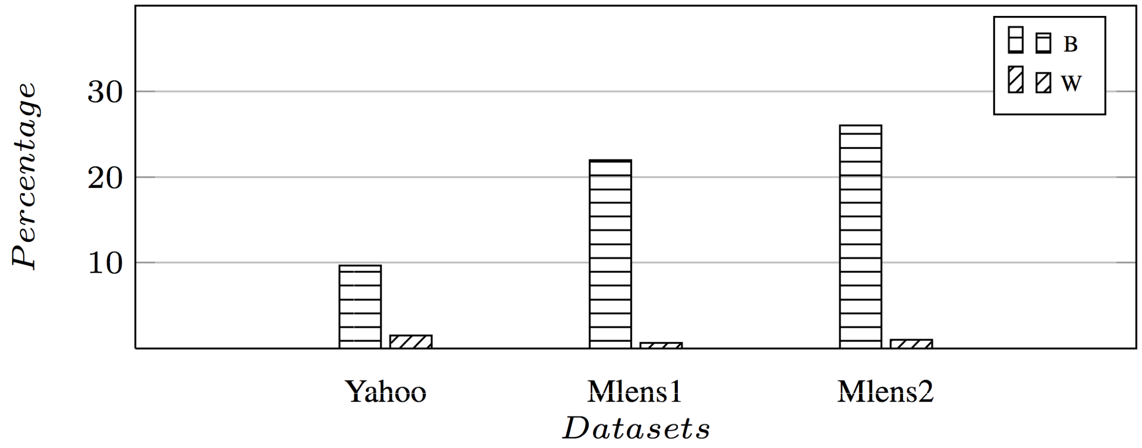
Experiments

Experiments

- We validated our proposal with two experiments:
  1. Percentage of times we out-/under-performed SVD++
  2. Accuracy improvement w.r.t. to SVD++
- Three datasets:
  - Two extracted from Movielens 10M
  - One extracted from Yahoo! Webscope (R4)
- Measure
  - *F1@n* (harmonic mean of Precision and Recall when *n* recommendations are generated)
  - We will present the difference of the F1 obtained by SVD++ and PBSVD++ (*F1-variation@n*)

Experiments

Results overview

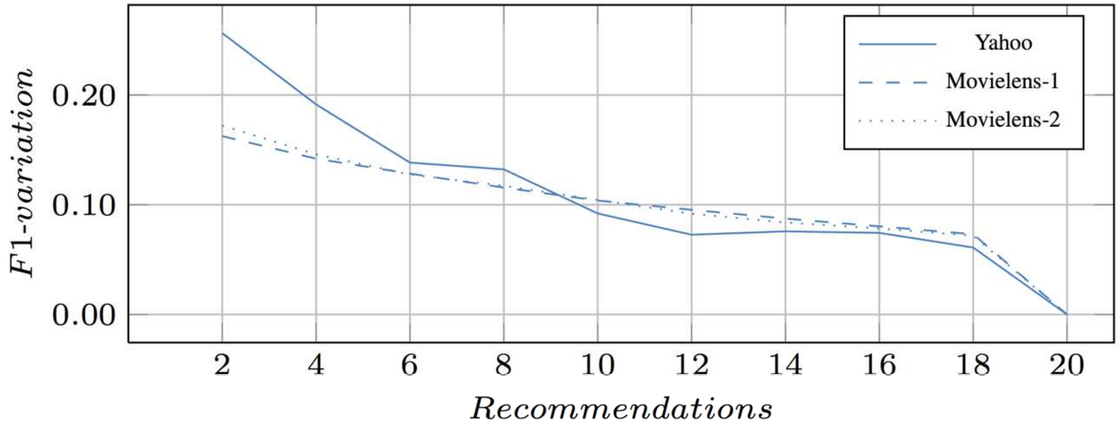


(a) General performance

- With all the three datasets, between 10 and 30% of the times our approach outperforms the state-of-the-art
- The additional information we employ for the re-ranking leads to significant improvements

Experiments

Improvement w.r.t. SDV++



(b) F1-Measure@n

- In all the three datasets there is a strong improvement (almost 30% for Yahoo and 2 recommended items)
- No improvement with the maximum number of recommendations (i.e., 20)
  - No matter how you re-rank them, the recommended items are the same

Conclusions and future work

Conclusions

- Semantics and popularity have proven to be useful sources of information in recommendation
  - But do not lead to accurate suggestions
- We proposed an algorithm (PBSVD++) that
  1. employs the state of the art in recommendation (SVD++)
  2. integrates two semantic and popularity indexes to modify the rating and re-rank the results
- Experimental results show the capability of our approach to outperform SDV++

Future work

- When an item belongs to more domains (e.g., movies with multiple genres) evaluate the popularity in each domain
- Introduction of other popularity metrics
  - Geographic
  - Demographic
  - …

Bibliography

[**Cremonesi et al.**] Cremonesi, Y. Koren, and R. Turrin, “Performance of recommender algorithms on top-n recommendation tasks,” in *Proceedings of the 2010 ACM Conference on Recommender Systems.*