Predicting the Performance of Collaborative Filtering Algorithms

Pawel Matuszyk and Myra Spiliopoulou

Knowledge Management and Discovery
Otto-von-Guericke University
Magdeburg, Germany

04. June 2014

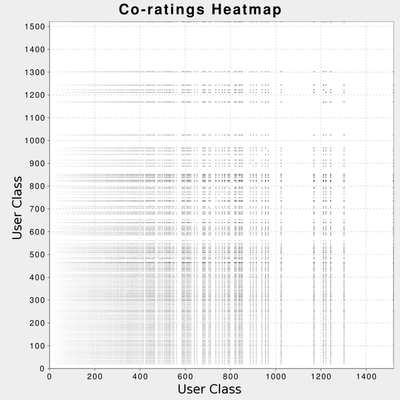
Motivation

- CF is a widely used family of algorithms for recommender systems
  - e.g. matrix factorization
  - neighbourhood-based methods
- not appropriate for all applications
- how do we know, if CF is applicable?
  - implementing a CF method $\rightarrow$ running expensive experiments $\rightarrow$ tuning $\rightarrow$ evaluation
    OR
  - **predicting the performance of CF** given a dataset

Visual Analysis

- a method for a visual assessment of dataset characteristics
- mapping of users into equivalence classes
- $[u] = \{u_x \in U || R(u_x)| = |R(u)|\}$
- $U =$ a set of users
- $R(u_x) =$ set of ratings of user u_x
- building of a co-rating matrix
  - one cell $=$ average number of co-ratings between $([u_x], [u_y])$
  - a heatmap as visualization
- histogram of cardinalities of user classes

MovieLens 1M Dataset



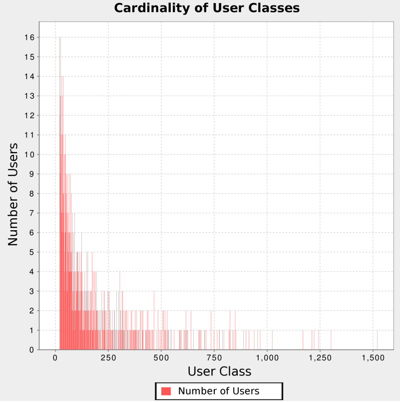
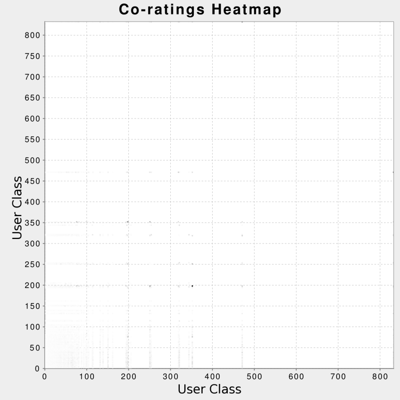


Figure: Visualization of the Movie Lens 1M dataset.

Epinions Dataset



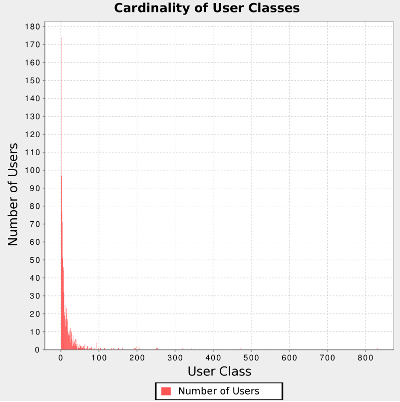


Figure: Visualization of the Epinions dataset.

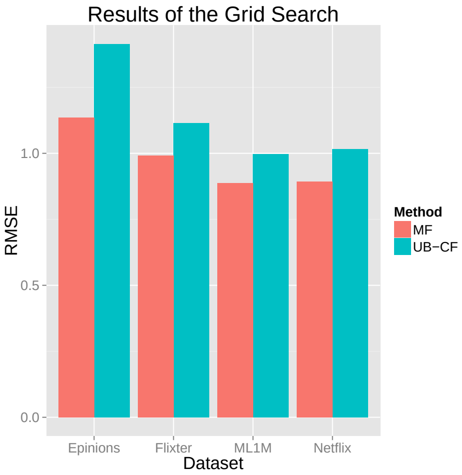
Characteristics of Datasets

- measures for quantifying the characteristics of dataset $\mathcal{D}$
- $sparsity(\mathcal{D}) = 1 - \frac{\sum_{u \in U} |R(u)|}{|U| \cdot |Items|}$ [AB11]
- quantifying the distribution of co-ratings (high values better)
- $Entropy(\mathcal{D}) = -\sum_{x,y} \frac{cor([u_x],[u_y])}{\sum_{x,y} cor([u_x],[u_y])} log_2 \left(\frac{cor([u_x],[u_y])}{\sum_{x,y} cor([u_x],[u_y])} \right)$
- $GiniIndex(\mathcal{D}) = 1 - \sum_{x,y} \left(\frac{cor([u_x],[u_y])}{\sum_{x,y} cor([u_x],[u_y])} \right)^2$

Building a Training Dataset

- mapping of the dataset measures to the RMSE values
- for each dataset $\mathcal{D}$: $\mathit{sparsity}(\mathcal{D}), \mathit{Entropy}(\mathcal{D}), \mathit{Gini}(\mathcal{D})$
- two CF methods:
  - user-based CF with cosine similarity
  - SVD++ [Kor08]
- two target attributes: $\{\mathit{RMSE}_{UB-CF}, \mathit{RMSE}_{MF}\}$
- target value = best value of RMSE found by a grid search
- 4 datasets $\rightarrow$ 4 learning instances

Results of the Grid Search



Correlation of our measures with RMSE

Table: Pearson's product moment correlation coefficients between our measures and the RMSE. Significance at level < 0.03 marked in red.

Measure	Correlation with RMSE		Alternative	p-value	
	UB-CF	MF	Hypothesis	UB-CF	MF
1-Gini	0.9969276652	0.9760339396	true correl > 0	0.001536	0.01198
Entropy	-0.9488311629	-0.9849657734	true correl < 0	0.02558	0.007517
Sparsity	0.737570808	0.7795095148	true correl > 0	0.1312	0.1102
(1-Gini) · Sparsity	**0.9969300691**	0.9758969096	true correl > 0	0.001535	0.01205
Entropy · Sparsity	-0.9409525839	-0.9733224195	true correl < 0	0.02952	0.01334

- strong linear correlation of our measures with the RMSE
- correlation based on only 4 instances, but p-values prove significance

CF-Performance Predictor

- linear regression (LR) as predictor of RMSE
- $RMSE = \alpha \cdot (1 - Gini) + \beta \cdot Sparsity + \gamma$
- evaluation: learn LR on $RMSE_{UB-CF}$ and use it to predict $RMSE_{MF}$ and vice versa
- evaluation measure: Pearson's product moment correlation between predictions and real values
- learnt parameters:

Method	α	β	γ
UB-CF	1158.0332	0.925	0.1004
MF	621.7	1.13	-0.2

Experimental Results

Regression on	UB-CF		MF	
	Corr.	p-value	Corr.	p-value
UB-CF	0.99967	0.000167	0.98455	0.00773
MF	0.99649	0.00175	0.98768	0.00616

Table: Pearson's product moment correlation coefficients of RMSE predictions with real values.

Conclusions

- based only on dataset statistics we built a performance predictor
- results highly and **significantly** correlate with real values
- alternative to implementing and running expensive experiments
- limitation: we tested the method on datasets with rating range between 1 and 5
- rating behaviour with a different rating rang is different (our future work)

References

[AB11] Deepa Anand and Kamal Bharadwaj.
Utilizing various sparsity measures for enhancing accuracy of collaborative recommender systems based on local and global similarities.
Expert Syst. Appl., 38(5):5101–5109, 2011.

[Kor08] Y. Koren.
Factorization meets the neighborhood: a multifaceted collaborative filtering model.
In *14th ACM SIGKDD*, pages 426–434. ACM, 2008.