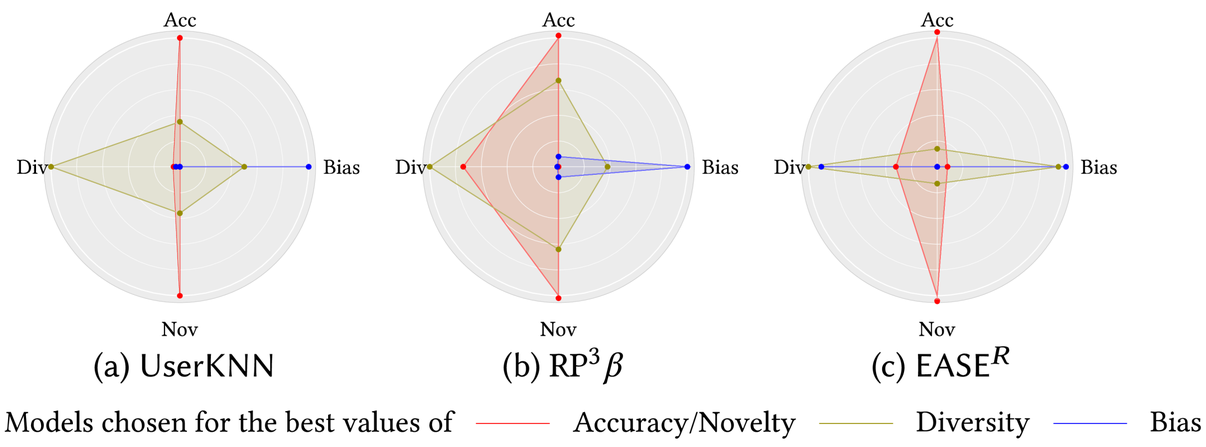
Broadening the Scope: Evaluating the Potential of Recommender Systems Beyond Prioritizing Accuracy

Vincenzo Paparella, Dario Di Palma, Vito Walter Anelli, Tommaso Di Noia
*firstname.lastname@poliba.it

Contributions



(a) UserKNN (b) RP$^3\beta$ (c) EASE^R

Models chosen for the best values of — Accuracy/Novelty — Diversity — Bias

- Simultaneously assess the models' performance on Accuracy, Diversity, Novelty, and Algorithmic Bias.
- Assess the entire set of Pareto-optimal configurations of 5 models, i.e. the Pareto frontiers, by exploiting the **Quality Indicators** (QIs).

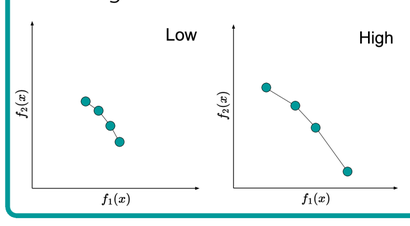
Questions

RQ1. How well do the models generate diverse Pareto-optimal configurations?

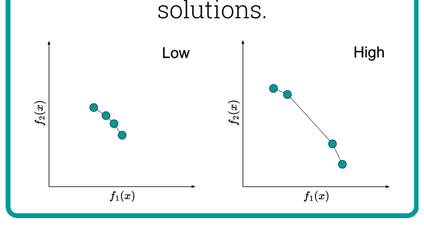
RQ2. Which model has the Pareto frontier that simultaneously offers better solutions on multiple metrics?

Quality Indicators

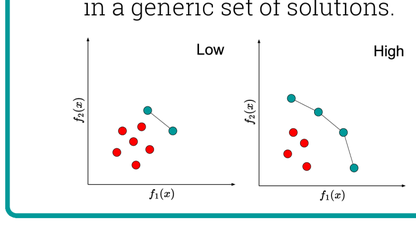
Maximum Spread (MS).
Range of a Pareto frontier.



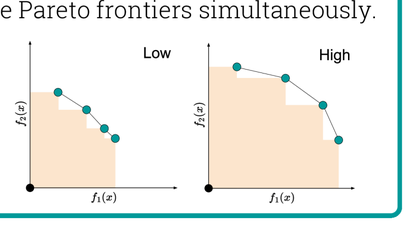
Spacing (SP).
Distribution of the Pareto-optimal solutions.



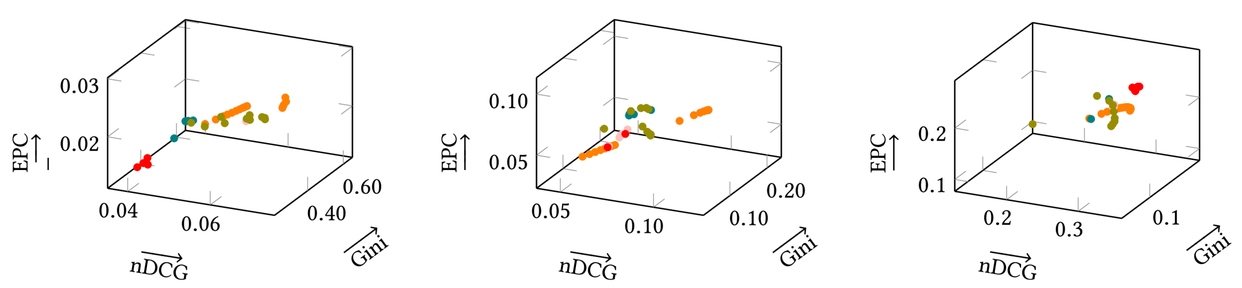
Error Ratio (ER).
Proportion of Pareto-optimal solutions in a generic set of solutions.



Hypervolume (HV).
Spread, uniformity, and cardinality of the Pareto frontiers simultaneously.



Results



(a) Amazon Music, nDCG/Gini/EPC (b) Goodreads, nDCG/Gini/EPC (c) Movielens1M, nDCG/Gini/EPC

(d) Amazon Music, nDCG/APLT (e) Goodreads, nDCG/APLT (f) Movielens1M, nDCG/APLT

- **UserKNN** provides diverse optimal solutions **balancing both scenarios**.
- **EASER** provides **many optimal solutions** but tends to cluster them.
- **RP3β balances accuracy and bias.**
- **LightGCN** and **MultiVAE** yield **inferior performance**. **RQ1.**

- **UserKNN** is the **superior** model **overall**.
- **RP3β** is a strong model when **balancing accuracy and bias**. **RQ2.**

Conclusions

Multi-objective evaluation through QIs **reshapes** the **performance ranking** of Recommender Systems:

- **EASER** was **outperformed** by other models.
- **USERKNN** demonstrated **superior performance** across diverse metrics.
- **RP3β** is effective in finding a **balance** between **nDGC and APLT** (bias) performance.

Read the Paper!

The 17th ACM Conference on Recommender Systems, Singapore, 18th-22nd September 2023