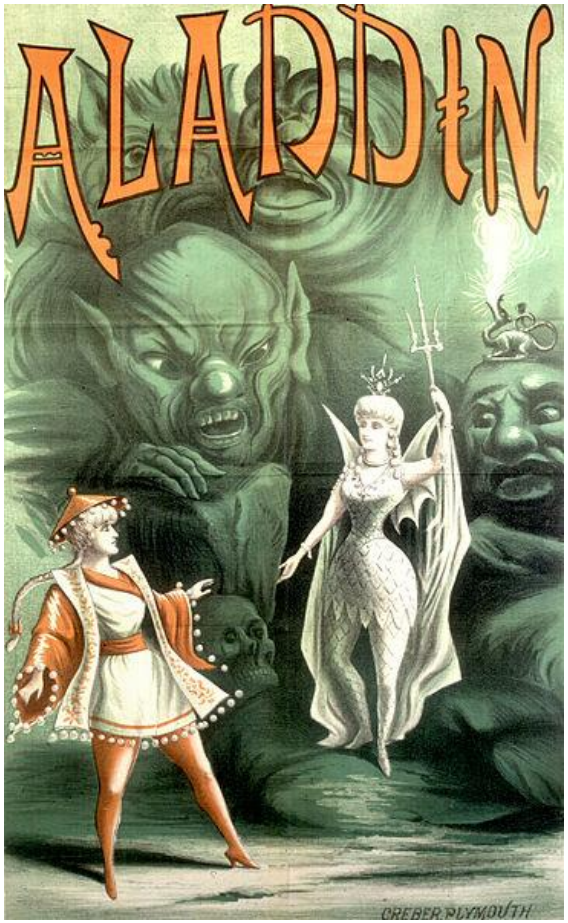


"An 1886 theatre poster advertising a production of the pantomime Aladdin" (Wikipedia), PD-US

ALADIN: A New Approach for Drug–Target Interaction Prediction

Krisztian Buza[a], Ladislav Peška[b]

[a] Knowledge Discovery and Machine Learning
Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

[b] Faculty of Mathematics and Physics
Charles University, Prague, Czech Republic

buza@cs.uni-bonn.de peska@ksi.mff.cuni.cz

Supplementary material: http://www.biointelligence.hu/dti

universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer IAIS

Outline

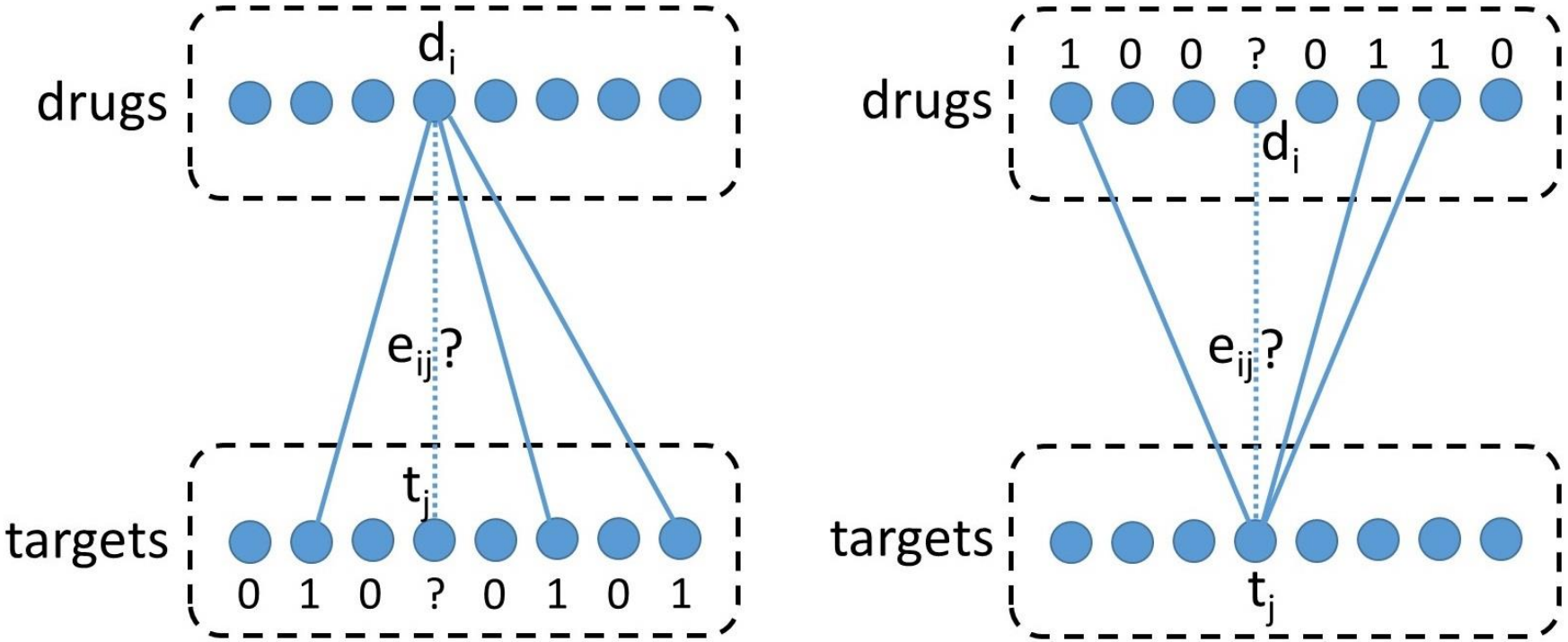
- Motivation
- Bipatite Local Models
- Our approach: Advanced Local Drug-Target Interaction Prediction (ALADIN)
- Experiments
- Outlook and Conclusion

Motivation

- Better understanding of the pharmacology of drugs
- Prediction of adverse effects
- Drug repurposing
  - use of an existing medicine to treat a disease that has not been treated with that drug yet
  - For example, *sildenafil* was designed to treat heart diseases, but it was not effective. However it turned out to be useful in case of erectile disorders → became known as *viagra*.
- drug discovery is expensive and needs long time
  (up to $1.8 billion, more than 10 years on average)

Morgan, S. et al.: The cost of drug development: a systematic review. Health Policy 100.1 (2011): 4-17.

Bipartite Local Models (BLM)



Bleakley, K., Yamanishi, Y.: Supervised prediction of drug–target interactions using bipartite local models. Bioinformatics 25(18), 2397–2403 (2009)

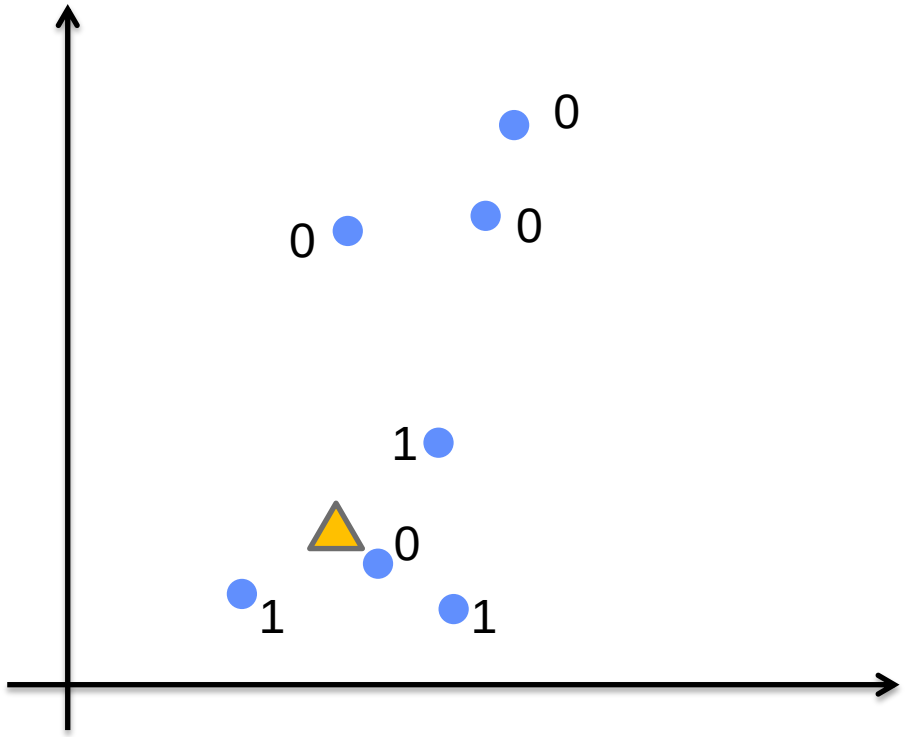
universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer
IAIS

Our approach: Advanced Local Drug–Target Interaction Prediction (ALADIN)

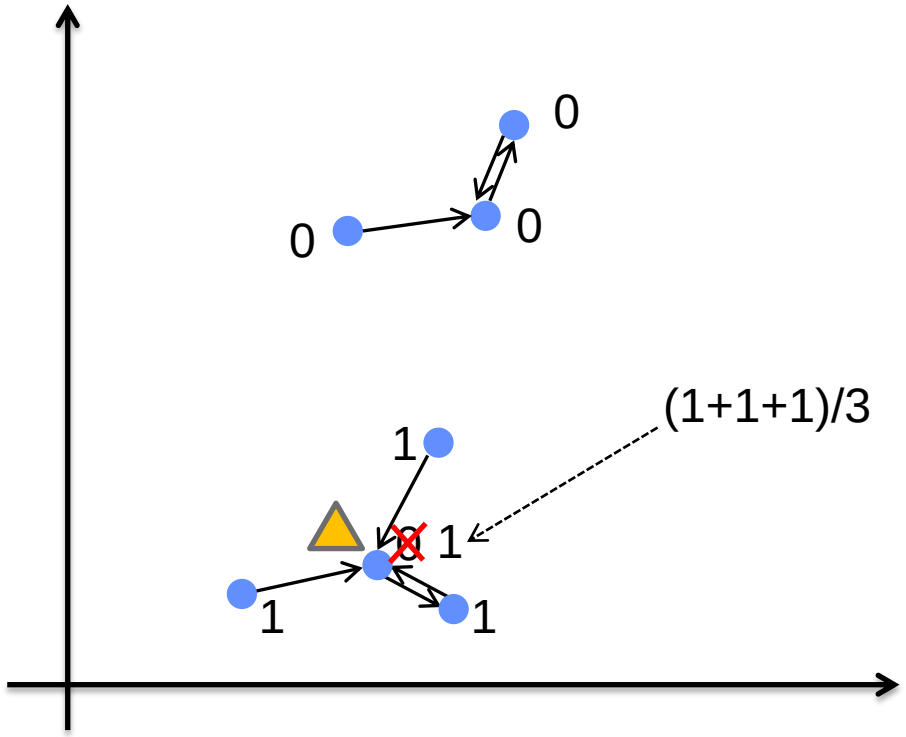
- Local model in BLM: EC*k*NN – a hubness-aware regressor
  - In case of “new” drugs/targets, BLM is inappropriate → use weighted profile
- Enhanced representation of drugs and targets in a multi-modal similarity space
- Projection-based ensemble

Local model: ECkNN – nearest neighbor regression with hubness-aware error correction (illustration with *k* = 1)



Buza, K., Nanopoulos, A., Nagy, G.: Nearest neighbor regression in the presence of bad hubs. Knowledge-Based Systems 86, 250–260 (2015)

Local model: ECkNN – nearest neighbor regression with hubness-aware error correction (illustration with *k* = 1)



Buza, K., Nanopoulos, A., Nagy, G.: Nearest neighbor regression in the presence of bad hubs. Knowledge-Based Systems 86, 250–260 (2015)

universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer IAIS

Enhanced similarity-based representation of drugs and targets

drug–target interaction matrix

	Target 1	Target 2	Target 3	Target 4	Target 5
Drug 1	1		1		
Drug 2			1		1
Drug 3				1	1
Drug 4		1		1	

Enhanced similarity-based representation of drugs

Drug 1	1	0.6	0.1	0.2	1	0.33	0	0
Drug 2	0.6	1	0.3	0.1	0.33	1	0.33	0
Drug 3	0.1	0.3	1	0.7	0	0.33	1	0.33
Drug 4	0.2	0.1	0.7	1	0	0	0.33	1

chemical similarities to all the drugs

Jaccard-similarities to all drugs (based on the interaction matrix)

Enhanced similarity-based representation of targets

Target 1	1	0.1	0.2	0.3	0.1	1	0	0.5	0	0
Target 2	0.1	1	0.2	0.5	0.1	0	1	0	0.5	0
Target 3	0.2	0.2	1	0.1	0.4	0.5	0	1	0	0.3
Target 4	0.3	0.5	0.1	1	0.6	0	0.5	0	1	0.3
Target 5	0.1	0.1	0.4	0.6	1	0	0	0.3	0.3	1

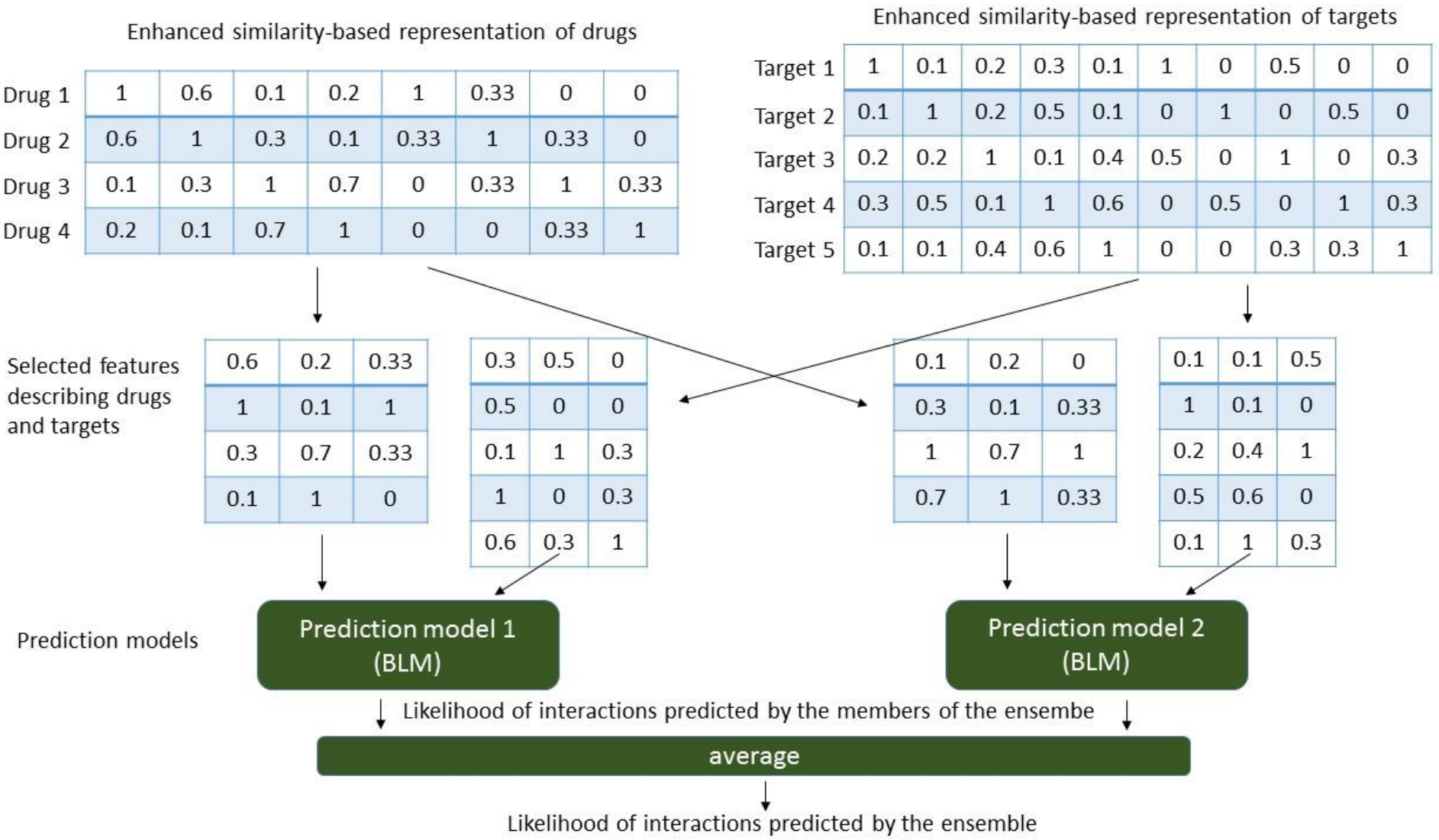
genomic similarities to all the targets

Jaccard-similarities to all targets (based on the interaction matrix)

universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer IAIS

Projection-based ensemble



universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer IAIS

Algorithm 1 Advanced Local Drug-Target Interaction Prediction (ALADIN)

Require: Drug–Target interaction matrix I, Drug–drug similarity matrix S^D, Target–target similarity matrix S^T, number of nearest neighbors k, ensemble size N, number of selected features F_D, F_T

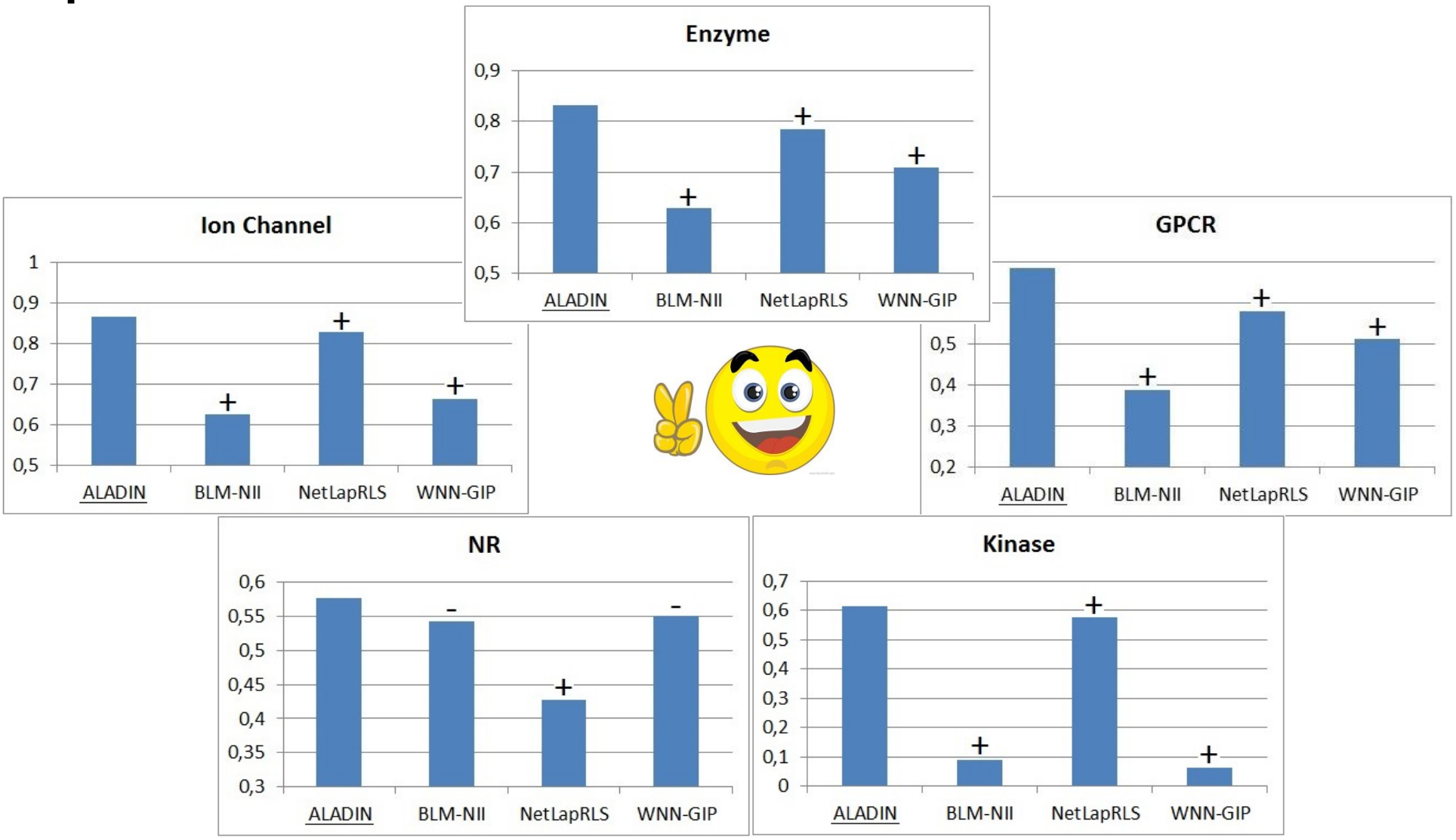
Ensure: Likelihood of drug–target interactions

1: $D \leftarrow$ enhanced similarity-based representations of drugs
2: $T \leftarrow$ enhanced similarity-based representations of targets
3: **for** $l = 1 \ldots N$ **do**
4: $D' \leftarrow$ random subset of D with F_D features
5: $T' \leftarrow$ random subset of T with F_T features
6: Predict interaction scores with BLM using ECkNN as local model and D' and T' as the representation of drugs and targets. (Use the weighted profile approach instead of BLM in case of new drugs/targets.)
7: **end for**
8: Average the predictions made in each execution of the loop

Experimental Settings

- Data: publicly available real-world drug-target interaction datasets: Enzyme, Ion Channel, G-protein coupled receptors (GPCR), Nuclear Receptors (NR), and Kinase
- Experimental protocol: 5x5 fold cross-validation
- Evaluation metrics:
  - Area under the ROC curve (AUC)
  - Area under Precision-Recall Curve (AUPR)
  - Statistical significance tests (t-test) at significance level of p=0.01
- Baselines:
  - BLM-NII: bipartite local models with „neighbor-based interaction-profile inferring“
  - NepLapRLS: „net Laplacian regularized least squares“
  - WNN-GIP: combination of weighted nearest neighbor and Gaussian interaction profile kernels
- Hyperparameters of ALADIN and the baselines were learned with grid search on the training data

Experimental Results



http://www.biointelligence.hu/dti

universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer
IAIS

Outlook: Recommender Systems for Drug–Target Interaction Prediction

ELSEVIER

Computer Methods and Programs in Biomedicine

Volume 152, December 2017, Pages 15–21

Drug-target interaction prediction: A Bayesian ranking approach

Ladislav Peska[a, b], Krisztian Buza[b, c], Júlia Koller[d]

Show more

https://doi.org/10.1016/j.cmpb.2017.09.003

Highlights

- Proposing BRDTI method for per-drug ranking of DTIs.
- Performed comparative evaluation of BRDTI w.r.t. AUC and per-drug nDCG.
- BRDTI achieved best average results on predicting new targets for existing drugs.

http://www.biointelligence.hu/dti

universitätbonn
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT

Fraunhofer IAIS

Conclusions

- Drug-target interaction prediction is one of the most prominent applications of machine learning in the pharmaceutical industry
- In our work, we extended bipartite local models (BLM) and showed that the resulting approach outperforms BLM and other drug-target interaction prediction techniques
- Prediction of drug-target interactions is related to those machine learning tasks that have been considered in the recommender systems community