metaMIMIC: an analysis of hyperparameter transferability for tabular data using MIMIC-IV database

Mateusz Grzyb[1], Zuzanna Trafas[2], Katarzyna Woźnica[1], Przemysław Biecek[1]

[1]Faculty of Mathematics and Information Science, Warsaw University of Technology
[2]Faculty of Computing and Telecommunications, Poznan University of Technology

Motivation

No Free Lunch Theorem [4] implies that for every machine learning problem, different ML algorithms maximise model performance. In hyperparameter transfer methods, it is assumed that we should use tuning histories for similar prediction problems to provide a targeted configuration. One approach to assess similarity between tabular datasets from **OpenML** [3] is to define a kind of distance between a priori defined statistical meta-features. However it is not clear which meta-feature definition would be most effective to apply for hyperparameter transfer.

- As an alternative, we propose extracting experience from experiments for **domain tasks** related to a specific domain and having a similar definition. Using the MIMIC-IV database [1], we create a **metaMIMIC - database of real-world medical prediction problems**.
- We define **consolidated learning** - scenarios of similarity between tasks reflecting real use-case when subsequent predictive models are often created for similar data from the same database and may also use similar variables.
- We test whether hyperparameter transfer for **metaMIMIC** domain tasks hits the optimal solution for a new prediction problem more effectively than transfer from **OpenML**.

Experiment scenarios

To create **metaMIMIC**, we determine the selection of patients cohort from MIMIC-IV and 58 features describing patient conditions. We define 12 binary classification tasks corresponding to the occurrence of each of the 12 selected types of diagnosis. Then we settle three scenarios of resemblance between tasks varying the selection of observations (the same and two independent samples of observations) and the available features (see Figure 1). We compare them with hyperparameter transfer based on tuning history for unrelated datasets from **OpenML** [3].

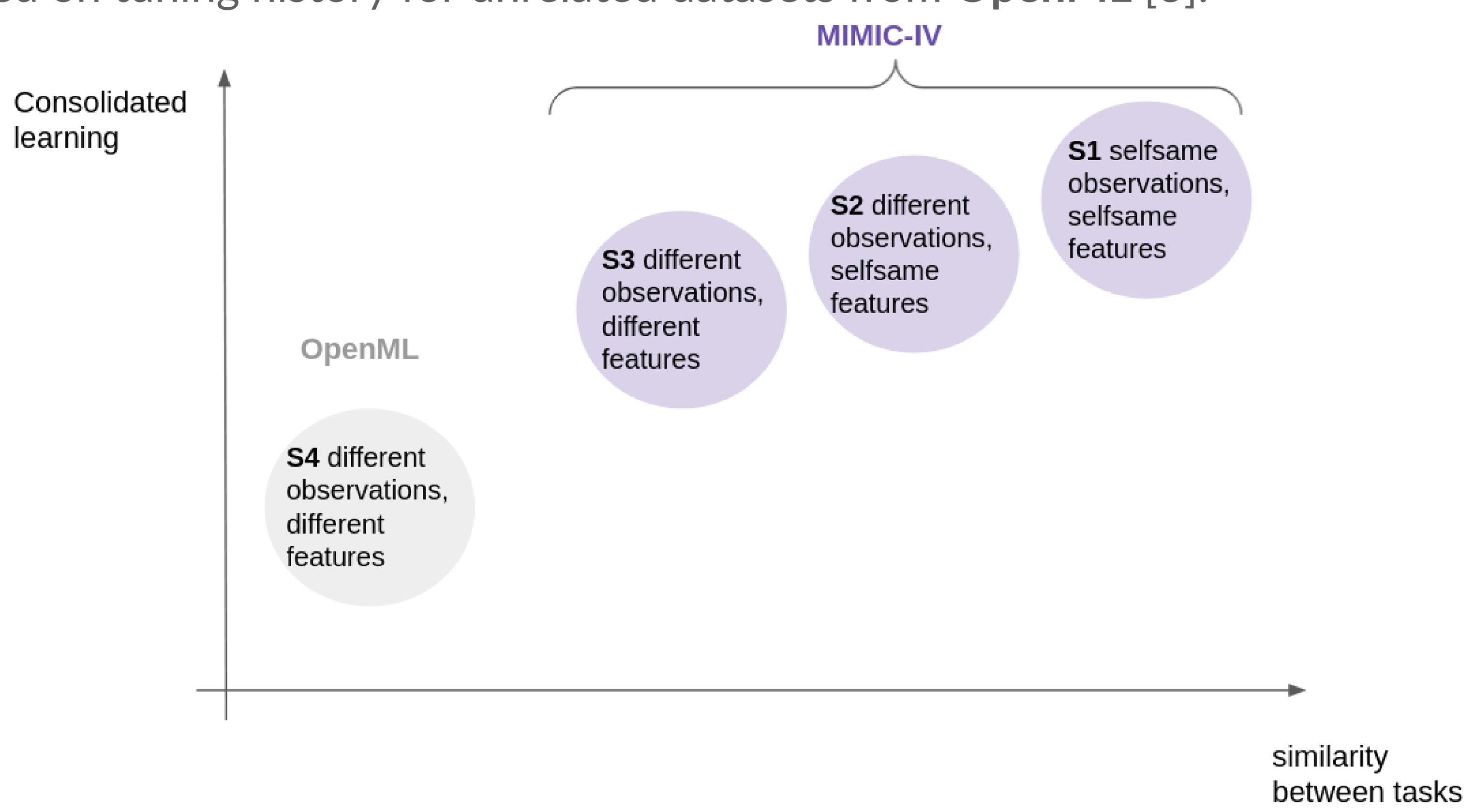

Figure 1: Three scenarios of similarity between tasks in **metaMIMIC** database. We suppose that the more similar tasks, the more effective hyperparameter transfer between them. **OpenML** scenario (S4) is baseline scenario, applied in previous meta-learning experiments.

Every scenario corresponds to real use-cases (consolidated learning): a data scientist working in a company builds machine learning models for multiple predictive tasks from the same database. Then he can work still on the same sample from data (S1) or get a new independent sample but once with the same set of explanatory variables (S2) or with the most important predictors relevant to the specific problem (S3).

Share of top settings

To verify whether prediction problems with similar definitions and common explanatory variables allow for more targeted tuning of hyperparameters we look at the consistency of the best hyperparameter configurations for each task pair (Figure 2. The comparison of the distributions of cover values shows that the number of shared best hyperparameter sets is significantly higher within **metaMIMIC** medical tasks than between **OpenML** and **metaMIMIC** problems.

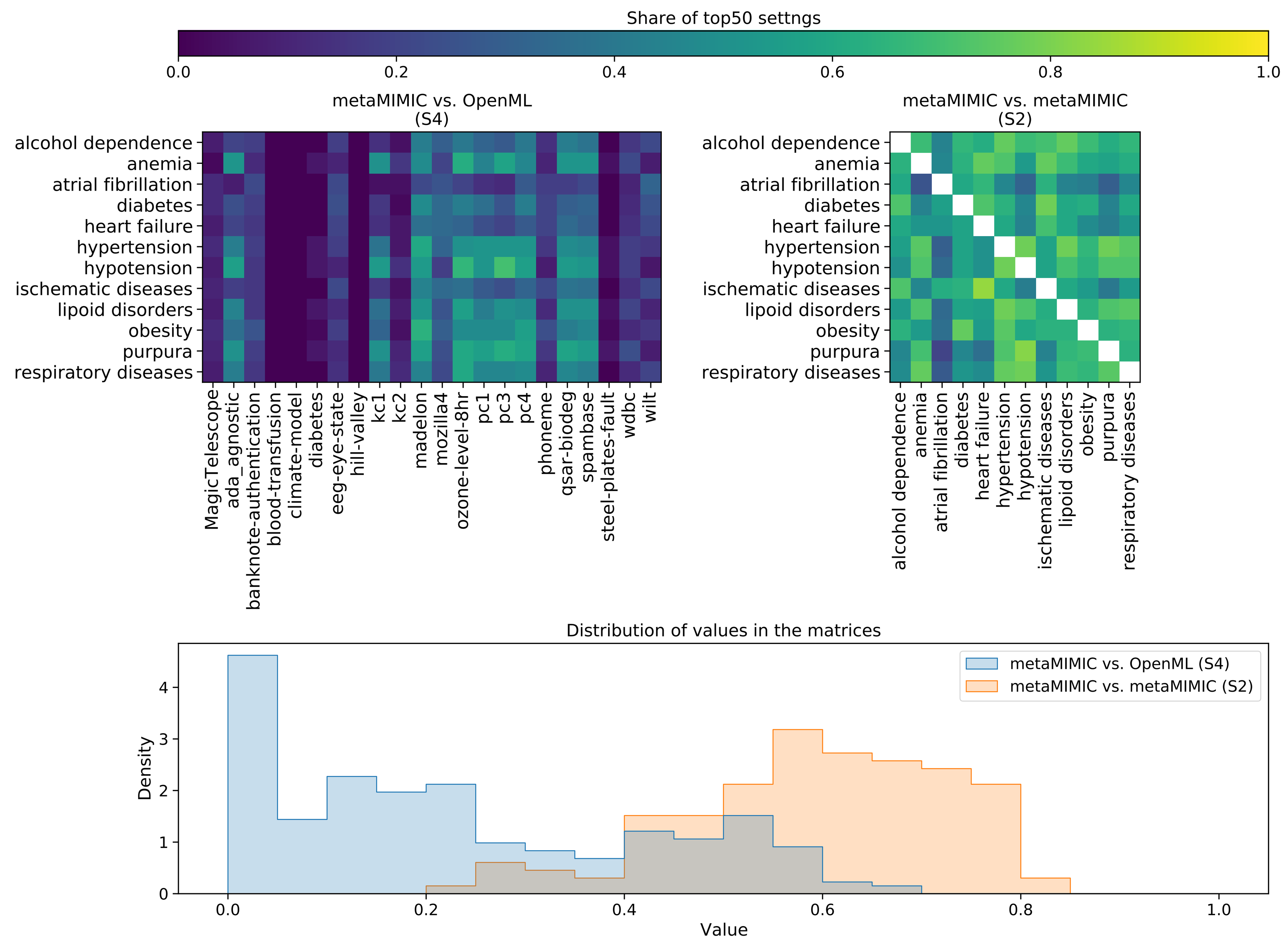

Figure 2: Cover of optimal hyperparameter configuration (regarding to the mean 4-CV ROC AUC measure) shared between tasks from **metaMIMIC** and **OpenML** (S4) and pairwise between tasks from **metaMIMIC** (S2). An individual cell of the matrix corresponds to the number of hyperparameter sets shared between a given pair of tasks. Histograms summarise the distribution of the values of each matrix. White colour on the diagonal means that the value is not considered.

When diversity between medical tasks increases because of different sets of features for available for every task in (S3), the cover of optimal hyperparmeters in **metaMIMIC** transfer decreases but it is still higher than in the transfer from the **OpenML**.

Transfered hyperparameters speed up tuning

For **metaMIMIC** tasks we perform simulations of hyperparameter tuning using random search, and Bayesian optimisation, as well as methods using hyperparameter ranking from tuning results for other tasks in S1, S2 and S4. Figure 3 shows hyperparameter tuning velocity (best performance obtained so far) of different methods and four benchmark tasks.
It can be seen that the methods involving hyperparameter transfer performs better than random search and Bayesian optimisation even when the searched ranking is based on tasks unrelated to the problem under consideration. Typically, however, a ranking based on related MIMIC-IV tasks provides a slightly higher speed relative to a ranking based on **OpenML**.

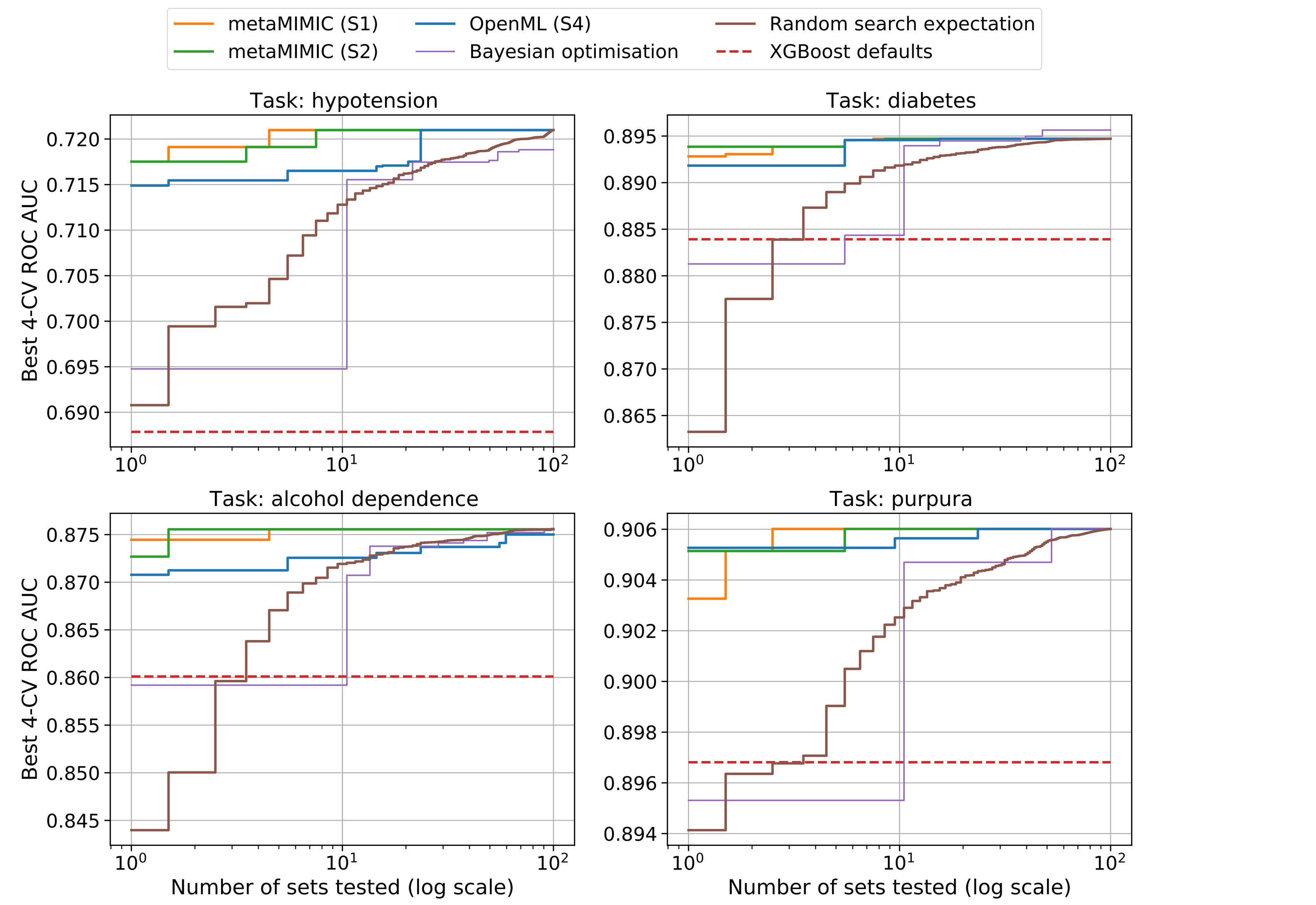

Figure 3: Hyperparameter tuning velocity of different methods and multiple benchmark tasks. Purpura is the only task for which **OpenML** initially significantly outperforms **metaMIMIC** among the 12 tasks considered.

Conclusions

- The presented results support hypothesis that the definition-based (domain) similarity of tasks is positively related to hyperparameters' transferability between them, both in terms of finding the best sets of hyperparameters and the speed of the optimisation performed.
- In this study, we do not assume any definitions of similarity based on statistical characteristics, only similarity based on tasks design. An open question remains whether we can define meta-features and similarity that capture these relationships or is data origin and domain knowledge necessary?
- To our knowledge, this is the first approach to creating a domain-based repository for meta-learning, in particular for medical data. Since machine learning for healthcare has been developing recently, this repository can be broadly used.
- For practitioners building multiple machine learning solutions for similar data, we advise to use the tuning history for new tasks.

- ✉ woznicakatarzyna22@gmail.com

- Preprint of article and codes to built **metaMIMIC**: https://github.com/ModelOriented/metaMIMIC

References

[1] A. Johnson, L. Bulgarelli, T. Pollard, S. Horng, L. Celi, and R. Mark. Mimic-iv (version 1.0). *PhysioNet*, 2020.

[2] W. Kretowicz and P. Biecek. MementoML: Performance of selected machine learning algorithm configurations on OpenML100 datasets, 2020.

[3] J. Vanschoren, J. N. Van Rijn, B. Bischl, and L. Torgo. OpenML: networked science in machine learning. *ACM SIGKDD Explorations Newsletter*, 15(2):49–60, 2014.

[4] D. H. Wolpert. The lack of a priori distinctions between learning algorithms. *Neural computation*, 8(7):1341–1390, 1996.