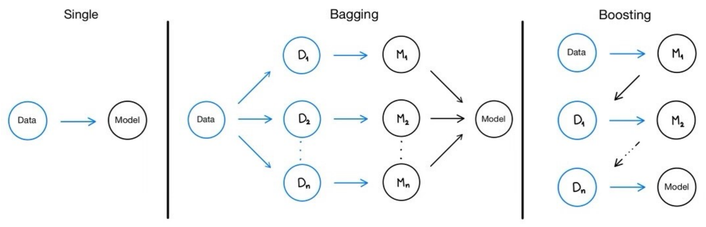
Analysis of machine learning models used in the process of building model ensembles for the regression task

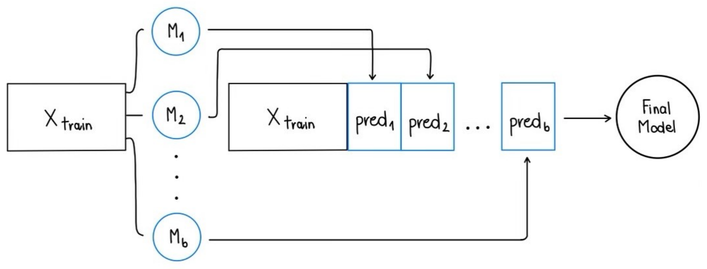
Jędrzej Ruciński, Anna Kozak
Warsaw University of Technology

Ensembles

- Ensemble models make predictions by aggregating the outputs of individual models.
- An ensemble model is the application of multiple models in order to obtain better performance.



Ensemble methods: bagging, boosting.



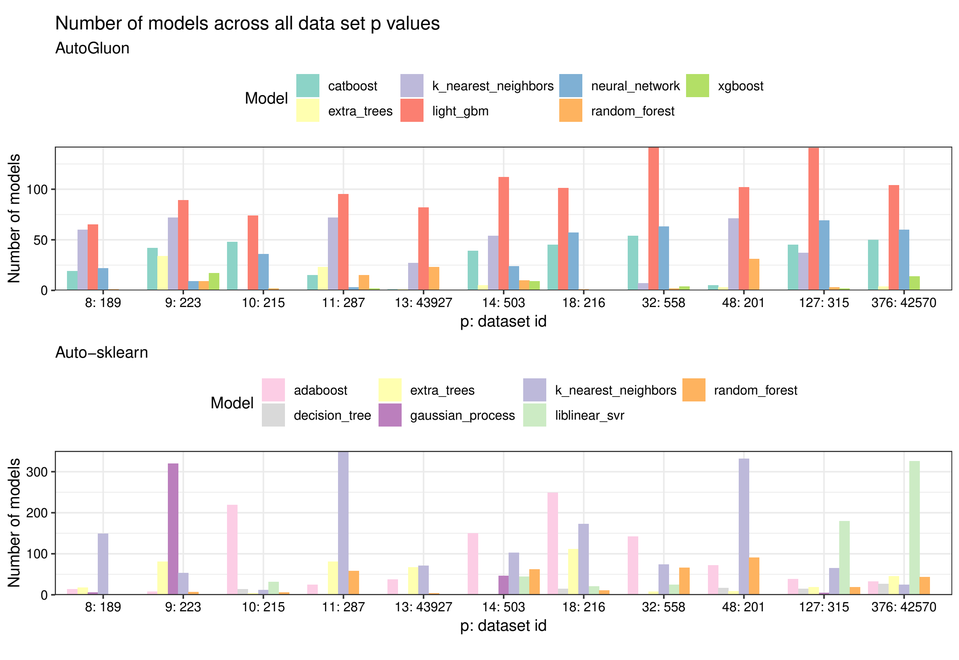
Ensemble methods: stacking.

Advantages? Disadvantages? Use cases?

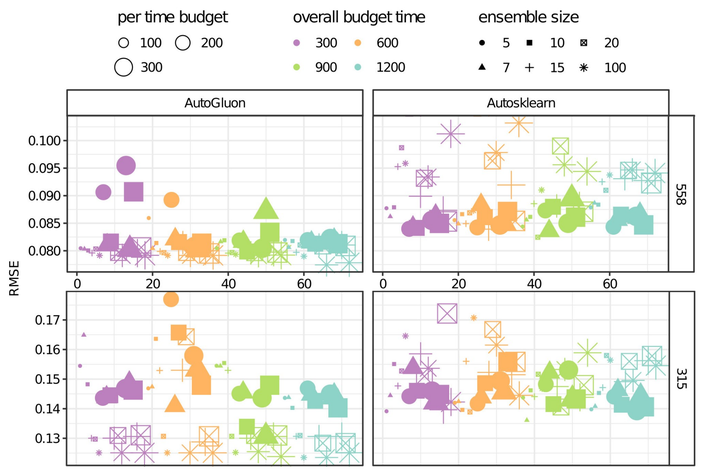
- Our goal is the comparison of the two AutoML frameworks to answer questions about model preference, popular model combinations, efficiency, accuracy, and more.
- The analysis was done over a selection of 11 OpenML data sets and three parameters passed to the frameworks. These parameters control the overall and per run time along with ensemble sizes.

Model selection

- The frameworks use a slightly overlapping selection of ML models, but rarely select create the same ensemble for a certain data set.
- As may be seen on the plot below Auto-sklearn is more effected by varying data set dimensions, while AutoGluon tends to stay persistent with its model selection.
- Another interesting result is present on the visible heat maps which show the most common model combinations within ensembles created by both frameworks.



Number of models across all data set p values.



Multiple parameters

- Here we may see our entire 3-dimensional parameter grid in use on two OpenML data sets.
- Every point represents one of the possible combinations of the three values seen in the plot legend.

References

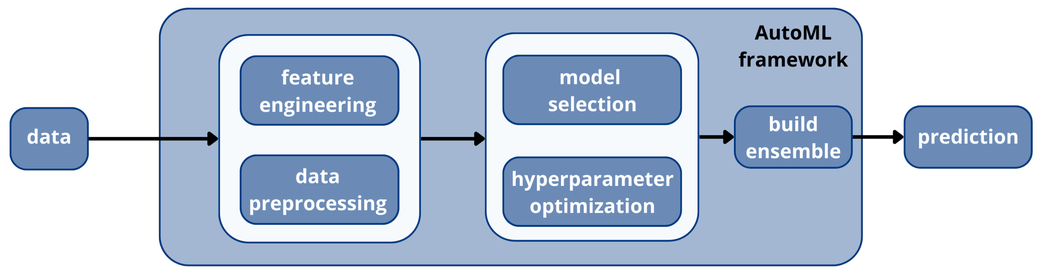
[1] Rich Caruana, Alexandru Niculescu-Mizil, Geoff Crew, and Alex Ksikes. Ensemble selection from libraries of models. In *Proceedings of the Twenty-First International Conference on Machine Learning*, ICML '04, page 18, New York, NY, USA, 2004. Association for Computing Machinery.

[2] Joaquin Vanschoren, Jan N. van Rijn, Bernd Bischl, and Luis Torgo. Openml: networked science in machine learning. *SIGKDD Explorations*, 15(2):49–60, 2013.

[3] Nick Erickson, Jonas Mueller, Alexander Shirkov, Hang Zhang, Pedro Larroy, Mu Li, and Alexander Smola. Autogluon-tabular: Robust and accurate automl for structured data. *arXiv preprint arXiv:2003.06505*, 2020.

[4] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. Efficient and Robust Automated Machine Learning. In *Advances in Neural Information Processing Systems*, 2015.

Automated Machine Learning

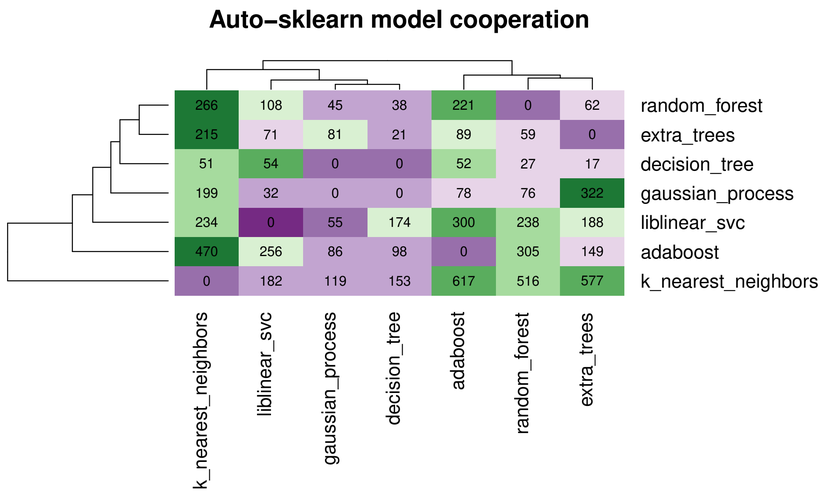


Auto-sklearn

- Developed in 2015.
- Based on the `scikit-learn` Python library.
- Uses Bayesian optimisation for hyperparameter tuning.
- At its release, the framework outperformed all other AutoML tools.

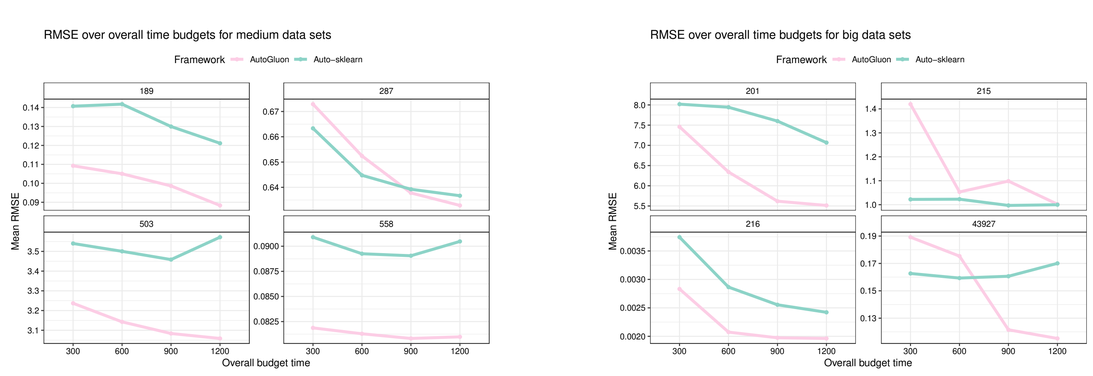
AutoGluon

- Developed in 2020.
- Based around the idea of *Achieving state-of-the-art results with 3 lines of Python code.*
- Uses multi-layer stacking along with k-fold bagging to create optimal model ensembles.



Performance

- AutoML is in general a very computationally expensive task with enormous resource consumption. This has been one of the main criticisms of this field of research throughout its development.
- Below we present results of how changing the overall time budget parameter given to each framework impacts the performance of the final model.



Conclusion

- While many various AutoML challenges are constantly being held during which AutoML frameworks compete in achieving high predictive accuracy, this work provides additional insight into the results of the training and ensembling processes of both frameworks.
- For future work, we should consider performing these experiments on a larger selection of data sets with an increased variety of sizes and dimensions along with a higher-dimensional parameter grid that also includes model hyperparameters.

Contact info

jedrek.rucinski@gmail.com
github.com/jedrzejrucinski/AutoML_Bachelors_thesis