A visual journey through AutoML.
cattleia: a tool for deep dive into ensembles

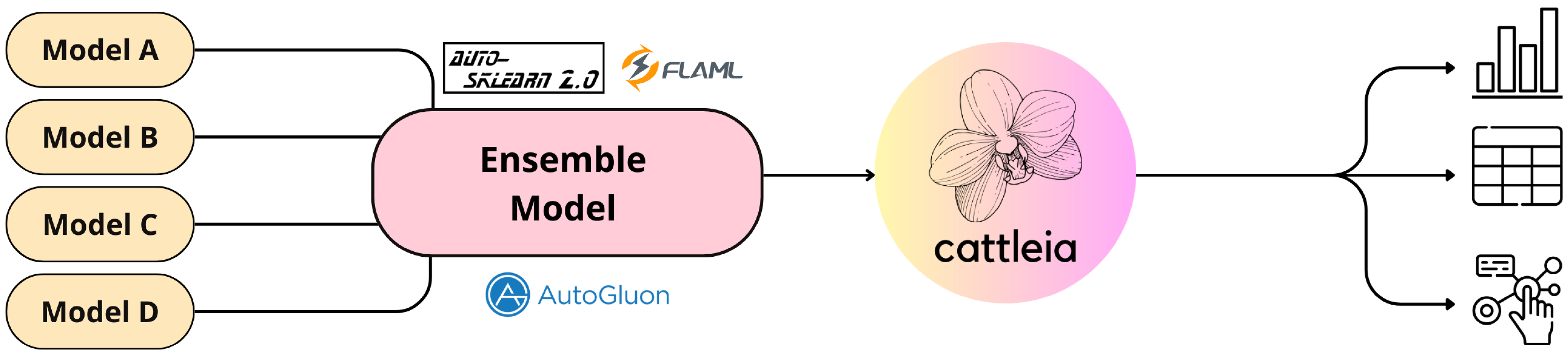
Jakub Piwko, Dominik Kędzierski, Malwina Wojewoda, Anna Kozak
Warsaw University of Technology

Motivation

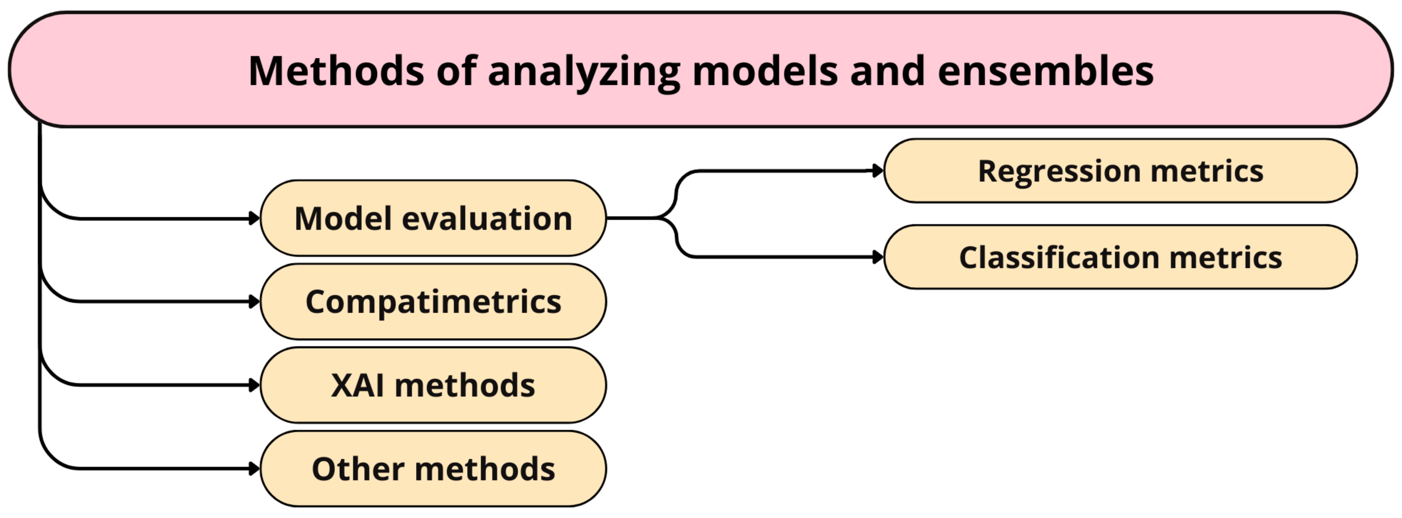
As data scientists experienced with AutoML frameworks we wanted to find answers to questions:

- **?** How component models of ensembles perform independently?
- **?** How similar to each other are these models?
- **?** What if corresponding weights change?

Complex **A**ccessible **T**ransparent **T**ool for **L**earning **E**nsemblers **I**n **A**utoML provides a solution!



Component models analysis



Model evaluation:

- Regression metrics:
  - MSE
  - MAE
  - MAPE
- Classification metrics:
  - Accuracy
  - Recall
  - Precision
  - F1 score

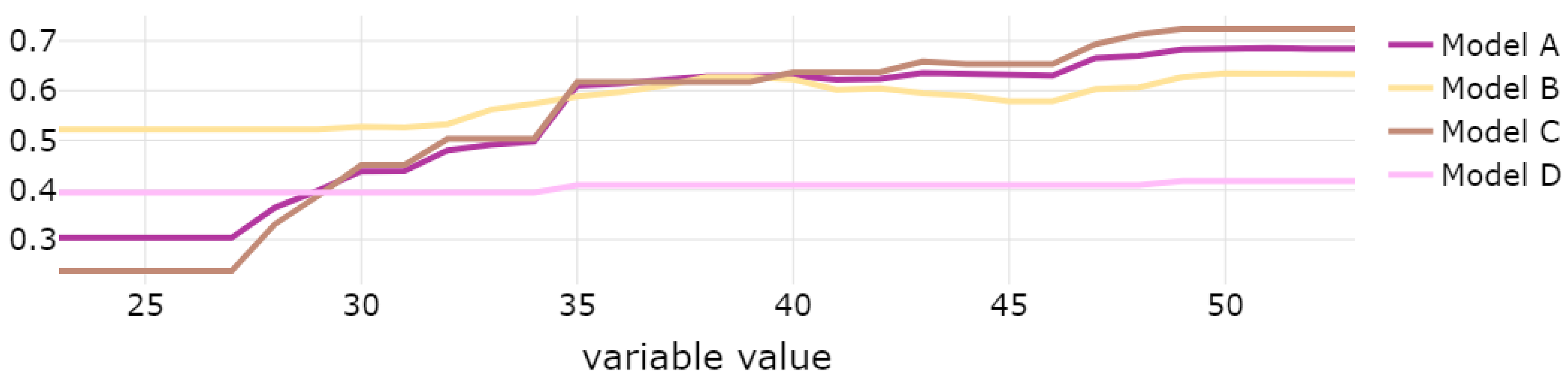
XAI method plots:

- Partial Dependence
- Feature Importance plots

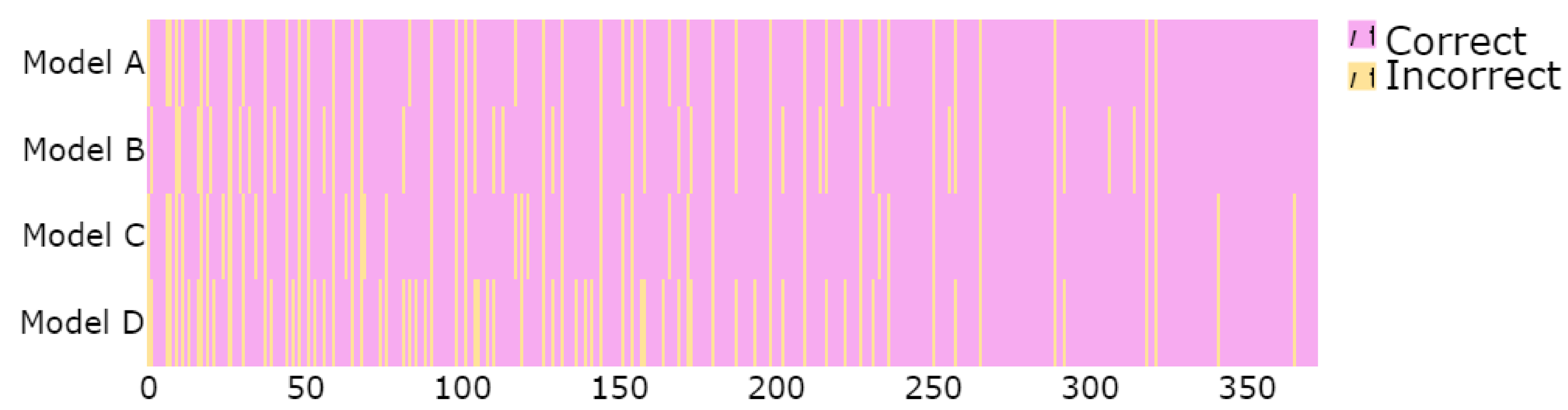
Other methods:

- Prediction correlation matrix
- Prediction compare matrix

Variable partial dependence



Models predictions compare

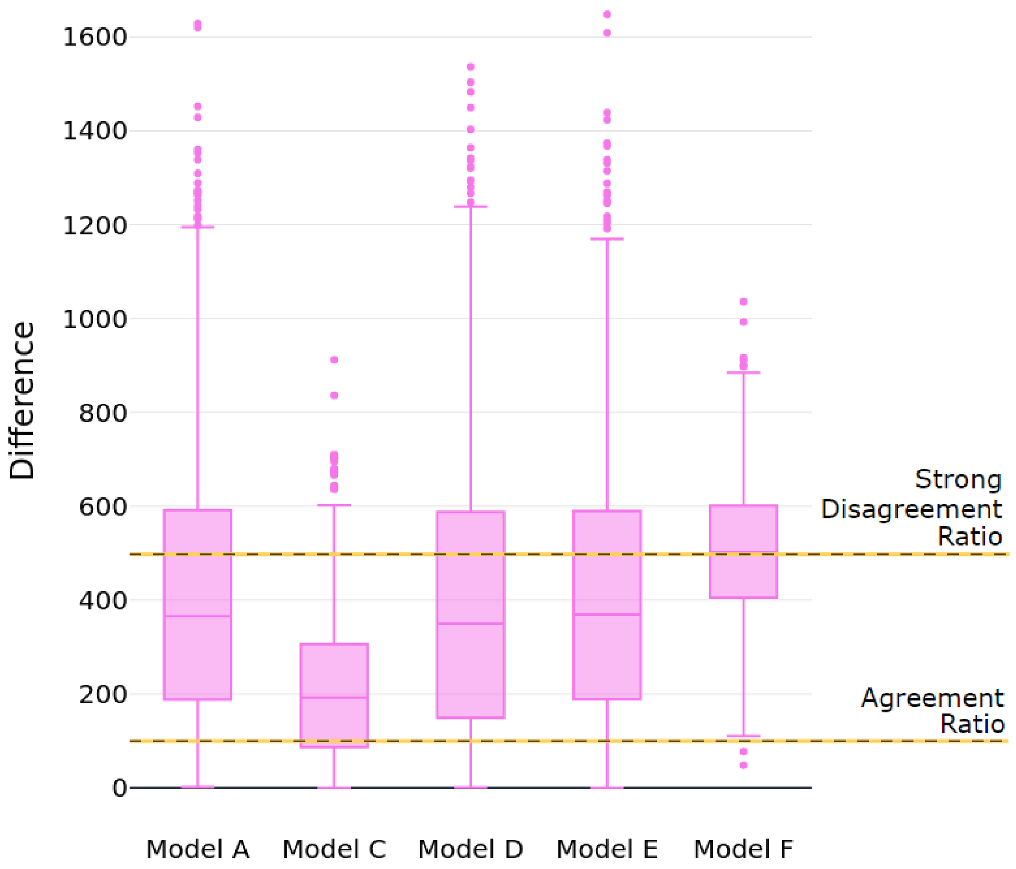


Compatimetrics

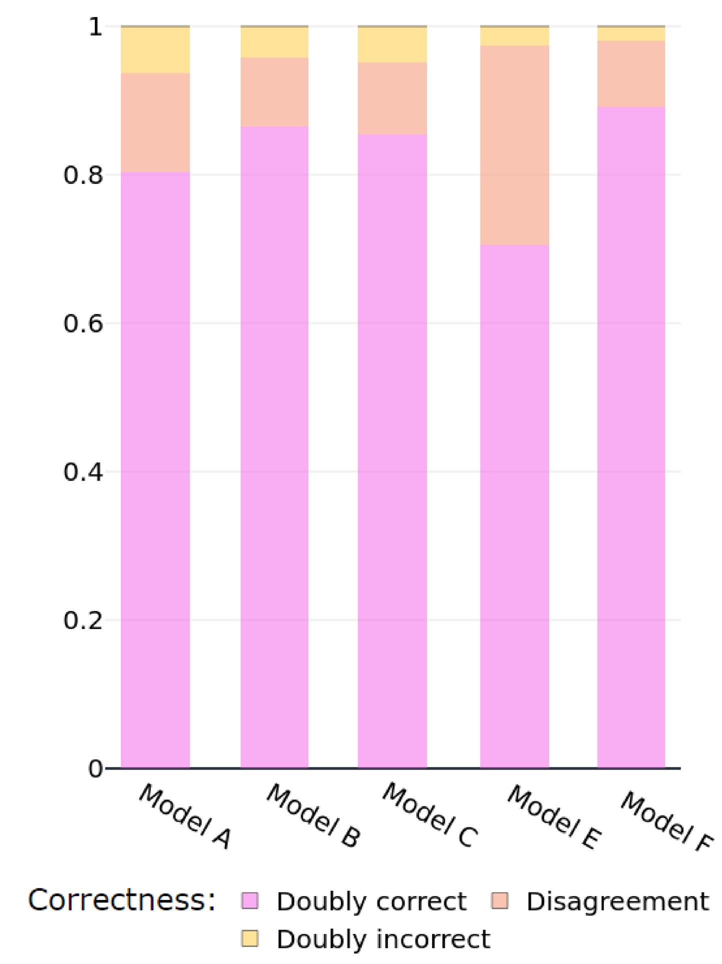
- ℹ Compatimetrics are novel indicators of models compatibility and similarity.
- ℹ They are based on simple heuristics and evaluation metrics.
- ℹ Compatimetrics can be a powerful tool in ensemble analysis and assist in everyday ML tasks to uncover the potential of different models.

- Regression:
  - Mean Squared Difference, Root Mean Squared Difference
  - Strong Disagreement Ratio, Agreement Ratio
- Classification:
  - Uniformity, Incompatibilty
  - Positive Disagreement Ratio, Negative Disagreement Ratio
  - Correctness Counter, Average Collective Score
  - Conjunctive Accuracy, Precision and Recall

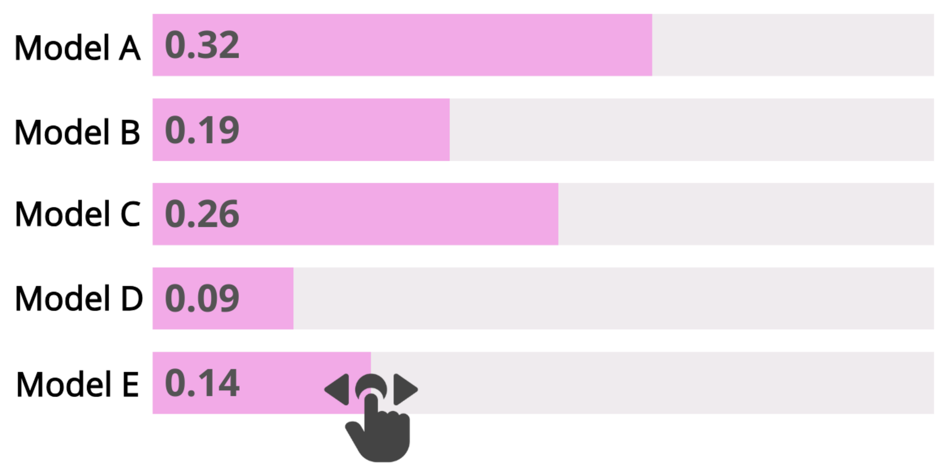
Distribution of difference between predictions of Model B and other models



Percentage of correct predictions of Model D with other models



Weight analysis



	Original model	Customized model
Accuracy	0.943	0.949
Precision	0.912	0.916
Recall	0.927	0.930
F1 score	0.919	0.921

cattleia enables the analysis of weights in an ensemble model. Users can easily adjust the weights in individual models using intuitive sliders and dynamically observe how the task-specific metrics change.

Contact info

✉ `dominik.kedzierski.stud@pw.edu.pl`
✉ `jakub.piwko2.stud@pw.edu.pl`
✉ `malwina.wojewoda.stud@pw.edu.pl`
`github.com/malwina0/cattleia`

Conclusion

- The *cattleia* tool is intuitive and easy to use, as is the process of creating AutoML models.
- It provides valuable insights into forecasting performance, decision-making process and impact of various factors on model performance.
- Its transparent visualisations make it easy for machine learning beginners and experienced data scientists to discover new relationships between models.

References

[1] Jaimie Drozdal, Justin D. Weisz, Dakuo Wang, Gaurav Dass, Bingsheng Yao, Changruo Zhao, Michael J. Muller, Lin Ju, and Hui Su. Trust in AutoML: exploring information needs for establishing trust in automated machine learning systems. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 2020.

[2] 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.

[3] 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.

[4] Chi Wang, Qingyun Wu, Markus Weimer, and Erkang Zhu. FLAML: A Fast and Lightweight AutoML Library. In *MLSys*, 2021.