Tabular Data's Best of Both Worlds: Fresh Art Meets Time-Honoured Craft

Andi Alidema, Rona Latifaj, Thejaswini Raju
AutoKSs

Modality 1/3

Motivation

Tree-based models are reliable workhorses for tabular data, valued for stability and interpretability, but newer deep-learning methods like TabNet and TabPFN uncover subtler patterns. Ensembling lets trees guard against overfitting and provide signals while neural components model complex interactions. The hybrid combines strengths and offsets weaknesses, yielding accurate, robust predictions.

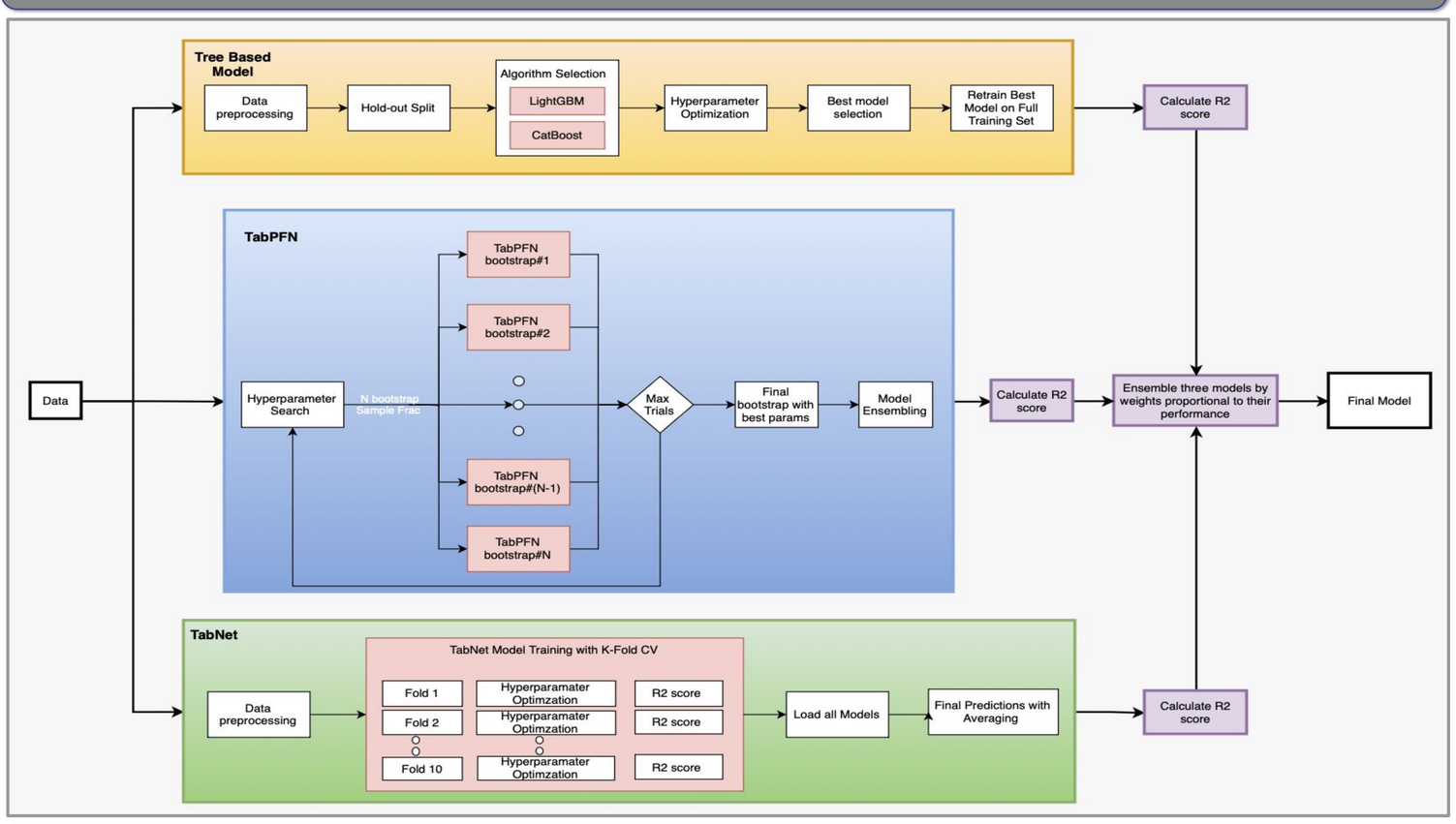
Contributions

- Andi Alidema: Developed the TabPFN pipeline and led the integration of the final weighted ensemble.
- Rona Latifaj: Designed and implemented the tree-based modeling pipeline efficiently.
- Thejaswini Raju: Built and tuned the TabNet module end-to-end for optimal performance

- [x] Week 1
- [x] Week 2
- [x] Week 3
- [x] Week 4
- [ ] Week 5
- [ ] Week 6
- [x] Week 7
- [x] Week 8
- [ ] Week 9
- [ ] Week 10
- [ ] Week 11
- [ ] Bonus
- [ ] Literature

Our Approach

Pipeline workflow



Final model

Each component model is fully tuned and validated on its own hold-out set to assess predictive quality. A final weighted ensemble is constructed, assigning each model a weight based on its validation R^2:

$$w_i = \frac{R_i^2}{\sum_j R_j^2}$$

This adaptive weighting balances contributions automatically, enhancing stability and overall accuracy. The resulting ensemble delivers robust, state-of-the-art performance across diverse tabular benchmarks.

Resources Used

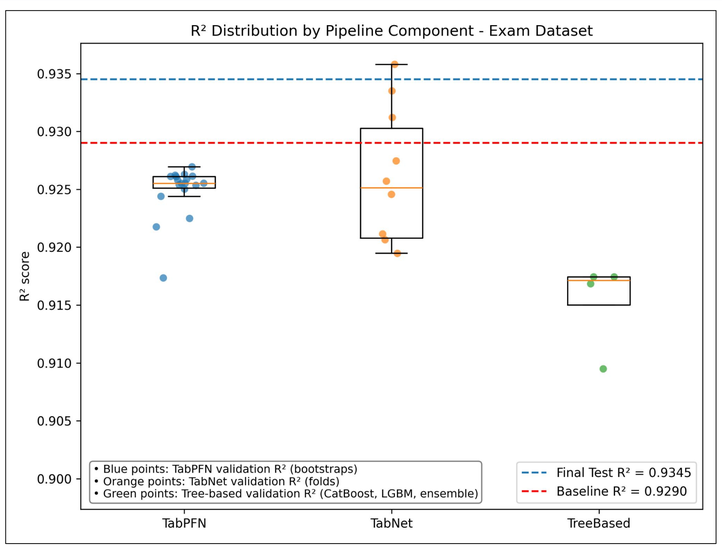
For development:
- 1 RTX2080 GPU
- 1 QEMU Virtual CPU version 2.5+ CPU
- Total compute estimate: 1000 CPU-h

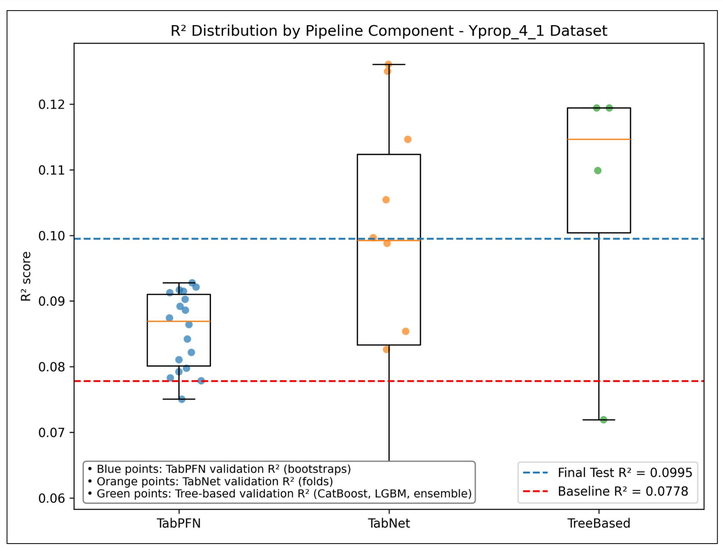
For AutoML:
- 1 RTX2080 GPU
- 7h 02 mins

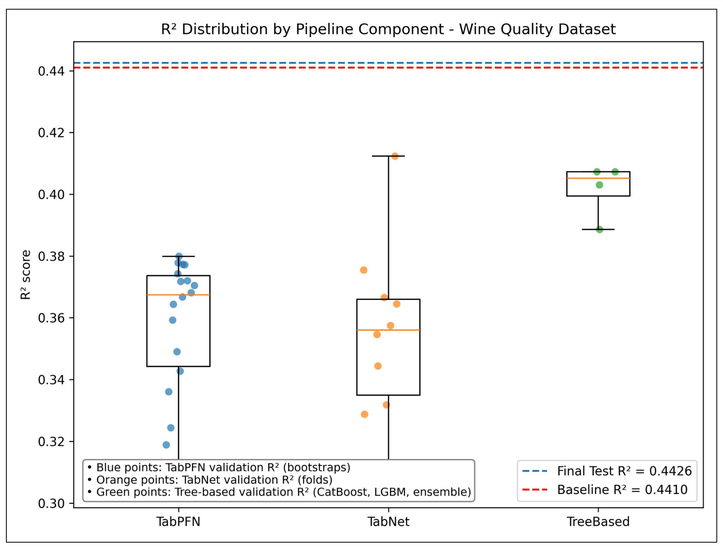
Workforce:
- 2 full week on average

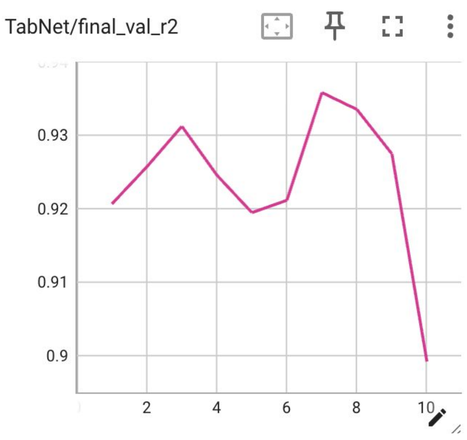
Number of queries for test score generation: **2**

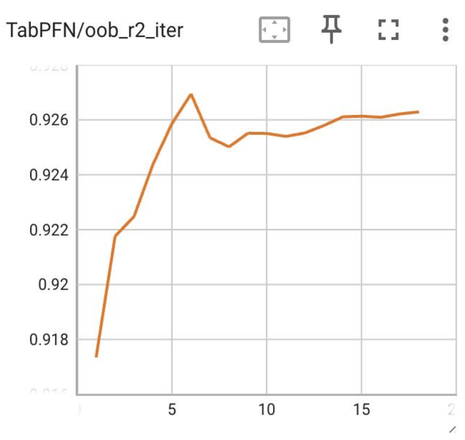
Empirical Results

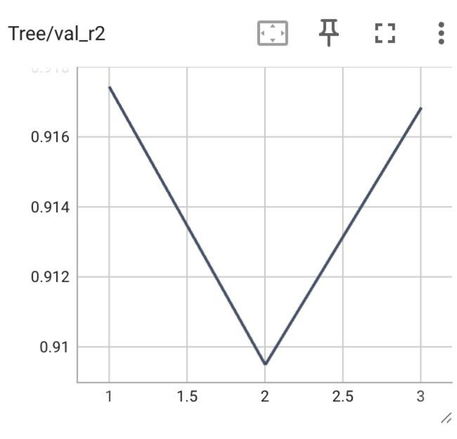












Dataset	Runtime	Baseline	TabPFN	TabNet	Tree-Based	Ensemble
exam_dataset	07:01:14	0.9290	0.9262	0.9239	0.9174	**0.9345**
yprop_4_1	18:19:24	0.0778	0.0921	0.0962	0.1194	**0.0995**
wine_quality	03:10:19	0.4410	0.3799	0.3615	0.4073	**0.4426**
brazilian_houses	02:37:50	0.9896	0.9948	0.9482	0.9953	**0.9891**
bike_sharing_demand	04:20:15	0.9457	0.9418	0.9375	0.9375	**0.9312**
superconductivity	09:30:18	0.9311	0.9140	0.8917	0.9255	**0.9057**

Table 1: Test R^2 across datasets: baseline, individual components, and final weighted ensemble.