LAMARR
INSTITUTE FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Automated Machine Learning for Tabular Data

Research Overview

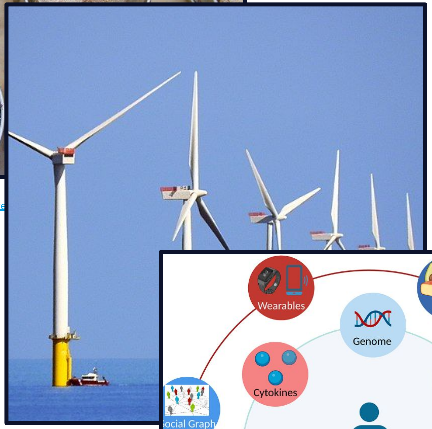
Katharina Eggensperger

Partner institutions:
technische universität dortmund
Fraunhofer IAIS
UNIVERSITÄT BONN
Fraunhofer IML

Institutionally funded by:
Bundesministerium für Forschung, Technologie und Raumfahrt
Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen

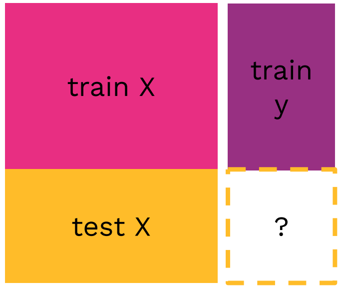
Machine learning for tabular data

credit: https://www.grida.no/re



credit: http://www.geogra

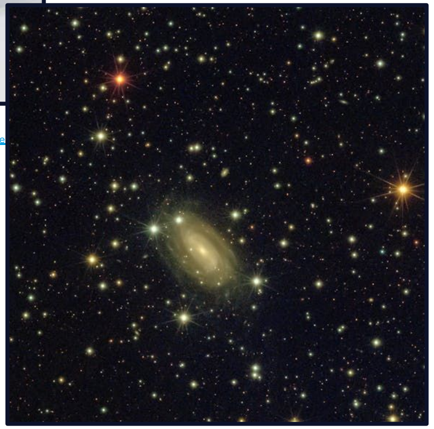
credit: Mohan Babu, Michael Snyder from
https://www.mcponline.org/article/S1535-9476(23)00071-3/fulltext



credit:https://pixnio.com/media/credit-european-investment-loan-money

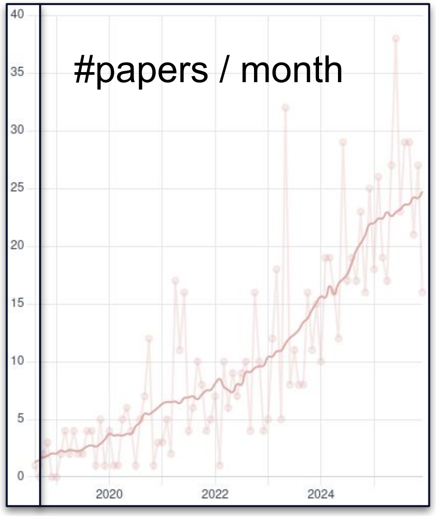
credit: Nataliya Vaitkevich
https://www.pexels.com/de-de/foto/smartphone-laptop-macbook-technologie-6214479/

One foundation model for all tabular tasks?



credit: Dennis Ariel
https://www.pexels.com/de-de/foto/galaxis-galaxie-sternsystem-sterne-10574859/

Machine learning for tabular data is pre-trained



4th Table Representation Learning Workshop @ ACL 2025

July 31st 2025, Vienna, Austria.

WORKSHOP · Fri, Jul 18, 2025 · 9:00 AM – 5:00 PM PDT

1st Workshop on Foundation Models for Structured Data (FMSD)

Nick Erickson · Xiyuan Zhang · Abdul Fatir Ansari · Boran Han · Mononito Goswami · Samuel Gabriel Müller · Lennart Purucker · Yuyang
Michael Mahoney

INNOVATION > VENTURE CAPITAL

From Text To Tables: Why Structured Data Is AI's Next $600 Billion Frontier

By Rocio Wu, Contributor. Rocio writes about how to Think Global, Act Local in innovation

Follow Author

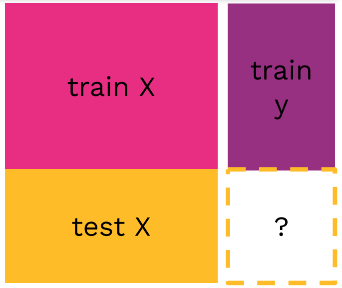
Published Jan 15, 2026, 02:46pm EST, Updated Jan 16, 2026, 01:39pm EST

EU-Startups

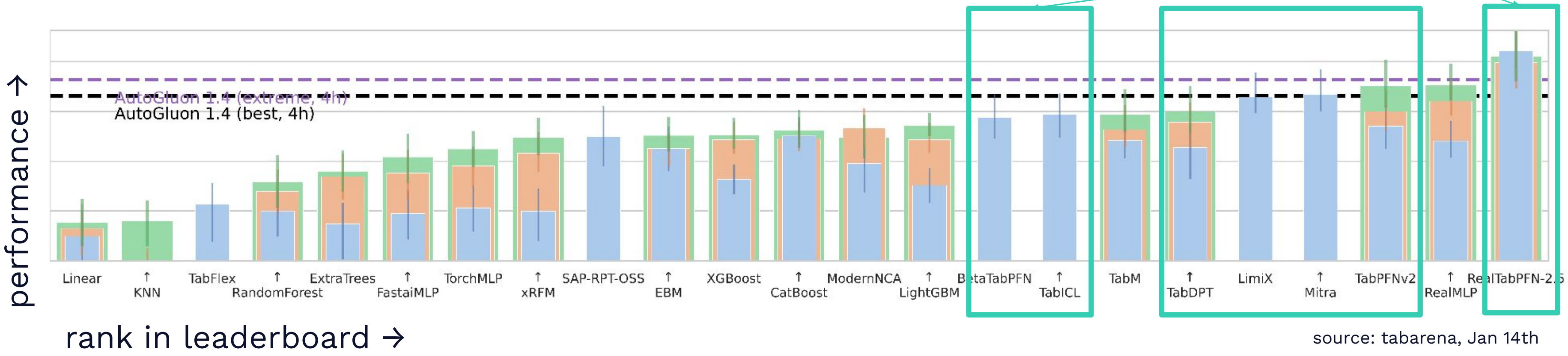
Prior Labs raises €9 million for foundation models for spreadsheets and databases

Prior Labs raises €9 million for foundation models for spreadsheets and databases ...
Freiburg-based Prior Labs, an AI startup innovating...

05.02.2025

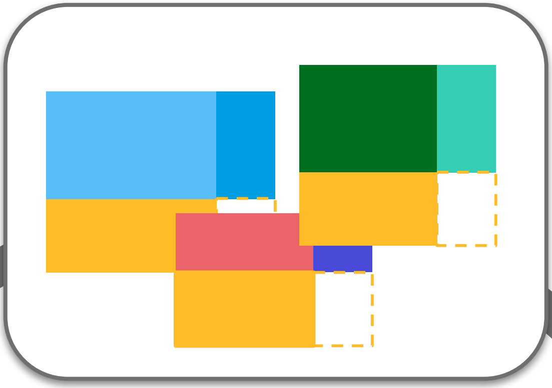


foundation models for predictive tabular tasks



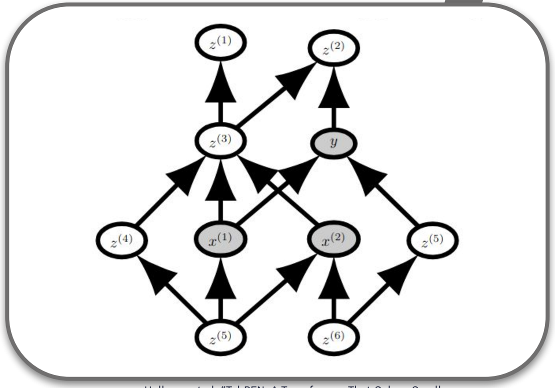
source: tabarena, Jan 14th

(Tabular) Prior-fitted networks



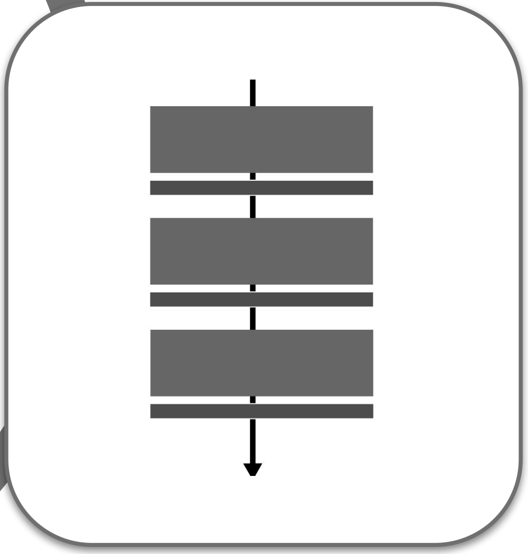
generate synthetic data and
split into train and test set

compute loss and
update weights



source: Hollman et al. "TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second" in ICLR'22

trained to "solve" tabular tasks via in-context learning



TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second
Hollmann, Müller, **KE** and Hutter. ICLR. (2023)

Transformers Can Do Bayesian Inference
Müller, Hollmann, Pineda, Josif and Hutter. ICLR. (2022)

Katharina
Eggensperger

LAMARR
INSTITUTE FOR MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

How can we

automatically,
systematically,
reproducibly

adapt ML given a (**tabular**) task?

Amir
Balef
(joined 2023)

Mykhailo
Koshil
(joined 2024)

+ 2.5 PhDs (to join soon).

How can we systematically assess models?

How can we build more flexible models?

AutoML for tabular data

How can we do model selection and hyperparameter optimization?

What have models learned?

Data leakage for benchmarking tabular foundation models

MSc (BSc)

Elephants Never Forget.

TL;DR. Prior work has shown that test data leakage for LLM evaluation has an impact, but it is probably not as bad as expected (since the models are very large and forget examples they've seen only once during training). Tabular foundation models are much smaller, so results might be different.

Goal. Set up an evaluation pipeline with tabular foundation models. Evaluate the data contamination of prominent benchmarks and in controlled experiments.

Research questions:

- Is data leakage an issue when evaluating tabular models?
- Does this change if we continue pre-training on a few tasks? Does the model get worse on others?
- Are our results similar to those found for LLMs?

Uncertainty Quantification for TabPFN

MSc (BSc)

Aleatoric Uncertainty.

TL;DR. TabPFN returns a bar distribution mixing aleatoric and epistemic uncertainty.

Goal. Implement and evaluate uncertainty quantification methods for TabPFN models.

Research questions:

- How do they compare against each other?
- How do they compare against models?
- Can we use this to measure prior mismatch?

Impact of Pre-training Data on Generalization Performance

MSc

Inter- and Extrapolation.

TL;DR. The performance of TabPFN depends on the “prior” and on whether a task at hand is likely under that prior.

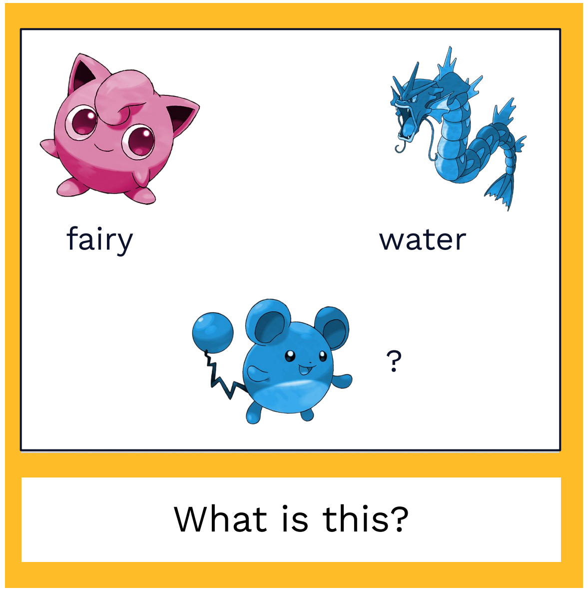
Goal. In controlled scenarios, evaluate the generalization behaviour of PFNs outside their prior.

Research questions:

- Does generalization depend on model size?
- Can the model interpolate and extrapolate priors?
- Does it generalize across input distributions?
- When trained on two complementary priors (e.g., RF vs. GP), do we get the best out of two worlds or a mix?

Out-of-distribution Generalization

MSc



TL;DR. It seems like Tabular foundation models can behave like exemplar or rule-based models; it is unclear whether this behaviour depends on pre-training data or context. Knowing this helps us to better understand how TabPFN works.

Goal. Define evaluation scenarios to study when a PFN behaves like what.

Research questions:

- Do all PFNs behave rule-based for small tasks and exemplar-based for large tasks?
- When trained on data generated from rule-based models, does it “strictly” behave rule-based?

More Topics

MSc BSc

Type I: “Reproduce-and-extend”

- Reproduce and extend amortized multi-objective Bayesian Optimization [link]
- Study Layer contribution for LLM-based models [link]
- Adversarial (continued) pre-training for PFNs [link]
- Sensitivity of fine-tuning hyperparameters for TabPFN? [link]

Type II: “Explorative-and-open-ended”

- Model-merging for PFNs
- Comparison of feature importance over input variations
- Analysis of (catastrophic) forgetting

Preliminaries and Sidenotes

MSc BSc

⏱️ timeline ⏱️

1. kickoff meeting (beginning of semester)
2. thesis proposal
3. registration

   … regular meetings …

4. intermediate presentation

   … regular meetings …

5. feedback on thesis draft
6. final presentation

✅ requirements ✅

- one or more ML lectures (= basics of ML and DL; WRUMS, PRIML, and DL)
- solid Python knowledge (= to implement and empirically evaluate algorithms; PyTorch)

⏭️ What's next ⏭️

→ Send me an email

+ your preference
+ 1-2 sentences about yourself
+ transcripts/course list