The Tabular Foundation Model TabPFN Outperforms Specialized Time Series Forecasting Models Based on Simple Features

Shi Bin Hoo, Samuel Müller, David Salinas, Frank Hutter.

NEURAL INFORMATION PROCESSING SYSTEMS

universität freiburg

ellis INSTITUTE TÜBINGEN

With only simple features, TabPFN-TS matches state-of-the-art Chronos-Large (65x larger) forecasting performance.

We demonstrate that the tabular foundation model TabPFN, when paired with minimal featurization, can perform **zero-shot forecasting**. Its performance on point forecasting matches or even slightly outperforms state-of-the-art methods.

Methodology

We frame time-series forecasting as a **tabular regression problem**, where each time series is treated as an independent table.

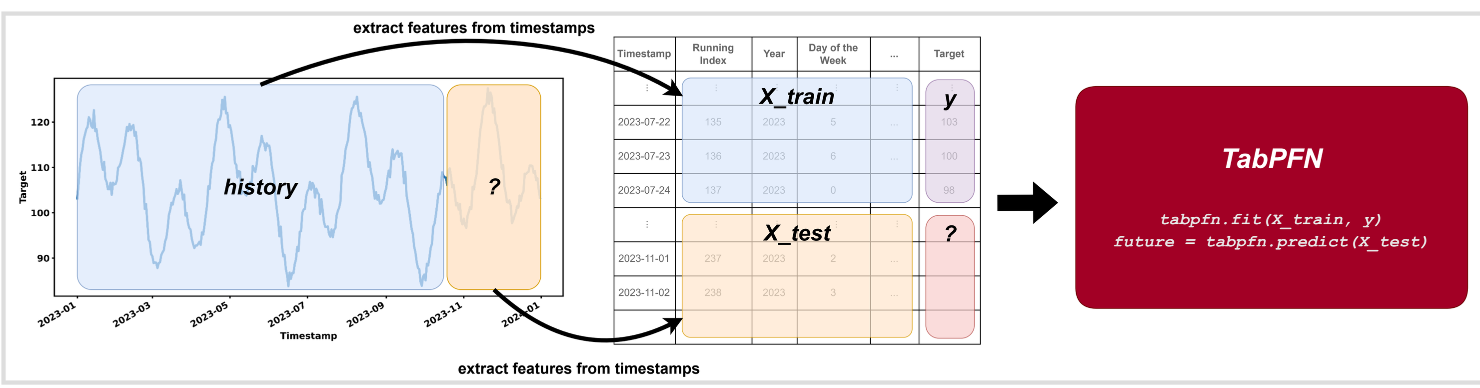

Figure 1: Overview of TabPFN-TS.

For each time series, we transform the sequence into a table, alongside with some features as new columns. This table is fed to TabPFN to perform regression on all future time steps (i.e. forecasting) in a single iteration — **multi-step-ahead forecasting**.

Featurization - we derive features directly from the timestamps, excluding lagged and autoregressive features (e.g. moving averages, lag terms).

Age Feature | Calendar Features in Sine and Cosine Encodings

Timestamp	target	age	year	month-of-the-year (sine)	month-of-the-year (cosine)	week-of-the-year (sine)	week-of-the-year (cosine)	day-of-the-year (sine)	day-of-the-year (cosine)	...
2023-07-22	300	1	2023	0.551	0.835	0.545	0.838	0.528	0.849	...
2023-07-23	305	2	2023	0.551	0.835	0.545	0.838	0.530	0.848	...
2023-07-24	308	3	2023	0.551	0.835	0.545	0.838	0.533	0.846	...
...	...	...	...	...	...	...	...	...	...	...

Figure 2: Featurization.

Results: TabPFN-TS matches Chronos-Large!

We evaluate the **point forecast accuracy** across 24 common datasets.

TabPFN-TS performs on par with, or slightly outperforms state-of-the-art methods.

- *surpasses **Chronos-Mini (20M) by 7.7%***
- *shows modest improvement over **Chronos-Large (710M) by 3.0%***

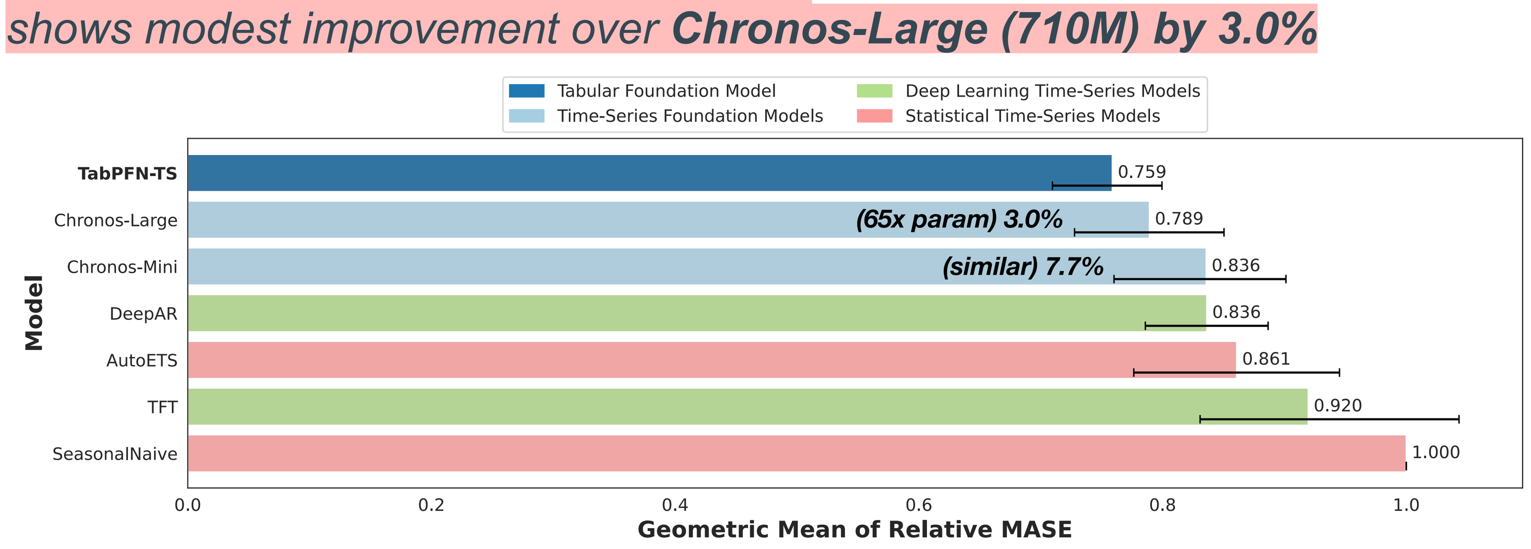

Figure 3: Point forecasting performance of various models. 95% confidence interval included. Lower is better.

We also aggregate the scores based on Chronos' **in-domain** and **zero-shot split.**

- Chronos outperforms TabPFN-TS on dataset it was pre-trained on.
- TabPFN-TS outperforms Chronos on unseen datasets.

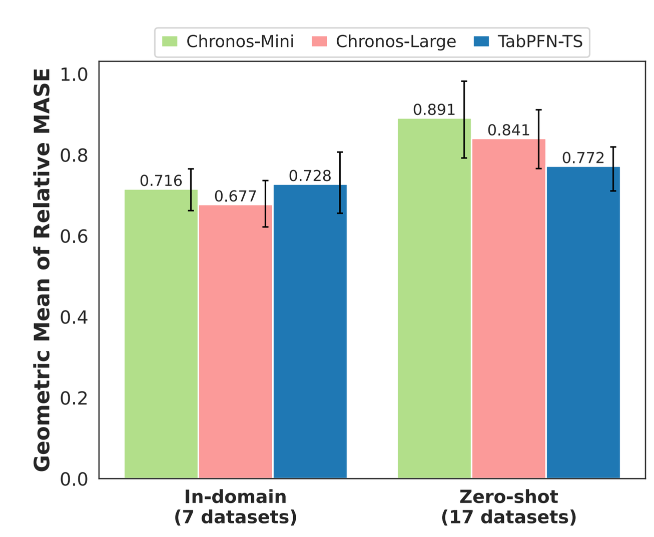

Figure 4: Forecasting Performance grouped by Chronos' in-domain vs zero-shot datasets split.

Relative MASE *of the method:* $RelativeMASE = \frac{MASE_{method}}{MASE_{seasonalnaive}}$

Ablation Study

Which features are essential?

- calendar features are crucial
- "age" alone results in poor performance
- sine & cosine encodings are not always beneficial

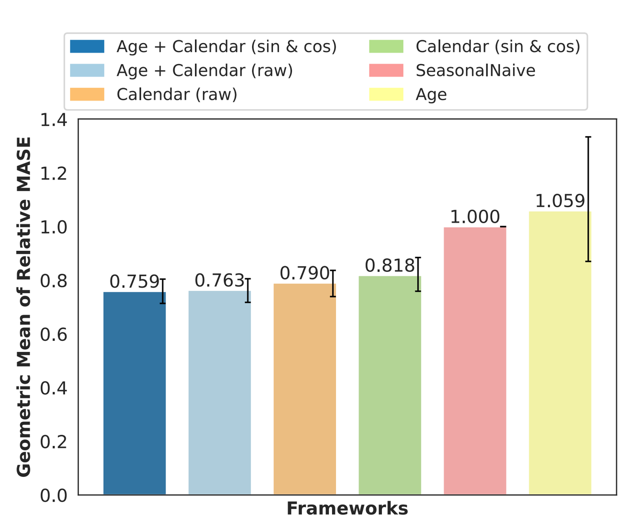

Figure 5: TabPFN-TS performance with different feature combinations.

Can any tabular regressor achieve this?

We use a default CatBoost regressor instead of TabPFN as the local forecasting model while keeping the rest of the pipeline unchanged.

CatBoost fails even to match Seasonal Naive, suggesting TabPFN's unique capability as a tabular foundation model for time series forecasting.

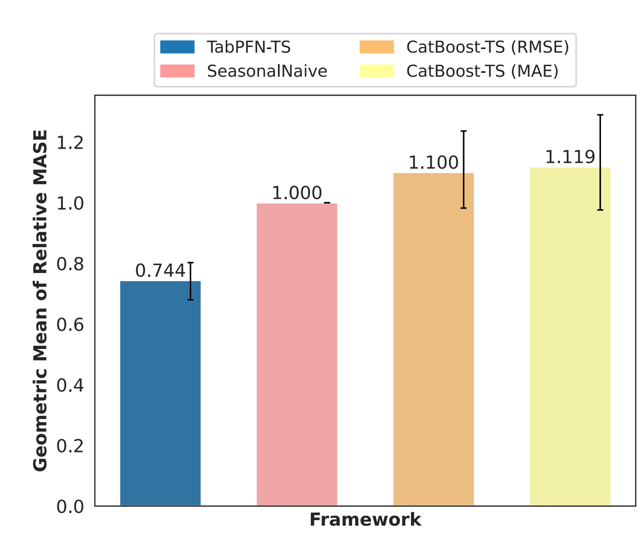

Figure 6: Performance of TabPFN vs CatBoost.

Quantitative Analysis

✅ **good at predicting seasonal patterns**

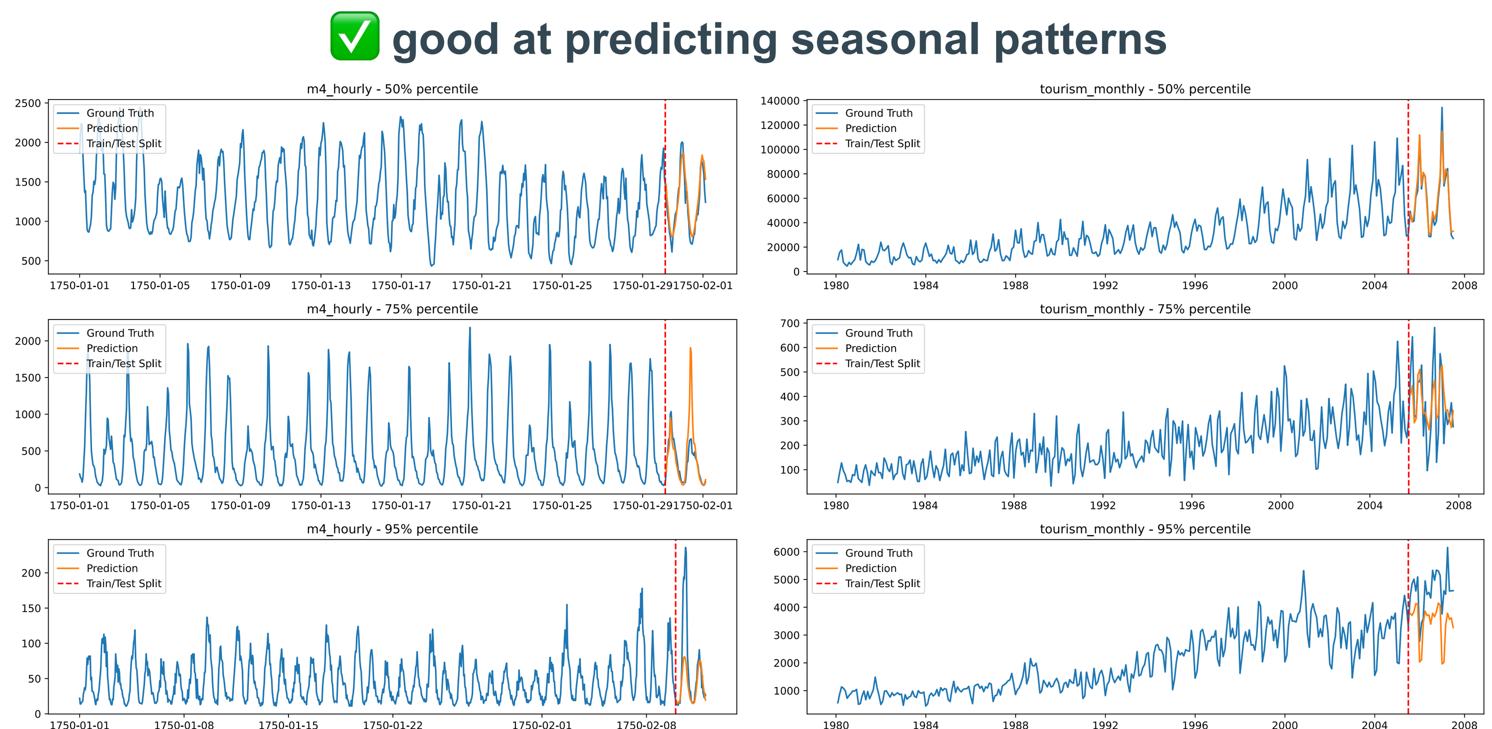

Figure 7: Visualization of TabPFN-TS's predictions on M4 Hourly and Tourism Monthly.

❌ **suffer at predicting trend → tendency to stagnate**

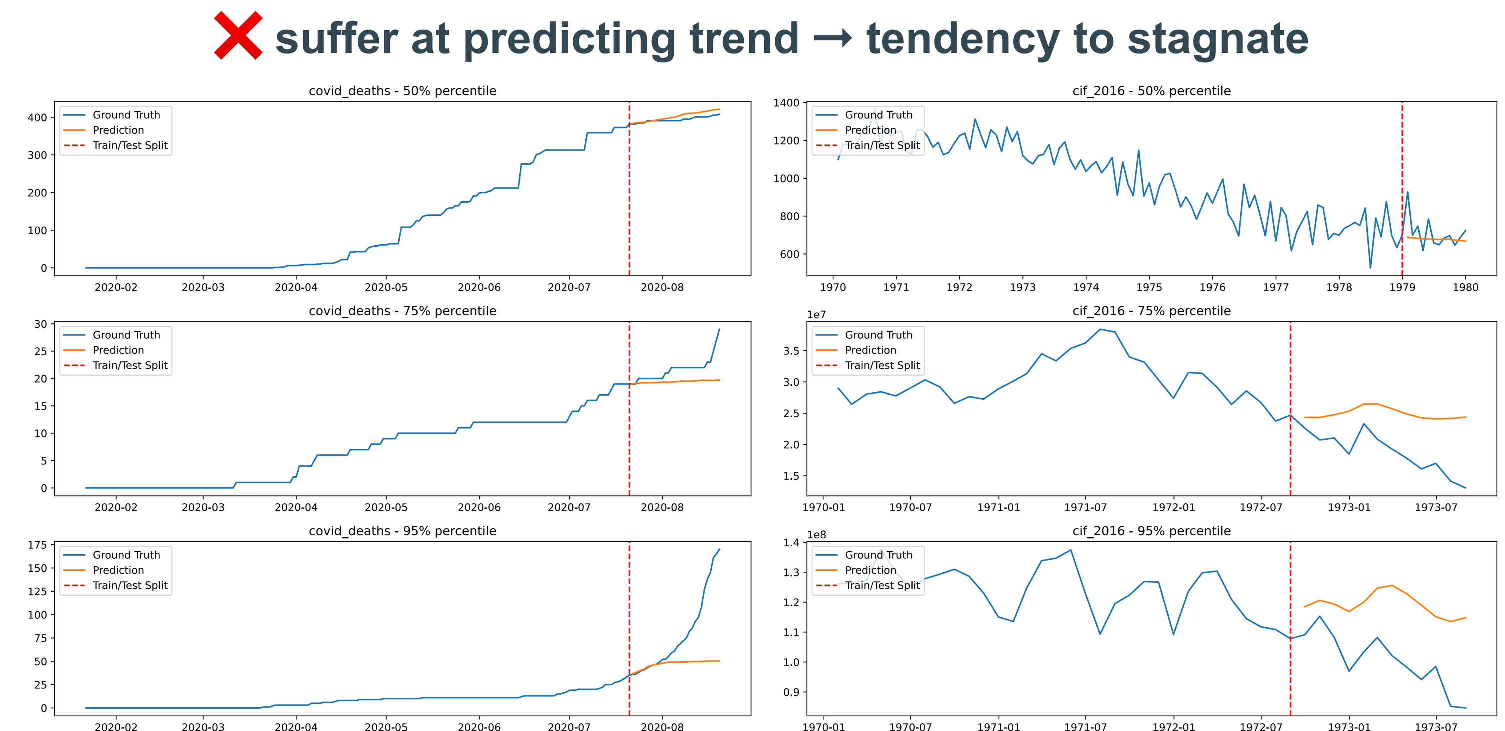

Figure 8: Visualization of TabPFN-TS's predictions on Covid Deaths and CIF 2016.

Takeaways

✅ An evidence for TabPFN, a tabular foundation model, being an incumbent for time series forecasting with minimal feature engineering.

✅ A hint towards the broader potential of tabular foundation models in advancing time series forecasting.

References

Hollman et al. "TabPFN: A transformer that solves small tabular classification problems in a second." In: The Eleventh International Conference on Learning Representations, 2023.

Shchur et al. "AutoGluon-TimeSeries: AutoML for probabilistic time series forecasting." In: International Conference on Automated Machine Learning, 2023.

Ansari et al. "Chronos: Learning the language of time series." In: Transactions on Machine Learning Research, 2024.

SCAN ME