Journal Club Session: TabPFN

Accurate predictions on small data with a tabular foundation model
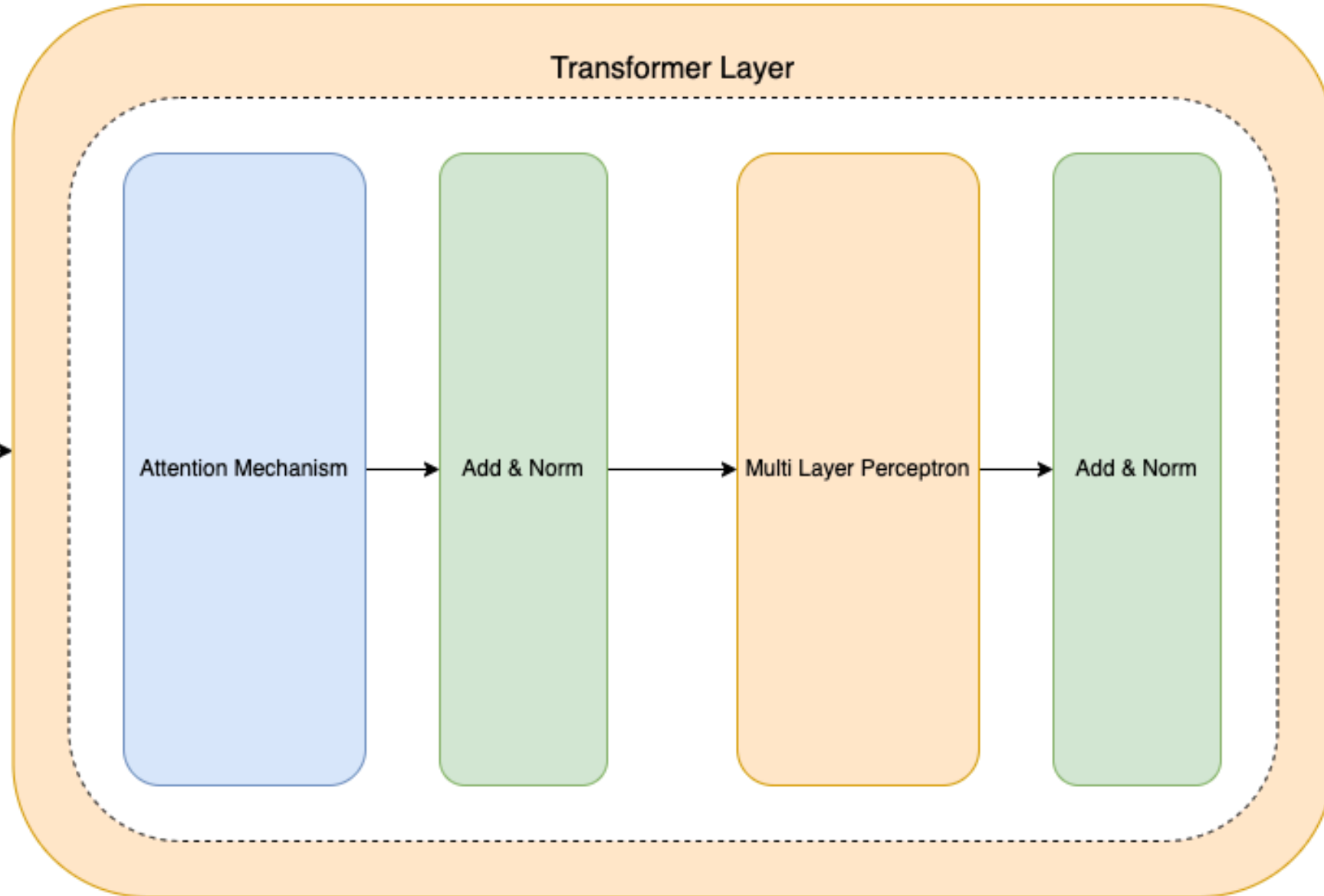
Autoren

- Noah Hollmann

- Samuel Müller

- Lennart Purucker

- Arjun Krishnakumar

- Max Körfer

- Shi Bin Hoo

- Robin Tibor Schirrmeister

- Frank Hutter

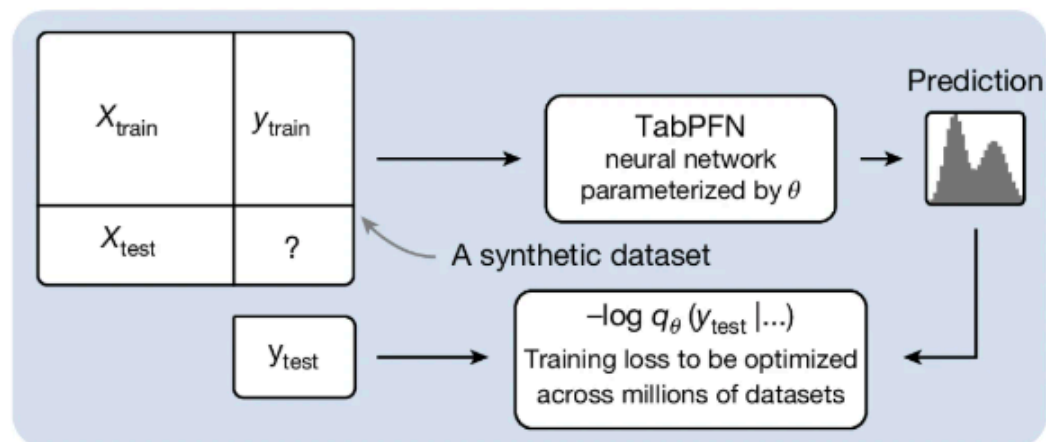# Transformer

Transformer Layer

Input

Attention Mechanism → Add & Norm → Multi Layer Perceptron → Add & Norm
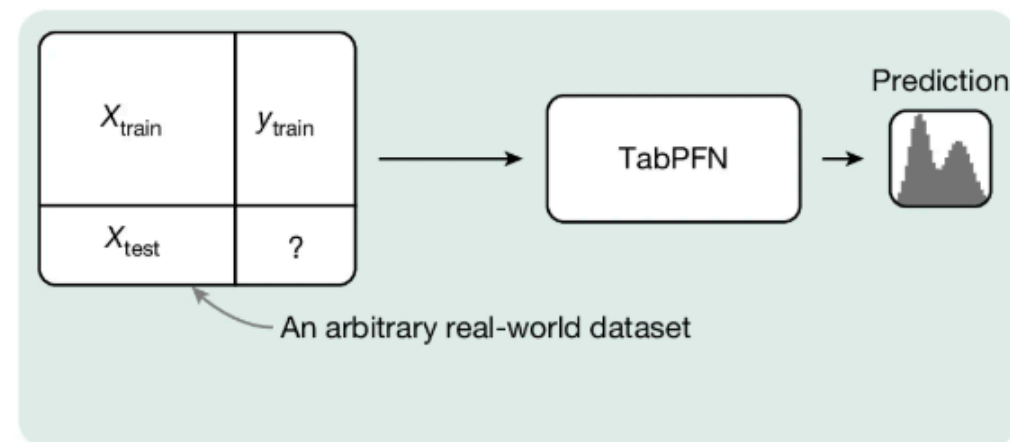
Output

Fig. 1: Overview of the proposed method.

a

TabPFN is trained on synthetic data to take entire datasets as inputs and predict in a forward pass

TabPFN can now be applied to arbitrary unseen real-world datasets

X_{train}	y_{train}
X_{test}	?

A synthetic dataset

TabPFN neural network parameterized by θ

Prediction

y_{test}

$-\log q_\theta (y_{test} |...)$
Training loss to be optimized across millions of datasets

X_{train}	y_{train}
X_{test}	?

An arbitrary real-world dataset

TabPFN

Prediction

b

Input dataset

	x_1	x_2	y
Training rows	1.2	6.1	3.0
	8.9	9.1	3.1
	1.0	2.9	6.7
Test	33.3	2.2	?

We predict this entry

2D TabPFN layer (12×)

1D feature attention

1D sample attention

MLP
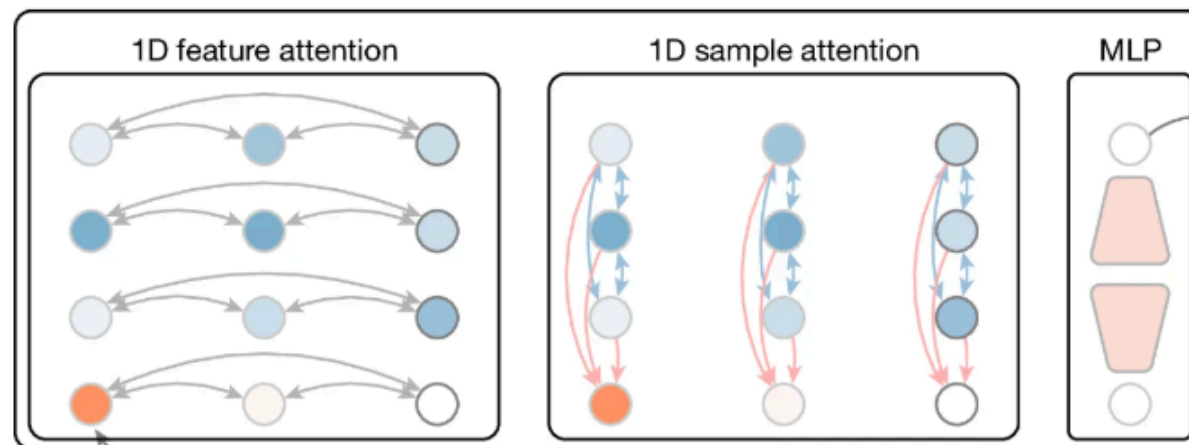
Each node represents one entry in the table

Predictions: $\hat{y}_{test}$

The vector is transformed to a piece-wise constant (Riemann) distribution with an MLP

Density

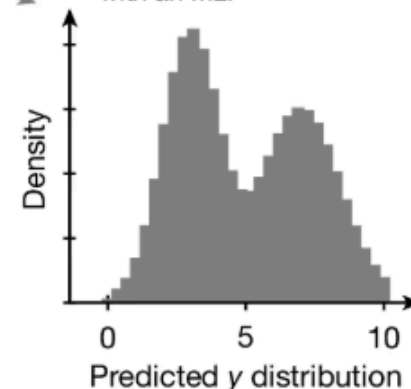0   5   10

Predicted y distribution

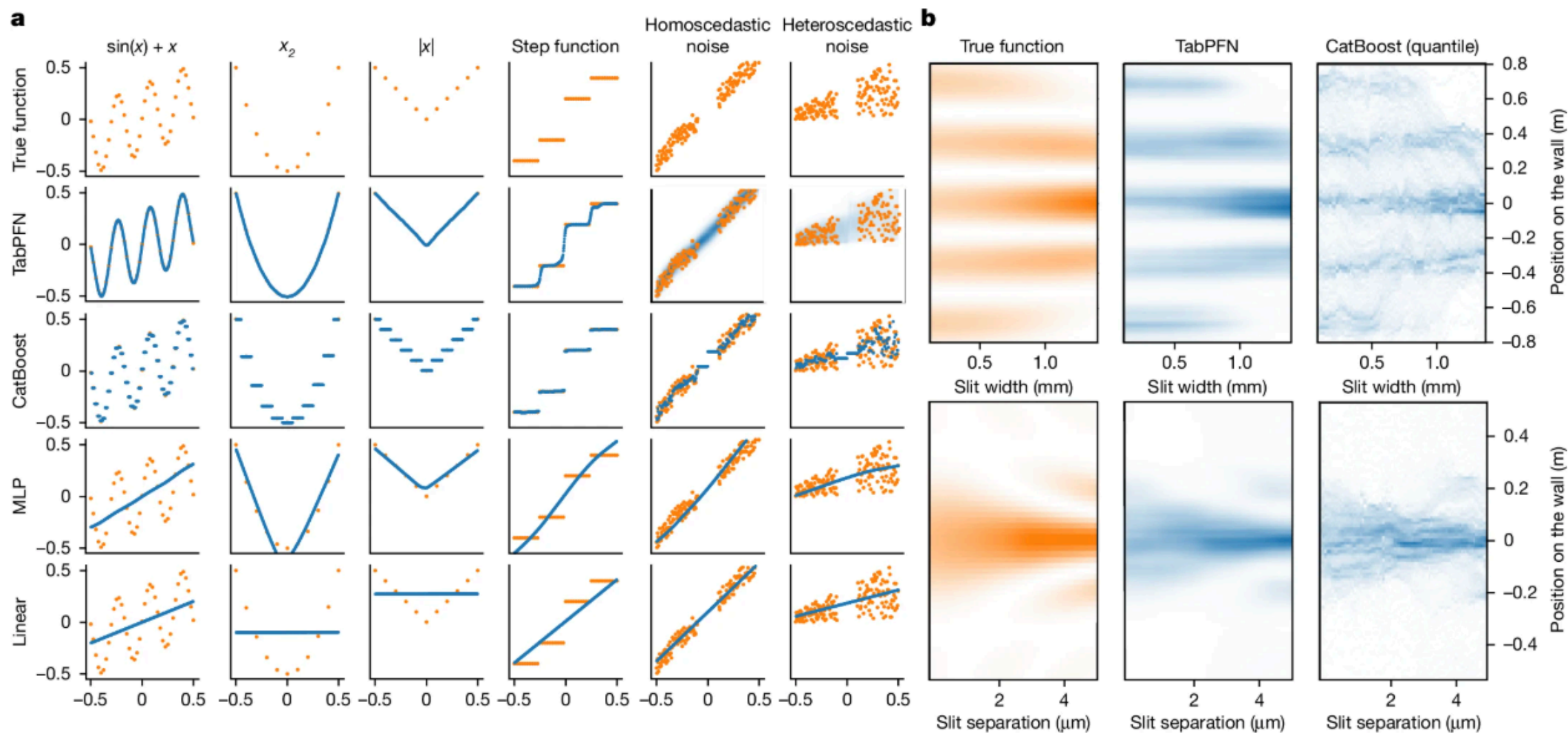# Fig. 3: The behaviour of TabPFN and a set of baselines on simple functions.

Fig. 4: Comparison of TabPFN on our test benchmarks, containing datasets with up to 10,000 samples and 500 features.