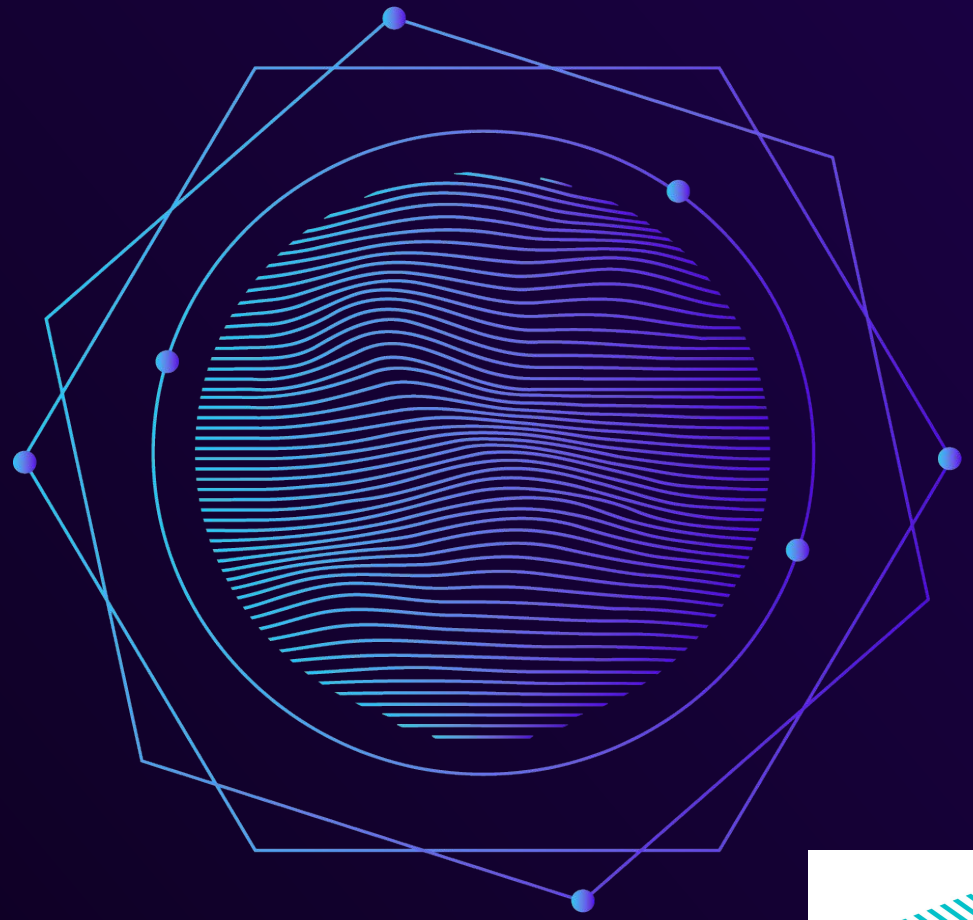
DATA SCIENCE SUMMIT ML EDITION



Occam's razor vs market:

How simple models can answer complex questions

Mateusz Panasiuk

Chief Scientific Officer, AI Investments

www.ml.dssconf.pl 15th-16th of June 2023 Warsaw + Online

ADVISOR

Organizer: Academic Partners FUNDACJA

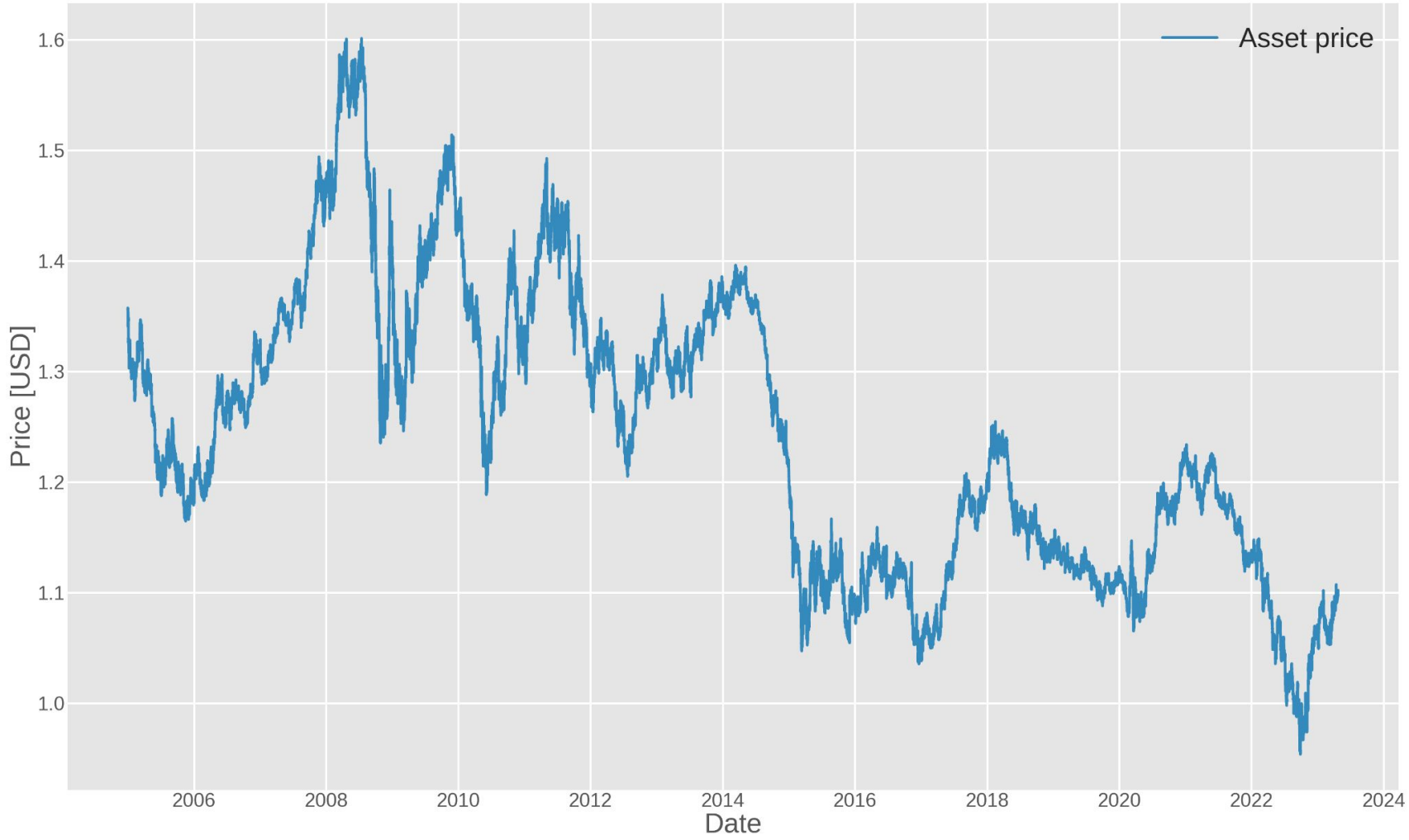
POIR.01.01.01-00-0144/18

A brief sprint through financial modelling

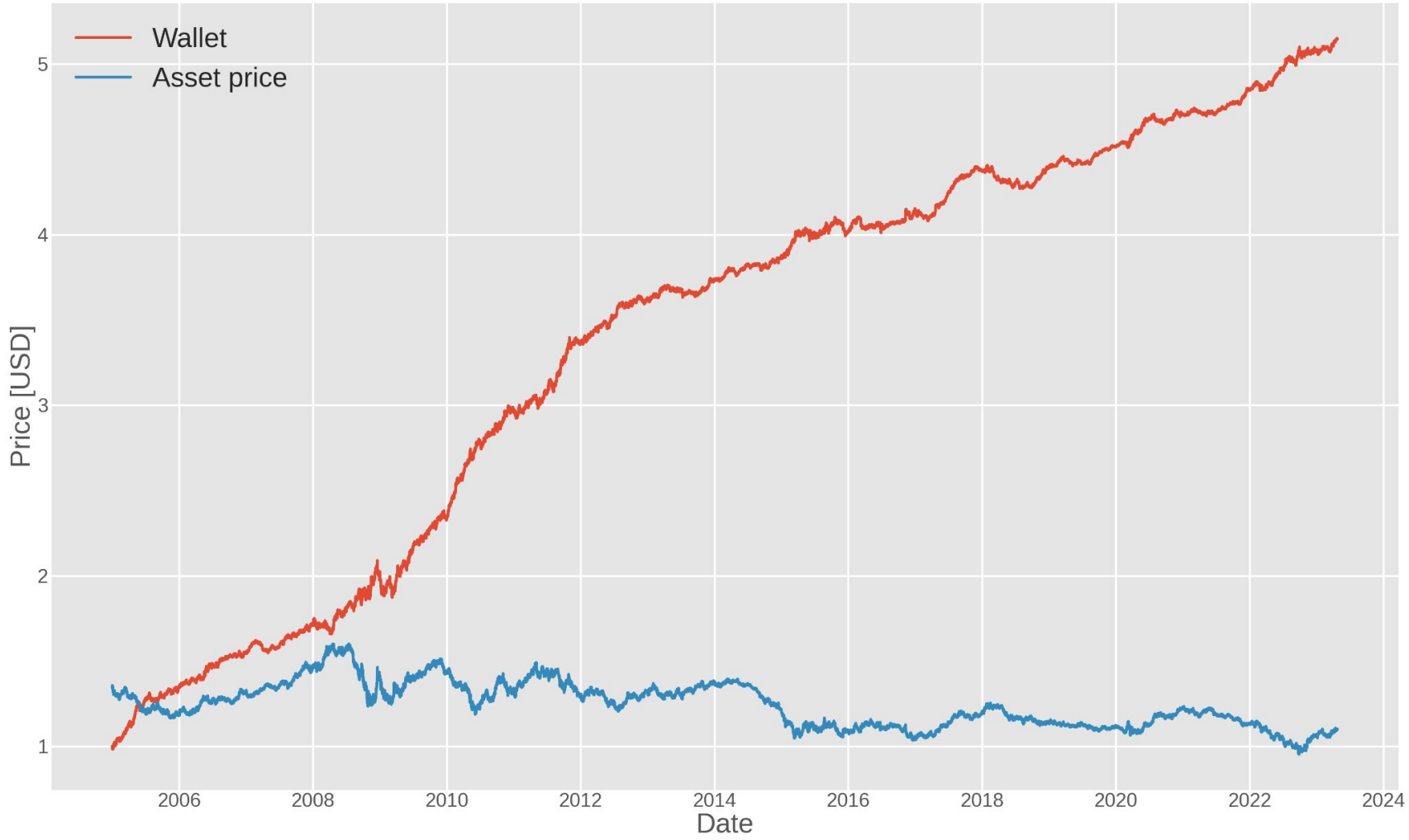
1. Challenges of financial forecasting
2. Problem of model complexity
3. Solutions for model complexity
4. Results

Is speculation hard?

Is speculation hard?



How good do we need to be to beat the market?



How to solve a complex problem?

Newest neural networks

Newest neural networks

More features

Engineer features

Newest neural networks

More features

Engineer features

Newest neural networks

Bigger neural networks

More features

GPT!

Engineer features

Newest neural networks

Bigger neural networks

More features

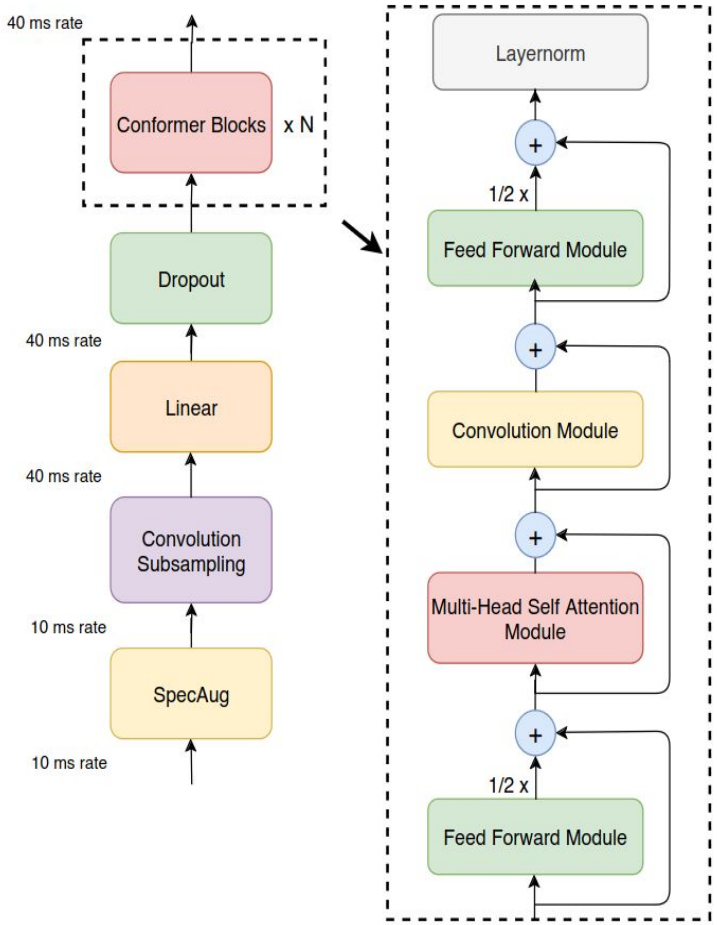
William’s razor

“Pluralitas non est ponenda sine necessitate.” - attributed to William of Ockham

“Objects should not be multiplied without necessity.”

wikipedia.org/wiki/William_of_Ockham#/media/File:William_of_Ockham.png; unedited, CC BY-SA 3.0

Leading neural network architecture



{"d_model": 32,

"N": 6,

"heads": 8,

"features_count": 24,

"length": 20}

300`000 parameters

Gulati, A. et al. (2020). Conformer: Convolution-augmented Transformer for Speech Recognition

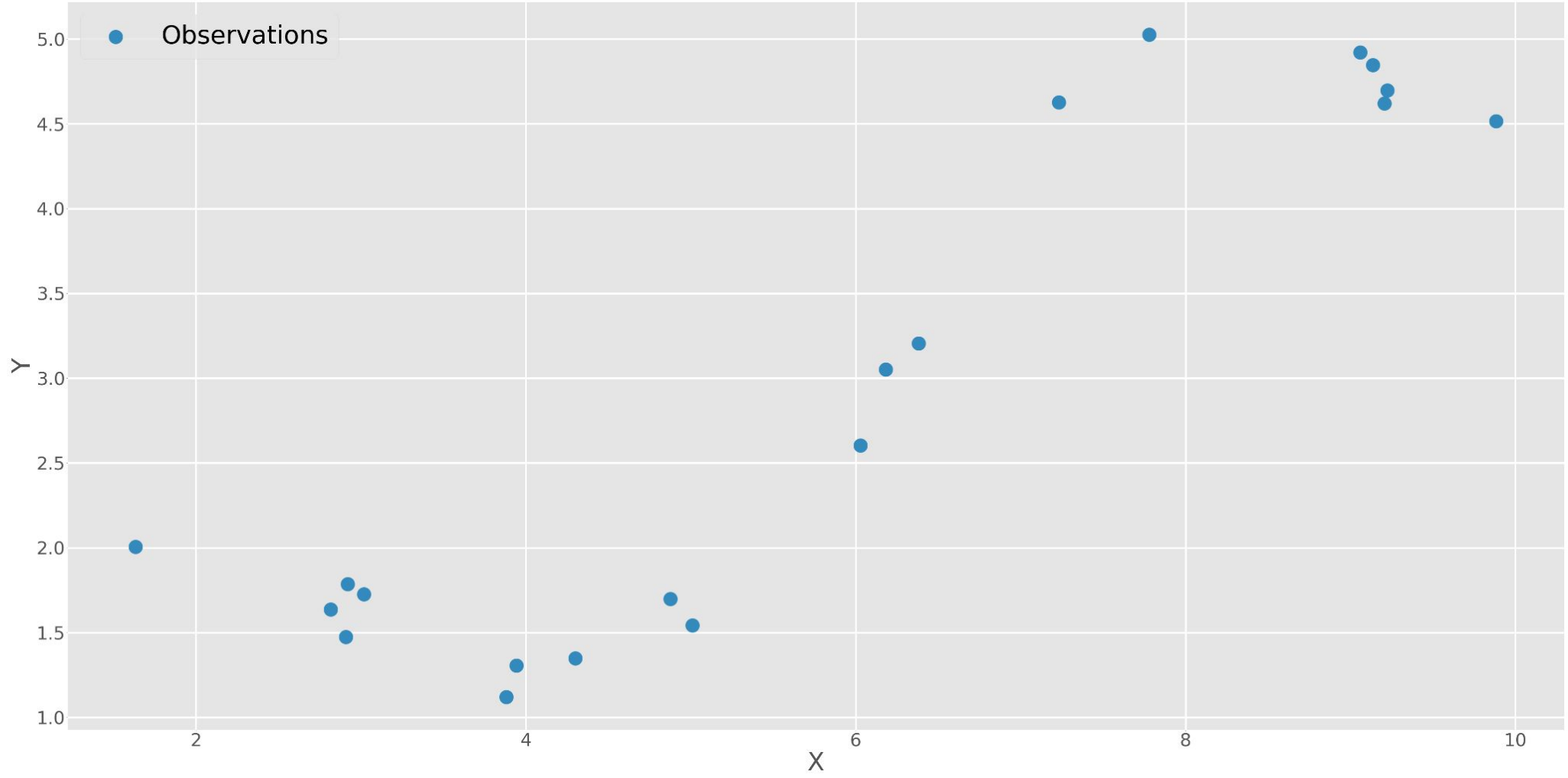
Number of observations

Standard Conformer was trained on 1000-1500 samples (daily), as higher frequency carries higher noise.

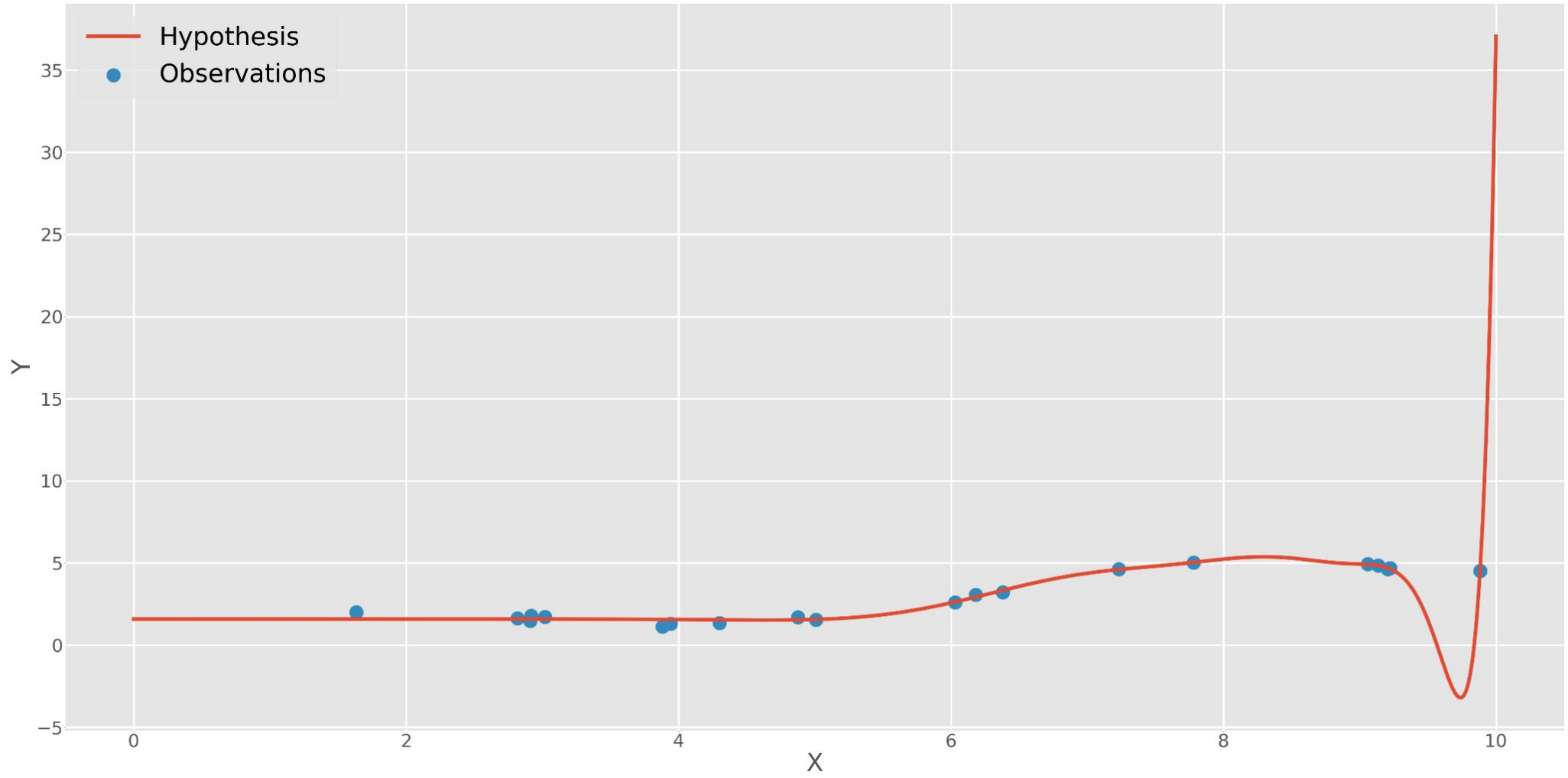
300`000 / 1’000 = 300

There are around 300 parameters for a single training sample.

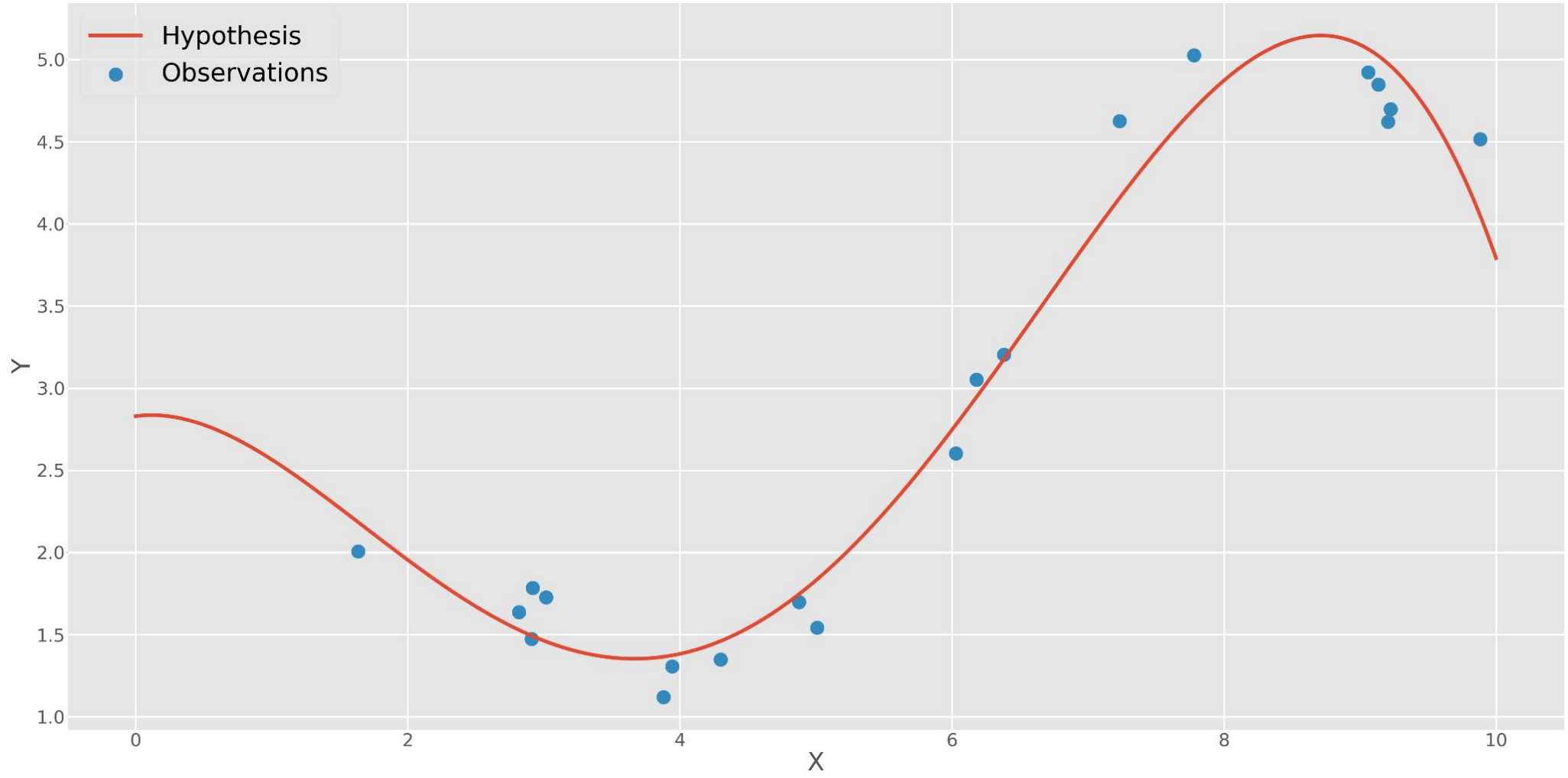
XIV-century philosopher won’t tell us what to do!



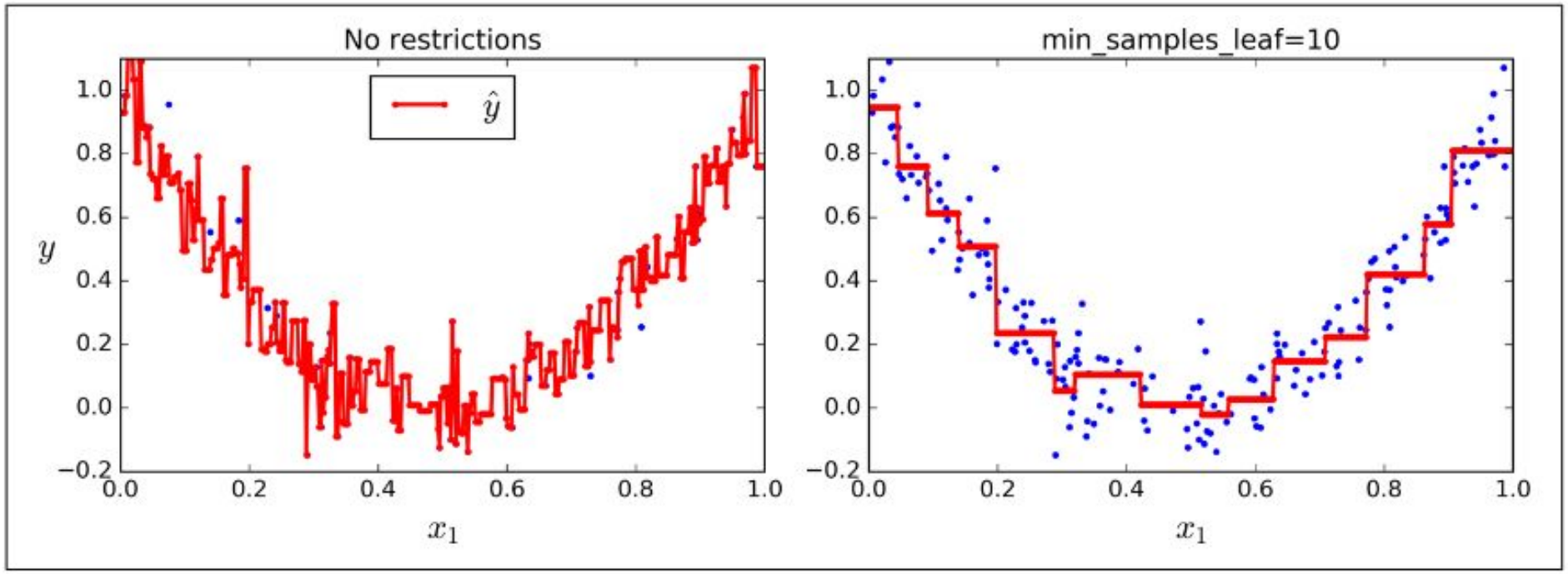
20 data points - only 25 parameters



20 data points - only 5 parameters

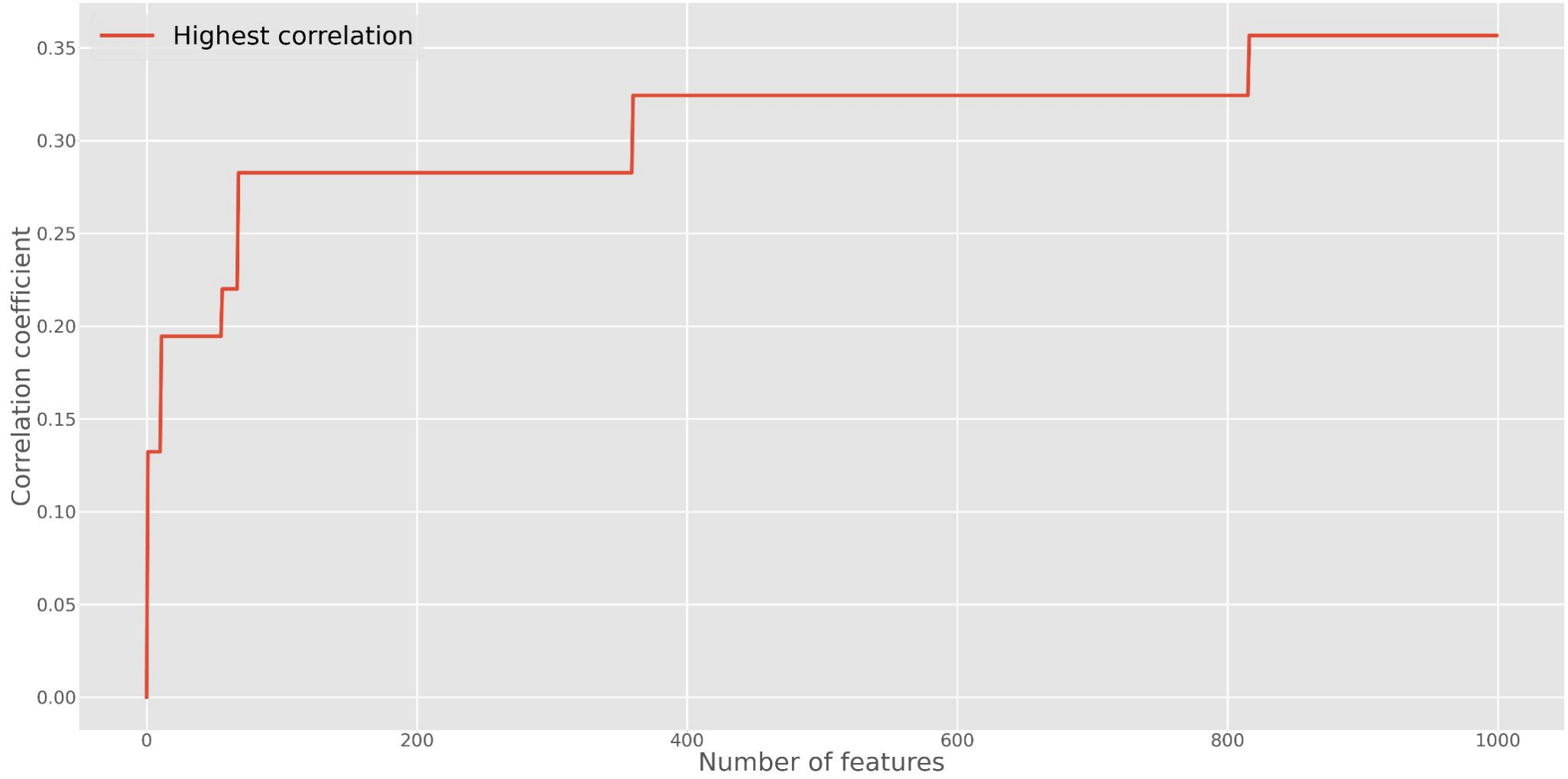


Bias vs. variance



Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow*

Is model complexity and number of samples all that there is to overfitting?



What can we realistically expect from a model?

Whom would you rather trust?

1. A person claiming that can predict the direction of price change tomorrow with 80% of accuracy?
2. A person claiming that can predict the direction of price change tomorrow with 52% of accuracy?

What is the usual accuracy to which neural network is trained?

Usually ~65%.

It is equivalent to person that has found a rule that works 65% of the time. Having a rule working so well, can we realistically expect it to describe a “hidden pattern” of the market? Or have we just found a spurious rule?

“Up to now, you have refuted everything which has been said. You have done nothing to point out the true Dharma to us.”

- a frustrated student to Huang Po

Blofeld, *Teachings of Huang Po*

Time for some real-life talk

- Problem of model complexity
- Problem of dataset size
- Problem of excessive number of features

Embrace simplicity!

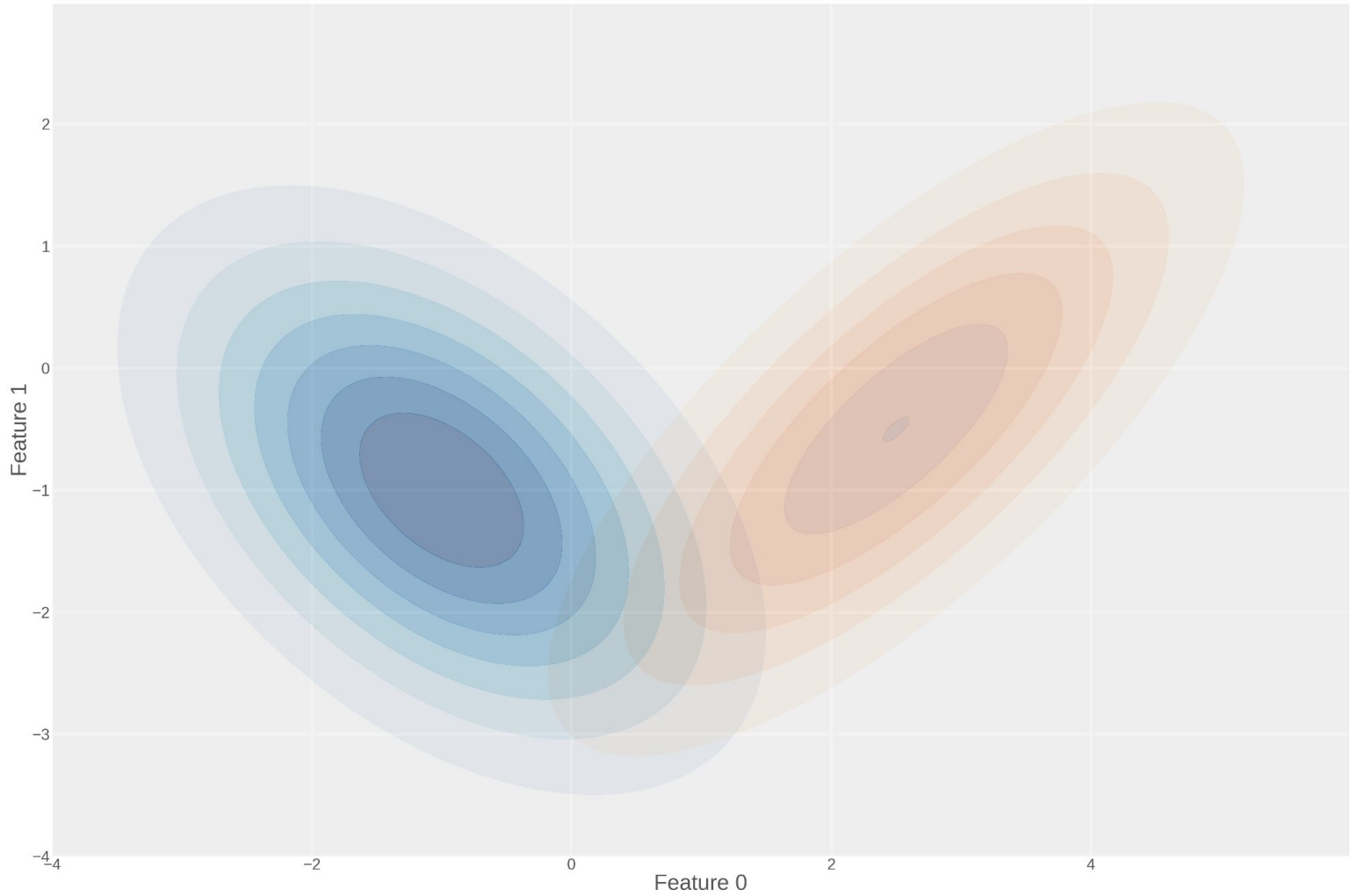
Linear classification model - f+1 parameters:

$$p_{\bar{a}}(\bar{x}) = \frac{1}{1 - e^{-\bar{x}\bar{a}}}$$

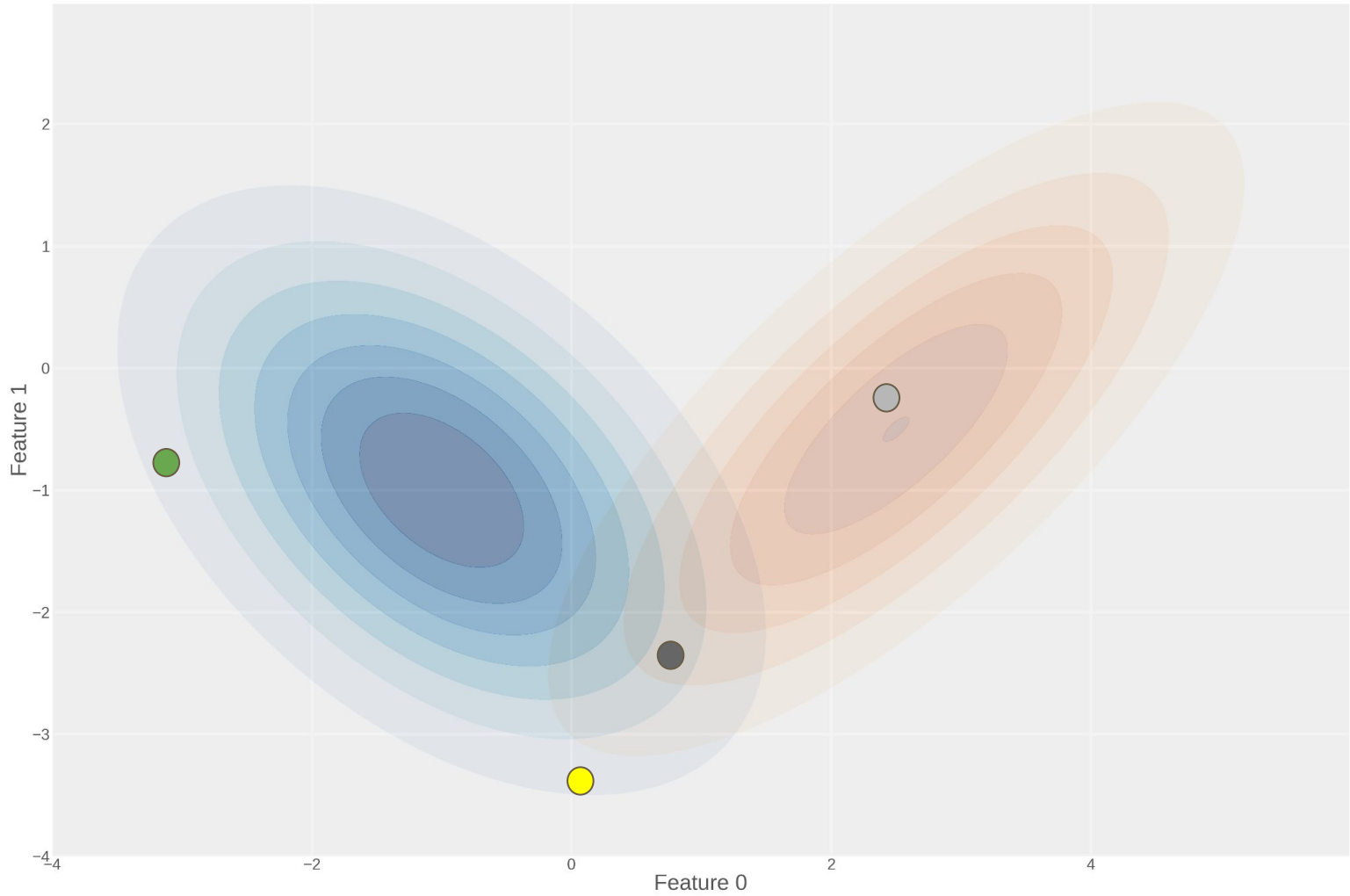
Bayesian classification model - 2*f*c parameters:

$$p(\bar{x}) = \prod_{i=1}^{k} \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\bar{\mu}_i - \bar{x}_i}{2\sigma}\right)^2}$$

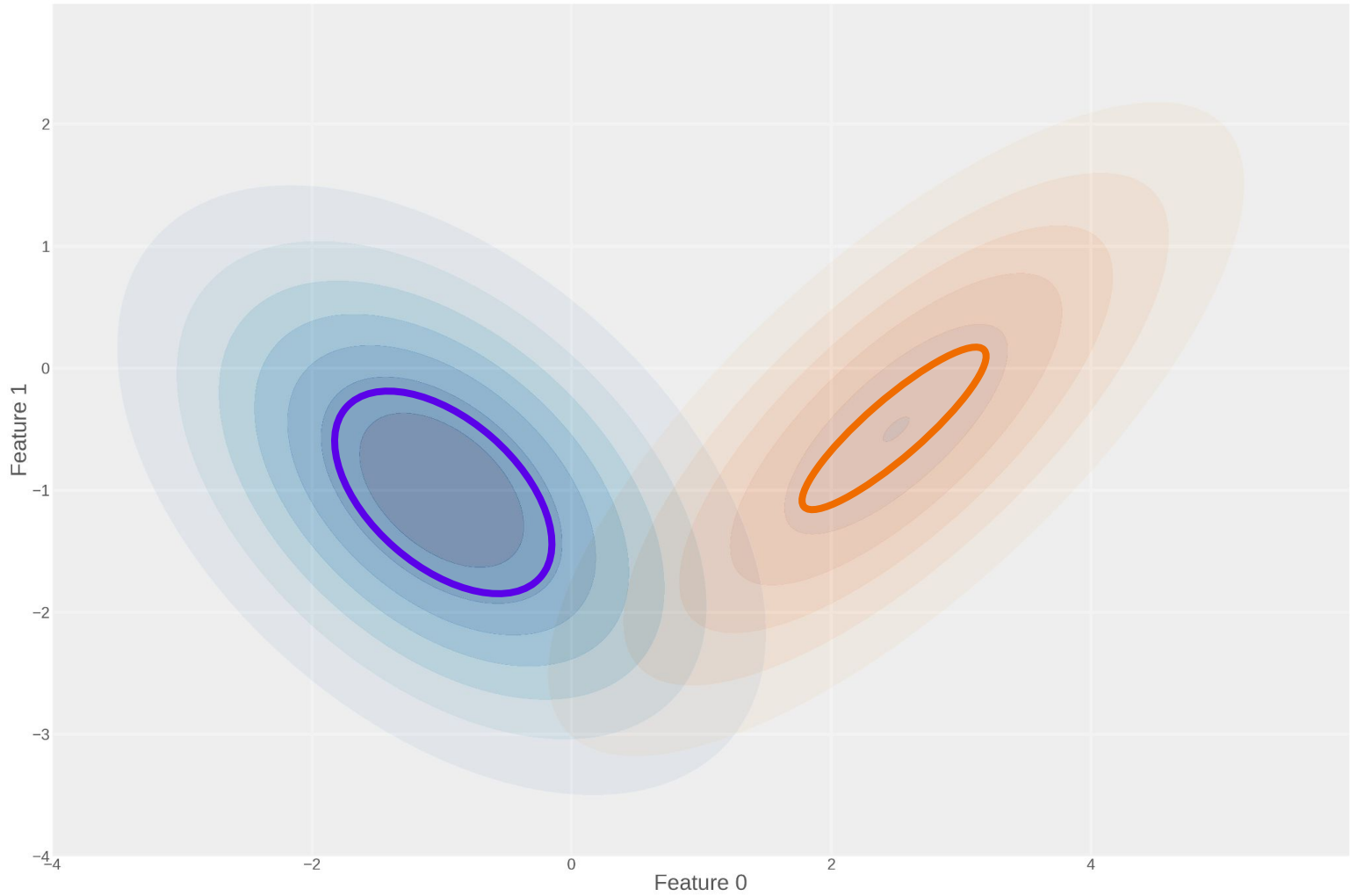
Bayesian classifier / Gaussian process classifier



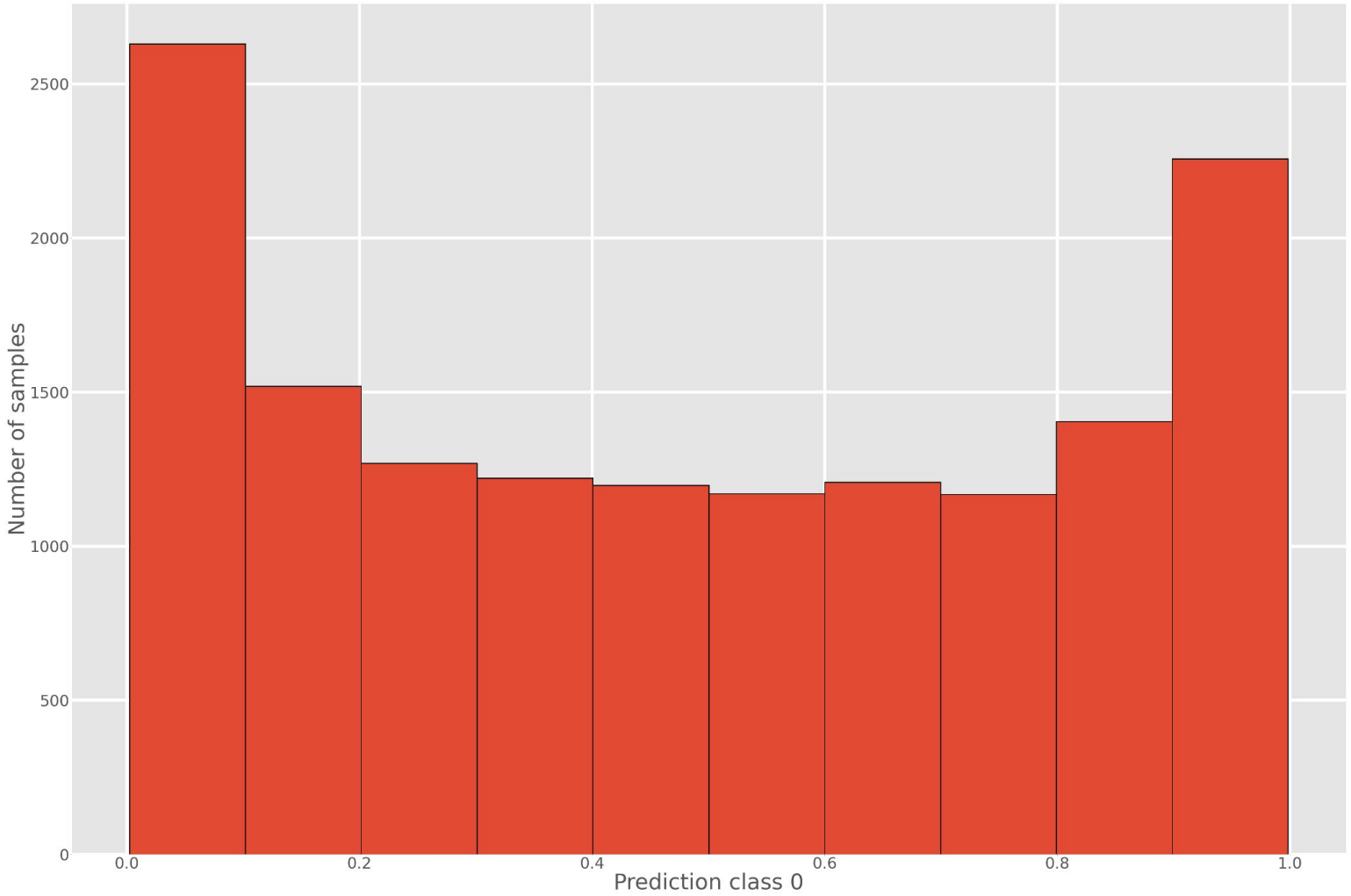
Bayesian classifier / Gaussian process classifier



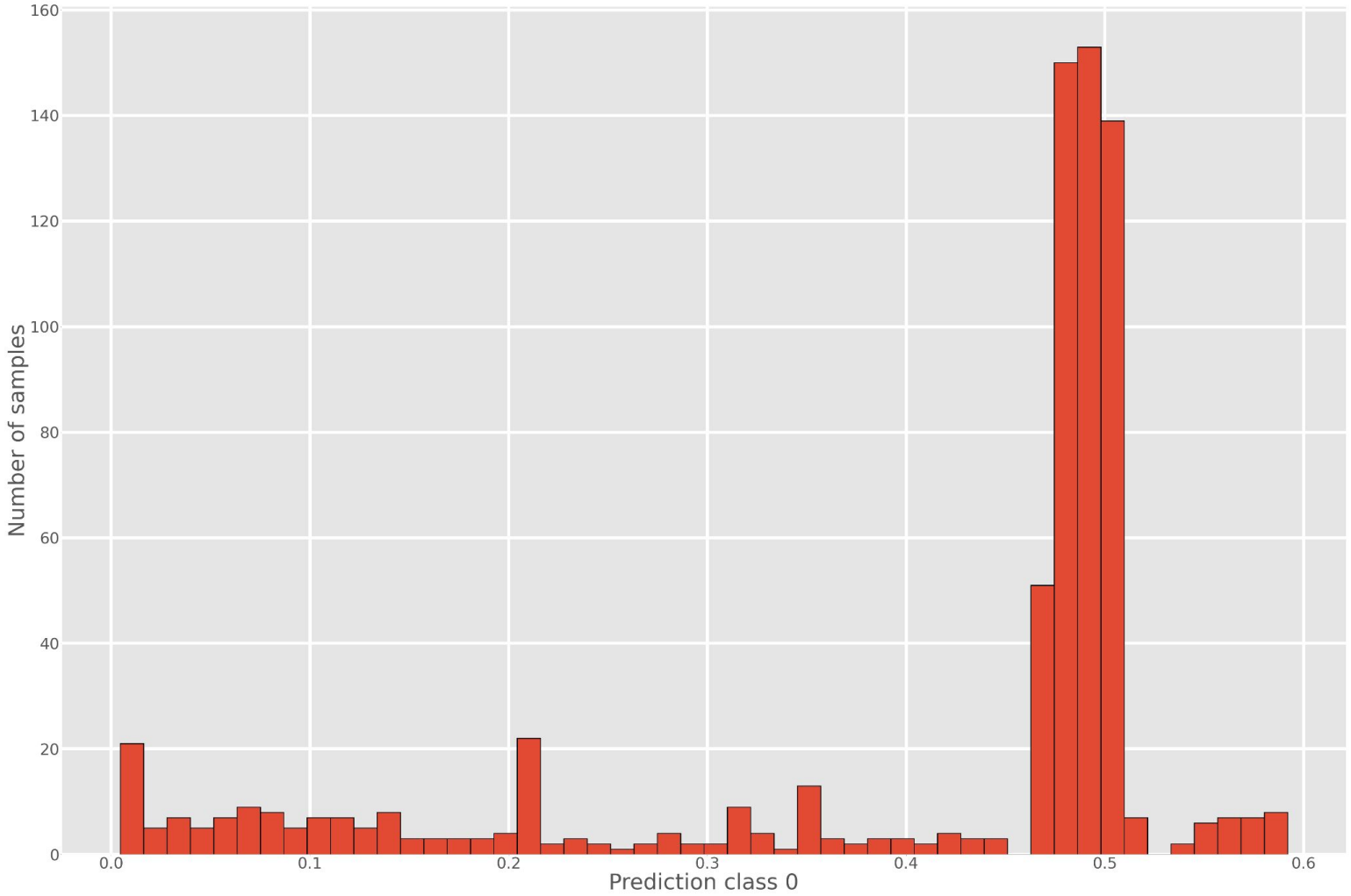
Bayesian classifier / Gaussian process classifier



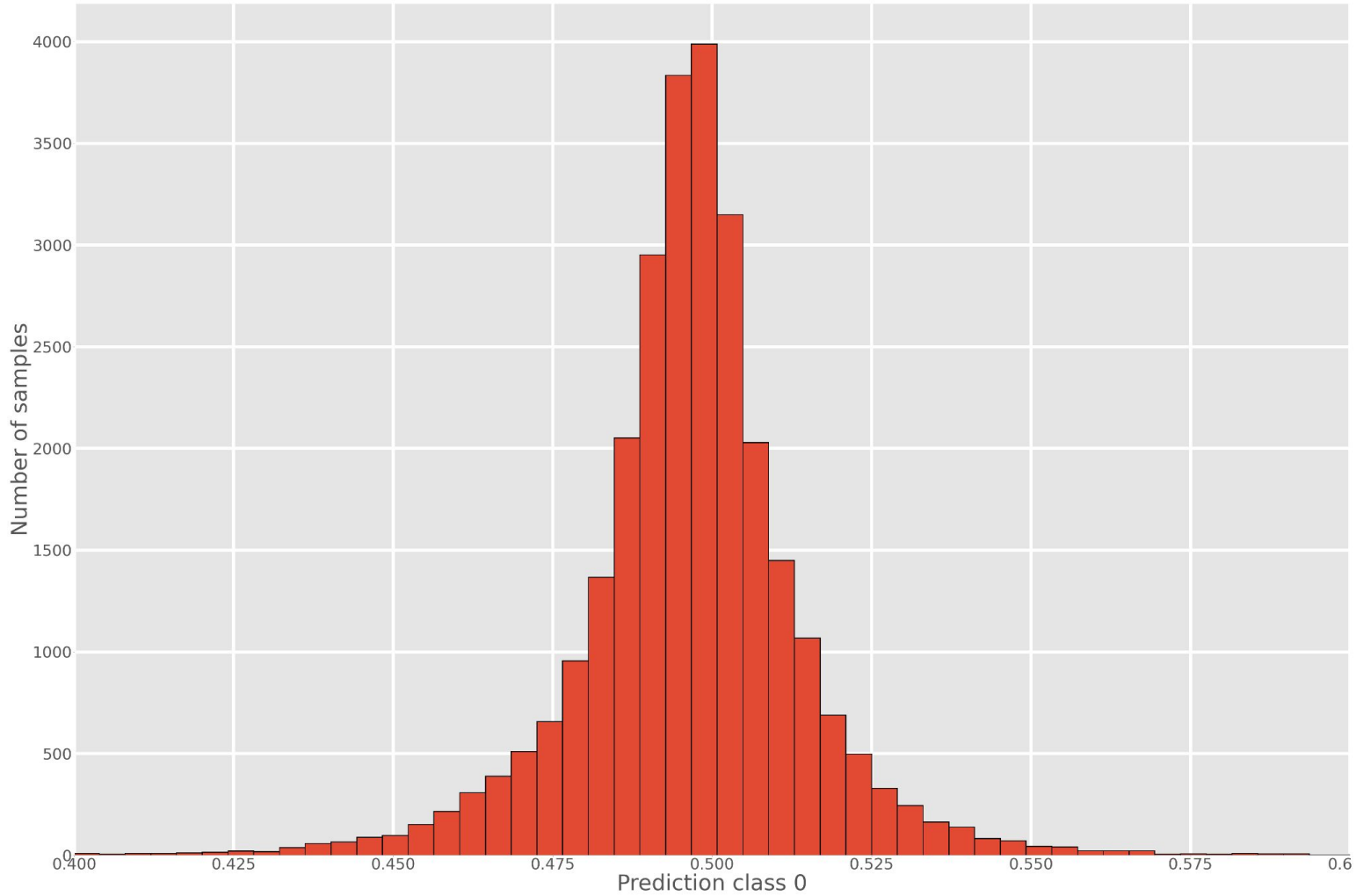
Neural network predictions ≠ probability



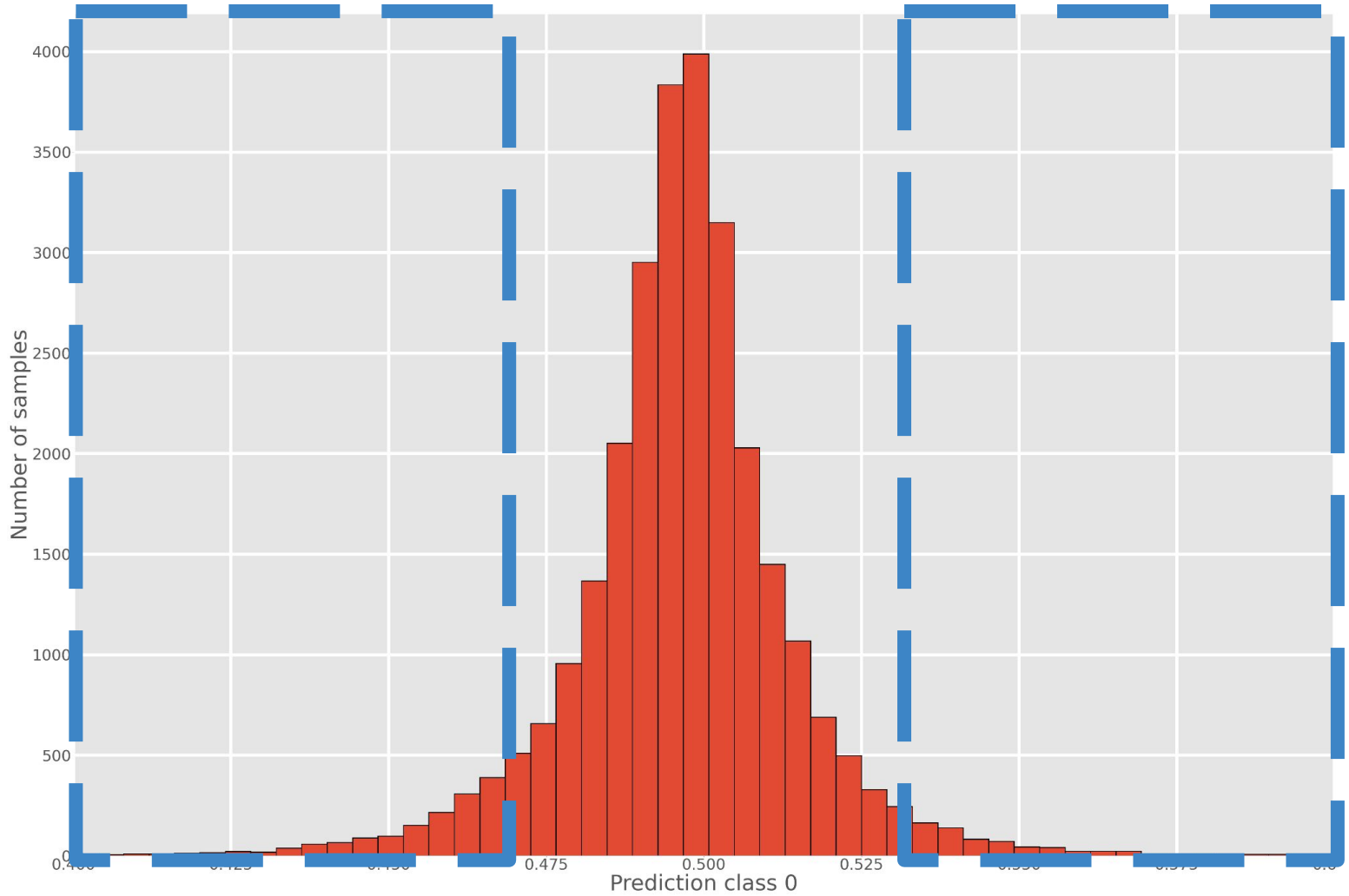
Neural network predictions ≠ probability



Statistical models predictions



Statistical models predictions

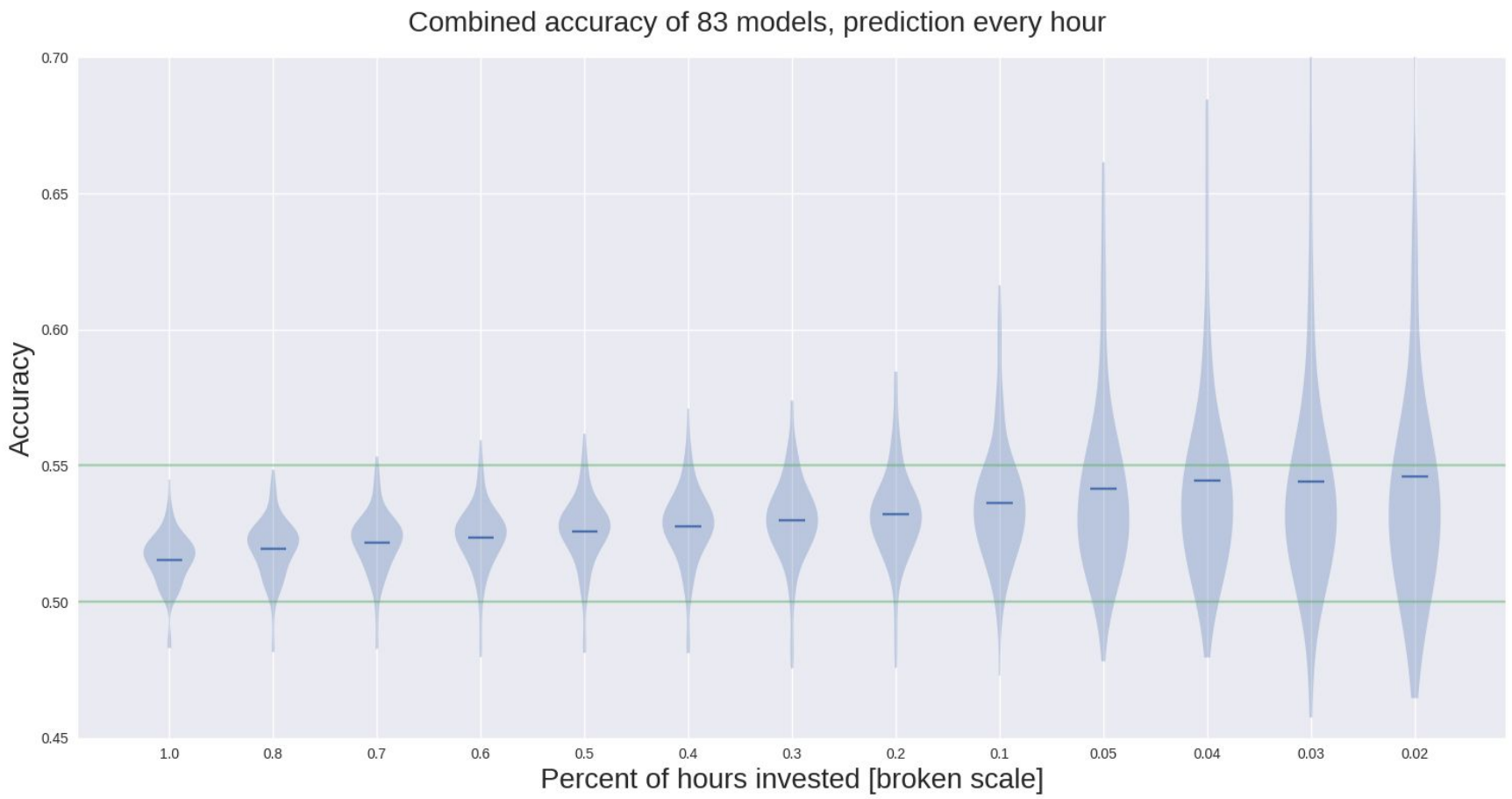


What we have done

- Shrunk the model (300’000 → 10s of parameters)
- Shrunk the feature space (10-100 → 4)
- Based our decision on model’s confidence

Did it work?

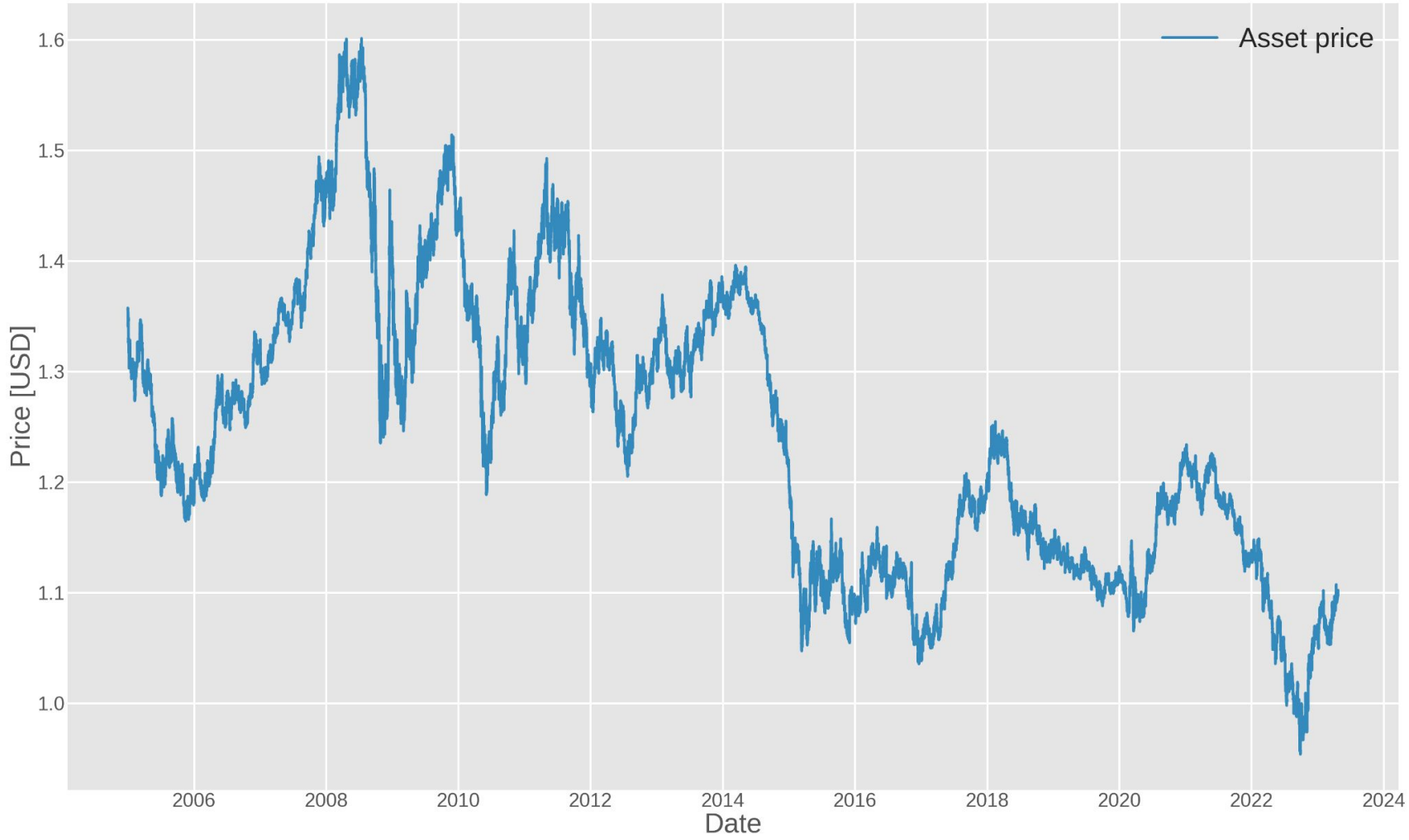
It does!


Combined accuracy of 83 models, prediction every hour
Accuracy
0.70
0.65
0.60
0.55
0.50
0.45
1.0
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.05
0.04
0.03
0.02
Percent of hours invested [broken scale]

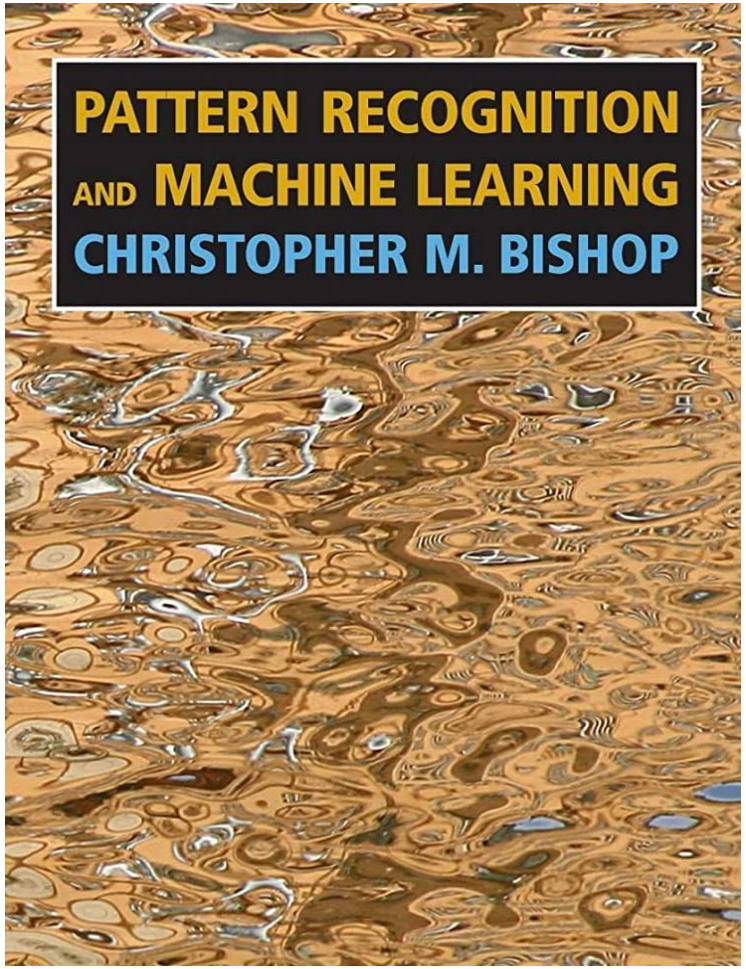
And it pays.

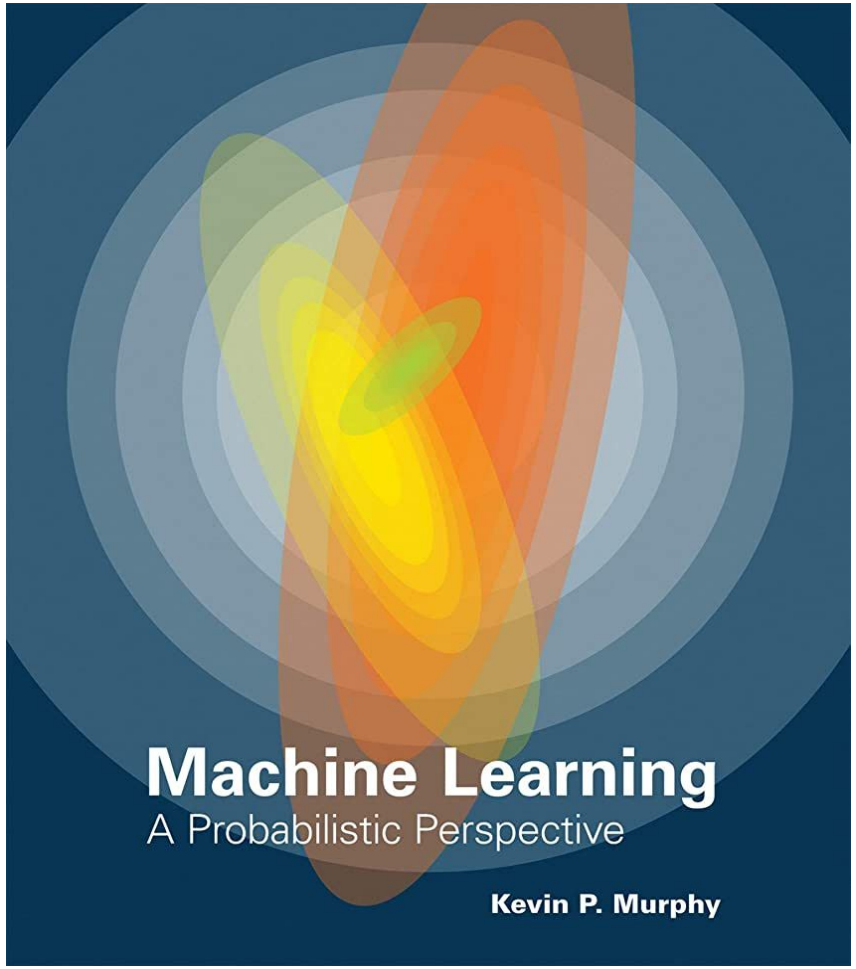
EURUSD profit and loss curve
600000
500000
400000
300000
200000
100000
0
Profit in dollars
2018
2019
2020
2021
2022
2023
Date

EURUSD


Asset price
Price [USD]
1.6
1.5
1.4
1.3
1.2
1.1
1.0
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024
Date

Interested?


PATTERN RECOGNITION
AND MACHINE LEARNING
CHRISTOPHER M. BISHOP


Machine Learning
A Probabilistic Perspective
Kevin P. Murphy

DATA SCIENCE SUMMIT ML EDITION

Thank you for watching!

Remember to rate the presentation and leave your questions in the section below.

mateusz.panasiuk@aiinvestments.pl
linkedin.com/in/mat-panasiuk

Academic Partners FUNDACJA