

Brzytwa Occama vs. rynek: Jak proste modele odpowiadają na złożone pytania

Mateusz Panasiuk
Chief Scientific Officer, AI Investments

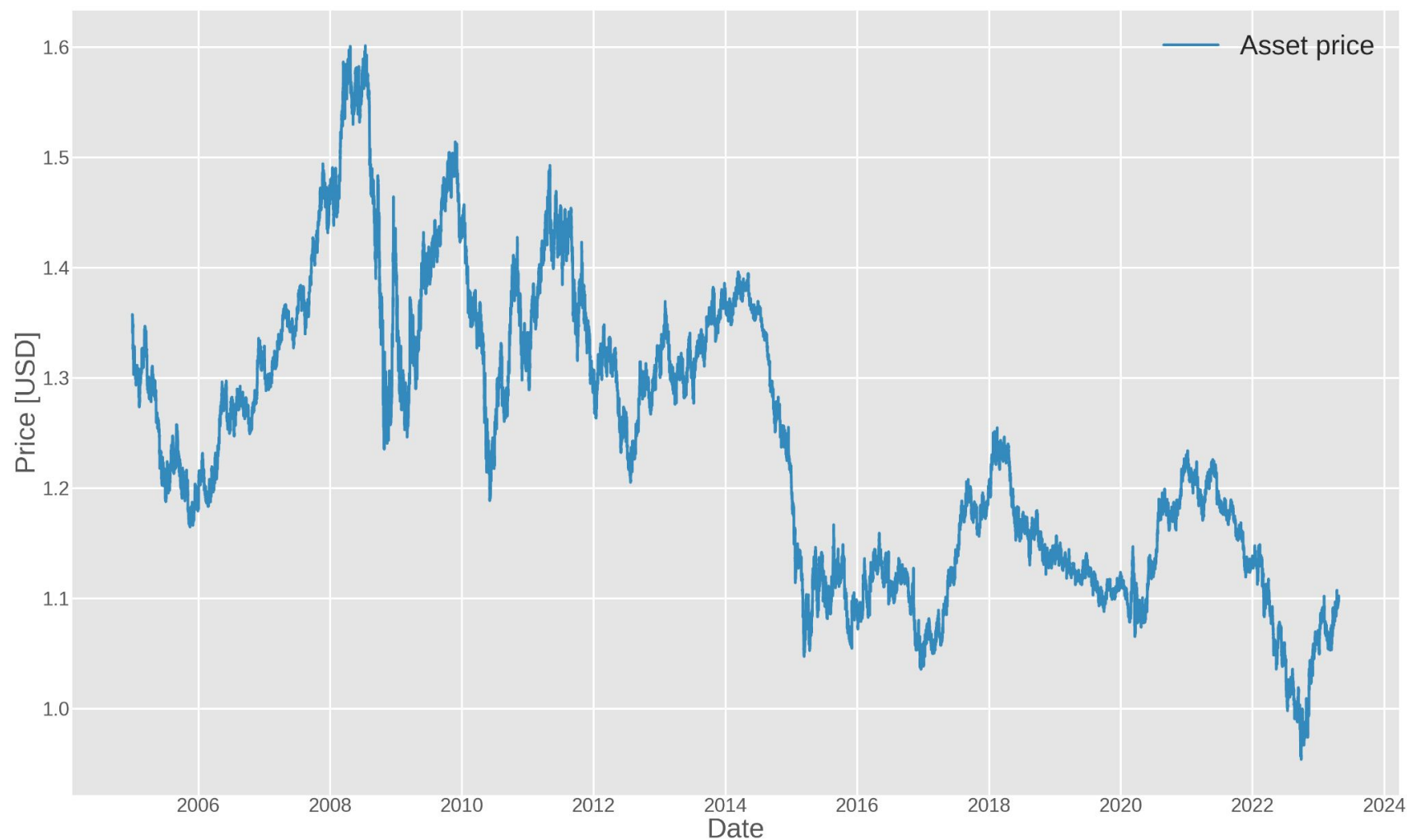


Krótki bieg przez modelowanie w finansach

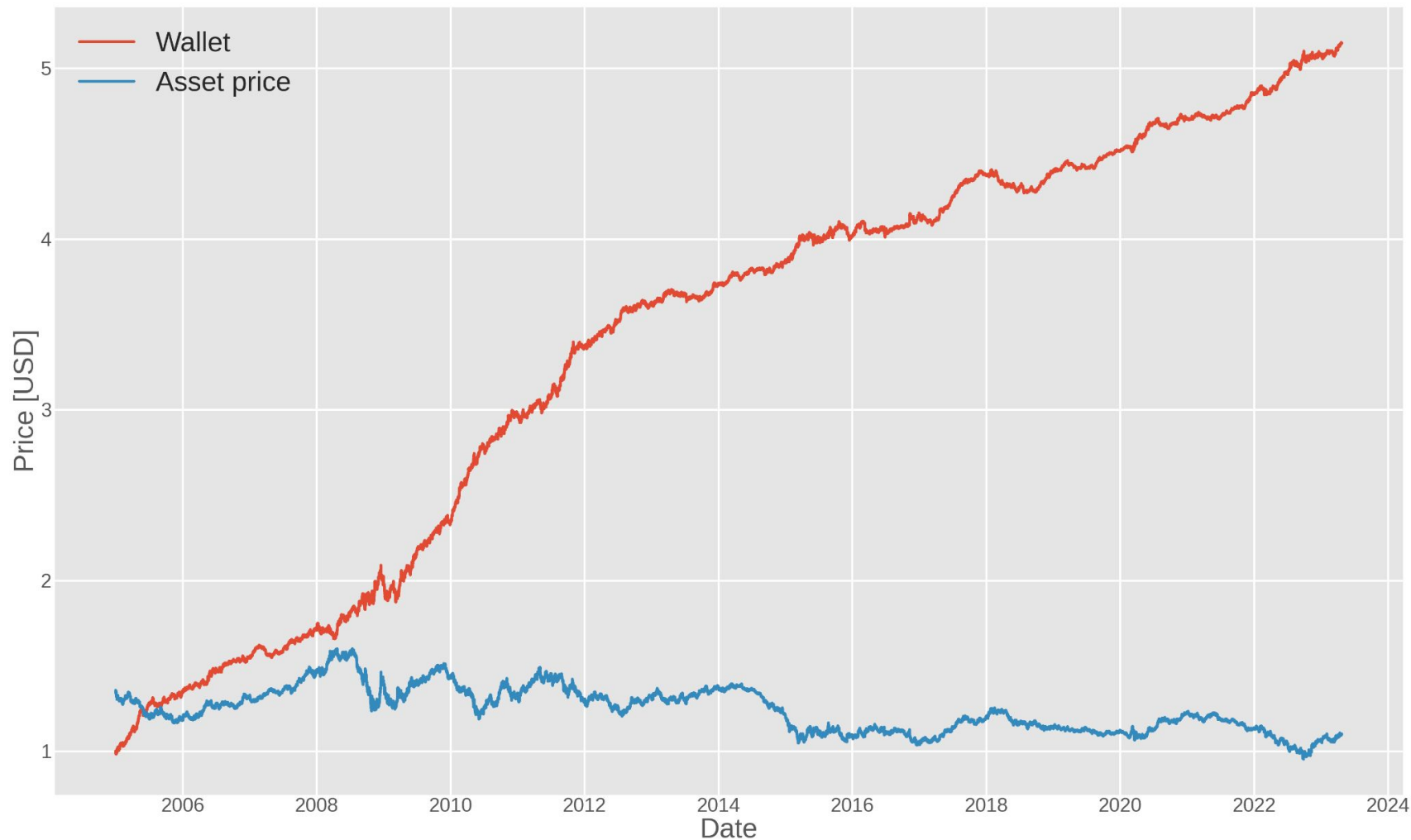
1. Wyzwania w przewidywaniach finansowych
2. Problem złożoności modelu
3. Rozwiązania problemu złożoności
4. Wyniki

Czy spekulacja jest trudna?

Czy spekulacja jest trudna?



Jak dobrzy musimy być żeby pobić rynek?



Ja rozwiązać złożony
problem?

Najnowsze sieci neuronowe

Najnowsze sieci
neuronowe

Więcej cech

Najnowsze sieci
neuronowe

Inżynieria cech
Więcej cech

Najnowsze sieci
Większe sieci
neuronowe
neuronowe
Inżynieria cech
Więcej cech

Najnowsze sieci
Większe sieci
neuronowe GPT
neuronowe
Inżynieria cech
Więcej cech

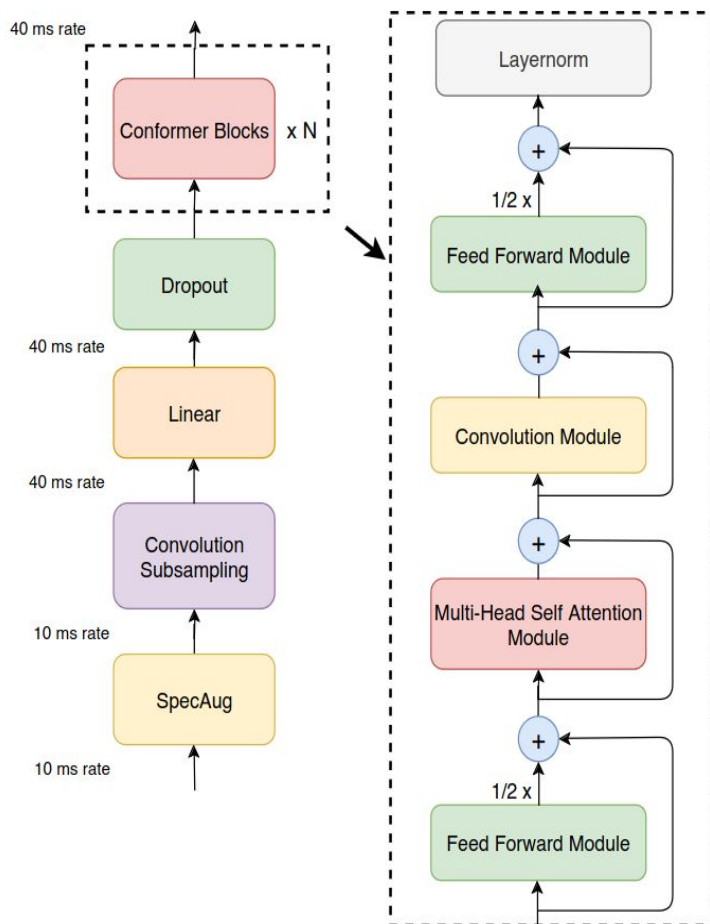
Brzytwa Williama



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

“Obiekty powinny być mnożone tylko przy wystąpieniu konieczności.”

Wiodące architektury



```
{"d_model": 32,  
"N": 6,  
"heads": 8,  
"features_count": 24,  
"length": 20}
```

300`000 parameters

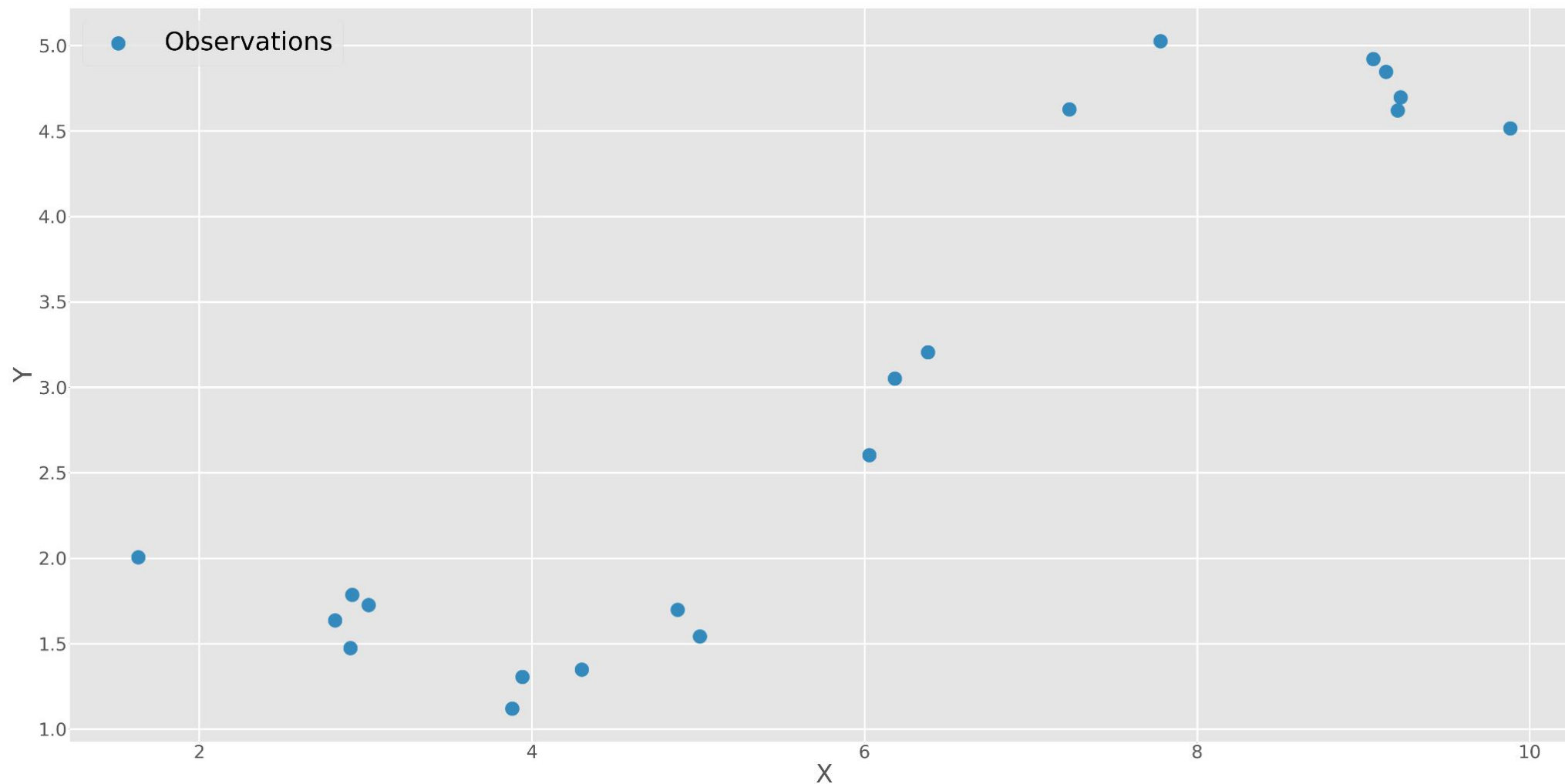
Ilość obserwacji

Standardowy Conformer był trenowany na 1000-1500 obserwacjach (daily), gdyż wyższe częstotliwości niosły szum.

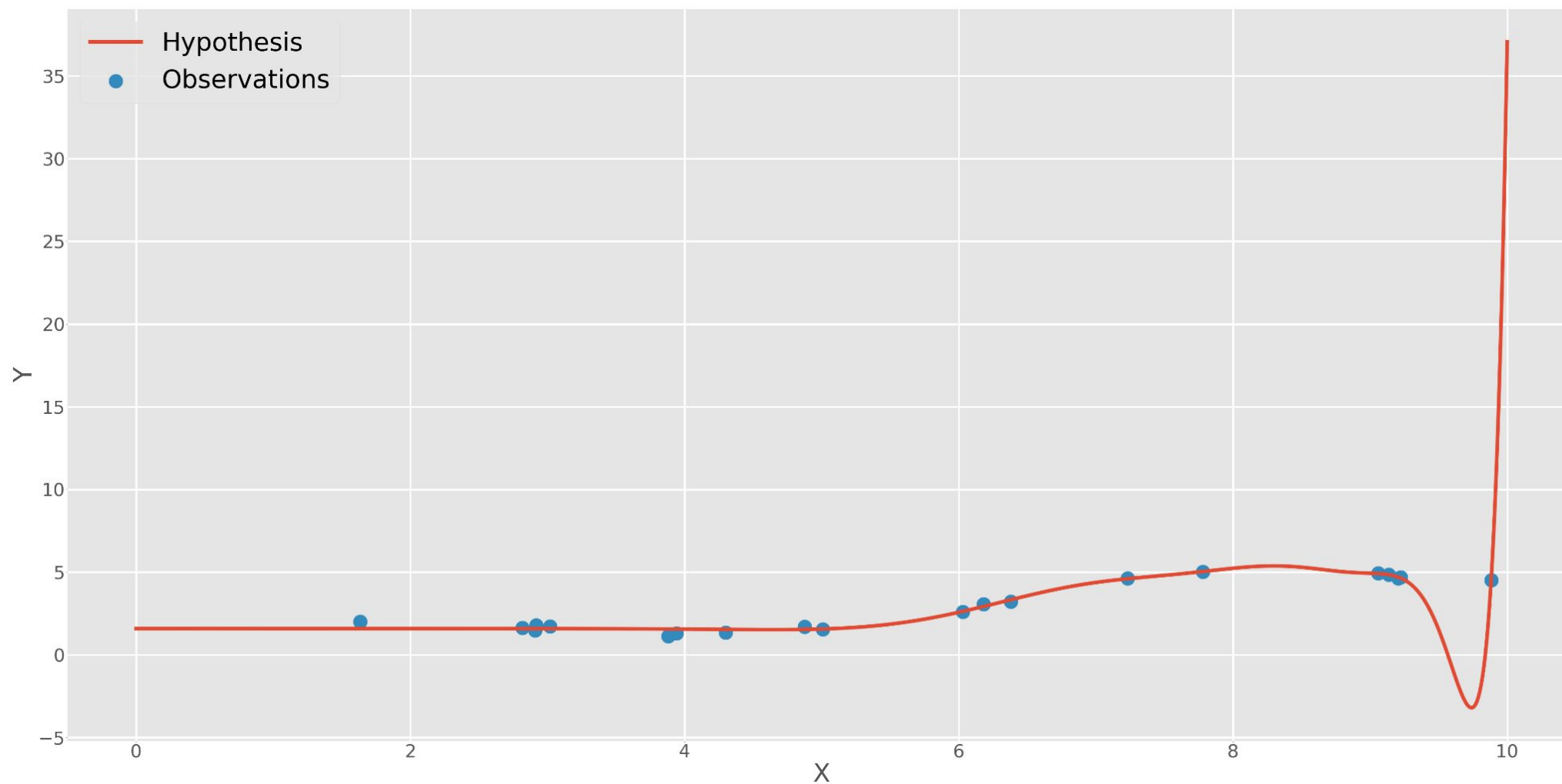
$$300\,000 / 1\,000 = 300$$

Skutkowało to tym, że na pojedynczą obserwację przypadało 300 parametrów.

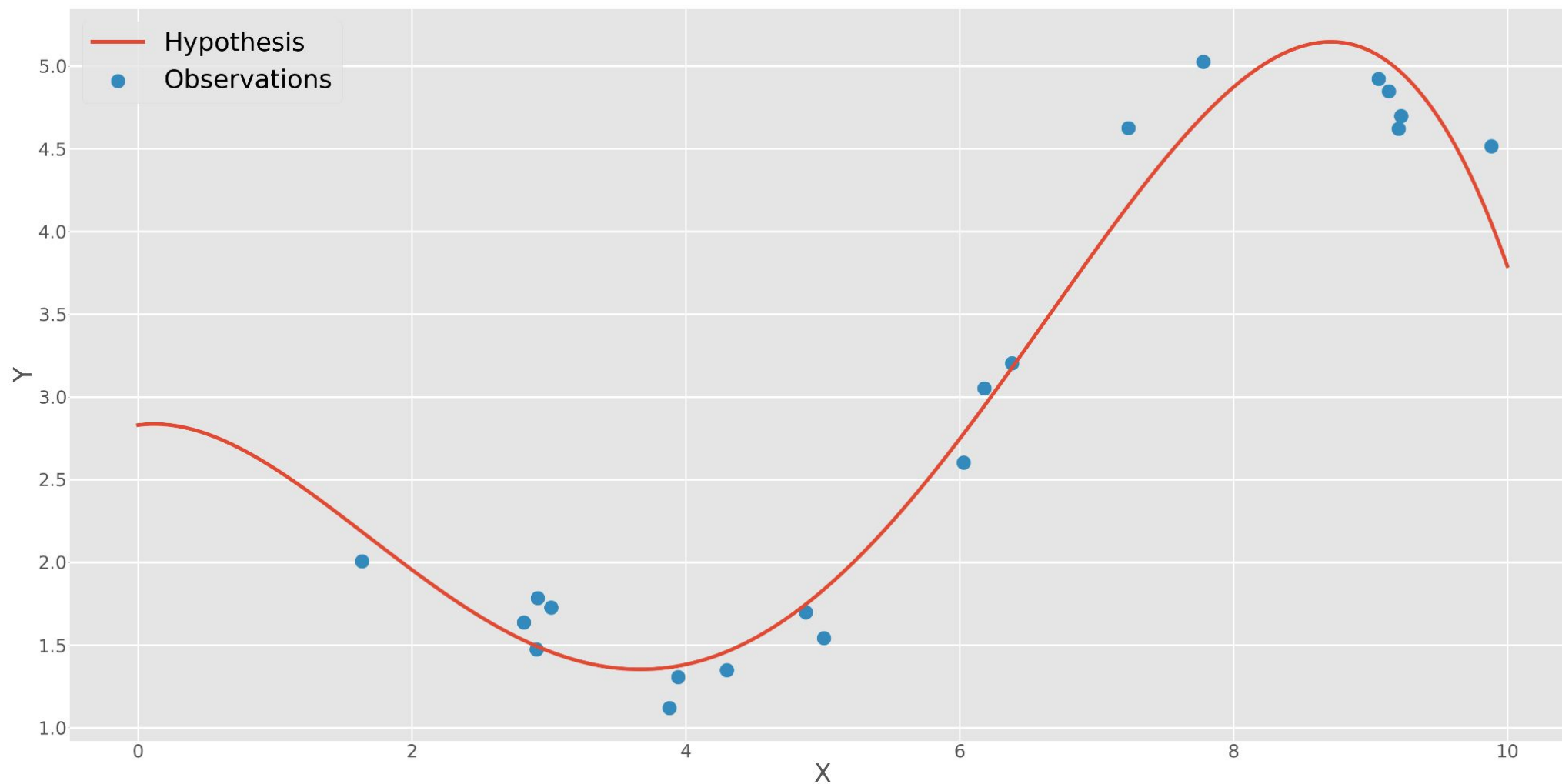
XIV-wieczny filozof nie będzie nam mówił co robić!



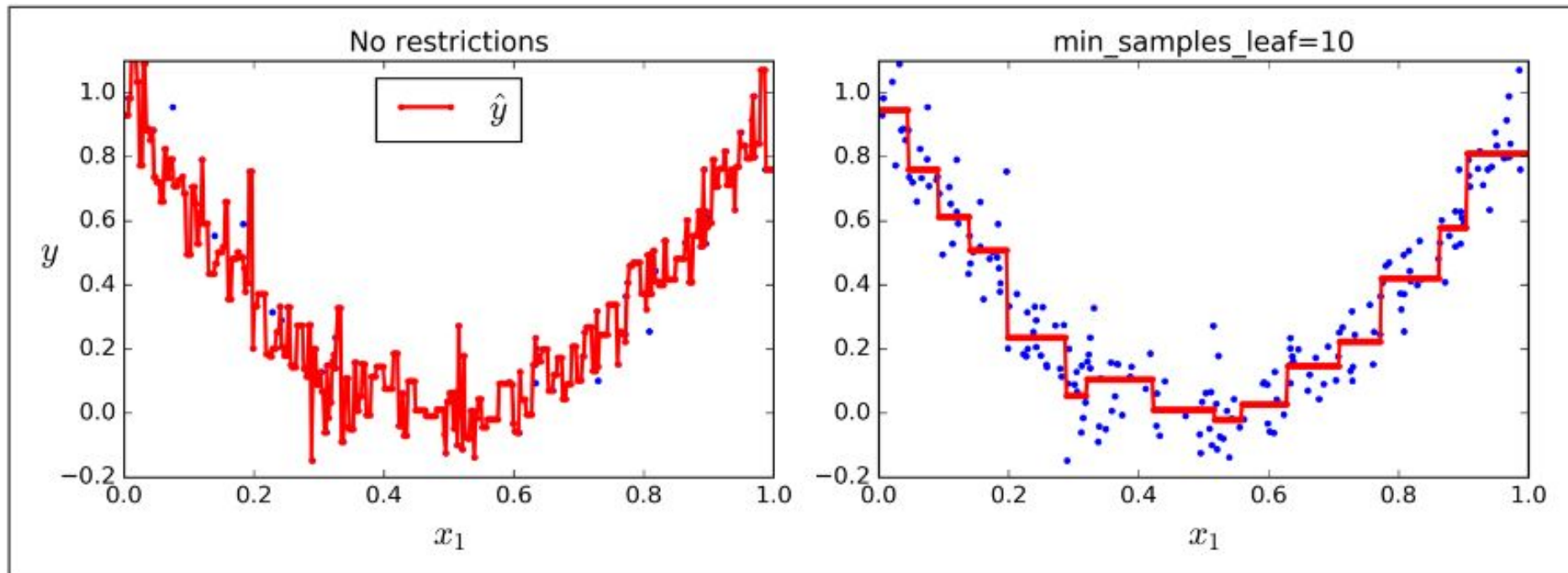
20 obserwacji - tylko 25 parametrów



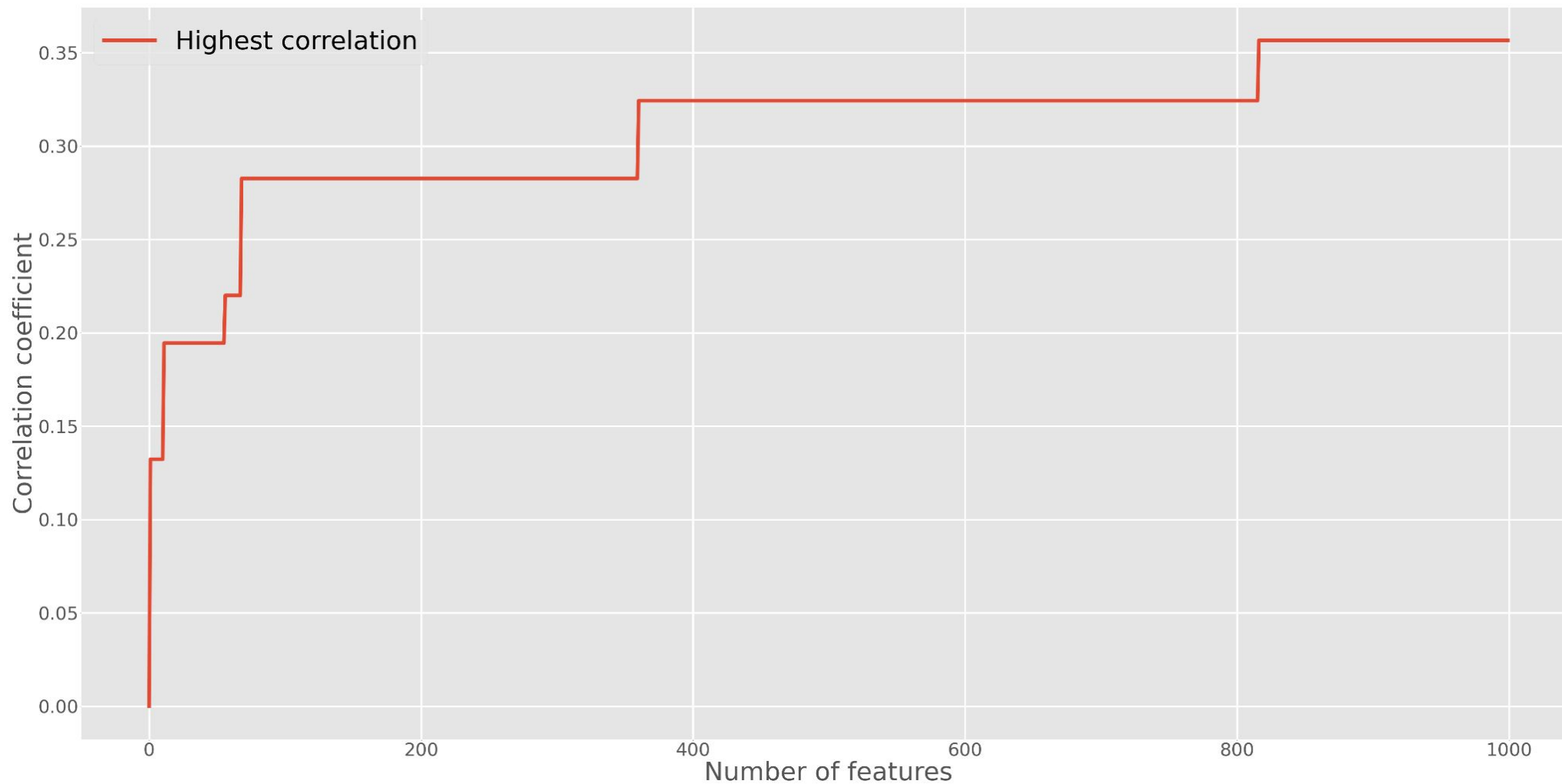
20 obserwacji - tylko 5 parametrów



Bias vs. variance



Czy złożoność modelu i ilość próbek to jedyne o czym powinniśmy myśleć przy overfittingu?



Czego możemy realistycznie oczekiwać od modelu?

Komu zdecydowałibyście się zaufać?

1. Osobie twierdzącej, że może przewidzieć kierunek zmiany ceny jutro z 80% dokładnością?
2. Osobie twierdzącej, że może przewidzieć kierunek zmiany ceny jutro z 52% dokładnością?

Jaka była dokładność do której trenowaliśmy sieci?

Zazwyczaj ~65%.

Jest to ekwiwalent osoby która znalazła regułę która działa w 65% przypadków. Rozważając regułę działającą tak sprawnie - czy jesteśmy w stanie oczekiwać, że faktycznie opisuje one "ukryte wzorce" rynku? Czy zwyczajnie znaleźliśmy nadmiarową korelację?

“Do tej pory odrzucałeś wszystko o czym rozmawialiśmy. Nie zrobiłeś niczego co wskazało by nam drogę prawdziwej Dharmy!”

- sfrustrowany uczeń zwracający się do Huang Po

Blofeld, *Teachings of Huang Po*

Czas na przykłady z prawdziwego życia

- Problem złożoności modelu
- Problem wielkości zbioru danych
- Problem nadmiernej ilości cech

Otwórzmy się na prostotę!

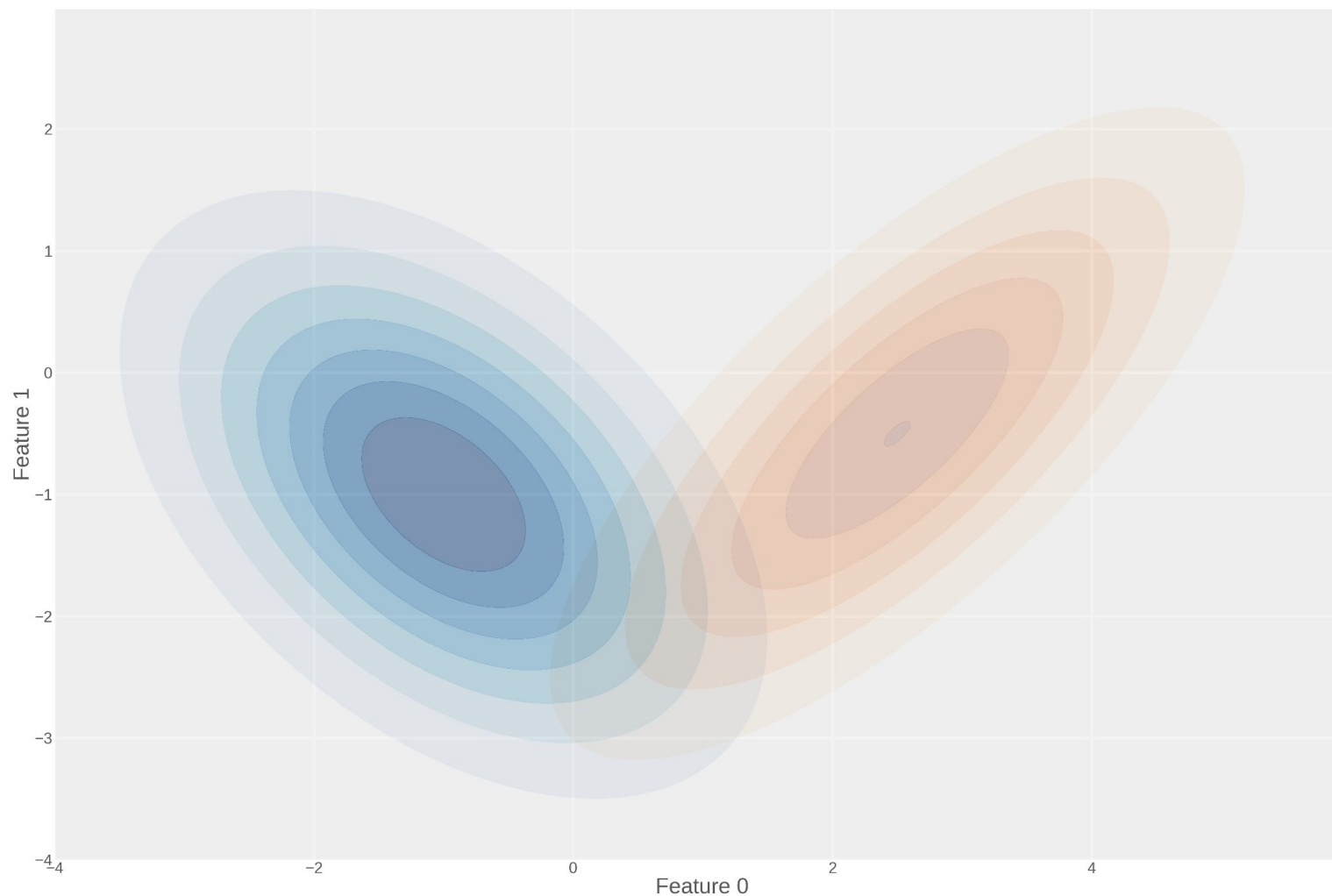
Model liniowy klasyfikacji - $f+1$ parametrów:

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

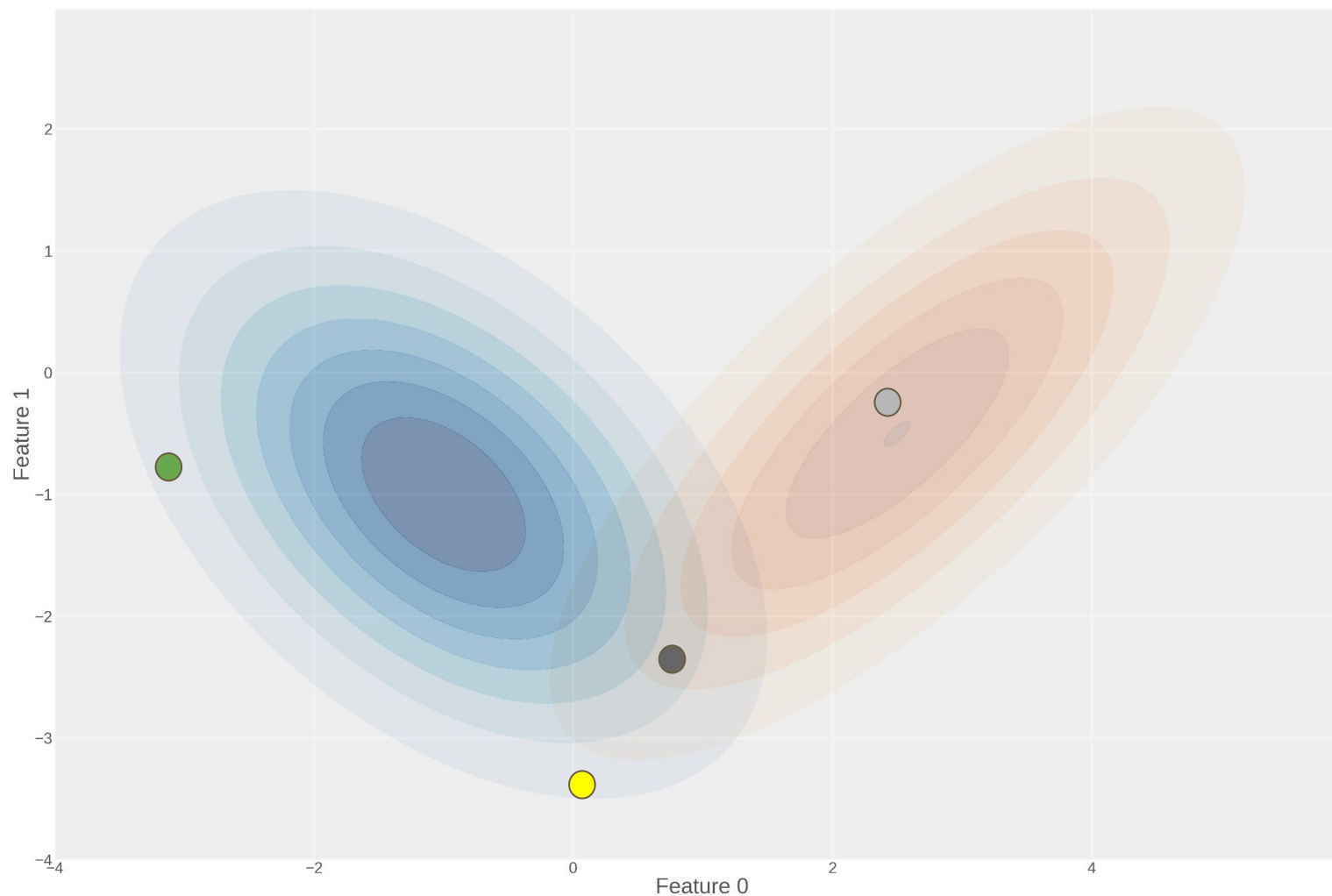
Model bayesowskiej klasyfikacji - $2*f*c$ parametrów:

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

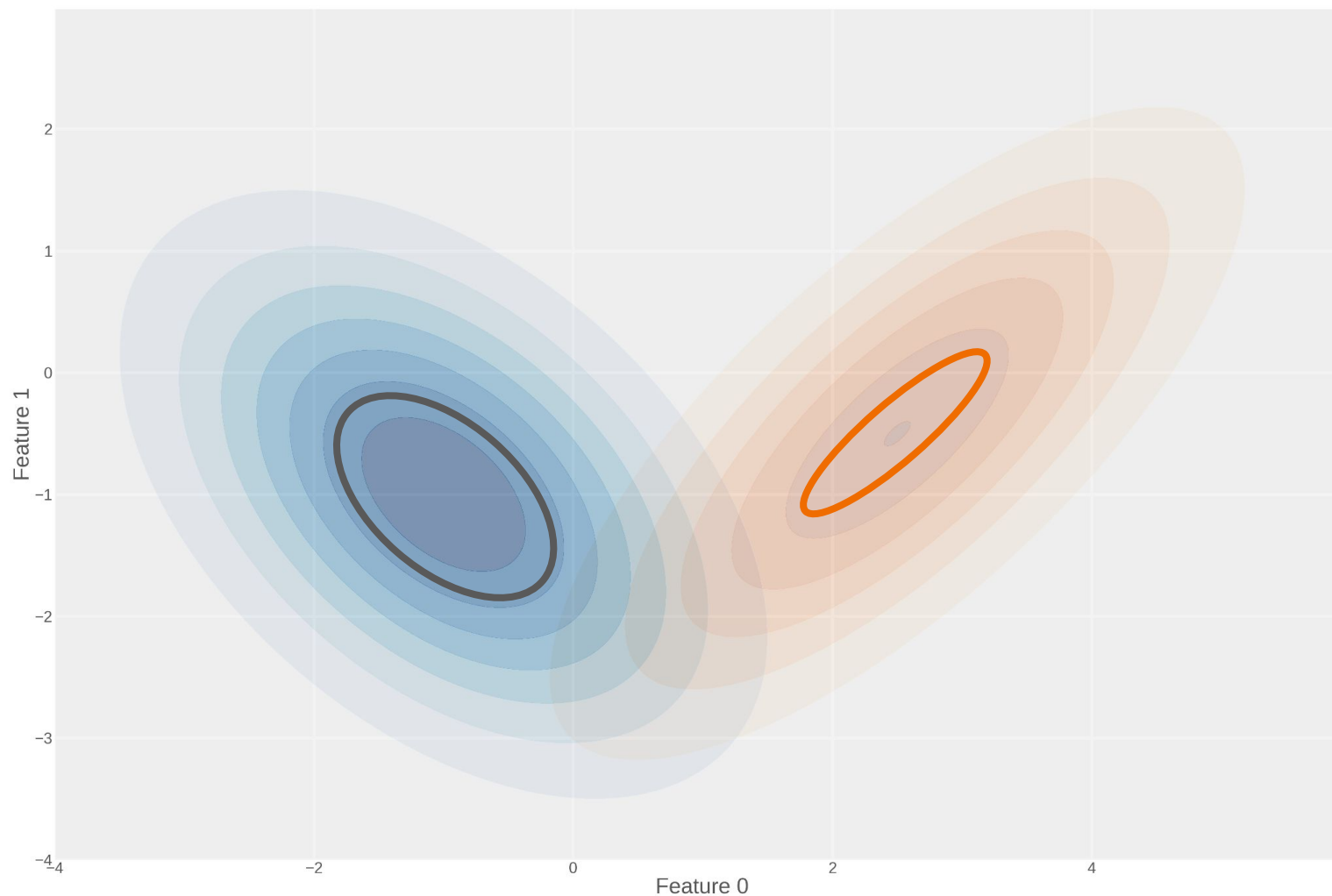
Klasyfikator bayesowski / Klasyfikator procesów Gaussa



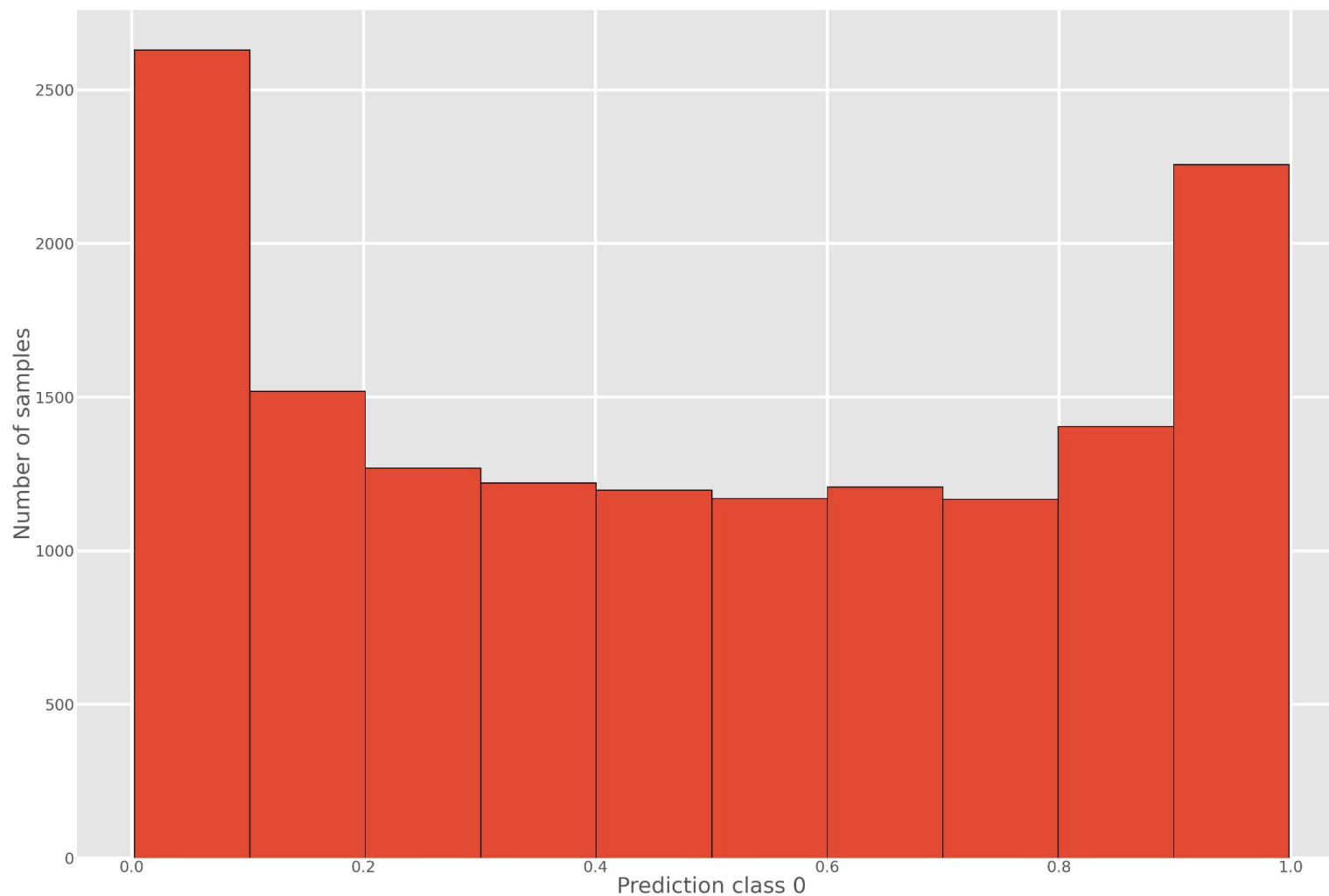
Klasyfikator bayesowski / Klasyfikator procesów Gaussa



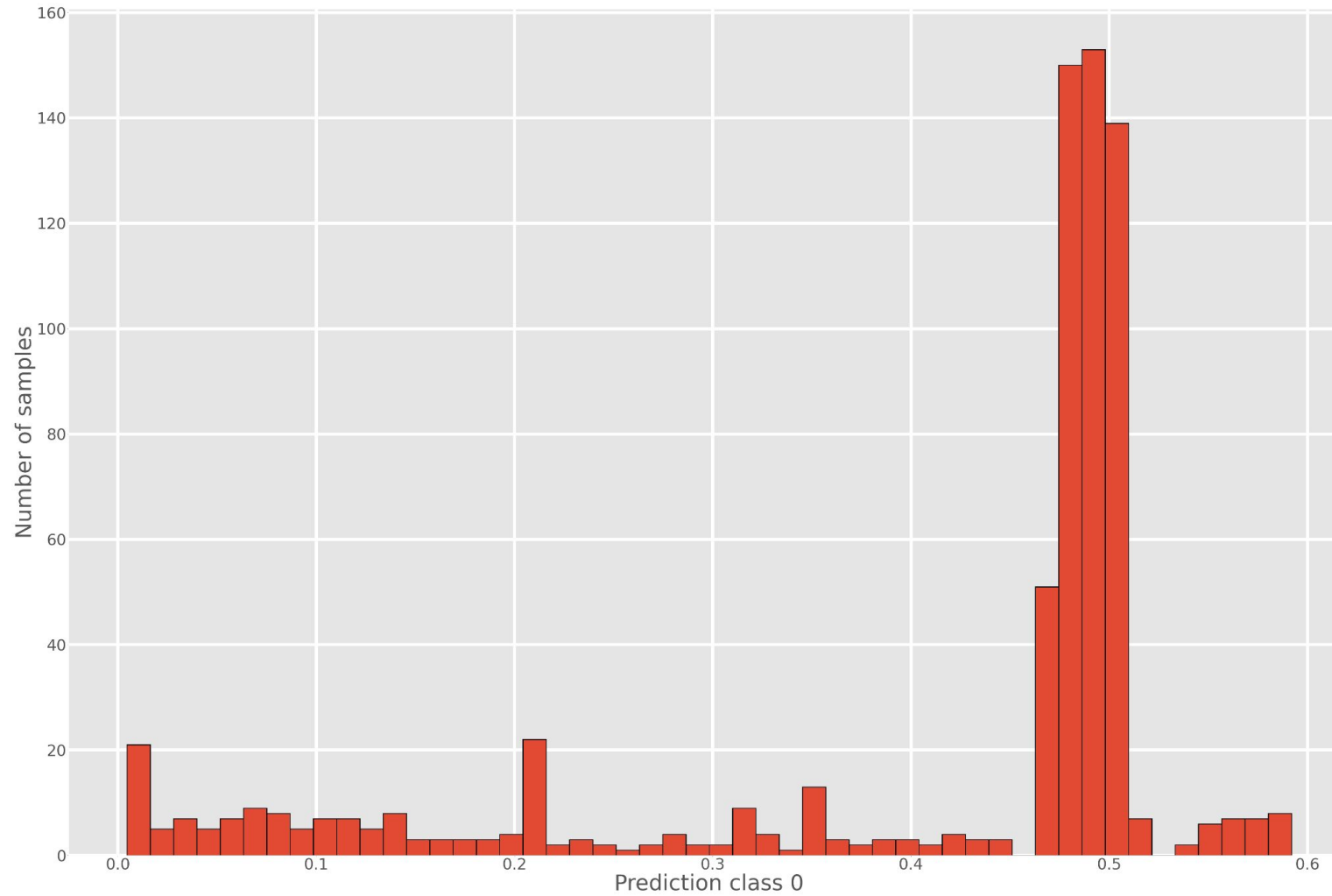
Klasyfikator bayesowski / Klasyfikator procesów Gaussa



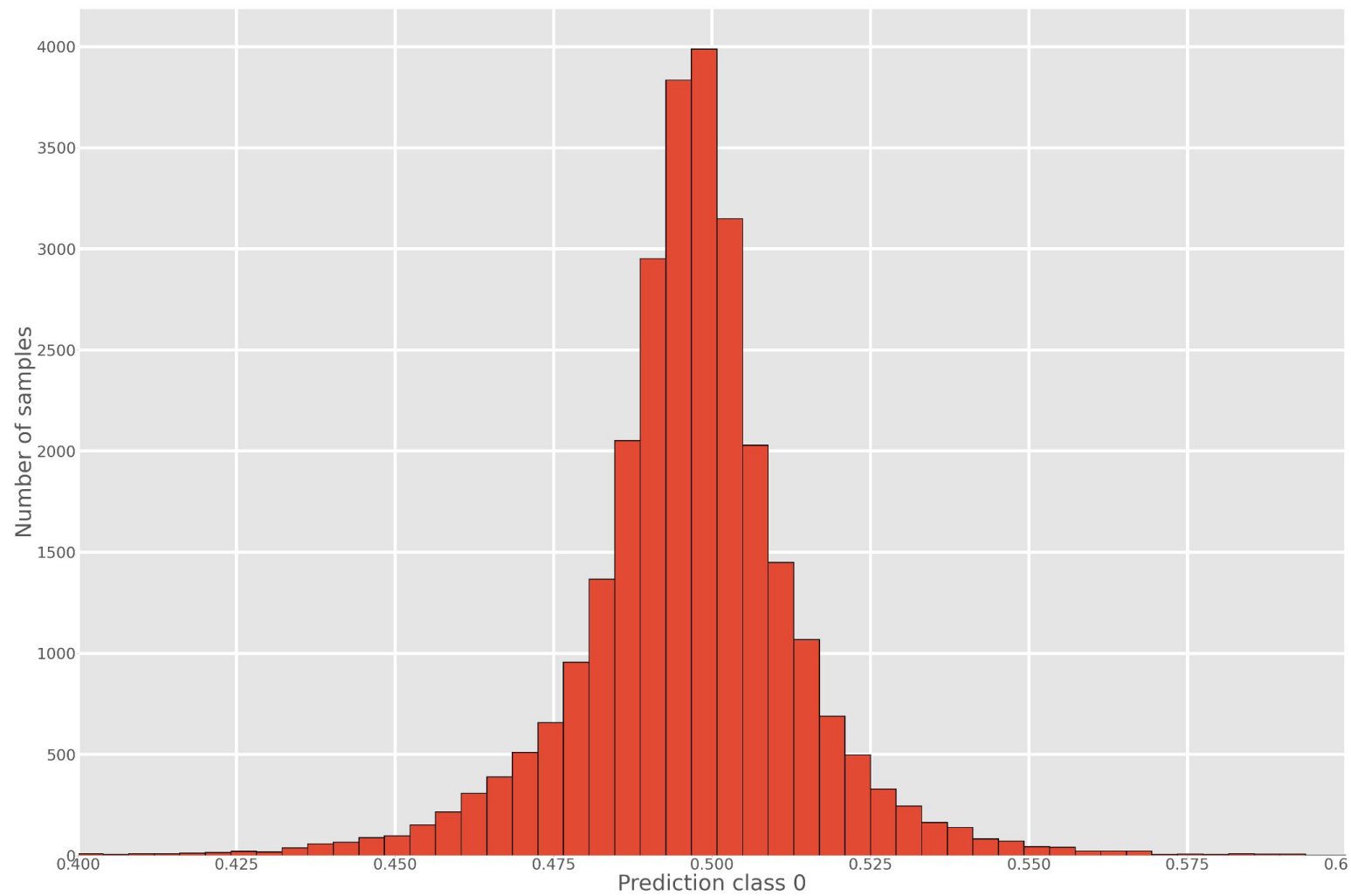
Predykcje sieci neuronowej ≠ prawdopodobieństwa



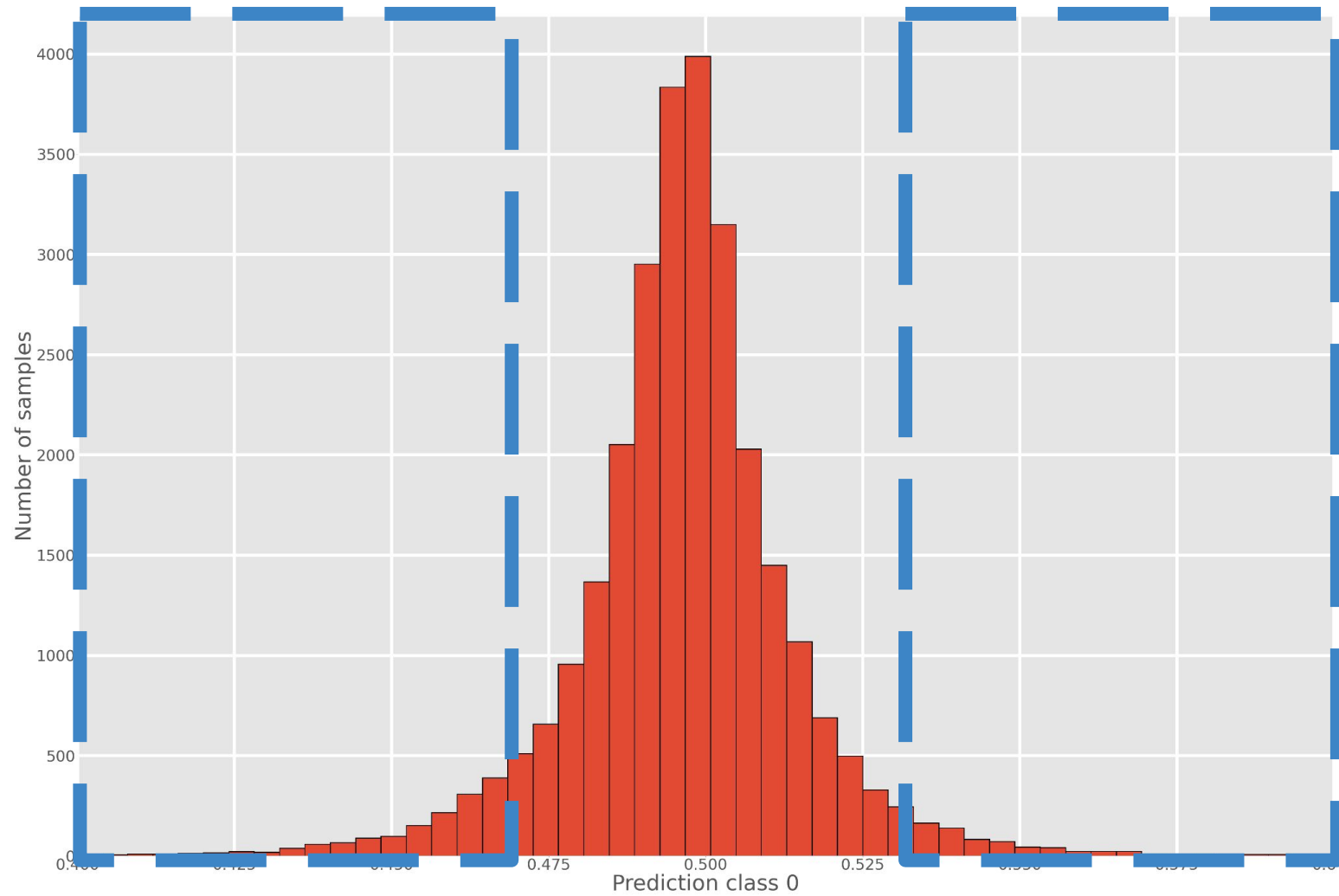
Predykcje sieci neuronowej \neq prawdopodobieństwa



Predykcje modeli statystycznych



Predykcje modeli statystycznych

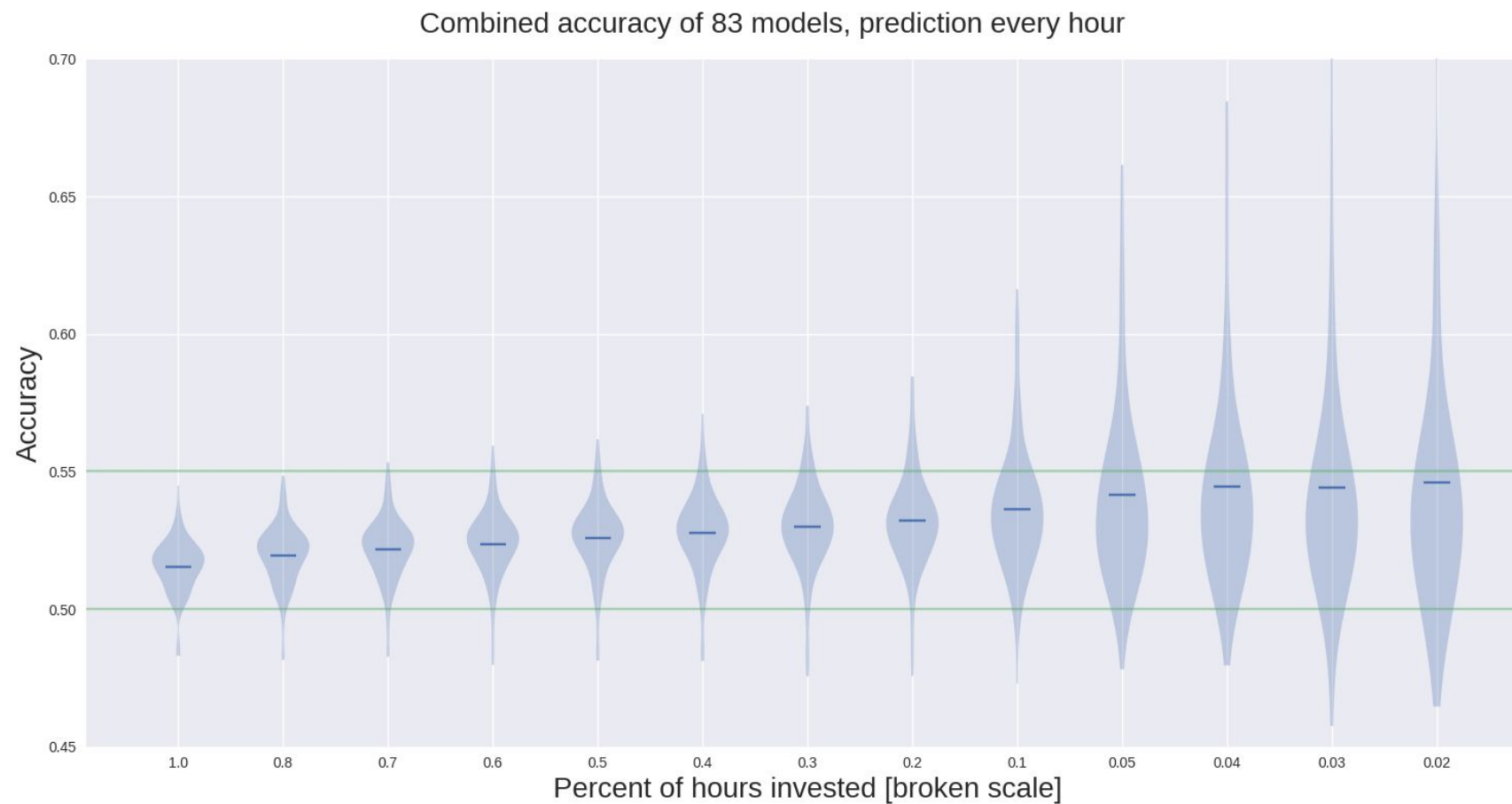


Co zrobiliśmy

- Zmniejszyliśmy model ($300'000 \rightarrow \sim 10$ parametrów)
- Zmniejszyliśmy przestrzeń cech ($10-100 \rightarrow 4$)
- Oparliśmy nasze decyzje o pewność modelu

Zadziałało?

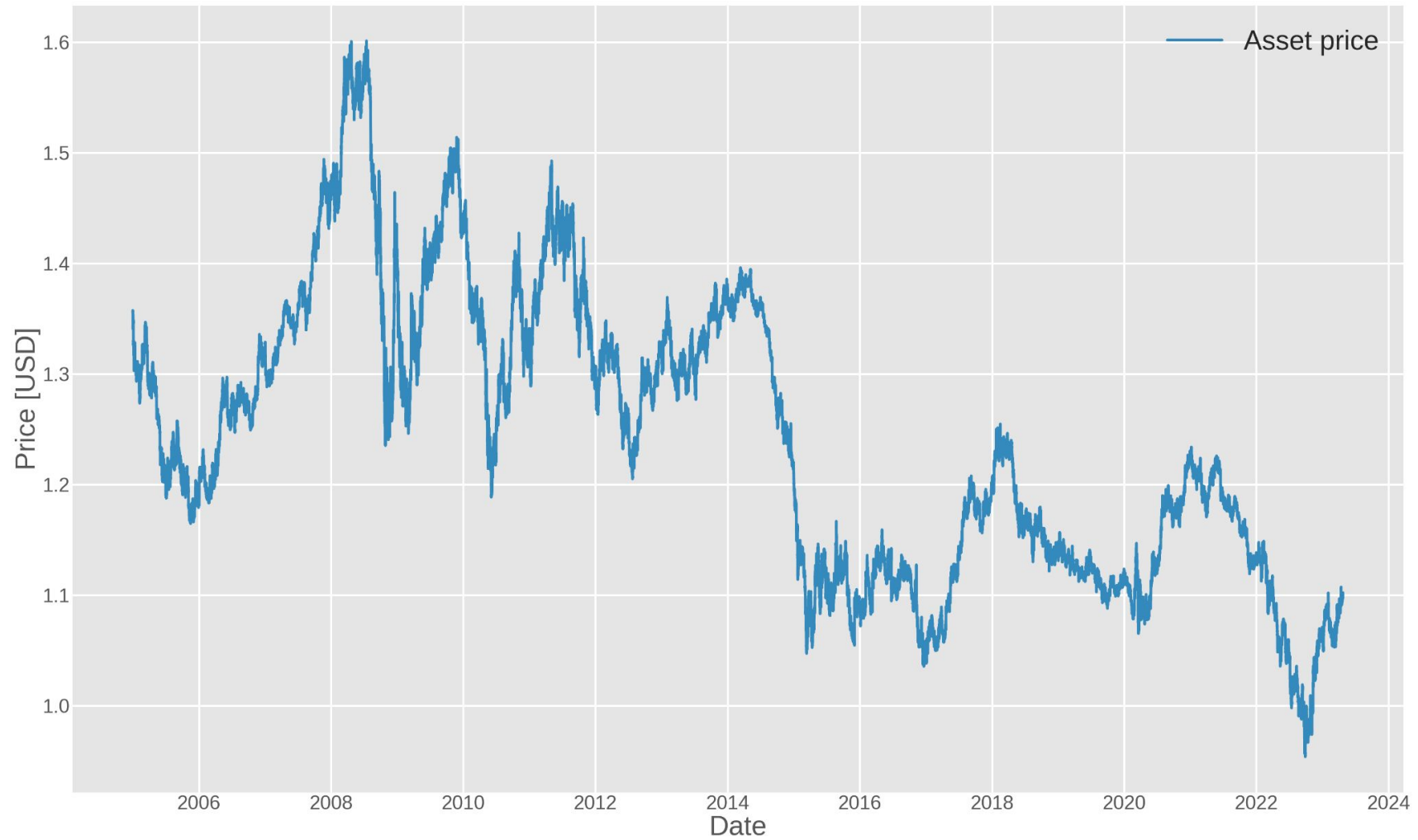
Tak!



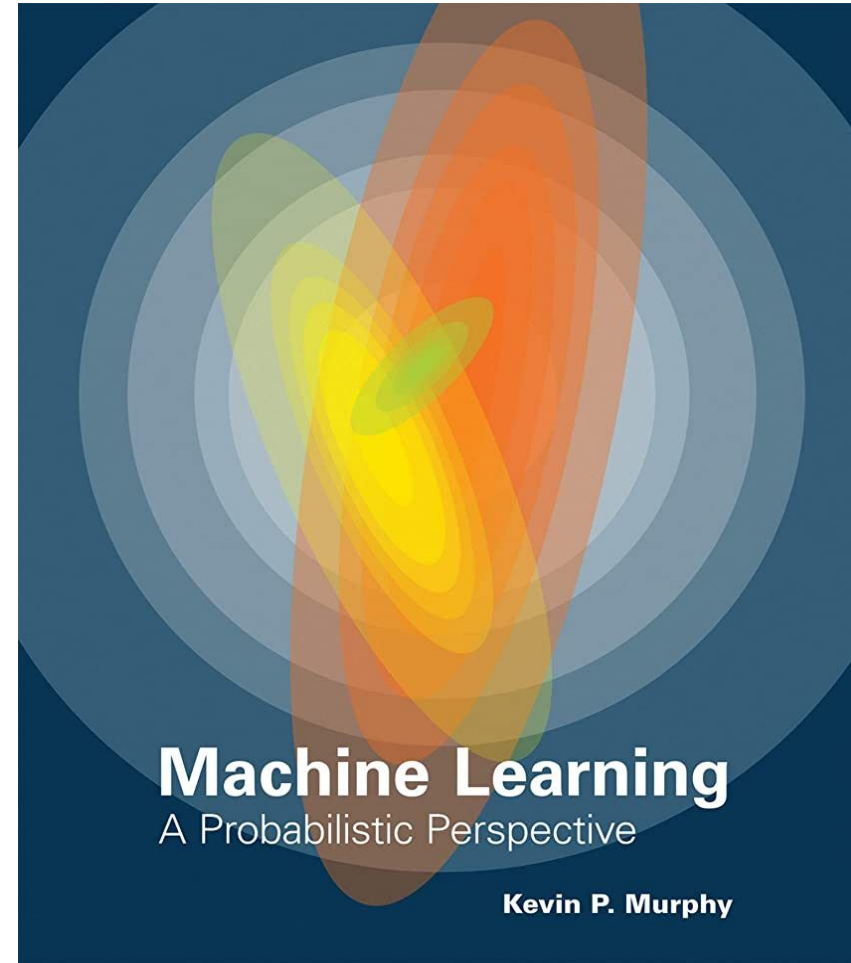
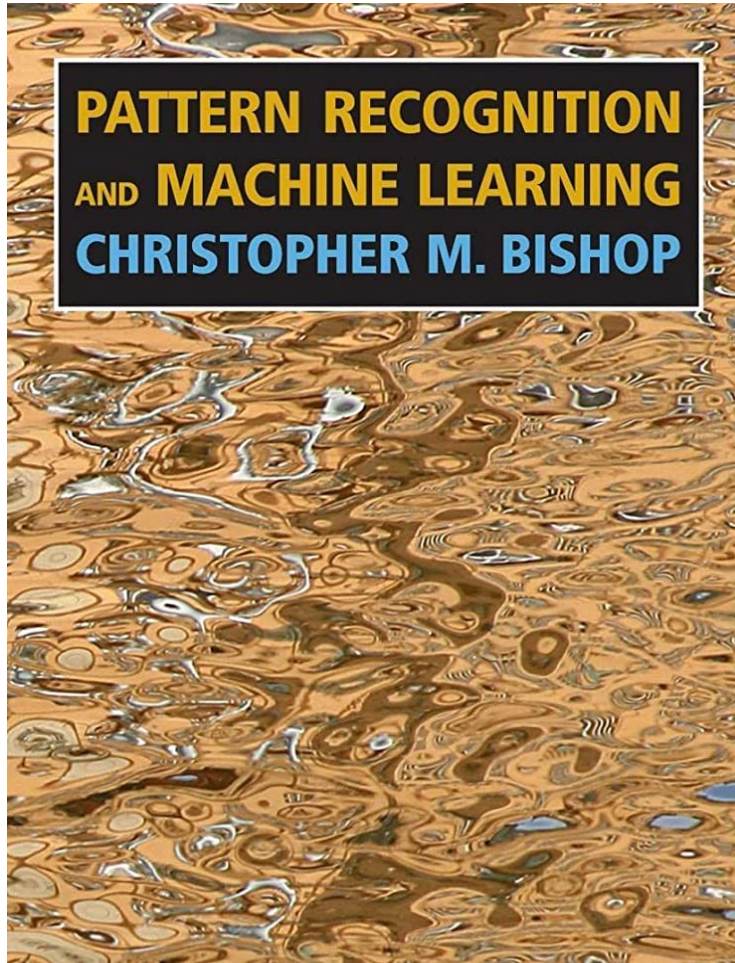
I zarabia.



EURUSD



Zainteresowani?



Dziękuję!

Mateusz Panasiuk
mateusz.panasiuk@aiinvestments.pl
[linkedin.com/in/mat-panasiuk/](https://www.linkedin.com/in/mat-panasiuk/)