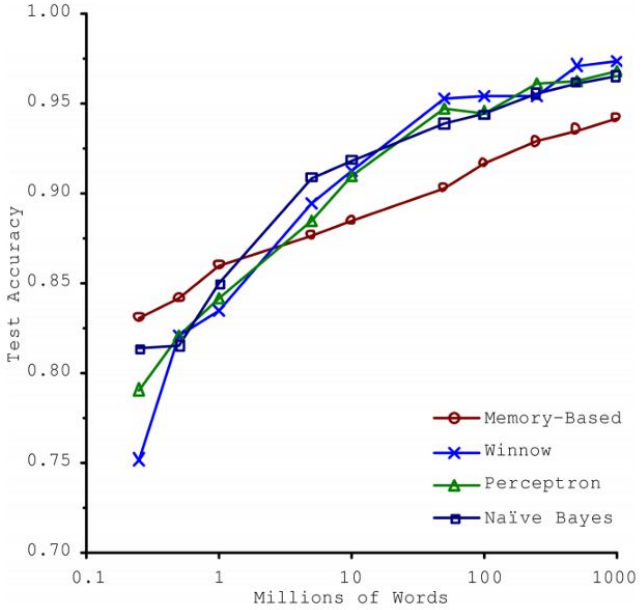
ML Problems

Insufficient training data



Hardware limitations

Can your computer train a model with 1,000,000 observations and 25 features? Try it! (save everything important first)

The rising need of computing power by AI:
https://www.technologyreview.com/2019/11/11/132004/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/

The curse of dimensionality

We want as much information as possible for our algorithms to find patterns, but more attributes mean:

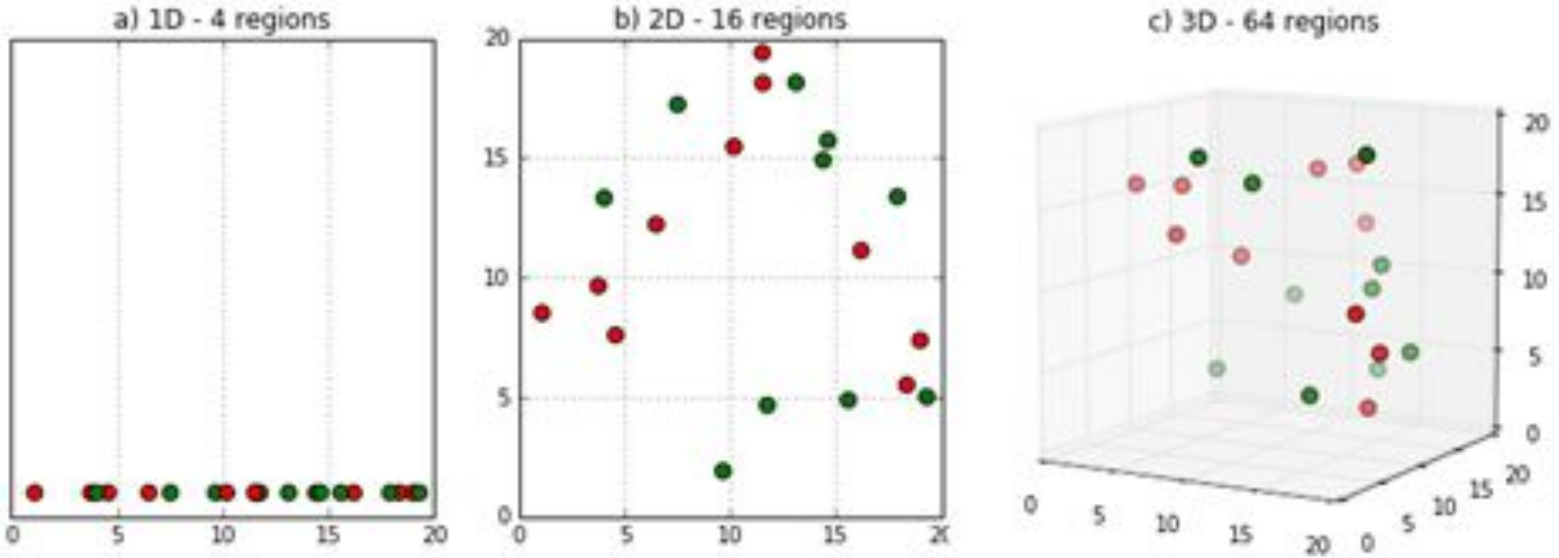
- More training time
- Difficult interpretation
- Possibly worse performance

Problem common in:

- Computer vision (each pixel is a data point)
- NLP (each word + each pair of words + each triplet of words + …)

The Master Algorithm, Pedro Domingos (186-190)

- High dimensionality = lower density
- With more “regions” and the same number of observations, we have fewer observations per region.



Solution? More observations! ...see previous problems.

The complexity monster

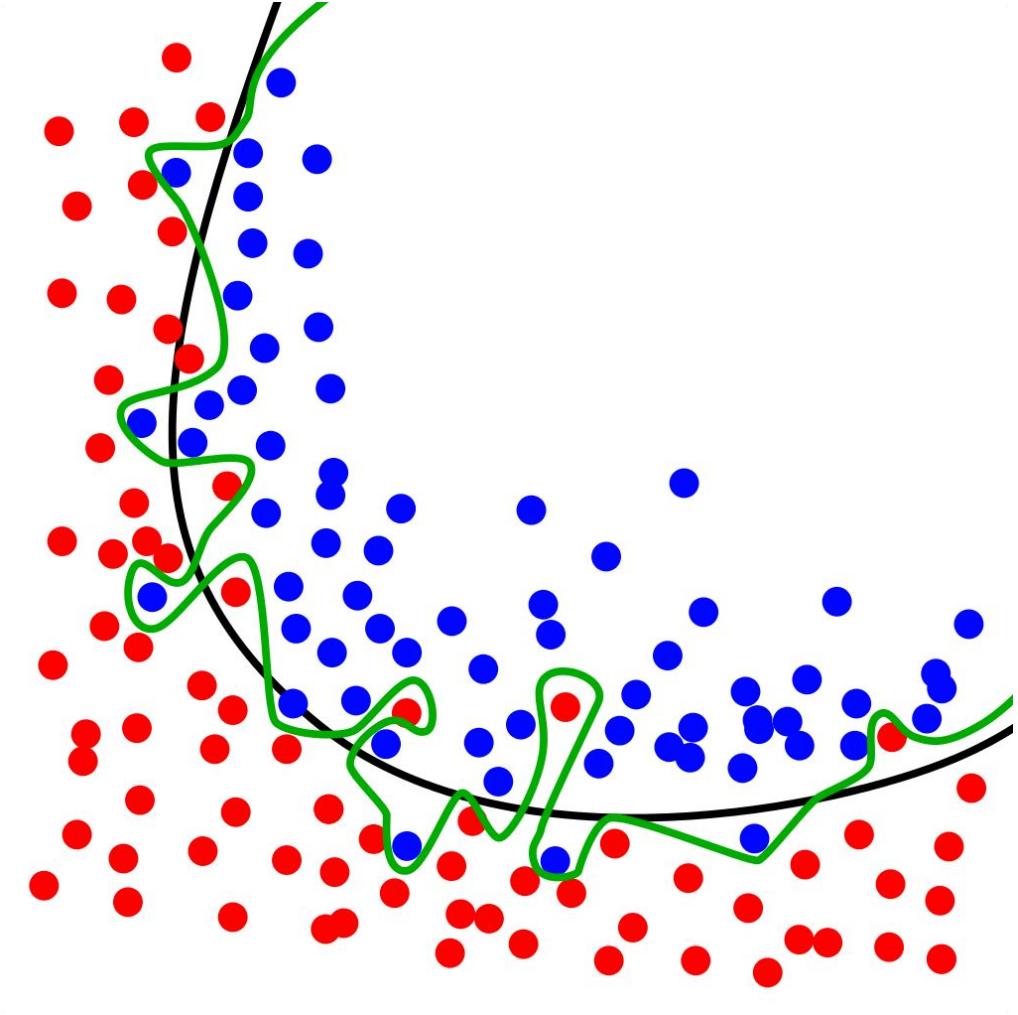
Your Convolutional Neural Network has analyzed thousands of features in candidates for a leading role in the company, including essays the candidates wrote and videos from the interviews they attended. The output is surprising: it picks the guy with the worse curriculum. How do you justify the choice to the stakeholders?

Introduction to Statistical Learning, “The Trade-Off Between Prediction Accuracy and Model Interpretability” (24 - 25)

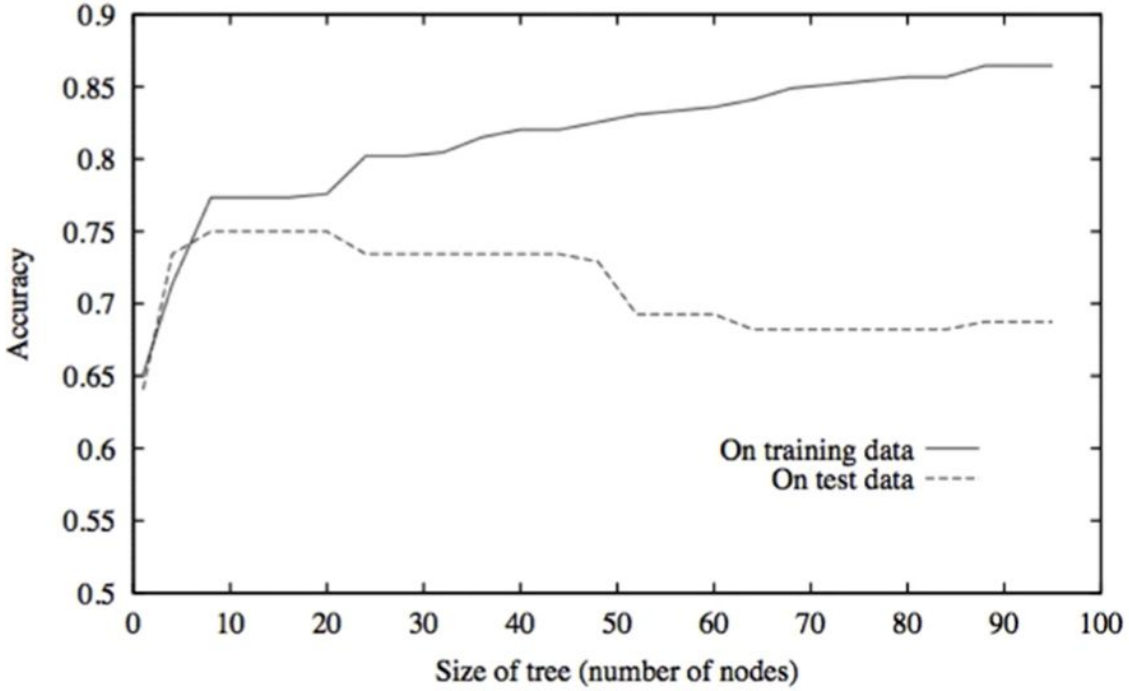
Bias vs Variance Tradeoff

Bias: error due to oversimplification. “All Germans are tall” → While it’s true that they are tall in general, this generalization will lead to many errors. The model is *under fitted*.

Variance: error due to having captured a pattern from what was actually noise. “Germans who have black eyes and live in Dresden and are older than 67 but younger than 69 and are called Jost, measure exactly 183cm.”



How do you detect overfitting?



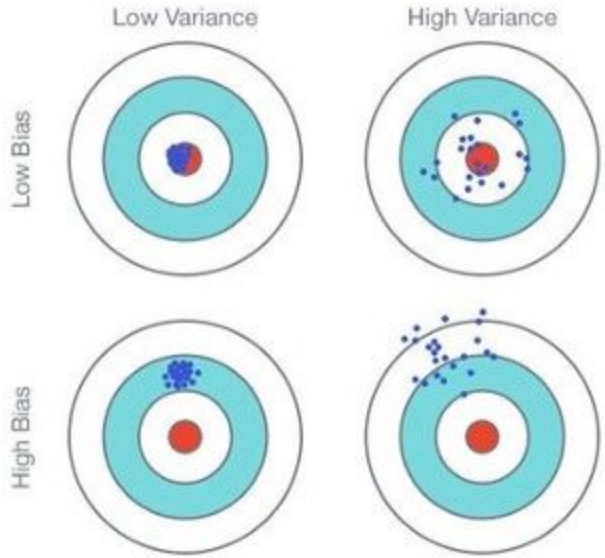
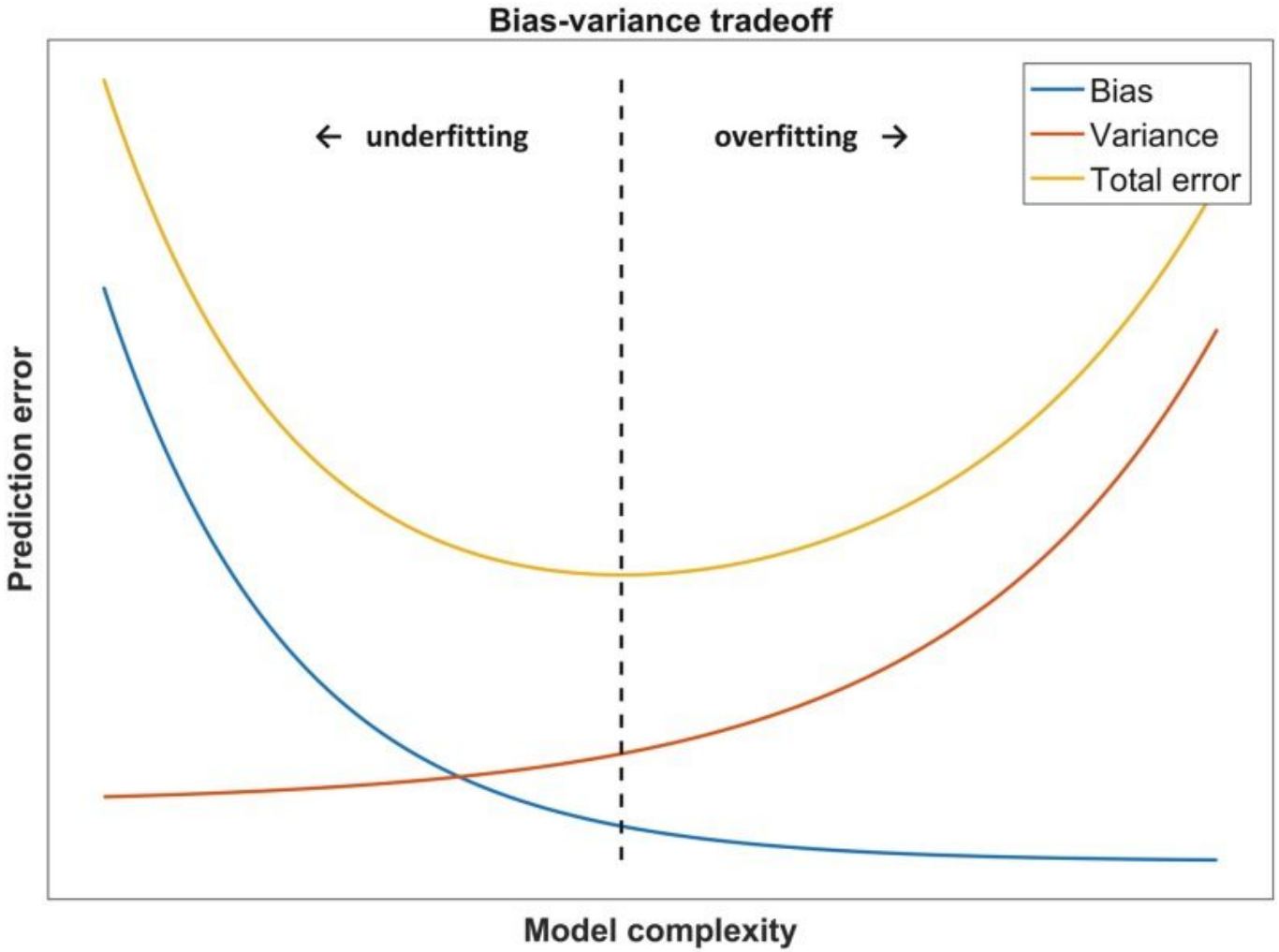


Fig. 1: Graphical Illustration of bias-variance trade-off , Source: Scott Fortmann-Roe., Understanding Bias-Variance Trade-off


Bias-variance tradeoff
← underfitting
overfitting →
Bias
Variance
Total error
Prediction error
Model complexity