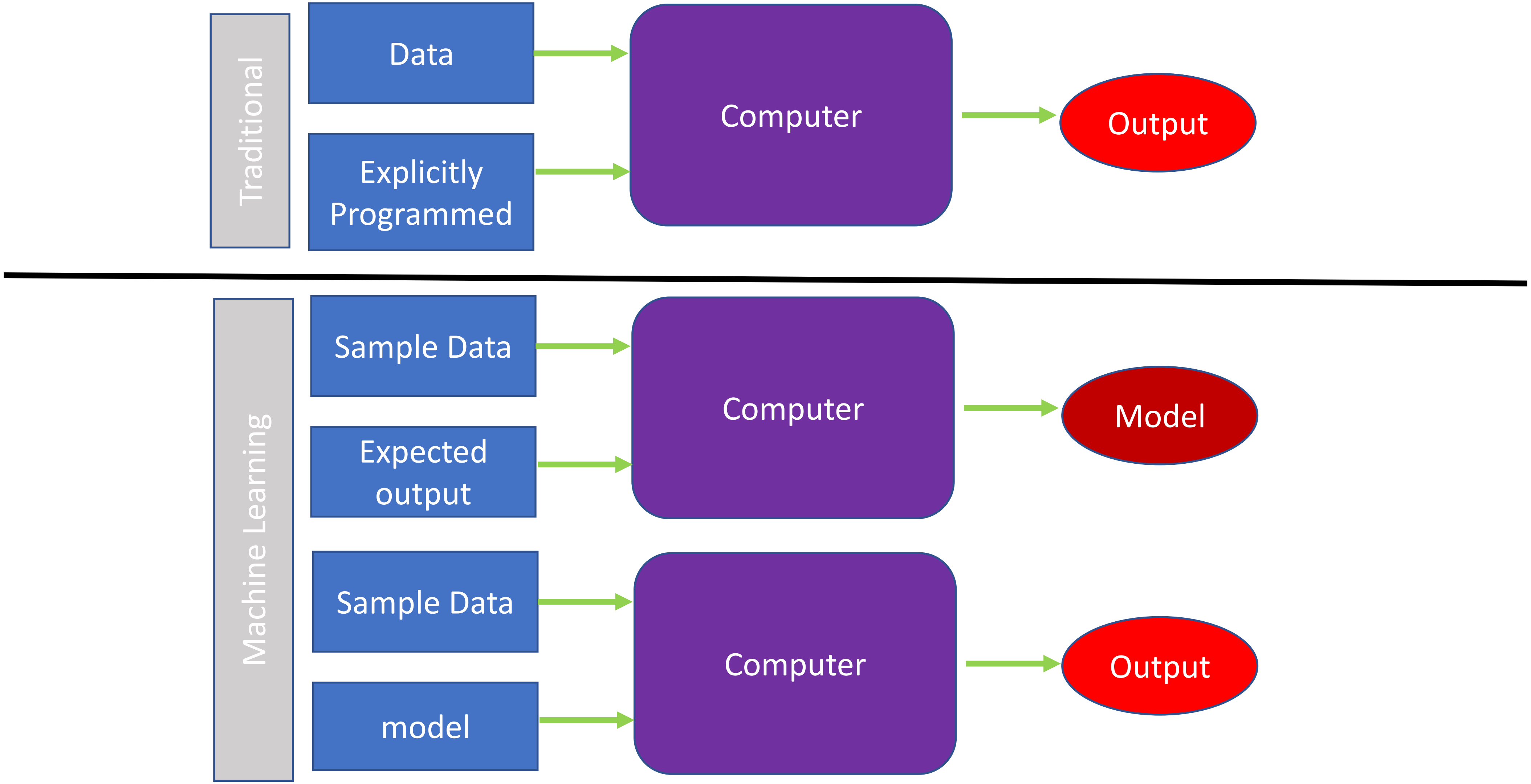
Machine Learning

Kazi Aminul Islam

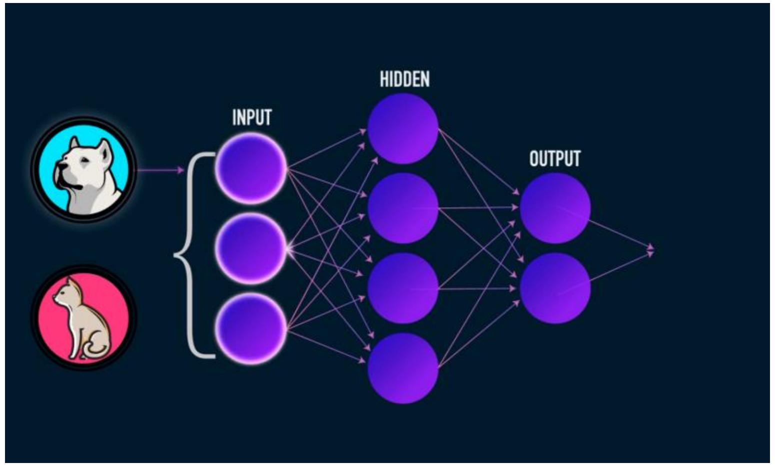
Machine Learning

•

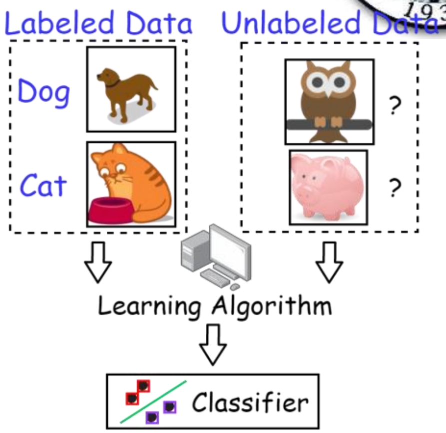


Examples of Applied Machine Learning Usage

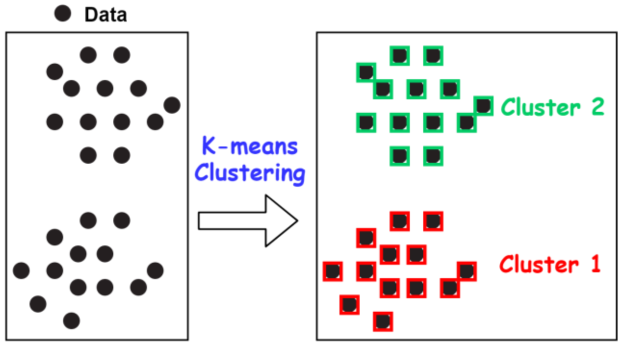
- ➢Supervised Learning, both the inputs and the outputs are provided.
- ➢Unsupervised Learning, no labeled data are provided
- ➢Semi-Supervised Learning. few labeled and unlabeled data are provided
- ➢Reinforcement Learning, learn in an interactive environment by trial and error



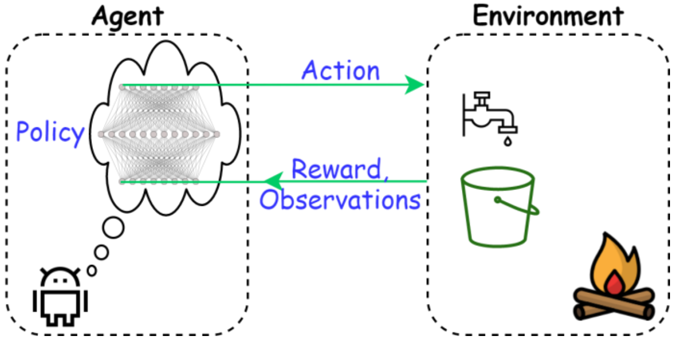
Supervised



Semi-supervised

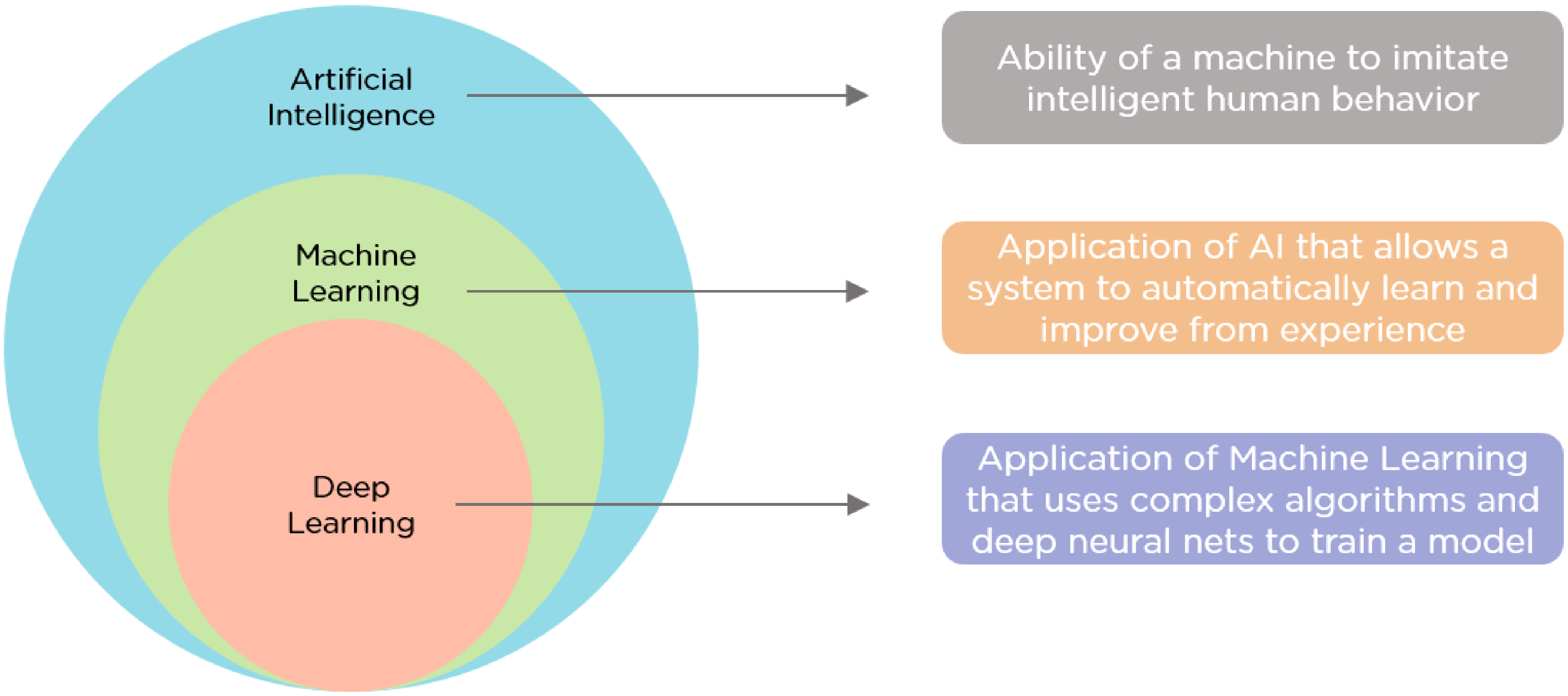


Unsupervised



Reinforcement Learning

Deep Learning



ML Basics

References:

Pattern Recognition and Machine Learning, by Christopher M. Bishop (2006)

Deep Learning by Ian Goodfellow and Yoshua Bengio and Aaron Courville

Machine Learning Algorithms

- Learn from experience E with respect to some class of tasks T and performance measure P,
- Its performance at tasks in T, as measured by P, improves with experience E.
- Three important pillars:
  - Tasks T
  - Performance Measure P
  - Experience E

The Task, T

- **Classification** : Asked to specify which of k categories some input belongs.

The learning algorithm is usually asked to produce a function $f:R^n \rightarrow \{1, . . . , k\}$

- **Classification with missing inputs:** Some of the inputs may be missing, rather than providing a single classification function, the learning algorithm must learn a set of functions
- **Regression**: Predict a numerical value given some input. To solve this task, the learning algorithm is asked to output a function $f: R^n \rightarrow R$.
- **Transcription**: machine learning system is asked to observe a relatively unstructured representation of some kind of data and transcribe the information into discrete textual form.
- **Machine translation:** the input already consists of a sequence of symbols in some language, and the computer program must convert this into a sequence of symbols in another language.

The Task, T

- **Anomaly detection:** the computer program sifts through a set of events or objects and flags some of them as being unusual or atypical.
    - An example of an anomaly detection task is credit card fraud detection. By modeling your purchasing habits, a credit card company can detect misuse of your cards.
    - If a thief steals your credit card or credit card information, the thief's purchases will often come from a different probability distribution over purchase types than your own
- **Synthesis and sampling: T**he machine learning algorithm is asked to generate new examples that are similar to those in the training data.
- **Imputation of missing values:** the machine learning algorithm is given a new example $x \in R^n$, but with some entries xi of x missing
- **Denoising**: The learner must predict the clean example from its corrupted version or more generally predict the conditional probability distribution
- Density estimation or probability mass function estimation

The Performance Measure, P

To evaluate the abilities of a machine learning algorithm, we must design a quantitative measure of its performance. this performance measure P is specific to the task T being carried out by the system.

- Accuracy: proportion of examples for which the model produces the correct output
- Error rate, the proportion of examples for which the model produces an incorrect output
- Mean Square Error
- RMSE Error

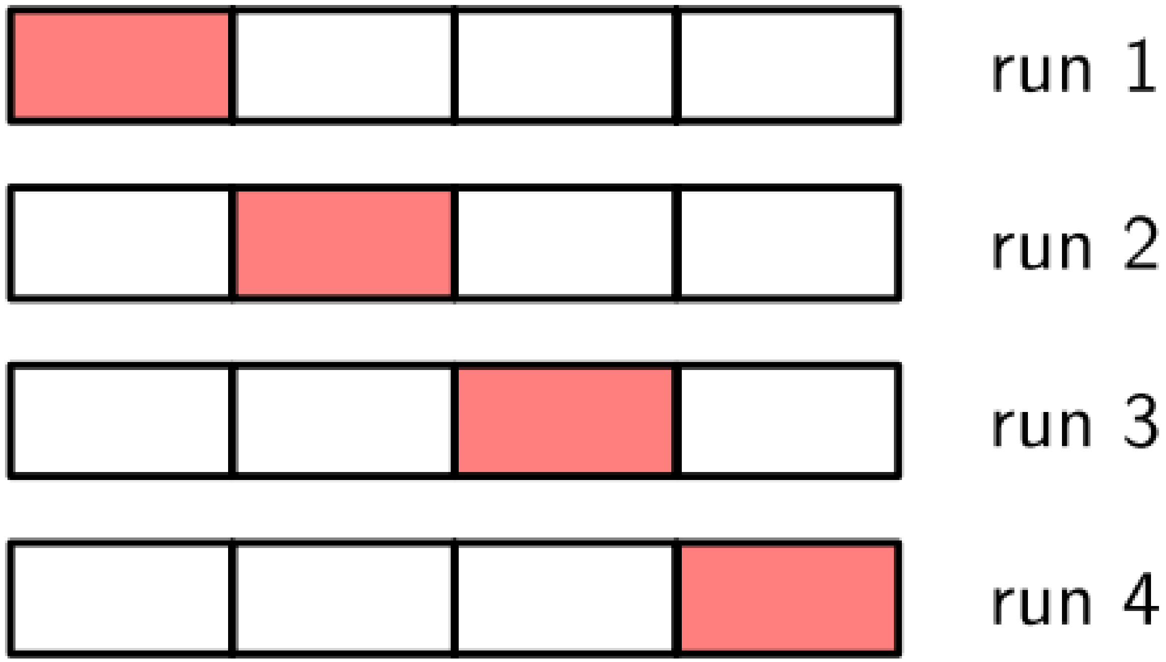
The choice of performance measure may seem straightforward and objective, but it is often difficult to choose a performance measure that corresponds well to the desired behavior of the system

The Experience, E

- **Unsupervised learning algorithms**: experience a dataset containing many features, then learn useful properties of the structure of this dataset
- **Supervised learning algorithms: E**xperience a dataset containing features, but each example is also associated with a label or target.

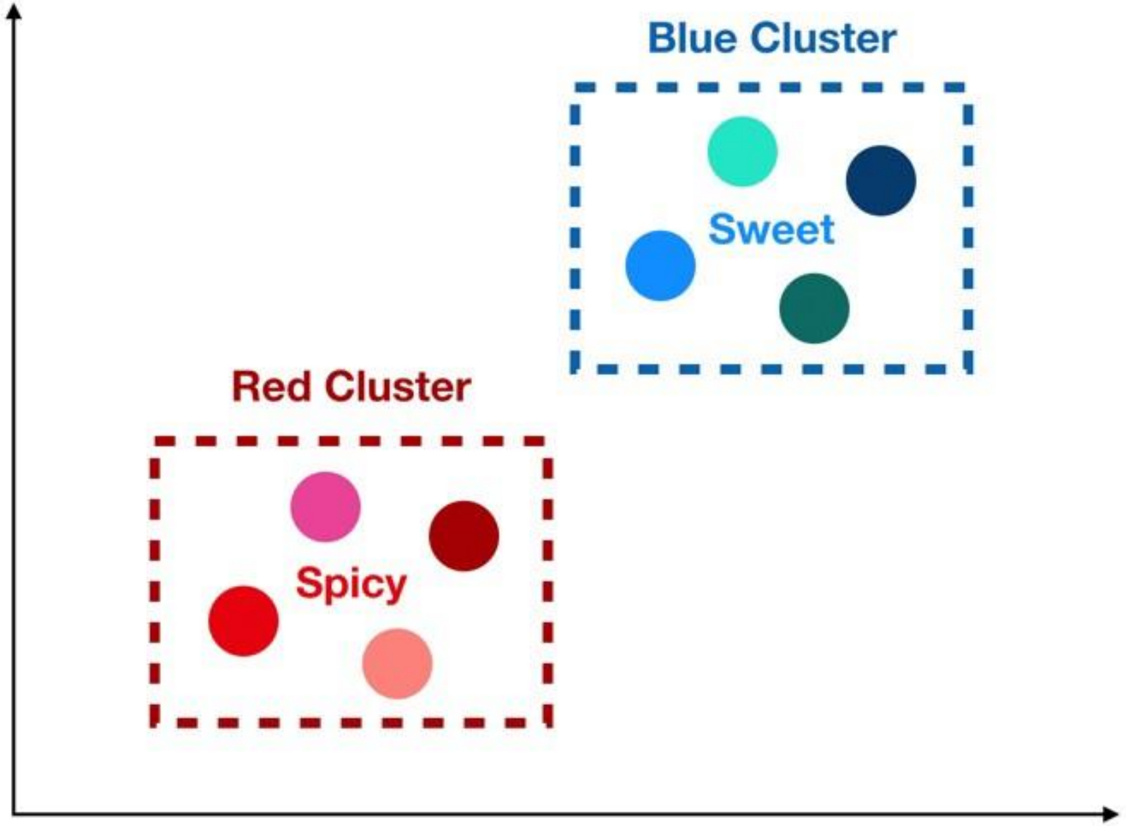
Model selection

- The technique of S-fold cross-validation, illustrated here for the case of S =4, involves taking the available data and partitioning it into S groups (in the simplest case these are of equal size).
- Then S − 1 of the groups are used to train a set of models that are then evaluated on the remaining group.
- This procedure is then repeated for all S possible choices for the held-out group, indicated here by the red blocks, and the performance scores from the S runs are then averaged



Curse of dimensionality

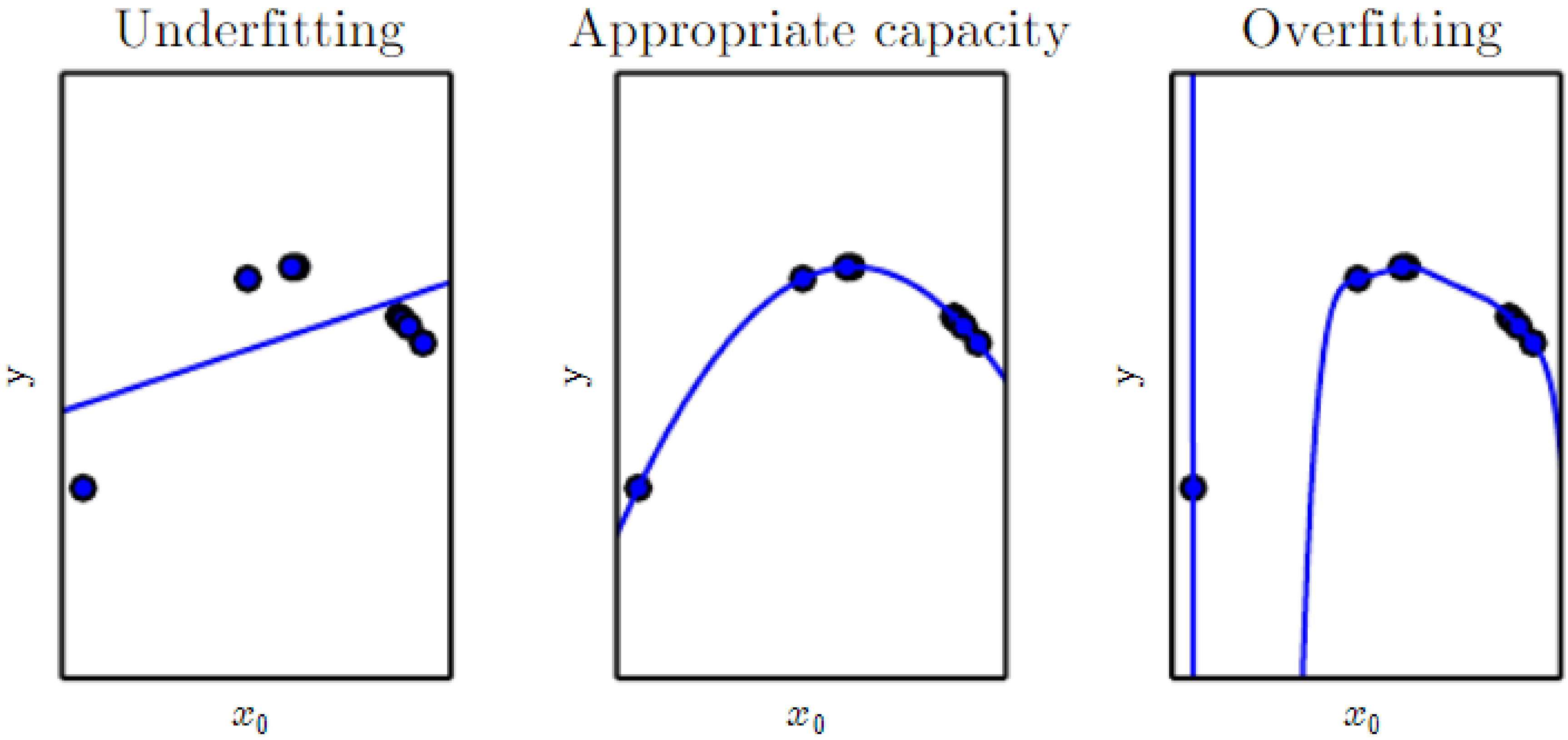
1. If we have more features than observations than we run the risk of massively overfitting our model — this would generally result in terrible out of sample performance.
2. When we have too many features, observations become harder to cluster — believe it or not, too many dimensions causes every observation in your dataset to appear equidistant from all the others



https://towardsdatascience.com/the-curse-of-dimensionality-50dc6e49aa1e

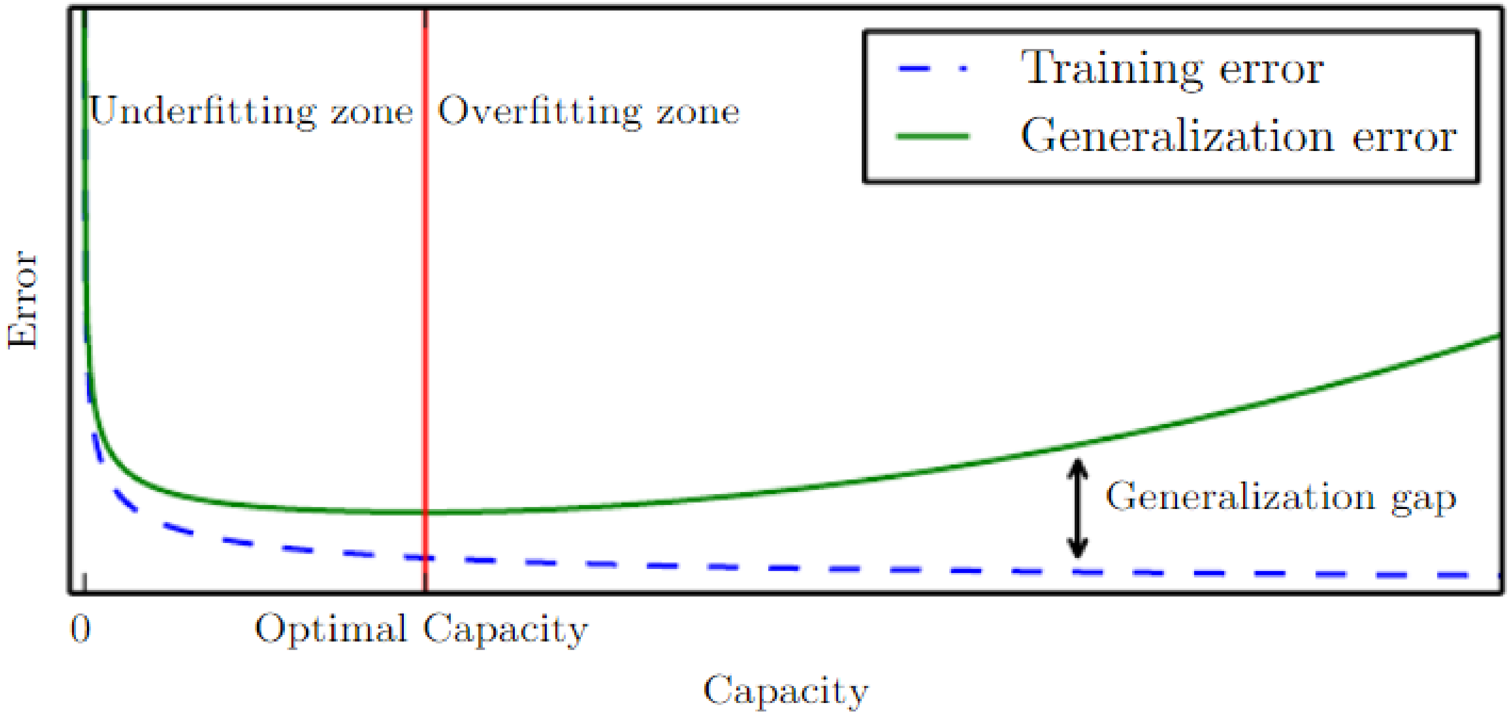
Capacity, Overfitting and Underfitting

- Capacity, Overfitting and underfitting



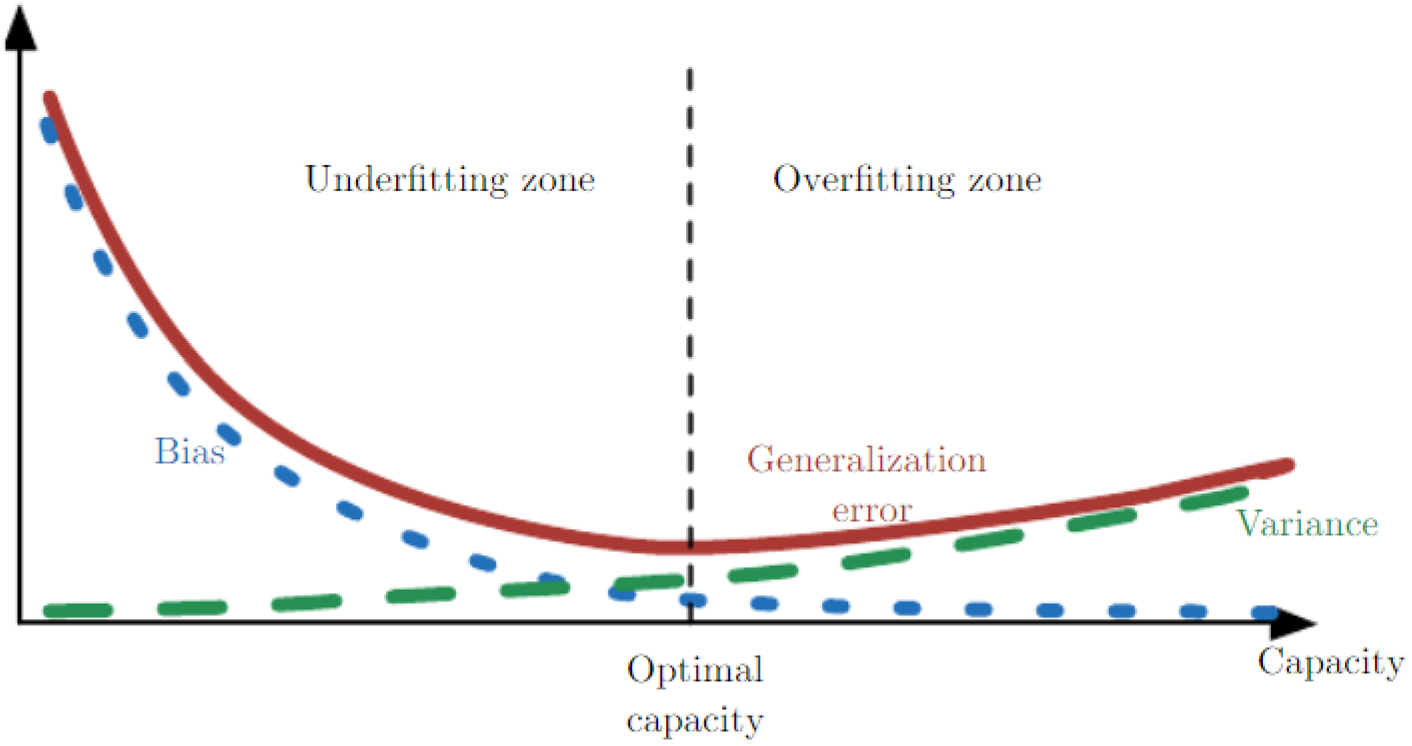
Typical relationship between capacity and error

- Training and test error behave differently. At the left end of the graph, training error and generalization error are both high. This is the underfitting regime
- As we increase capacity, training error decreases, but the gap between training and generalization error increases.
- Eventually, the size of this gap outweighs the decrease in training error, and we enter the Overfitting regime, where capacity is too large, above the optimal capacity.



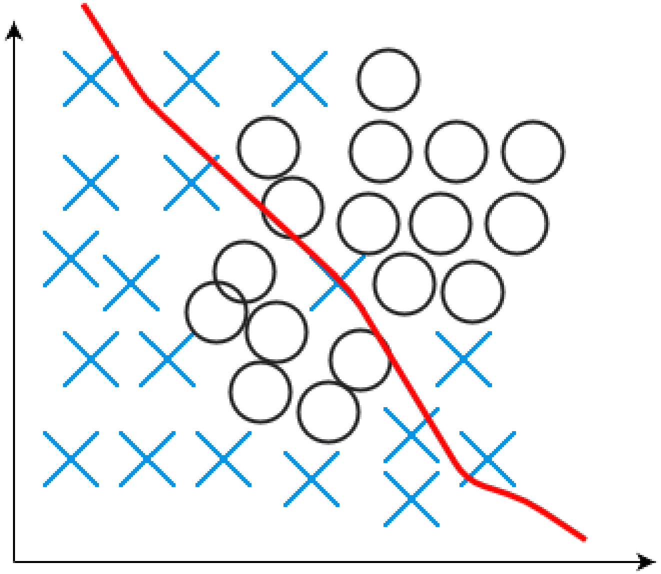
Trading off Bias and Variance to Minimize Mean Squared Error

- As capacity increases (x - axis), bias (dotted) tends to decrease and variance
- (dashed) tends to increase, yielding another U-shaped curve for generalization error (bold curve).
- If we vary capacity along one axis, there is an optimal capacity, with underfitting
- when the capacity is below this optimum and overfitting when it is above.

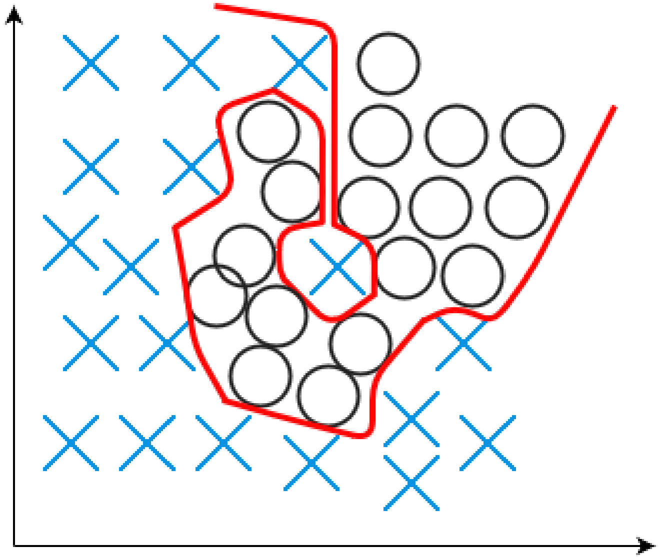


Bias vs Variance Tradeoff

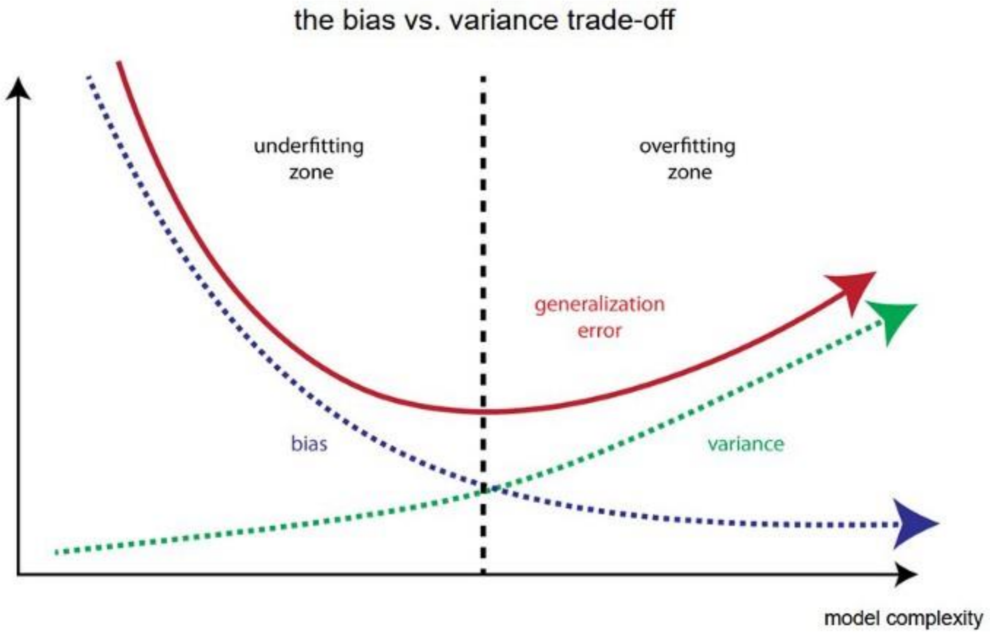
- Build accurate model and avoid the mistake of overfitting and underfitting.
- Bias vs Variance



Underfitting/High Variance



Overfitting/High Variance



https://towardsdatascience.com/bias-and-variance-but-what-are-they-really-ac539817e171

Thank You

Questions?