An Introduction to Machine Learning

What does it mean for a machine to learn anyways?

What is it?

Machine Learning is a type of Artificial Intelligence.

All Machine Learning is Artificial Intelligence, but not all Artificial Intelligence is Machine Learning. For example, Answer Set Programming.

Similarly, all Deep Learning is Machine Learning, but not all Machine Learning is Deep Learning.

Artificial Intelligence != Artificial Consciousness

Formal Definition

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.”

-Tom Mitchell

Task, T

Common Machine Learning Tasks include(but are not limited to)

- Classification
- Regression
- Translation
- Denoising

Performance Measure, P

P is usually specific to the task at hand.

One common method is to measure the accuracy(How often does the model produce the correct output.)

Similarly information can be obtained by measuring the error rate(How often does the model produce the incorrect output.)

We are primarily interested in how well the algorithm performs on new data.

The best choice is not always clear.

Experience, E

In general, we can describe a learning algorithm as being allowed to experience a dataset.

The type and format of your data plays a large role in categorizing your algorithm.

Data is often structured as a design matrix.

Experience is made up of examples, where each example is a collection of features.

Categories

Machine Learning Algorithms are often split into three categories.

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Generalization

We are mostly concerned about generalization error.

i.i.d. assumptions:

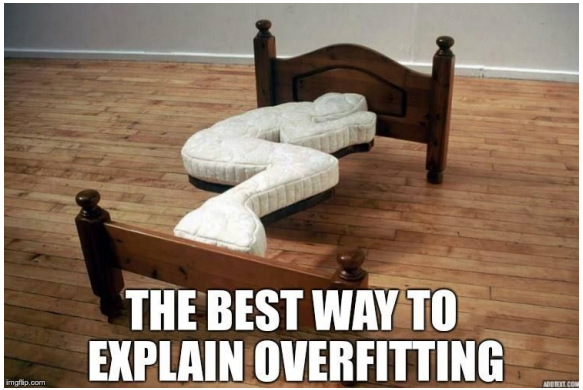
We assume that examples in each dataset are independent and that the training set and test set are identically distributed, drawn from the same probability distribution.

Probability Theory and the i.i.d. assumptions are used to mathematically study the relationship between training error and test error.

Overfitting and Underfitting

Underfitting is what happens when the model performs poorly on the training set.

Overfitting is what happens when the model performs well on the training set but poorly on the test set.



Capacity

A model's capacity refers to the variety of functions it can fit.

A larger/smaller capacity can effects the likelihood that your model will overfit or underfit.

The hypothesis space is the set of functions a learning algorithm is allowed to select as being the solution.

No Free Lunch

The No Free Lunch Theorem:

Averaged over all possible data-generating distributions, every classification algorithm has the same error rate when classifying previously unobserved points.

Regularization

Regularization refers to modifications made to a learning algorithm intended to reduce generalization error but not training error.

General Construction

While it is not always obvious, most Machine Learning algorithms can be constructed by combining the following parts.

- Dataset
- Cost Function
- Optimization Procedure
- Model

Next time we will examine a variety of examples and how they can be described this way.

Resources for getting started

http://www.deeplearningbook.org/

https://www.hackerrank.com/

https://developers.google.com/machine-learning/crash-course/