What is Machine Learning?

Alex Thayn

Machine Learning Is:

- A subset of Artificial Intelligence
- Training of a system to do a task without explicit instructions
- Development of programs that that use data to learn for themselves without human intervention
- Categorized as supervised or unsupervised

Artificial Intelligence

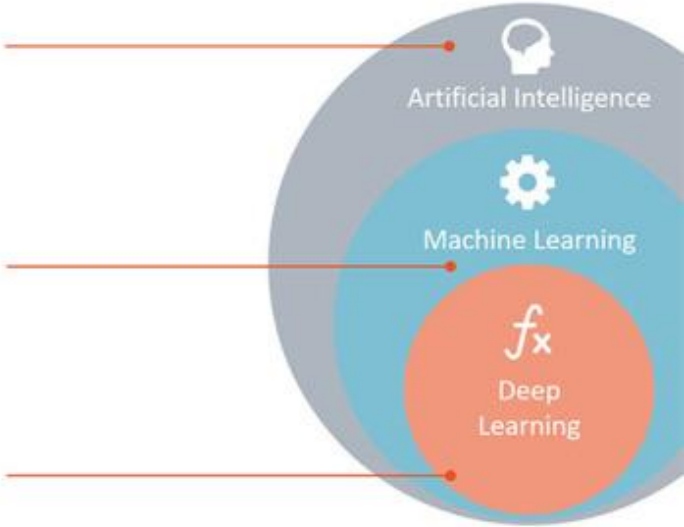
Any technique which enables computers to mimic human behavior.

Machine Learning

Subset of AI techniques which use statistical methods to enable machines to improve with experiences.

Deep Learning

Subset of ML which make the computation of multi-layer neural networks feasible.



Source: https://rapidminer.com/blog/artificial-intelligence-machine-learning-deep-learning/

Types of Learning

Supervised Learning

- We know what the correct output should look like
- Categorized into two types of problems:
  - Regression: trying to predict results with continuous output, mapping input variables to a continuous function
  - Classification: trying to predict results with discrete output, mapping input variables into discrete categories

Unsupervised Learning

- We have little or no idea what out results should look like
- We derive structure from data by clustering the data based on relationships among the variables in the data.

Source: https://www.coursera.org/learn/machine-learning/home/welcome

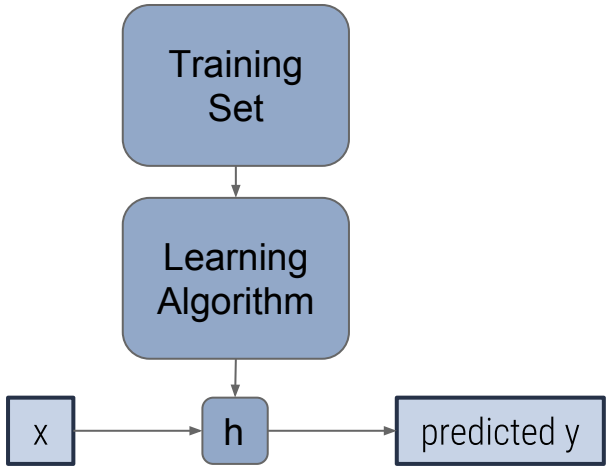
1

Supervised Learning

Formal Description of Supervised Learning

- Goal: Given a training set, learn a function h: X ⟶ Y so that h(x) is a “good” predictor for the corresponding value of y.
  - h(x) = hypothesis

- Example: Given a training set of data containing handwritten text, convert the hand-written images to digital text



https://www.coursera.org/learn/machine-learning/supplement/cRa2m/model-representation

Cost Function

- We use the cost function to measure the accuracy of our hypothesis function
- The cost function is also known as the “squared error function”, or “mean squared error”.
- It can be calculated using the following equation:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2 = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2$$

Gradient Descent Algorithm

- The Gradient Descent algorithm was the first ML algorithm I learned.
- It is used to find parameters that best fit θ_0 and θ_1 of the hypothesis function
- The algorithm picking a starting point (red star) and involves taking steps (arrows) toward the direction of steepest descent until we reach a minimum point on the graph.