Outline

Introduction

What is Machine Learning?

- Field of study that gives computers the ability to learn without being explicitly programmed [Authur Samuel (1959)]
- A computer program is said to learn from experience E with respect to some task T and some performance on T, as measured by P, improves with experience E [Tom Mitchell (1998)]

Consider your e-mail program watches which e-mails you do or do not mark as spam, and based on that learns how to better filter spam

- Classifying e-mails as spam or not spam $\rightarrow T$
- Watching you label e-mails as spam or not spam $\rightarrow E$
- The number (or fraction) of e-mails correctly classified as spam/not spam $\rightarrow P$

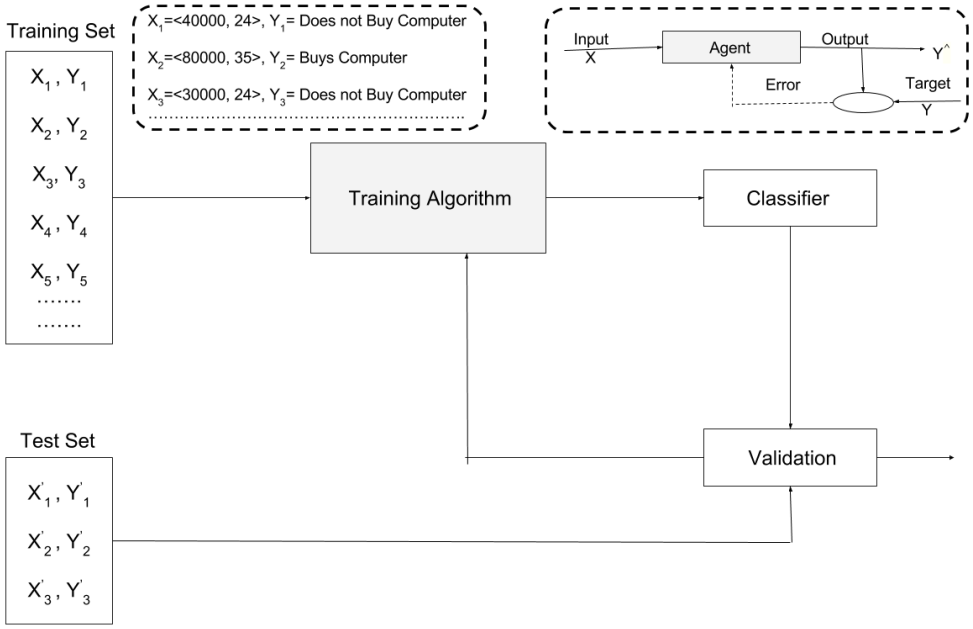
Different Types of Learning

- Supervised Learning: Learn an input and output map
  - Classification: categorical output
  - Regression: continuous output
- Unsupervised Learning: Discover patterns in the data
  - Clustering: cohesive grouping
  - Association: frequent cooccurrence
- Reinforcement Learning: Learning Control

Supervised Learning

Supervised Learning

- Need to generalize→Assumptions about lines!
- In general, Inductive bias
  - Language bias
  - Search bias
- The process



Supervised Learning

Applications

- Credit Card fraud detection
  - Valid transaction or not
- Sentiment Analysis
  - Opinion mining, buzz analysis etc.
- Part of Speech Tagging
  - Word in a sentence is noun, pronoun, verb etc.
- Entity identification and categorization
  - Entity is a person name, location name, company name etc.

Thank You!