Machine Learning **Note**

MIT 6.036

1 Introduction

1.1 Problem classes

1.1.1 Supervised learning

In *supervised learning*, the learning system is given inputs and their expected outputs to learn their relations. If the output is discrete, we call it *classification*. Otherwise, for continuous outcomes, we call it *regression*. Classification can be further divided into *two-class* if the output can only be two possible values, and *multi-class* otherwise.

More formally, the learning system is given training data $\mathcal{D}_n = \{(x^{(1)}, y^{(1)}), \ldots, (x^{(n)}, y^{(n)})\}$. The goal is, given a new data $x^{(n+1)}$, to predict the value of $y^{(n+1)}$.

1.1.2 Unsupervised learning

In *unsupervised learning*, the learning system is not given expected outputs, but expected to find some inherent property or pattern in it. There are several problem classes:

Density estimation. The goal is, given samples $x^{(1)}, \ldots, x^{(n)}$ drawn i.i.d. from the same distribution, to predict the probabaility of $P(x^{(n+1)})$ appearing as next input.

Clustering. Given samples $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^D$, the goal is to found partitioning, or clusters that groups "similar" data together.

Dimensionality reduction. Given samples $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^D$, the goal is to re-represent them in smaller dimensional space while losing little informations.