Lecture 2

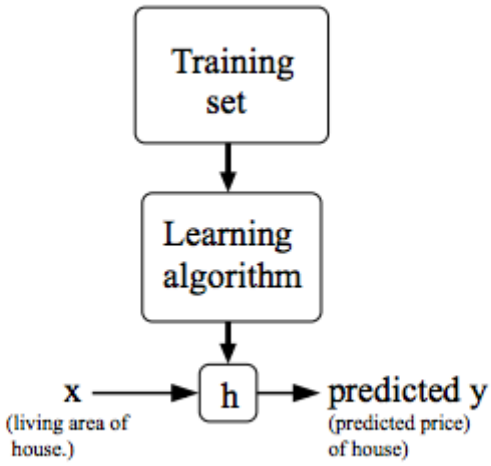
Notation ¶

- $x^{(i)}$: input (features) variables
- $y^{(i)}$: output (target) variable, i.e., what we are trying to predict
- a pair $\left(x^{(i)}, y^{(i)}\right)$ is called a training example
- a list of m training examples $\left(x^{(i)}, y^{(i)}\right); i = 1, \ldots, m$ is called a training set, i.e., the dataset we'll be using to learn
- X: space of input values
- Y: space of output values

Supervised Learning (formally)

- **goal:** given a training set, learn a function $h : X \to Y$ so that $h(x)$ is a "good predictor" of y.

> For historical reasons, the function h is called a hypothesis.



- **regression problem:** when the target variable is continuous.
- **classification problem:** when y can take on only a small number of discrete values.