

Welcome

✓ **Video:** Welcome to Machine Learning!
1 min

✓ **Reading:** Machine Learning Honor Code
8 min

Introduction

✓ **Video:** Welcome
6 min

✓ **Video:** What is Machine Learning?
7 min

✓ **Reading:** What is Machine Learning?
5 min

✓ **Reading:** How to Use Discussion Forums
4 min

✓ **Video:** Supervised Learning
12 min

✓ **Reading:** Supervised Learning
4 min

✓ **Video:** Unsupervised Learning
14 min

✓ **Reading:** Unsupervised Learning
3 min

✓ **Reading:** Who are Mentors?
3 min

✓ **Reading:** Get to Know Your Classmates
8 min

✓ **Reading:** Frequently Asked Questions
11 min

Review

✓ **Reading:** Lecture Slides
20 min

✓ **Quiz:** Introduction
5 questions

Model and Cost Function

✓ **Video:** Model Representation
8 min

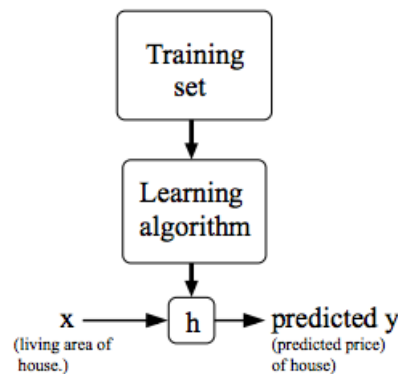
✓ **Reading:** Model Representation
~



Model Representation

To establish notation for future use, we'll use $x^{(i)}$ to denote the "input" variables (living area in this example), also called input features, and $y^{(i)}$ to denote the "output" or target variable that we are trying to predict (price). A pair $(x^{(i)}, y^{(i)})$ is called a training example, and the dataset that we'll be using to learn—a list of m training examples $(x^{(i)}, y^{(i)}); i = 1, \dots, m$ —is called a training set. Note that the superscript "(i)" in the notation is simply an index into the training set, and has nothing to do with exponentiation. We will also use X to denote the space of input values, and Y to denote the space of output values. In this example, $X = Y = \mathbb{R}$.

To describe the supervised learning problem slightly more formally, our goal is, given a training set, to learn a function $h : X \rightarrow Y$ so that $h(x)$ is a "good" predictor for the corresponding value of y . For historical reasons, this function h is called a hypothesis. Seen pictorially, the process is therefore like this:



When the target variable that we're trying to predict is continuous, such as in our housing example, we call the learning problem a regression problem. When y can take on only a small number of discrete values (such as if, given the living area, we wanted to predict if a dwelling is a house or an apartment, say), we call it a classification problem.

✓ Complete

Go to next item

