



Classification and Representation

- ✓ **Video:** Classification
8 min
- ✓ **Reading:** Classification
2 min
- ✓ **Video:** Hypothesis Representation
7 min
- ✓ **Reading:** Hypothesis Representation
3 min
- ✓ **Video:** Decision Boundary
14 min
- ✓ **Reading:** Decision Boundary
3 min

Logistic Regression Model

- ✓ **Video:** Cost Function
10 min
- ✓ **Reading:** Cost Function
3 min
- ✓ **Video:** Simplified Cost Function and Gradient Descent
10 min
- ✓ **Reading:** Simplified Cost Function and Gradient Descent
3 min
- ▶ **Video:** Advanced Optimization
14 min
- 📖 **Reading:** Advanced Optimization
3 min

Multiclass Classification

- ▶ **Video:** Multiclass Classification: One-vs-all



Classification

To attempt classification, one method is to use linear regression and map all predictions greater than 0.5 as a 1 and all less than 0.5 as a 0. However, this method doesn't work well because classification is not actually a linear function.

The classification problem is just like the regression problem, except that the values we now want to predict take on only a small number of discrete values. For now, we will focus on the **binary classification problem** in which y can take on only two values, 0 and 1. (Most of what we say here will also generalize to the multiple-class case.) For instance, if we are trying to build a spam classifier for email, then $x^{(i)}$ may be some features of a piece of email, and y may be 1 if it is a piece of spam mail, and 0 otherwise. Hence, $y \in \{0, 1\}$. 0 is also called the negative class, and 1 the positive class, and they are sometimes also denoted by the symbols “-” and “+.” Given $x^{(i)}$, the corresponding $y^{(i)}$ is also called the label for the training example.

✓ Complete

Go to next item

