

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

6 min
- ✓

Reading: Multiclass Classification: One-vs-all

3 min

Review

- ✓

Reading: Lecture Slides

10 min
- Quiz: Logistic Regression

5 questions

Solving the Problem of Overfitting

- Video: The Problem of Overfitting

9 min
- Reading: The Problem of Overfitting

3 min
- Video: Cost Function

Multiclass Classification: One-vs-all

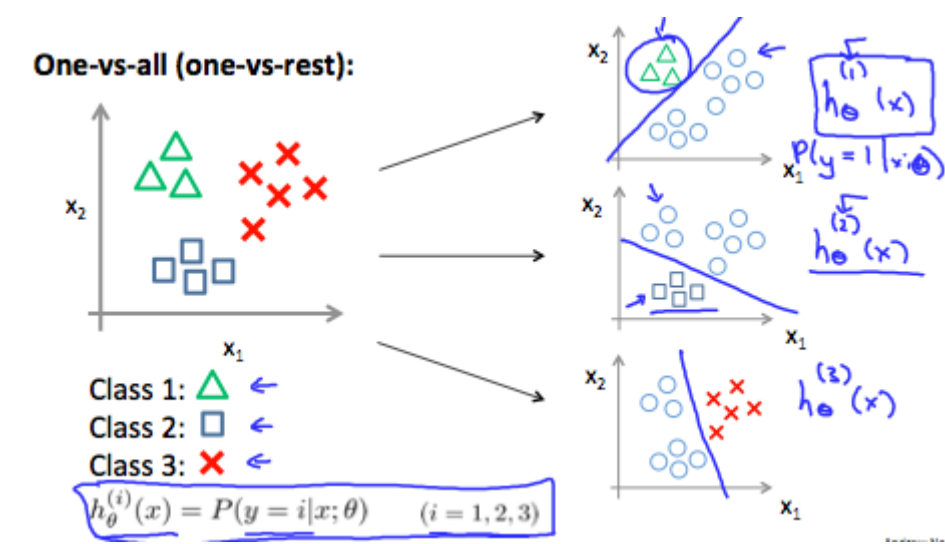
Now we will approach the classification of data when we have more than two categories. Instead of  $y = \{0,1\}$  we will expand our definition so that  $y = \{0,1...n\}$ .

Since  $y = \{0,1...n\}$ , we divide our problem into  $n+1$  (+1 because the index starts at 0) binary classification problems; in each one, we predict the probability that 'y' is a member of one of our classes.

$$y \in \{0,1...n\}$$
$$h_{\theta}^{(0)}(x) = P(y = 0|x;\theta)$$
$$h_{\theta}^{(1)}(x) = P(y = 1|x;\theta)$$
$$\dots$$
$$h_{\theta}^{(n)}(x) = P(y = n|x;\theta)$$
$$\text{prediction} = \max_i(h_{\theta}^{(i)}(x))$$

We are basically choosing one class and then lumping all the others into a single second class. We do this repeatedly, applying binary logistic regression to each case, and then use the hypothesis that returned the highest value as our prediction.

The following image shows how one could classify 3 classes:



To summarize:

Train a logistic regression classifier  $h_{\theta}(x)$  for each class to predict the probability that  $y = i$ .

To make a prediction on a new  $x$ , pick the class that maximizes  $h_{\theta}(x)$

✓ Complete

Go to next item