Writing exercise 3 in machine learning

Kyrie

崔颖琦

1730004002

1. Build a 3-layer neural network for XOR, as the one for XNOR in the lecture notes. Please explain why it works well for XOR.

Answer:

-15 -5

-5

10 -20

10

10 10

10

Layer 1 Layer 2 Layer 3

When x1 and x2 are 0, a11=g(-15+0+0)=g(-15)0, a12=g(-5+0+0)+g(-5)0. Output=g(-5+0+0)=g(-5)0.

When x1 is 0 and x2 is 1, a11=g(-15+0+10)=g(-5)0, a12=g(-5+0+10)=g(5)1, output=g(-5+0+10)=g(5)1.

When x1 is 1 and x2 is 0, a11=g(-15+10+0)=g(-5)0, a12=g(-5+10+0)=g(5)1, output=g(-5+0+10)=g(5)1.

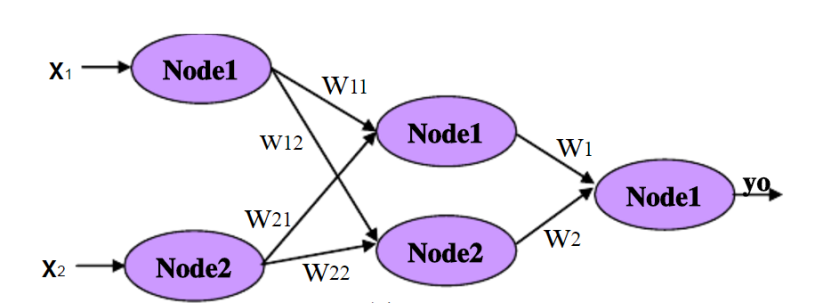
When x1 and x2 are 1, a11=g(-15+10+10)=g(5)1, a12=g(-5+10+10)=g(15)1, output=g(-5-20+10)=g(-15)0.

So it works well for XOR.

1. Express the derivative of a sigmoid function in terms of the sigmoid function itself.

Answer:

1. We plan to train a neural network with the structure below using the forward and back propagation method. Suppose the initial weights for all arcs are 0.5. And we use the sigmoid function as the activation function for each node. Suppose we use data (x1=1,x2=1, y=0) the input, what is the a value and δ value for each node.



Answer:

1. Describe the process when we want to select the model out of 10 models given a set of training data.

Answer:

Divide the data set into three parts. Training set, cross validation set and test set. The proportion of the three sets is 6:2:2. Use the training set to train the models to minimize the cost function. And then use cross validation set to compute the cost of each model. Select the model which cost computed by cross validation set is the minimum. Finally, using test set to compute the cost of the model.

1. Describe the effect on bias/variance when we increase the λ value in the cost function when we are doing linear regression.

Answer:

When we increase the λ value in the cost function in linear regression, it will cause high bias.

1. Describe how do we use neural network to do spam classification.

Answer:

Collect lots of data.

Develop sophisticated features based on email routine information.

Develop sophisticated features for message body.

Develop sophisticated features to detect misspellings.

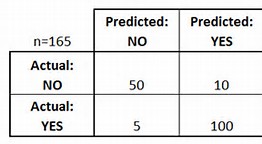
And then start will a simple neural network than can implement fast and test it on cross validation data set.

Plot learning curve to decide if more data or more features are likely to help.

Use error analysis to find out the trend in what type of examples it is making errors on.

And then improve the algorithm.

1. What are the precision, recall and F1-score for the following confusion matrix?



Precision=100/(100+10)=100/110

Recall=100/(100+5)

F1=