coursera

Discussion Forums

Week 4

SUBFORUMS

All

Assignment: Multi-class Classification and Neural Networks

Week 4

ex3 tutorial for predict()

Tom Mosher Mentor · a year ago · Edited

Here is an outline for forward propagation using the vectorized method. This is an implementation of the formula in Figure 2 on Page 11 of ex3.pdf.

1. Add a column of 1's to X (the first column), and it becomes 'a1'.
2. Multiply by Theta1 and you have 'z2'.
3. Compute the sigmoid() of 'z2', then add a column of 1's, and it becomes 'a2'
4. Multiply by Theta2, compute the sigmoid() and it becomes 'a3'.
5. Now use the max(a3, [], 2) function to return two vectors - one of the highest value for each row, and one with its index. Ignore the highest values. Keep the vector of the indexes where the highest values were found. These are your predictions.

Note: When you multiply by the Theta matrices, you'll have to use transposition to get a result that is the correct size.

Note: The predictions must be returned as a column vector - size (m x 1). If you return a row vector, the script will not compute the accuracy correctly.

Note: Not getting the correct results? In the hidden layer, be sure you use sigmoid() first, then add the bias unit.

------ dimensions of the variables ---------

a1 is (m x n), where 'n' is the number of features including the bias unit

Theta1 is (h x n) where 'h' is the number of hidden units

a2 is (m x (h + 1))

Theta2 is (c x (h + 1)), where 'c' is the number of labels.

a3 is (m x c)

p is a vector of size (m x 1)

36 Upvote · Follow 18 · Reply to Tom Mosher

This thread is closed. You cannot add any more responses.

Earliest **Top** **Most Recent**

Krishnang Dalal · 8 months ago

I get test case oneVSall correct. I get 94.62% prediction accuracy for predictoneVSall. But when I am submitting the entire assignment it does not calculate scores for the regularized Logistic Regression and One-vs-All Classifier Training. It even doesn't give wrong feedback. What could be the issue?

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 8 months ago

The issue is that your code does not work correctly for the test case that the submit grader uses. Your code must work correctly for any data set.

I recommend you read the Tutorials and use the additional test cases in the Resources menu.

0 Upvote

H

Harpreet · 8 months ago

How do I calculate Theta1 and Theta2, if it was not provided to us.

1 Upvote · Hide 3 Replies

Tom Mosher Mentor · 8 months ago

That is what we learn in Week 5.

1 Upvote

L

L.L. · 8 months ago · Edited by moderator

Thanks for your reply,but I has a little question,you said that a1 is m x n,and Theta1 is (h x n),so I think they can not be mutiplied,because maybe n≠h or m≠n,so it should be **{mentor edit: code removed}**

1 Upvote

Tom Mosher Mentor · 8 months ago

Perhaps you overlooked this line in the turorial:

Note: When you multiply by the Theta matrices, you'll have to use transposition to get a result that is the correct size.

The tutorial leaves a few things for the student to figure out . Otherwise, it would be merely a cookbook.

3 Upvote

A

Alexandr · 10 months ago

Hi Tom,

Thanks for helpful tutorials. I finished assignment but did not understand why we use indices instead of max values of a3.

Thanks,

0 Upvote · Hide 6 Replies

Tom Mosher Mentor · 10 months ago

We don't really care what the maximum value is.

We want to find which classifier gives us the maximum value.

2 Upvote

Julius Jacobsen · 9 months ago

JJ

My result for this test is off by one value. My vector returns the correct values save for the last number of my vector p. I get a '1' instead of a '2'. Any ideas as to what could be causing this miscalculation?

0 Upvote

Julius Jacobson · 9 months ago

JJ

Here is the offending row of a3:

0.633861 0.582628 0.522118 0.458987

0 Upvote

Tom Mosher Mentor · 9 months ago

Here are my results for that test case:

```
>> p = predict(Theta1, Theta2, X)
stopped in C:\Users\Tom\ml-od\machine-learning-ex3\ex3\predict.m
    at line 29
29: [v p] = max(a3, [], 2);
m =  8
debug> a3
a3 =

   0.53036   0.54588   0.55725   0.56352
   0.54459   0.54298   0.53754   0.52875
   0.49979   0.49616   0.49288   0.49024
   0.41357   0.42199   0.43736   0.45844
   0.37321   0.40368   0.44349   0.48911
   0.42073   0.45935   0.50210   0.54464
   0.50962   0.53216   0.55173   0.56659
   0.54882   0.55033   0.54738   0.54021
```

2 Upvote

JJ

Julius Jacobson · 9 months ago

Thanks again for the reply! I figured out what was wrong with my code. I am sure there is a place in Machine Learning heaven reserved for you!

0 Upvote

Tom Mosher Mentor · 9 months ago

Good news!

0 Upvote

XL

Xinghou Liu · 10 months ago

Hi Tom,

I am working on the predict.m, following you steps, I got the correct output:

'Training Set Accuracy: 97.520000'

I couldn't find any errors from the codes and I am pretty sure they are logically correct with your tips.

But after I submit the codes, they don't passed the assignment.

Any ideas on how can this happen?

0 Upvote · Hide 1 Reply

Tom Mosher Mentor · 10 months ago

There is an error in your code. It does not work correctly with the test case that the submit grader uses.

Try the additional test case, in this thread:

https://www.coursera.org/learn/machine-learning/discussions/iyd75Nz_EeWBhgpcuSIffw

0 Upvote

LL lennie leong · a year ago

HI Tom / mentors,

I am working on predict.m currently.

Using Tom's tutorial as a guide for the assignment, In Tom's tutorial guide, it mentions " Add a column of 1's to X (the first column), and it becomes 'a1'."

Currently, I have X that's a 16x2 matrix. By adding a column of 1's to X, does this mean that it still remains as 16x2 matrix with the first column filled with 1's.

Or, is it a matrix with 16x3 with the first column filled with 1's?

Thank you mentors for help in advanced!

0 Upvote · Hide 1 Reply

 Tom Mosher Mentor · a year ago

Add a column of 1's. Don't overwrite an existing feature with 1's.

0 Upvote

IC Irena Chanis · a year ago

Hi Tom,

i don't understand exactly how the size of the hidden layer (h) is determined (25 in the example). it seems it's arbitrary & not connected to input or output data. Couldn't understand it through the course videos also...

Another question: is it OK that i get the all history cost details of each iteration (attached Pic) for both NN and Multi class classification sections? (was able to submit all the tasks).

thanks,

```
Iteration     5 | Cost: 3.611087e-01
Iteration     6 | Cost: 3.598573e-01
Iteration     7 | Cost: 3.598359e-01
Iteration     8 | Cost: 3.598130e-01
Iteration     9 | Cost: 3.598122e-01
Iteration    10 | Cost: 3.598119e-01
Iteration    11 | Cost: 3.598119e-01
```

Irena.

0 Upvote · Hide 1 Reply

 Tom Mosher Mentor · a year ago

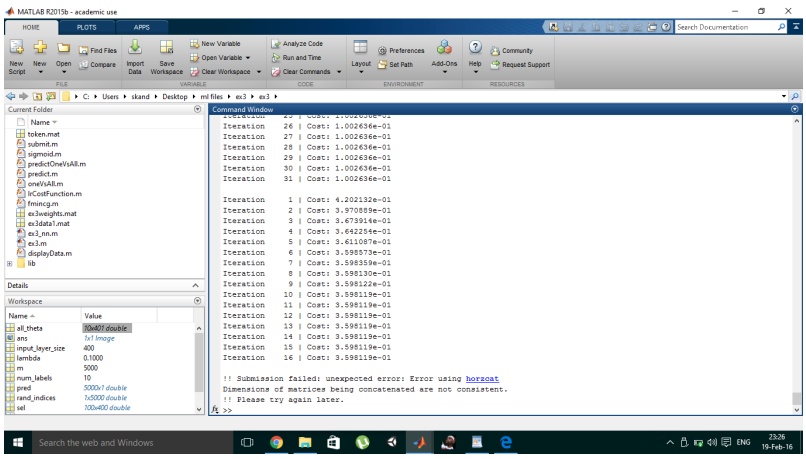
The course doesn't discuss how to set the number of hidden layer units. It is more art than science. There are lots of rules-of-thumb and no general agreement, other than it is typically a value between the number of input features and the number of output classifications. Experimentation is a good idea.

The iteration and cost values from fmincg() are supposed to all be printed on one line. It doesn't matter if they are on separate lines, but perhaps it is being caused by a missing semicolon somewhere in your cost function.

Just guessing, but it might also be due to how your command console is configured. You didn't mention whether you were using MATLAB or Octave, or what version, or if you're using the GUI or the CLI interface. You may be able to adjust some configuration setting.

4 Upvote

Sunil Skanda · a year ago



This is what I got when I tried to submit my code

0 Upvote · Hide 6 Replies

Tom Mosher Mentor · a year ago

Don't add bias to the Theta matrices - just to the input and hidden layer features.

The horzcat() error means you are not adding the columns of 1's correctly - it means "horizontal concatenation".

0 Upvote

Tom Mosher Mentor · a year ago · Edited

I recommend you study the tutorial. It tells you exactly what to do.

All you need to figure out is the order of operands and where to use transposition.

1 Upvote

Tom Mosher Mentor · a year ago

Also try testing your code using the test cases in this thread:

https://www.coursera.org/learn/machine-learning/discussions/0SxufTSrEeWPACIACw4G5w

0 Upvote

Sunil Skanda · a year ago

Thanks. So, z2 equals a1 multiplied by Theta1 right ?

0 Upvote

Tom Mosher Mentor · a year ago

Yes. You will need to use the correct order of the operands and add a transposition.

0 Upvote

Sunil Skanda · a year ago

1.Thanks. All I had was a minor error in using parantheses. Thanks for the help Tom!.

2. The pdf suggested that some interaction session will launch the images. I didn't find any.

3. We wrote a code similar to the predictOnevsAll function. But, I didn't get the prediction result(Training Set Accuracy) as expected.

Again Thanks for the help.

0 Upvote

Sunil Skanda · a year ago

Do we need to add bias unit to theta and theta as well?

0 Upvote · Reply

Sunil Skanda · a year ago

How man hidden units are there ?

0 Upvote · Hide 5 Replies

Tom Mosher Mentor · a year ago · Edited

Everything you need to know is provided via the Theta matrices.

0 Upvote

Sunil Skanda · a year ago

Ok. I wrote the code for predict.m as instructed and ran it in Matlab. But, It just stops after the prediction results for the previous function.

0 Upvote

Tom Mosher Mentor · a year ago

Per the tutorial, we're just doing a couple of matrix multiplications and running the sigmoid function. There's nothing there that should cause execution to stop.

If you're using for-loops, all sorts of havoc can occur.

1 Upvote

Sunil Skanda · a year ago

Yea, That's what I thought. I think I must have made a mistake in adding ones to the matrix X and the others. Can you tell me how to do that?

0 Upvote

Tom Mosher Mentor · a year ago

Add a column of 1's to a matrix:

Add a column of 1's to a matrix.

```
Q = magic(3)
m = size(Q,1)
Q = [ones(m,1) Q]
```

3 Upvote

Bernadberoy Remi · a year ago

I think there is a small error at point 2. Should be:

2. Multiply by Theta1and it becomes 'z2'. Take the sigmoid and it becomes 'a2'.

3. Add a column of 1's to 'a2'

0 Upvote · Hide 5 Replies

Tom Mosher Mentor · a year ago · Edited

I disagree. It is not a2 until you add the bias units, just as X becomes a1 after you add its bias unit.

2 Upvote

Bernadberoy Remi · a year ago

Ok but z2= theta1 * a1 without sigmoid. We take the sigmoid afterward, don't we?

0 Upvote

Tom Mosher Mentor · a year ago

Add the bias unit after computing the sigmoid. Otherwise, sigmoid(1) would turn it into 0.73106.

6 Upvote

jeevesh sharma · a year ago

Thnx this comment got me out of my agony, i was adding the bias before calculating sigmoid.

2 Upvote

AG

Atul Kumar Gupta · 8 months ago

Your comment helped me too . Thanks

0 Upvote

Hyatt Baker · a year ago

Well that was very helpful, thanks!

0 Upvote · Reply

CJ

Carsten Just · a year ago

Hi Tom a matlab technical question that i have had trouble with: when i, in the EX3 assignment, perform the predictOneVsAll function, from where do i get the all_theta input? when i run the OneVsAll funtion, the output is not stored anywhere.... Can you helP?

Best

Carsten

0 Upvote · Hide 2 Replies

Tom Mosher Mentor · a year ago

Read the ex3.m script. It calls your oneVsAll() function, saves the return value, and passes it to the prediction function.

1 Upvote

CJ

Carsten Just · a year ago

Oki thanks. Tried to just run the function itself... thanks!

0 Upvote