

Microsoft: DAT236x Deep Learning Explained

Help

Course > 2 | Multi-class Classification using Logistic Regression > Tutorials > Homework Assignment

Homework Assignment

☐ Bookmark this page

Homework Assignment

This homework assignment is based on Lab2_LogisticRegression.ipynb notebook.

Before doing this assignment, you should first complete the following tutorials, in sequence:

- Loading MNIST Dataset
- Reading MNIST Dataset
- Creating and Training Logistic Regression Model
- Predicting New Data

Homework Questions

Run the notebook and observe the average test error rate. Right now, your average test error hovers around 7.4x% to 7.6x%. Let's take this as our baseline and start experimenting.

Important: For the following questions, when asked to "train the model" or "run the notebook", you should re-run your whole notebook from the beginning so that you are building the model from scratch each time you change a parameter.

Question 1

0/1 point (graded)

First, let's experiment with the minibatch_size . Change the minibatch_size parameter from 64 to 512 during training, run the notebook and observe the resulting average test error. What is the observed average test error rate in comparison with the baseline?

Using the minibatch_size of 512 DECREASES the average test error ★

Homework Assignment Tutorials DAT236x Courseware edX
Using the minibatch_size of 512 INCREASES the average test error
Using the minibatch_size of 512 DOES NOT CHANGE the average test error
Submit You have used 1 of 1 attempt
Once you've finished experimenting with minibatch_size, revert back to the original setting, that is minibatch_size = 64.
Changing, adding, or removing features is one of the approaches in ML to explore different modeling options. Currently, the input features are scaled between 0-1 range. Let's now experiment removing the feature scaling.
Question 3 1/1 point (graded) Modify the code to remove this feature scaling, run the notebook and observe the resulting average test error. What is the observed average test error rate in comparison with the baseline?
Removing feature scaling DECREASES the average test error
■ Removing feature scaling INCREASES the average test error ✓
Removing feature scaling DOES NOT CHANGE the average test error
Submit You have used 1 of 1 attempt
Question 4
1/1 point (graded)

Which of the following reason suggests scaling of input features to be a desirable feature?

Efficient computation of model weights

Submit

	Homework Assignment Tutorials DAT236x Courseware edX
	■ Improve effectiveness of the learner (optimizer)
	Deal with more 0's than 1's in the input digits images
	Features scaling reduce the number of features
	Submit You have used 2 of 2 attempts
	Ince you've finished experimenting with removing feature scaling, revert back to the riginal setting.
c (b	urrently our features are specified as the expression input/255.0 (which we pass to the reate_model() function). This expression evaluates to a small CNTK computational graph built from tensors and functions), that is then extended with the layers specified within the
•	reate_model() call.
	reate_mode1() call. Question 5
1, C p th	
1, C p th	Question 5 /1 point (graded) hange the features passed to create_model() to be the square of each normalized ixel value. You can use the CNTK square() function for squaring values in a tensor. Run he notebook and observe the resulting average test error. Which of the following is true in
1, C p th	Question 5 (1 point (graded)) (hange the features passed to create_model() to be the square of each normalized ixel value. You can use the CNTK square() function for squaring values in a tensor. Run he notebook and observe the resulting average test error. Which of the following is true in comparison with the baseline? Using the square of each pixel values as the features DECREASES the average test

You have used 1 of 1 attempt

Question 6

1/1 point (graded)

Now change the features to the square root of each normalized pixel value. You can use the CNTK sqrt() function for this operation. Run the notebook and observe the resulting average test error. Which of the following is true in comparison with the baseline?

- Using the square root of each pixel values as the features DECREASES the average test error 🗸
- Using the square root of each pixel values as the features INCREASES the average test error
- Using the square root of each pixel values as the features DOES NOT CHANGE the average test error

Submit

You have used 1 of 1 attempt

Correct (1/1 point)

Question 7

1/1 point (graded)

Now, let's use all three of these sets of features at once (normalized pixel value, square of normalized value, and square root of normalized value). You can use the CNTK splice() function to concatenate these three sets of features together. Run the notebook and observe the resulting average test error. Which of the following is true in comparison with the baseline?

- Using the three sets of features DECREASES the average test error ✓
- Using the three sets of features INCREASES the average test error
- Using the three sets of features DOESNT CHANGE the average test error

Submit

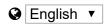
You have used 1 of 1 attempt

Question 8

Homework Assignment | Tutorials | DAT236x Courseware | edX 1/1 point (graded) With the addition of these features (square and square root), what is the shape of the input to the logistic regression model? (784, 3)(784*3,) (3,784)(, 784*3) Submit You have used 1 of 2 attempts Discussion **Show Discussion** Topic: Mod2-3 Tutorials / Homework Assignment

© All Rights Reserved





© 2012–2017 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open edX logos are registered trademarks or trademarks of edX Inc. | 粤ICP备17044299号-2

















