

Unit 7 - Week 6

Course outline

How does an NPTEL online course work?

Week 1

Week 2

Week 3

Week 4

Week 5

Week 6

Estimator API

Logistic Regression

Boosted Trees

Introduction to word embeddings

Quiz : Assignment 6

Week 6 Feedback

Week 7

Week 8

Text Transcripts

Download Videos

Assignment 6

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

Due on 2020-03-11, 23:59 IST.

1) Your first task is to perform regression using estimators on Boston Housing data. Please visit [this](#) notebook for answering the following questions. **1 point**
Follow the instructions given in the colab notebook to train a linear regression model on the data for 3000 steps. On evaluating on the test data, what is the range of average loss of the trained model?

- 50-60
 20-30
 40-50
 30-40

No, the answer is incorrect.
Score: 0

Accepted Answers:
30-40

2) Now train a DNN regressor model on the data for 3000 steps. Your network should have one hidden layer of 10 neurons. Leave the other parameters (except config) on their default values. On evaluating on the test data, what is the range of average loss of the trained model? **1 point**

- Between 150-350
 Between 50-150
 More than 300
 Less than 50

No, the answer is incorrect.
Score: 0

Accepted Answers:
More than 300

3) Train the model of Q2 for 6000 steps instead of 3000 steps. Evaluate the trained model on both the training data and the test data. What is the range of difference between the average loss on the training data and the test data? **1 point**

- More than 100
 Between 10-50
 Between 50-100
 Less than 10

No, the answer is incorrect.
Score: 0

Accepted Answers:
Between 50-100

4) Train a boosted trees regressor on the data for 50 steps. Set n_batches_per_layer as 1, center_bias as True and leave the other parameters (except config) on their default values. Evaluate the trained model on both the training data and the test data. What is the range of difference between the average loss on the training data and the test data? **1 point**

- More than 100
 Between 10-50
 Between 50-100
 Less than 10

No, the answer is incorrect.
Score: 0

Accepted Answers:
Less than 10

5) In the next 3 questions, you will observe the plotted graphs in the notebook and answer simple questions about DFCs. We have plotted the feature contributions for the 15th example of the test data. Which feature has the largest contribution (positive or negative) to the predicted value? **1 point**

- LSTAT
 AGE
 INDUS
 CHAS

No, the answer is incorrect.
Score: 0

Accepted Answers:
AGE

6) In continuation of Q5, if we increase the value of RM keeping the contributions constant, what happens to our predicted value? **1 point**

- Increases
 Decreases
 Not enough information

No, the answer is incorrect.
Score: 0

Accepted Answers:
Increases

7) How does the contribution of the RM feature change with an increase in its value from 6 to higher values? **1 point**

- Remains constant
 Increases
 Decreases

No, the answer is incorrect.
Score: 0

Accepted Answers:
Increases

8) Which of the following data augmentation techniques can be done using tf.keras.preprocessing.image.ImageDataGenerator? **1 point**

- Rotation
 Height shift
 Width shift
 Zoom
 Shearing
 Horizontal flip
 Vertical flip
 Brightness change

No, the answer is incorrect.
Score: 0

Accepted Answers:
Rotation
Height shift
Width shift
Zoom
Shearing
Horizontal flip
Vertical flip
Brightness change

9) What are the input and output shapes of an embedding layer with vocab_size = 1000 and embedding dimension = 25? **1 point**

- Input shape: (samples, sequence_length, 1000)
 Output shape: (samples, sequence_length, 1000, 25)
 Input shape: (samples, sequence_length)
 Output shape: (samples, sequence_length, 25)
 Output shape: (samples, sequence_length, 1000)

No, the answer is incorrect.
Score: 0

Accepted Answers:
Input shape: (samples, sequence_length)
Output shape: (samples, sequence_length, 25)

10) When we learn embeddings from a large corpus of data, we might learn embeddings that are biased in a certain way. A good set of embeddings should be free of any bias. Let e_x be the embedding for word x . Which of the options is correct about the following statements? **1 point**

- i. $e_{girl} - e_{boy}$
 ii. $e_{aunt} - e_{uncle}$
 iii. $e_{brother} - e_{sister}$

- i and ii should be approximately equal
 ii and iii should be approximately equal
 i, ii and iii should be approximately equal
 i, ii and iii should not be equal

No, the answer is incorrect.
Score: 0

Accepted Answers:
i and ii should be approximately equal