



Modules



Grades



Inbox



Discuss



Calc

Classifier

Video

ⓘ AQ 7.6: Activity Question 6 - Not Graded
Assignment

🔵 Practice Assignment 7 - Not Graded
Assignment

🔵 Graded Assignment 7
Assignment

🔵 Week 8 ▾

🔵 Week 9 ▾

🔵 Week 10 ▾

🔵 Week 11 ▾

🔵 Week 12 ▾

🔵 MOCK Unit ▾

🔵 Supplementary Contents ▾

Graded Assignment 7

The due date for submitting this assignment has passed.

Due on 2024-03-17, 23:59 IST.

You may submit any number of times before the due date. The final submission will be considered for grading.

You have last submitted on: 2024-03-17, 21:52 IST

Please note that the graded assignment is a continuation of the practice assignment.

1) Train the LogisticRegression model using SGDClassifier() with the following common settings.

1 point

1. No Regularization
2. random_state : 10
3. Iteration : 30

Capture the loss for each iteration and plot the iteration vs loss curve. For which of the following settings, the iteration vs loss curve converged quickly to zero loss?

- A. Set Learning rate : 0.01 and plot the curve and fit the model with `x_train_69` .
- B. Set learning rate to 0.000001 and fit the model with `x_train_69` .
- C. Keep the learning rate as 0.01. Scale the samples using StandardScaler() and fit the model with the pre-processed samples.
- D. Use the "invscaling" strategy for the learning rate with power_t = 1. Fit the model with x_train_69 (without pre-processing).

- ☐ A
- ☐ B
- ☒ C
- ☐ D

No, the answer is incorrect.

Score: 0

Accepted Answers:

A

2) In the above question, enter the iteration number for which the loss becomes zero and remains zero for the rest of the iterations.

6

No, the answer is incorrect.

Score: 0

Accepted Answers:

(Type: Numeric) 14

1 point

3) Create the classifier with the following settings

1. No Regularization
2. eta0 = 0.5
3. learning_rate = 'inv_scaling'
4. power_t = 0.5
5. iterations = 10
6. Shuffle = True
7. random_state = 10

Train the classifier with x_train_69. Answer the following questions?

How many false positives (FP) are there in the predictions made on training samples?

12

No, the answer is incorrect.

Score: 0

Accepted Answers:

(Type: Numeric) 7

1 point

4) Display all the False Positive samples. Could a human recognize them correctly?. Get the index of all FP samples in ascending order. Enter the first index.

80

No, the answer is incorrect.

Score: 0

Accepted Answers:

(Type: Numeric) 2167

1 point

5) Load the train subset of '20newsgroups' data.

1 point

Vectorize the data using TfidfVectorizer. Which of the following options represent the shape of the fitted and transformed dataset??

- ☐ (11310, 130507)
- ☐ (21314, 190807)
- ☐ (11514, 160107)
- ☒ (11314, 130107)

Yes, the answer is correct.

Score: 1

Accepted Answers:

(11314, 130107)

6) Split the training subset of fetch_20newsgroups data into train and validation sets using train_test_split with test_size = 0.3 and random_state = 0.

Use MultinomialNB to train the model.

Compute score on validation set.

0.8427

Yes, the answer is correct.

Score: 1

Accepted Answers:

(Type: Range) 0.8 , 0.85

1 point