CS-E4740 - Federated Learning D, Lectures, 28.2.2024-29.5.2024

Started on	Thursday, 7 March 2024, 1:39 PM
State	Finished
Completed on	Thursday, 7 March 2024, 1:58 PM
Time taken	18 mins 45 secs
Grade	**11.00** out of 11.00 (**100**%)

Question 1

Flag question Mark 2.00 out of 2.00 Correct

Linear regression learns the parameters $\mathbf{w} \in \mathbb{R}^d$ of a linear map by minimizing the average squared error loss incurred on a training set. Can the objective function of linear regression always be written as a quadratic function,

$$f(\mathbf{w}) = \mathbf{w}^T\mathbf{Q}\mathbf{w} + \mathbf{q}^T\mathbf{w} + c \text{ with some } \mathbf{Q} \in \mathbb{R}^{d\times d}, \mathbf{q} \in \mathbb{R}^d, c \in \mathbb{R}?$$

- a. Yes. ✔
- b. No.

Your answer is correct.

See discussion around Eq. 2.7 of the lecture notes (found here).

The correct answer is:
Yes.

Question 2

Flag question Mark 2.00 out of 2.00 Correct

Ridge regression learns the parameters $\mathbf{w} = \left(w_1, \ldots, w_n\right)^T$ of a linear map by minimizing the sum of the average squared error loss on a training set and the penalty term $\alpha\|\mathbf{w}\|_2^2$. Is it possible to formulate ridge regression as the minimization of a quadratic function of the form

$$f(\mathbf{w}) = \mathbf{w}^T\mathbf{Q}\mathbf{w} + \mathbf{q}^T\mathbf{w} \ ?$$

- a. Yes. ✔
- b. No.

Your answer is correct.

See Sec. 2.6. of the lecture notes (found here).

The correct answer is:
Yes.

Question 3

Flag question Mark 2.00 out of 2.00 Correct

Consider some ERM-based ML method that learns model parameters by minimizing the average loss on a training set. After completing this training process, we compute the resulting average loss on the training set (the training error) and the average loss on a validation set (the validation error). Can it happen that the validation error is smaller than the training error?

- a. Yes, for some training and validation sets. ✔
- b. No, the validation error is **always (for any dataset)** at least as high as the training error.
- c. Yes. Moreover, the validation error is **always (for any dataset)** smaller than the training error.

Your answer is correct.

See discussion around Figure 2.3. of the lecture notes (found here).

The correct answer is:
Yes, for some training and validation sets.

Question 4

Flag question Mark 1.00 out of 1.00 Correct

*This question refers to **the student task #1** in the "ML Basics" assignment.*

The coding assignment required you to read temperature measurements from a CSV file. Each temperature measurement is a data point characterized by seven features (location and time-stamp of measurement) and the measured temperature value as its label. The assignment required training a linear model on the first 100 data points and evaluating the validation error on the remaining data points. What were the observed values for training and validation errors in student task #1?

- a. The value of the **training** error was between 0 and 5.
- b. The value of the **validation** error was between 40 and 50.
- c. The value of the **validation** error was between 15 and 20. ✔
- d. The value of the **training** error was between 15 and 20. ✔

Your answer is correct.

Please, join the Slack channel (link) if you have any questions regarding the assignment.

The correct answers are:
The value of the **training** error was between 15 and 20.,

The value of the **validation** error was between 15 and 20.

Question 5

Flag question Mark 2.00 out of 2.00 Correct

*This question refers to **the student task #2** in the "ML Basics" assignment.*

The coding assignment required you to study the effect of augmenting the given features (latitude, longitude, day, month, year, hour, minute of measurement) by their polynomial combinations up to a given maximum degree d.
For each choice of $d = 1, 2, 3$, you had to learn the parameters of a linear model and then compute the training and validation errors (average squared error loss) of the learnt model parameters. How do the training and validation errors change with increasing polynomial degrees d?

- a. The validation error might increase with increasing d. ✔
- b. The training error might increase with increasing degree d.
- c. The validation error never increases with increasing d.
- d. The training error never increases with increasing d. ✔

Your answer is correct.

Increasing the degree d results in a larger hypothesis space and, in turn, can only result in a decrease of the training error but never (ignoring imperfections such as numerical errors during optimization) in an increase of the training error. However, increasing d might result in increased validation error due to overfitting. More discussion can be found in Ch. 6 and 7 of AJ, "Machine Learning: The Basics", Springer, 2022 (preprint here).

The correct answers are:
The training error never increases with increasing d.,

The validation error might increase with increasing d.

Question 6

Flag question Mark 2.00 out of 2.00 Correct

*This question refers to **the student task #3** in the "ML Basics" assignment.*

The coding assignment required you to use ridge regression to learn the parameters of a linear model using the original features and their polynomial combinations. You had to determine the resulting training and validation errors for different choices for the regularization parameter α of ridge regression. How do they vary with increasing values of α?

- a. The training error always decreases with increasing α.
- b. The validation error always increases with increasing α.
- c. The training error might increase with increasing α. ✔
- d. The validation error might decrease with increasing α. ✔
- e. The training error does not change with increasing α.

Your answer is correct.

Increasing the regularization parameter α in ridge regression corresponds to a shrinking of the effective hypothesis space and, in turn, can only result in an increase of the training error. However, increasing α might result in a reduced validation error due to making the training more robust against overfitting. More discussion can be found in Ch. 7 of AJ, "Machine Learning: The Basics", Springer, 2022 (preprint here).

The correct answers are:
The training error might increase with increasing α.

,

The validation error might decrease with increasing α.