

Machine Learning for Bioinformatics

Exercise Sheet

Freie Universität Berlin, SoS 2024

Week 10 · Assignment on 19.06.2024. Submit until 30.06.2024 11pm.

Please note that the jupyter notebook must be submitted
along with the exercise sheet!¹

Name:

Matriculation no.:

Name:

Matriculation no.:

Model selection

Assignment 1. *PRESS for linear and ridge regression*

1. Complete the implementation of *MyRidgeRegressor*. Points: 0
2. What is the minimum and maximum value of the PRESS statistic for varying p ? (Format: N.NN) Points: 1

Minimum PRESS:

Maximum PRESS:

3. What is the minimum and maximum value of LOO-CV with mean squared error for varying p ? (Format: N.NN) Points: 1

Minimum LOO-CV:

Maximum LOO-CV:

Assignment 2. *Degrees of freedom (DF)*

1. Complete the implementation of the function *df*. Points: 0
2. What is the degree of freedom for $\alpha = 0.5$? (Format: N.NN) Points: 1

DF:

3. What is the degree of freedom for $p = 1, 2, \dots, 50$? Points: 1

n : p : $\max(n, p)$: $\min(n, p)$:

¹No points are awarded if an answer is only partially correct. Answers must be supported by results from the jupyter notebook.

Assignment 3. *Regularization paths*

1. At the beginning of the ℓ_1 regularization path all features have coefficients equal to zero. Which is the first feature (1-based column index in F) with non-zero coefficient? Points: 1

Feature:

2. Which of the following statements holds for the ℓ_2 regularization path? Points: 1

The sum of absolute values of the coefficients increases with α .

Correct:

Some of the coefficients change sign.

Correct:

The curves are smoother than for the ℓ_1 path.

Correct:

Similar to the ℓ_1 path, there are α -regions of non-zero length where coefficients return back to zero.

Correct:

Assignment 4. *Implicit regularization*

1. Complete the implementation of the function *compute_noisy_polynomial_features*. Points: 0
2. For which noise level(s) (scale / standard deviation) do you observe the minimum average mse after the interpolation threshold? Points: 2

0.01: 0.02: 0.05:

Maximum number of points: 8