

Which of the following describes underfitting?

- A. High variance and high bias
- B. High variance and low bias
- C. Low variance and high bias
- D. Low variance and low bias

How does Ridge regularization mitigate overfitting?

- A. Reduce bias and variance
- B. Increase bias a bit and reduce variance
- C. Reduce bias and increase variance a bit
- D. Increase bias and variance

How will training set's R^2 move, as the model complexity increases?

- A. It goes down and then goes up
- B. It goes up and then goes down
- C. It always goes down
- D. It always goes up

Gloria managed to get last year's final exam paper. She decided to memorize the solutions. However, she eventually performed badly on this year's final exam paper. What happened here?

- A. Overfitting
- B. Train set's loss is much larger than test set's loss
- C. She did not aim to minimize train set's loss function
- D. Underfitting

How do we construct different trees for random forest?

- A. Use different observations and different input variables for each tree
- B. Use the same observations and different input variables for each tree
- C. Use different observations and the same input variables for each tree
- D. Use the same observations and the same input variables for each tree

The target variable y is binary (1 or 0). The percentage of $y=0$ in the data set is p . Which p value gives the highest Gini impurity?

- A. 0%
- B. 50%
- C. 100%
- D. It depends

Why is Linear Regression bad for binary classification problem?

- A. The equation with linear combination is too simple to fit the data set
- B. Target variable can be $(-\infty, \infty)$ while binary classification outcome is just 0 or 1
- C. The accuracy rate tends to be poor
- D. The cut-off value for regression output to be considered 1 or 0 is fixed to 0.5, which is a bit restrictive

Why do we NOT minimize test set's loss?

- A. Because test set is meant to be future data, and we cannot minimize unseen data
- B. Because minimizing test set's loss will lead to overfitting
- C. Because test set is irrelevant to the model's performance

D. Because train set's loss is all we care about

The target variable y is binary with 30% in category 0 and 70% in category 1. If you are required to predict all targets with only one category (either 0 or 1), what is your best predictions' training accuracy rate?

- A. 30%
- B. 70%
- C. 50%
- D. It depends

Answers:

C B D A A B B A B