03. Self-assessment questions

September 26, 2025

1. What does it mean that a model has an excessive variance? What is, instead, the variance of a feature (write the formula)? Are the variance of a model and the variance of a feature the same thing? Or are they different concepts?
2. Which methods to do feature selection have we seen up to now? Tell at least three of them.
3. In the types of models you have seen so far, what are the methods and the hyper-parameters that allow to increase or decrease their complexity?
4. How can you understand if a model is overfitting or underfitting? How can you improve the model, in case of overfitting or underfitting?
5. What is the goal of regularization?
6. Describe a correct procedure to choose a good regularization coefficient.
7. In Ridge regression, which regressions coefficients do you obtain when you fix the regularization coefficient $\alpha = 0$? And if $\alpha \to \infty$
8. In order to regularize a model, we modify the loss function adding a "regularization term". Is this claim correct? If yes, should this term be added during training, or test or both?
9. In case of polynomial regression, describe a correct procedure to choose a good pair of regularization coefficient and polynomial degree
10. Let us consider a polynomal regression model and suppose to increase the polynomial degree p. What happens to the training error? Does it increase or decrease or does it depend? And what about the test error?
11. Is scaling strictly necessary or useful when applying linear regression? And when applying Ridge Regression? Why?
12. Consider a linear regression model and suppose that our dataset is not scaled. Is it correct to affirm that the features whose coefficients are the smallest are the least important? If you scale the dataset, is this affirmation correct? Why? Provide an example.

13. Regarding the gradient descent in logistic regression, we compute the gradient of which function, exactly?

14. When using logistic regression, we use the gradient descent during training or during inference?

15. In logistic regression, the model gives directly the predicted class of a sample or the probability? In the latter case, how do we go from the probability to the predicted class?

16. What is the difference between full/stochastic/batch gradient descent?

17. Is logistic regression a linear or non-linear classifier? Why? What does "linear classifier" mean?

18. What is the softmax function? Write its formula. Why is it employed?

19. Write the formula of the cross-entropy.

20. Suppose you have an unbalanced dataset and you directly train a model on top of it. Suppose you obtain a good accuracy, close to 99.9%. Would you say you have a good model? Why? Give an example in which, despite a very high classification accuracy, your model is bad.