1. Supervised learning problem statement. Regression and classification problems. What's the difference?
2. Linear model for regression problem in matrix notation. Mean Squared Error loss function.
3. What is the gradient? How is it used in optimization?
4. Write down gradient descent step for linear model and MSE for one-dimensional case.
5. What is validation? Cross validation?
6. What is regularization? How does L1 regularization differ from L2?
7. What are precision and recall metrics?
8. How does the bagging work? What is Random Forest? What's the difference between Random forest and Bagging?
9. How are parameters different from hyperparameters? E.g. what are parameters in linear models and decision trees? Hyperparameters?
10. What is boosting. Gradient boosting? How should a model be trained on step t+1 in a gradient boosting ensemble?
11. What is backpropagation? How does it work? E.g. how would gradient propagate through a linear layer? Through ReLU?
12. How does convolutional layer work? What are the kernels (filters) in the covolutional layer? Are they independent?
13. What is dropout? How does it work in a neural network? Does it change its behaviour on the inference (test) stage?
14. What is batch normalization? How does it work? How does it affect the learning rate? Does it change its behaviour on the inference (test) stage?
15. How does RNN work? Can you combine CNN and RNN? What is the difference between Vanilla RNN and LSTM?
16. State the unsupervised problem statement. What is clustering? How does k-means algorithm work?