1. What is the loss function used in regression and classification
2. What are assumptions in linear regression
3. How can the feature set be reduced
4. Why L1=0 but L2 -> 0
5. What is central limit theorem
6. How to tackle imbalanced data
7. What are metrics in multiclass classification in unbalanced data
8. What is variance bias tradeoff
9. How to prevent overfitting
10. How to split in decision tree
11. What is backpropagation
12. How to do dimentionality reduction
13. What is the cost function used in logistic regression
14. What layers are in deep learning
15. Why is bias term used in linear regression
16. Why is LSTM better than RNN
17. How is hypothesis testing done
18. What is grid search
19. What are other methods of hyperparameter optimization/tuning
20. Parameter vs hyperparameter
21. Why is gini impurity better than entropy
22. When to use transfer learning and when not to
23. What are anomaly detection methods
24. In 5 fold cross validation how is metrics decided finally
25. How to make AUCROC plot from data
26. Is feature scaling required in Logistic regression
27. Is logistic regression sensitive to missing values
28. What are hyperparameter tuning techniques for decision trees
29. What is impact of outliers on decision trees
30. Does decision tree require scaling
31. What is stopping condition in k means clustering
32. How is xgboost better than gbm
33. Mention steps for outlier treatment
34. What is pruning in decision tree
35. How is choice of cluster size in k means clustering done
36. Why Relu is used in inner layer and softmax in last layer
37. How to work embeddings work
38. Mention some graph based algorithms
39. What is reinforcement learning, give example
40. Correlation vs causation
41. Mention some graph based algorithms and how do they work