Machine Learning

1. What are the different types of machine learning?

Primarily there are 3 types of ML

a. **Supervised** Supervised learning algorithms are generally used for solving classification and regression problems.
   Some of the most popularly used supervised learning algorithms are:

   1. Linear Regression
   2. Logistic Regression
   3. Support Vector Machine
   4. K Nearest Neighbor
   5. Decision Tree
   6. Random Forest
   7. Naive Bayes

b. **Unsupervised** learning finds patterns and understands the trends in the data to discover the output. So, the model tries to label the data based on the features of the input data.
   Some of the common examples of unsupervised learning are:

   1. K Means Clustering
   2. Hierarchical Clustering
   3. DBSCAN
   4. Principal Component Analysis

c. Reinforcement

2. What is deep learning, and how does it contrast with other machine learning algorithms?
3. What are the differences between machine learning and deep learning?
4. Explain the confusion matrix with respect to machine learning algorithms.
5. What is the difference between artificial intelligence and machine learning?
6. What's the trade-off between bias and variance?
7. Explain the difference between L1 and L2 regularization.
8. What's your favorite algorithm, and can you explain it to me in less than a minute?
9. How is KNN different from k-means clustering?
10. What is cross validation and what are different methods of using it?

11. Explain how a ROC curve works.
12. What's the difference between probability and likelihood?
13. What's the difference between a generative and discriminative model?
14. How is a decision tree pruned?
15. How can you choose a classifier based on a training set size?
16. What methods for dimensionality reduction do you know and how do they compare with each other?
17. Define precision and recall.
18. What's a Fourier transform?
    Well, it is a mathematical concept. Basically, it is a generic method that is used for decomposing the generic functions into a superposition of other functions that are generally symmetric in nature. When it comes to finding the speeds of cycles and amplitudes, they are widely adopted in machine learning. Fourier Transformation is also used for solving some of the very complex problems of mathematics.

19. What's the difference between Type I and Type II error?
20. When should you use classification over regression?
21. How would you evaluate a logistic regression model?
22. What is Bayes' Theorem? How is it useful in a machine learning context?
23. Describe a hash table.

Technical Skills Questions

24. How would you handle an imbalanced dataset?
25. How do you handle missing or corrupted data in a dataset?
26. Do you have experience with Spark or big data tools for machine learning?
27. Pick an algorithm. Write the pseudocode for a parallel implementation.
28. Which data visualization libraries do you use? What are your thoughts on the best data visualization tools?
29. Given two strings, A and B, of the same length n, find whether it is possible to cut both strings at a common point such that the first part of A and the second part of B form a palindrome.
30. How would you build a data pipeline?
31. How would you implement a recommendation system for our company's users?
32. Can you explain your approach to optimizing auto-tagging?
33. Suppose you are given a data set that has missing values spread along 1 standard deviation from the median. What percentage of data would remain unaffected and why?
34. Suppose you found that your model is suffering from low bias and high variance. Which algorithm do you think could tackle this situation and why?

35. You are given a data set. The data set contains many variables, some of which are highly correlated and you know about it. Your manager has asked you to run PCA.
36. Would you remove correlated variables first? Why?

37. What are the advantages and disadvantages of neural networks?
38. How would you go about understanding the sorts of mistakes an algorithm makes?
39. Explain the steps involved in making decision trees.

Personal Questions

40. How do you keep informed of developments in machine learning?
41. How do you think quantum computing will affect machine learning?
42. Is machine learning a science or an art?
43. What are you passionate about?
44. How do you handle stress and pressure?
45. What makes you unique?
46. What motivates you?
47. Tell me about yourself.
48. How would you describe yourself?
49. How do you evaluate success?
50. What is your greatest weakness?
51. What is your greatest strength?
52. Describe your work ethic.
53. Why do you want to work here?

Leadership and Communication

54. Tell me about a time when you had to convince others to take your position on a specific matter. What was the outcome?
55. How do you make sure projects and tasks stay on schedule?
56. How do you handle disagreements on your team?
57. Tell me about a time when something went wrong at work and you took control.
58. How do you deal with people who disagree with you?
59. How would you go about simplifying a complex issue in order to explain it to a client or colleague?
60. How would you go about persuading someone to see things your way at work?
61. How would you go about explaining a complex idea/problem to a client who was already frustrated?
62. What would you do if there was a breakdown in communication at work?
63. Talk about a successful presentation you gave and why you think it did well.
64. Talk about a time when you made a point that you knew your colleagues would be resistant to.
65. Is it more important to be a good listener or a good communicator?

Behavioral

66. Give me an example of how you've used your data analysis to change behavior. What was the impact, and what would you do differently in retrospect?
67. Give an example of a problem you solved (or tried to solve) with machine learning.
68. Tell me about a time when you had to think outside the box to complete a task. Were you successful?
69. Can you describe a time when you had to develop a complex algorithm?
70. Can you tell me about a major success you had with a machine learning project?
71. What's the most difficult decision you've had to make recently and how did you come to that decision?
72. Tell me about a time you were under a lot of pressure. What was going on, and how did you get through it?
73. Tell me about a time you had a conflict at work.
74. Give an example of when you made a mistake at work.
75. Describe a time when you disagreed with a client. How did you handle it?
76. Tell me about a time you set a goal for yourself. How did you go about ensuring that you would meet your objective?
77. Describe a time when you saw a problem and took the initiative to correct it rather than waiting for someone else to do it.
78.
79. Questions From Top Companies (Amazon, Google, Facebook, Microsoft)
80.
81. What are the differences between generative and discriminative models?
82. How would you weigh nine marbles three times on a balance scale to select the heaviest one?
83. What's the difference between MLE and MAP inference?
84. Why did you use this particular machine learning algorithm in your project?
85. What is K-means algorithm?
86. Describe a time when you let go of a short-term goal for a long-term goal.
87. What's the difference between the summaries of a Logistic Regression and SVM?
88. Explain ICA and CCA. How do you get a CCA objective function from PCA?
89. What is the relationship between PCA with a polynomial kernel and a single layer autoencoder? What if it is a deep autoencoder?
90. What is A/B testing in machine learning?
91. What is activation function in machine learning?
92. How would you build, train and deploy a system to detect if multimedia and/or ad content being posted violated terms or contained offensive materials?
93. How do you solve a disagreement with a team member?
94. What is the bias-variance tradeoff? How is it expressed using an equation?
95. Describe the idea behind boosting. Give an example of one method and describe one advantage and disadvantage.
96. Formulate the background behind an SVM, and show the optimization problem it aims to solve.