Question Bank_ML(303105353)

1. Define Machine Learning and explain its importance in modern applications.
2. Differentiate between Artificial Intelligence, Machine Learning, and Deep Learning.
3. What are the main types of learning paradigms in Machine Learning?
4. Explain the concept of PAC (Probably Approximately Correct) learning with an example.
5. State and explain the components of a Machine Learning system.
6. Describe the steps involved in building a Machine Learning model.
7. What are version spaces in hypothesis learning?
8. Explain the role of probability in Machine Learning.
9. Give an example to demonstrate conditional probability in a classification problem.
10. Write a short note on bias-variance tradeoff.
11. Explain the importance of data collection in ML projects.
12. What are missing values? How can they be handled?
13. Explain data normalization and its types.
14. Give an example of feature scaling using Min-Max normalization.
15. What is outlier analysis? How does Z-Score help in detecting outliers?
16. Define feature selection and its importance.
17. What are hyperparameters? Give examples.
18. Discuss model selection and cross-validation techniques.
19. Explain how to tune hyperparameters using Grid Search and Random Search.
20. Describe the importance of visualization in Machine Learning.
21. Define supervised learning with suitable examples.
22. Differentiate between linear and nonlinear models with examples.
23. Explain multi-class and multi-label classification with real-life examples.
24. Write the equation for simple linear regression and explain its parameters.
25. What is the difference between linear and multiple linear regression?
26. Compute the regression line for the data: X=[1,2,3], Y=[2,4,5].
27. Explain Naïve Bayes Classifier with an example.
28. Define entropy and information gain used in Decision Trees.
29. What is the difference between ID3 and CART algorithms?
30. Explain the concept of error bounds in supervised learning.
31. What is the K-Nearest Neighbors (KNN) algorithm?
32. Compute Euclidean distance between points (2,3) and (5,7).
33. What is logistic regression? How does it differ from linear regression?
34. Explain the concept of perceptron learning rule.
35. Differentiate between single-layer and multi-layer perceptron.
36. Compute the perceptron output for inputs x1=1, x2=0 with weights w1=0.5, w2=0.3, bias=0.2.
37. What is Support Vector Machine (SVM)? Explain linear and nonlinear SVM.
38. Describe the use of kernel trick in SVM.
39. Compare SVM and Logistic Regression models.
40. Explain the concept of semi-supervised learning with an example.
41. What is clustering? Mention its types.

42. Explain the K-Means clustering algorithm with an example.
43. Compute one iteration of K-Means for data points [1,2,3,10,11,12] with k=2.
44. Differentiate between K-Means and K-Modes clustering.
45. Explain hierarchical clustering and dendrograms.
46. What is density-based clustering? Give one example algorithm.
47. Explain the working of Self Organizing Maps (SOM).
48. What is Expectation Maximization (EM) algorithm?
49. Describe Principal Component Analysis (PCA) and its use in dimensionality reduction.
50. Calculate the mean and covariance matrix for given data X=[[2,3],[4,5],[5,6]].
51. Define reinforcement learning and its main components.
52. What is the difference between supervised and reinforcement learning?
53. Explain reward, policy, and value functions.
54. What is Q-learning? Write its update rule.
55. Give an example of reinforcement learning in real-world applications.
56. Define exploration and exploitation trade-off.
57. What is temporal difference learning?
58. Compute the Q-value update for α=0.5, γ=0.9, reward=10, next max Q=15, old Q=8.
59. Explain the difference between model-based and model-free RL.
60. What is the Bellman equation in RL?
61. Define accuracy, precision, recall, and F1-score.
62. Construct a confusion matrix for a binary classification example.
63. What is ROC curve? How is AUC interpreted?
64. Compute precision and recall for TP=50, FP=10, FN=5.
65. Explain K-Fold Cross Validation with example.
66. What is the purpose of significance tests in ML model evaluation?
67. Define overfitting and underfitting with examples.
68. What is the role of regularization in preventing overfitting?
69. Explain error correction in perceptrons.
70. Describe bias, variance, and irreducible error.
71. What is ensemble learning? Why is it used?
72. Explain bagging and boosting with examples.
73. Describe the working of Random Forest algorithm.
74. Explain AdaBoost and how weak learners are improved.
75. What is XGBoost and why is it popular?
76. Compare bagging and boosting approaches.
77. Compute a simple average ensemble output for models giving predictions [0.8, 0.6, 0.9].
78. Explain how feature importance is calculated in Random Forest.
79. What is stacking in ensemble learning?
80. Discuss advantages and limitations of ensemble methods.
81. What are the main challenges in applying ML to real-world data?
82. Explain the role of data preprocessing in model performance.
83. What is feature engineering? Give examples.
84. What are dimensionality reduction techniques apart from PCA?
85. Define one-hot encoding and label encoding with examples.
86. Write a Python code snippet to split data into training and test sets using sklearn.

87. How is model performance evaluated using cross-validation?
88. Explain the use of pipelines in scikit-learn.
89. What are the ethical concerns related to Machine Learning applications?
90. Describe the workflow of a typical ML project.
91. Compute mean squared error for actual=[2,3,4], predicted=[2.5,3.5,3.0].
92. Compute correlation coefficient for X=[1,2,3], Y=[2,4,6].
93. Calculate Z-score for value=70, mean=60, std=5.
94. What is gradient descent? Derive its update rule.
95. Compute one step of gradient descent for cost function J= (y - wx)^2 at x=2, y=4, w=1, learning rate=0.1.
96. Explain stochastic, mini-batch, and batch gradient descent.
97. What is regularization? Compute cost with L2 regularization for w=[2,3], λ=0.1.
98. Compute accuracy for a model with 85 correct predictions out of 100.
99. Discuss how to handle class imbalance in classification tasks.
100. What are the future trends and applications of Machine Learning?
101. Consider the following **confusion matrix**:

	Predicted Positive	**Predicted Negative**
Actual Positive	50	10
Actual Negative	5	35

Calculate **Accuracy** and **Precision.**

102. In a classification problem, a **logistic regression** model gives the following probabilities for five observations: [0.8, 0.3, 0.6, 0.9, 0.2]. If the cutoff is **0.5**, determine which observations are classified as positive and compute **accuracy**, given the actual labels are [1, 0, 1, 1, 0].

103. Perform one iteration of **K-Means clustering** for data points (2,4), (3,5), (10,8), (11,9) with initial centroids (2,4) and (10,8).
Assign points to clusters and compute new centroids.

104. In **K-Means clustering**, the coordinates of two cluster centers are (2,3) and (8,7). Find the **Euclidean distance** between them.