Software Engineering

S.Y.B.Sc. – I.T. Semester - IV

Introduction to Machine Learning

1. What is Machine Learning?
   A) The study of mechanical systems
   B) The study of computer algorithms that improve automatically through experience
   C) The study of human learning behavior
   D) The study of machine intelligence

   Answer: B) The study of computer algorithms that improve automatically through experience

2. Which of the following is NOT a type of Machine Learning?
   A) Supervised Learning
   B) Unsupervised Learning
   C) Reinforcement Learning
   D) Deterministic Learning

   Answer: D) Deterministic Learning

3. Which algorithm is commonly used for classification problems in Machine Learning?
   A) Linear Regression
   B) K-Means
   C) Decision Trees
   D) Apriori

   Answer: C) Decision Trees

4. Which type of learning algorithm predicts outputs based on input-output pairs?
   A) Supervised Learning
   B) Unsupervised Learning
   C) Semi-supervised Learning
   D) Reinforcement Learning

   Answer: A) Supervised Learning

5. Which of the following is used for reducing the dimensionality of data in Machine Learning?
   A) Principal Component Analysis (PCA)
   B) K-Nearest Neighbors (KNN)

C) Random Forest
D) Gradient Descent

Answer: A) Principal Component Analysis (PCA)

6. Which algorithm is used for finding hidden patterns or grouping similar instances together?
   A) Decision Trees
   B) K-Means
   C) Support Vector Machines (SVM)
   D) AdaBoost

   Answer: B) K-Means

7. What is the main goal of Regression analysis in Machine Learning?
   A) Predicting categorical variables
   B) Predicting continuous variables
   C) Grouping similar data points together
   D) Classifying data into different categories

   Answer: B) Predicting continuous variables

8. Which of the following is an example of a classification problem?
   A) Predicting house prices
   B) Identifying spam emails
   C) Predicting stock prices
   D) Predicting temperature

   Answer: B) Identifying spam emails

9. Which algorithm aims to find the best fitting line to describe the relationship between input and output variables?
   A) Decision Trees
   B) Logistic Regression
   C) K-Means
   D) Random Forest

   Answer: B) Logistic Regression

10. What does the term "overfitting" mean in Machine Learning?
    A) The model generalizes well to unseen data
    B) The model captures noise and irrelevant patterns from the training data
    C) The model doesn't learn anything from the training data
    D) The model is too simple to capture the underlying patterns in the data

    Answer: B) The model captures noise and irrelevant patterns from the training data

11. Which of the following is NOT a measure of model performance in Machine Learning?
    A) Accuracy
    B) Precision
    C) Recall
    D) Diversity

    Answer: D) Diversity

12. Which method is used for evaluating the performance of a classification model?
    A) Mean Absolute Error (MAE)
    B) Root Mean Squared Error (RMSE)
    C) Confusion Matrix

D) R-squared

Answer: C) Confusion Matrix

13. What does the term "bias" refer to in Machine Learning?
    A) The difference between predicted and actual values
    B) The variance in the data
    C) The error introduced by approximating a real-world problem
    D) The tendency of a model to consistently learn the wrong thing

    Answer: D) The tendency of a model to consistently learn the wrong thing

14. What is the purpose of feature scaling in Machine Learning?
    A) To remove irrelevant features from the dataset
    B) To reduce the number of features in the dataset
    C) To standardize the range of features in the dataset
    D) To increase the complexity of the model

    Answer: C) To standardize the range of features in the dataset

15. Which algorithm is used for solving both classification and regression problems?
    A) Decision Trees
    B) K-Means
    C) Gradient Boosting
    D) Support Vector Machines (SVM)

    Answer: C) Gradient Boosting

16. Which type of Machine Learning algorithm is suitable for predicting future stock prices?
    A) Supervised Learning
    B) Unsupervised Learning
    C) Reinforcement Learning
    D) Semi-supervised Learning

    Answer: A) Supervised Learning

17. Which technique is used to avoid overfitting in Machine Learning models?
    A) Regularization
    B) Feature engineering
    C) Cross-validation
    D) All of the above

    Answer: D) All of the above

18. Which of the following is NOT a step in the Machine Learning pipeline?
    A) Data preprocessing
    B) Model training
    C) Model testing
    D) Model deployment

    Answer: C) Model testing

19. What does the term "ensemble learning" refer to in Machine Learning?
    A) Training multiple models independently and then combining their predictions
    B) Training a single model with multiple features
    C) Training a single model with multiple hidden layers
    D) Training multiple models sequentially

    Answer: A) Training multiple models independently and then combining their predictions

20. Which algorithm is used for reducing the complexity of a decision tree model?
   A) Bagging
   B) Boosting
   C) Pruning
   D) Stacking

   Answer: C) Pruning

21. Which evaluation metric is suitable for imbalanced datasets?
   A) Accuracy
   B) Precision
   C) Recall
   D) F1-score

   Answer: D) F1-score

22. What is the purpose of regularization techniques in Machine Learning?
   A) To reduce the computational cost of the model
   B) To increase the complexity of the model
   C) To prevent overfitting by penalizing large coefficients
   D) To speed up the training process

   Answer: C) To prevent overfitting by penalizing large coefficients

23. Which algorithm is used for solving the exploration-exploitation trade-off problem in reinforcement learning?
   A) Q-Learning
   B) Deep Q-Networks (DQN)
   C) Policy Gradient methods
   D) Monte Carlo Tree Search (MCTS)

   Answer: D) Monte Carlo Tree Search (MCTS)

24. What is the primary difference between bagging and boosting algorithms?
   A) Bagging combines multiple weak learners sequentially, while boosting combines them in parallel.
   B) Bagging combines multiple weak learners in parallel, while boosting combines them sequentially.
   C) Bagging uses a weighted average of weak learners' predictions, while boosting uses a simple average.
   D) Bagging focuses on reducing bias, while boosting focuses on reducing variance.

   Answer: B) Bagging combines multiple weak learners in parallel, while boosting combines them sequentially.

25. Which algorithm is suitable for handling high-dimensional data?
   A) Decision Trees
   B) K-Means
   C) Support Vector Machines (SVM)
   D) Logistic Regression

   Answer: C) Support Vector Machines (SVM)

26. What is the main difference between stochastic gradient descent and batch gradient descent?
   A) Stochastic gradient descent updates the model parameters after processing the entire training dataset, while batch gradient descent updates them after each data point.
   B) Stochastic gradient descent updates the model parameters after processing a subset of

the training dataset, while batch gradient descent updates them after each data point.

    C) Stochastic gradient descent updates the model parameters after processing each data point, while batch gradient descent updates them after processing the entire training dataset.

    D) Stochastic gradient descent updates the model parameters using a fixed learning rate, while batch gradient descent adjusts the learning rate dynamically.

    Answer: C) Stochastic gradient descent updates the model parameters after processing each data point, while batch gradient descent updates them after processing the entire training dataset.

27. Which of the following is NOT a disadvantage of using deep learning models?
    A) They require large amounts of labeled data for training.
    B) They are computationally expensive and require powerful hardware.
    C) They are prone to overfitting, especially with small datasets.
    D) They are interpretable and provide insights into the decision-making process.

    Answer: D) They are interpretable and provide insights into the decision-making process.

28. What is the purpose of the activation function in a neural network?
    A) To compute the gradient of the loss function
    B) To normalize the input data
    C) To introduce non-linearity into the model
    D) To regularize the weights of the network

    Answer: C) To introduce non-linearity into the model

29. Which type of neural network architecture is used for image recognition tasks?
    A) Convolutional Neural Networks (CNNs)
    B) Recurrent Neural Networks (RNNs)
    C) Long Short-Term Memory (LSTM) networks
    D) Autoencoder networks

    Answer: A) Convolutional Neural Networks (CNNs)

30. Which algorithm is commonly used for natural language processing tasks such as language translation and sentiment analysis?
    A) Recurrent Neural Networks (RNNs)
    B) Convolutional Neural Networks (CNNs)
    C) Transformer networks
    D) Autoencoder networks

    Answer: C) Transformer networks

31. Which technique is used for generating synthetic data to augment the training dataset in deep learning?
    A) Dropout
    B) Batch normalization
    C) Data augmentation
    D) Transfer learning

    Answer: C) Data augmentation

32. What is the purpose of the softmax function in a neural network?
    A) To compute the probability distribution over multiple classes
    B) To normalize the output of the network
    C) To introduce non-linearity into the model
    D) To regularize the weights of the network

    Answer: A) To compute the probability distribution over multiple classes

33. Which of the following is a drawback of using Recurrent Neural Networks (RNNs) for long sequences?
    A) They are computationally expensive and require powerful hardware.
    B) They suffer from the vanishing gradient problem.
    C) They are prone to overfitting, especially with small datasets.
    D) They cannot capture long-term dependencies effectively.

    Answer: D) They cannot capture long-term dependencies effectively.

34. What is the purpose of the attention mechanism in Transformer networks?
    A) To compute the gradient of the loss function
    B) To normalize the input data
    C) To focus on relevant parts of the input sequence
    D) To regularize the weights of the network

    Answer: C) To focus on relevant parts of the input sequence

35. Which type of neural network architecture is used for sequential data processing tasks such as time series forecasting and speech recognition?
    A) Convolutional Neural Networks (CNNs)
    B) Recurrent Neural Networks (RNNs)
    C) Long Short-Term Memory (LSTM) networks
    D) Autoencoder networks

    Answer: B) Recurrent Neural Networks (RNNs)

36. Which algorithm is used for training generative models such as Generative Adversarial Networks (GANs)?
    A) K-Means
    B) Reinforcement Learning
    C) Variational Autoencoders (VAEs)
    D) Expectation-Maximization (EM)

    Answer: C) Variational Autoencoders (VAEs)

37. Which of the following is NOT a step in the training process of a neural network?
    A) Forward propagation
    B) Backward propagation
    C) Weight initialization
    D) Hyperparameter tuning

    Answer: D) Hyperparameter tuning

38. What is the purpose of the loss function in a neural network?
    A) To compute the gradient of the loss function
    B) To normalize the input data
    C) To measure the difference between predicted and actual values
    D) To regularize the weights of the network

    Answer: C) To measure the difference between predicted and actual values

39. Which type of neural network architecture is used for learning dense representations of input data?
    A) Convolutional Neural Networks (CNNs)
    B) Recurrent Neural Networks (RNNs)
    C) Autoencoder networks
    D) Transformer networks

Answer: C) Autoencoder networks

40. Which algorithm is commonly used for detecting anomalies in data?
    A) K-Means
    B) Support Vector Machines (SVM)
    C) Isolation Forest
    D) Expectation-Maximization (EM)

    Answer: C) Isolation Forest

41. Which technique can be used for reducing the dimensionality of text data before feeding it into a neural network?
    A) Principal Component Analysis (PCA)
    B) Term Frequency-Inverse Document Frequency (TF-IDF)
    C) Word Embeddings
    D) Recurrent Neural Networks (RNNs)

    Answer: B) Term Frequency-Inverse Document Frequency (TF-IDF)

42. Which of the following is a drawback of using unsupervised learning algorithms?
    A) They require large amounts of labeled data for training.
    B) They are computationally expensive and require powerful hardware.
    C) They may not produce interpretable results.
    D) They cannot handle missing values in the dataset.

    Answer: C) They may not produce interpretable results.

43. Which algorithm is commonly used for collaborative filtering in recommendation systems?
    A) K-Means
    B) Support Vector Machines (SVM)
    C) Matrix Factorization
    D) Apriori

    Answer: C) Matrix Factorization

44. Which algorithm is used for anomaly detection in time series data?
    A) K-Means
    B) Isolation Forest
    C) DBSCAN
    D) Hierarchical clustering

    Answer: B) Isolation Forest

45. Which algorithm is used for finding the shortest path between nodes in a graph?
    A) Breadth-First Search (BFS)
    B) Depth-First Search (DFS)
    C) Dijkstra's algorithm
    D) A* algorithm

    Answer: C) Dijkstra's algorithm

46. Which of the following is NOT a type of ensemble learning method?
    A) Bagging
    B) Boosting
    C) Stacking
    D) Dimensionality Reduction

    Answer: D) Dimensionality Reduction

47. Which technique is used for finding the most similar documents in a corpus?
   A) Latent Semantic Analysis (LSA)
   B) Term Frequency-Inverse Document Frequency (TF-IDF)
   C) Word Embeddings
   D) All of the above

   Answer: A) Latent Semantic Analysis (LSA)

48. Which technique is used for dimensionality reduction in text data?
   A) Principal Component Analysis (PCA)
   B) Latent Semantic Analysis (LSA)
   C) Word Embeddings
   D) All of the above

   Answer: D) All of the above

49. Which of the following is NOT a clustering algorithm?
   A) K-Means
   B) DBSCAN
   C) Decision Trees
   D) Hierarchical clustering

   Answer: C) Decision Trees

50. Which of the following is NOT a hyperparameter in K-Means clustering?
   A) Number of clusters (K)
   B) Distance metric
   C) Initialization method
   D) Learning rate

   Answer: D) Learning rate

51. Which algorithm is used for finding patterns in transactional datasets?
   A) Apriori
   B) DBSCAN
   C) K-Means
   D) Hierarchical clustering

   Answer: A) Apriori

52. Which of the following is NOT a type of recommendation system?
   A) Content-based filtering
   B) Collaborative filtering
   C) Hybrid filtering
   D) Reinforcement learning

   Answer: D) Reinforcement learning

53. Which of the following is an example of unsupervised learning?
   A) Spam email detection
   B) Image classification
   C) Customer segmentation
   D) Predicting stock prices

   Answer: C) Customer segmentation

54. Which algorithm is used for minimizing errors in the predictions made by a model?

A) Gradient Descent
B) Apriori
C) AdaBoost
D) K-Means

Answer: A) Gradient Descent

55. Which of the following is a hyperparameter in Machine Learning algorithms?
A) Weight
B) Bias
C) Learning rate
D) Activation function

Answer: C) Learning rate

56. Which of the following is an example of a semi-supervised learning algorithm?
A) K-Means
B) Support Vector Machines (SVM)
C) Decision Trees
D) Random Forest

Answer: A) K-Means

57. Which of the following is an example of a kernel function used in Support Vector Machines (SVM)?
A) Sigmoid
B) ReLU
C) Softmax
D) Tanh

Answer: A) Sigmoid

58. Which of the following techniques can be used to address the problem of vanishing gradients in deep learning models?
A) Batch normalization
B) Dropout regularization
C) Weight initialization
D) All of the above

Answer: D) All of the above

59. Which of the following techniques can be used to prevent overfitting in deep learning models?
A) Dropout regularization
B) Batch normalization
C) Weight initialization
D) All of the above

Answer: D) All of the above

60. Which of the following is a common activation function used in hidden layers of neural networks?
A) Linear
B) Sigmoid
C) Softmax
D) ReLU

Answer: D) ReLU

61. Which technique is used for reducing the computational cost of training deep learning models?
    A) Stochastic Gradient Descent (SGD)
    B) Mini-batch Gradient Descent
    C) Transfer learning
    D) All of the above

    Answer: D) All of the above

62. Which technique is used for automatically generating captions for images?
    A) Object detection
    B) Image segmentation
    C) Image classification
    D) Image captioning

    Answer: D) Image captioning

63. Which of the following is NOT a type of recommendation system?
    A) Content-based filtering
    B) Collaborative filtering
    C) Hybrid filtering
    D) Reinforcement learning

    Answer: D) Reinforcement learning

64. Which technique is used for detecting and removing outliers from a dataset?
    A) Interquartile Range (IQR)
    B) Z-score
    C) Box plot
    D) All of the above

    Answer: D) All of the above

65. Which of the following is NOT a distance metric used in clustering algorithms?
    A) Euclidean distance
    B) Manhattan distance
    C) Hamming distance
    D) Chi-squared distance

    Answer: D) Chi-squared distance

66. Which algorithm is commonly used for identifying frequent itemsets in transactional datasets?
    A) K-Means
    B) Apriori
    C) Hierarchical clustering
    D) DBSCAN

    Answer: B) Apriori

67. Which technique is used for finding the optimal number of clusters in a dataset?
    A) Elbow method
    B) Silhouette method
    C) Davies-Bouldin index
    D) All of the above

    Answer: D) All of the above

68. Which of the following is NOT a type of neural network architecture?
   A) Feedforward Neural Networks (FNNs)
   B) Convolutional Neural Networks (CNNs)
   C) Recurrent Neural Networks (RNNs)
   D) Naive Bayes Networks (NBs)

   Answer: D) Naive Bayes Networks (NBs)

69. Which technique is used for generating synthetic samples from an existing dataset?
   A) Data augmentation
   B) SMOTE (Synthetic Minority Over-sampling Technique)
   C) Principal Component Analysis (PCA)
   D) Expectation-Maximization (EM)

   Answer: B) SMOTE (Synthetic Minority Over-sampling Technique)

70. Which algorithm is used for reducing the dimensionality of a dataset while preserving its structure?
   A) K-Means
   B) Principal Component Analysis (PCA)
   C) DBSCAN
   D) Hierarchical clustering

   Answer: B) Principal Component Analysis (PCA)

71. Which technique is used for visualizing high-dimensional data in lower dimensions?
   A) Principal Component Analysis (PCA)
   B) K-Means clustering
   C) Hierarchical clustering
   D) DBSCAN

   Answer: A) Principal Component Analysis (PCA)

72. Which algorithm is used for detecting communities or clusters in a graph?
   A) K-Means
   B) DBSCAN
   C) Louvain algorithm
   D) Hierarchical clustering

   Answer: C) Louvain algorithm

73. Which of the following is NOT a type of graph traversal algorithm?
   A) Breadth-First Search (BFS)
   B) Depth-First Search (DFS)
   C) Dijkstra's algorithm
   D) A* algorithm

   Answer: C) Dijkstra's algorithm

74. Which algorithm is used for finding the shortest path between nodes in a weighted graph?
   A) Breadth-First Search (BFS)
   B) Depth-First Search (DFS)
   C) Dijkstra's algorithm
   D) A* algorithm

   Answer: C) Dijkstra's algorithm

74. Which of the following is NOT a clustering algorithm?
    A) K-Means
    B) DBSCAN
    C) Decision Trees
    D) Hierarchical clustering

    Answer: C) Decision Trees

75. Which algorithm is used for density-based clustering?
    A) K-Means
    B) DBSCAN
    C) Hierarchical clustering
    D) K-Nearest Neighbors (KNN)

    Answer: B) DBSCAN

76. Which technique is used for detecting outliers in a dataset?
    A) Interquartile Range (IQR)
    B) Z-score
    C) Box plot
    D) All of the above

    Answer: D) All of the above

77. Which of the following is NOT a hyperparameter in K-Means clustering?
    A) Number of clusters (K)
    B) Distance metric
    C) Initialization method
    D) Learning rate

    Answer: D) Learning rate

78. Which algorithm is used for finding patterns in transactional datasets?
    A) Apriori
    B) DBSCAN
    C) K-Means
    D) Hierarchical clustering

    Answer: A) Apriori

79. Which technique is used for handling missing values in a dataset?
    A) Deleting rows with missing values
    B) Deleting columns with missing values
    C) Imputation
    D) All of the above

    Answer: D) All of the above

80. Which algorithm is used for finding the most similar items to a given item in a recommendation system?
    A) K-Means
    B) Apriori
    C) Collaborative filtering
    D) K-Nearest Neighbors (KNN)

    Answer: D) K-Nearest Neighbors (KNN)

81. Which of the following is NOT a type of recommendation system?
    A) Content-based filtering

B) Collaborative filtering
C) Hybrid filtering
D) Reinforcement learning

Answer: D) Reinforcement learning

82. Which algorithm is used for dimensionality reduction in text data?
   A) Principal Component Analysis (PCA)
   B) Latent Semantic Analysis (LSA)
   C) Word Embeddings
   D) All of the above

   Answer: D) All of the above

83. Which algorithm is used for density-based clustering?
   A) K-Means
   B) DBSCAN
   C) Hierarchical clustering
   D) K-Nearest Neighbors (KNN)

   Answer: B) DBSCAN

84. Which technique is used for detecting outliers in a dataset?
   A) Interquartile Range (IQR)
   B) Z-score
   C) Box plot
   D) All of the above

   Answer: D) All of the above

85. Which technique is used for handling missing values in a dataset?
   A) Deleting rows with missing values
   B) Deleting columns with missing values
   C) Imputation
   D) All of the above

   Answer: D) All of the above

86. Which algorithm is used for finding the most similar items to a given item in a recommendation system?
   A) K-Means
   B) Apriori
   C) Collaborative filtering
   D) K-Nearest Neighbors (KNN)

   Answer: D) K-Nearest Neighbors (KNN)

87. Which algorithm is used for detecting communities or clusters in a graph?
   A) K-Means
   B) DBSCAN
   C) Louvain algorithm
   D) Hierarchical clustering

   Answer: C) Louvain algorithm

88. Which of the following is NOT a type of graph traversal algorithm?
   A) Breadth-First Search (BFS)
   B) Depth-First Search (DFS)

C) Dijkstra's algorithm
D) A* algorithm

Answer: C) Dijkstra's algorithm

89. Which technique is used for finding the most similar documents in a corpus?
   A) Latent Semantic Analysis (LSA)
   B) Term Frequency-Inverse Document Frequency (TF-IDF)
   C) Word Embeddings
   D) All of the above

   Answer: A) Latent Semantic Analysis (LSA)

90. Which algorithm is used for finding the shortest path between nodes in a weighted graph?
   A) Breadth-First Search (BFS)
   B) Depth-First Search (DFS)
   C) Dijkstra's algorithm
   D) A* algorithm

   Answer: C) Dijkstra's algorithm

91. Which technique is used for dimensionality reduction in text data?
   A) Principal Component Analysis (PCA)
   B) Latent Semantic Analysis (LSA)
   C) Word Embeddings
   D) All of the above

   Answer: D) All of the above

92. Which of the following is NOT a clustering algorithm?
   A) K-Means
   B) DBSCAN
   C) Decision Trees
   D) Hierarchical clustering

   Answer: C) Decision Trees

96. Which algorithm is used for density-based clustering?
   A) K-Means
   B) DBSCAN
   C) Hierarchical clustering
   D) K-Nearest Neighbors (KNN)

   Answer: B) DBSCAN

97. Which technique is used for detecting outliers in a dataset?
   A) Interquartile Range (IQR)
   B) Z-score
   C) Box plot
   D) All of the above

   Answer: D) All of the above

98. Which of the following is NOT a hyperparameter in K-Means clustering?
   A) Number of clusters (K)
   B) Distance metric
   C) Initialization method
   D) Learning rate

Answer: D) Learning rate

99. Which algorithm is used for finding patterns in transactional datasets?
    A) Apriori
    B) DBSCAN
    C) K-Means
    D) Hierarchical clustering

    Answer: A) Apriori

100. Which technique is used for handling missing values in a dataset?
    A) Deleting rows with missing values
    B) Deleting columns with missing values
    C) Imputation
    D) All of the above

    Answer: D) All of the above

-X-X-X-