**Section – A ( 20 questions)**

**Q1. What is StandardScaler used for ?**

a) to normalize the input datab) to standardize the input data

c) to rationalize the input datad) for changing the gaussian of whole dataset

**Q2. What is joint maximization in machine learning?**

a) method to maximize the two parameters at the same time

b) minimize the two parameters at the same time

c) maximize the weights and bias at the same time d) minimize the labels and features

**Q3. What is the simplest machine learning model?**

a) Support vector machine b) Decision Tree Classifier

c) Logistic Regression d) Linear Regression

**Q4. What is Linear Regression in Machine learning?**

a) prescription of continuous equal interval parameters b) prediction of continous target variable

c) predicton of discrete target variable d) prediction of categorical target variable

**Q5. What is the formula of bayes theoram?**

a) prior = posterior \*likelihood / marginalized b) posterior = likelihood \* prior

c) posterior = likelihood \* prior / marginalized d) Marginalized = likelihood \* prior

**Q6. Which function is used in logistic regression**

a) softmax b) ReLu c) tan(x) d) sigmoid

**Q7. Which accuracy of a machine learning model is best among the following for classification?**

a) 50% b) 55% c) 60% d) 20%

**Q8. How can you make Logistic Regression a classification model since it is a regression model?**

a) output is continous but become discrete by applying some threshold on it

b) output is based on continous function of feautures in a model

c) output does not convert into dicrete ,always remains the continous

d) output of non- continous target variable in 3-d plane

**Q9. Which loss function is used in Logistic regression?**

a) Mean Squared error b) Binary Cross Entropy Loss

c) Hinge Loss d) Root mean squared error

**Q10. Which loss is used in Support Vector Machine?**

a) Mean Squared error b) Binary Cross Entropy Loss

c) Hinge Loss d) cross Entropy loss

**Q11. Which of the following is a pure statistical machine learning model?**

CNNLogistic regressionPCAK-means clustering

**Q12. What is ratio of testing to training dataset for splitting ?**

80:20 70:30 20:30 30:70

**Q13. Which machine learning model has no learnable parameters?**

Linear regression KNN K-mean clustering SVM

**Q14. Why the value k in KNN has to be odd?**

Order to ensure that there is no tie in voting It is difficult to predict the class to which it belongs to

To underfit the model To overfit the model

**Q15. Which machine learning model among them is purely probabilistic model?**

K-means clustering KNN ANN PCA

**Q16. How can we reduce the dimension in PCA?**

By the use of mathematical operatorsBy the use of Eigen Vectors

By SVM classifierBy KNN algorithm use

**Q17. In machine learning, what is the purpose of the validation set?**

a) To train the model parameters on labeled data.

b) To fine-tune the hyperparameters of the model.

c) To test the generalization ability of the model.

d) To evaluate the performance of the model during training.

**Q18. Which of the following is NOT a common preprocessing step in preparing data for machine learning models?**

a) Feature scaling b) Dimensionality reduction

c) Normalization d) Random initialization

**Q19. What is the primary goal of using regularization techniques in machine learning?**

a) To increase the model complexity b) To reduce the risk of overfitting

c) To decrease the computational cost d) To improve model interpretability

**Q20. Which of the following is a characteristic of unsupervised learning?**

a) The model learns from labeled data.

b) The model makes predictions without any input data.

c) The model is trained using input-output pairs.

d) The model clusters data points based on similarities.

**Section – B (2 marks Each)**

**Q21. What is the primary purpose of using the activation function in neural networks?**

a) To transform the input features into a higher-dimensional space.

b) To calculate the cost function. c) To introduce non-linearity into the model.

d) To regularize the model parameters.

**Q22. Which of the following evaluation metrics is used to assess the performance of a regression model?**

a) Accuracy b) Precision c) R-squared d) F1-score

**Q23. What does the term "one-hot encoding" refer to in machine learning?**

a) Encoding categorical variables as integers.

b) Encoding categorical variables as binary vectors.

c) Encoding ordinal variables as continuous values.

d) Encoding continuous variables as categorical values.

**Q24. What is the primary purpose of using cross-validation in machine learning?**

a) To maximize the training accuracy of the model.

b) To minimize the computational cost of model training.

c) To assess the generalization performance of the model.

d) To optimize the hyperparameters of the model.

**Q25. Which of the following algorithms is NOT a type of supervised learning?**

a) Decision Trees b) K-Means Clustering

c) Support Vector Machines d) Linear Regression

**Q26. What is the primary objective of dimensionality reduction techniques in machine learning?**

a) To increase the computational cost of model training.

b) To decrease the interpretability of the model.

c) To remove irrelevant features and reduce the complexity of the dataset.

d) To introduce non-linearity into the model.

**Q27. Which of the following metrics is used to evaluate the performance of a binary classification model when both false positives and false negatives are equally important?**

a) Precision b) Recall c) F1-score d) Accuracy

**Q28. What is the primary purpose of using logistic regression in machine learning?**

a) To predict continuous outcomes. b) To perform dimensionality reduction.

c) To classify data into discrete categories. d) To handle missing values in the dataset.

**Q29. What is the primary goal of using feature scaling in machine learning?**

a) To increase the dimensionality of the dataset.

b) To reduce the computational cost of model training.

c) To ensure that all features contribute equally to the learning process.

d) To introduce non-linearity into the model.

**Q30. Which of the following is a characteristic of supervised learning?**

a) The model learns from unlabeled data. b) The model makes predictions without any input data.

c) The model is trained using input-output pairs.

d) The model clusters data points based on similarities.

**Q31. What does the term "bias" refer to in the context of machine learning models?**

a) The difference between predicted and actual values.

b) The degree of flexibility of the model.

c) The error introduced by approximating a real problem.

d) The sensitivity of the model to changes in the input data.

**Q32. Which of the following is NOT a classification evaluation metric?**

a) R-squared b) Confusion Matrix c) Precision d) Accuracy

**Q33. What is the primary purpose of using the F1-score as an evaluation metric?**

a) To measure the overall performance of a classifier.

b) To balance precision and recall in binary classification.

c) To assess the goodness-of-fit in regression models.

d) To determine the optimal number of clusters in clustering algorithms.

**Q34. In machine learning, what does the term "regularization" aim to prevent?**

a) Underfitting b) Overfitting c) Model convergence d) Model interpretability

**Q35. Which of the following algorithms is primarily used for regression tasks?**

a) K-Means Clustering b) Decision Trees

c) Linear Regression d) Support Vector Machines