SREENIVASA INSTITUTE OF TECHNOGY AND MANAGEMENT STUDIES

(AUTONOMOUS)

Murukambattu, Chittoor

MCA DEPARTMENT

QUESTION BANK

For

24MCA211 - MACHINE LEARNING

Regulation – 24

Academic Year 2026 – 27

Prepared by

Mr. M Rajesh, Assistant Professor

ISREENIVASA INSTITUTE OF TECHNOLOGY AND MANAGEMENT STUDIES.
(AUTONOMOUS)
MCA DEPARTMENT
QUESTION BANK

SUBJECT NAME : MACHINE LEARNING **SUBJECT CODE : 24MCA211**

YEAR &SEM : II & I **Academic Year : 2026-27**

UNIT I - INTRODUCTION AND DATA PREPROCESSING		
Basic concept of Machine Learning- Types of Machine Learning-– Supervised Learning –Unsupervised Learning- Reinforcement learning Applications and challenges-A Brief Review of Probability Theory- Baye's Theorem. Data Preprocessing- Data Cleaning, Data Integration, Data Transformation, Data Reduction or Dimensionality Reduction - Principal component Analysis.		
PART –A		
Q.No.	**Questions**	**Blooms Taxonomy Level**
1.	Define Machine Learning	L1
2.	List the types of Machine Learning	L1
3.	State any two applications of Machine Learning.	L1
4.	Mention two challenges of Machine Learning.	L1
5.	Define Probability.	L1
6.	State Bayes' Theorem.	L1
7.	What is Data Preprocessing?	L1
8.	Define Data Cleaning.	L1
9.	What is Data Integration?	L1
10.	Define Principal Component Analysis (PCA).	L2
PART –B		
1.	Explain the basic concepts and workflow of Machine Learning with suitable examples.	L2
2.	Compare Supervised, Unsupervised, and Reinforcement Learning with examples and applications	L3
3.	Discuss the applications and challenges of Machine Learning	L4
4.	Explain the basics of Probability Theory and Bayes' Theorem with a suitable example.	L2
5.	Describe the steps involved in Data Preprocessing.	L2
6.	Explain Data Cleaning, Data Integration, Data Transformation and Data Reduction	L2
7.	Explain Principal Component Analysis (PCA) and its advantages in dimensionality reduction.	L3
8.	Analyze the importance of Data Preprocessing in Machine Learning	L4
UNIT II - MODEL EVALUATION AND FEATURE ENGINEERING:		
Selecting a Model, training a Model, Model Representation and Interpretability, Evaluating Performance of a Model, Improving performance of a Model. Basics of Feature Engineering-Feature Transformation, Feature Subset Selection.		
PART –A		
1.	What is model selection?	L1
2.	Define model training.	L1
3.	What is model interpretability?	L2
4.	Define model evaluation.	L1
5.	What is overfitting?	L2
6.	What is underfitting?	L2
7.	Define Feature Engineering.	L1
8.	What is Feature Transformation?	L2
9.	Define Feature Subset Selection.	L1

10.	Mention two techniques to improve model performance.	L2
	PART –B	
1.	Explain the process of selecting and training a Machine Learning model.	L2
2.	Describe Model Representation and Model Interpretability with examples.	L2
3.	Explain different methods used to evaluate the performance of a Machine Learning model.	L2
4.	Discuss techniques used to improve the performance of a Machine Learning model.	L3
5.	Explain the concept and importance of Feature Engineering.	L2
6.	Describe Feature Transformation techniques with suitable examples.	L3
7.	Explain Feature Subset Selection methods and their advantages.	L3
8.	Analyze the impact of Feature Engineering on model performance.	L4
	UNIT III - REGRESSION	
Introduction to Supervised Learning and Regression, Linear Regression, Evaluation of Model Estimators, Regularization-Ridge regression, LASSO regression, Multi Linear Regression, Gradient Based Methods. Cost function- Minimizing the Cost Function for a Single Variable Function, Minimizing the Cost Function for a Two-Variable Function, Evaluation Metrics (Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Root Mean Squared Log Error (RMSLE), R Squared (R2), Adjusted R Squared).		
	PART – A	
1.	Define Supervised Learning.	L1
2.	What is Regression in Machine Learning?	L1
3.	Define Linear Regression.	L1
4.	State the equation of Simple Linear Regression.	L1
5.	What is a Model Estimator?	L2
6.	Define Regularization.	L1
7.	What is Ridge Regression?	L1
8.	What is LASSO Regression?	L1
9.	Define Multiple Linear Regression.	L1
10.	What is Gradient Descent?	L2
11.	Define Cost Function.	L2
12.	What is meant by minimizing the Cost Function?	L2
13.	Define Mean Absolute Error (MAE).	L1
14.	Define Mean Squared Error (MSE).	L1
15	What is Root Mean Squared Error (RMSE)?	L1
16	Define Root Mean Squared Log Error (RMSLE).	L1
17	What is R-Squared (R^2)?	L2
18	Define Adjusted R-Squared.	L2
19	List any two evaluation metrics used for regression models.	L1
20	Mention two applications of Linear Regression.	L2
	PART –B	
1.	Explain the concept of Supervised Learning and Regression with suitable examples.	L2
2.	Explain Linear Regression with its mathematical model, assumptions, advantages, and limitations.	L2
3.	Describe the evaluation of model estimators and explain the characteristics of a good estimator.	L2

4.	Explain Regularization and compare Ridge Regression and LASSO Regression with suitable examples.	L4
5.	Explain Multiple Linear Regression with mathematical formulation and real-world applications.	L3
6.	Explain Gradient-Based Methods and describe the Gradient Descent algorithm with suitable examples.	L3
7.	Explain the Cost Function in Linear Regression and describe the process of minimizing the Cost Function for a single-variable function.	L3
8.	Explain the minimization of the Cost Function for a two-variable function using Gradient Descent.	L3
9.	Explain the regression evaluation metrics MAE, MSE, RMSE, RMSLE, R^2, and Adjusted R^2 with suitable examples.	L2
10.	Compare MAE, MSE, RMSE, RMSLE, R^2, and Adjusted R^2 and discuss their significance in evaluating regression models.	L4
11.	Analyze the advantages and limitations of Ridge Regression and LASSO Regression in handling overfitting.	L4
12.	Evaluate the performance of different regression models using appropriate evaluation metrics.	L5
13.	Compare Simple Linear Regression and Multiple Linear Regression with suitable examples.	L4
14.	Discuss the importance of Cost Functions and Gradient-Based Optimization techniques in Machine Learning.	L3
15	Evaluate the effectiveness of different regression evaluation metrics for various real-world applications.	L5
UNIT IV: CLASSIFICATION		
Introduction to Classification, Logistic Regression-Building Logistic Regression Model (Logit Function), Maximum Likelihood Estimation. Decision Tree-Steps to Construct a Decision Tree, Classification Using Decision Trees, Issues in Decision Trees, Ensemble Learning-Random Forest. Bayesian Classification-Naive Bayes Classifier, k-Nearest Neighbor (KNN). Neuron-Neural Networks-The Perceptron Multilayer Perceptron(MLP), Support Vector Machines- Linear Support Vector Machines, Optimal Hyperplane, Radial Basis Functions, Evaluation Metrics (Accuracy, Confusion Matrix, Precision, Recall, F1 Score)		
PART –A		
1.	Define Classification in Machine Learning.	L1
2.	What is Logistic Regression?	L1
3.	Define the Logit Function.	L1
4.	What is Maximum Likelihood Estimation (MLE)?	L2
5.	Define Decision Tree.	L1
6.	List the steps involved in constructing a Decision Tree.	L2
7.	What is Information Gain?	L1
8.	Define Entropy in Decision Trees.	L1
9.	Mention any two issues in Decision Trees.	L2
10.	What is Ensemble Learning?	L1
11.	Define Random Forest.	L1
12.	What is Bayesian Classification?	L1
13.	Define Naïve Bayes Classifier.	L1
14.	What is the assumption of the Naïve Bayes algorithm?	L2
15.	Define K-Nearest Neighbor (KNN).	L1
16.	What is the role of the value of K in the KNN algorithm?	L2
17.	Define a Neuron in an Artificial Neural Network.	L1
18.	What is a Perceptron?	L1

SUBJECT NAME : MACHINE LEARNING **SUBJECT CODE : 24MCA211**

YEAR &SEM : II & I **Academic Year : 2026-27**

19.	Define Multilayer Perceptron (MLP).	L1
20.	What is a Support Vector Machine (SVM)?	L1
21.	Define Optimal Hyperplane.	L2
22.	What is a Radial Basis Function (RBF)?	L1
23.	Define Accuracy.	L1
24.	What is a Confusion Matrix?	L2
25.	Define Precision.	L1
26.	Define Recall.	L1
27.	What is the F1 Score?	L2
28.	List any two applications of Classification.	L2
29.	Mention any two advantages of Random Forest.	L2
30.	State one difference between Logistic Regression and Linear Regression.	L2
	PART –B	
1.	Explain the concept of Classification in Machine Learning with suitable examples.	L2
2.	Explain Logistic Regression and describe the steps involved in building a Logistic Regression model using the Logit Function.	L3
3.	Explain the concept of Maximum Likelihood Estimation (MLE) and its role in Logistic Regression.	L3
4.	Explain the steps involved in constructing a Decision Tree with a suitable example.	L3
5.	Explain the process of Classification using Decision Trees and discuss the issues associated with Decision Trees.	L3
6.	Discuss Ensemble Learning and explain the Random Forest algorithm with its advantages and limitations.	L2
7.	Explain Bayesian Classification and describe the Naïve Bayes Classifier with suitable examples.	L2
8.	Explain the K-Nearest Neighbor (KNN) algorithm with suitable examples and applications.	L3
9.	Explain the architecture and working of Artificial Neural Networks, including the Perceptron and Multilayer Perceptron (MLP).	L2
10.	Explain Support Vector Machines (SVM), Linear SVM, Optimal Hyperplane, and Radial Basis Function (RBF) with suitable diagrams.	L3
11.	Compare Logistic Regression, Decision Tree, Naïve Bayes, KNN, and SVM based on their working principles, advantages, and limitations. L4 – Analyze	L4
12.	Explain the evaluation metrics Accuracy, Confusion Matrix, Precision, Recall, and F1 Score with suitable examples.	L2
13.	Compare Precision, Recall, and F1 Score, and discuss their importance in evaluating classification models.	L4
14.	Analyze the advantages and limitations of Decision Trees and Random Forest in classification problems.	L4
15.	Evaluate the performance of different classification algorithms using appropriate evaluation metrics.	L4
	UNIT V: CLUSTERING	
Introduction to Unsupervised Learning Algorithms, Clustering- Types of Clustering, Partitioning Methods of Clustering, Hierarchical Methods, Density-based Methods and Grid-Based Methods- K- means Clustering-Choosing number of Clusters		
	PART –A	
1.	Define Unsupervised Learning.	L1
2.	What is Clustering?	L1
3.	List the different types of clustering.	L1

SUBJECT NAME : MACHINE LEARNING **SUBJECT CODE : 24MCA211**

YEAR &SEM : II & I **Academic Year : 2026-27**

4.	Define Partitioning Clustering.	L1
5.	What is Hierarchical Clustering?	L1
6.	Define Density-Based Clustering.	L1
7.	What is DBSCAN?	L2
8.	Define Grid-Based Clustering.	L1
9.	What is K-Means Clustering?	L1
10.	What is the objective of the K-Means algorithm?	L2
11.	What is the Elbow Method?	L2
12.	Mention any two applications of Clustering.	L2
	PART –B	
1.	Explain the concept of Unsupervised Learning and discuss the importance of clustering in Machine Learning.	L2
2.	Explain the different types of clustering with suitable examples.	L2
3.	Describe the Partitioning Method of Clustering and explain the K-Means Clustering algorithm with a suitable example.	L3
4.	Explain Hierarchical Clustering and compare Agglomerative and Divisive clustering methods.	L4
5.	Explain Density-Based Clustering and discuss the DBSCAN algorithm with its advantages and limitations.	L3
6.	Describe Grid-Based Clustering methods and explain their advantages and applications.	L2
7.	Explain different methods used for choosing the optimal number of clusters, including the Elbow Method and Silhouette Score.	L3
8.	Compare Partitioning, Hierarchical, Density-Based, and Grid-Based clustering methods based on their characteristics, advantages, and limitations.	L4
9.	Analyze the advantages, limitations, and real-world applications of K-Means Clustering.	L4
10.	Evaluate the performance of clustering algorithms and justify the selection of an appropriate clustering method for a given dataset.	L5

L1: Remember, L2: Understand, L3: Apply, L4: Analyze, L5: Evaluate, L6: Create