Course Code: 23CS0523

R23

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY: PUTTUR (AUTONOMOUS)

Siddharth Nagar, Narayanavanam Road – 517583

QUESTION BANK (DESCRIPTIVE)

Subject with Code: Machine Learning (23CS0523) **Regulation:** R23

Course & Branch: B.Tech – CSE **Year & Sem:** III Year & II Sem

UNIT – I

INTRODUCTION TO MACHINE LEARNING

1.	**a)**	Define Machine Learning What are the advantages of Machine Learning.	**[CO1][L1]**	**5M**
	b)	How do you select Machine Learning model.	**[CO1][L3]**	**5M**
2.		Classify various types of Machine Learning & explain any ML model with an example.	**[CO1][L2]**	**10M**
3.		Explain the applications of Machine Learning.	**[CO1][L2]**	**10M**
4.	**a)**	Explain concepts of learning by Rote & Induction with an example.	**[CO1][L2]**	**5M**
	b)	Briefly discuss various Dataset‟s in Machine Learning	**[CO1][L2]**	**5M**
5.	**a)**	Briefly explain the concepts of Reinforcement Algorithm	**[CO1][L2]**	**5M**
	b)	Discuss about various types of Data	**[CO1][L1]**	**5M**
6.		Explain different stages in Machine Learning	**[CO1][L2]**	**10M**
7.	**a)**	Briefly describe the procedure to handle duplicate records	**[CO1][L3]**	**5M**
	b)	Explain about feature engineering in Machine Learning	**[CO1][L3]**	**5M**
8.		What is the role of searching and learning in Machine Learning Explain Grid search & Random search with an example	**[CO1][L4]**	**10M**
9.		Explain different Data collection Methods	**[CO1][L2]**	**10M**
10.	**a)**	What is the need of label encoding	**[CO1][L1]**	**2M**
	b)	What is meant by supervised and Unsupervised learning in ML	**[CO1][L1]**	**2M**
	c)	Define machine learning.	**[CO1][L1]**	**2M**
	d)	What is meant by Label Dataset in ML	**[CO1][L1]**	**2M**
	e)	What is meant by regression problem	**[CO1][L1]**	**2M**

R23

Course Code: 23CS0523

UNIT – II

NEAREST NEIGHBOR-BASED MODELS

1.	**a)**	Explain the steps involved in Nearest Neighbour Models.	**[CO2][L2]**	**5M**
	b)	List out the applications of Manhattan Distance in ML.	**[CO2][L1]**	**5M**
2.		Explain Euclidian Distance measure with one example.	**[CO2][L2]**	**10M**
3.		Explain Manhattan Distance measure with example.	**[CO2][L2]**	**10M**
4.	**a)**	Explain the concept of Murkowski distance.	**[CO2][L2]**	**5M**
	b)	Explain hamming distance for any two binary and DNA patterns.	**[CO2][L2]**	**5M**
5.		Explain Non Metric Proximity measures with one example.	**[CO2][L2]**	**10M**
6.		Explain the steps involved in K-Nearest Neighbors algorithm	**[CO2][L2]**	**10M**
7.		Classify a new customer as "high-spending" or "low-spending" based on their age and income by using KNN algorithm.	**[CO2][L3]**	**10M**

Customer	Age	Income (Rupees)	Spending Class
A	25	30,000	Low
B	45	80,000	High
C	35	50,000	Low
D	50	90,000	High
E (New)	40	60,000	

8. A bank wants to classify customers as "Low-Risk" or "High-Risk" for loans. Classify Customer E using Manhattan Distance and Radius Distance Nearest Neighbour Algorithm. Assume suitable radius. **[CO2][L3]** **10M**

Customer	Age	Income ($)	Credit Score	Loan Amount (Rupees)	Risk Level
A	25	30,000	700	10,000	Low
B	45	80,000	600	40,000	High
C	35	50,000	750	20,000	Low
D	55	90,000	580	50,000	High
E (New)	40	60,000	680	25,000	

9.		Explain KNN Regression with one example. Also list out advantages	**[CO2][L2]**	**10M**
10.	**a)**	List out the performance measures of Regression.	**[CO2][L1]**	**2M**
	b)	List out the performance measures of Classification.	**[CO2][L1]**	**2M**
	c)	Define MAE and R^2.	**[CO2][L1]**	**2M**
	d)	How does KNN regression differ from KNN classification	**[CO2][L2]**	**2M**
	e)	List out common distance measures used in machine learning.	**[CO2][L1]**	**2M**

R23

Course Code: 23CS0523

UNIT – III

MODELS BASED ON DECISION TREES & THE BAYES CLASSIFIER

1.		Explain the key steps involved in building a Decision Tree for classification.	**[CO4][L2]**	**10M**
2.		Discuss how impurity is measured in decision trees using Gini Index and Entropy with examples.	**[CO3][L4]**	**10M**
3.		Describe the steps involved in Decision Tree is built for regression with one example.	**[CO3][L3]**	**10M**
4.		Describe the steps involved in training a Random Forest for classification.	**[CO4][L2]**	**10M**
5.		Apply Random Forest Regression to a sample dataset and explain how it predicts continuous target values. Mention any evaluation metrics used.	**[CO3][L3]**	**10M**
6.		Explain the key concepts of Bayes‟ Rule and how it is used in classification. Provide an example.	**[CO4][L2]**	**10M**
7.		Describe the steps involved in building a Bayes Classifier for binary classification using Bayes‟ Rule.	**[CO4][L3]**	**10M**
8.		What is the Naive Bayes Classifier Explain the assumption of class conditional independence and how it simplifies computation.	**[CO4][L2]**	**10M**
9.		Apply the Naive Bayes Classifier to a real-world multi-class problem such as spam detection. Outline the model building, prediction, and evaluation steps.	**[CO4][L3]**	**10M**
10.	**a)**	What is the main criterion used to split nodes in a decision tree during classification	**[CO4][L1]**	**2M**
	b)	State one key difference between decision trees used for classification and for regression.	**[CO3][L2]**	**2M**
	c)	What is the role of the bias–variance trade-off in decision trees	**[CO3][L4]**	**2M**
	d)	In the context of Bayes‟ Rule, what is meant by the 'prior probability'	**[CO3][L1]**	**2M**
	e)	What assumption does the Naive Bayes Classifier make about features	**[CO4][L1]**	**2M**

R23

Course Code: 23CS0523

UNIT – IV

LINEAR DISCRIMINANTS FOR MACHINE LEARNING

1.		Explain the steps involved in training a Perceptron classifier with a simple example.	**[CO5][L3]**	**10M**
2.		What is Linear Discriminant Analysis (LDA) Explain LDA steps for classification.	**[CO5][L2]**	**10M**
3.		Explain the Perceptron Learning Algorithm with steps. How does the algorithm converge for linearly separable data	**[CO5][L3]**	**10M**
4.		Describe the working of Support Vector Machines (SVMs) for linearly separable data. Explain margin, hyper plane, and support vectors.	**[CO5][L2]**	**10M**
5.		What happens when data is not linearly separable in SVM Explain how soft margin and Kernel Trick help.	**[CO5][L4]**	**10M**
6.		Explain the concept of the Kernel Trick in SVM with an example. Why is it useful for non-linear classification	**[CO5][L3]**	**10M**
7.		Compare and contrast Logistic Regression and Linear Regression. Give one example each.	**[CO5][L4]**	**10M**
8.		What is a Multi-Layer Perceptron (MLP) Describe its architecture and how it extends the single-layer perceptron.	**[CO5][L2]**	**10M**
9.		Explain the Backpropagation algorithm for training an MLP. Include forward pass, error calculation, and weight update steps.	**[CO5][L3]**	**10M**
10.	**a)**	What is the main objective of the Perceptron Learning Algorithm	**[CO5][L1]**	**2M**
	b)	Define the term „Support Vector“ in Support Vector Machines (SVM).	**[CO5][L1]**	**2M**
	c)	What is the role of the activation function in a Multi-Layer Perceptron	**[CO5][L2]**	**2M**
	d)	Mention one key difference between Linear Regression and Logistic Regression.	**[CO5][L1]**	**2M**
	e)	What does the Kernel Trick do in the context of SVM	**[CO5][L2]**	**2M**

R23

Course Code: 23CS0523

UNIT – V

CLUSTERING

1.		What is clustering Explain the types of clustering methods with simple examples.	**[CO6][L2]**	**10M**
2.		Describe the K-Means clustering algorithm. What are the steps involved in the algorithm Provide an example.	**[CO6][L3]**	**10M**
3.		What is Fuzzy C-Means clustering How does it differ from K-Means clustering	**[CO6][L3]**	**10M**
4.		Explain Agglomerative and Divisive hierarchical clustering methods. How are they different	**[CO6][L3]**	**10M**
5.		What is Soft Clustering How is it used in clustering methods like Fuzzy C-Means	**[CO6][L4]**	**10M**
6.		What is Matrix Factorization in clustering How does it help in clustering large datasets	**[CO6][L3]**	**10M**
7.		Explain Rough Clustering and the Rough K-Means algorithm. How is it different from traditional K-Means	**[CO6][L4]**	**10M**
8.		What is the Expectation Maximization (EM) algorithm How does it work for clustering data	**[CO6][L3]**	**10M**
9.		Explain Spectral Clustering. How does it work, and when is it useful for clustering	**[CO6][L4]**	**10M**
10.	**a)**	What is the difference between Agglomerative and Divisive clustering methods	**[CO6][L1]**	**2M**
	b)	What is a centroid in K-Means clustering	**[CO6][L1]**	**2M**
	c)	What is Soft Clustering Give an example.	**[CO6][L2]**	**2M**
	d)	How does Fuzzy C-Means differ from K-Means	**[CO6][L2]**	**2M**
	e)	What is Matrix Factorization in clustering Why is it useful	**[CO6][L2]**	**2M**