UNIT-1:

Sr.No	Question	Marks
1	Explain clustering and association in brief.	3/4
2	Give the difference between supervised learning and unsupervised learning.	3/4
3	Explain the concept of penalty and rewards in reinforcement learning.	3/4
4	Explain human learning in brief.	3/4
5	Define Machine Learning. Also, compare it with Human Learning.	3/4
6	Explain classification and regression in brief.	3/4
7	Write a short note on machine learning in finance and banking	3/4
8	Identify main three features for following well-posed problem: 1. Fruit prediction Problem 2. Handwriting recognition Problem	3/4
9	Give the difference between Python and R.	3/4
1	Compare the different types of machine learning.	7
2	Define well-posed problem. Explain important features that are required to well-define a learning problem.	7
3	Explain different tools and technology used in Machine Learning.	7
4	Describe basic concept of Machine Learning and its application.	7

1	Compare the different types of machine learning.	7
2	Define well-posed problem. Explain important features that are required to well-define a learning problem.	7
3	Explain different tools and technology used in Machine Learning.	7
4	Describe basic concept of Machine Learning and its application.	7
5	Define machine learning. Explain any two business applications of Machine Learning in detail.	7
6	Explain types of machine learning in detail.	7

UNIT-2:

Sr.No	Question	Marks
1	What is the main purpose of matplotlib in Machine Learning ?	3/4
2	Compare and contrast Pandas and Numpy	3/4
3	Give the purpose of Numpy in Machine Learning.	3/4
4	What are the advantages of using Scikit-learn?	3/4
5	How to load dataset using Numpy? Explain.	3/4
6	How to load dataset using Panda ?	3/4
7	How to plot a vertical line and a horizontal line using Matplotlib	3/4
1	Explain features and applications of Matplotlib.	7
2	Explain features and applications of Numpy.	7
3	Explain features and applications of Pandas.	7
4	Write a Python program to load the iris data from a given csv file into a dataframe and print the shape of the data, type of the data and first 3 rows using Scikit-Learn.	7
5	Explain features and applications of Scikit-Learn.	7

1	Explain features and applications of Matplotlib.	7
2	Explain features and applications of Numpy.	7
3	Explain features and applications of Pandas.	7
4	Write a Python program to load the iris data from a given csv file into a dataframe and print the shape of the data, type of the data and first 3 rows using Scikit-Learn.	7
5	Explain features and applications of Scikit-Learn.	7
6	While predicting malignancy of tumour of a set of patients using a classification model, following are the data recorded: 1. Correct predictions 20 malignant, 70 benign 2. Incorrect predictions - 4 malignant, 6 benign Create confusion matrix for the same. And, calculate the accuracy, error rate, sensitivity, specificity, precision, recall and F-measure of the model.	7

UNIT-3:

Sr.No	Question	Marks
1	State any four real-world problems solved by predictive models. Explain any one brief	3/4
2	Give the difference between predictive model and descriptive model.	3/4
3	Define model. How can you train a model ?	3/4
4	State any four real-world problems solved by descriptive models. Explain any one i brief.	3/4
5	Give the difference between Bagging and Boosting	3/4
6	Draw a detailed diagram to show the approach of 10-fold cross validation.	3/4
7	Write down steps to use holdout method for model training	3/4
8	Write a short note on bias-variance trade-off in context of model fitting.	3/4
9	Define overfitting. When does it happen? Define underfitting. When does it happen?	3/4
10	Write a brief note on stacking.	3/4
11	State various ways to improve performance of a model	3/4

1	Explain different types of model.	7
2	While predicting malignancy of tumour of a set of patients using a classification model, following are the data recorded: 1. Correct predictions - 15 malignant, 75 benign 2. Incorrect predictions - 4 malignant, 6 benign Create confusion matrix. Calculate the sensitivity, specificity, precision, Recall and F-measure of the model.	7
3	Discuss the kNN model in detail.	7
4	Describe k-fold cross validation in detail.	7
5	Explain bagging, boosting and stacking in detail. Describe Ensemble learning approach in detail.	7
6	Write short note on support vector machine .	7

Sr.No	Question	Marks
1	Define following terms: Support Vectors, Hyperplane, Margin	3/4
2	Give any three examples of supervised learning in the field of healthcare.	3/4
3	Explain classification model in brief.	3/4
4	Give any three examples of supervised learning in Industry 4.0	3/4
5	Give the difference between classification and regression.	3/4
6	Explain different types of logistic regression.	3/4
7	Draw the flowchart which shows the classification learning process.	3/4
8	List applications of SVM algorithm.	3/4
9	State advantages and disadvantages of k-NN algorithm.	3/4
10	Compare and contrast Single linear regression and multiple linear regression.	3/4

1	Write and discuss k-NN Algorithm.	7
2	Discuss the SVM model in detail with its pros and cons.	7
3	Write a short note on Single Linear Regression. Also, state applications of it.	7
4	Write a short note on Multiple Linear Regression. Also, state applications of it.	7
5	Define Classification. Explain classification learning steps in detail.	7
6	Explain logistic regression with advantage and disadvantage.	7
7	Explain any three applications of classification in detail	7

UNIT-4:

Sr.No	Question	Marks
1	Differentiate clustering with classification.	3/4
2	Write pseudo code of k-means clustering algorithm.	3/4
3	State any four applications of unsupervised learning.	3/4
4	Give the difference between supervised learning and unsupervised learning.	3/4
5	Write strength and weakness of k-means clustering algorithm:	3/4
6	How unsupervised learning is useful in fraud detection?	3/4
7	Define: Support, Confidence	3/4
8	State apriori property.	3/4
9	Write strength and weakness of apriori clustering algorithm	3/4
10	Explain any two applications of apriori algorithm.	3/4

1	Write and explain k-means clustering approach in detail.	7
2	Write and explain apriori algorithm in detail.	7
3	Write and explain applications of unsupervised learning.	7
4	Generate frequent itemsets and generate association rules based on it using apriori algorithm. Minimum support is 50% and minimum confidence is 70%	7

TID	ITEM
T1	1,5,4
T2	2,3
T3	1,3,5
T4	1,4,5

UNIT-5:

Sr.No	Question	Marks
1	Give the difference between histogram and box plot with example.	3/4
2	Give the difference between qualitative data and quantitative data.	3/4
3	State the activities involved in model preparation stage.	3/4
4	State different measures of central tendency	3/4
5	Explain PCA in brief.	3/4
6	What are the factors which lead to the data quality issues?	3/4
7	State various strategies to handle missing values.	3/4
8	What is IQR? How it is meseaured?	3/4
9	Find mean, median, mode and standard deviation for the following data: 1,2,3,4,5,5,4,3,2,,5,8,8,7,9,6,8,7,6,2	3/4
10	Define outliers. How can we take care of outliers in data?	3/4

12	Define following terms: Data pre-processing, Data remediation, Outliers, Imputation, Standard Deviation, Ratio Data, Ordinal Data	7
3	Explain steps to create box plot with suitable example.	7
4	Explain steps to create histogram with suitable example. Also, state the difference between bar-chart and histogram	7
5	Write a short note on dimensionality reduction.	7
6	Write a short note on feature subset selection.	7
7	Describe machine learning activities in detail.	7
8	Define data pre-processing. Explain various methods used in data pre-processing.	7
9	Explain data types in machine learning with example.	7