Department of Information Technology
Academic Year 2025-26(SEM I)

Class: TE **Subject: LP I -ML Code: 314448**
Teaching Scheme: 2 Hours/Week **Credits: 02**
Examination Scheme: Term Work: 25 Marks **Practical: 25 Marks**

Course Objectives:

1. The objective of this course is to provide students with the fundamental elements of machine learning for classification, regression, clustering.
2. Design and evaluate the performance of a different machine learning models.

Course Outcomes:
On completion of the course, students will be able to–
CO1: Implement different supervised and unsupervised learning algorithms.
CO2: Evaluate performance of machine learning algorithms for real-world applications.

Sr. No	List of Assignments
1.	Perform following operations on given dataset. a) Load the Data set b) Find Shape of Data c) Find the summary of the data d) Find data type of each column e) Find Missing Values f) Finding out Zero's g) Find Mean h) Replace the missing values i) Draw the pair plot j) Create subsets as per give instructions k) divide dataset in training (75%) and testing (25%).
2.	For the given data set, apply linear regression model to predict the value of the given Label 1. Assess the performance of regression models using MSE, MAE and R-Square metrics 2. Visualize simple regression model. 3. Assess the performance of Multivariate regression model using MSE, MAE and R-Square metrics
3.	Every year many students give the GRE exam to get admission in foreign Universities. The data set contains GRE Scores (out of 340), TOEFL Scores (out of 120), University Rating (out of 5), Statement of Purpose strength (out of 5), Letter of Recommendation strength (out of 5), Undergraduate GPA (out of 10), Research Experience (0=no, 1=yes), Admitted (0=no, 1=yes). Admitted is the target variable. The counselor of the firm is supposed check whether the student will get an admission or not based on his/her GRE score and Academic Score. So to help the counselor to take appropriate decisions build a machine learning model classifier using Decision tree to predict whether a student will get admission or not.
4.	The customer dataset gives the data of Income and money spent by the customers visiting a Shopping Mall. The data set contains Customer ID, Gender, Age, Annual Income, Spending Score. Therefore, as a mall owner you need to find the group of people who are the profitable customers

	for the mall owner. Apply at least two clustering algorithms (based on Spending Score) to find the group of customers.
5.	The dataset comprises the list of transactions of a retail company over the period of one week. It contains a total of 7501 transaction records where each record consists of the list of items sold in one transaction. Using this record of transactions and items in each transaction, find the association rules between items. (There is no header in the dataset and the first row contains the first transaction, so mentioned header = none here while loading the dataset.) Perform the following steps: 1. Data Pre-processing 2. Generate the list of transactions from the dataset 3. Train Apriori algorithm on the dataset 4. Visualize the list of rules 5. Generated rules depend on the values of hyper parameters. By increasing the minimum confidence value and find the rules accordingly
6.	A SMS unsolicited mail (every now and then known as cell smartphone junk mail) is any junk message brought to a cellular phone as textual content messaging via the Short Message Service (SMS). Use probabilistic approach (Naive Bayes Classifier / Bayesian Network) to implement SMS Spam Filtering system. SMS messages are categorized as SPAM or HAM using features like length of message, word depend, unique keywords etc. a. Apply Data pre-processing (Label Encoding, Data Transformation….) techniques if necessary b. Perform data-preparation (Train-Test Split) c. Apply a Machine Learning Algorithm and Evaluate Models

Subject In-charge	Academic Coordinator	HOD
Prof. Mrs. M. R .Apsangi	Prof. K.S.Balbudhe	Dr.S.A.Mahajan