Course Title	Machine Learning lab				Course Type		HC	
Course Code	B20EF0507	Credits	2		Class		V Semester	
Course Structure			Contact Hours	Work Load	Total Number of Classes Per Semester		Assessment in Weightage	
	TLP	Credits						
	Theory	-	-	-				
	Practice	1	2	2	Theory	Practical	CIE	SEE
	-	-	-	-				
	Total	1	2	2	-	26	25	25

COURSE OVERVIEW:

This course gives an overview of many machine learning concepts, techniques and algorithms beginning with topics such as classification and linear regression and ending up with more recent topics such support vector machines. Also, students will learn to analyze the data and predict the results.

COURSE OBJECTIVE (S):

The objectives of this course are to:

1. Apply the mathematical and statistical prospective of machine learning algorithms through python programming.
2. Analyze machine learning algorithms to predict individual income, medical data & more
3. Execute Python program to design and implement classifiers for machine learning applications.
4. Analyze model generated from various dataset.

COURSE OUTCOMES (COs)

After the completion of the course, the student will be able to:

CO#	Course Outcomes	POs	PSOs
CO1	Implement various data processing operations and evaluate prediction results.	1,3,4,5,8,9,1012	1,2,3
CO2	Execute K-mean clustering to perform segmentation.	1,3,4,5,8,9,1012	1,2,3
CO3	Evaluate supervised machine learning algorithm to predict classes	1 to 5,8,9,10,12	1,2,3
CO4	Assess regression algorithms for predicting continuous values.	1 to 5,8,9,10,12	1,2,3

CO5	Examine efficiency of linear classifier on aforesaid data set	1 to 5,8,9,10,12	1,2,3
CO6	Create a mini project to address data optimization techniques in retail industry	1 to 5,9,10,11,12	1,2,3

BLOOM'S LEVEL OF THE COURSE OUTCOMES

CO#	Bloom's Level					
	Remember (L1)	Understand (L2)	Apply (L3)	Analyze (L4)	Evaluate (L5)	Create (L6)
CO1			✓			
CO2			✓			
CO3					✓	
CO4					✓	
CO5			✓			
CO6						✓

COURSE ARTICULATION MATRIX

CO# / POs	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2		1	2	2			1	2	2		2	3	2	2
CO2	2		1	1	2			1	2	2		2	3	2	2
CO3	3	2	2	2	2			1	2	2		2	3	2	2
CO4	3	2	2	2	2			1	2	2		2	3	3	2
CO5	3	2	2	2	2			1	2	2		2	3	2	2
CO6	3	3	3	3	3			1	3	2	3	3	3	3	3

No	Title of the Experiment	Tools and Techniques	Expected Skill /Ability
	Part-A		
1.	Write a python program to predict income levels of adult individuals using a linear regression model. The process includes training, testing and evaluating the model on the Adult dataset. In this sample experiment you need to train a classifier on the Adult dataset, to predict whether an individual's income is greater or less than $50,000. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Adult Income Dataset: https://raw.githubusercontent.com/jbrownlee/Datasets /master/adult-all.csv	Anaconda, PyCharm	Python Programming skills
2.	Write a python code to cluster similar companies into same group given their Wikipedia articles and can be used to assign cluster to new company. The students are expected to demonstrate how to use the K-Means clustering algorithm to perform segmentation on companies from the Standard & Poor (S&P) 500 index, based on the text of Wikipedia articles about each company. Refer the following link (S&P) 500 index: https://en.wikipedia.org/wiki/List_of_S%26P_500_comp anies	Anaconda, PyCharm	Python Programming skills
3.	Write a python program to classify the medical dataset using KNN. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Medical Data set: https://lionbridge.ai/datasets/18-free-life-sciences-medical-datasets-for-machine-learning/	Anaconda, PyCharm	Python Programming skills
4.	Write a python code to predict the real estate sales price of a house based upon various quantitative features about the house and sale. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the real estate sales price of a house Data set: https://data.world/datasets/real-estate	Anaconda, PyCharm	Python Programming skills

5.	Write a python program to predict income levels of adult individuals using Decision Tree Regression Model. The process includes training, testing and evaluating the model on the Adult dataset. In this sample experiment you need to train a classifier on the Adult dataset, to predict whether an individual's income is greater or less than $50,000. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Adult Income Dataset: https://raw.githubusercontent.com/jbrownlee/Datasets /master/adult-all.csv	Anaconda, PyCharm	Python Programming skills
6.	Write a python program to predict income levels of adult individuals using Support Vector Machine Regression Model. The process includes training, testing and evaluating the model on the Adult dataset. In this sample experiment you need to train a classifier on the Adult dataset, to predict whether an individual's income is greater or less than $50,000. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Adult Income Dataset: https://raw.githubusercontent.com/jbrownlee/Datasets /master/adult-all.csv	Anaconda, PyCharm	Python Programming skills
7.	Write a python program to classify the medical dataset using Multilayer Perceptron Classifier. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Medical Data set: https://lionbridge.ai/datasets/18-free-life-sciences-medical-datasets-for-machine-learning/	Anaconda, PyCharm	Python Programming skills

8.	Write a python program to classify the medical dataset using Gaussian Naïve Bayes Classifier. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Medical Data set: https://lionbridge.ai/datasets/18-free-life-sciences-medical-datasets-for-machine-learning/	Anaconda, PyCharm	Python Programming skills
9.	Compare the performances of the classification models that you implemented to classify the medical dataset. The students are expected to demonstrate how you can perform basic data processing operations, split the dataset into training and test sets, train the model, score the test dataset, and evaluate the predictions. Refer the following link to download the Medical Data set: https://lionbridge.ai/datasets/18-free-life-sciences-medical-datasets-for-machine-learning/	Anaconda, PyCharm	Python Programming skills
10.	Consider the publicly available banknote authentication Data Set on the UCI machine learning repository. Implement the deep learning classifier to classify the aforesaid data set. Refer the following link to download the Medical Data set: https://archive.ics.uci.edu/ml/datasets/banknote+authentication	Anaconda, PyCharm	Python Programming skills
Part - B			
	Mini Project		

| 1 | **Retail Price Optimization using Machine Learning**

In this machine learning pricing project, implement a retail price optimization algorithm using regression trees. This is one of the first steps to building a dynamic pricing model.

Pricing a product is a crucial aspect in any business. A lot of thought process is put into it. There are different strategies to price different kinds of products. There are products whose sales are quite sensitive to their prices and as such a small change in their price can lead to noticeable change in their sales. While there are products whose sales are not much affected by their price - these tend to be either luxury items or necessities (like certain medicines). This machine learning retail price optimization project will focus on the former type of products.

Price elasticity of demand (Epd), or elasticity, is the degree to which the effective desire for something changes as its price changes. In general, people desire things less as those things become more expensive. However, for some products, the customers desire could drop sharply even with a little price increase, and for other products, it could stay almost the same even with a big price increase. Economists use the term elasticity to denote this sensitivity to price increases. More precisely, price elasticity gives the percentage change in quantity demanded when there is a one percent increase in price, holding everything else constant. | To implement a retail price optimization algorithm using regression trees | Python Programming skills |
|---|---|---|---|

In this machine learning pricing optimization case study, we will take the data of a cafe and based on their past sales, identify the optimal prices for their items based on the price elasticity of the items. For each item, first the price elasticity will be calculated and then the optimal price will be figured. While this is taking a particular cafe data, this work can be extended to price any product.

Briefly, price optimization uses data analysis techniques to pursue **two main objectives**:

Understanding how customers will react to different pricing strategies for products and services, i.e., understanding the elasticity of the demand.

Finding the best prices for a given company, considering its goals.

Pricing systems have evolved since the early 1970s until now, from applying very simple strategies, such as a standard markup to base cost, to being capable of predicting the demand of products or services and finding the best price to achieve the set KPI.

Price optimization techniques can help retailers **evaluate the potential impact of sales promotions** or **estimate the right price** for each product if they want to **sell it in a certain period of time**.

Modules:
Competition
Weather
Season
Special events / holidays
Macroeconomic variables
Operating costs
Warehouse information

To **determine**:
The initial price
The best price
The discount price
The promotional price