Course Code : **CN304**
Course Title : **MACHINE LEARNING LAB**
Pre-requisite(s) : **Python Programming Lab**
Co- requisite(s) : **Introduction to Machine Learning**
Credits: 2 L:0 T:0 P:4
Class schedule per week : **04**
Class : **BCA**
Semester / Level : **V/3**
Branch : **Bachelor of Computer Applications**

Course Objectives
This course enables the students to:

A.	Make use of datasets in implementing the machine learning algorithms.
B.	Understand the basic concepts and techniques of machine learning through programming.
C.	Develop skills of machine learning packages for solving practical problems.
D.	Understand the importance of feature selection and feature transformation.
E.	Implement machine learning concepts and algorithms in any language.

Course Outcomes
After the completion of this course, students will be able to:

CO1	Describe the implementation procedures for the machine learning algorithms.
CO2	Understand the features of data to apply on real-world problems.
CO3	Apply appropriate datasets to the machine learning algorithms.
CO4	Perform machine learning experiments to solve real-world problems.
CO5	Predict/classify using machine learning algorithms.

Syllabus

List of Programs as Assignments:

1. Introduction to the language – Importing datasets – Data visualization.
2. Implement Linear Regression problem. For example, based on a dataset comprising of existing set of prices and area/size of the houses, predict the estimated price of a given house.
3. Based on multiple features/variables perform Linear Regression. For example, based on a number of additional features like number of bedrooms, servant room, number of balconies, number of houses of years a house has been built – predict the price of a house.
4. Implement a classification/ logistic regression problem. For example, based on different features of student's data, classify, whether a student is suitable for a particular activity. Based on the available dataset, a student can also implement another classification problem like checking whether an email is spam or not.
5. Use some function for regularization of dataset based on problem 4.
6. Use some function for neural networks, like Stochastic Gradient Descent or backpropagation - algorithm to predict the value of a variable based on the dataset of problem 4.
7. Write a program to demonstrate the working of the decision tree based ID3 algorithm.
8. Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a CSV file.
9. Assuming a set of documents that need to be classified, use the naïve Bayesian Classifier model to perform this task. Calculate the accuracy, precision, and recall for your data set.

10. Write a program to construct a Bayesian network considering medical data. Use this model to demonstrate the diagnosis of heart patients using standard Heart Disease Data Set.
11. Implement dimensionality reduction using PCA method.
12. Implement classification using SVM to classify emails into spam or non-spam categories.
13. Apply K-Means algorithm to cluster similar documents together.
14. Implement agglomerative and divisive clustering approach.
15. For a given dataset, analyze which regression model performs better.

TEXT BOOKS:

1. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems by Aurélien Géron.
2. C. Müller and Sarah Guido, Introduction to Machine Learning with Python: A Guide for Data Scientists by Andreas.

Gaps in the Syllabus (to meet Industry/Profession requirements)

1. Industry oriented projects.

POs met through Gaps in the Syllabus – PO2, PO4

Topics beyond syllabus/Advanced topics/Design

1. Analysis of Time series data and Stream data.

POs met through Topics beyond syllabus/Advanced topics/Design – PO4

Course Outcome (CO) Attainment Assessment Tools & Evaluation Procedure

Direct Assessment

Assessment Tool	% Contribution during CO Assessment
First Quiz	10
Second Quiz	10
Viva voce	20
Day to day performance	30
Exam Evaluation performance	30

Indirect Assessment

1. Student Feedback on Faculty
2. Student Feedback on Course Outcome

Mapping between COs and Program Outcomes

Course outcome	Program Outcomes (POs)												Program Specific Outcomes		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CO1	3	1	1	1	1	1	1	1	1	1	1	2	3	1	1
CO2	3	2	2	2	1	1	1	1	1	1	1	2	3	1	1
CO3	3	3	3	3	3	3	2	2	2	2	2	2	3	3	2
CO4	3	3	3	3	3	2	2	2	2	2	2	2	3	3	2
CO5	3	3	3	3	3	3	2	2	2	2	2	2	3	3	2

Correlation Levels 1, 2 or 3 as defined below:
1: Slight (Low) 2: Moderate (Medium) 3: Substantial (High)

Mapping Between COs and Course Delivery (CD) methods

CD Code	Course Delivery Methods	Course Outcome	Course Delivery Method Used
CD1	Lecture by use of Boards/LCD Projectors	CO1	CD5, CD8, CD9
CD2	Tutorials/Assignments	CO2	CD5, CD9
CD3	Seminars	CO3	CD5, CD9
CD4	Mini Projects/Projects	CO4	CD5, CD6, CD9
CD5	Laboratory Experiments/Teaching Aids	CO5	CD5, CD6, CD9
CD6	Industrial/Guest Lectures		
CD7	Industrial Visits/In-plant Training		
CD8	Self- learning such as use of NPTEL Materials and Internets		
CD9	Simulation		