BCSE0105: MACHINE LEARNING

Objective: *To introduce students to the basic concepts and techniques of Machine Learning. To develop skills of using recent machine learning software for solving practical problems. To gain experience of doing independent study and research.*

Credits: 03   L–T–P-J: 3–0–0-0

Module No.	Content	Teaching Hours
I	**Introduction:** Machine Learning basics, Hypothesis space and inductive bias, training and test set, and cross-validation. **Introduction to Statistical Learning:** Bayesian Method. **Machine Learning:** Supervised (Regression, Classification) vs. Unsupervised (Clustering) Learning. **Data Preprocessing:** Imputation, Outlier management, One hot encoding, Dimensionality Reduction- feature extraction, Principal Component Analysis (PCA), Singular Value Decomposition **Supervised Learning:** Regression- Linear regression, Polynomial regression, Classification- Logistic regression, k-nearest neighbor classifier,	20
II	**Supervised Learning:** Decision tree classifier, Naïve Bayes classifier Support Vector Machine (SVM)Classifier, **Unsupervised Learning:** k-means clustering, Hierarchical clustering **Underfitting vs Overfitting:** Regularization and Bias/Variance. **Ensemble methods:** Bagging, Boosting, Improving classification with Ada-Boost algorithm.	20

Text Book:

- Tom M. Mitchell, Machine Learning. Tata McGraw-Hill Education, 2013.
- Alpaydin, E. . Introduction to machine learning. MIT press, 2009.

Reference Books:

- Harrington, P. , " Machine learning in action", Shelter Island, NY: Manning Publications Co, 2012.
- Bishop, C. M. . Pattern recognition and machine learning (information science and statistics) springer-verlag new york. Inc. Secaucus, NJ, USA. 2006

Outcome: After completion of Lab, student will be able to:

- CO1: Apply the basic concepts of machine learning.
- CO2: Apply the concepts of regression and re-sampling methods.
- CO3: Design supervised and re-enforcement learning based solution.
- CO4: Apply the ensemble methods for improving classification.
- CO5: Identify the ways of feature extraction, reduction and selection.
- CO6: Design the applications of machine learning algorithms.

Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

COs	POs/PSOs
CO1	PO1,PO2/PSO3,PSO4
CO2	PO1,PO2/PSO1,PSO3
CO3	PO1,PO3,PO5/PSO1,PSO3
CO4	PO1 /PSO1
CO5	PO2/PSO3
CO6	PO1,PO2,PO3/PSO1,PSO2,PSO4

BCSE0133: MACHINE LEARNING LAB

Credits: 01   **L–T–P-J: 0–0-2-0**

S.No	PROGRAM
1	Estimate parameters of a model based on Linear Regression method using a given set of training data set.
2	Estimate parameters of a model based on maximum likelihood estimation method using a given set of training data set.
3	Compute weights of ANN based on back propagation method using a given training dataset.
4	Compute probability of a person to be diabetic based on a given dataset of diabetic persons using Naïve Bayesian classifier.
5	Classify a person as male or female based on a given dataset using naïve Bayesian Classifier, and calculate accuracy, precision, and recall for your data set.
6	Write a program to implement *k*-Nearest Neighbour method to classify the iris data set. Print both correct and wrong predictions. Use Java/Python ML library classes
7	Predicts whether the bank should approve the loan of an applicant, based on his profile using Ensemble learning method.
8	Apply Ensemble learning to cluster a set of data stored in a .CSV file. Use the same dataset for clustering using *k*-Means method. Compare the results of these two algorithms and comment on the quality of clustering. You can add Java/Python ML library classes/API in the program.
9	The stock prediction data is used to predict, whether the stock will go up or down. Perform the task of feature selection with the help of wrapper method.
10	Identify principal components of Big Mart sales data using Principal component analysis (PCA). Also plot the result of PCA, and give inferences.

Outcome: After studying the subject, the students will be able to:
- CO1: Applythe machine learning algorithms in the area of text, audio and image processing.
- CO2: Apply classification algorithms to design complex problems.
- CO3: Design solution to societal issues using machine learning algorithms.
- CO4: Analyze the view problem in the perspective of machine learning.

Mapping of Course Outcomes (COs) with Program Outcomes (POs) and Program Specific Outcomes (PSOs):

CO	PO/PSO
CO1	PO3,PO5/PSO4
CO2	PO1,PO5/PSO1,PSO3,PSO4
CO3	PO3 /PSO1,PSO2,PSO4
CO4	PO2,PO3/PSO1,PSO2,PSO4

BCSE0141: MACHINE LEARNING PROJECT

Credits: 02   **L–T–P-J: 0–0–0-0**

1. Iris Flowers Classification ML Project: Learn about Supervised Machine Learning Algorithms

Objective: The goal of this machine learning project is to classify the flowers into among the three species – virginica, setosa, or versicolor based on length and width of petals and sepals.

2. BigMart Sales Prediction ML Project: Learn about Unsupervised Machine Learning Algorithms

Objective: The goal of the BigMart sales prediction ML project is to build a regression model to predict the sales of each of 1559 products for the following year in each of the 10 different BigMart outlets.

3. Social Media Sentiment Analysis using Twitter Dataset

Objective: Working with the twitter dataset will help you understand the challenges associated with social media data mining and also learn about classifiers in depth.

4. Sales Forecasting using Walmart Dataset

Objective: The goal of this machine learning project is to forecast sales for each department in each outlet to help them make better data driven decisions for channel optimization and inventory planning.

5. Learn to build Recommender Systems with MovieLens Dataset

Objective: The goal of this machine learning project is to start working with this dataset by building a world-cloud visualization of movie titles to build a movie recommender system.

6. Stock Prices Predictor

Objective: Stock prices predictor is a system that learns about the performance of a company and predicts future stock prices.

7. Boston Housing Price Prediction ML Project

Objective: The goal of this machine learning project is to predict the selling price of a new home by applying basic machine learning concepts on the housing prices data.

8. MNIST Handwritten Digit Classification

Objective: The goal of this machine learning project is study, analyze and recognize elements in the images.

9. Human Activity Recognition using Smartphone Dataset

Objective: The goal of this machine learning project is to build a classification model that can precisely identify human fitness activities. Working on this machine learning project will help you understand how to solve multi-classification problems.

10. Use classification and clustering techniques to deal with the Turkiye Student Evaluation Data Set

Objective: This dataset is based on an evaluation form filled out by students for different courses. It has different attributes including attendance, difficulty, score for each evaluation question, among others. This is an unsupervised learning problem.

11. Predict height and weight from Heights and Weights dataset.

Objective: The goal of this project is to learn the concept of Regression.

12. Text Mining Data Set

Objective: The goal of this project is classify the documents according to their labels.

13. Predict the income class of any country population

Objective: The goal of this project is to work with imbalanced classification problem.

14. Classify the type of sound from the audio.

Objective: The goal of this project is to introduce the student to audio processing in the usual classification scenario.