SEMESTER-VI

COURSE 15 B: MACHINE LEARNING

Theory **Credits: 3** **3 hrs/week**

Course Objectives:

1. Understand fundamental concepts, types, and applications of machine learning.
2. Develop, evaluate, and optimize machine learning models through preprocessing, training, and feature engineering techniques.
3. Apply supervised and unsupervised learning algorithms to real-world problems using appropriate tools and methods.

Course Outcomes:

Upon successful completion of this course, students will be able to:

1. Describe various machine learning paradigms, data types, and the overall structure of a machine learning pipeline.
2. Perform data preprocessing, feature engineering, and evaluate models using appropriate metrics.
3. Implement and analyze supervised learning algorithms for regression and classification tasks.
4. Apply unsupervised learning techniques for clustering and identify suitable machine learning approaches for specific application domains

Unit 1. Introduction to Machine Learning:
Introduction to Machine Learning: Types of human learning, What is machine learning?, Types of machine learning: supervised, unsupervised, semi-supervised and reinforcement learning, machine learning activities, applications of machine learning. Types of data in machine learning, Structure of data

Unit 2. Model Preparation, Evaluation and feature engineering:
Data pre-processing, Model selection and training (for supervised learning), Model representation and interpretability, Evaluating machine learning algorithms and performance enhancement of models. What is feature engineering?, Feature transformation, Feature subset selection. Principal component analysis.

Unit 3. Supervised Learning-Regression:
Regression: Introduction of regression, Regression algorithms: Simple linear regression, Multiple linear regression, Polynomial regression model, Logistic regression, Maximum likelihood estimation.

Unit 4. Supervised Learning- Classification:
Introduction of supervised learning, Classification model and learning steps, Classification algorithms: Naïve Bayes classifier, k-Nearest Neighbour (kNN), Decision tree, Support vector machines, Random Forest.

Unit 5. Unsupervised Learning:
Introduction of unsupervised learning, Unsupervised vs supervised learning, Application of unsupervised learning, Clustering and its types, Partitioning method: k-Means and KMedoids, Hierarchical clustering, Density-based methods – DBSCAN.

Case-study of ML applications: Image recognition, speech recognition, Email spam filtering, Online fraud detection and other.

Textbooks:

1. Introduction to Machine Learning, Ethem Alpaydin, MIT Press, Fourth Edition, 2020.
2. Machine Learning: Theory and Practice, M N Murthy, V.S Ananthanarayana, Universities press
3. Machine Learning, S. Sridhar, M. Vijayalakshmi, Second Edition, Oxford University Press

Reference Books:

1. Machine Learning: An Algorithmic Perspective, Second Edition, Stephen Marsland, CRC Press, 2014
2. Machine Learning, Tom Mitchell, McGraw Hill, 3rd Edition.
3. Python Machine Learning, Sebastain Raschka, Vahid Mirjalili , Packt publishing 3rd Edition, 2019.

Activities:

Outcome: Describe various machine learning paradigms, data types, and the overall structure of a machine learning pipeline.

Activity: Prepare a detailed comparative report/chart explaining supervised, unsupervised, and reinforcement learning paradigms, data types, and step-by-step machine learning pipeline stages.
Evaluation Method: Rubric-based assessment evaluating completeness, clarity, correctness, and presentation quality - scored on a 10-point scale.

Outcome: Perform data preprocessing, feature engineering, and evaluate models using appropriate metrics.

Activity: Conduct a hands-on lab exercise using a real dataset to perform data cleaning, normalization, feature extraction/selection, and evaluate model performance using metrics like accuracy, precision, recall, and F1-score.

Evaluation Method: Practical assessment including code correctness, applied techniques, and interpretation of evaluation metrics; assessed with a rubric out of 10.

Outcome: Implement and analyze supervised learning algorithms for regression and classification tasks.

Activity: Implement at least two supervised learning algorithms (e.g., Linear Regression and Decision Trees) to solve prediction tasks, followed by comparative analysis of their performance on test datasets.

Evaluation Method: Code and report evaluation focusing on implementation accuracy, performance comparison, and analysis depth; scored on a 10-point rubric.

Outcome: Apply unsupervised learning techniques for clustering and identify suitable machine learning approaches for specific application domains.

Activity: Perform clustering (e.g., K-Means, Hierarchical) on a given dataset and prepare a case study selecting and justifying machine learning methods suited for different application scenarios.
Evaluation Method: Lab practical combined with a written case study; assessed for correct algorithm application, cluster interpretation, and justification of approach - evaluated on a 10-point rubric.

SEMESTER-VI

COURSE 15 B: MACHINE LEARNING

Practical **Credits: 1** **2 hrs/week**

Lab Experiments:

1. Write a python program to import and export data using Pandas library functions.
2. Demonstrate various data pre-processing techniques for a given dataset
3. Implement Dimensionality reduction using the Principal Component Analysis (PCA) method.
4. Write a Python program to demonstrate various Data Visualization Techniques.
5. Implement MLE on a Dataset
6. Implement Simple and Multiple Linear Regression Models.
7. Develop Logistic Regression Model for a given dataset.
8. Develop Decision Tree Classification model for a given dataset and use it to classify a new sample.
9. Implement Naïve Bayes Classification in Python.
10. Develop K-Means for a Given Dataset
11. Build KNN Classification model for a given dataset.
12. Develop DBSCAN on a given Dataset