BCE-P614

MACHINE LEARNING –II

MM : 100 **Sessional : 30**

Time : 3Hr **ESE : 70**

L T P **Credit : 3**

3 0 0

Prerequisite: Knowledge of basic machine learning concept**,** Understanding of Basic Programming Concept and Mathematics (probability and statistics).

Objectives:

The course has following objectives

- To learn the fundamentals of Advanced Machine Learning.
- To understand basic components of an intelligence system for regression & classification problems.
- To explore applications and implementation of advanced machine learning.
- To understand different types of machine learning algorithms, framework and tools.
- To learn how to use machine learning models to solve real world problems.

Outcomes:

On completion of course, student will be able to:

- List various approaches of Advanced Machine Learning.
- Describe machine learning algorithms to solve the real-world problems.
- Develop Hypothesis and machine learning models.
- Identify appropriate models for solving machine learning problems.
- Apply learning techniques to solve real world machine learning problems.
- Evaluate and interpret the results of the algorithms.
- Choose an appropriate machine learning model for a defined problem.

NOTE: The question paper shall consist of two sections A and B. Section A contains 10 short type questions of 6 marks each and student shall be required to attempt any five questions. Section B contains 8 long type questions of ten marks each and student shall be required to attempt any four questions. Questions shall be uniformly distributed from the entire syllabus.

UNIT I

Brief Recap of Fundamentals of Machine Learning, Overview of problem – Regression, Classification and clustering, Types of data – Low dimensional Data and High dimensional Unstructured Data. Machine Learning Tools, Frameworks, and popular public Data repositories. Basics of Linear Algebra, Pandas and Numpy Libraries.



UNIT II

Statistical decision theory - Bias, Variance, Bias-Variance trade-off, overfitting and underfitting, pre-processing of data - Feature extraction and selection, Dimensionality Reduction, Train-Test splitting strategy. Cross validation Techniques. Data exploration, Performance metrics for ML Algorithm, Fine tuning of model – Grid Search, Randomized Search, Ensemble Method. Concept of Bagging & Boosting.

UNIT III

Introduction to regression and classification problems, Linear regression in One Variable: Introduction, Gradient Descent Algorithm, Gradient Descent Implementation, Gradient Descent Update Rule for Regression, Data Preparation, Scoring, Surface Plots and Contours, Visualizing Loss function and Gradient Descent Trajectory. Linear regression in multiple variables: Gradient Descent for Multiple Variables, Features and Polynomial Regression.

UNIT IV

Dimensionality Reduction and Feature Selection: Data Compression and Visualization, Principal Component Analysis Problem Formulation, PCA Algorithms, Reconstruction from Compressed Representation, Choosing number of Principal Components. K Nearest Neighbors: Distance Metrics, 1-NN algorithm, K-NN algorithm, Weighted KNN algorithm, Performance of NN as data grows, Issues with high dimensions, high data scarcity and computational complexity, Classification using KNN.

UNIT V

K-Means Clustering: Intro to Clustering, Intro to Hierarchical Clustering, DBSCAN. Decision Trees and Random Forests: Intro to Decision Trees, Decision Trees implementation, Decision Forests Visualization using, Random Forests Ensembles. Bayes Theorem Formula and Proof, Naïve Bayes Classifier and Introduction and Implementation of Support Vector Machine (SVM).

Project Task - End-to-End Machine learning Model Development – Housing Price Prediction, Stock Price Prediction, MNIST Digit Classification.

Text Books:

- Introduction to Machine Learning with Python: A Guide for Data Scientists, Andreas C. Müller and Sarah Guido, O'Reilly, 2016.
- Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, Aurelien Geron, O'Reilly, 2017.

Suggested Readings:

- Elements of Machine Learning, Pat Langley Morgan Kaufmann Publishers, Inc. 1995.
- Machine Learning: The Art and Science of Algorithms that Make Sense of Data, Peter Flach, Cambridge University Press, 2012.
- Pattern recognition and machine learning by Christopher Bishop, Springer Verlag, 2006.
- Machine Learning, Tom Mitchell, McGraw Hill, 1997.