UML501: MACHINE LEARNING

L	T	P	Cr
3	0	2	4.0

Course Objectives: This course provides a broad introduction to machine learning and statistical pattern recognition. It offers some of the most cost-effective approaches to automated knowledge acquisition in emerging data-rich disciplines and focuses on the theoretical understanding of these methods, as well as their computational implications.

Introduction: Introduction to Machine Learning, Basic Concepts, Issues, Applications, Types of machine learning: Supervised learning, Unsupervised learning, Semi-supervised learning, Reinforcement learning, Transfer Learning .

Data Collection: Structured and Unstructured Data, Data Collection using web scraping, data collection using APIs

Data Pre-processing: Need of Data Pre-processing, Data Pre-processing Methods: Data Cleaning, Data Integration, Data Transformation, Data Reduction; Feature Scaling (Normalization and Standardization), Splitting dataset into Training and Testing set.

Regression: Linear Regression, Multiple Linear Regression and Polynomial Regression, Evaluating Regression Models' Performance (RMSE, Mean Absolute Error, Correlation, RSquare), Regularization Methods

Classification: Need and Applications of Classification, Logistic Regression, , Naïve Bayes algorithm; K-Nearest Neighbours (K-NN), Support Vector Machine (SVM), Overview of Tree based methods, Bagging: Random Forests, Boosting: AdaBoost, XGBoost, Evaluating Classification Models' Performance (Sensitivity, Specificity, Precision, Recall, etc).

Clustering: Hierarchical methods, Density-based methods.

Association Rules Learning: Need and Application of Association Rules Learning, Basic concepts of Association Rule Mining, Naïve algorithm, Apriori algorithm.

Introduction to Deep Learning: Introduction to Artificial Neural Networks (ANNs), Artificial Neurons, Layers, Perceptron, Multilayer Perceptron, Advanced Deep Neural Networks (DNNs), Batch Normalization, Hyperparameter tuning, Activation Functions, Metrics, Optimization, Regularization.

Laboratory Work:

Implement data preprocessing, Simple Linear Regression, Multiple Linear Regression, , Random forest classification, AdaBoost, Naïve Bayes algorithm; K-Nearest Neighbors (K-NN), Support Vector Machine , Apriori algorithm and Shallow and Deep Neural Networks in Python (using inbuilt libraries like Pandas/NumPy/Sklearn/PyTorch/Tensorflow and from scratch).

Course Learning Outcomes (CLOs) / Course Objectives (COs):

After the completion of the course, the student will be able to:
1. Analyze methods and theories in the field of machine learning and understand the data collection, pre-processing and analytics pipeline.
2. Comprehend and apply various supervised learning techniques for regression and classification tasks .

3. Comprehend and apply unsupervised learning techniques for clustering and association learning tasks.
4. Understand the concept of Neural Networks and its implementation in the context of Machine Learning.

Text Books:
1. Mitchell M., T., Machine Learning, McGraw Hill (1997) 1st Edition.
2. Alpaydin E., Introduction to Machine Learning, MIT Press (2014) 3rd Edition.
3. Vijayvargia Abhishek, Machine Learning with Python, BPB Publication (2018)

Reference Books:
1. Bishop M., C., Pattern Recognition and Machine Learning, Springer-Verlag (2011) 2nd Edition.
2. Michie D., Spiegelhalter J. D., Taylor C. C., Campbell, J., Machine Learning, Neural and Statistical Classification. Overseas Press (1994).