Semester	3
Course Title	Machine Learning using Python
Course Code	MMCA311C (IPCC)
Credits	3
Total Hours of Pedagogy	40
L–T–P-S	3-0-2-0
CIE	50
SEE	50
TOTAL	100
Exam Type	Theory
Exam Hours	3 Hrs

Course Learning Objectives:

1. To understand the basic theory underlying machine learning, types, and the process.
2. To become familiar with data and visualize univariate, bivariate data using statistical techniques.
3. To understand various machine learning algorithms such as similarity-based learning, regression, decision trees.
4. Developing the skills required for decision-making in dynamic environments.

Module I

Introduction to Machine Learning: Need for Machine Learning, Machine Learning Explained, Machine Learning in Relation to Other Fields, Types of Machine Learning, and Challenges of Machine Learning, Machine Learning Process, and Machine Learning Applications.
Data Preprocessing: Data cleaning and transformation, Handling missing values and outliers, Feature scaling and normalization.

Text Book	1
Chapter	1-1.7
RBT	L3

Module II

Perspectives and issues in Machine Learning: concept Learning: concept Learning task, concept learning as search, find –S Algorithm, Version Space, Candidate Elimination Algorithm.
Understanding DATA: Introduction, Big Data Analytics and Types of Analytics, Big Data Analysis Framework, Descriptive Statistics, Univariate Data Analysis and Visualization, Bivariate Data and Multivariate Data.

Text Book	1
Chapter	2-2.6
RBT	L3

Module III

Basics of Learning Theory: Introduction to Learning and its Types, Introduction to Computation Learning Theory, Design of a Learning System, Introduction to Concept Learning, Induction Biases, Modelling in Machine Learning.
Similarity-based Learning: Introduction to similarity or instance=based Learning, Nearest-Neighbor Learning, Weighted K-Nearest-Neighbor Algorithm.

Text Book	1
Chapter	3-3.6,4-4.3
RBT	L3

Module IV

Regression Analysis: Introduction to Regression, Introduction to Linearity, Correlation, and Causation,

Introduction to Linear Regression, Validation of Regression Methods. **Decision Tree Learning:** Introduction to Decision Tree Learning Model, Structure a Decision Tree, Fundamentals of Entropy, Decision Tree Induction Algorithms: C4.5 Construction, Validating and Pruning of Decision Trees.	
Text Book	1
Chapter	5-5.4,6-6.3
RBT	L3

Module V

Bayesian Learning: Introduction to Probability-based Learning, Fundamentals of Bayes Theorem, Basics of Naïve Bayes Algorithm. **Artificial Neural Networks:** Introduction, Biological Neurons, Artificial Neurons: Simple Model of Artificial Neurons, Artificial Neural Network Structure, Types of Artificial Neural Network.	
Text Book	1
Chapter	9-9.2,10-10.3.2,10.5
RBT	L3

PRACTICAL COMPONENT OF IPCC

Sl. No.	Experiments
1	Implement and demonstrate the FIND-S Algorithm for finding the most specific hypothesis based on a given set of training data samples. Read the training data from a .CSV file.
2	Write a program to implement the naïve Bayesian classifier for a sample training data set stored as a .CSV file Compute the accuracy of the classifier, considering few test data sets.
3	Build an Artificial Neural Network by implementing the Back propagation algorithm and test the same using appropriate data sets.
	Case Studies on Big Data Analytics across different domains:
4	**Healthcare – Predictive Diagnosis** Use of Big Data analytics on electronic health records and medical imaging to predict diseases (e.g., early detection of cancer or heart disease). **Retail – Personalized Recommendations** Amazon and Walmart use customer purchase history, clickstream data, and browsing behavior to recommend products and optimize inventory. **Banking & Finance – Fraud Detection** Real-time Big Data systems analyze transaction patterns to detect anomalies and prevent fraudulent credit card or online transactions. **Social Media – Sentiment Analysis** Twitter and Facebook data analyzed to understand public opinion, customer feedback, and political trends using natural language processing. **Transport & Logistics – Route Optimization** UPS and FedEx use sensor and GPS data to optimize delivery routes, reduce fuel costs, and predict delays. **Smart Cities – Traffic & Energy Management** IoT and Big Data analytics applied to monitor traffic flow, predict congestion, optimize energy usage, and improve city planning. **Telecommunications – Customer Churn Prediction** Telecom companies analyze call records, usage data, and complaints to predict which customers are likely to leave and design retention strategies. **Sports Analytics – Performance Optimization** Teams like NBA and EPL clubs use Big Data to analyze player performance, injury risks, and game strategies