SEMESTER 4

FOUNDATION OF MACHINE LEARNING

Course Code	**PBAIT404**	**CIE Marks**	**60**
Teaching Hours/Week (L: T:P: R)	**3-0-0-1**	**ESE Marks**	**40**
Credits	**4**	**Exam Hours**	**2 Hrs 30 Mns**
Prerequisites (if any)	**None**	**Course Type**	**Theory**

Course Objectives:

1. To equip the learners to understand the basic and advanced concepts and algorithms in machine learning.
2. To enable the learners to use standard and most popular supervised and unsupervised learning algorithms.

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	**Introduction to ML:** Machine Learning vs. Traditional programming, Machine learning paradigms - supervised, semi-supervised, unsupervised, reinforcement learning. Basics of parameter estimation - maximum likelihood estimation (MLE) and maximum aposteriori estimation (MAP), Bayesian formulation, Least Square Estimation (LS), Minimum Mean Square Estimation (MMSE), The Gradient Descent Method and its variants (Simple Gradient Descent Method, Stochastic Gradient Method, Batch Gradient Method), Illustrative examples of parameter estimation using different methods.	8
2	**Supervised Learning Models:** Regression - Regression Models (Simple Regression, Multivariable linear Regression, Polynomial Regression, Logistic Regression), Illustrative examples of designing regression models. **Decision Trees:** Working Principle, Decision Tree formation using ID3 Algorithm, The Classification and Regression Trees Algorithm (CART). Illustrative examples on designing decision trees. **Generalisation and Overfitting:** Idea of overfitting, LASSO and RIDGE regularization, Idea of Training, Testing, Validation.	8

KTU SPECIAL

	Dimensionality reduction: Principal Component Analysis, Multidimensional scaling.	
3	**Clustering:** Similarity measures, Hierarchical Clustering - Agglomerative Clustering, partitional clustering, K-means clustering. Naive Bayes Classifier - Working Principle, Types of Naive Bayes Classifiers, Illustrative examples. **More Supervised Learning Methods:** Support Vector machines (SVM) - Maximum Margin Classification, Mathematics behind Maximum Margin Classification, Hard Margin SVM, Soft Margin SVM, Non-Linear SVM, The Kernel Trick, Different Kernel Functions (Linear Kernel, Polynomial Kernel, RBF Kernel, Sigmoid Kernel).	8
4	**Ensemble Learning:** Basic Techniques (Voting, Simple and Weighted Averaging Techniques), Advanced Techniques (Bagging, Boosting, Stacking) - The Random Forest Classifier. **Resampling methods**: Bootstrapping, Cross Validation. Practical aspects -Bias-Variance trade-off **Assessment of Classification Models:** Performance measures- Confusion Matrix, Accuracy, Precision, Recall, F1-Score, Receiver Operating Characteristics Curve (ROC), Area Under the Curve (AUC), Illustrative examples on Classification model assessment.	8

KTU SPECIAL

Suggestion on Project Topics

Students may identify a real-life problem that can be solved using the concepts of machine learning. They may either use a standard dataset or build their own, for the purpose of applying machine learning algorithms. Besides, they may make use of various evaluation parameters to measure the efficiency of the methods.

A few sample questions are given below.

1. Classification of Handwritten Digits: Classify handwritten digits (0-9) from the MNIST dataset using various machine learning algorithms and compare their performance. This will

involve preprocessing the data, applying different classification techniques and using ensemble methods to improve accuracy.

2. Predicting Student Examination Performance Based on Internal Marks: Develop a real life dataset using the historical data of your institution. Use various machine learning algorithms to build models to predict students' final examination scores based on their internal assessment marks and evaluate their performance. Use ensemble methods to improve accuracy of prediction. Also, segment the students using clustering techniques to identify potential students at academic risk so that the educators can make timely interventions.

Course Assessment Method
(CIE: 60 marks,ESE: 40 marks)

Continuous Internal Evaluation Marks (CIE):

Attendance	Project	Internal Examination-1	Internal Examination- 2	Total
5	30	12.5	12.5	60

KTU SPECIAL

End Semester Examination Marks (ESE)

In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions.

Part A	Part B	Total
• 2 Questions from each module. • Total of 8 Questions, each carrying 2 marks **(8x2 =16 marks)**	• 2 questions will be given from each module, out of which 1 question should be answered. • Each question can have a maximum of 2 subdivisions. • Each question carries 6 marks. **(4x6 = 24 marks)**	**40**

Course Outcomes (COs)

At the end of the course students should be able to:

	Course Outcome	Bloom's Knowledge Level (KL)
CO1	Explain the fundamental concepts of machine learning, including different paradigms, parameter estimation techniques, and optimization methods.	**K2**

CO2	Describe various supervised learning models such as regression, decision trees and dimensionality reduction techniques, and their role in predictive modeling.	**K3**
CO3	Implement clustering algorithms, including hierarchical and partitional clustering, and apply Naïve Bayes classifiers for classification tasks.	**K3**
CO4	Evaluate classification models using performance metrics such as confusion matrix, precision, recall, F1-score, and ROC-AUC to assess model effectiveness.	**K3**

Note*: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create*

CO-PO Mapping Table (Mapping of Course Outcomes to Program Outcomes)

	PO1	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PO8**	**PO9**	**PO10**	**PO11**	**PO12**
CO1	3	3	3	3								2
CO2	3	3	3	3	3							2
CO3	3	3	3	3	3							2
CO4	3	3	3	3	3							2

Note: *1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High), -: No Correlation*

KTU SPECIAL

Text Books

Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Introduction to Machine Learning	Ethem Alpaydin	MIT Press	4/e, 2020
2	Data Mining and Analysis: Fundamental Concepts and Algorithms	Mohammed J. Zaki, Wagner Meira	Cambridge University Press	1/e, 2016

Reference Books

Sl. No	**Title of the Book**	**Name of the Author/s**	**Name of the Publisher**	**Edition and Year**
1	Machine Learning	Tom Mitchell	McGraw-Hill	1/e, 1997
2	Neural Networks for Pattern Recognition	Christopher Bishop	Oxford University Press	1/e, 1995
3	Machine Learning: A Probabilistic Perspective	Kevin P Murphy	MIT Press	1/e, 2012

4	The Elements of Statistical Learning	Trevor Hastie, Robert Tibshirani, Jerome Friedman	Springer	2/e, 2007
5	Foundations of Machine Learning	Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar	MIT Press	1/e, 2012

Video Links (NPTEL, SWAYAM…)	
Module No.	**Link ID**
1	https://youtu.be/fC7V8QsPBec?si=8kqBn-7x1RG5V1J
2	https://youtu.be/g__LURKuIj4?si=Xj10NPfMfpQSOhVx
3	https://youtu.be/yG1nETGyW2E?si=ySlxpeWuFAUQBf7-
4	https://youtu.be/zop2zuwF_bc?si=W7TpSHLdi4rykva4

KTU SPECIAL

PBL Course Elements

L: Lecture (3 Hrs.)	**R: Project (1 Hr.), 2 Faculty Members**		
	Tutorial	**Practical**	**Presentation**
Lecture delivery	Project identification	Simulation/ Laboratory Work/ Workshops	Presentation (Progress and Final Presentations)
Group discussion	Project Analysis	Data Collection	Evaluation
Question answer Sessions/ Brainstorming Sessions	Analytical thinking and self-learning	Testing	Project Milestone Reviews, Feedback, project reformation (If required)

Guest Speakers (Industry Experts)	Case Study/ Field Survey Report	Prototyping	Poster Presentation/ Video Presentation: Students present their results in a 2 to 5 minutes video

Assessment and Evaluation for Project Activity

Sl. No	Evaluation for	Allotted Marks
1	Project Planning and Proposal	5
2	Contribution in Progress Presentations and Question Answer Sessions	4
3	Involvement in the project work and Team Work	3
4	Execution and Implementation	10
5	Final Presentations	5
6	Project Quality, Innovation and Creativity	3
	Total	**30**

KTU SPECIAL

1. **Project Planning and Proposal (5 Marks)**

   - Clarity and feasibility of the project plan
   - Research and background understanding
   - Defined objectives and methodology

2. **Contribution in Progress Presentation and Question Answer Sessions (4 Marks)**

   - Individual contribution to the presentation
   - Effectiveness in answering questions and handling feedback

3. **Involvement in the Project Work and Team Work (3 Marks)**

   - Active participation and individual contribution
   - Teamwork and collaboration

4. **Execution and Implementation (10 Marks)**

    - Adherence to the project timeline and milestones
    - Application of theoretical knowledge and problem-solving
    - Final Result

5. **Final Presentation (5 Marks)**

    - Quality and clarity of the overall presentation
    - Individual contribution to the presentation
    - Effectiveness in answering questions

6. **Project Quality, Innovation, and Creativity (3 Marks)**

    - Overall quality and technical excellence of the project
    - Innovation and originality in the project
    - Creativity in solutions and approaches

KTU SPECIAL