Subject Code	Subject Name	L	T	P		C
PCCAIML 502	Introduction to Machine Learning	3	0	0		3
Pre-requisite	NIL					

Course Objectives:

1. Ability to comprehend the concept of supervised and unsupervised learning techniques
2. Differentiate regression, classification and clustering techniques and to implement their algorithms.
3. To analyze the performance of various machine learning techniques and to select appropriate featuresfor training machine learning algorithms.

Expected Course Outcome:

1. Understand the concepts of various machine learning strategies.
2. Handle computational data and learn ANN learning models.
3. Solve real world applications by selecting suitable learning model.
4. Boost the performance of the model by combining results from different approaches.
5. Recognize and classify sequencing patterns using HMM.
6. Infer the association and relationship between the data objects.
7. Construct machine learning model for unseen data and can solve real world application.

Module:1	Introduction to Machine Learning	3 hours

Introduction to Machine Learning (ML); Feature engineering; Learning Paradigm, Generalization of hypothesis, VC Dimension, PAC learning, Applications of ML.

Module:2	Data Handling and ANN	4 hours

Feature selection Mechanisms, Imbalanced data, Outlier detection- Artificial neural networks including backpropagation- Applications

Module:3	ML Models and Evaluation	6 hours

Regression: Multi-variable regression; Model evaluation; Least squares regression; Regularization; LASSO; Applications of regression, Classification – KNN, Naïve Bayes, SVM, Decision Tree; Training and testing classifier models; Cross-validation; Model evaluation (precision, recall, F1-mesure, accuracy, area under curve); Statistical decision theory including discriminant functions and decision surfaces

Module:4	Model Assessment and Inference	4 hours

Model assessment and Selection – Ensemble Learning – Boosting, Bagging, Model Inference and Averaging, Bayesian Theory, EM Algorithm

Module:5	Hidden Markov Models	3 hours

Hidden Markov Models (HMM) with forward-backward and Vierbi algorithms; Sequence classification using HMM; Conditional random fields; Applications of sequence classification such as part-of-speech tagging

Module:6	Association Rules	3 hours

Mining Association Rules in Large Databases. Mining Frequent Patterns-- basic concepts - Efficient and scalable frequent item set mining -methods, Apriori algorithm, FP-Growth algorithm		
Module:7	**Clustering**	**5 hours**
K Means, Hierarchical Clustering – Single, complete, Average linkage; Ward's algorithm; Minimum spanning tree clustering; BIRCH clustering		
Module:8	**Recent Trends**	**2 hours**
Recent Trends and case study		
	Total Lecture hours:	**30 hours**

Text Book(s)	
1.	Ethem Alpaydin, Introduction to Machine Learning, MIT Press, Pearson, Third Edition, 2014.
2.	Friedman Jerome, Trevor Hastie, and Robert Tibshirani. The Elements of Statistical Learning. Springer-Verlag, 2nd Edition, 2013.
3.	Jeeva Jose, Introduction to Machine Learning, Khanna Book Publishing.
Reference Books	
1.	Kevin P. Murphy, "Machine Learning: A Probabilistic Perspective", MIT Press, 2012.
2.	Peter Flach, "Machine Learning: The Art and Science of Algorithms that Make Sense of Data", Cambridge University Press, 2012.
3.	Rajiv Chopra, Machine Learning, Khanna Book Publishing.