Name of the Programme : **M.Sc. in Data Science**
Course Code : **CSD-502**
Title of the Course : **Machine Learning (Theory)**
Number of Credits : **2(2L-0T-0P)**
Total Contact Hours : **30 hours (30L-0T-0P) Effective from AY** : **2023-24**

Pre-requisites for the course	Familiarity with linear algebra, statistics & probability theory	
Course Objectives:	This course provides students with 1. In-depth introduction to three main areas of Machine Learning: supervised and unsupervised and reinforcement learning. 2. This course will cover some of the main models and algorithms for regression, classification, clustering and Markov decision processes. Topics will include linear and logistic regression, regularisation, SVMs and kernel methods, ANNs, clustering, and dimensionality reduction, sequential learning Like HMM and deep learning CNN and RNN	
Content:	**Unit 1:** **Introduction:** well posed learning problem, designing a learning system, perspectives and issues in machine learning- types of learning - supervised, unsupervised and reinforcement learning **Concept learning:** concept learning task, notation, inductive learning hypothesis, concept learning as search, version space and candidate elimination algorithm, decision tree, random forest. **Linear regression:** logistic regression-Support vector machine kernel, Model selection and feature selection-Ensemble methods: Bagging, boosting, Evaluating and debugging learning algorithms. **Continuous Latent Variables:** Principal Component Analysis, Maximum variance formulation, Minimum error formulation, Applications of PCA, PCA for high-dimensional data. **Neural Networks:** -Feed-forward Network, Functions, perceptron, Weight-space symmetries, Network Training, Parameter optimization, Local quadratic approximation, Use of gradient information, Gradient descent optimization, Error Backpropagation, Evaluation of error-function derivatives, Efficiency of backpropagation.	**15 hours**

	Unit 2: **Deep learning:** Deep Feedforward Networks, Gradient-Based Learning, Hidden Units, -Architecture Design, CNN and RNN (simple RNN and LSTM). **Unsupervised learning;** Clustering, K-means, EM.Mixture of Gaussians. **Sequential Data:** Markov Models, Hidden Markov Models, Maximum likelihood for the HMM, The forward-backward algorithm, The sum-product algorithm for the HMM, Scaling	**15 hours**

	factors, -The Viterbi algorithm. **Reinforcement learning:** introduction- learning task-Q learning, non-deterministic rewards and actions-temporal difference learning.	
Pedagogy:	Lectures/ Tutorials/Hands-on assignments/Self-study/Flipped classroom	
References/ Readings	Main Reading:- 1. Alpaydin, E. (2020). Introduction to machine learning. MIT press. 2. Bishop, C. M. (2006). Pattern recognition and machine learning: springer New York.. 3. Flach, P. (2012). Machine learning: the art and science of algorithms that make sense of data. Cambridge university press. 4. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press. 5. Hart, Peter E., David G. Stork, and Richard O. Duda.(2000) Pattern classification. Hoboken: Wiley, 2000. 6. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.	
Course Outcomes	1. Develop an appreciation for what is involved in learning from data. 2. Understand a wide variety of learning algorithms. 3. Understand how to apply a variety of learning algorithms to data. 4. Understand how to perform evaluation of learning algorithms and model selection and Have a basic understanding of deep learning.	

(Back to Index)