Savitribai Phule Pune University

(Formerly University of Pune)

Two year M.Sc. Degree Program in Computer Science

(Faculty of Science & Technology)

M.Sc.- II (Computer Science)

Choice Based Credit System Syllabus

To be implemented from Academic Year

2020-2021

CSUT232 Machine Learning Total Credits - 4

Pre-requisites:

- Familiarity with Probability Theory, Multivariable Calculus, Linear Algebra
- Programming in Python (NumPy, SciPy, Pandas, Matplotlib, Seaborn, SciKit-Learn, StatsModel)

Course Objectives:

- To introduce students to the basic concepts and techniques of Machine Learning.
- To write python programs using machine learning algorithms for solving practical problems.
- To understand about Machine Learning Library and use cases.
- To understand about the process of deploying ML model.

Course Outcomes:

- Recognize the characteristics of machine learning that make it useful to real-world problems.
- Process available data using python libraries and predict outcomes using Machine Learning algorithms to solve given problem.
- Able to estimate Machine Learning models efficiency using suitable metrics.
- Design application using machine learning techniques.

Chapter No.	Topics	# Lectures
1.	**Introduction to Machine Learning**	**7**
1.1	Data Science, Artificial Intelligence and Machine Learning	
1.2	Why Learn and What is Learning, What is Machine Learning Traditional Programming Vs. Machine Learning, Machine Learning Process, Types of Data, Key Elements of Machine Learning (Representation, Evaluation and Optimization), Dimensionality Reduction (Feature Reduction)	
1.3	Descriptive and Inferential Statistics: Probability, Distribution, Distance Measures (Euclidean and Manhattan), Correlation and Regression, Hypothesis Testing.	

1.4	Creating our own dataset, Importing the dataset, Handling Missing Data, Splitting the dataset into the Training set and Test set, Feature Scaling	
2.	**Machine Learning Models**	**6**
2.1	Type of Learning- Supervised, Unsupervised and Semi-Supervised Learning	
2.2	Components of Generalization Error (Bias, Variance, underfitting, overfitting)	
2.3	A Learning System Cycle and Design Cycle	
2.4	Metrics for evaluation viz. accuracy, scalability, squared error, precision and recall, likelihood, posterior probability	
2.5	Classification Accuracy and Performance	
3.	**Regression Models**	**10**
3.1	Linear Regression - Simple , Multiple, Polynomial	
3.2	Non-linear Regression – Decision Tree, Support Vector, Random Forest	
4.	**Classification Models**	**10**
4.1	K – Nearest Neighbours (KNN)	
4.2	Logistic Regression	
4.3	Naive Bayes Theorem	
4.4	Support Vector Machine	
4.5	Decision Forest Classification	
4.6	Random Tree Classification	
5.	**Clustering Models**	**7**
5.1	K-means	
5.2	Hierarchical Clustering (Agglomerative, Divisive), Dendrogram	
5.3	Selecting optimal number of clusters: Within Clusters Sum of Squares (WCSS) by Elbow Method	
6.	**Association Rules**	**5**
6.1	Key Terms: Support, Confidence and Lift	
6.2	Apriori Algorithm	
7.	**Reinforcement Learning**	**7**
7.1	Upper Confidence Bound	
7.2	Thompson Sampling	
7.3	Q-Learning	

8. Deep Learning **8**

- **8.1** Artificial Neural Network
- **8.2** Convolution Neural Network
- **8.3** Recurring Neural Network

References:

1. Mitchell, Tom M. "Machine learning. WCB." (1997).
2. Rogers, Simon, and Mark Girolami. *A first course in machine learning*. CRC Press, 2015.
3. Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*. Vol.1. Springer, Berlin: Springer series in statistics, 2001.
4. Witten, Ian H., and Eibe Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.
5. Machine learning course material by Andrew Ng, Stanford university
6. Sutton, Richard S., and Andrew G. Barto. *Reinforcement learning: An introduction*. Vol. 1. No. 1. Cambridge: MIT press, 1998.
7. Iba, Takashi, et al. "Learning patterns: A pattern language for active learners." *Conference on Pattern Languages of Programs (PLoP)*. 2009.