Shri Vile Parle Kelavani Mandal's
DWARKADAS J. SANGHVI COLLEGE OF ENGINEERING
(Autonomous College Affiliated to the University of Mumbai)
NAAC Accredited with "A" Grade (CGPA : 3.18)

Proposed scheme for Minor in Data Science

(Academic Year 2022-2023)

Sr.	Course Code	Course	Teaching Scheme (hrs.)				Continuous Assessment (A) (marks)			Semester End Assessment (B) (marks)					(A+B)	Total Credits
			Th	P	T	Credits	Th	T/W	Total CA (A)	Th / Cb	O	P	O & P	Total SEA (B)		
Sem V																
1	DJ19MN4C1	Foundations of Data Analysis	4	--	--	4	25	--	25	75	--	--	--	75	100	**4**
Sem VI																
2	DJ19MN4C2	Machine Learning	4	--	--	4	25	--	25	75	--	--	--	75	100	**4**
	DJ19MN4L1	Machine Learning Laboratory	--	2	--	1	--	25	25	--	--	--	--	25	25	**1**
Sem VII																
3	DJ19MN4C3	Advanced Machine Learning	4	--	--	4	25	--	25	75	--	--	--	75	100	**4**
4	DJ19MN4L2	Advanced Machine Learning Laboratory	--	2	--	1	--	25	25	--	--	--	--	25	25	**1**
Sem VIII																
5	DJ19MN4C4	Big Data Engineering	4	--	--	4	25	--	25	75	--	--	--	75	100	**4**
		Total	**16**	**4**	**0**	**18**	**100**	**50**	**150**	**300**	**0**	**0**	**--**	**350**	**450**	**18**

Minor in Data Science **Semester: VI**

Program: Common for All Programs (except Computer Science and Engineering (Data Science)

Course: Machine Learning (DJ19MN4C2)

Pre-requisite:

1. Data Structures
2. Basic Probability and Statistics

Objectives:

1. To introduce the concepts of computation learning theory and techniques of Machine Learning.
2. To become familiar with regression, classification and clustering tasks.

Outcomes: On completion of the course, the learner will be able to:

1. Classify given problems into classification, clustering and regression problems
2. Apply machine learning techniques for a given problem.
3. Examine the dataset, choose appropriate algorithm and evaluate the results.
4. Design applications using machine learning algorithms.

Detailed Syllabus: (unit wise)

Unit	Description	Duration
1.	**Introduction to Machine Learning:** Types of Machine Learning, Issues in Machine Learning, Application of Machine Learning, Steps involved in developing a Machine Learning Application, Hypothesis and Inductive Bias	**04**
2.	**Regression:** Linear Regression, Least Minimum Slope (LMS) algorithm, Gradient Descent, Lasso and Ridge Regression. Polynomial Regression. Logistic Regression, Maximum Likelihood Function.	**08**
3.	**Trees:** Introduction to decision tree, Learning Decision tree using ID3 and Gini index; CART, Overfitting. Ensemble methods: Bagging (Random Forest) and Boosting (XG Boost).	**08**
4.	Classification: Bayesian Learning, Naïve Bayes, Bayesian Network: Representation in Bayesian Belief Network, Inference in Bayesian Network, Applications of Bayesian Network	**06**
5.	Introduction to Support Vector Machine: Support Vectors, Functional Margin, Geometric Margin, Optimization problem, Lagrange Duality, KKT condition, Maximum margin with noise, Non-linear SVM and Kernel Function.	**06**
6.	Clustering: K-means, Adaptive hierarchal Clustering, Gaussian Mixture Models, Expectation Maximization.	**08**

Books Recommended:

Text books:

1. Ethem Alpaydın, "Introduction to Machine Learning", MIT Press, 4th Edition, 2020
2. Peter Harrington, "Machine Learning in Action", DreamTech Press, 1st Edition, 2012
3. Tom M.Mitchell, "Machine Learning", McGraw Hill, 1997.

Reference Books:

1. Han Kamber, "Data Mining Concepts and Techniques", Morgann Kaufmann Publishers, 4th Edition, 2022.
2. Stephen Marsland, "Machine Learning An Algorithmic Perspective", CRC Press, 2nd Edition, 2014.
3. Kevin P. Murphy, "Machine Learning — A Probabilistic Perspective", MIT Press, Illustrated Edition, 2012.
4. Andreas C. Müller, Sarah Guido, "Introduction to Machine Learning with Python: A Guide for Data Scientists", O'reilly, 1st Edition, 2016.

Weblinks:

1. Towards Data Science: https://towardsdatascience.com
2. Machine Learning — Andrew Ng, Stanford University: https://youtube.com/playlist?list=PLLssT5z_DsKh9vYZkQkYNWcItqhlRJLN
3. Commonly used Machine Learning Algorithms: https://www.analyticsvidhya.com/blog/2017/09/common-machinelearning-algorithms/
4. A Tour to Machine Learning Algorithms: https://machinelearningmastery.com/a-tour-of-machine-learningalgorithms/

Suggested List of Experiments:

Students should be encouraged to write these programs from scratch to develop better understanding of the algorithms. Last 30 mins of the laboratory should be utilized as a discussion on available python libraries and hyperparameters.

Sr. No.	Title of the Experiments
1.	Perform Linear Regression.
2.	Perform Logistic Regression.
3.	Perform Decision Tree using GINI.
4.	Perform CART decision tree algorithm.
5.	Perform Ensemble methods.
6.	Perform Bayesian Classification.
7.	Perform Support Vector Machine.
8.	Perform K-means clustering.
9.	Perform Expectation -Maximization.
10.	Mini project based on any machine learning application.

Prepared by Checked by Head of the Department Principal