Program: PGPDS		**Trimester:**	I/II/III
Course:	Advance Machine Learning (ML2- AML)	Credits	3.00 / 2.00 / 1.50
		Hours	30
Faculty:	Prof. Gourab Nath (Bangalore) Prof. Subhasis Dasgupta (Kolkata)	Batch:	Jan / July
		Campus:	Kolkata; Bangalore

Teaching Scheme				***Evaluation Scheme***	
Weekly Class Discussions	Weekly Practical Workshops	Weekly Tutorials Sessions	Total Program Weeks	Internal Continuous Assessment (ICA) **Marks = 70**	Term End Examination (TEE) **Marks = 30**
				Marks Scaled to 70	**Marks Scaled to 30**

Design Philosophy	
Course Rationale	*The Course has been designed with the following Rationale:*
	To introduce the students to solving real-life problems of higher complexities.
Course Objectives	*The course has been offered by the Instructor to satisfy the following objectives:*
1	To make students understand the advanced machine learning algorithms
2	To make students understand the model building processes using python
3	To infer the outcomes from the machine learning models wherever possible.
Learning Outcomes	*At the end of the course, a student shall:*
1	Should be able apply Machine Learning in large-scale multidimensional data problems
2	Get a clear understanding of how to evaluate machine learning models
3	
Prerequisite(s):	
1	Working knowledge of Python programming
2	
3	
Pedagogy:	
1	Classroom interactions covering theoretical & practical aspects of Data Mining and Machine Leaning
2	Practical training using student's own systems
3	Class tests and quizzes to reinforce student's learnings

Learning Compendium

	Textbook(s):	Publisher
1	Data Mining & Predictive Analytics by Larose & Larose	Wiley
2		

	Reference Books:	Publisher
1	Data Mining Concepts and Techniques by Han, Kamber & Pei	Morgan Kaufmann
2	Data Mining and Analysis – Fundamental Concepts and Algorithms by Zaki & Meira	Cambridge University Press
3	Elements of Statistical Learning	Springer
4		

	Journal Article & Research Papers
1	
2	
3	
4	
5	

	Websites	Topics
1		
2		
3		
4		
5		

	Videos:	Topics
1		
2		
3		

Tentative Session Plan

Session	Topic (including subtopics)	Learning Outcomes	Pedagogy	Case Study / Readings / Group Assignment
Beginning Module 1	**Module 1: Ensembles & Dimension Reduction**			
1	Ensemble Modeling (Bagging and Ada-Boosting)	• Bagging concepts • Boosting concepts • Python based modeling Model Deployment	Class Discussion and Hands-On	

2	Ensemble Modeling (Random Forest)	• Basic concepts • Why it works • Model building Model deployment	Class Discussion and Hands-On	
3	Ensemble Modeling (GBM)	• Basic concepts • Why it works • Model building Model deployment	Class Discussion & Hands-On	
4	Ensemble Modeling (Xgboost)	• Taylor expansion • Working principles • Python based modeling Model deployment	Class Discussion and Hands-On	
5	Dimension Reduction (PCA)	• Eigenvalue • Eigenvectors • Python based modeling	Class Discussion and Hands-On	
6	Dimension Reduction (SVD)	• Eigenvalue • Eigenvectors • Singular values • Python based modeling	Class Discussion and Hands-On	
Evaluation				
Module 1 Notes:				
Beginning Module 2	**Module 2: Kernel based techniques**			
7	Kernel and Kernel spaces	• Basic concept • Simple calculations in Kernel Space • Basic python based demo	Class Discussion and Hands-On	
8	Support Vector Machines (SVM)	• Margin • Objective function • Optimization • Python based modeling	Class Discussion and Hands-On	
9	Manifold Learning (t-SNE)	• Dimension embedding • K-L Divergence Python based modeling	Class Discussion and Hands-On	
10	Evolutionary Search	• Genetic algorithm • Use of GA in parameter tuning Python based example	Class Discussion and Hands-On	
11	Hands-on exercise	Comprehensive revision	Lab Session	
Evaluation				
Module 2 Notes:				
Beginning Module 3:	**Module 3: Advanced Clustering**			
12	Two stage clustering	• Combining K-Means and Hierarchical clustering	Class Discussion and Hands-On	

		• Python implementation		
13	DBSCAN	• Density chain • Epsilon neighborhood • Python based modeling	Class Discussion and Hands-On	
14	Maximum Likelihood Estimation	• Likelihood function • Maximization of likelihood function • Use of MLE in simple regression Python implementation	Class Discussion and Hands-On	
15	EM clustering	• Likelihood function • Maximization of likelihood function • Use of MLE in simple regression Python implementation	Class Discussion and Hands-On	
16	Hands-On exercise		Lab Session	
Evaluation				
Module 3 Notes:				
Beginning Module 4:	**Module 4: Recommendation Systems**			
17	Matrix Factorization	• UV decomposition • Usage in Recommendation system Python implementation	Class Discussion and Hands-On	
18	KNN Based Recommendation System	• UBCF • IBCF Python implementation	Class Discussion and Hands-On	
19	Class Test		Hands-on	
20		Class Test	Discussion	
Evaluation				
Module 4 Notes:				

Evaluation Scheme			
Sl No.	Component	Evaluation timeframe & Methodology	Weightage (%)
1	**Class Participation**		10

2	**Class Participation / Quiz**		30
3	**Mid-Term Examination**		30
4	**End-Term Examination**		30
TOTAL			100
Evaluation Notes:			

Signature
(Prepared by Concerned Faculty/HOD)

Signature
(Approved by Director)