Birla Institute of Technology & Science, Pilani
Work Integrated Learning Programmes Division
Second Semester 2024-2025

Digital Learning Handout

Part A: Content Design

Course Title	Machine Learning
Course No(s)	AIMLCZG565
Credit Units	4
Credit Model	3-1-0
Course Author	Sugata Ghosal
Version No:	2.0
Date:	26/02/2025

Course Description:

Introduction to Machine Learning, Various kinds of learning, Supervised Learning, Unsupervised Learning, Model Selection; Bayesian Learning, MAP Hypothesis, MDL Principle, Bias Variance Decomposition, Bayes Optimal Classifier, Naive Bayes Classifier; Linear Models for Regression, Linear Models for Classification; Non-Linear models, Decision trees; Instance Based Learning, KNN Algorithm, CBR Learning; Support Vector Machines, VC Dimension; Neural Networks, Perceptron Learning, Back Propagation Algorithm; Introduction to Genetic Algorithms.

Course Objectives

No	Course Objective
CO1	Introduce students to the basic concepts and techniques of Machine Learning.
CO2	To gain experience of doing independent study and research in the field of Machine Learning
CO3	To develop skills of using recent machine learning software tools to evaluate learning algorithms and model selection for solving practical problems

Text Book(s):

T1	Tom M. Mitchell, Machine Learning, The McGraw-Hill Companies, Inc. Indian Edition 1997

Reference Book(s) & other resources:

R1	Christopher M. Bishop, Pattern Recognition & Machine Learning, Springer, 2006
R2	PANG-NING TAN, MICHAEL STEINBACH, VIPIN KUMAR, Introduction To Data Mining, Pearson, 2nd Edition.
	CHRISTOPHER J.C. BURGES: A Tutorial on Support Vector Machines for Pattern Recognition, Kluwer Academic Publishers, Boston, pp. 1–43.

Learning Outcomes: Students will be able to

LO1	A strong understanding of the foundations of Machine Learning algorithms
LO2	Able to solve Machine Learning problems using appropriate learning techniques
LO3	Evaluate machine learning solutions to problems
LO4	Identify appropriate tools to implement the solutions to machine learning problems

Modular Content Structure

1. **Introduction**
   - Introduction to ML
   - Objective of the course
   - Taxonomy (types) of Machine Learning
   - Design a Learning System
   - Challenges in Machine Learning
2. **Machine learning Workflow**
   - Role of Data
   - Data Preprocessing, wrangling
   - Data skewness removal (sampling)
   - Model Training
   - Model Testing and performance metrics
3. **Linear models for Regression**
   - Direct Solution Method
   - Iterative Method – Gradient Descent (batch/stochastic/mini-batch)
   - Linear basis function models
   - Bias-variance decomposition
4. **Linear models for classification**
   - Discriminant Functions
   - Decision Theory
   - Probabilistic Discriminative Classifiers
   - Logistic Regression
5. **Decision Tree**
   - Information Theory
   - Entropy Based Decision Tree Construction
   - Avoiding Overfitting
   - Minimum Description Length
   - Handling Continuous valued attributes, missing attributes
6. **Instance-based Learning**
   - k-Nearest Neighbor Learning
   - Locally Weighted Regression (LWR) Learning
   - Radial Basis Functions
7. **Support Vector Machines**
   - Linearly separable data
   - Non-linearly separable data

- Kernel Trick (Mercer)
- Applications to both structured and unstructured data

8. **Bayesian Learning**
   - MLE Hypothesis
   - MAP Hypothesis
   - Bayes Rule
   - Optimal Bayes Classifier
   - Naïve Bayes Classifier
   - Probabilistic Generative Classifiers
   - Bayesian Linear Regression
9. **Ensemble Learning**
   - Combining Classifiers
   - Bagging
   - Random Forest
   - Boosting
     - o ADABoost
     - o Gradient Boosting
     - o XGBoost
10. **Unsupervised Learning**
    - K-means Clustering and variants
    - Review of EM algorithm
    - GMM based Soft Clustering
    - Applications
11. **Machine Learning Model Evaluation/Comparison**
    - Comparing Machine Learning Models
    - Emerging requirements e.g., bias, fairness, interpretability of ML models

Part B: Learning Plan

Contact Session	List of Topic Title	Sub-Topics	Reference
1	**Introduction**	Objective, what is Machine Learning? Application areas of Machine Learning, Why Machine Learning is important? Design a Learning System, Issues in Machine Learning	T1 – Ch1
2	**Mathematical Preliminaries** In M. Tech. AIML **Machine learning Workflow**	Linear Algebra, Calculus, Probability theory, Decision Theory, Information Theory In M. Tech. AIML Role of Data, Data Pre-processing, wrangling, Data skewness removal (sampling), Model Training, Model Testing and performance metrics	R2 – Ch2, Ch3 Lecture Notes

3	**Linear models for Regression**	Direct Solution Method, Iterative Method – Gradient Descent (batch/stochastic/mini-batch), Linear basis function models, Bias-variance decomposition	R1 - Ch3
4	**Linear models for classification**	Discriminant Functions, Decision Theory, Probabilistic Discriminative Classifiers, Introduction to Logistic Regression	R1 - Ch. 3, 4
5	**In Logistic Regression**	Logloss Function, Gradient Descent, multi-class classification	R1 - Ch. 4 R2 – Ch. 4
6	**Decision Tree**	Information Theory, Entropy Based Decision Tree Construction, Avoiding Overfitting, Minimum Description Length, Handling Continuous valued attributes, missing attributes	T1 – Ch. 3 R2 - Ch. 3
7	**Instance-based Learning**	k-Nearest Neighbor Learning, Locally Weighted Regression (LWR) Learning, Radial Basis Functions	T1 – Ch. 8
8	**Review**	Review of Session 1 to 7	Books, Web references and Slides
9	**Support Vector Machine**	Linearly separable data, Non-linearly separable data, Kernel Trick (Mercer), Applications to both structured and unstructured data	R2 - Ch. 4 R3
10	**Bayesian Learning**	MLE Hypothesis, MAP Hypothesis, Bayes Rule, Optimal Bayes Classifier	T1 - Ch. 6 R2 – Ch. 4
11	**Bayesian Learning**	Naïve Bayes Classifier, Probabilistic Generative Classifiers, Bayesian interpretation of Linear Regression	T1 - Ch. 6 R2 – Ch. 4 R1 – Ch. 4
12	**Ensemble Learning**	Combining Classifiers, Bagging, Random Forest, Boosting	R2 – Ch. 4
13	**Ensemble Learning**	ADABoost, Gradient Boosting, XGBoost	R2 – Ch. 4 Lecture Notes
14	**Unsupervised Learning**	K-Means clustering, Mixture Models for probabilistic clustering, Review of EM algorithm, Applications	T1 – Ch. 6

15	**ML model evaluation**	Comparing Machine Learning Models, Emerging requirements e.g., bias, fairness, interpretability of ML models	T1 - Ch. 5 Lecture Notes
16	**Review**	Review of session 9 to 15	Books, Web references and Slides

Experiential Learning Components:

Describe objective, outcome of Experiential Learning Component and the lab infrastructure needed (virtual, remote, open source etc...) number of lab exercises needed, etc.

1. Lab work: 5
2. Project work: 0
3. Case Study: 4 Webinars
4. Simulation: 0
5. Work Integrated Learning Assignment- 2 Assignments
6. Design work/ Field work: 0

Objective of Experiential Learning Component:
Hands on sessions on implementation of fundamental machine learning algorithms using state of art tools
Scope of Experiential Learning Component:
Programming language - Python
Tools and libraries: Jupyter, ScikitLearn, etc.
Lab Infrastructure:
Online/ Open source/Google Colab
List of Experiments:

Lab No	Lab Objective	Session Reference
1	End to End Machine Learning	2
2	Linear Regression and Gradient Descent	3, 4
3	Logistic Regression classifier	5
4	Decision Tree and Random Forest	6, 12
5	Naïve Bayes Classification	11

Evaluation Scheme:
Legend: EC = Evaluation Component; AN = After Noon Session; FN = Fore Noon Session

Evaluation Component	Name (Quiz, Lab, Project, Mid-term exam, End semester exam, etc.)	Type (Open book, Closed book, Online, etc.)	Weigh t	Duration	Day, Date, Session, Time
EC – 1*	Quiz	Online	10%	1 week	To be announced

	Assignment/Lab Assignment / Lab Exams	Online	20 %	10 days	To be announced
EC - 2	Mid-Semester Test	Closed Book	30%	2 hours	To be announced
EC - 3	Comprehensive Exam	Open Book	40%	2 ½ Hours	To be announced

EC1* (20% - 30%): Quiz (optional): 5-10 %, Lab Assignment/Assignment: 20% - 30%
Syllabus for Mid-Semester Test (Closed Book): Topics in Contact session: 1 to 8
Syllabus for Comprehensive Exam (Open Book): All topics

Important Links and Information:

eLearn Portal: https://elearn.bits-pilani.ac.in

Students must visit the eLearn portal regularly and stay updated with the latest announcements and deadlines.

Contact Sessions: Students should attend the online lectures as per the schedule provided on the eLearn portal.

Evaluation Guidelines:

1. EC-1 consists of either two Assignments or three Quizzes. Students will attempt them through the course pages on the eLearn portal. Announcements will be made on the portal in a timely manner.
2. For Closed Book tests: No books or reference material of any kind will be permitted.
3. For Open Book exams: "open book" means text/ reference books (publisher copy only) and does not include any other learning material. No other learning material will be permitted during the open book examinations. For Detailed Guidelines refer to the attached document.
   EC3 Guidelines
4. If a student is unable to appear for the Regular Test/Exam due to genuine exigencies, the student should follow the procedure to apply for the Make-Up Test/Exam, which will be made available on the eLearn portal. The Make-Up Test/Exam will be conducted only at selected exam centres on the dates to be announced later.

It shall be the responsibility of the individual student to be regular in maintaining the self-study schedule as given in the course handout, attend the online lectures, and take all the prescribed evaluation components such as Assignments/Quizzes, Mid-Semester Tests and Comprehensive Exams according to the evaluation scheme provided in the handout.

****