



Course Syllabus
Gyanmanjari College of Computer Application
Semester-2 (MCA)

Subject: Machine Learning - MCAAI11512

Type of course: Minor Stream

Prerequisite: Python, Mathematical Concepts.

Rationale:

Machine Learning (ML) has gained significant traction in recent years due to its capacity to revolutionize various sectors. By automating routine tasks, ML enhances efficiency and productivity, allowing human resources to focus on more complex endeavors. Its ability to analyze vast datasets aids in informed decision-making, while personalized recommendations based on customer data boost satisfaction and loyalty.

Additionally, ML contributes to cost savings by automating processes and reducing the reliance on human labor. Beyond economic benefits, ML fuels innovation in fields like healthcare and finance, promising to reshape industries and improve lives on a global scale.

Teaching and Examination Scheme:

Teaching Scheme			Credits	Examination Marks					Total Marks
CI	T	P		Theory Marks		Practical Marks		CA	
			ESE	MSE	V	P	ALA		
3	0	2	4	60	30	10	20	30	150

Legends: CI-Class Room Instructions; T – Tutorial; P - Practical; C – Credit; ESE - End Semester Examination; MSE- Mid Semester Examination; V – Viva; CA - Continuous Assessment; ALA- Active Learning Activities.

Course Content:

Sr. No	Course content	Hrs	% Weightage
1	Introduction to Machine Learning: <ul style="list-style-type: none"> What is human learning? What is Machine Learning? 	5	15



	<ul style="list-style-type: none"> • Human learning versus machine learning • Types of machine learning • Applications of machine learning • Tools for machine learning. 		
2	Preparing to Model , Feature Engineering: <ul style="list-style-type: none"> • Machine Learning activities • Basic Types of data in Machine Learning • Structures of data • Data Quality and Remediation • Data Pre-Processing • Introduction to Feature Engineering • Feature Transformation • Feature Subset Selection 	9	20
3	Modeling and Evaluation: <ul style="list-style-type: none"> • Selecting a Model • Training a Model • Model Representation and Interpretability • Evaluating Performance of a Model • Improving Performance of a Model 	8	15
4	Supervised Learning: Classification and Regression: <ul style="list-style-type: none"> • Example of Supervised Learning • Classification Model • Classification Learning Steps • Common Classification Algorithms (k-Nearest Neighbor (kNN), Decision tree) • Example of Regression • Common regression Algorithms (Simple linear regression, Multiple linear regression) 	12	25
5	Unsupervised Learning: <ul style="list-style-type: none"> • Unsupervised Learning versus Supervised Learning • Applications of Unsupervised Learning Clustering(Partitioning methods(k-means)) • Finding Pattern using Association Rules 	11	25

Continuous Assessment:

Sr. No	Active Learning Activities	Marks
1	Study Synopsis : Students have to explore different industries, domains, and contexts where machine learning is applied to solve problems. Make predictions, automate tasks, or gain insights from data. Students have to prepare document file and upload it on GMIU web portal. Perform this activity in group (Maximum 4 students in one group).	10



2	Idea Revolve: Students have to identify potential domains or contexts where machine learning is not currently applied but could offer valuable solutions or improvements. Students have to prepare document file and upload it on GMIU web portal.	10
3	Review Paper: Students have to prepare review paper for machine learning and prepare PDF and upload it on GMIU web portal. Students have to perform this activity in group(Maximum 4 students in one group)	10
Total		30

Suggested Specification table with Marks (Theory):60

Distribution of Theory Marks (Revised Bloom’s Taxonomy)						
Level	Remembrance (R)	Understanding (U)	Application (A)	Analyze (N)	Evaluate (E)	Create (C)
Weightage	25%	30%	25%	10%	10%	0

Note: This specification table shall be treated as a general guideline for students and teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Course Outcome:

After learning the course the students should be able to:	
CO1	Understand basic concept of machine learning and its applications.
CO2	Identify data and perform pre-processing on data. Data is in a format that the algorithm can understand and that it is free of errors or outliers that can negatively impact the model's performance
CO3	Understand features and uncover the hidden patterns in the data and boosts the predictive power of machine learning.
CO4	Analyze for a variety of real-world problems at scale, they should to discover hidden and interesting patterns in unlabeled data
CO5	Navigate practical domains such as classification, decision tree, clustering, acquiring the ability to design and implement ML solutions and communicate their results effectively.



List of Practical

Sr. No.	Descriptions	Unit No.	Hrs																		
1	Introduction to Jupyter Notebook	1	2																		
2	Write a python code to Perform following task 1. Installing Libraries. 2. Loading the dataset. 3. Summarizing the dataset. 4. Visualizing the dataset. 5. Evaluating some algorithms. 6. Making some predictions.	2	4																		
3	Write a python code to Perform data preprocessing on following csv file <table border="1" style="margin-left: 20px;"> <thead> <tr> <th>Name</th> <th>Age</th> <th>Gender</th> </tr> </thead> <tbody> <tr> <td>Alice</td> <td>25</td> <td>Female</td> </tr> <tr> <td>Bony</td> <td>30</td> <td>Male</td> </tr> <tr> <td>Carol</td> <td></td> <td>Female</td> </tr> <tr> <td>David</td> <td>22</td> <td>Male</td> </tr> <tr> <td>Eve</td> <td>28</td> <td></td> </tr> </tbody> </table>	Name	Age	Gender	Alice	25	Female	Bony	30	Male	Carol		Female	David	22	Male	Eve	28		2	2
Name	Age	Gender																			
Alice	25	Female																			
Bony	30	Male																			
Carol		Female																			
David	22	Male																			
Eve	28																				
4	Write a python code to perform feature engineering on following data set and add new column "Income per Age" <table border="1" style="margin-left: 20px;"> <thead> <tr> <th>Age</th> <th>Income</th> </tr> </thead> <tbody> <tr> <td>30</td> <td>50000</td> </tr> <tr> <td>25</td> <td>60000</td> </tr> <tr> <td>35</td> <td>75000</td> </tr> <tr> <td>22</td> <td>45000</td> </tr> </tbody> </table>	Age	Income	30	50000	25	60000	35	75000	22	45000	2	2								
Age	Income																				
30	50000																				
25	60000																				
35	75000																				
22	45000																				
5.	Write a python code to perform feature engineering on following data set and convert categorical data(gender) to numeric data <table border="1" style="margin-left: 20px;"> <thead> <tr> <th>Gender</th> <th>Age</th> <th>Income</th> </tr> </thead> <tbody> <tr> <td>Male</td> <td>30</td> <td>50000</td> </tr> <tr> <td>Female</td> <td>25</td> <td>60000</td> </tr> <tr> <td>Male</td> <td>35</td> <td>75000</td> </tr> <tr> <td>Female</td> <td>22</td> <td>45000</td> </tr> </tbody> </table>	Gender	Age	Income	Male	30	50000	Female	25	60000	Male	35	75000	Female	22	45000	2	2			
Gender	Age	Income																			
Male	30	50000																			
Female	25	60000																			
Male	35	75000																			
Female	22	45000																			
6.	Write a python code to perform Preparing to Model for the following dataset <table border="1" style="margin-left: 20px;"> <thead> <tr> <th>Age</th> <th>Income</th> </tr> </thead> <tbody> <tr> <td>30</td> <td>50000</td> </tr> <tr> <td>25</td> <td>60000</td> </tr> <tr> <td>35</td> <td>75000</td> </tr> <tr> <td>22</td> <td>45000</td> </tr> </tbody> </table>	Age	Income	30	50000	25	60000	35	75000	22	45000	3	2								
Age	Income																				
30	50000																				
25	60000																				
35	75000																				
22	45000																				
7.	Write a python code to perform k-nearest neighbors (KNN) on popular Iris dataset	4	2																		
8.	Write a python code to implement decision tree for below given dataset. Identify the root node and all subpart or children of node and draw the tree.	4	2																		



Item no	Age	Income	Student	Credit Rating	Buys-Computer
1	Youth	High	No	Fair	No
2	Youth	High	No	Excellent	No
3	Middle	High	No	Fair	Yes
4	Senior	Medium	No	Fair	Yes
5	Senior	Low	Yes	Fair	Yes
6	Middle	Low	Yes	Excellent	No
7	Senior	Low	Yes	Excellent	Yes
8	Youth	Medium	No	Fair	No
9	Youth	Low	Yes	Fair	Yes
10	Senior	Medium	Yes	Fair	Yes
11	Youth	Medium	Yes	Excellent	Yes
12	Middle	Medium	No	Excellent	Yes
13	Middle	High	Yes	Fair	Yes
14	Senior	Medium	No	Excellent	No

9	Build a prediction model using regression technique for following dataset.			4	2
	X	Y			
	10	11			
	11	13			
	12	12			
	13	15			
	14	17			
	15	18			
	16	18			
	17	19			
	18	20			
	19	22			

10.	Build a prediction model using regression technique for (1) Boston house-prices (from sklearn.dataset import load_boston) (2) Diabetes (from sklearn.dataset import load_diabetes) datasets. Also, evaluate the model.	4	4																								
11.	Implement unsupervised machine learning algorithm (Clustering – K Means) in python on Breast Tumour dataset to cluster data (use Breast Tumour dataset) by removing the class label.	5	2																								
12	Apply k-means clustering approach with k = 2 to the following dataset.	5	2																								
	<table border="1"> <thead> <tr> <th>X</th> <th>Y</th> <th>Z</th> </tr> </thead> <tbody> <tr><td>-0.154</td><td>0.376</td><td>0.099</td></tr> <tr><td>-0.103</td><td>0.476</td><td>-0.027</td></tr> <tr><td>0.228</td><td>0.036</td><td>-0.251</td></tr> <tr><td>0.33</td><td>0.013</td><td>-0.251</td></tr> <tr><td>-0.114</td><td>0.482</td><td>0.014</td></tr> <tr><td>0.295</td><td>0.084</td><td>-0.297</td></tr> <tr><td>0.262</td><td>0.042</td><td>-0.304</td></tr> </tbody> </table>			X	Y	Z	-0.154	0.376	0.099	-0.103	0.476	-0.027	0.228	0.036	-0.251	0.33	0.013	-0.251	-0.114	0.482	0.014	0.295	0.084	-0.297	0.262	0.042	-0.304
X	Y			Z																							
-0.154	0.376			0.099																							
-0.103	0.476			-0.027																							
0.228	0.036			-0.251																							
0.33	0.013			-0.251																							
-0.114	0.482			0.014																							
0.295	0.084	-0.297																									
0.262	0.042	-0.304																									



13	Write a python code to implement Apriori algorithm , apply join and prune method and find frequent item set. ['milk', 'bread', 'nuts', 'apple'], ['milk', 'bread', 'nuts'], ['milk', 'bread'], ['milk', 'bread', 'apple'], ['bread', 'nuts', 'apple']	5	2
----	---	---	---

Instructional Method:

The course delivery method will depend upon the requirement of content and need of students. The teacher in addition to conventional teaching method by black board, may also use any of tools such as demonstration, role play, Quiz, brainstorming, MOOCs etc.

Students will use supplementary resources such as online videos, NPTEL/SWAYAM videos, e-courses, Virtual Laboratory.

The internal evaluation will be done on the basis of Active Learning Assignment.

Practical/Viva examination will be conducted at the end of semester for evaluation of performance of students in laboratory.

Reference Books:

- [1] Machine Learning by Saikat Dutt, Subramanian Chandramouli, Amit Kumar Das
- [2] Introduction of Machine Learning: V.D.Patel, J.B.Patel, P.J.Joshi, Atul Prakashan
- [3] Machine Learning: Tom M Mitchell
- [4] Machine Learning: Anuradha Srinivasaraghavan, Vincy Joseph
- [5] Machine Learning in Action: Peter Harrington

