Program Name : **Diploma in Artificial Intelligence and Machine Learning**
Program Code : **AN**
Semester : **Sixth**
Course Title : **Advanced Algorithms in AI & ML**
Course Code : **22683**

1. RATIONALE

Machine Learning is a branch of Computer Science that uses algorithms to imitate the way in which humans learn. It uses statistical methods to train algorithms and make predictions. Machine learning is one of the most in-demand Data Science skills, which allows data scientists to increase the accuracy of predictions of software applications, without explicitly programming them to do so. These algorithms make use of historical data to predict output values and these insights and predictions enable businesses to make smart decisions

2. COMPETENCY

The aim of this course is to help the student to attain the following ***industry identified*** competency through various teaching learning experiences:

- **Implement AI-ML algorithm using Python.**

3. COURSE OUTCOMES (COs)

The theory, practical experiences and relevant soft skills associated with this course are to be taught and implemented, so that the student demonstrates the following industry oriented COs associated with the above mentioned competency:

a. Implement preprocessing steps on data to make it ready for analysis
b. Implement Naive Bayes and Random Forest algorithm.
c. Describe Support Vector Machines and K Nearest Neighbors.
d. Apply clustering algorithm and Dimensionality Reduction.
e. Describe ANN and basic hyper parameters of Deep Learning.
f. Classify Sequential and Image Data of Deep Learning.

4. TEACHING AND EXAMINATION SCHEME

Teaching Scheme (In Hours)				Examination Scheme												
				Theory						Practical						
L	T	P	Credits (L+T+P)	Paper Hrs.	ESE		PA		Total		ESE		PA		Total	
					Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
3	0	2	5	3	70	28	30*	0	100	40	25@	10	25	10	50	20

(**) *marks should be awarded on the basis of internal end semester theory exam of 50 marks based on the specification table given in S. No. 9.*

(~²): *For the **practical only courses,** the PA has two components under practical marks i.e. the assessment of practicals (seen in section 6) has a weightage of 60% (i.e. 30 marks) and micro-project assessment (seen in section 12) has a weightage of 40% (i.e. 20 marks). This is designed to facilitate attainment of COs holistically, as there is no theory ESE.*

Legends: L-Lecture; **T** – Tutorial/Teacher Guided Theory Practice; **P** -Practical; **C** – Credit, **ESE** -End Semester Examination; **PA** - Progressive Assessment, **'#':** No Theory Examination

5. COURSE MAP (with sample COs, Learning Outcomes i.e. LOs and topics)

This course map illustrates an overview of the flow and linkages of the topics at various levels of outcomes (details in subsequent sections) to be attained by the student by the end of the course, in all domains of learning in terms of the industry/employer identified competency depicted at the centre of this map.

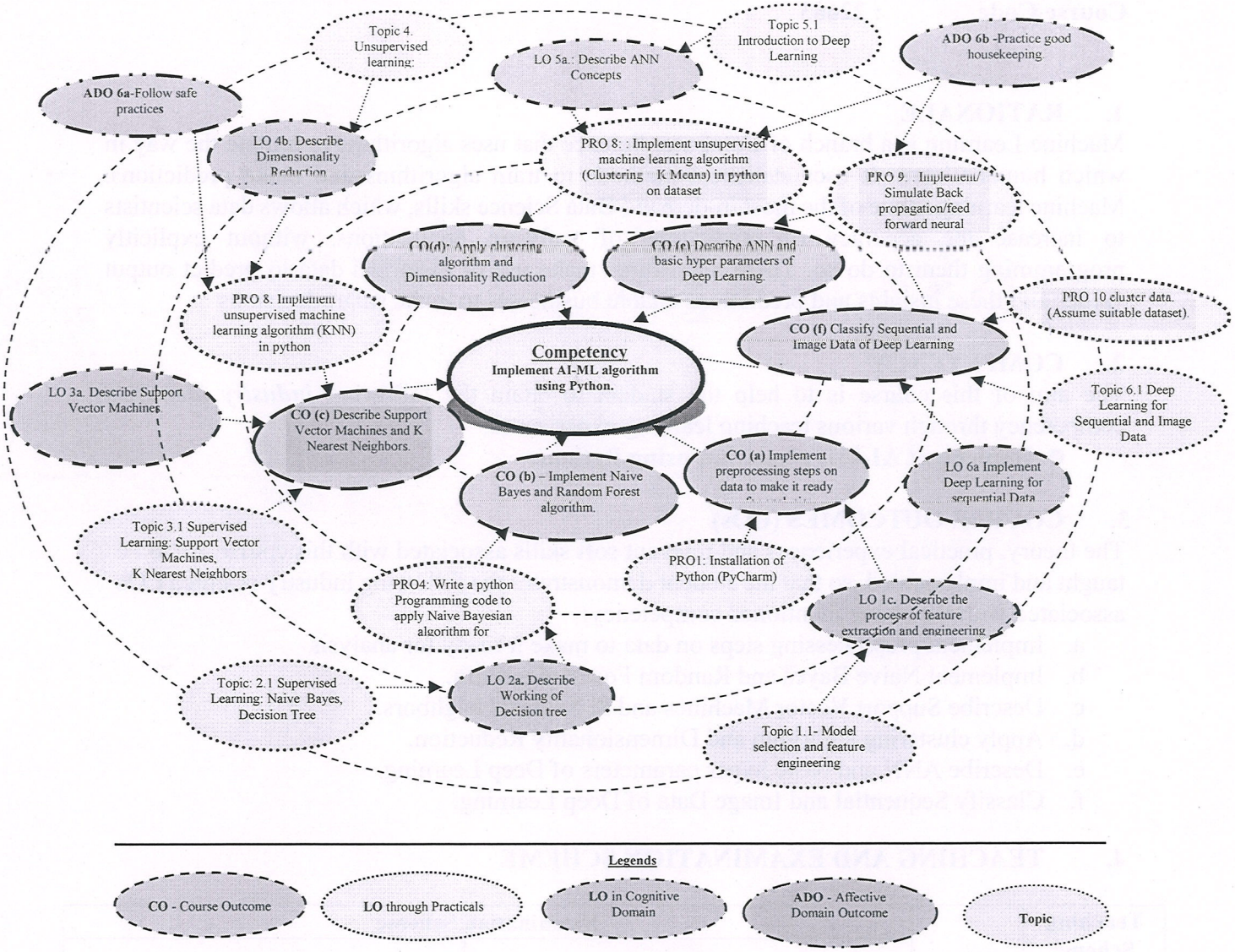

Figure 1 - Course Map

6. SUGGESTED PRACTICALS/ EXERCISES

The practicals/exercises/tutorials in this section are psychomotor domain LOs (i.e. sub-components of the COs) are to be developed and assessed in the student to lead to the attainment of the competency.

Sr. No.	Practical Exercises (Learning Outcomes to be achieved through practicals)	Unit No.	Approx. Hrs. Required
1	a. Installation of Python (PyCharm) b. Installation of Python scikit learn for ML c. Installation of Tools and Libraries (Jupyter Notebook/Matplotlib/Numpy/Pandas) d. Use of google colab (https://colab.research.google.com/)	I	02*

Sr. No.	Practical Exercises (Learning Outcomes to be achieved through practicals)	Unit No.	Approx. Hrs. Required
2	Study Different datasets such as Iris dataset ,Titanic dataset,Imdb Movies dataset	I	02
3	Perform following operations :(Assume suitable data/dataset if needed). a. Write program to read dataset (Text,CSV,JSON,XML) b. Which of the attributes are numeric and which are categorical? c. Performing Data Cleaning, Handling Missing Data, Removing Null data d. Rescaling Data v. Encoding Data e. Feature Selection	I	04*
4	Write a python Programming code to apply Naive Bayesian algorithm for classification using suitable data/dataset	II	04*
5	Write a python Programming code to implement decision tree for classification using suitable data/dataset.	II	02*
6	Implement the Random Forest Algorithm using following Steps a. Data Preprocessing Step b. Fitting the Random Forest Algorithm to the Training Set c. Predicting the Test Set Result d. Creating the confusion Matrix e. Visualizing the training set result f. Visualizing the test set result	II	04
7	Implement unsupervised machine learning algorithm (KNN) in python on dataset to cluster data. (Assume suitable dataset)	III	02
8	Implement unsupervised machine learning algorithm (Clustering – K Means) in python on dataset to cluster data. (Assume suitable dataset).	IV	04*
9	Implement/Simulate Back propagation/feed forward neural network.	V	04
10	Study of Classification of Dog images and Cat images categories using suitable dataset (foe ex imagenet dataset)	VI	04
	Total		**32**

'': compulsory practicals to be performed.*

Note

i. *Given in above tables is suggestive list of practical exercises. Teachers can design other similar exercises.*

ii. *Assessment of the 'Process' and 'Product' related skills in the laboratory/workshop/field work should be done as per suggested sample below:*

S.No.	Performance Indicators	Weightage in %
1	Import packages and Libraries of Python.	20
2	Use Python to create, edit, assemble and link the programs.	40
3	Debug, test and execute the programs	20
4	Able to answer oral questions.	10
5	Submission of report in time.	10
	Total	**100**

Additionally, the following affective domain LOs (social skills/attitudes), are also important constituents of the competency which can be best developed through the above mentioned laboratory/field based experiences:

a. Handle command prompt environment.
b. Experiment with Python
c. Plan, develop, assemble, link, debug and test the programs.
d. Demonstrate working as a leader/a team member.
e. Maintain tools and equipment.
f. Follow ethical practices.

The development of the attitude related LOs of Krathwohl's 'Affective Domain Taxonomy', the achievement level may reach:

- 'Valuing Level' in 1st year
- 'Organizing Level' in 2nd year and
- 'Characterizing Level' in 3rd year.

7. MAJOR EQUIPMENT/ INSTRUMENTS REQUIRED

The major equipment with broad specification mentioned here will usher in uniformity in conduct of experiments, as well as aid to procure equipment by authorities concerned.

S. No.	Equipment Name with Broad Specifications	PrO. Sr. No.
1	Hardware: Personal computer, (i3 preferable), RAM minimum 4 GB onwards.	For all Experiments
2	Operating system: Windows 7 onward	
3	Libraries of Python	

8. UNDERPINNING THEORY COMPONENTS

The following topics/subtopics should be taught and assessed in order to develop LOs in cognitive domain for achieving the COs to attain the identified competency.

Unit	Major Learning Outcomes (in cognitive domain)	Topics and Sub-topics
Unit I Model selection and feature engineering	1a. Select a suitable model for the given data with justification. 1b. Describe the process of using supervised learning on the given data. 1c. Describe the process of feature extraction and engineering on the given data. 1d. Compare Feature Engineering for the given type of data. 1e. Select feature scaling, feature selection, dimensionality reduction in the given situation with justification.	1.1. Introduction: Selecting a model 1.2. Training a model for supervised learning Features – understand your data better, Feature extraction and engineering, 1.3. Feature engineering on – numerical data & categorical data & text data 1.4. Feature scaling, Feature selection
Unit-II	2a. Classify the given data	2.1 Naive Bayes

Unit	Major Learning Outcomes (in cognitive domain)	Topics and Sub-topics
Supervised Learning: Naive Bayes, Decision Tree	using Bayesian method with stepwise justification. 2b. Describe Working of Decision tree algorithm 2c. Enlist application of Random Forest Algorithm.	• Bayes Theorem, Working of Naive Bayes • Bayes classifier, Applying Bayes Theorem, • Advantages and Disadvantages of Naive Bayes classifier, • Application of Naive Bayes, • Implementation of Naive Bayes classifier 2.2 Decision Tree • Decision Tree Diagram, Why Used decision tree? • Working of Decision tree algorithm, attributes selection Measures (ASM), • Advantages and Disadvantages of Decision tree, • Implementation of Decision Tree 2.3 Random Forest • Why use Random Forest? • Working of Random Forest Algorithm, • Application of Random Forest Algorithm, • Advantages and Disadvantages of Random Forest algorithm, • Implementation of Random Forest algorithm
Unit-III Supervised Learning: Support Vector Machines, K Nearest Neighbors	3a. Describe Support Vector Machines. 3b. Enlist advantages and disadvantages of KNN algorithm.	3.1 Support Vector Machines: • Types of SVM, • How does SVM work? • Advantages and Disadvantages of Decision tree, • Implementation of SVM 3.2 K Nearest Neighbors: • Need of KNN algorithm, • Working of KNN Algorithm, • Advantages and Disadvantages of KNN algorithm, • Implementation of KNN algorithm
Unit– IV Unsupervised learning: Clustering Algorithms	4a. Describe the performance analysis of clustering for the given situation. 4b. Describe Dimensionality Reduction	4.1 K-Means Clustering: • What is K-means Clustering? • Working of K-means Algorithm, • Failure of K-means algorithm, • Implementation of K-means

Unit	Major Learning Outcomes (in cognitive domain)	Topics and Sub-topics
		algorithm 4.2 Dimensionality Reduction: • Introduction to Dimensionality Reduction, Subset Selection, • Introduction to Principal Component Analysis.
Unit V Introduction to Deep Learning	5a. Describe ANN Concepts 5b. Learn Hyper parameter basics.	5.1 Introduction • Artificial Neural Network, • Perceptron EX-OR problem, • Feed Forward and Back Propagation, Losses • Activation Function, GPU Training 5.2 Basics Hyper parameter • Selecting number of Neurons, • Activation Functions, • Layers using Greedy Search and Random Access
Unit VI Deep Learning for Sequential and Image Data	6a. Implement Deep Learning for sequential Data 6b. Implement Deep Learning for Image Data	6.1Sequential Data: • RNN, LSTM, LSTM-GRU, • Introduction to Transformers, GPT 6.2 Image Data : • CNN, (Resnet , VGG) Pre-trained • Neural Networks, Transfer Learning, • Fine Tuning

***Note**: To attain the COs and competency, above listed Learning Outcomes (LOs) need to be undertaken to achieve the 'Application Level' of Bloom's 'Cognitive Domain Taxonomy'.*

9. SUGGESTED SPECIFICATION TABLE FOR QUESTION PAPER DESIGN

Unit No.	Unit Title	Teaching Hours	Distribution of Theory Marks			
			R Level	U Level	A Level	Total Marks
I	Model selection and feature engineering	06	4	4	--	08
II	Supervised Learning: Naive Bayes, Decision Tree	08	4	4	6	14
III	Supervised Learning: Support Vector Machines, K Nearest Neighbors	10	2	4	6	12
IV	Unsupervised learning: Clustering Algorithms	10	2	4	6	12
V	Introduction to Deep Learning	08	2	4	6	12
VI	Deep Learning for Sequential and Image Data	06	2	6	4	12
	Total	**48**	**16**	**26**	**28**	**70**

***Legends:** R=Remember, U=Understand, A=Apply and above (Bloom's Revised taxonomy)*

Note: This specification table provides general guidelines to assist student for their learning and to teachers to teach and assess students with respect to attainment of LOs. The actual distribution of marks at different taxonomy levels (of R, U and A) in the question paper may vary from above table.
This specification table also provides a general guideline for teachers to frame internal end semester practical theory exam paper which students have to undertake.

10. SUGGESTED STUDENT ACTIVITIES

Other than the classroom and laboratory learning, following are the suggested student-related ***co-curricular*** activities which can be undertaken to accelerate the attainment of the various outcomes in this course:

a. Prepare journals based on practical performed in laboratory.
b. Library/E-Book survey regarding Advanced Algorithm in AI & ML used in Computer industries.
c. Prepare power point presentation for showing different types of Advanced Algorithm in AI & ML Applications.

11. SUGGESTED SPECIAL INSTRUCTIONAL STRATEGIES (if any)

These are sample strategies, which the teacher can use to accelerate the attainment of the various outcomes in this course:

a. Massive open online courses (***MOOCs***) may be used to teach various topics/sub topics.
b. ***'L' in item No. 4*** does not mean only the traditional lecture method, but different types of teaching methods and media that are to be employed to develop the outcomes.
c. About ***15-20% of the topics/sub-topics*** which is relatively simpler or descriptive in nature is to be given to the students for ***self-directed learning*** and assess the development of the LOs/COs through classroom presentations (see implementation guideline for details).
d. With respect to item No.10, teachers need to ensure to create opportunities and provisions for ***co-curricular activities***.
e. Guide student(s) in undertaking micro-projects.
f. No. of practical's selection to be performed should cover all units.

12. SUGGESTED MICRO-PROJECTS

Only one micro-project is planned to be undertaken by a student assigned to him/her in the beginning of the semester. S/he ought to submit it by the end of the semester to develop the industry oriented COs. Each micro-project should encompass two or more COs which are in fact, an integration of practicals, cognitive domain and affective domain LOs. The micro-project could be industry application based, internet-based, workshop-based, laboratory-based or field-based. Each student will have to maintain dated work diary consisting of individual contribution in the project work and give a seminar presentation of it before submission. The total duration of the micro-project should not be less than ***16 (sixteen) student engagement hours*** during the course.

In the first four semesters, the micro-project could be group-based. However, in higher semesters, it should be individually undertaken to build up the skill and confidence in every student to become problem solver so that s/he contributes to the projects of the industry.

A suggestive list is given here. Similar micro-projects could be added by the concerned faculty:

a. Credit Card Fraud Detection Using Classification Algorithm
b. Anomaly detection Using KNN Machine learning
c. Image Classification using Support Vector machine
d. Classification based on Decision Tree
e. Application of SVM algorithm for Cancer detection

13. SUGGESTED LEARNING RESOURCES

Sr. No.	Title of Book	Author	Publication
1	Introduction to Machine Learning with Python	Andreas C. Müller & Sarah Guido	O'Reilly Media, Inc
2	Machine Learning in Action	Peter Harrington,	DreamTech, First Edition, 2012 13: 978- 1617290183
3	Machine Learning	Tom M Mitchell,	McGraw Hill, First edition, 1997 13: 978- 0070428072
4	Machine Learning Step-by-Step Guide To Implement Machine Learning Algorithms with Python	Rudolph Russell	CreateSpace Independent
5	Mastering Machine Learning with Python in Six Steps A Practical Implementation Guide to Predictive Data Analytics Using Python	Manohar Swamynathan	Apress Publication ISBN-13 (pbk): 978-1-4842-2865-4
6	Practical Machine Learning with Python A Problem-Solver's Guide to Building Real-World Intelligent Systems	Dipanjan Sarkar, Raghav Bali, Tushar Sharma	Apress publication ISBN-13 (pbk): 978-1-4842-3206-4 ISBN-13 (electronic): 978-1-4842-3207-1
7	Machine Learning using Python	Manaramjan Pradhan, U Dinesh Kumar	Wiley India ISBN: 978-81-265-7990-7
8	Deep Learning with Python	FRANÇOIS CHOLLET	Manning Publications ISBN 9781617294433
9	Python Deep Learning Second Edition	Ivan Vasilev, Daniel Slater	Packt Publishing ISBN 978-1-78934-846-0

14. SOFTWARE/LEARNING WEBSITES

a. https://www.pdfdrive.com/machine-learning-for-absolute-beginners-e188007429.html
b. https://www.geeksforgeeks.org/ml-fuzzy-clustering/
c. https://www.pdfdrive.com/machine-learning-step-by-step-guide-to-implement-machine-learning-algorithms-with-python-d158324853.html
d. https://www.pdfdrive.com/machine-learning-for-absolute-beginners-e188007429.html
e. https://www.geeksforgeeks.org/ml-fuzzy-clustering/
f. https://www.pdfdrive.com/machine-learning-step-by-step-guide-to-implement-machine-learning-algorithms-with-python-d158324853.html
g. https://machinelearningmastery.com/classification-as-conditional-probability-and-the-naive-bayes-algorithm/ (Practical)
h. https://www.geeksforgeeks.org/naive-bayes-classifiers/(Practical)
i. https://www.javatpoint.com/machine-learning-naive-bayes-classifier (Practical)
j. https://www.javatpoint.com/machine-learning-random-forest-algorithm (Random Forest)
k. https://learning.oreilly.com/library/view/deep-learning-with/9781617296864/
l. https://www.deeplearningbook.org/