GUJARAT TECHNOLOGICAL UNIVERSITY

Programme Name: Post Graduate Diploma in Data Science

Level: PG Diploma

Branch: Data Science

Subject Code : DS01080031

Subject Name : Fundamentals of Machine Learning

w. e. f. Academic Year:	2025-26
Semester:	1
Category of the Course:	Compulsory Subject

Prerequisite:	To effectively learn and apply machine learning, a solid foundation in mathematics, particularly linear algebra and calculus, is essential. Additionally, programming skills, especially in languages like Python, are necessary for implementing and testing machine learning algorithms.
Rationale:	Machine learning allows the user to feed a computer algorithm an immense amount of data and have the computer analyze and make data-driven recommendations and decisions based on only the input data.

Course Outcome:

After Completion of the Course, Student will able to:

No	Course Outcomes	RBT Level
01	Understand the concept of Machine learning and range of problems that can be solved by machine learning.	UN
02	Understand and analyze different method and Techniques.	UN
03	Compare different types of learning algorithms and apply machine learning concepts in real life problems.	AN
04	Use of Different datasets analyze the data.	AN

Revised Bloom's Taxonomy (RBT)

Teaching and Examination Scheme:

Teaching Scheme (in Hours)			Total Credits L+T+ (PR/2)	Assessment Pattern and Marks				Total Marks
				Theory		Tutorial / Practical		
L	**T**	**PR**	**C**	ESE (E)	PA / CA (M)	PA/CA (I)	ESE (V)	
4	0	2	5	70	30	00	50	150

GUJARAT TECHNOLOGICAL UNIVERSITY

Programme Name: Post Graduate Diploma in Data Science

Level: PG Diploma

Branch: Data Science

Subject Code : DS01080031

Subject Name : Fundamentals of Machine Learning

Course Content:

Unit No.	Content	No. of Hours	% of Weightage
1.	**Introduction** Learning Problems, designing a learning system, Issues with machine learning. Concept Learning, Version Spaces and Candidate Eliminations, Inductive bias, Supervised, Unsupervised and Semi supervised Learning	08	10
2.	**Supervised Learning** Decision Tree Representation, Appropriate problems for Decision tree learning, Algorithm, Hypothesis space search in Decision tree learning, inductive bias in Decision tree learning, Issues in Decision tree learning, Overfitting, Underfitting, Pre-pruning, Post-pruning	12	25
3.	**Artificial Neural Networks** Neural Network Representation, Appropriate problems for Neural Network Learning, perceptron, Cost function, Gradient descent, Hypothesis formation, Multilayer Networks and Back Propagation Algorithms, Remarks on Back Propagation Algorithms Case Study: face Recognition	10	25
4.	**Bayesian Learning** Bayes Theorem, Bayes Theorem and Concept Learning, Maximum Likelihood and Least squared Error Hypothesis, Maximum likelihood hypothesis for Predicting probabilities, Minimum Description Length Principle, Bayes Optimal Classifier, Gibbs Algorithm, Naïve Bayes Classifier , Bayesian Belief Network, EM Algorithm **Case Study:** Learning to classify text	12	25
5.	**Overview of typical application areas such as Recommender System etc**	6	15
	Total	**48**	**100**

Suggested Specification Table with Marks (Theory):

Distribution of Theory Marks (in %)					
R Level	**U Level**	**A Level**	**N Level**	**E Level**	**C Level**
15	**30**	**25**	**20**	**25**	**10**

Where R: Remember; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (as per Revised Bloom's Taxonomy)

GUJARAT TECHNOLOGICAL UNIVERSITY

Programme Name: Post Graduate Diploma in Data Science

Level: PG Diploma

Branch: Data Science

Subject Code : DS01080031

Subject Name : Fundamentals of Machine Learning

References/Suggested Learning Resources:

(a) Books:

1. Machine Learning
by Tom M Mitchell
McGraw Hill, ISBN: 0070428077

2. Machine Learning in Action
by Peter Harrington
Manning Publications Co., ISBN 9781617290183

3. The Elements of Statistical Learning: Data Mining, Inference, and Prediction
by Trevor Hastie, Robert Tibshirani, Jerome H. Friedman
Springer Series in Statistics, Second edition, ISBN 978-0-387-84858-7

4. Real-World Machine Learning
by Henrik Brink, Joseph Richards, Mark Fetherolf
Manning Publications Co., ISBN-10: 9781617291920,ISBN-13: 978-1617291920

(b) Open source software and website:

1. OPENML: https://www.openml.org/
2. TensorFlow: https://www.tensorflow.org/
3. https://www.coursera.org/learn/foundations-of-machine-learning
4. https://www.coursera.org/learn/fundamentals-of-machine-learning-and-artificial-intelligence

Suggested Course Practical List:

1. Import the real world dataset from any of the public dataset repository. Apply various data functions to clean the data (i.e. handling missing values, normalization).
2. For cleaned dataset, construct the decision tree, visualize it and use it for classification of the new samples.
3. Apply various decision tree pruning methods by tuning the parameters of functions. Make a comparison chart based on it for five different real world datasets.
4. Measure the precision, recall, accuracy, and F-score of the decision tree classifier on any text classification using 10-fold cross-validation. Also apply tree pruning option. Compare the results without pruning and with pruning.
5. For cleaned dataset, construct the random forest, visualize it and use it for classification of the new samples.
6. On cleaned dataset, apply Naïve Bayes classification and compare its result with decision tree and random forest.

GUJARAT TECHNOLOGICAL UNIVERSITY

Programme Name: Post Graduate Diploma in Data Science

Level: PG Diploma

Branch: Data Science

Subject Code : DS01080031

Subject Name : Fundamentals of Machine Learning

7. Develop a code to classify spam mail with Naïve Bayes.
8. Estimate the accuracy of Naïve Bayes algorithm using 10-fold cross validation on the house-votes-84 data set.
9. Develop a feed forward neural network with backpropagation function to improve a handwritten character recognition system.
10. Develop a neural network, use learning functions and tune the parameters to reduce the mean square error for recognizing the face.

* * * * * * *