GUJARAT TECHNOLOGICAL UNIVERSITY

Bachelor of Engineering Syllabus

Subject Code : 3154602

Subject Name : Machine Learning

WEF Academic Year :	2022 - 23
Semester :	5
Category of the Course :	Professional Elective

Prerequisite :	Linear Algebra, Probability, Statistics
Rationale :	The students will become familiar with the ideas of machine learning, the algorithms that are used in machine learning, and the process of designing applications that make use of machine learning for a variety of different domains.

Course Scheme :

Teaching Scheme			Total Credits	Assessment Pattern and Marks				Total Marks
				Theory		Practical		
L	T	PR	C	ESE (E)	PA(M)	ESE (V)	PA (I)	
3	0	2	4	70	30	30	20	150

Course Content :

Sr. No.	Course Content	No. of Hours	% of Weightage
1	**INTRODUCTION TO MACHINE LEARNING :** Overview of Human Learning and Machine Learning, Machine Learning basics, Applications of ML, Data Mining Vs Machine Learning vs. Big Data Analytics. Types of Machine Learning: Supervised learning, Unsupervised learning, Reinforcement learning, Applications of Machine Learning.	6	15
2	**FEATURE SELECTION AND DIMENSION REDUCTION :** **Feature selection**: Subset Selection, Ranking, Decision tree - Entropy reduction and information gain, Exhaustive, best first, Greedy forward & backward. **Dimension Reduction :** Principal Components Analysis, Linear Discriminant Analysis.	8	20
3	**SUPERVISED MACHINE LEARNING :** **Regression**: Simple linear regression, Multiple linear regression, Polynomial regression, evaluating regression fit. **Classification** : Logistic Regression, Decision tree, Random forest, Naive Bayes, Support vector machine.	10	25
4	**UNSUPERVISED MACHINE LEARNING :** Grouping unlabeled items using k-means clustering, Anomaly detection, Hierarchical Clustering.	6	15

5	**DEVELOPMENT AND EVALUATION OF ML MODELS :** Steps in ML modeling, Data Collection, Data pre-processing, Model Selection, Cross validation, Model evaluation: confusion matrix, Accuracy, Precision, Recall, True positive, false positive etc., Hyper parameter Tuning, Grid Search, Predictions, Bias/variance tradeoff.	6	15
6	**ENSEBMLE TECHNIQUES :** Combining multiple learners: Model combination schemes, Voting, Ensemble Learning - bagging, boosting, stacking.	4	10
	Total		**100**

Reference Book :

- Tom Mitchell, “Machine Learning”, First Edition, McGraw- Hill, 1997.
- Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, MIT Press, 2012.
- Saikat Dull, S. Chjandramouli, Das, “Machine Learning”, Pearson.
- Mark Fenner, “Machine Learning with Python for Everyone”, Pearson.

Course Outcome :

After Completion of the Course, Student will able to :

No.	Course Outcomes	RBT Level*
01	Demonstrate fundamentals of machine learning.	RM
02	Understand the importance of feature selection and dimension reduction techniques to deal with curse of dimensionality.	UN
03	Apply machine learning algorithms to solve various problems.	AP
04	Devise and develop a machine learning model using various steps.	CR
05	Evaluate the performance of machine learning models.	EL

*RM: Remember, UN: Understand, AP: Apply, AN: Analyze, EL: Evaluate, CR: Create

Suggested Course Practical List :

1. Learn about basic packages needed for machine learning.
2. Find statistical properties of the given vectors. Also compute covariance and correlation matrix of given dataset.
3. Use Boston house prediction data set and develop a simple linear regression model to predict the price of the house.
4. Use Boston house prediction data set and develop a multiple linear regression model to predict the price of the house.
5. Write a program to compute various distance measures and similarity measure between the samples.

6. Load iris flower data set and implement decision tree model and test its performance.
7. Load iris flower data set and implement kNN model and test its performance.
8. Use titanic dataset and apply Principle Component Analysis to reduce the dimension of it. Use logistic regression to perform classification on reduced dataset.
9. Given a dataset for classification task. Write a program to implement Support Vector Machine and estimate it test performance.
10. Write a program to implement K means clustering algorithm. Select your own dataset to test the program. Demonstrate the nature of output with varying value of K.

List of Laboratory/Learning Resources Required :

https://codecrucks.com/
https://www.machinelearningisfun.com/
https://machinelearningmastery.com/blog/
http://fastml.com/
https://netflixtechblog.com/tagged/machine-learning?gi=6fab2ff50309

****