GUJARAT TECHNOLOGICAL UNIVERSITY

BACHELOR OF ENGINEERING SYLLABUS

5th Semester, Subject Code : 3154203

Subject Name : Machine Learning Essentials

Category : Professional Core.

Prerequisite : Linear Algebra, Probability, Programming Language.

Rationale: Understanding and analyzing data in the era of data is trivial. Industries in this age are data driven. This course enable to understand various kinds of data, how to model them, and how to evaluate those models.

Course Objectives :

1. Understand importance of machine learning analysis of data.
2. Should be able to various kind of ML models.
3. Should be able to evaluate and compare various ML models.

Teaching and Examination Scheme :

Teaching Scheme			Credits	Examination Marks				Total Marks
				Theory Marks		Practical Marks		
L	T	P	C	ESE (E)	PA (M)	ESE Viva (V)	PA (I)	
4	0	2	5	70	30	30	20	150

Sr.	Description	Total Hrs.	Weightage (%)
1	**Introduction:** History, Human Learning, Machine Learning, Types of Machine Learning.	04	10%
2	**Data Preprocessing:** Machine Learning activities, Types of data in Machine Learning, Structures of data, Descriptive statistics: Mean, Median, Mode, Standard, deviation, Variance Data visualization: Scatter plot, Histogram, Box plot, Pre-Processing: Dimensionality reduction, Feature subset selection.	06	15%
3	**Evaluating ML models:** Predictive model, Descriptive model, Training a Model for supervised learning: Holdout method, K-fold cross validation, Bootstrapping Overfitting, overfitting, bias variance tradeoff	06	15%

	Performance evaluation of supervised and unsupervised models		
4	**Feature Engineering :** Introduction, Feature transformation, Feature Construction, Principal Component Analysis, Singular Value Decomposition, Linear Discriminant Analysis, Feature subset selection: Measure of feature relevance, measure of feature redundancy Overall feature selection process	06	15%
5	**Supervised Learning :** Supervised Learning, Classification vs Regression, Linear Regression, Logistic Regression, Decision Tree, K-nearest neighbor, Support Vector Machine, Artificial Neural Networks	10	25%
6	**Unsupervised Learning :** Supervised vs. unsupervised learning, Clustering, Anomaly detection, Association rules	06	15%
7	**Ensemble Learning Models :** Bagging, Boosting, Random Forest Classifier	04	10%

Reference books :

1. Machine Learning, Saikat Dutt, S. Chandramouli, Amit Kumar Das, Pearson, 1st edition.
2. Machine Learning with Python For Everyone, Mark E. Fenner, Addison Weseley Data & Analytics Series.
3. Python for Data Analysis, Wes McKinney, O'Reilly.
4. Introduction to Machine Learning with Python, A. C. Muller, Sarah Guido, O'Reilly.

Course Outcomes :

1. Understand the importance of machine earning over human learning.
2. Understand characteristics of various types of data with the help of statistic descriptive analysis.
3. Understand the metric used for evaluating models.
4. Understand the importance of feature engineering.
5. Understand the use of ML models for various applications.

List of practical :

1. Find and analyze mean, median and mode of given data.
2. Understand structure of data using various visualization methods like scatter plot, histogram and boxplot.
3. Implement Logistic Regression on Iris dataset and evaluate its performance.
4. Implement Decision Tree on Iris dataset and evaluate its performance.
5. Implement K-NN on Iris dataset and evaluate its performance.
6. Implement SVM on Iris dataset and evaluate its performance.
7. Implement ANN on Iris dataset and evaluate its performance.
8. Perform K-means clustering on Irish dataset and evaluate its performance.
9. Read Wisconsin Breast Cancer Dataset (WBCD) and implement various ensemble models on it.
10. Apply various dimension reduction methods on WBCD and evaluate their performance.