



National Institute of Technology Meghalaya

An Institute of National Importance

CURRICULUM

	National Institute of Technology Meghalaya An Institute of National Importance	CURRICULUM														
Programme	Bachelor of Technology in Computer Science and Engineering	Academic Year of Regulation 2018-19														
Department	Computer Science and Engineering	Semester VIII														
Course Code	Course Name	Credit Structure														
		Marks Distribution														
		L T P C INT MID END Total														
CS 422	Data Mining	3 0 0 3 50 50 100 200														
Course Objectives	This course illustrates the need of data mining and data pre-processing techniques	Course Outcomes	CO1	Able to experiment with different data pre-processing techniques												
	This course explains the different types of data mining techniques as association rule mining, classification, clustering and outlier detection techniques		CO2	Able to estimate and compare different data mining techniques												
	This course explains the different data mining techniques to real life applications		CO3	Able to design data mining solution framework for real life problems												
No.	COs	Mapping with Program Outcomes (POs)												Mapping with PSOs		
		PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
1	CO1	1	1	-	-	-	-	-	-	-	-	-	-	2	1	-
2	CO2	1	1	-	-	-	-	-	-	-	-	-	-	2	1	-
3	CO3	1	1	2	-	2	-	-	-	-	-	-	-	2	1	-
SYLLABUS																
No.	Content													Hours	COs	
I	Introduction: Data Mining, Motivation, Applications, Data Mining Approaches, Data Types, Data Objects and Attributes, Challenges in Data Mining, Data Similarity and Dissimilarity Measures Data —Preprocessing: Data Quality Issues, Data Cleaning, Data Integration, Data Reduction, Data Transformation and Data Discretization													06	CO1	
II	Mining Frequent Pattern Mining and Association Rules: Basic Concepts, Apriori Algorithm, Frequent Pattern growth (FP-growth) Algorithm, Mining Closed and Max Patterns, Pattern Evaluation Methods, Constraint-Based Frequent Pattern Mining													08	CO1	
III	Classification Techniques: Basic Concepts, Decision Tree Classifier, Rule-Based Classifier, Nearest Neighbor Classifiers, Naive Bayes Classifier, Artificial Neural Network (ANN), Support Vector Machine (SVM), Model Over fitting, Model Evaluation and Selection													10	CO2	
IV	Clustering Techniques: Overview, Types of Clustering Methods, Partitioning Methods, Hierarchical Methods, Density-Based Methods, Grid-Based Methods, Performance Parameters, Clustering with Constraints Outlier Detection: Basic Concepts, Outlier Detection Methods, Statistical Approaches, Proximity-Based Approaches, Clustering-Based Approaches, Classification-Based Approaches													12	CO2 & CO3	
Total Hours													36			
Essential Readings																
1. J. Han, J. Pei, and M Kamber. "Data mining: concepts and techniques". Elsevier, 3 rd edition, 2011																
2. P.N. Tan, M. Steinbach, A. Karpatne, and V. Kumar. "Introduction to data mining". Pearson Education India, 2 nd edition, 2016.																
3. C.C. Aggarwal. "Data mining: the textbook". Springer, 1 st edition, 2015.																
Supplementary Readings																
1. C.C. Aggarwal, and C. Zhai. "Mining text data". Springer, 1 st edition, 2012.																
2. J. Leskovec, A. Rajaraman, and J.D. Ullman. "Mining of massive datasets". Cambridge University Press, 3rd edition, 2019																
3. J. Dean. "Big data, data mining, and machine learning: value creation for business leaders and practitioners". John Wiley & Sons, 1 st edition, 2014.																