DATA MINING

PROF. PABITRA MITRA
Department of Computer Science and Engineering
IIT Kharagpur

TYPE OF COURSE : Rerun | Elective | UG
COURSE DURATION : 8 weeks (15 Feb’ 21 - 09 Apr’ 21)
EXAM DATE : 24 Apr 2021

INTENDED AUDIENCE : Any engineering discipline and mathematics, physics.
INDUSTRIES APPLICABLE TO : TCS, Infosys, CTS, Accenture

COURSE OUTLINE :
Data mining is study of algorithms for finding patterns in large data sets. It is an integral part of modern industry, where data from its operations and customers are mined for gaining business insight. It is also important in modern scientific endeavors. Data mining is an interdisciplinary topic involving, databases, machine learning and algorithms. The course will cover the fundamentals of data mining. It will explain the basic algorithms like data preprocessing, association rules, classification, clustering, sequence mining and visualization. It will also explain implementations in open source software. Finally, case studies on industrial problems will be demonstrated.

ABOUT INSTRUCTOR :
Prof. Pabitra Mitra is an Associate Professor of Computer Science and Engineering at Indian Institute of Technology Kharagpur. He did his BTech in Electrical Engineering from IIT Kharagpur and PhD from ISI Calcutta. He was a Scientist at Centre for Artificial Intelligence and Robotics, Bangalore and an Assistant Professor at IIT Kanpur. He received the INAE Young engineering Award, IBM Faculty Award and Yahoo Faculty Award. He has authored a book on Data mining and about 50 papers in international journals.

COURSE PLAN :

Week 1: Introduction, Data Preprocessing
Week 2: Association Rule Mining, Classification Basics
Week 3: Decision Tree, Bayes Classifier, K nearest neighbor
Week 4: Support Vector Machine, Kernel Machine
Week 5: Clustering, Outlier detection
Week 6: Sequence mining
Week 7: Evaluation, Visualization
Week 8: Case studies