

Dept Number	CS 533		Course Title	Data Mining and Big Data Analysis						
Semester Hours	3		Course Coordinator	Dunren Che						
Catalog Description	This course provides a series of comprehensive and in-depth lectures on the core techniques in data mining and knowledge discovery; addresses the unique issues of big data; and discusses potential applications of data mining particularly on big data analysis. Major topics include: data preparation, association mining, classification (and prediction), clustering, characteristics and challenges of big data, and strategies of big data mining and analysis.									
Textbooks										
Introduction to Data Mining (1st Ed), Pang-Ning Tan, Michael Steinbach, Vipin Kumar (2018). ISBN-13: 978-0-13-312890-1. Published by Pearson.										
References										
Mining of Massive Datasets , by Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman (manuscript available in PDF, unpublished)										
Course Learning Outcomes										
To learn the core techniques of data mining, including: <ul style="list-style-type: none"> • Association analysis • Classification/Prediction • Clustering (cluster analysis) • Anomaly detection and analysis • And their application/adaptation to Big Data 										
Assessment of the Contribution to Student Outcomes										
Outcome →	1	2	3	4	5	6	7	8	9	10
Assessed →	X	X			X					
Prerequisites by Topic										
CS 330 and 430 with grades of C or better or consent of instructor.										

Major Topics Covered in the Course

1. Introduction to Data Mining and Bioinformatics {4 classes}
2. Data Cleaning/Transformation/Preparation {6 classes}
3. Association Rule Mining {6 classes}
4. Classification/Prediction Techniques {6 classes}
5. Clustering Techniques {6 classes}
6. Anomaly detection and analysis {4 classes}
7. Special issues of Big Data Mining and Analysis {8 classes}