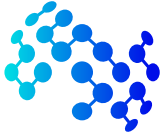




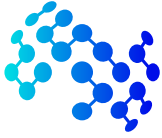
Core Courses Syllabi

DS701 - Data Mining

Title	Data Mining
Code	DS701
Loading	4 Credit-hours
Prerequisites	<ul style="list-style-type: none">Discrete Mathematics, Probability and StatisticsProficiency in Java or Python
Catalog Description	This course is an introductory course on data mining, which is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.
Goal	The aim of this course is to provide students with the comprehensive understanding of the modern development of data mining foundations and techniques. Students will be able to develop advanced skills to solve a wide range of unsupervised learning problems, such as frequent pattern mining and data clustering.
Content	This course introduces the basic concepts, principles, methods, implementation techniques, and applications of data mining, with a focus on two major data mining functions: (I) pattern discovery and (II) cluster analysis.
Recommended Textbooks	Data Mining: Concepts and Techniques (The Morgan Kaufmann Series in Data Management Systems) 3rd Edition by Jiawei Han, Micheline Kamber, Jian Pei
Recommended References & Supplemental Material	Relevant research papers, tech reports, and surveys for each topic, where needed, are identified in the teaching plan ahead.



Teaching Week	Topics
1	Efficient Pattern Mining Methods Lecture <ul style="list-style-type: none">• Course Orientation;• Course Part 1 Pattern Discovery Overview;• Pattern Discovery Basic Concepts;• Efficient Pattern Mining Methods: Apriori, ECLAT, and FPgrowth. Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
2	Pattern Evaluation Lecture <ul style="list-style-type: none">• Pattern Evaluation Measures: Lift, chisquare, cosine, Jaccard, and Kulczynski• Association Analysis• Correlation Analysis Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
3	Mining Diverse Frequent Patterns Lecture <ul style="list-style-type: none">• Pattern mining multi-level and multi-dimensional space• Mining qualitative patterns• Mining rare patterns and negative correlations• Mining long (colossal) patterns. Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
4	Sequential Pattern Mining Lecture <ul style="list-style-type: none">• Sequential Pattern Mining methods: GSP, SPADE, PrefixSpan, and CloSpan;• Mining Spatiotemporal and Trajectory Patterns Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
5	Constraint-Based Mining Lecture <ul style="list-style-type: none">• Data and pattern-based constraints• Anti-monotone and monotone constraints• Succinct, convertible, and multiple constraints. Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week



Teaching Week	Topics
6	Graph Pattern Mining Lecture <ul style="list-style-type: none">• Subgraph pattern mining methods: gSpan, CloseGraph,• Mining top-k large structural patterns in a single large network• Graph indexing and similarity search in graph databases. Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
7	Pattern-Based Classification Lecture <ul style="list-style-type: none">• Decision Tree Induction• Rule-based Classification• K-nearest-neighbor classifier Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week• Midterm Exam preparation
8	Text Data Mining Lecture <ul style="list-style-type: none">• Mining Quality Phrases from Text Data;• Sentiment Analysis• Advanced Topics on Pattern Discovery Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
9	Cluster Analysis Lecture <ul style="list-style-type: none">• Course Part 2 Cluster Analysis Overview;• Cluster Analysis Introduction;• Similarity Measures for Cluster Analysis Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
10	Partition-Based Clustering Methods Lecture <ul style="list-style-type: none">• K-Means: A Centroid-Based Technique• K-Medoids: A Representative Object-Based Technique• KMedians and the Kernel K-Means Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week



Teaching Week	Topics
11	Hierarchical Clustering Methods Lecture <ul style="list-style-type: none">• BIRCH: Multiphase Hierarchical Clustering Using Clustering Feature Trees• Chameleon: Multiphase Hierarchical Clustering Using Dynamic Modeling• Probabilistic hierarchical clustering Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
12	Density-Based Clustering Methods Lecture <ul style="list-style-type: none">• DBSCAN: Density-Based Clustering Based on Connected Regions with High Density• OPTICS: Ordering Points to Identify the Clustering Structure• DENCLUE: Clustering Based on Density Distribution Functions Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
13	Grid-Based Clustering Methods Lecture <ul style="list-style-type: none">• Organizing regions of the data space into grid-like structures• STING: STatistical INformation Grid• CLIQUE: An Apriori-like Subspace Clustering Method Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
14	Clustering Validation Lecture <ul style="list-style-type: none">• Assessing Clustering Tendency• Determining the Number of Clusters• Measuring Clustering Quality Lab <ul style="list-style-type: none">• Instructor-led demonstration related to the topics taught in the week
15	Guest Lecture & Review Lecture <ul style="list-style-type: none">• Guest Lecture on Trending Topic Lab <ul style="list-style-type: none">• Review and Final Exam Preparation