

<Introduction to Data Mining>

Description

Data mining refers to theories and techniques for finding useful patterns from massive amount of data. Data mining has been used in high impact applications including web analysis, fraud detection, recommendation system, cyber security, etc. This course covers important algorithms and theories for data mining. Main topics include mapreduce, finding similar items, mining frequent patterns, link analysis, data stream mining, clustering, graphs, and mining big data.

Topics to be covered

- Map-Reduce and the New Software Stack
- Finding Similar Items
- Mining Data Streams
- Link Analysis
- Frequent Itemset
- Clustering
- Advertising on the Web
- Recommendation System
- Graphs
- Dimensionality Reduction

Textbook

Mining of Massive Datasets by Jure Leskovec, Anand Rajaraman, Jeff Ullman.