Course: Data Mining

Instructor: C.W. Hsiao

Course description

Data mining, or called knowledge discovery, is the processes of analyzing data from different perspectives, deriving useful information, and finally acquiring knowledge. The applications of data mining are now prevalent in varied domains, such as stock prediction, customer behavior analysis, and social network. In this course, we will learn what data mining can do and how to do it. Some important concepts and techniques, including association rules, clustering, classification, and artificial intelligence, will be discussed.

References

The Top Ten Algorithms in Data Mining,
Xindong Wu, Vipin Kumar,
Chapman and Hall/CRC

Data Mining: Practical Machine Learning Tools and Techniques (Second Edition),
Ian H. Witten, Eibe Frank,
Morgan Kaufmann

Introduction to Data Mining,
Pang-Ning Tan, Michael Steinbach, Vipin Kumar,
Addison-Wesley

Course Schedule

Introduction

Support Vector Machines

Support Vector Machines

K-Means

Expectation Maximization

Apriori Algorithm

Holiday

k-Nearest Neighbors

Mid-term Exam Week

PageRank

Naive Bayes

C4.5

AdaBoost

CART

Paper Presentation

Paper Presentation

Paper Presentation

Final Exam Week

Course evaluation

Passing score for graduate course is 70. In general, score is allocated between class attendance, homework, mid-term written exam, final written exam and student oral presentation. Course instructor reserves the right to adjust the grading scheme.