CS 415: DATA MINING (3-0-0: 3)

Introduction to Data Mining
Basic concepts in data mining, Data measurement, exploratory data analysis, data visualization.

Basic Principles of Data Mining
Predictive modeling: classification and regression, model fitting as optimization, evaluation of predictive performance, over-fitting, regularization, other data mining tasks: clustering and pattern detection, Lasso method for regularized regression.

Text Mining
Information retrieval and search, Text classification, Unsupervised learning.

Data Mining Applications

Recommender System: Recommender data, Netflix prize data, nearest neighbor algorithms, matrix decomposition algorithms, efficient algorithms for large data sets, modeling systematic effects.

Web Data Analysis: Web data, collection and interpretation, analyzing user browsing behavior, learning from click through data, predictive modeling and online advertising, link analysis and the Page-Rank algorithm..

Data Mining Applications

Social Network Analysis: descriptive analysis of social networks, network embedding and latent space models, network data over time: dynamics and event-based networks, link prediction.

Time Series Analysis and Anomaly Detection: Basic concepts in time-series analysis, principles of Markov and hidden Markov models, event data and Poisson models, techniques for detecting anomalies, events, changes, motifs, etc, case study: analysis of large-scale traffic data**.**

Text Books:
1. D. J. Hand, H. Mannila, P. Smyth*, "Principles of Data Mining*" MIT Press.
2. J. Leskovec, A. Rajaraman, J. D. Ullman , "Mining Massive Data Sets" , Cambridge University Press.

References:

1. C.D. Manning, P. Raghavan, and H. Schutze, "Introduction to Information Retrieval " Cambridge University Press.