Syllabus for Course Text Mining

Lecturer: Ronen Feldman

We will cover the following topics:

1. Introduction to Text Mining
2. Term Extraction
3. Text Categorization
   a. Decision Trees
   b. RIPPER
   c. Naive Bayes
   d. SVM
4. Information Extraction
   a. General Architecture
   b. HMM
   c. Knowledge Based Systems
      i. DIAL
   d. CRYSTAL
   e. Boot Strapping
   f. Unsupervised Learning
      i. Knowitall
      ii. SRES
      iii. CARE
5. Analytics
   a. Maximal Association Rules
   b. Trend Analysis
   c. Distribution Analysis
   d. Comparing Profiles
6. Link Analysis
7. Visualizations
   a. Circle Graphs
   b. Spring Graphs
8. Sentiment Analysis
   a. Document Level
   b. Sentence Level
   c. Aspect Level
   d. Building Sentiment Lexicons
   e. Comparative Sentiment Analysis
9. Applications
   a. BioTech
   b. Competitive Intelligence
   c. Anti Terror Applications
   d. Financial Applications
      i. News Analysis
      ii. StockTwits
10. Text/Data Mining Packages
    a. Stanford NLP Tools
    b. OpenNLP
    c. Rapid Miner
d. Weka

References


*The Handbook of Data Mining* (N. Ye, ed.). Lawrence-Erlbaum Associates.

Handbook of Data Mining and Knowledge Discovery Edited by WILLI KLÖSGEN, Fraunhofer Institute for Autonomous Intelligent Systems, Sankt Augustin, Germany, and the late JAN M. ZYTKOW

KDD-2002 Conference on Knowledge Discovery and Data Mining, proceedings and CD-ROM, ACM Press, 2002


KDD-99 Conference on Knowledge Discovery and Data Mining, ed. S. Chaudhuri, D. Madigan. ACM Press, 1999

