Project Guideline

CIS 4930 / 6930 Data Mining

Fall 2007
Outline

- Select team members
  - Ideally the class needs to organize itself in teams of three or four members
- Pick a thread of data mining research
  - Use selected list of papers as a start point
  - Or pick your own from recent KDD, ICDM, TKDD or TKDE
- Comprehensive literature survey in your selected thread
  - Use Citeseer, Google Scholar extensively
  - Read papers carefully. Identify key ideas
  - Compare and contrast various approaches
Choose a few approaches that you will implement and experiment with
  - Follow the strengths of your team
  - Justify your choices

Implement these approaches
  - It is OK to use existing data mining code / toolkits, but the key algorithm / approach should be your own implementation
  - Clearly identify and acknowledge the code / toolkit that is not your own work
Outline ...

- Identify data sets to evaluate your implementations
  - Synthetic datasets
  - Real world datasets

- Perform the actual experiments

- Conclusions
Time Line and Deliverables

- Teams and research thread (Oct 29)
- Project plan (Nov 5)
- In-class project presentations (Nov 30 and Dec 3)
- Project report (Dec 5)