CMSC 828G Project Part 1

Data Mining,CMSC 828G, Spring 2002

Due Date
Writeup: Wednesday February 13, before noon
bring by AVW 3205 or slide under my door
Presentation: Thursday February 14

Writeup

Please turn in a one page description of the data set you will use for your project.

You should:

- Describe the problem domain.
- Describe the procedure you will use for obtaining the data (if relevant).
- Describe the attributes in your domain. Describe how the attributes are measured. Discuss any issues with representations and possible alternate representations.
- Describe the number of instances, any bias that there may be in the sample and whether the sample is representative.
- Describe possible data mining goals.

Presentation

In class, each person will give a 3-5 minute presentation describing their problem domain. Please describe your data set and the items listed above.

You may use overheads if you like. In order to ensure that we can hear about everyone's data set choice, your 5 minute time limit will be enforced.