Data Mining
資料探勘

Project 2

Hung-Yu Kao, Fall 2018

Classification

- **Goal**
  - Understand what classification systems do and the difference between real behavior of classification model and observed data

- Description
  - Construct a classification model to observe the difference between real 'right' data and modeled data

- **Due**: 11/20 9am

Data Mining

Flow

- **<u>Step 1</u>**: Design a set of rules to classify data, e.g., classify students with good performance.
  - You should design **k** features/attributes for your problems first.
  - Use 'absolutely right' rules to generate your positive and negative data (the number of data = **M**)
- **<u>Step 2</u>**: Use the data generated in Step 1 to construct your classification model
  - Decision tree is basic requirement, you can add more classification models.
- **<u>Step 3</u>**: Compare the rules in the decision tree from Step 2 and the rules you used to generate your 'right' data
- **<u>Step 4</u>**: Discuss anything you can

Data Mining