

Data Mining

Week 1 Exercise Sheet

1. Discuss whether or not each of the following activities is a data mining task.
 - (a) Dividing the customers of a company according to their gender.
 - (b) Dividing the customers of a company according to their profitability.
 - (c) Computing the total sales of a company.
 - (d) Sorting a student database based on student identification numbers.
 - (e) Predicting the outcomes of tossing a fair pair of dice.
 - (f) Predicting the future stock price of a company using historical records.
 - (g) Monitoring the heart rate of a patient for abnormalities.
 - (h) Monitoring seismic waves for earthquake activities.
 - (i) Extracting the frequencies of a sound wave.
2. Suppose that you are employed as a data mining consultant for an Internet search engine company. Describe how data mining can help the company by giving specific examples of how techniques, such as clustering, classification, association rule mining and anomaly detection can be applied.
3. For each of the following data sets, explain whether or not data privacy is an important issue.
 - (a) Census data collected from 100-1950.
 - (b) IP addresses and visit times of Web users who visit your website.
 - (c) Images from Earth-orbiting satellites.
 - (d) Names and addresses of people from the telephone book.
 - (e) Names and email addresses collected from the Web.