**Chapter 3**

**Exercises**

1 Describe benefits and drawbacks of a source-driven architecture for gathering of data at a data warehouse, as compared to a destination-driven architecture.

2. Suppose half of all the transactions in a clothes shop purchase jeans, and one third of all transactions in the shop purchase T-shirts. Suppose also that half of the transactions that purchase jeans also purchase T-shirts. Write down all the (nontrivial) association rules you can deduce from the above information giving support and confidence of each rule.

3. The organization of parts, chapters, sections and subsections in a book is related to clustering. Explain why and to what form of clustering.

4. Suggest how predictive mining techniques can be used by a sports team, using your favorite sport as an example.

5 Suppose you want to find documents that contain at least k of a given set of n keywords. Suppose also you have a keyword index that gives you a (shorted) list of identifiers of documents that contain a specified keyword. Give an efficient algorithm of find the desired set of documents

Using some toll to finish the DM