Business Intelligence Lab

Second mid-term 28/5/2015

Deliver solutions to ex. 1, 2, 3 within 3 h

Notice: use your own SQL Server credentials (the lbi account is disabled)

Exercise 1 (7 pts). The average fraction of female customers in a country is defined as:

$$I = \frac{1}{n} \cdot \sum_{i=1}^{n} \cdot \frac{m_i}{t_i}$$

where i ranges over customer cities with at least one customer with non-zero sales, m_i is the number of female customers in city i with non-zero sales, t_i is the number of customers in city i with non-zero sales.

Write a MDX query on the Sales cube of the `ruggieri_foodmart` OLAP database to report the value I for each country.

What to deliver: MDX query and a brief comment about it, a PowerPoint file with the screenshot of the MDX query result.

Exercise 2 (3 pts). Answer the question of Ex. 1 using an SQL query over the `foodmart` datawarehouse. The SQL query may use analytic functions.

What to deliver: SQL query and a brief comment about it, a PowerPoint file with the screenshot of the SQL query result.

Exercise 3 (6 pts). Consider the `census.arff` dataset splitted into 60% training and 40% test. Predictions of a classifier are non-discriminatory on the test set if $P(income = >50\text{K}|sex = female) = P(income = >50\text{K}|sex = male)$. Build a J48 classifier and then make predictions on the test set so that they are non-discriminatory. Compare the resulting accuracy with the one of predictions of the pure J48.

What to deliver: either a Weka knowledge flow `.kfml` file or a PowerPoint file with screenshots of Weka explorer or a Java program with Weka API calls, and a brief description of the steps of the analysis.

How to deliver: send an e-mail with a single <your surname>.zip file attached to ruggieri@di.unipi.it, including your name, surname, student ID, and computer IP address (http://www.whatismyip.com).

Results and oral exam. Results will be published on-line by tomorrow, including the date and time for those who are admitted to the oral exam. Oral exams will start on Friday 6 June.