Business Intelligence Lab

# Second mid-term 16/12/2015

**Deliver solutions within 3 h**

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

**Exercise 1 (8 pts).** Answer the following business questions with MDX over the *Sales* cube of `ruggieri_foodmart` project:

(a) for every product category, percentage of customers buying that product category on Tuesday's over the total number of customers buying that category;

(b) for every customer city, number of different product categories bought by customers of that city;

(c) for every product category and customer country, the store with the highest percentage of sales over the total sales of that category and country. Result must be ordered by percentage descending.

**What to deliver:** Power Point file with MDX queries (as text) and with a brief comment about them, and with screenshots of SQL Management Studio with query results.

**Exercise 2 (2 pts).** Answer the business question (c) of Exercise 1 with SQL with analytic functions over the `foodmart` datawarehouse.

**What to deliver:** Power Point file with SQL query (as text) and with a brief comment about it, and with screenshot of SQL Management Studio with query result.

**Exercise 3 (6 pts).** Consider a dataset with predictive attributes $S = \{a_1, \ldots, a_n\}$ and a class attribute $c$, and the following statement:

*For all $S_1 \subseteq S_2 \subseteq S$, the J48 classifier learned from $S_1$ attributes only has an accuracy lower or equal than the J48 classifier learned from $S_2$ attributes only.*

Discuss whether the statement is true or not, and experiment on the `census.arff` dataset using Weka APIs.

**What to deliver:** a powerpoint with the discussion and with a description of the experiments done, and a Java program with Weka API calls.

**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 and oral dates will be published on-line by Tuesday 22nd December 2015.