**Assignment 3 Decision Trees 2020**

### The objective of these practical exercises is to make you familiar in practice with decision trees and association rules.

**Decision Trees**

**Exercise 2.1:** (compute the answers manually)

The following table consists of data from a simplified employee database:

No Department Status Age Salary

1 sales senior 31…40 high

2 sales junior 21…30 low

3 sales junior 31…40 low

4 systems junior 21…30 high

5 systems senior 31…40 high

6 systems junior 21…30 high

7 systems senior 41…50 high

8 marketing senior 31…40 high

9 marketing junior 31…40 low

10 secretary senior 41…50 low

11 secretary junior 21…30 low

Let *salary* be the class label attribute.

The decision tree depicted in Fig. 1 is constructed based on the employee data on the table above.

low

not senior

senior

not systems

systems

high

31…40

high

low

not 31…40

Fig. 1. Decision tree based on the employee data.

A. Compute the confusion matrix for the model presented in Fig.1.

B. By using the model (tree) in Fig.1, predict the salary for an employee with the following

characteristics:

**Department Status Age**

sales senior 41…50

# Exercise 2.2:

* Create *Houses\_nominal*.arff using the Houses\_nominal Excel file
* Open file *Houses nominal.arff* in WEKA
* Run *J48* classifier ("*trees*"subfolder).

**Tasks:**

**\*** A. Present the results in the form of tree.

**\*** B. Construct two decision rules based on the resulting tree.

**\*** C. How many instances are correctly classified by the model as class “a = cheap”? (Hint: see the confusion matrix at the end of the window output)