**Homework assignment in Week 7**

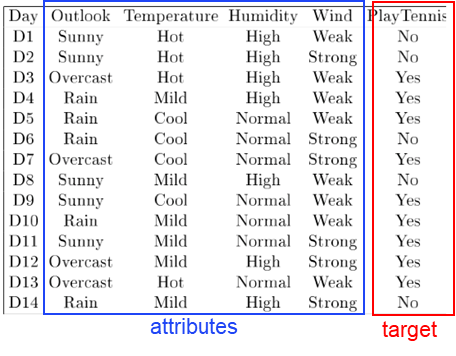
* Due Friday (17:00, 19th, Nov.), submit electronically to Manaba+R.
* File name: Q1 and Q2 🡪YourStudentID\_W07\_Q12.pdf and Q3🡪YourStudentID\_W07\_Q3.py

Q1: For a classification problem with 4 classes (A,B,C,D), calculate the Entropy of a set which has 4 instances of A class, 2 instances of B class, 4 instances of C class, and 3 instances of D class.

The calculation process must be included in the answer.

Q2: Imagine you play tennis, and you invite your friend. Your friend sometimes comes to join but sometimes not. For your friend, it depends on a number of factors, for example, weather, temperature, humidity, and wind. Please use the right dataset to build a decision tree which can predict whether or not your friend will join you to play tennis.

You must list the calculation process (to build the decision tree) and plot the decision tree.



Q3:

Programming in python to build logistic regression model for the following sample data.

• This is a binary classification problem because y has two values (0 or 1), and X (feature) has 4 dimensions.

• Build a logistic regression model to project X to y (classify X into two categories: 0 or 1).

• Your source code should follow the structure of the “Pseudocode of Logistic Regression” in the slide (modification is necessary).

• The initialization is: w1 = 0, w2 = 0, w3 = 1, w4 = 1, b = 0, Learning rate is 3.

• The number of iterations is 200.

Your code must include your own comments for all code. Go line-by-line. Comments in your program must be full sentences and reﬂect your understanding of the code.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| X | | | | y |
| 0.3 | 0.3 | 0.2 | 0.2 | 0 |
| 0.3 | 0.4 | 0.5 | 0.7 | 1 |
| 0.6 | 0.6 | 0.3 | 0.2 | 0 |
| 0.6 | 0.7 | 0.8 | 0.9 | 1 |