**Homework-1**

# Question 1

You are given the following training data:

| | WRITABLE | UPDATED | SIZE | CLASS |
|---|---|---|---|---|
| 1 | yes | no | small | text |
| 2 | yes | yes | large | text |
| 3 | no | yes | med | text |
| 4 | no | no | med | executable |
| 5 | yes | no | large | executable |
| 6 | no | no | large | executable |

The target attribute is CLASS.

### 1

Compute the decision tree generated by first considering WRITABLE, then UPDATED, then SIZE.

### 2

What tree is generated by using ID3?

### 3

How would these trees classify the following two instances:

| | WRITABLE | UPDATED | SIZE | CLASS |
|---|---|---|---|---|
| 1 | yes | yes | large | ? |
| 2 | no | no | small | ? |

# Question 2

You are given the following training data.

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| label | A | A | A | A | B | A | A | A | A | B | B | B | B | A | B | B | B | B |

**1.** What would be the classification of a test sample with $x = 4.2$ according to 1-NN ?

**Answer:** *A / B*

**2.** What would be the classification of a test sample with $x = 4.2$ according to 2-NN ?

**Answer:** *A / B*

**3.** What would be the classification of a test sample with $x = 4.2$ according to 3-NN ?

**Answer:** *A / B*

**4** Use "leave-one-out" cross validation to estimate the error of 1-NN. If you need to choose between two or more examples of identical distance, make your choice so that the number of errors is maximized.

**Answer:** $\frac{?}{18}$

**5** Use "leave-one-out" cross validation to estimate the error of 2-NN. Whenever you need to make a choice between equal distance data or determining a majority, make your choice so that the number of errors is maximized.

**Answer:** $\frac{?}{18}$

**6** Use "leave-one-out" cross validation to estimate the error of 3-NN. Whenever you need to make a choice between equal distance data or determining a majority, make your choice so that the number of errors is maximized.

**Answer:** $\frac{?}{18}$

**7** Use "leave-one-out" cross validation to estimate the error of 4-NN. Whenever you need to make a choice between equal distance data or determining a majority, make your choice so that the number of errors is maximized.

**Answer:** $\frac{?}{18}$

**8** Use "leave-one-out" cross validation to estimate the error of 17-NN. Whenever you need to make a choice between equal distance data or determining a majority, make your choice so that the number of errors is maximized.

**Answer:** $\frac{?}{18}$