

Comparing Classifiers - Exercise

Problem 1: Given the training dataset presented on Table 1, classify the following record using k-Nearest Neighbors (K=1; Euclidian Distance) and Naïve Bayes classification.

outlook	temperature	humidity	windy	play
overcast 3	cool 3	normal 2	true 2	?

Table 1: Training Data from "weather.nominal" data set

	outlook	temperature	humidity	windy	play
1	sunny 1	hot 1	high 1	false 1	no 1
2	sunny 1	hot 1	high 1	true 2	no 1
3	rainy 2	mild 2	high 1	false 1	yes 2
4	rainy 2	cool 3	normal 2	true 2	no 1
5	sunny 1	cool 3	normal 2	false 1	yes 2
6	sunny 1	mild 2	normal 2	true 2	yes 2
7	overcast 3	mild 2	high 1	true 2	yes 2
8	overcast 3	hot 1	normal 2	false 1	yes 2
9	rainy 2	mild 2	high 1	true 2	no 1
10	rainy 2	cool 3	normal 2	false 1	yes 2
11	sunny 1	mild 2	high 1	false 1	no 1
12	overcast 3	hot 1	high 1	false 1	yes 2
13	rainy 2	mild 2	normal 2	false 1	yes 2

Problem 2: The "J48 pruned tree" for the training data set presented on Table 1 was obtained using the WEKA software. The decision tree is given below.

```

Outlook = sunny
|  humidity = high: no (3.0)
|  humidity = normal: yes (2.0)
outlook = overcast: yes (3.0)
outlook = rainy
|  windy = true: no (2.0)
|  windy = false: yes (3.0)
    
```

- How the test record from Problem 1 is classified using this decision tree?
- ✓ What are the classification rules that can be obtained from this tree?