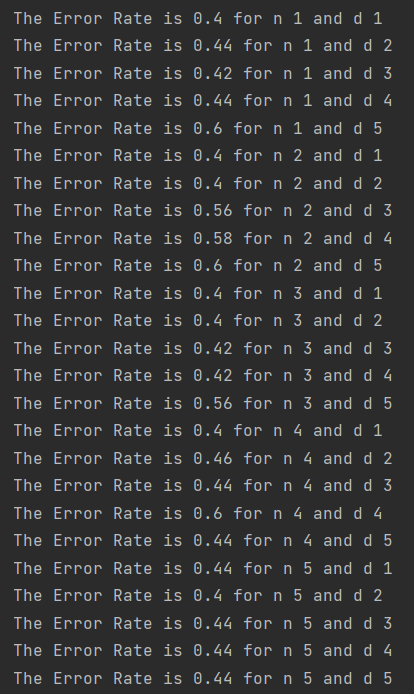
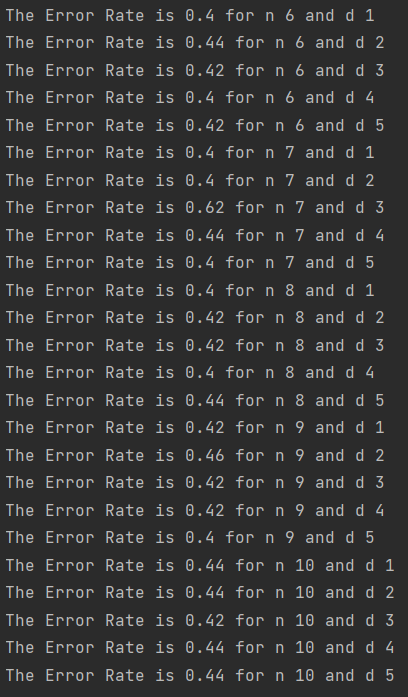
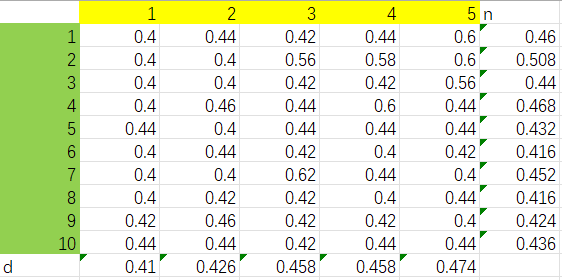
**Related file: XU\_YUHAN\_Assign\_11\_forest.py&attempt1\_forest.xlsx&Error\_Rate.pdf**

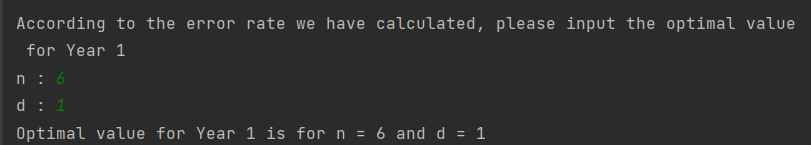
**Assignment 11 Random Forest**

**Notice: Because all the classifiers are random, all the results come from the one attempt**

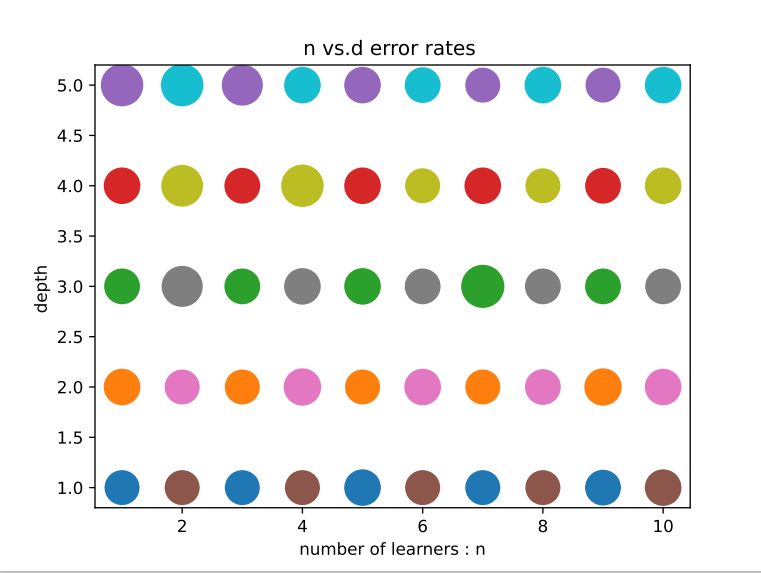
## **1. take N = 1, . . . , 10 and d = 1, 2, . . . , 5. For each value of N and d construct a random tree classifier (use ”entropy” as splitting criteria - this is the default) use your year 1 labels as training set and compute the error rate for year 2. Plot your error rates and find the best combination of N and d.**

****

****

****

According to the output, we can find that the optimal value of n is 6 or 8, and the d is 1. We use n =6 and d =1 to find the answer of the remaining questions



## **2. using the optimal values from year 1, compute the confusion matrix for year 2**

## 

| TP | 10 |
| --- | --- |
| FP | 11 |
| FN | 14 |
| TN | 17 |

## 

## **3. what is true positive rate and true negative rate for year 2?**

## 

| TPR | 0.5484 |
| --- | --- |
| TNR | 0.4762 |

## **4. implement a trading strategy based on your labels for year 2 and compare the performance with the ”buy-and-hold” strategy. Which strategy results in a larger amount at the end of the year?**

Because the second year is a big drop compared to the first year,it is impossible to hold for a long time.

Therefore, the strategies we can take are simple, all short-term actions. Buy when it falls compared to the previous day,

and sell when it rises compared to the previous day,so as to maximize the benefits.

Compared to buy-and-hold, my strategy will result in a larger amount at the end of the year.