MAL1 Exercises 5

(These exercises will take a long time by hand. Try utilizing Python, Excel, or similar.)

# Exercise 5.1 - Decision Trees

Consider the following data set

|  |  |  |
| --- | --- | --- |
| x1 | x2 | y |
| 3 | 1 | A |
| 3 | 2 | B |
| 3 | 1 | B |
| 4 | 2 | B |
| 5 | 1 | A |

(y is the target)

1. Compute the Gini coefficient for the data set
2. Create a decision tree of max\_depth=1 for the data set

# Exercise 5.2 – Random Forest

Use the data set from exercise Using a source of random numbers (Python, random.org, ...) to create a bag of the same size as the original data set.

# Exercise 5.3 – Gradient Boosted Regressor

Consider the following data set

|  |  |  |
| --- | --- | --- |
| x1 | x2 | y |
| 3 | 1 | 4 |
| 3 | 2 | 7 |
| 3 | 1 | 6.5 |
| 4 | 2 | 5.3 |
| 5 | 1 | 4.6 |

(y is the target)

1. Compute the value (mean) and MSE for the data set
2. Create a decision tree (regressor) of max\_depth=1 for the data set
3. Create the residual data set for the decision tree
4. Using learning\_rate = .1 create a decision tree (regressor) for the residual data set
5. What does your ensemble output on x1=4, x2=1?