Optimal decision tree pipeline: proof of concept.

Phases:

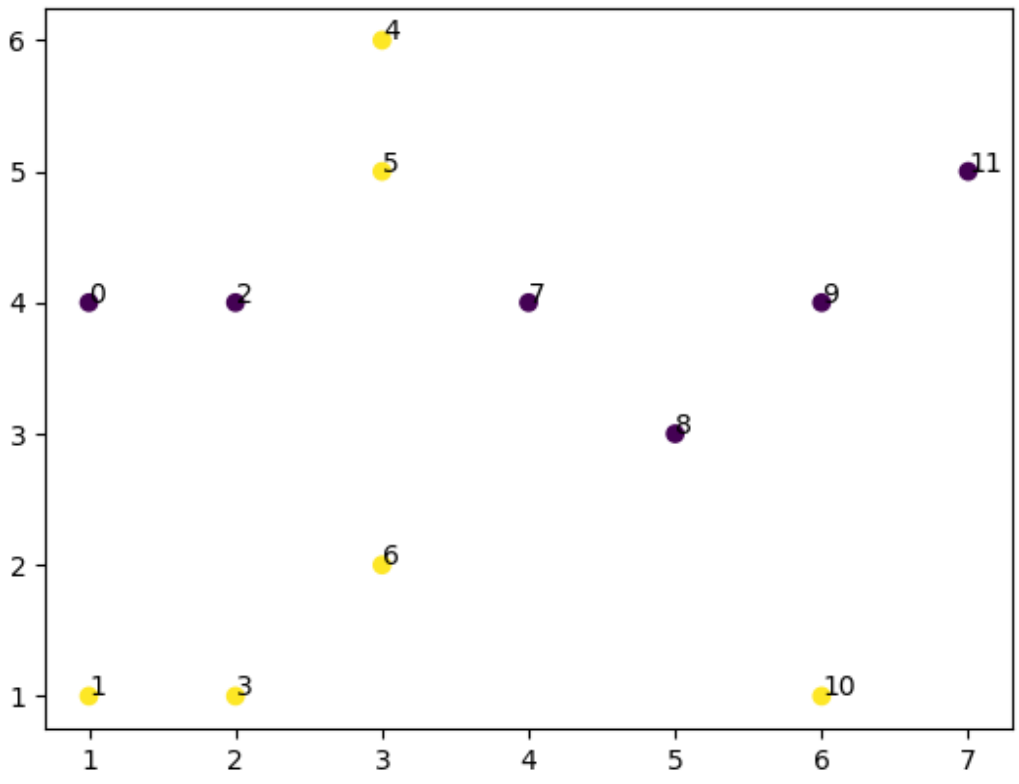
1) Maximal hyperbox clustering

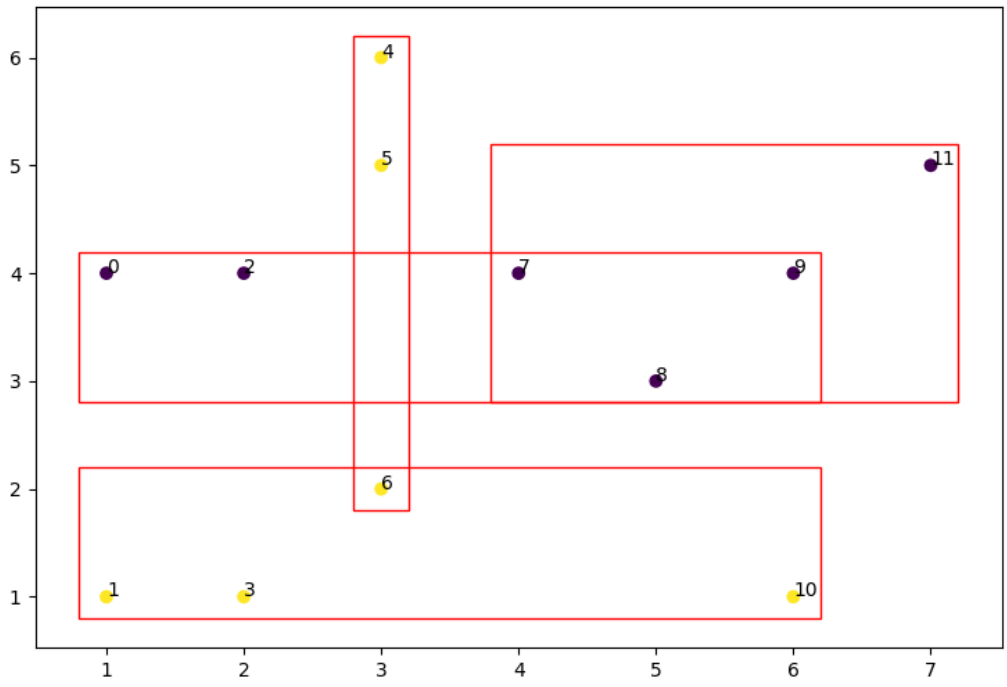
2) Minimal cutset

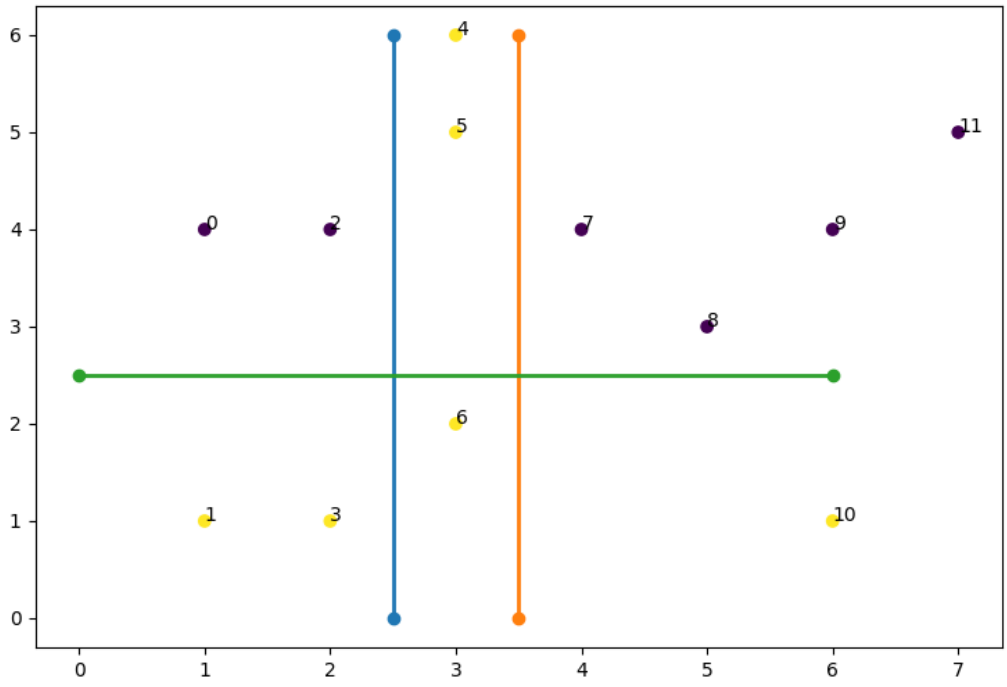
3) Minimal decision tree

Here phases 1 and 2. Five testsets: four tiny artificial (few points, 2 dimensions), fifth is iris dataset (https://archive.ics.uci.edu/dataset/53/iris), attributes only Sepal Width and Petal Length, well known to separate iris setosa from the other two species (trivial).

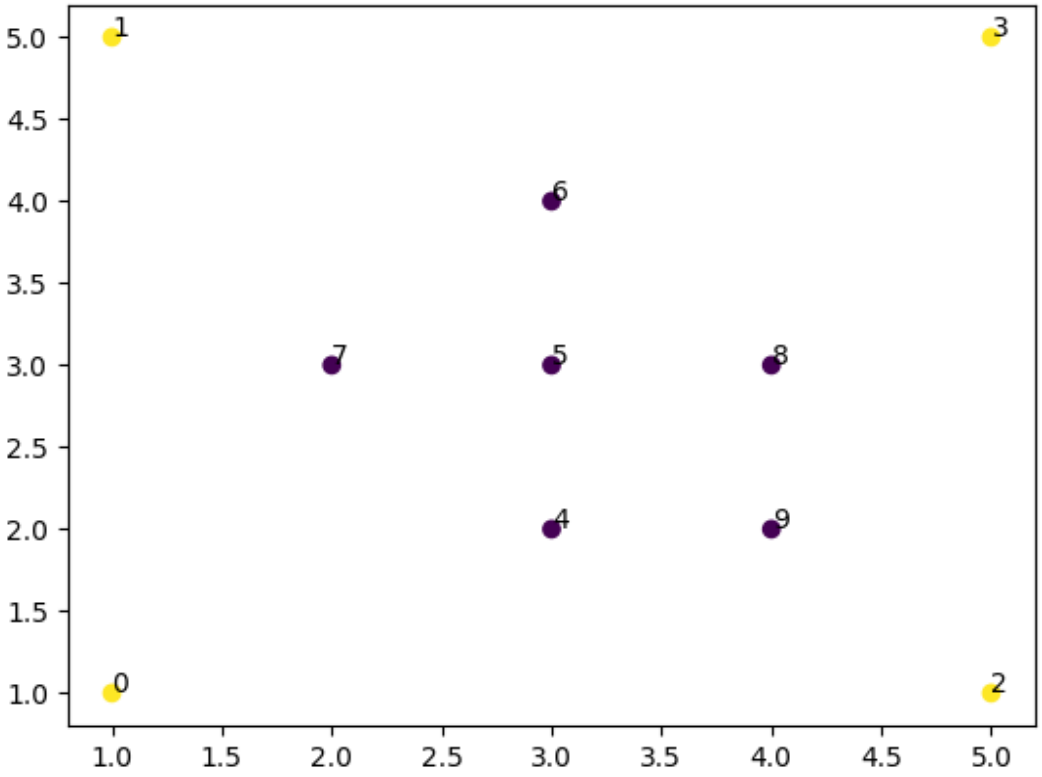
Test1

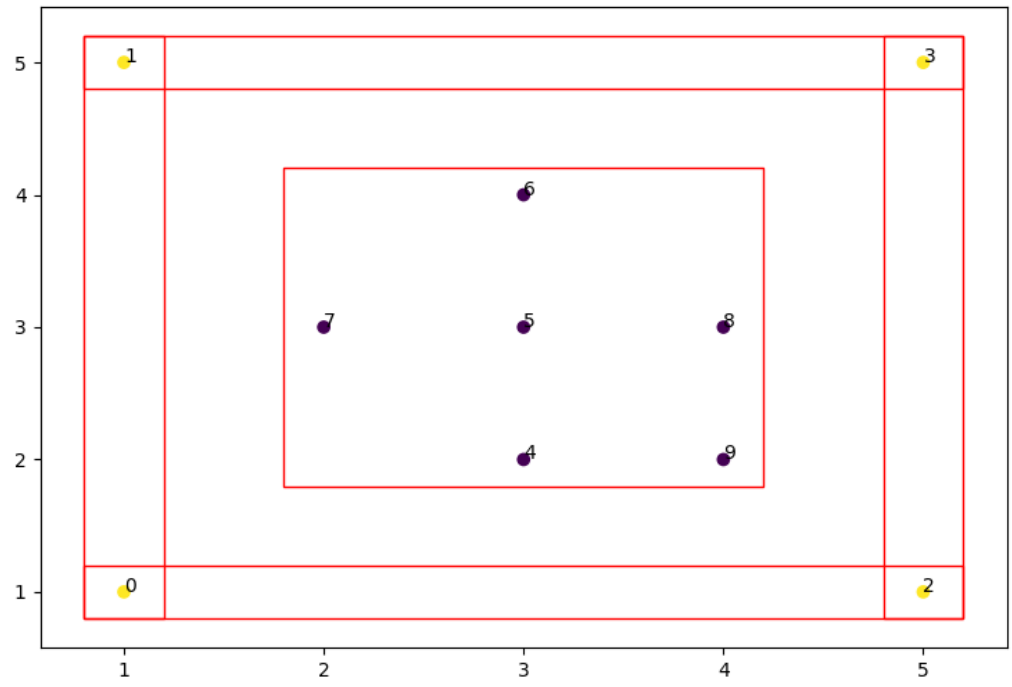


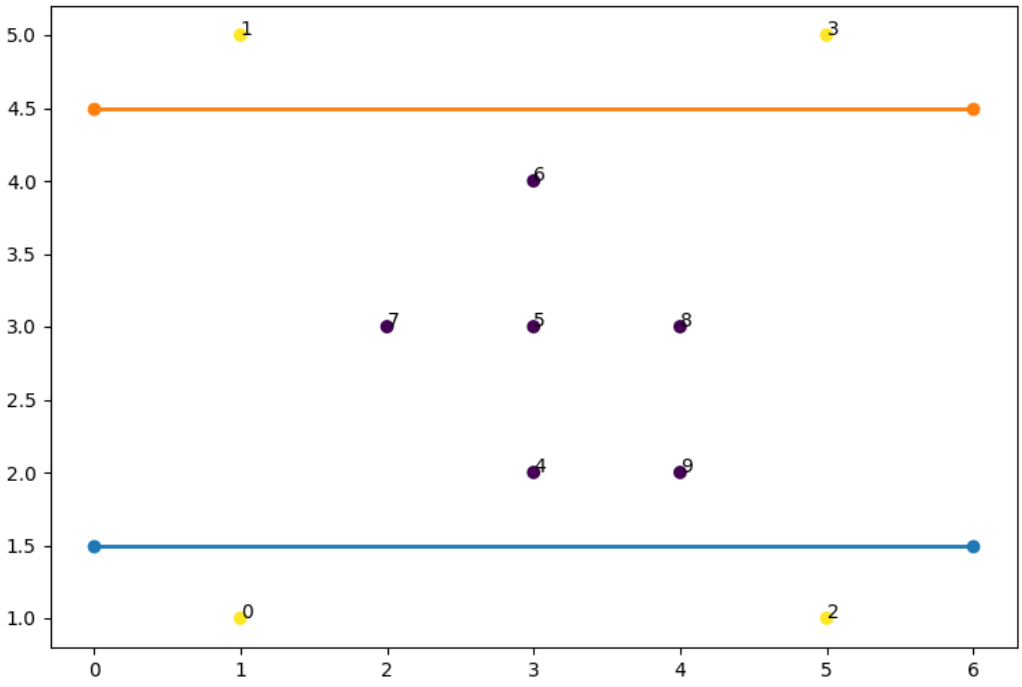

0
1
2
3
4
5
6
7
8
9
10
11


0
1
2
3
4
5
6
7
8
9
10
11

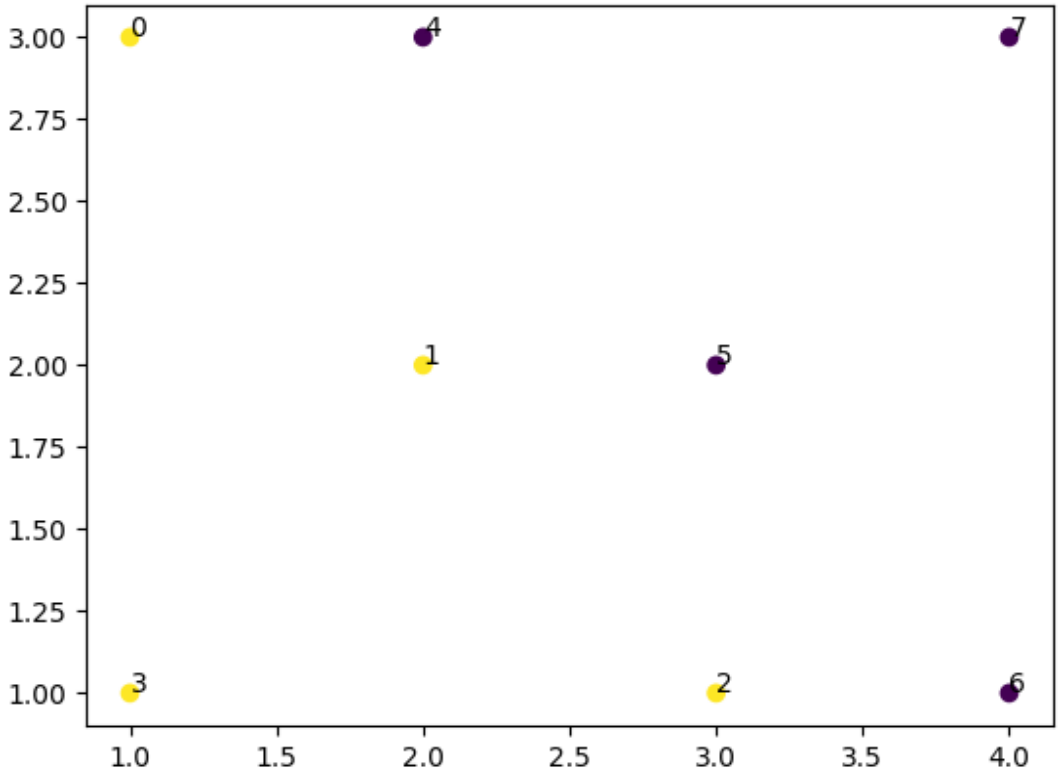
Test2

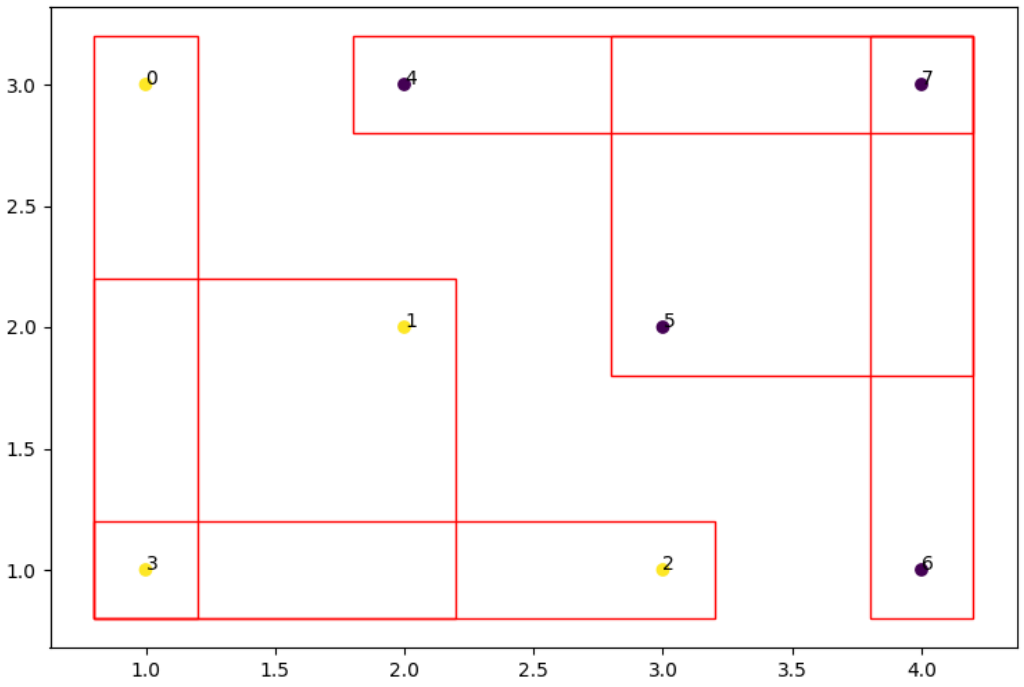


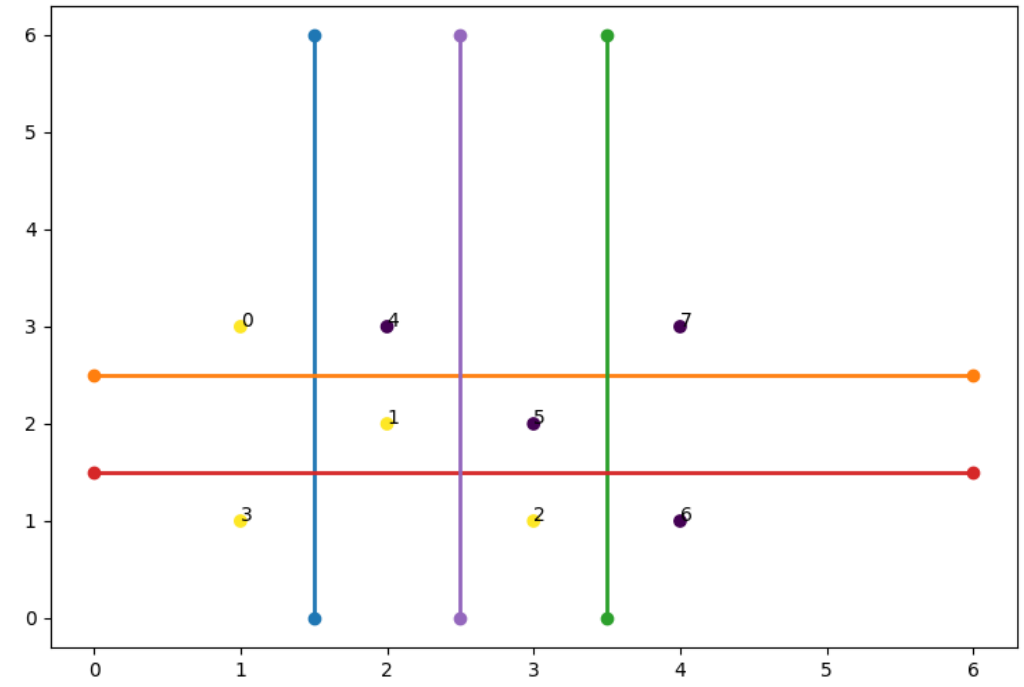

1
3
6
7
5
8
4
9
0
2
5
4
3
2
1
1
2
3
4
5


1
3
6
7
5
8
4
9
0
2
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
1
2
3
4
5
6

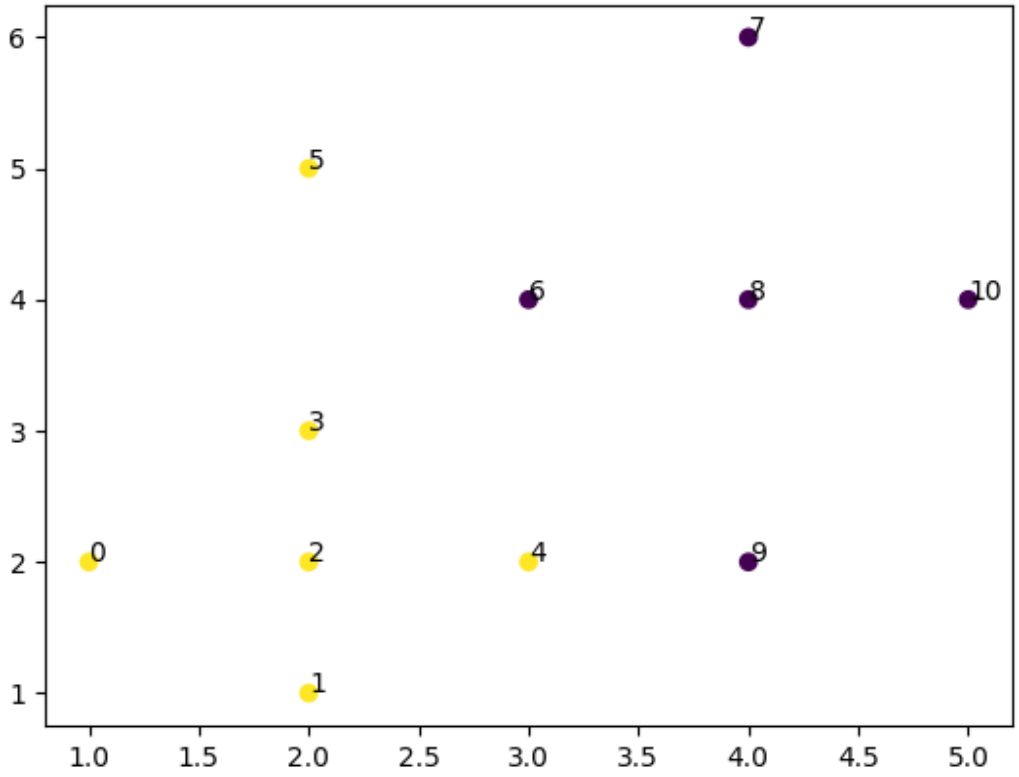
Test 3

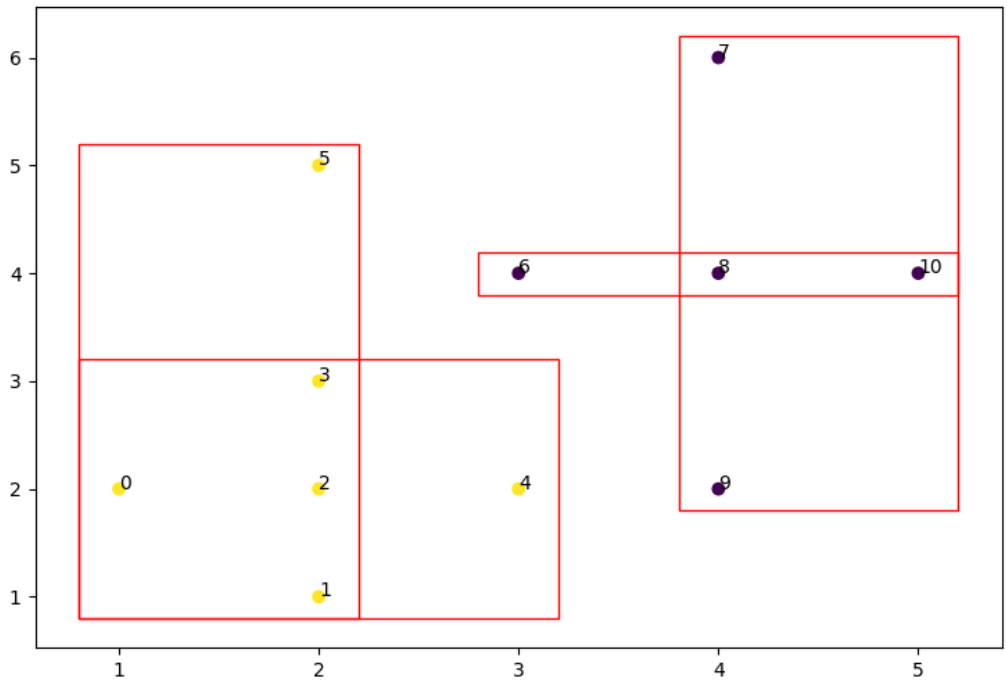


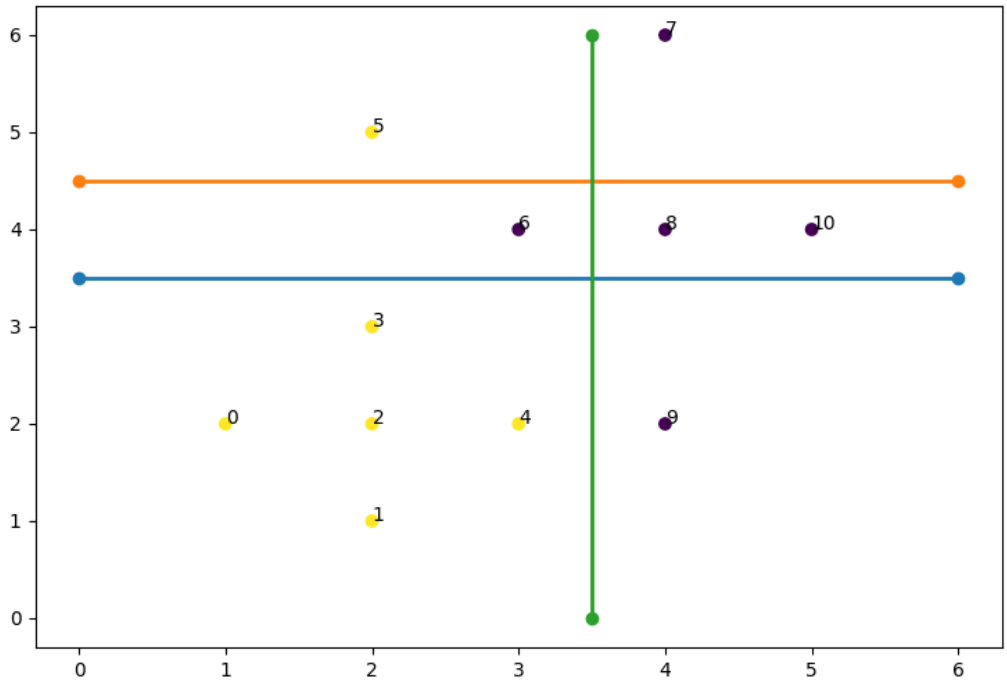
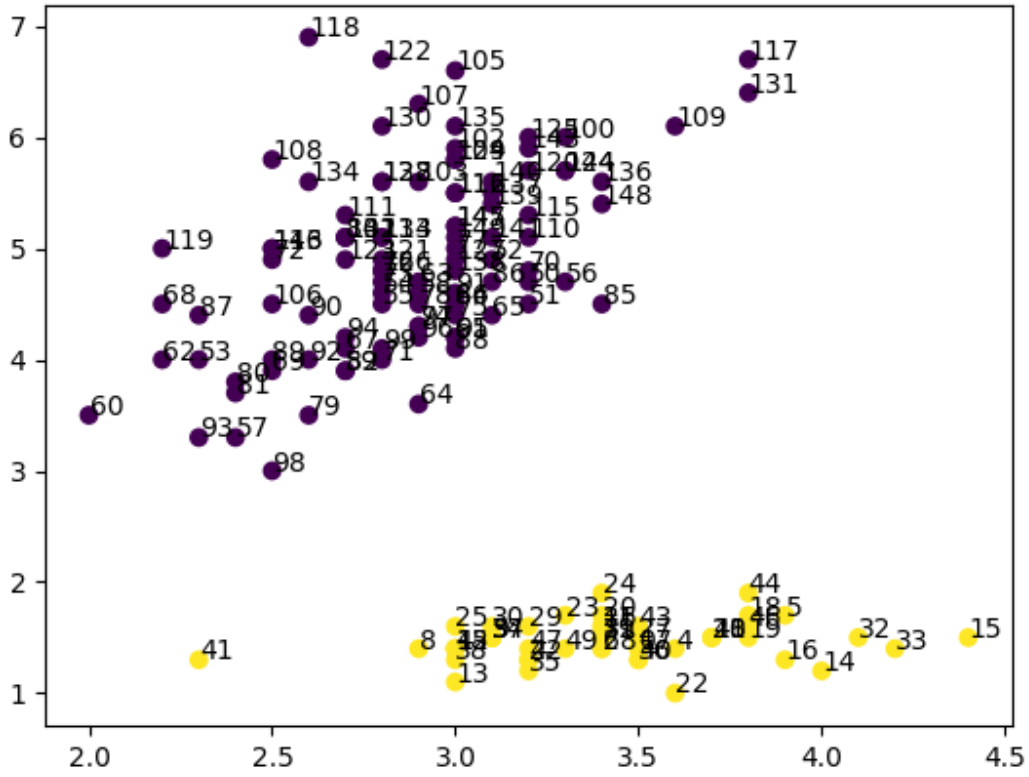
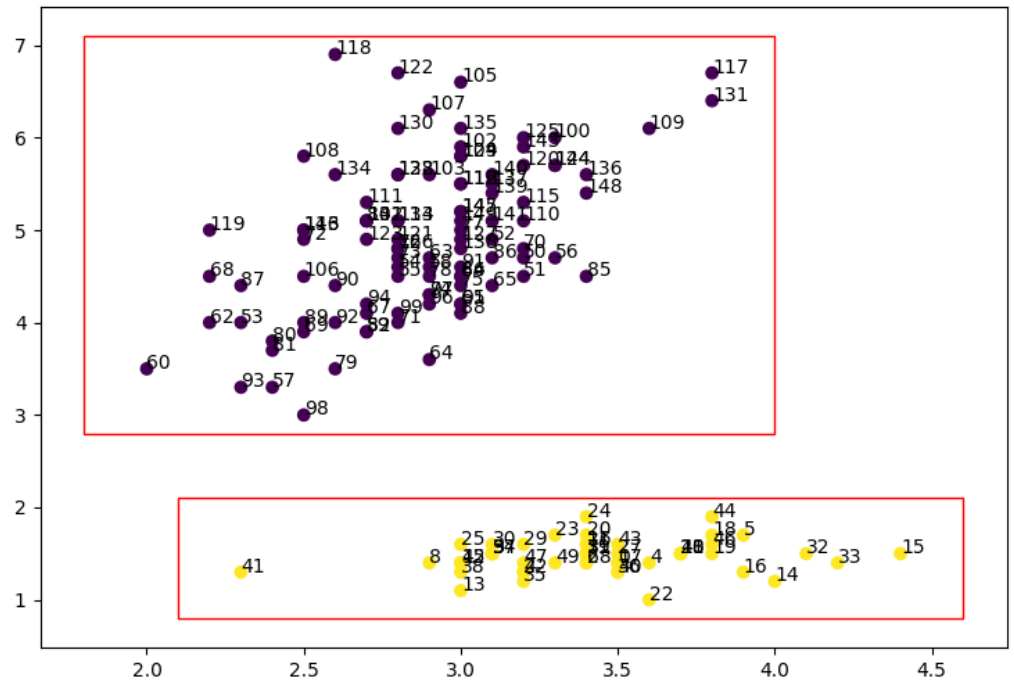
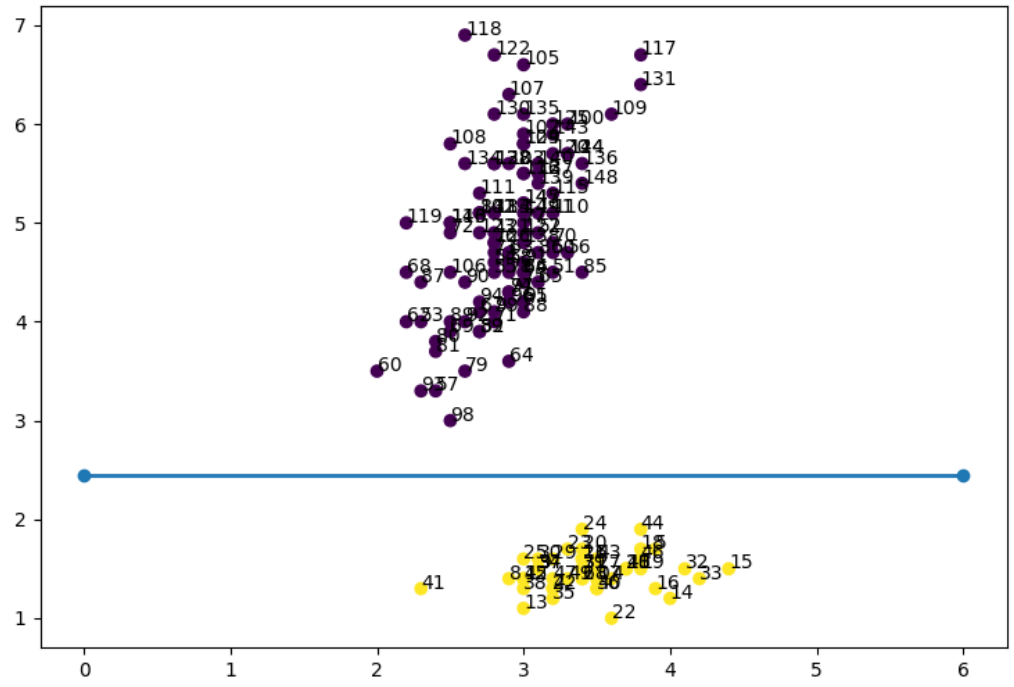

0
4
7
1
5
3
2
6
3.0
2.5
2.0
1.5
1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0


0
4
7
1
5
3
2
6
6
5
4
3
2
1
0
0
1
2
3
4
5
6

Test 4




7
5
6
8
10
3
0
2
4
9
1
1
2
3
4
5


7
5
6
8
10
3
0
2
4
9
1
0
1
2
3
4
5
6

Iris, dimensions: SepalWidthCm, PetalLengthCm