readme

Devin Johnson

January 24, 2020

Problem 1

- mallet import-svmlight –input train.vectors.txt –output train.vectors
  mallet import-svmlight –input test.vectors.txt –output test.vectors –use-pipe-from train.vectors
  vectors2classify –training-file train.vectors –testing-file test.vectors –trainer DecisionTree
- Summary:
  train accuracy mean = 0.6377777777777778
  stddev = 0.0
  stderr = 0.0
  Summary:
  test accuracy mean = 0.5233333333333333
  stddev = 0.0
  stderr = 0.0

Problem 2

Depth	Train	Test
1	0.45	0.42
2	0.52	0.53
4	0.64	0.52
10	0.75	0.60
20	0.86	0.68
50	0.97	0.70

We can draw the conclusion that the further the depth of the tree goes, the more accurate our training and testing is. This makes sense because we will be able to segment the data into better classification regions.

Problem 4

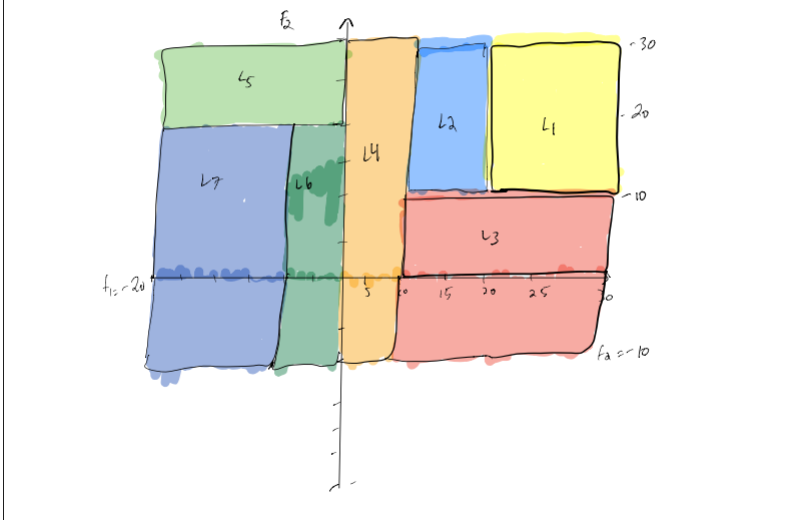

Table 2 (0.0):

Depth	Train	Test	Time
1	0.45	0.41	0.19
2	0.52	0.53	0.49
4	0.64	0.53	1.57
10	0.75	0.61	10.71
20	0.84	0.66	24.44
50	0.97	0.67	43.57

Table 3 (0.1):

Depth	Train	Test	Time
1	0.45	0.42	0.30
2	0.52	0.53	1.05
4	0.60	0.54	0.93
10	0.60	0.54	3.34
20	0.60	0.54	0.93
50	0.60	0.54	1.84