Ling571

Zoe Fang

**Q1.**

1. mallet import-file --input examples/train.vectors.txt --output train.vectors

mallet import-file --input examples/test.vectors.txt --output test.vectors --use-pipe-from train.vectors

vectors2classify --training-file train.vectors --testing-file test.vectors --trainer DecisionTree > dt.stdout 2>dt.stderr

cat dt.stdout

1. Training acc: 0.1781

Testing acc: 0.1

**Q2.**

1. Table 1

|  |  |  |
| --- | --- | --- |
| Depth | Training accuracy | Test accuracy |
| 1 | 0.1393 | 0.1033 |
| 2 | 0.1419 | 0.12 |
| 4 | 0.1781 | 0.1 |
| 10 | 0.6285 | 0.1133 |
| 20 | 0.9970 | 0.1367 |
| 50 | 1 | 0.1367 |

1. Decision trees yield higher training accuracies as depth increases; however, there is little correlation between depth and test accuracies.

Q3.

Table 2 min\_gain = 0

|  |  |  |  |
| --- | --- | --- | --- |
| Depth | Training Acc | Test Acc | CPU |
| 1 | 0.453 | 0.417 | 5 |
| 2 | 0.521 | 0.53 | 6 |
| 4 | 0.638 | 0.527 | 8 |
| 10 | 0.751 | 0.593 | 11 |
| 20 | 0.854 | 0.673 | 27 |
| 50 | 0.963 | 0.693 | 39 |

Table 3min\_gain = 0.1

|  |  |  |  |
| --- | --- | --- | --- |
| Depth | Training Acc | Test Acc | CPU |
| 1 | 0.601 | 0.54 | 4 |
| 2 | 0.52 | 0.53 | 4 |
| 4 | 0.601 | 0.54 | 5 |
| 10 | 0.601 | 0.54 | 7 |
| 20 | 0.601 | 0.54 | 7 |
| 50 | 0.601 | 0.54 | 8 |

A picture containing diagram

Description automatically generatedQ4.

Notes:

Scripts should run to completion and yield expected results.