**ML - HW 2 – NXN141730**

**Dataset 1**

**Accuracy on the test set for decision trees:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| L | K | Information Gain | | Variance Impurity | | Random Attribute | |
| Before Pruning | After Pruning | Before Pruning | After Pruning | Before Pruning | After Pruning |
| 10 | 0 | 75.85 | 75.85 | 78.0 | 78.0 | 55.45 | 55.45 |
| 10 | 5 | 75.85 | 76.05 | 78.0 | 78.45 | 59.2 | 59.4 |
| 10 | 10 | 75.85 | 75.95 | 78.0 | 78.95 | 53.55 | 53.95 |
| 10 | 15 | 75.85 | 75.8 | 78.0 | 78.15 | 52.55 | 53.15 |
| 10 | 20 | 75.85 | 77.35 | 78.0 | 78.8 | 53.5 | 53.5 |

**Tree Parameters:**

|  |  |  |
| --- | --- | --- |
|  | Average Depth | Number of nodes(Leaf + Internal) |
| Tree constructed using ID3(Entropy) | 8 | 138+137 = 275 |
| Tree constructed using ID3(Variance Impurity) | 9 | 193+192 = 385 |
| Tree constructed using Random Attribute Selection | 3 | 578+576 = 1154 |

**Dataset 2**

**Accuracy on the test set for decision trees:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| L | K | Information Gain | | Variance Impurity | | Random Attribute | |
| Before Pruning | After Pruning | Before Pruning | After Pruning | Before Pruning | After Pruning |
| 10 | 0 | 72.33 | 72.33 | 69.0 | 69.0 | 54.0 | 54.0 |
| 10 | 5 | 72.33 | 72.33 | 69.0 | 68.5 | 53.67 | 53.83 |
| 10 | 10 | 72.33 | 72.33 | 69.0 | 69.33 | 56.5 | 57.0 |
| 10 | 15 | 72.33 | 73.67 | 69.0 | 69.0 | 61.67 | 61.17 |
| 10 | 20 | 72.33 | 72.33 | 69.0 | 67.83 | 51.67 | 52.5 |

**Tree Parameters:**

|  |  |  |
| --- | --- | --- |
|  | Average Depth | Number of nodes(Leaf + Internal) |
| Tree constructed using ID3(Entropy) | 8 | 142+141 = 283 |
| Tree constructed using ID3(Variance Impurity) | 9 | 209+208 = 417 |
| Tree constructed using Random Attribute Selection | 3 | 589+587 = 1176 |