**Machine Learning Coursework 2 Report**

**Group B members:**

|  |  |
| --- | --- |
| Name | Student ID |
| Koo Wai Kun | 025254 |
| Darren Cheng Sum Jun | 025299 |
| Ong Kok Ting | 025288 |
| Loh Jin Xian | 016763 |

**Table of content**

|  |
| --- |
| 1. Result output   1.1 Decision tree for the whole data set  1.2 Cross validation classification result   1. Questions a   2.1 What is pruning  2.2 Node that will be pruned from the decision tree  2.3 Difference between an original and a pruned tree   1. Question b - Using decision tree for multiclass problem |
|
|

**1.1 Decision Tree for the whole data set**

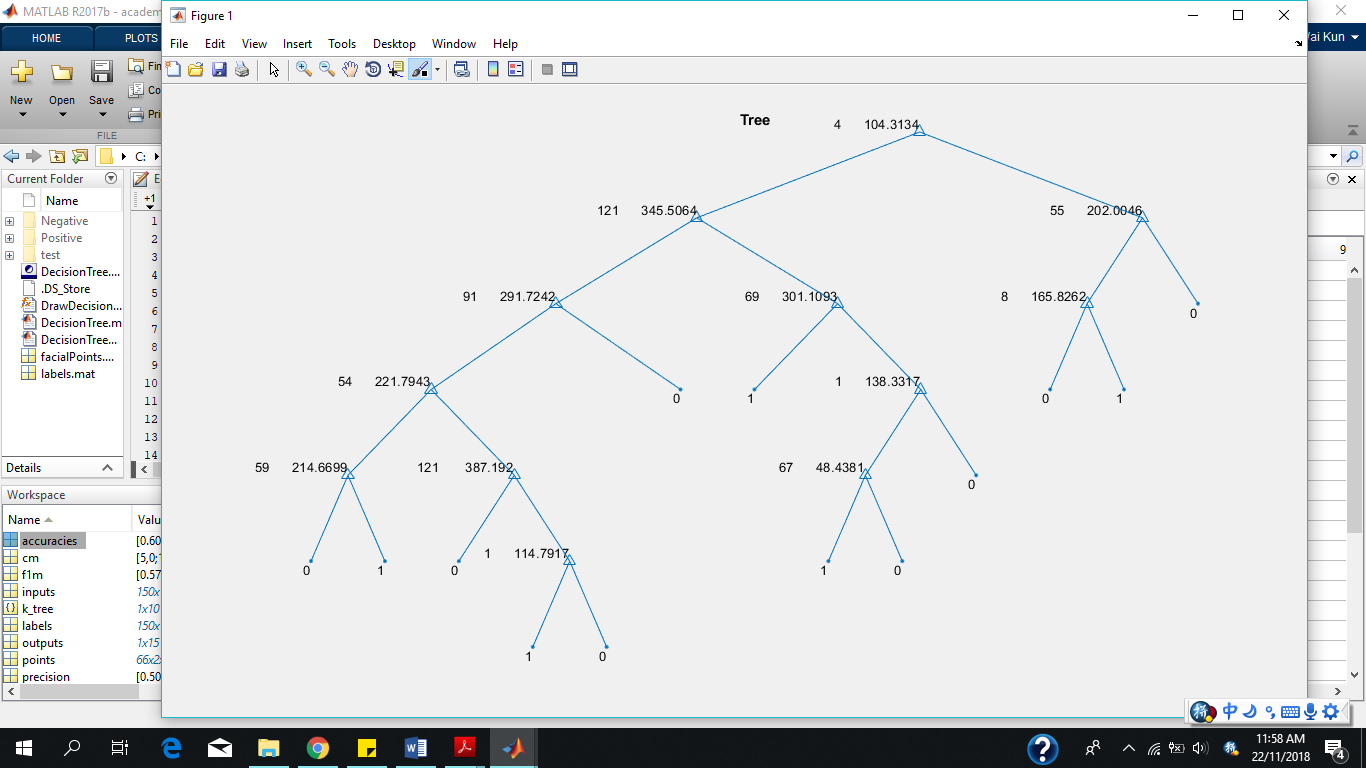


Diagram 1.1 Tree.op shows the best feature and best threshold chosen to split the tree.

**1.2 Cross Validation Classification Results**

|  |  |  |  |
| --- | --- | --- | --- |
| **Fold(k)** | **Recall Rate(%)** | **Precision Rate(%)** | **F1 measure(%)** |
| 1 | 66.67 | 50 | 57.14 |
| 2 | 80 | 80 | 80 |
| 3 | 100 | 100 | 100 |
| 4 | 100 | 71.43 | 83.33 |
| 5 | 75 | 60 | 66.67 |
| 6 | 100 | 100 | 100 |
| 7 | 80 | 50 | 61.54 |
| 8 | 80 | 100 | 88.89 |
| 9 | 66.67 | 66.67 | 66.67 |
| 10 | 100 | 83.33 | 90.91 |

**2. Questions Section**

**2.1 What is Pruning?**

In the realm of machine learning and data mining, there is a technique that is deeply associated with decision tree which is called pruning. Just like the traditional way of pruning a physical tree by the removal of certain branches, buds and roots, pruning a decision tree reduces the size of a decision tree resulting in increased predictive power due to the lower possibility of overfitting. This is done by removing nodes from the decision tree that do not contribute power to classify instances.

Pruning can be done in a top down fashion which will traverse the nodes and trim subtrees starting at the root, or in a bottom up fashion which will start with the leaf nodes. Overfitted decision tree being built can also be avoided by using pre-pruning which stop growing the tree before it perfectly classifies the training set, or by using post-pruning which allows the tree to perfectly classify the training set before being pruned. There are many pruning algorithms available but the 2 most popular algorithm are *Reduced error pruning* and *Cost complexity pruning.*

**2.2 Node that will be pruned in the decision tree**

The node that will be pruned from the decision tree is the twig of the decision tree that has the least information gain when the desired number of leaves in the tree exceeds a desired number via pruning by information gain.

**2.3 Difference between an original and a pruned tree**

The pruned tree will be simpler than the original tree in terms of complexity. The size of the pruned tree will also be smaller. At the same time, it will has a slight increase of accuracy and perform better in the whole distribution.

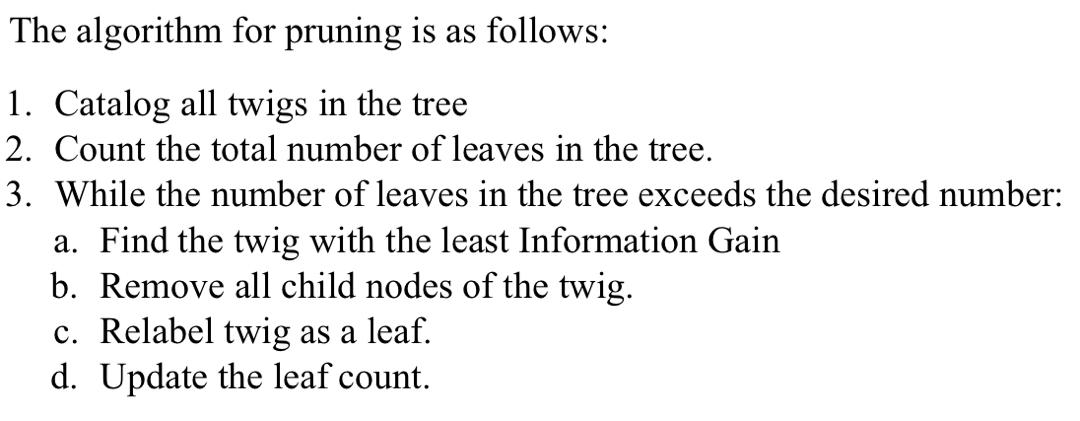


Diagram 2.1 Pruning via Information Gain

**3 Using decision trees for multiclass problem**

Just as decision trees can be used in binary classification, they can also be used in multiclass problems. Instead of using 0 and 1 for labels in binary classifications problems, we assign one unique integer for each basic emotion. Hence, integers 0,1,2,3,4,5 can be used to represent the six basic emotions.

At root node, we split the data into 2 subtrees after acquiring the best feature and best threshold, and split again at every node until it reaches the leaf nodes. We can also determine the range of threshold for 6 classes before doing the decision tree.

|  |  |
| --- | --- |
| Threshold/ X | Class label representing emotions |
| X : { 0 - x1 } | 0 |
| X : { x1 - x2 } | 1 |
| X : { x3 - x4 } | 2 |
| X : { x4 - x5 } | 3 |
| X : { x5 - x6 } | 4 |
| X : { x6 - x7 } | 5 |

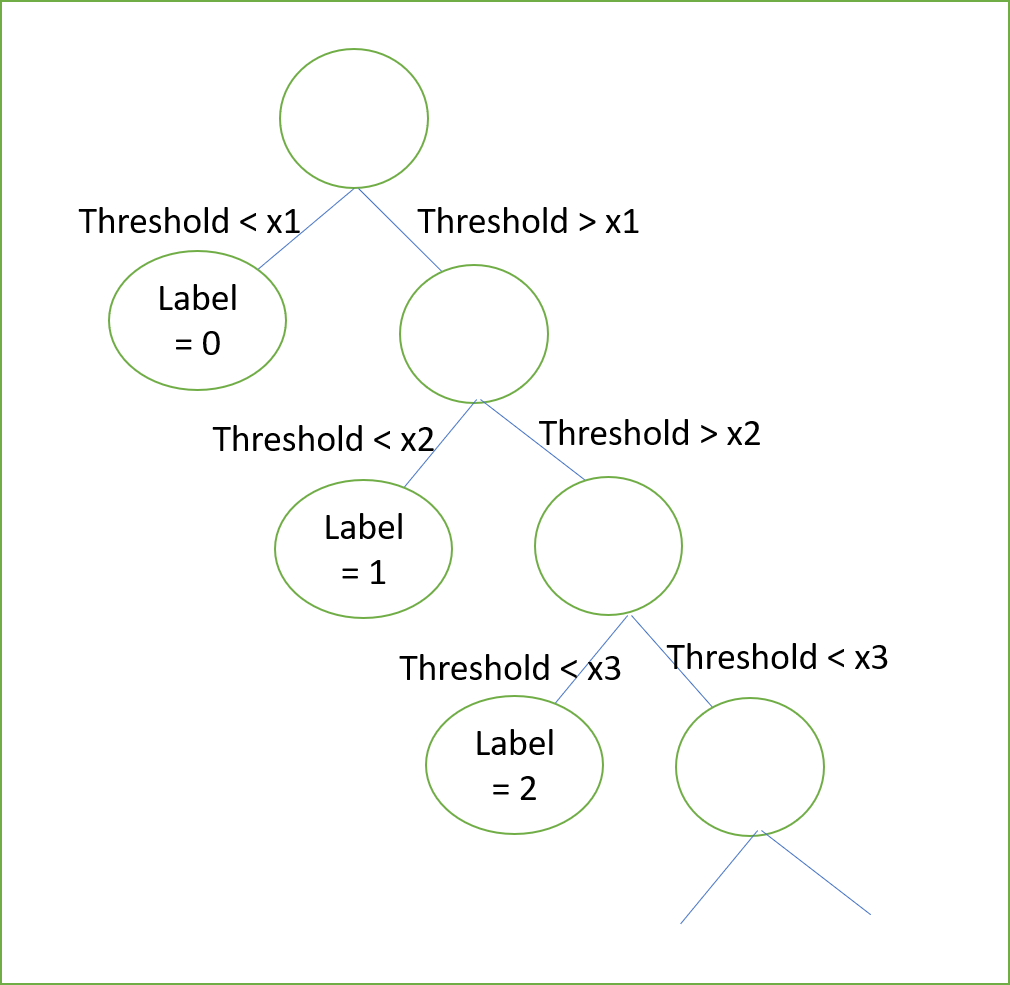


Diagram 3.1 Example of decision tree for multiclass problem