**Q1(4)**

**Yes, over fitting occurs in the asked decision tree. Over fitting occurs when the accuracy of “TEST SET” starts to go down while the accuracy of “TRAINING SET ” still increases.**

**So, there will be two points of over fitting for each type of Priority Queue.**

**IG as Priority Queue:**

**Node 27 with accuracy of 82.60%**

**IG \* No. Of Training Leaves as Priority Queue:**

**Node 16 with accuracy of 84.72%**

Q1 (5)

In general, we can see that for both type of queues the word features selected contains the very common words like **That, who, use, writes etc**. which may be in any label.

Also, there is a little difference in the selection of the word features by the queues. It seems to be that second queue is still making more fair selection than the first one. As in the decision tree for the first queue, we can observe that when it reaches “ARCHIVE ” Node it gives back the class of Label 1 and Label 2 by selecting only from “2 documents ” each. Which is not a very concrete selection.

On the other and, the second priority queue is giving the classification from the set of reasonable number of documents which is more efficient than the first one. The conclusion is that we cannot generalize from analyzing only, say, 2 documents we need more than that to give the generalization.

The word features selected by both the queues do not really make any sense. First, the words like THAT, USE, WHO etc. are most common words and do not classify perfect to any one label. These are not very distinctive words. Also, there is a point that these word features only make sense if they selected in particular order as described by decision tree. Independently, they are ambiguous.

The second Priority queue is better than the first one because by looking at the efficiency data the maximum efficiency of 2nd Priority Queue is 84.72% and reaches this efficiency fast as well i.e. at **Node 16.**But maximum efficiency of the first Priority Queue is 82.60% and reaches only at **NODE 27.**