Assignment 2

Shubham Jangid, Vivek Raj, Abhijeet Roy Vasu, Kushal Maheshwari, Kunal Saini

5th April 2023

1 Question 1

One effective method, although slightly expensive, is to use lookahead when implementing a decision tree. Instead of selecting the split that produces the maximum entropy reduction at the present moment, it is better to choose a split that will result in maximum entropy reduction two or three levels down the tree. For instance, In most cases a greedy split can sometimes lead to unsatisfactory results,Nevertheless, using a large lookahead can significantly increase the training time. Therefore, it is advisable to implement lookahead after reducing the number of words in the nodes to smaller numbers, and not to apply lookahead at the root, but rather a few levels below it.

In the dummy code we removed the random function for choosing query and instead we used entropy for choosing the word. We are choosing the word with maximum entropy

And when the number of words become very small then we considered all those words.

These are the only changes that we made to the Dummy code given by the instructor. The dummy code given was very inefficient as it was using random queries at each node and hence it was giving poor results. Other than this we removed a few attributes that were useless and in order to decrease the model size we removed these attributes.

$$E(S) = \sum_{i=1}^{c} - p_i \log_2 p_i$$