Question 2, Part B:

Output from validation code yields the following:

Information Gain accuracy for max_depth 2: 0.7402862985685071
Gini accuracy for max_depth 2: 0.7402862985685071

Information Gain accuracy for max_depth 3: 0.7402862985685071
Gini accuracy for max_depth 3: 0.7402862985685071

Information Gain accuracy for max_depth 5: 0.7402862985685071
Gini accuracy for max_depth 5: 0.7402862985685071

Information Gain accuracy for max_depth 11: 0.7402862985685071
Gini accuracy for max_depth 11: 0.7402862985685071

Information Gain accuracy for max_depth 17: 0.7402862985685071
Gini accuracy for max_depth 17: 0.7402862985685071

Question 2, Part C:

Since all accuracies were approximately the same, the following shows the tree with *criterion* Information Gain and *max_depth* of 5:

Note 1: the following values extracted from the vocabulary in the vectorizer for interpretation of the tree:
- X[5303] = "trump"
- X[4187] = "reject"
- X[5143] = "the"
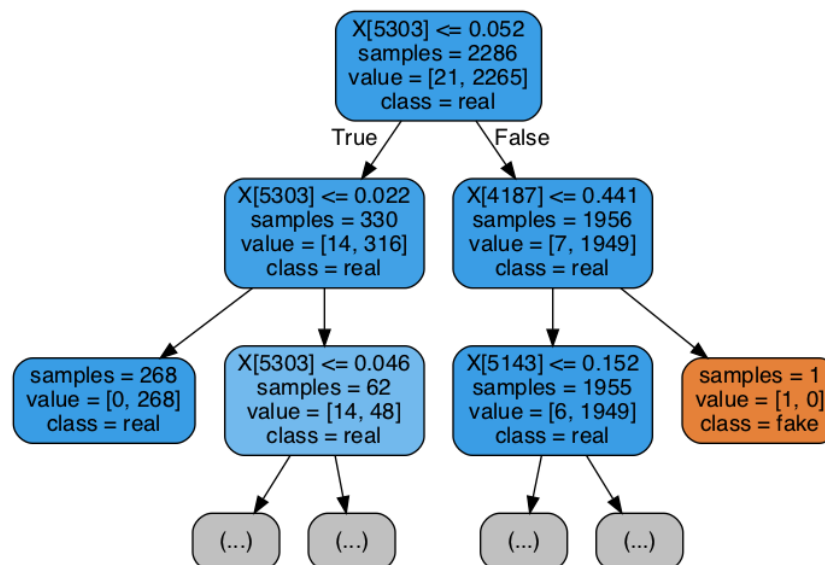
Question 2, Part D:

The following is the output for the word "trump" chosen to be the root:
- Information Gain for word "trump" is: 0.04370012447843627

Here are a few other sample keywords' computations for Information Gain:
- Information Gain for word "hillary" is: 0.047341783016359024
- Information Gain for word "debate" is: 0.08555126897517264
- Information Gain for word "the" is: 0.06649755612035199