1, step 1 is the program: it will take input from the user to get the value of training size: this input number should be some integer value S that is a multiple of 250, within bounds 250 ≤ S ≤ 1000. Then the program will choose the S distinct examples at random from the entire dataset. S is removed from the dataset so that any testing of the tree is not on the data used to train!

Step 2. Then the problem asks for the user to input training set increment I: this should be some integer value I ∈ {10,25,50} (that is, one of those three values only). For example, if I choose S=250 and I = 25, it will choose 250 examples from all data and put them aside. Then it will take 25 elements of those and use them to build tree.  Check the sample output for details: iteration.

* 2, Step 2, this program needs to build a tree. I am partially done with the part, but not sure why it is not the same as the sample output (maybe printout is different, see step 5). But this can be left at the end! If the program can pass and show same result in the output sample file (images, and also prediction accuracy after each, run, this is not priority. Can leave as it is now). The program needs to ask the user to choose using heuristic  method (shown in output file)

3. After the tree, the program will run through all remaining data and check accuracy (4 decimal places, 0.9999)

4. Iteration: program increases increment each time on the size of the set used to build tree.

5, Print out tree: attributes used for splitting at each node (both name and number). Also print value of attribute on each branch (single letter code or full name) . Leaves on tree should be edible or poison.

6, run the program twice each for two different heuristics (counting vs. entropy. Will this work? I only did entropy). With different set of training parameters (4 runs in total). Save results in two text files. One graph for each run.