I tried to resolve some of the current issues before sending this. MATLAB will allow me to specify categorical attributes, as well as specify a title for the variables, but as I originally thought, it will not be trivial to streamline this process. I get the sense that specifying categorical attributes and specifying titles will be resolved in tandem.

There are two basic types of problems to resolve: the generic and the specific. The specific are the more pressing issues, as we want to maintain a continuity of results with Dr. Kline and the Alzheimer’s-specific issues can be resolved as a step towards generic resolutions. I will first talk about the specific issues we are facing.

1.) We want to find predictor importance for the following tests:

a.) Healthy patients who either remain healthy or transition to AD [We have insufficient data for this correct?]

b.) Mild patients who either remain mild or transition to severe [Focus on this]

c.) Patients who wind up severe, having begun either as mild or severe [This is a party favor. Not really to focus on]

d.) We would also like to run a–c using the Tau-family variables, though this means removing many of the observations [This is important]

2.) We want to figure out, for a particular predictor, whether the prediction is positive or negative (e.g. does high BMI mean more advanced Alzheimer’s or less) [tree should tell us shouldn’t it?]

3.) Ideally, the program should be able to run categorical attributes and should have the titles of our attributes. I will meet with Juan to figure out how to tackle this problem.

4.) Once this is complete my code should spit out a picture/text view of the tree (with attribute names); I should specify predictor importance as well as the meaning (e.g. does high BMI correlate with advanced disease)

5.) I should also make a second package, which will complement the first package. The inputs are simply a confusion matrix (whose p-value we are testing) and the target attribute as a column vector. The output will be a rough p-value. Upon further reflection I realized that I can easily make this generic the first time around.

Finally, there is one issue we would like to resolve regarding how generic my code is.

6.) This will be where I send you a zip file of the code and you will need to have a data matrix to input [Let’s talk through this]

7.) I should send you a readme file that instructs the user on: how to specify the target, explain the guidelines for dealing with NULL cells (ignore row or column) and a default option [Yes, I want to see that.]