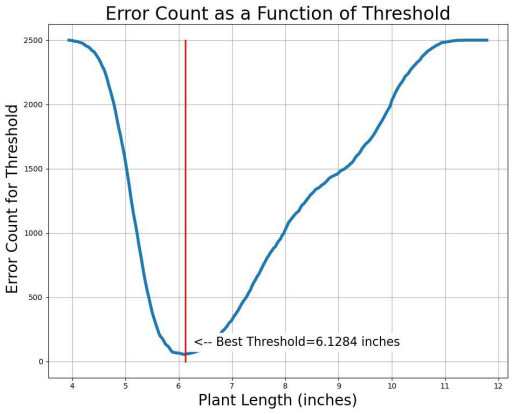
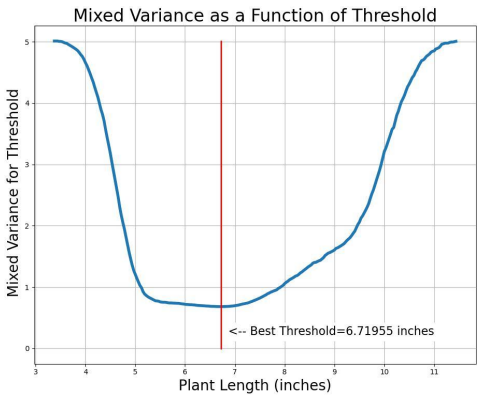
Threshold classifier: 6.1284
Threshold clustering: 6.71955

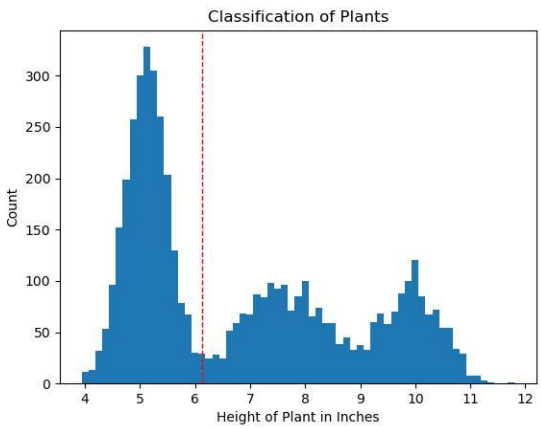
While for this homework I did not find two people to talk to about this (honestly didn't know how to get to know my classmates in this online class), I thought it would be an interesting experiment to try and explain these concepts used for the assignment to two of my friends outside of this course in a way that could be understood without knowledge in the materials of the course. It proved to be a helpful exercise, mainly as preparation for the midterm, as it forced me to really get the meaning behind why we are doing what we are doing to the data. If the point of Data Mining is at the end of the day giving and extrapolating meaning from data, then being able to do this task is a big necessity if I want to pursue a career in Data Mining.

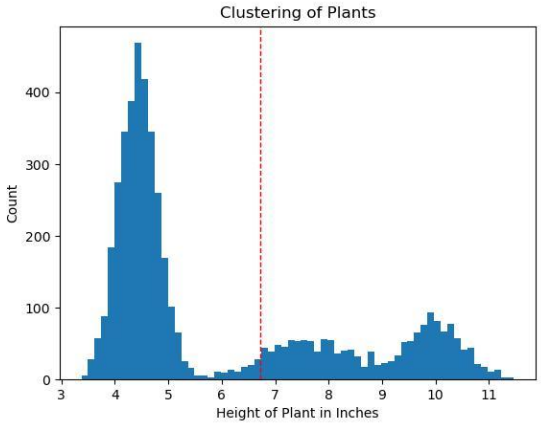
Did you notice anything about where this threshold classifier fell?
Interestingly enough, the classifier hit right around one of the relative min's of the histogram, which logically makes sense since most of class 1 would be located right around the first leftmost mode (grass is smaller than weeds).
Is the number of fall alarms the same as the number of false rejections?
Not likely, since we expect weeds to have a larger height that grass, there will still be a smaller amount of weeds that get classified as grass (class 1), and some smaller amount of grass that gets classified as not grass. Something to keep in mind is that there are also more than two classes, both of which tend to have a larger height than our target class.

If you ran Otsu's method on the clustering data, would you get the same result?
Because of the fact that there are more than two classes, even by coincidence I wouldn't expect the clustering to produce a similar threshold to the classifier.









The two classes with a greater average height increase the amount of false positives given by the clustering based on the weighted variance. Unless the clustering would create two thresholds for the three classes, the variance is always likely to be an inaccurate way to cluster the data into grass and not grass. The classifier, since it uses the classes to create the threshold, ends up much more accurate for the data in separating the grass from not grass. However it might be premature to say it is the best way to distinguish grass and not grass, as neither method has been tested on testing data. We can only compare how the two act on this set of data, where the classifying just simply works on the data itself.