EEE591 Additional Assignment

**NOTE: This assignment is for EEE591 students only. If you are in EEE419, then you do not turn this assignment in!**

In Project 1, you analyzed the ability to predict heart issues using a number of different methods. Copy your Project 1 part B solution to a new script called hw7.py. Your assignment is to use ensemble learning to see if you can get better results. In Project 1, you used 6 different methods. Here's what you need to do:

Step 1: Pick the 3 methods which got the best accuracy for you in Project 1 and use them to vote. Does your accuracy improve?

Step 2: Now add the 4th best method. Does accuracy improve? What do you decide to do in the case of a tie?

Step 3: Now add the 5th best method. Does accuracy improve?

Deliverables:

1. Your script, named hw7.py

2. Print the accuracy of each individual method as found in Project 1: e.g. Random Forest: 0.81  
NOTE: it's ok to print this as 81%.

3. Print the accuracy obtained via each of the steps above: e.g. Ensemble with three methods: 85%  
For 4 methods, print to the screen whether ties were counted as Yes or No.

Further notes:

Do NOT change your results from Project 1 unless you did not finish the project... in which case, you'll need to finish it.

For this assignment, use the **test** results, **not** the combined results.

Finding the vote outcome is simple: you'll have an array of answers from each method. You just need to add them! That is, in the CSV file, 1 meant no heart disease, 2 meant it was present. All you need to do is add the predicted value arrays for the methods you are using. Then, those entries that are greater than 1/2 between the number of methods and twice the number of methods have heart disease. One way to handle this is to convert them back to 0s and 1s via something like this:

results = np.where(sum\_of\_methods > thresh, 2, 1)

where you have set thresh equal to the midpoint. (For 3 methods, thresh would be 4.5. For 4 methods, does 6 count as a yes or a no? The answer will influence whether you use > or >= to compare with thresh. And for 5 methods, thresh would be 7.5.)