

### **Part 1 executive Summary**

There seems to be some strong correlations that can be used to predict heart disease. The thal parameter shows strong correlation with heart disease. Its not strong enough on its own to predict heart disease but with the aid of several other parameters we have a very good chance to predict heart disease with a high degree of accuracy. We may not need every parameter from this data to predict heart disease. Dest's and opt's have a fairly strong correlation with each other and may only produce noise in our predictions. All in all this effort is worth pursuing as it could very well likely aid doctors in diagnosing and help patients get better care.

### **Part 2 executive Summary**

Seems like we have had a breakthrough. The effort to has gone very well. After looking at the results it seems our best bet will be the random forest classifier. We achieved a high level of accuracy of 94% on our test data. We achieved 99% accuracy when testing our model on all the data. This can be a very strong tool for doctors. Some of our other classifiers performed fairly strong as well but did not have as high results on the full set of data. The perceptron for example may have achieved 95% accuracy on the test data but when the data only 83% accuracy on the full set of data. We may be able to achieve higher results if we tried a voting system with the kernel SVM and the Knearest neighbor classifiers as those classified stronger on the entire data sets as well. But the current recommendation is to go with the Random Forest classifier as the tool to aid the doctors.