Aaron Ballard

Data Mining Project #1

I used C++ to create my Relief Algorithm program. In creating my program, I used two datasets throughout writing to ensure I was getting correct results. One was from page 67, with 2 classes, 4 samples, and 2 features. This was the data I used the most since it had only four samples and was easy to calculate step by step to make sure my intermediate functions were correct. Although my final weight for the second feature came out different from the book, I saw that two of the distances I calculated myself were different than what the book used, so I assumed there was an error. The second set I used was from the pdf given in the assignment, which I got correct results for throughout the program’s functions. I then added more samples, more features, and added in another class, to make sure these changes would not produce errors, which they did not. The only obstacle I came across was in listing the features sorted by their weights. Although I can sort them and display the values of each weight, I have the weights in a vector, and I do not know how to show the sorted indexes for the values.