Robert Herriott
Naive Bayes Classifier
Machine Learning
Spring 2014
Undergraduate

I will start with the test results:

- 10 Bins: 96% Accuracy
- 20 Bins: 98% Accuracy
- 50 Bins: 91% Accuracy
- 100 Bins: 92% Accuracy

To run the program, just double click the JAR file. No special libraries or anything special needed, just the latest version of Java.

So I start by reading in the data, mostly reusing code from the KNN project. Then I build up my data cube using a method I wrote called FindBin which takes a min and max value, the value to be binned, and the desired number of bins, and returns the correct bin for that item. After the data cube is built I use this method again for each test fruit to find what bins its attributes belong in. I get all the values from those bins and get the fractions for each type of fruit as discussed in class. Then I multiply those across and assign the highest value to the test fruit.

To handle test instances outside the bin bounds I did the same thing you did, just stuck them into the end bins.

I didn't have any issues and the program runs as expected.