MUSI 6201 – Assignment 5

Ananya Bhardwaj

A.2

1. Tiebreaking rules in kNN:

This is when there are multiple equidistant neighbors, and not all of them can be selected since k is smaller than the remaining number of neighbors.

My approach to break the tie is to basically go with random choice. However, this necessitates that the training data and labels be shuffled before I do this, to avoid any bias due to the order in the training data, which is true in our case.

So, I shuffle both the training data and labels in unison, and let fate decide!

It’s a reasonable approach, not the best, but it’s hard to say what is the best.

Expanding to include all neighbors till there is no tie? It is almost certainly better, since the more the neighbors, the closer to bayes’s risk a kNN would approach. But, is it “k”NN anymore, if the k keeps changing? And all the extra computation to an already slow inference method? I prefer randomness.

B.2

1. Best feature was Spectral Flux Std in my testing. But since I was shuffling my data, on one occurrence I did get Spectral Crest Mean as the best (k = 3, num\_folds=3).

C.2

1. Best feature set was all 10 features in the following order:

[8 9 4 5 3 7 1 6 0 2]

Spectral Flux Std

Spectral Crest Std

Spectral Crest Mean

Root Mean Square Std

Spectral Flux Mean

Specral Centroid Std

Zero Crossing Rate Mean

Zero Crossing Rate Std

Root Mean Square Mean

Specral Centroid Mean

A diagram of a confusion matrix

Description automatically generatedA graph with a line

Description automatically generated

D.1

avg accuracy for k = 1 is: 0.746

avg accuracy for k = 3 is: 0.524

avg accuracy for k = 7 is: 0.336

Confusion matrices for all 3 cases: (Sorry about the formatting, never used word on Mac before, it’s weird)

Insights:

Label 0 – Classical is least confused and the one to most easily be classified.

Country tends to get misclassified the most, both with Jazz and Hiphop.

Metal is pretty well classified too, so it’s distinct from everything else

Jazz and Hiphop also get confused with each other.

A diagram of a confusion matrix

Description automatically generatedA diagram of a confusion matrix

Description automatically generatedA diagram of a confusion matrix

Description automatically generated