Zach Kondak and Ryan Rose

MUSI 6201

13 November 2017

Assignment 5: Genre Classification

Part A

**Explain how you solve the equal distance problem in your implementation and why do you think your approach makes sense.**

In order to break ties on distance measures, we look to the training set for more information. When a tie happens, we look at the number of classifications of both classes in the training set, and classify the given point in the class with a higher number. This makes sense because it helps the estimated classifications conform to what was learned from the training set. However, this approach relies on the assumption that the training set represents the distribution of real world data (which a good training set should do). Furthermore, as it stands, if matched classes in the training set have the same number of classifications, our algorithm reverts to picking the lowest class number. If we were allowed to make this non-deterministic, the class picking in the previous case would be random, which would then match the training set on average.

Part B

**Find the feature which performs the best. Report your result.**

Our results were as follows. As you can see, standard deviation of spectral flux performed best.

1 Root Mean Square Mean, 0.3581

2 Zero Crossing Rate Mean, 0.32999

3 Specral Centroid Mean, 0.28225

4 Spectral Flux Mean, 0.38217

5 Spectral Crest Mean, 0.36804

6 Root Mean Square Std, 0.38179

7 Zero Crossing Rate Std, 0.32795

8 Specral Centroid Std, 0.35196

**9 Spectral Flux Std, 0.39977**

10 Spectral Crest Std, 0.35611

Part C

**Run your feature selection method with K = 3 and num\_folds = 3. Report the best set of features. Plot how the accuracy changes with the feature selection iterations.**

Our results were as follows. The best set was the set of all features minus feature 5, mean of spectral crest.

1 Root Mean Square Mean,

2 Zero Crossing Rate Mean,

3 Specral Centroid Mean,

4 Spectral Flux Mean,

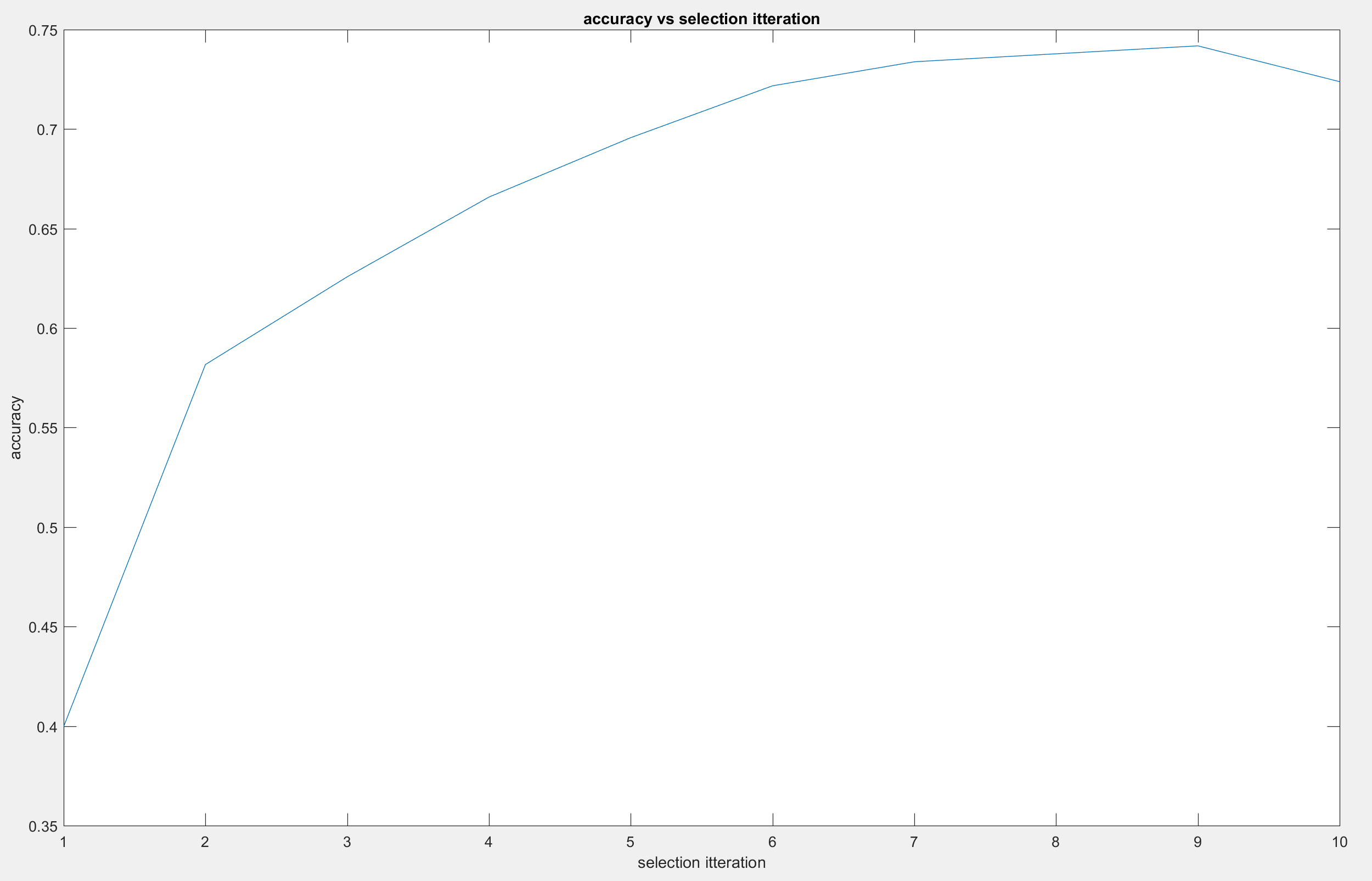
6 Root Mean Square Std,

7 Zero Crossing Rate Std,

8 Specral Centroid Std,

9 Spectral Flux Std,

10 Spectral Crest Std,



Part D

**Report the results of your cross validation (accuracy and confusion matrices) using the best set of features obtained in Part C.**

Our results were as follows.

k = 1

accuracy = 0.74

86 6 1 6 1

0 64 12 17 7

0 13 74 5 8

8 9 4 69 10

0 14 5 4 77

k = 3

accuracy = 0.722

83 6 1 9 1

1 67 10 12 10

0 12 73 5 10

10 20 4 58 8

0 8 8 4 80

k = 7

accuracy = 0.732

83 7 1 7 2

0 72 5 11 12

0 11 76 3 10

11 24 4 55 6

0 12 6 2 80

**Discuss the confusion matrices obtained in terms of what genres are confused with each other. How does K affect the classification performance?**

Our results show that there is no clear correlation between overall accuracy and K, with K=1 being the most accurate, but K=7 being more accurate than K=5. The largest difference with higher values of K is that for K=5 and K=7, values in class 4 (Jazz) were often being incorrectly estimated as class 2 (Country). In general, our classifier seemed to overestimate data points as Country, with Jazz and HipHop being classified as Country most often. This behavior also became exaggerated as K increased. All classifiers had least trouble classifying Classical music.

Part E

**Perform your k-means clustering using K = 5 for the genre data provided to you. Contrast the cluster labels with the ground truth genre labels. Discuss your results in terms of how closely genre relates to music similarity (based on this particular set of the features).**

In order to try to evaluate the cluster labels as class labels, we evaluated the accuracy of this cluster-classifier for all permutations of cluster label to class label mappings. When this best “ground truth” is used, the highest accuracy via this method was only 0.248. The confusion matrix is below.

49 21 17 4 9

27 25 27 10 11

41 18 25 10 6

27 21 24 15 13

35 21 25 9 10

Because of this low accuracy value, it follows that music similarity for this set of features and classification using these features in a kNN do not have a high degree of correlation.