ECSE 526
Assignment 2
Music Genre Classification

Sean Stappas
260639512

October 19th, 2017

1 What assumptions about the data do we make when we model the data using a Gaussian distribution?

Fundamentally, when we model the data using a Gaussian distribution, we assume that it has a Gaussian shape. This mainly implies the following:

1. The data has a mean.

2. The data has a standard deviation.

3. The data is symmetric around its mean.

2 When do you expect that a Gaussian will work well and when do you think it will not work well?

A Gaussian classifier will work well when the data can be modeled with a Gaussian distribution, with the features listed in the previous question. This can be seen for instance in Figure 1.
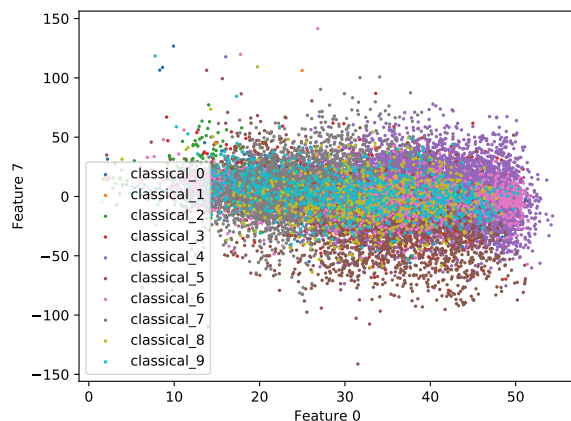

Figure 1: Projection of feature 0 vs. feature 7 for 10 songs from the "classical" genre. The distribution appears Gaussian.

A Gaussian classifier will not work in many situations. First, if the data itself cannot be modeled as Gaussian (using the assumptions in Question 1), then a Gaussian classifier will clearly not predict the correct values. For example, if the distribution is not symmetric, a Gaussian model will not work well. This can be seen in Figure 2, where two features were plotted against each other for ten songs from the "metal" genre.

An example where a Gaussian model would not be appropriate is if, within a genre, there are multiple clusters of values. Here, multiple Gaussian distributions may be appropriate, but certainly not one. This can be seen in Figure 3, where two features were plotted against each other for ten songs from the "kids" genre.
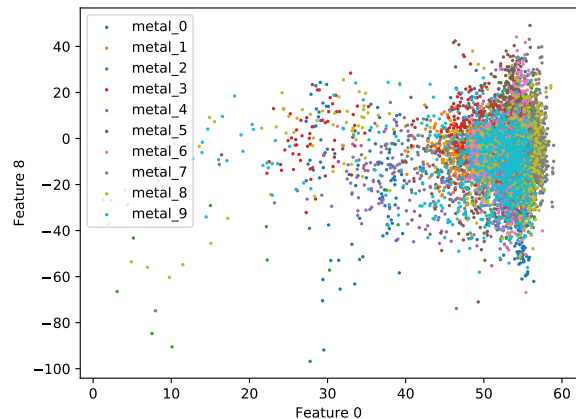

Figure 2: Projection of feature 0 vs. feature 8 for 10 songs from the "metal" genre. The distribution is not symmetric.
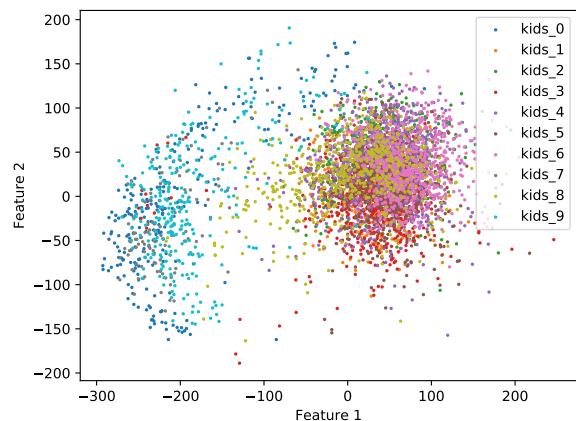

Figure 3: Projection of feature 1 vs. feature 2 for 10 songs from the "kids" genre. Two separate clusters of values can be seen, on the left and the right.

Secondly, if the feature vector distributions of two genres are very similar, then their corresponding Gaussian distributions will be very close, perhaps intersecting. In this situation, a Gaussian classifier can have a hard time differentiating between these two genres. This can be seen in Figure 4, where the "latin" and "rnb" genres have very similar distributions.

Also, a Gaussian classifier will not work well for values that stray too far from the mean. In situations like this, kNN has the clear advantage. This can be seen in Figure 5, where two features were plotted against each other for ten songs from the "latin" genre.
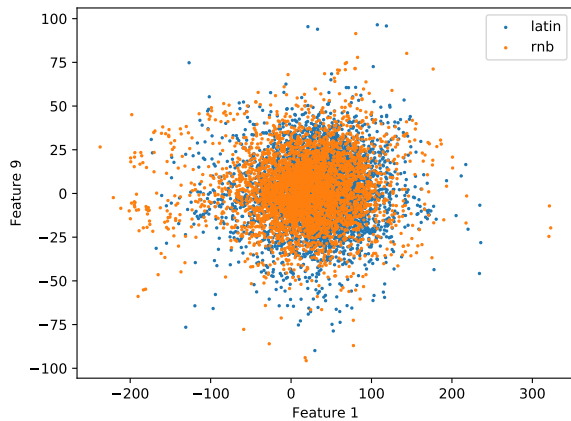
1

Figure 4: Projection of feature 1 vs. feature 9 for 5 songs from the "latin" genre and 5 songs from the "rnb" genre. The distributions are very similar.
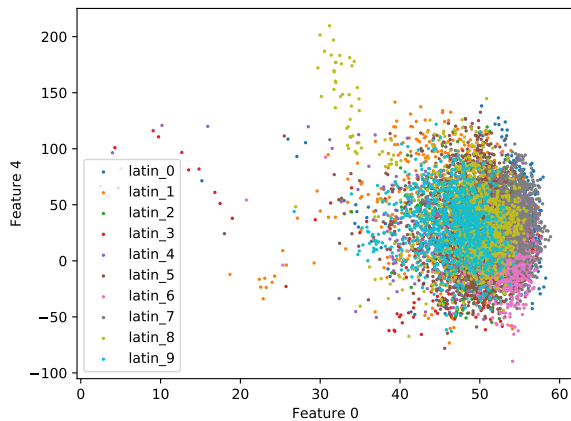

Figure 5: Projection of feature 1 vs. feature 2 for 10 songs from the "latin" genre. There are many extreme values for which kNN would have an easier time classifying than Gaussian.

3 What values of k work best for the kNN classifier?

A common rule-of-thumb for the value of k is that it should be equal to $\sqrt{N}$, where N is the number of training samples.

After testing on various values of k, the best one was found to be

4 Based on your results from this assignment, which classifier (Gaussian or kNN) works best for the task of Music Genre Classification? Why?

The results of both classifiers can be seen in

Based on the above results, the kNN classifier works best for music classification. This can be at-

tributed to the simple fact that the Gaussian classifier retains much less information than the kNN for each genre. While the Gaussian classifier only stores a 12x12 covariance matrix and 12x1 mean vector for each genre, the kNN classifier stores every training example and uses these examples to predict every genre. This additional information can allow the kNN classifier to classify new songs with much greater accuracy.

Also, kNN can take into account very complex distributions of feature vectors, whereas the Gaussian classifier will attempt to fit a Gaussian distribution to the data set, which may not be appropriate.

2

Table 1: *Confusion matrix for the Gaussian classifier. Each row represents the actual genre, and each column the predicted genre. The ordering of the column and row genres are the same.*

	kids	classical	rnb	jazz	metal	pop	edm˙dance	latin	rock	country
kids	31	3	34	11	0	2	0	0	0	0
classical	16	54	0	7	0	0	0	0	0	0
rnb	3	0	72	2	0	0	0	0	0	0
jazz	2	0	11	66	0	0	0	0	0	0
metal	14	0	29	0	0	22	4	0	5	0
pop	1	0	68	0	0	1	0	0	0	0
edm˙dance	3	0	69	0	0	3	0	0	0	0
latin	0	0	70	0	0	0	0	7	0	0
rock	1	0	63	0	0	8	0	0	0	0
country	1	0	65	0	0	8	0	0	0	0

Table 2: *Confusion matrix for the kNN classifier, with $k = 1$. Each row represents the actual genre, and each column the predicted genre. The ordering of the column and row genres are the same.*

	kids	classical	rnb	jazz	metal	pop	edm˙dance	latin	rock	country
kids	26	14	1	9	7	1	6	13	0	4
classical	0	76	0	1	0	0	0	0	0	0
rnb	0	2	29	8	2	7	7	21	0	1
jazz	0	2	1	73	2	0	1	0	0	0
metal	0	0	0	0	73	0	0	0	0	1
pop	0	0	9	7	9	5	14	17	1	8
edm˙dance	0	3	1	4	2	1	58	5	0	1
latin	0	0	0	0	0	0	4	72	0	1
rock	0	0	0	2	38	1	12	4	3	12
country	0	0	0	3	16	1	2	5	1	46