Results & Conclusion

Results

The dataset contains converted data from bird song samples. The datapoints contain information about the intensity of certain tones in these samples. The CSV data is supplied by Edoardo Ferrante on Kaggle. This data was created using the Librosa package for Python. Librosa outputs the intensity of a certain tone at different time intervals from the provided sound file. There is a Python script available from the author that explains how the data was transformed from sound files to numeric data points, but this script has a specific shortcoming that will be explained later.

Data Pre-Processing

First let's have a look at how Librosa normally outputs the chromogram data:

	0	1	2	3	4	5	6
chromogram_0	0.68661	0.67378	0.65758	0.66149	0.68533	0.72239	0.76395
chromogram_1	0.91368	0.88148	0.85024	0.82476	0.82282	0.83024	0.83908
chromogram_2	0.98221	0.97060	0.95834	0.94729	0.94785	0.95189	0.95408
chromogram_3	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
chromogram_4	0.96223	0.95790	0.95436	0.95403	0.94404	0.93285	0.91552
chromogram_5	0.92098	0.89960	0.88007	0.86290	0.84262	0.82280	0.80381
chromogram_6	0.87591	0.85544	0.83320	0.80870	0.78241	0.75819	0.73456
chromogram_7	0.79397	0.79418	0.79848	0.80535	0.80741	0.80750	0.80913
chromogram_8	0.62856	0.64859	0.68255	0.72441	0.76565	0.80078	0.82178
chromogram_9	0.41881	0.41914	0.44003	0.48525	0.54615	0.61608	0.68581
chromogram_10	0.36895	0.38299	0.40678	0.43114	0.45138	0.47237	0.52333
chromogram_11	0.33855	0.34676	0.35929	0.37378	0.38602	0.39931	0.42756

As you can see, the output is a neat array containing the values of 12 different tones at different time intervals. This data is sorted and can be read by Librosa.

Now onto the issue; the provided dataset contains this data in a stacked order, so each sample only takes up one row. This is a good idea, but due to sorting by alphabetical order the original order is lost. The order is important because we are working with data over time. This is not a problem if the trained model is only used on the provided test data, but we want the trained model to work in as many situations as possible.

Here is a look at the provided data:

id	chromogram__0__0	chromogram__0__1	chromogram__0__10	chromogram__0__11
0	0.997943662321316	0.832392210770135	0.7653861625931	0.70427464132375
1	0.996254885931866	0.839119599044146	0.760416790506312	0.705141765139875
2	0.970810156116343	0.823539694937237	0.759508104372184	0.709057883677716
3	1	0.855558393364941	0.752038009313116	0.710976936190937
4	1	0.884304523555434	0.741884532311754	0.714775207828629
5	0.971867873978603	0.824311712155432	0.755293860709407	0.71448132195049
6	1	0.835499361583387	0.751917158063063	0.717361992854453
7	0.978929855885584	0.827216718543843	0.751072631712318	0.718400862681119
8	1	0.895339720206626	0.733409813021178	0.722747412968086
9	0.967651828343747	0.823697857901917	0.746005680687241	0.721194823494439
10	0.993699774531599	0.847257121555946	0.734368883301346	0.72620069139032
11	0.00947350497274455	0.00699383738737368	0.372026644035831	0.0516494292032762
12	0.00982270123521504	0.00712337798131429	0.371129653847745	0.051631441504244

Each row contains a stack of chromogram data in a non-sequential order. The end of the array also contains the species of the corresponding bird and some spectral centroid data. These collumns need to be deleted since we want to predict using only the chromogram data.

Figure 1: Comparison of the data order

Figure 1 shows a comparison of the order of provided data and the ideal order the data should be sorted in. A Python script was used to order the data into this format.
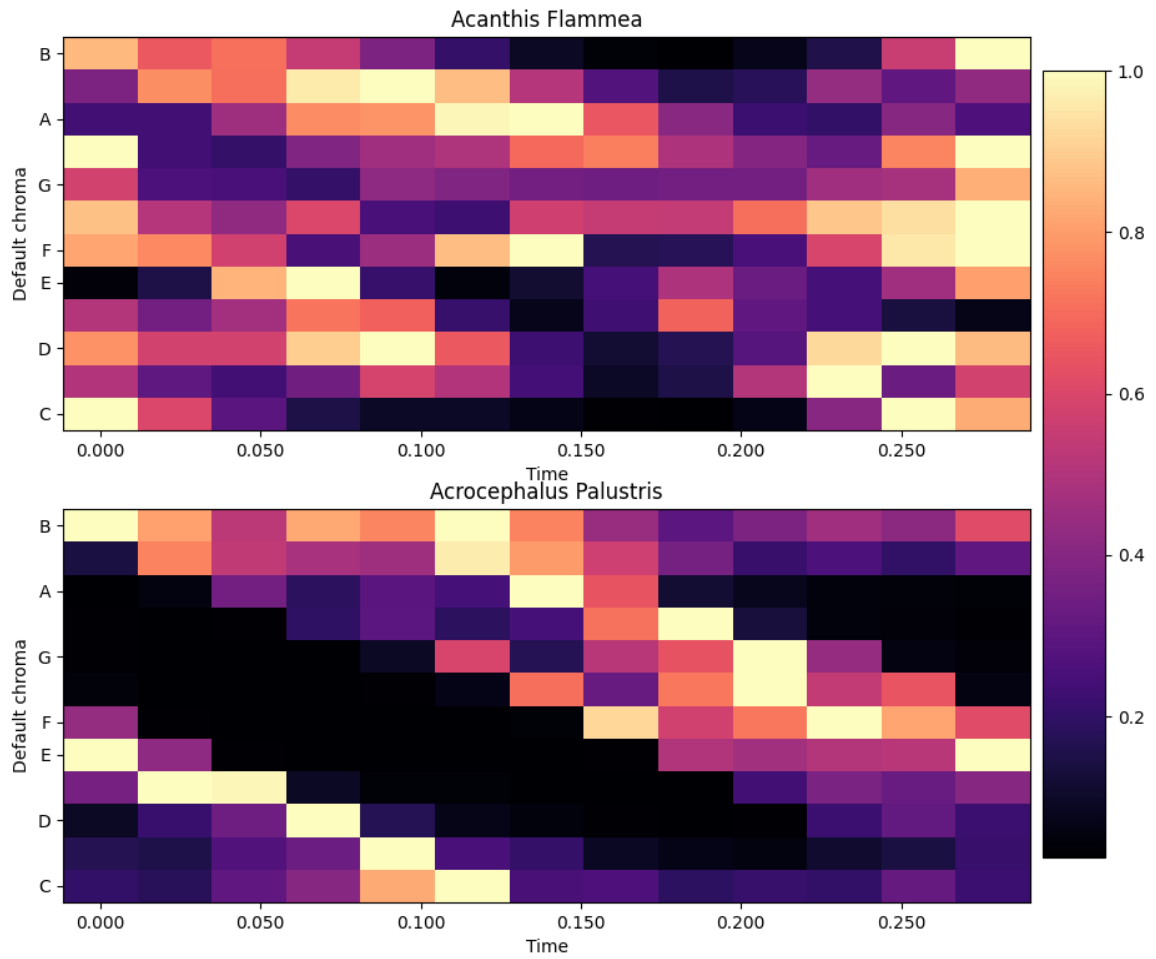
Figure 2: Chroma signature comparison for 2 fragments of different bird-species songs

Figure 2 shows the first result of the data exploration. There is a clear difference in signature between the two species. This difference can be explored further by analyzing one tone at the time.
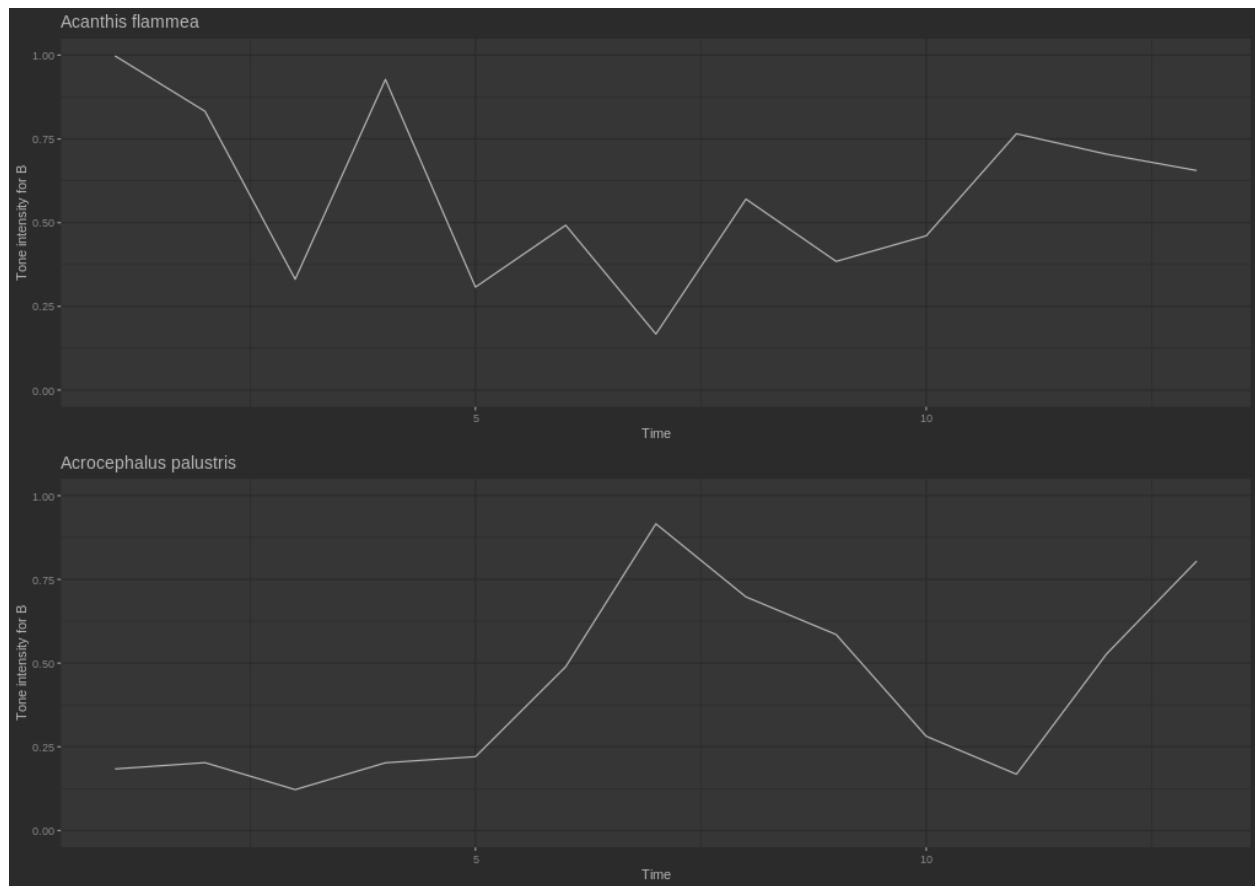
Figure 3: Tone intensity comparison at tone B for 2 fragments of different bird-species songs

Figure 3 also shows a clear difference in tone intensity over the duration of the sound fragment. Tone B was chosen as an example, but multiple tones exhibit this behaviour as seen in Figure 2.
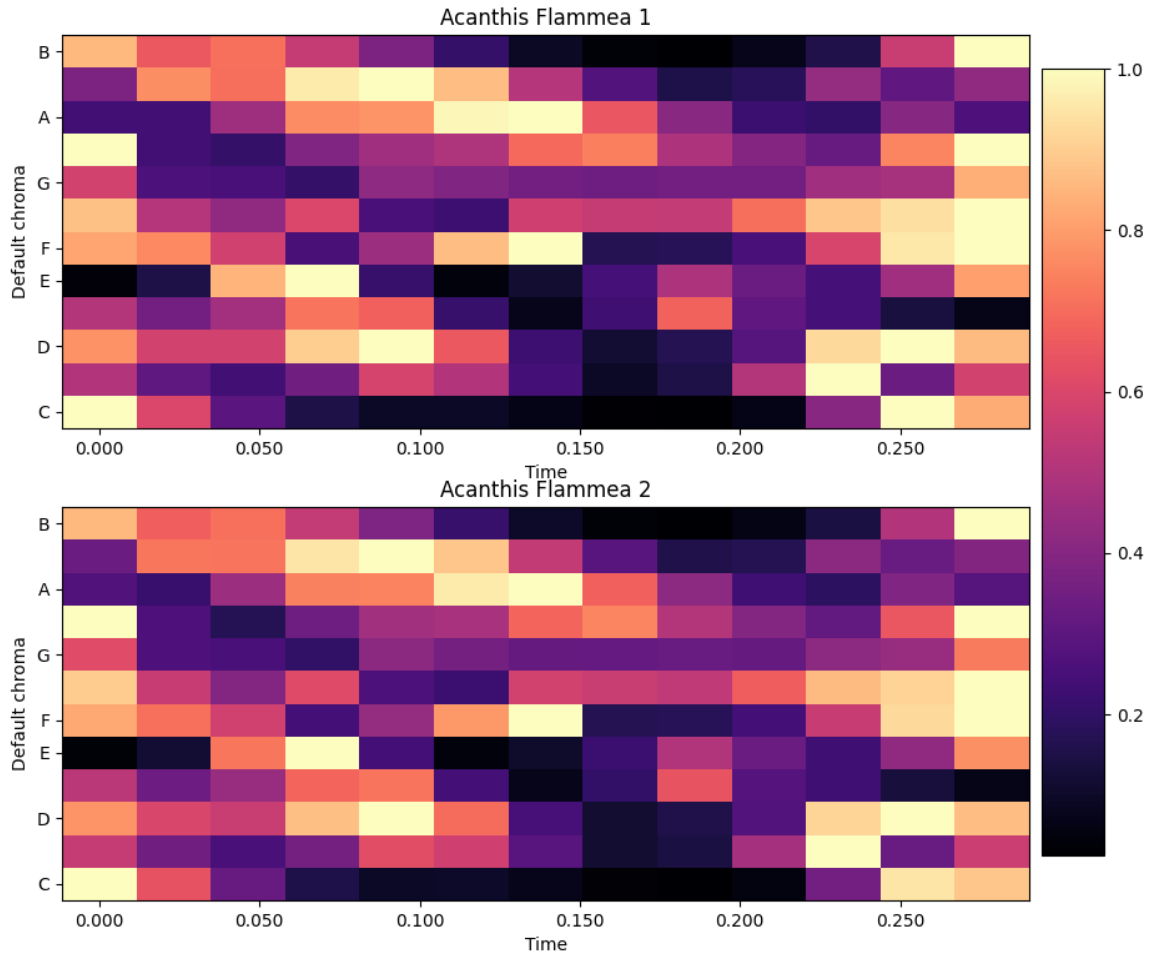
Figure 4: Chroma signature comparison for 2 fragments of the same bird-species songs

Lastly, figure 4 shows the signature comparison of 2 bird songs of the same species. These signatures look alike, but there are subtle differences to be seen.

Conclusion

Concluding from the data, it seems there is a high probability of success when training a machine learning model on this data. The signatures of the same species align with each other nicely and the signatures of other species are different enough to be separated. Although the signatures of the same species look alike, there are subtle differences in the intensity. We can also conclude from this that birds of the same species sing the same songs.

Discussion

The data is in fact pretty messy, but once it is cleaned it makes for a nice machine learning dataset.