SPEAKER IDENTIFICATION

Machine Learning to identify speakers

The Problem

- Given a set of data containing 370 utterances of an vowel by different speakers, the task is to train a machine learning model which can identify the speaker given the utterance.
- The train data consists of LPC coefficients of the voice data
- Labels are provided for the train data.
- Each sample is a block in the train data set.

Loading data and preprocessing

Reading the data requires special processing.

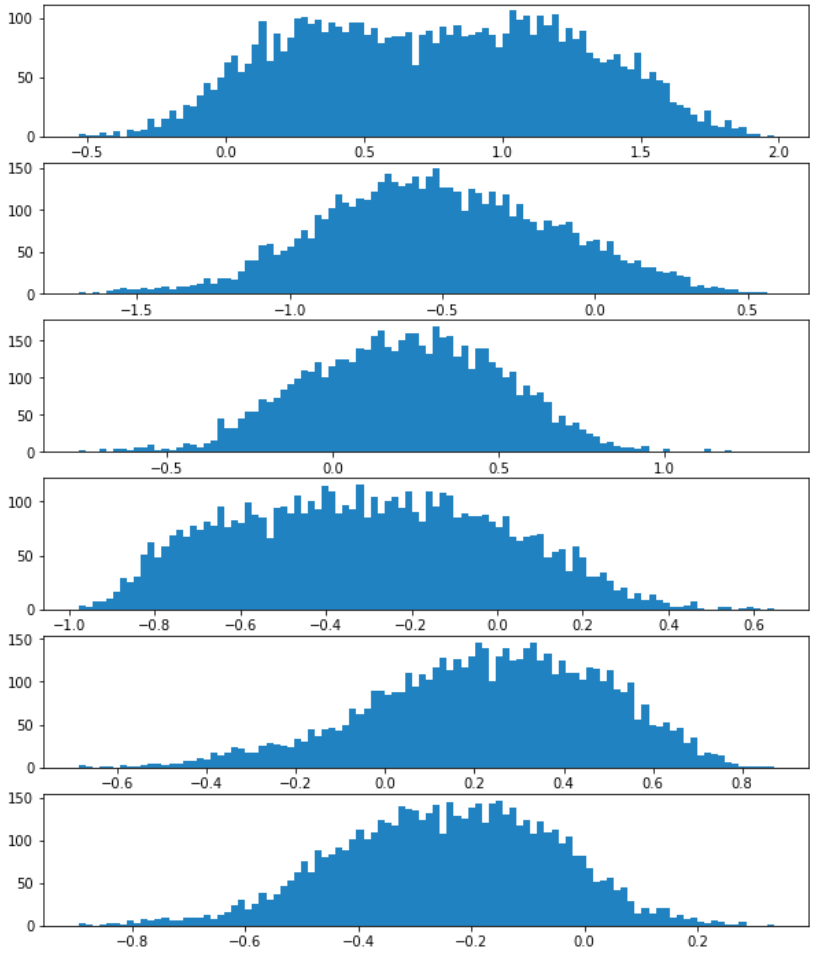
Each block in the train file needs to be separated into individual sample.

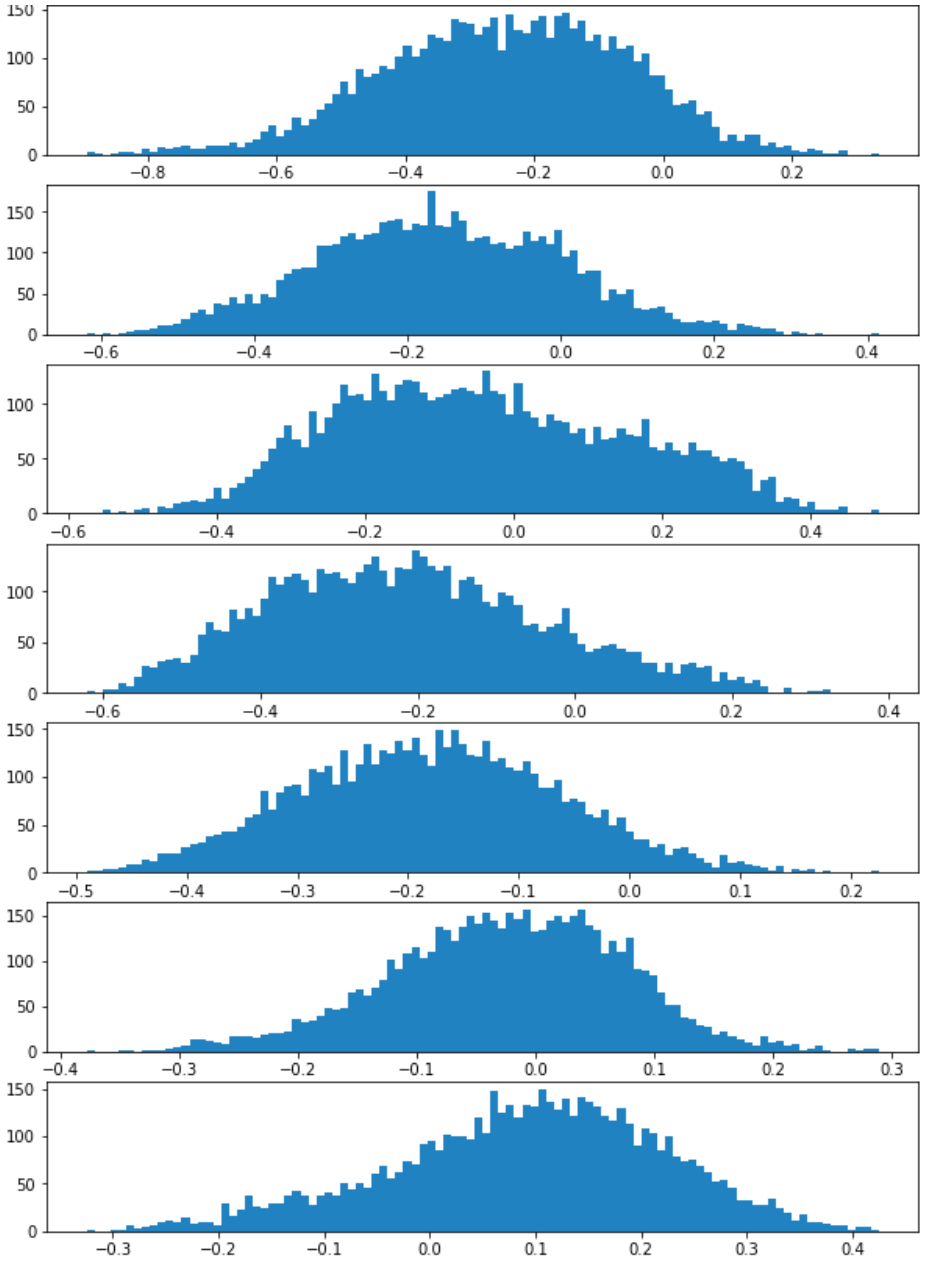
Not all sample have same length.

So padding has to be added to make all of them equal in length.

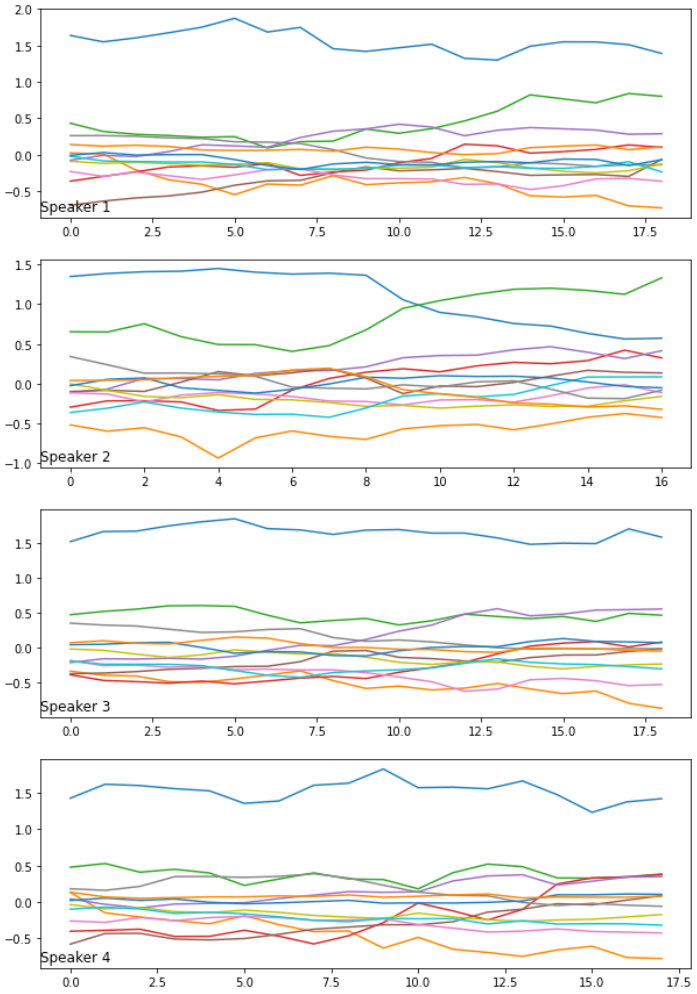
Data Exploration

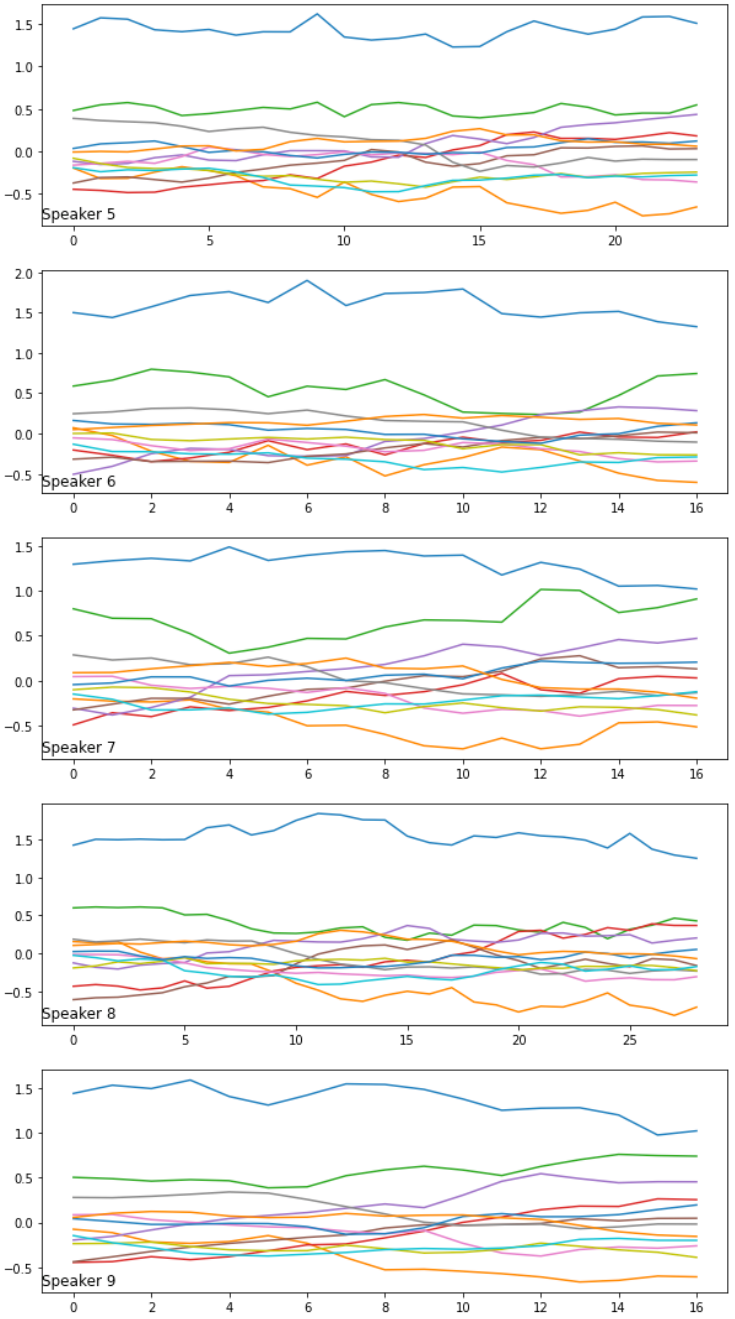
- Histograms of LPC Coefficients



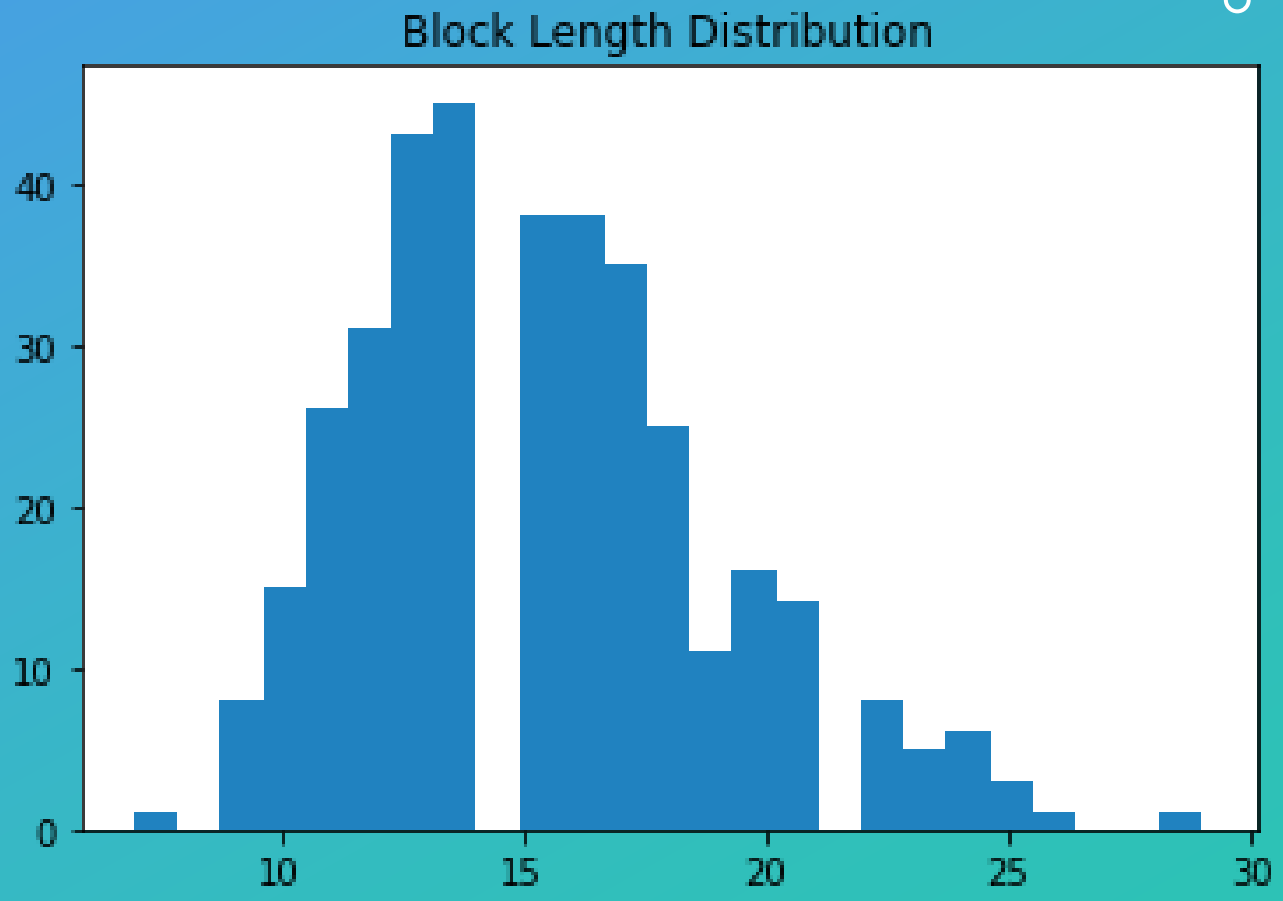


Visualizing LPC data

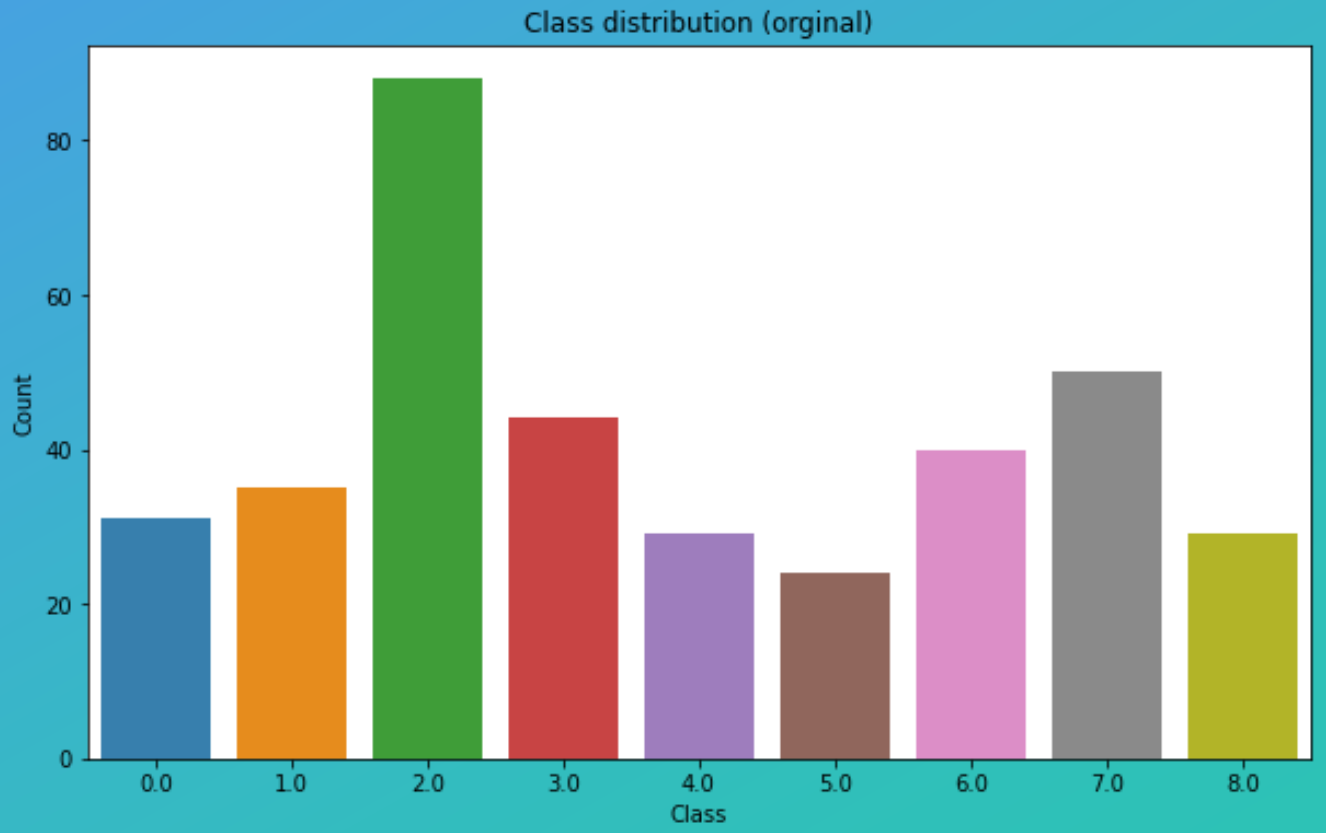




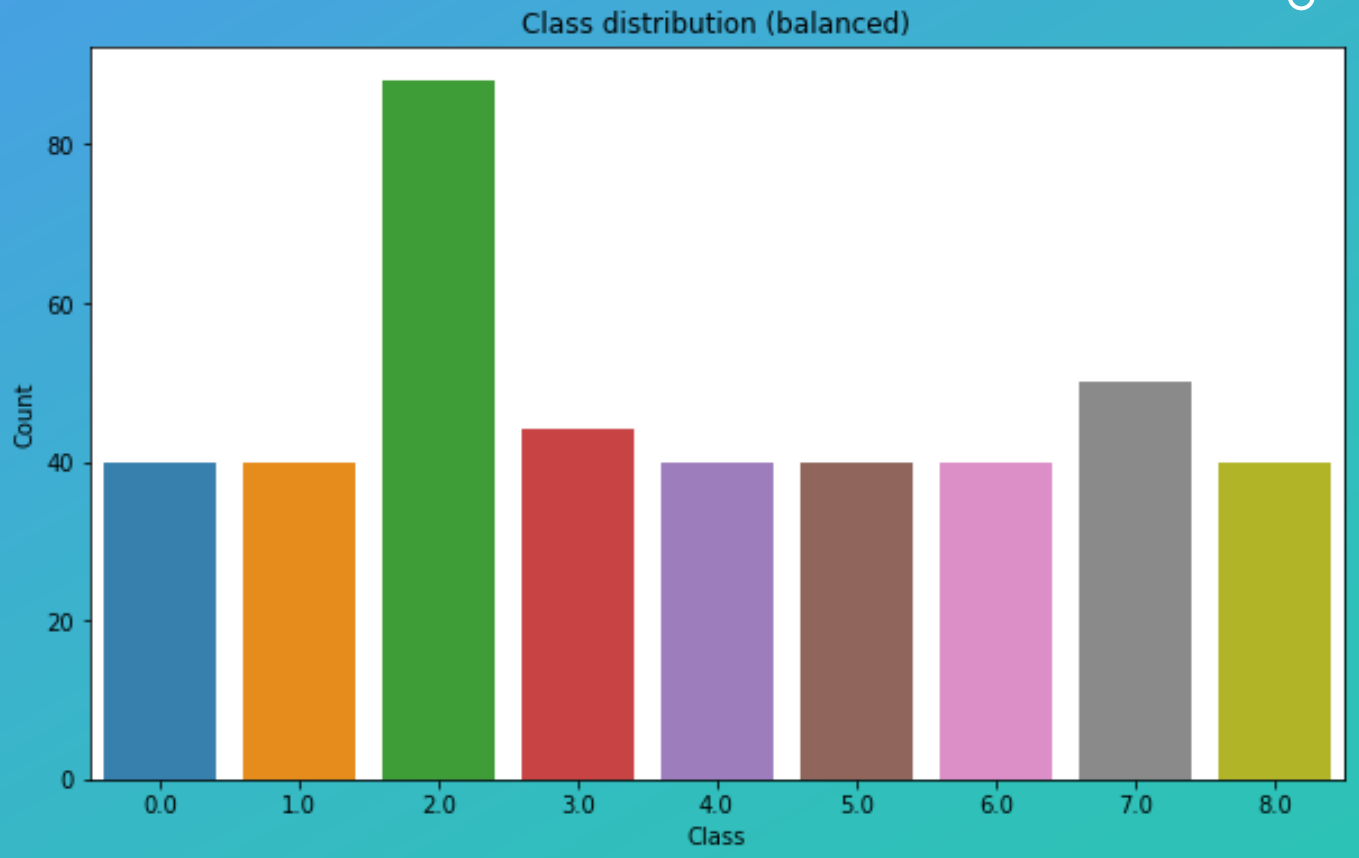
BLOCK LENGTH DISTRIBUTION



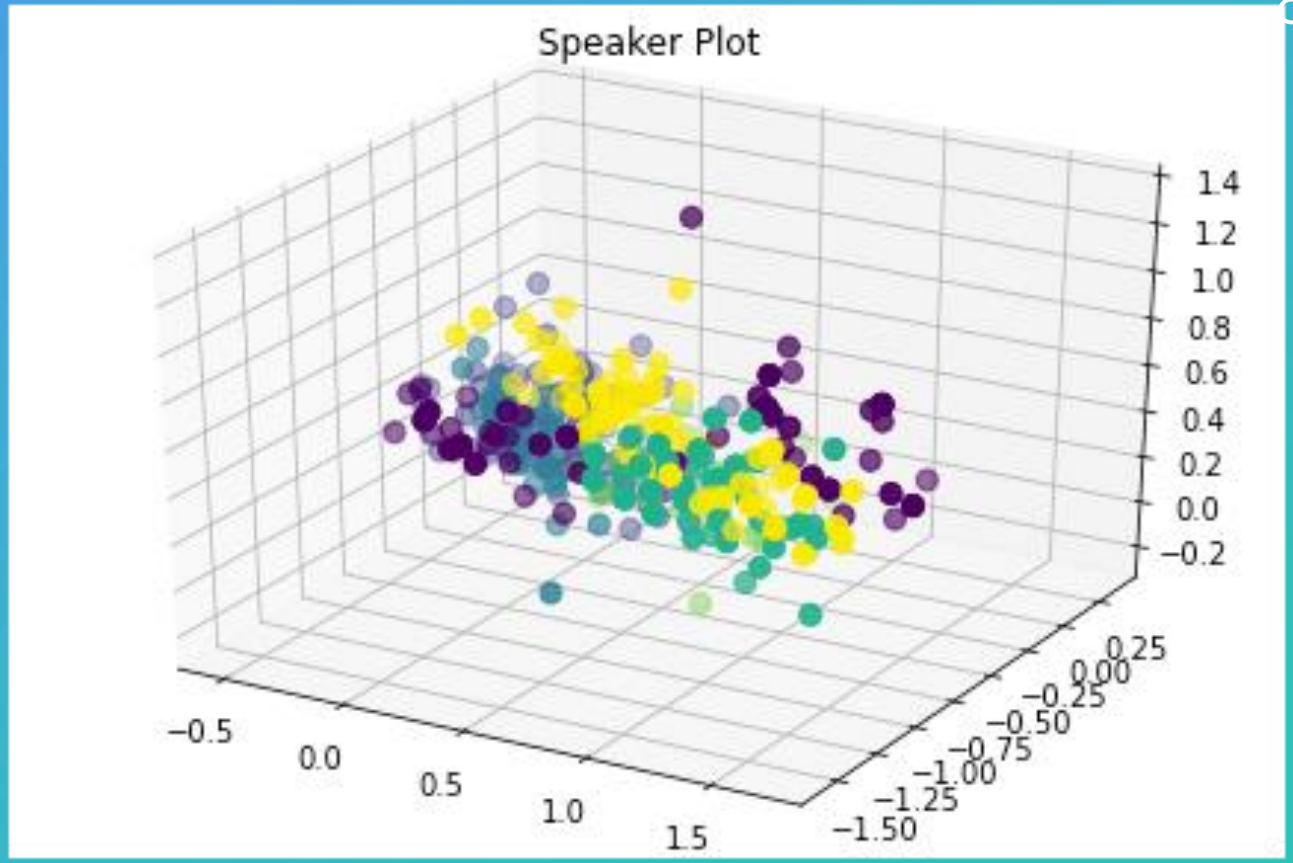
CLASS DISTRIBUTION



BALANCING CLASSES



VISUALIZING THE SPEAKERS IN SCATTER PLOT



Accuracy using different ML models

- KNN 96.443% +/-1.309
- LR 98.106% +/-1.420
- CART 82.457% +/-3.179
- SVM 98.342% +/-1.212
- RF 96.443% +/-1.848

- SVM seem to be working well, so fine tuning it further.

SVM fine tuned 97.027% +/-1.986
Best Parameters: {'C': 1, 'gamma': 0.001, 'kernel': 'linear'}

Results using SVM before balancing

	precision	recall	f1-score	support
0.0	1.00	0.97	0.98	31
1.0	0.94	0.94	0.94	35
2.0	0.98	0.97	0.97	88
3.0	1.00	0.95	0.98	44
4.0	0.97	1.00	0.98	29
5.0	1.00	1.00	1.00	24
6.0	1.00	0.97	0.99	40
7.0	0.94	0.98	0.96	50
8.0	0.90	0.97	0.93	29
accuracy			0.97	370
macro avg	0.97	0.97	0.97	370
weighted avg	0.97	0.97	0.97	370

Confusion Matrix:

```
[[30  0  0  0  0  0  0  0  1]
 [ 0 33  0  0  0  0  0  1  1]
 [ 0  1 85  0  0  0  0  1  1]
 [ 0  1  0 42  0  0  0  1  0]
 [ 0  0  0  0 29  0  0  0  0]
 [ 0  0  0  0  0 24  0  0  0]
 [ 0  0  1  0  0  0 39  0  0]
 [ 0  0  1  0  0  0  0 49  0]
 [ 0  0  0  0  1  0  0  0 28]]
```

Results using SVM after balancing

- SVM got score of 0.96 for train data

	precision	recall	f1-score	support
0.0	1.00	1.00	1.00	40
1.0	0.97	0.95	0.96	40
2.0	0.98	0.98	0.98	88
3.0	1.00	0.98	0.99	44
4.0	1.00	1.00	1.00	40
5.0	1.00	1.00	1.00	40
6.0	1.00	0.97	0.99	40
7.0	0.96	0.98	0.97	50
8.0	0.95	1.00	0.98	40
accuracy			0.98	422
macro avg	0.98	0.98	0.98	422
weighted avg	0.98	0.98	0.98	422

Confusion Matrix:

```
[[40  0  0  0  0  0  0  0  0]
 [ 0 38  0  0  0  0  0  1  1]
 [ 0  0 86  0  0  0  0  1  1]
 [ 0  1  0 43  0  0  0  0  0]
 [ 0  0  0  0 40  0  0  0  0]
 [ 0  0  0  0  0 40  0  0  0]
 [ 0  0  1  0  0  0 39  0  0]
 [ 0  0  1  0  0  0  0 49  0]
 [ 0  0  0  0  0  0  0  0 40]]
```