

Colm Mooney - 20325583

The amount of correct predictions is dependent on MFCCs. The accuracy of the model varies due to how close the mfccs is close to 20. Test 1 has achieved the highest accuracy of the observed, getting 81%, and the lowest amount of accuracy is 32% with MCC being 2. If changed the MCC to 1, the accuracy would have been even lower.

The Accuracy for the model continues to drop the higher the mfcc gets. When the mfccs is 50, 200, 100, or 250 it shows that there isn't a point in going any higher than 50. The accuracy drops completely. It's being overfitted, and when the mfcc was 2, it was being underfitted.

Changing the FFT window did not seem to affect the accuracy. The values used were 128, 256, 512, 1024

For the best results, I believe consistently retraining the model when mfcc is between 18-22 would lead to the best result, with 20 most likely being the highest. This range offers a good compromise between complexity & accuracy.

I tested changing the fmin, where I changed it from 0 to 10. Doing so made the program less accurate on when mfcc = 20. I got a result of .76. Changing fmin to 20 made it even less accurate, going to .73

Setting fmin to 0, and fmax = 100, The accuracy dramatically got worse. Going to 0.23 accuracy. Setting fmax = 500 only increased it marginally to 0.36. This discouraged me from changing fmin & fmax at all as it only made the model less accurate.

Hop length affects the time resolution but has no noticeable effects when tested in this code. That isn't until it is paired with win_length. I did some tests where the hop length was half of win_length so there was good overlap between frames. And after training the data, there were no noticeable changes when the window & hop length were decently large. However, there was a significant increase in accuracy when the window length & hop length were short. With a window length of 30 & a hop length of 15, the accuracy got slightly better, going up 2-4%. This also made the time it took for the code to process, however, take around 2 minutes per try. This low win_length is useful when trying to recognize speech.

The center being true or false was tested, but it had no noticeable impact on the accuracy. Whether it was on or off, the level of accuracy was the same.