Emotion Classification Using Machine Learning: Project Proposal

Nathaniel Lounsbury, Qiyang Ma, Christopher Woloshyn

March 3, 2020

1 Objectives of Proposed Research

Our goal is to analyze a set of audio data, and classify the emotions present using various methods of pre-processing and classification. The aim is to begin with three different methods of pre-processing: MFCC, LPCC, and Wavelet analysis, and attempt to recognize the emotions present in the audio. Our starting point will be Support Vector Machines (SVM), with further plans to expand onto other classification methods upon successful implementation of the SVM algorithm.

As we discussed during the first lecture of class, the goal of the third wave of artificial intelligence is to train computers to understand context and meaning that may exist outside of pure statistical observation. For example, the ability to infer emotion from human speech is key to understanding the full context and meaning of spoken language in our society. With our project, we hope to take a small step towards improving the interaction between humans and computers.

2 Approach

The foundation of this project will be based on a similar paper titled "Speech Emotion Recognition using Support Vector Machine," by Manas Jain and colleagues[1]. In their paper, they reference several pre-processing and classification methods published by other researchers. These methods include MFCC and LPCC for pre-processing and feature extraction, and SVM for emotion detection. The researchers used two data sets. One is from Linguistic Data Consortium (LDC), which is a website where access to their data is not free. The second is a dataset from UGA. Since the LDC data set is difficult to access, we plan on instead using a free data set from the University of Toronto called the Toronto emotional speech set (TESS). If successful, we would also like to expand this idea by creating our own audio in different languages. Some languages could include: Mandarin, Spanish, Japanese, etc. Additionally, We would like to experiment with other classification methods; possibilities include Random Forrest, Logistic Regression, LDA, etc.

3 Expected Outcomes

Overall, we expect to achieve reasonable results for emotion recognition using different pre-processing and classification techniques. Results may vary, but we plan to identify the best combination of pre-processing and classification algorithms that maximize the accuracy on our dataset.

References

[1] Jain, M., Narayan, S., Balaji, P., Bhowmick, A., & Muthu, R. K. (2020). Speech emotion recognition using support vector machine. arXiv preprint arXiv:2002.07590.