

Research Assessment: Final Product

Kavan Mehta

Ms. Whitcomb

ISM 2

28 January 2023

Kavan Mehta

Ms. Whitcomb

ISM 2

28 January 2023

Research Assessment #16: Final Product

Date: 28 January 2023

Subject: Machine Learning in Acoustics

MLA citation(s):

Bianco, Michael J., et al. "Machine Learning in Acoustics: Theory and Applications." *ArXiv.org*, 1 Dec. 2019, <https://doi.org/10.48550/arXiv.1905.04418>.

Assessment:

After deciding on my Final Product idea of utilizing both audio and video to recognize speech, I look forward to understanding speech acoustics. Speech acoustics already covers some of my knowledge of natural language processing (NLP) through my Original Work, but I still need to comprehend the various applications of machine learning in speech acoustics. From my past interview, I know that I need to see the "Big Picture" of understanding speech before I begin implementing my Final Product idea. To understand the roots of the problems and necessities for optimization of speech recognition, I found a scholarly article, "Machine Learning in Acoustics: Theory and Applications," which went over the various applications of machine learning in acoustics and algorithms that can be used to help solve and optimize speech and sound recognition.

Through the journal article, I first refreshed my memory on the utilization of machine learning to transform certain text and sound features to understand speech and acoustics

effectively. There are many applications of machine learning in acoustics, including “speaker localization in reverberant environments, source localization in ocean acoustics, bioacoustics, and reverberation and environmental sounds in everyday scenes” (Bianco et al. 2). I think that understanding these applications of machine learning can help us see the acoustics of various sounds that factor into speech recognition. I also learned about various algorithms, including linear regression, logistic regression, and deep neural networks (feed-forward and recursive) that play into speech recognition. I think another interesting fact that will serve useful with supervised and unsupervised learning is the need to “approximate data relationships (between features and labels) of a particular complexity is its capacity“ in the generalization of speech recognition and other applications (Bianco et al. 3). As mentioned in my previous interviews, it will be crucial to understand all the stakeholders and the various points of influence that would impact the analysis of speech. My Original Work projects already explore text classification and speech recognition of simple phrases, which are very crucial to understanding the user’s intent and really emphasize the study of data. From my recent interview with Dr. Golden from the University of Texas at Dallas, I learned about Bayesian networks, which I grew more interested in through the journal article. After understanding some ML theory, an application that I explored was acoustic reverberation, where ML was used to understand “sounds presented from different locations in the same space” and where algorithms “leverage knowledge about the acoustical structure of sources” (Bianco et al. 30). I think that we could utilize speech and its very foundation to understand the elements of it, and by utilizing various ways of representation of speech, I could help optimize current solutions with both audio and video of lip movements and sounds.

By understanding the various factors that impact speech recognition, I hope to utilize my machine learning knowledge to identify current weaknesses that I could improve through my Final Product idea. This could be groundbreaking in assisting us efficiently in everyday applications of voice assistants, speech recognition, and advancements in future speech-based technologies, saving millions of hours and dollars. I hope to continue to learn about methods of implementation I could use to utilize NLP technology to solve the optimization problem in speech acoustics. Moreover, I hope to finalize my mentor and create my Final Product by understanding the speech acoustics area through a professional lens that has experience in effectively implementing NLP technologies!