

# Unsupervised Speaker Identification

Thivakkar Mahendran  
University of Massachusetts Amherst  
tmahendran@umass.edu

Varun Prasad  
University of Massachusetts Amherst  
varunanantha@umass.edu

## 1. Abstract

Filled by Varun

## 2. Introduction

A voice recorder makes it super convenient to take notes or preserve the minutes of a meeting. In order to make the experience better, there are tools that can be used to automatically convert speech to text. One area where this tool currently fails at is identifying the speaker. The problem that we are trying to solve is to identify the speaker. Our ultimate goal is to build a project where we are able to record a meeting and which particular person speaks at a specific time and what they speak.

It would be cumbersome and time consuming to record each speaker's voice beforehand and train a speaker recognition model on the speakers' voice. The goal is to make this tool predict the speaker without prior training on the speaker's voice. This would be an unsupervised learning project.

For example, if two speakers were talking in a meeting we would like our tool to give an output:

**Speaker1:** Hey, How are you?

**Speaker2:** I'm doing well.

**Speaker2:** I was able to complete the tasks that we talked about last week

**Speaker1:** That's great.

These two speakers' voices have not been trained before but still the program would identify the two speakers just by getting the features from their voice.

## 3. Background/Related Work

Filled by Varun

## 4. Approach

Filled by Thivakkar

## 5. Experiment

Filled by Varun

## 6. Conclusion

Filled by Varun

## References

- [1] Barnard, Dom. "Average Speaking Rate and Words per Minute." VirtualSpeech, VirtualSpeech, 20 Jan. 2018, [virtualspeech.com/blog/average-speaking-rate-words-per-minute](https://virtualspeech.com/blog/average-speaking-rate-words-per-minute).
- [2] Sharma, Usha, et al. "Study of Robust Feature Extraction Techniques for Speech Recognition System." 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 2015, doi:10.1109/ablaze.2015.7154944.
- [3] "Analytical Review of Feature Extraction Technique for Automatic Speech Recognition." International Journal of Science and Research (IJSR), vol. 4, no. 11, 2015, pp. 2156–2161., doi:10.21275/v4i11.nov151681.