1. Introduction

Background

Accent is a distinctive manner of expression such as a way of speaking on typical of particular group of people and especially of the native or residents of a region and an individual distinctive or characteristic influential, tone, or choice of words [1]. In sociolinguistics, an accent is a manner of pronunciation peculiar to a particular individual, location, or nation [2]. Researchers at Cisco, the Moscow Institute of Physics and Technology, and The Higher School of Economics present a possible solution in a new paper "Foreign English Accent Adjustment by Learning Phonetics Patterns" published on the preprint server Arxiv.org [3]. Hence, we can see that researchers attempted to create accent detection are the evidence that technology for accent detection is not useless in the meantime. According to Arifin, is the form of Sundanese priangan accent same as non-priangan Sundanese accent, like Cirebonese accent, Bantenese accent, and the other Sundanese accent? [4]. Nowadays, voice can be deepfake using AI and resembles more human-like and having accent. According to the Post, cybersecurity firm Symantec says it has come across at least three cases of deepfake voice fraud used to trick companies into sending money to a fraudulent account [5]. In this matter of facts, accent detection will be in need on accent recognition if AI tries to deepfake or someone who tries to fakes accent and resembles someone who we knew in family members.

Accent detector that developed by Roshni and Annie (2020) already can recognize some of accent in Malayalam language [6]. On the other hand, accent detector that developed by Ashok Kumar, Karthika Renuka, and Raajkumar (2021) already can recognize Indian accent using Listen, Attend, and Spell (LAS) model with end-to-end automatic speech recognition (ASR) [7]. To conclude accent detection related field, Accent detection using Sundanese accent have not been done. Limitations on Sundanese accent detection system is mainly on reference that related to Sundanese accent detection. Accent detection related research already referenced in the previous paragraph have not specifically refer to Sundanese accent detection.

As cover up statement above, the reason on creating accent detector is that accent detector having sort of limitation in the research and not specified to Sundanese accent itself. Therefore, the purpose of this project is to creating possibility on accent detection and a system to validate Sundanese accent specifically between Native Sundanese accent and Non-Native Sundanese accent.

Problem Statement

As the current condition and reason that explained in the background above, there are problems that will be raised in this study, mainly how accent detection system detecting the certain local language specifically in Indian, Malay Pontianak, and Garut Sundanese accent [7, 14]. Furthermore, accent detection system that specifically detects Sundanese accent has not been fully implemented from the previous accent detection. As the relevance to the Sundanese accent detection system that was previously implemented, accent detection system that specifically detecting Sundanese accent has not been implemented in the process.

Objective

The objective of this final project is to develop local accent detection system that specifically detecting Sundanese accent by implementing Mel-Frequency Cepstral Coefficient (MFCC) feature extraction method and Hidden Markov Model (HMM) as the model. Furthermore, the reason of using MFCC-HMM implementation due to MFCC-HMM method are the most common method on speech recognition and accent detection system.