

ABSTRAK

Seiring perkembangan zaman, umat Islam di dunia khususnya di Indonesia terus mengalami perkembangan. Hingga kini, Islam menjadi agama dengan pemeluk terbanyak di Indonesia. Tak dapat dipungkiri pula, setiap agama memiliki pedoman hidupnya masing-masing. Al-Qur'an adalah pedoman pertama dan utama bagi umat Islam yang diberikan oleh Allah SWT [1]. Membaca Al-Qur'an pun sudah menjadi rutinitas bagi kalangan umat Muslim. Namun, membaca Al-Qur'an bukan menjadi perkara mudah bagi masyarakat Indonesia karena Al-Qur'an menggunakan Bahasa Arab. Kesalahan-kesalahan dalam membaca pun sulit untuk dihindari seperti kesalahan pengucapan huruf hijaiyah bertanda baca. Permasalahan ini akan coba diselesaikan menggunakan teknologi *Speech Recognition*. *Speech Recognition* akan membantu pengguna untuk melakukan pengecekan apakah ucapan huruf hijaiyah pengguna tersebut sudah benar atau tidak.

Pada tugas akhir ini dibangun sebuah sistem *Speech Recognition* dengan menggunakan metode ekstraksi ciri *Mel-Frequency Cepstrum Coefficients* (MFCC) dan menggunakan metode *K-Nearest Neighbour* (KNN) untuk melakukan klasifikasi. MFCC telah *common* digunakan dalam memroses data suara [9]. Selain itu, untuk meningkatkan kinerja dan performansi sistem yang dibangun, digunakan *Principal Component Analysis* (PCA) untuk melakukan reduksi dimensi pada dataset yang digunakan. Pada tugas akhir ini juga dilakukan komparasi hasil pengujian jika terdapat penggunaan PCA dan tanpa penggunaan PCA. Pengujian dilakukan menggunakan skema *Speaker-Dependent* dan *Speaker-Independent*.

Hasil pengujian menunjukkan MFCC, PCA dan KNN berhasil diimplementasikan dengan performansi *Micro Average F1-Score* tertinggi sebesar 78,32% untuk *Speaker-Dependent* dengan parameter uji *sample rate* 16000 Hz, jumlah *filterbank* sebanyak 26, jumlah *principal components* sebanyak 64 dan nilai *k* sebesar 1. Sedangkan untuk *Speaker-Independent*, *Micro Average F1-Score* tertinggi sebesar 18,97% dengan parameter uji *sample rate* 16000 Hz, jumlah *filterbank* sebanyak 26, jumlah *principal components* sebanyak 104 dan jumlah *k* sebanyak 1.

Kata Kunci: Pengucapan Huruf Hijaiyah, Tanda Baca, *Speech Recognition*, MFCC, KNN, PCA