

ABSTRAK

Di era globalisasi ini, salah satu efek yang ada adalah terbentuknya bahasa internasional yaitu bahasa Inggris. Untuk menghadapi masalah tersebut sekaligus dapat memberikan dasar fondasi bagi generasi muda maka aplikasi *Speech to Text* ini dibuat dengan tujuan memperkenalkan bahasa Inggris. Aplikasi ini nantinya berguna untuk memperkenalkan kata-kata dalam bahasa Inggris. Selain dapat memperbanyak kosakata, aplikasi ini juga diperuntukkan untuk melatih melafalkan kata-kata tersebut dalam bahasa Inggris.

Aplikasi *Speech to Text (STT)* ini menggunakan metode *Hidden Markov Models (HMM) Hybrid* dengan *Gaussian Mixture Model (GMM)*. Tahap awal dari *Hidden Markov Models* adalah ketika ada suara, maka suara tersebut akan dikenali sebagai *Speech Signal*. Kemudian menggunakan *Feature extraction* yaitu *Mel-frequency cepstral coefficients (MFCC)* signal tersebut disimpan ke dalam *frame-frame* dan dicari nilai koefisien *cepstral*-nya. Selanjutnya tiap vector di kuantisasi yang menghasilkan output simbol observasi (*codebook*). Setiap kata yang tidak dikenal maka akan dimodelkan dengan *HMM/GMM* sehingga mendapatkan model kata. Untuk proses pengenalan kata maka akan dihitung probabilitas kemiripan pola dari tiap model *HMM/GMM* yang dimiliki dengan hasil dari observasi. Hasil probabilitas paling maksimum kemudian ditetapkan sebagai kata yang dikenali.

Pengujian ini dilakukan dengan mengubah nilai *feature MFCC* dan nilai *mixture GMM*. Performansi sistem diukur berdasarkan akurasi yang didapat dari parameter *WER (Word Error Rate)*. Setelah dilakukan pengujian terhadap sistem dengan beberapa skenario, diperoleh akurasi terbaik 100% dalam mengenali 10 kata. Akurasi ini diperoleh dari hasil pengujian dengan *MFCC 13 Feature* dan *GMM 6 mixture*.

Kata Kunci: *Speech to Text, Hidden Markov Model, Mel-frequency cepstral coefficient, Gaussian Mixture Model*