

# BAB I PENDAHULUAN

## 1.1 Latar Belakang Masalah

Perkembangan teknologi bisa memudahkan pekerjaan manusia. Pekerjaan manusia secara manual dapat digantikan dengan sistem otomatis. Salah satunya adalah sistem pengenalan ucapan manusia. Pengenalan suara dibagi menjadi dua jenis, yaitu *speech recognition* dan *speaker recognition*. *Speech recognition* adalah proses yang dilakukan komputer untuk mengenali ucapan yang diucapkan oleh seseorang, sedangkan *speaker recognition* yaitu merupakan pengenalan identitas orang dari suara yang diucapkannya.

Dengan pengenalan suara, manusia bisa berkomunikasi dengan sistem dengan suara. Hal ini memudahkan manusia berhubungan dengan sistem. Beberapa implementasi pengenalan ucapan adalah *Speech Dialing System*, *Spoken Dialogue System*, *Speech to Text Translation System* dan lain sebagainya [16]. Pengenalan suara merupakan salah satu bidang yang mempunyai sifat unik. Untuk setiap bahasa mempunyai karakteristik yang berbeda-beda. Penelitian yang dilakukan pada satu bahasa, belum tentu bisa digunakan pada bahasa yang lainnya.

Teknik yang paling terkenal untuk membangun pengenalan ucapan adalah *Hidden Markov Model* (HMM). HMM bagus untuk pengenalan ucapan karena karakteristik dari sinyal suara bisa dimodelkan dengan baik oleh HMM yaitu memodelkan sinyal suara dalam bentuk state-state yang sebelumnya diubah ke dalam bentuk vektor. HMM merupakan teknik yang paling bagus saat ini, karena berhasil memberikan *error rate* yang paling kecil dibandingkan dengan teknik lain [8].

Ada beberapa teknik *speech recognition*, diantaranya *speech recognition based on word* dan *speech recognition based on phoneme*. Pengenalan suara berbasis kata berarti satu kata dianggap satu pola, sedangkan pengenalan suara berbasis fonem, satuan terkecil dari pengenalan kata adalah fonem. Permasalahan yang masih ada adalah pengenalan ucapan dengan perbedaan kecepatan berbicara, pengenalan ucapan dengan berbeda speaker dan pengenalan ucapan di tempat yang berderau. Hal tersebut bisa meningkatkan *error rate*. Pada tugas akhir ini membahas bagaimana perbedaan kecepatan berbicara berpengaruh pada akurasi. Selain itu juga membahas bagaimana perbedaan usia speaker berpengaruh dan bagaimana penanganannya menggunakan data training yang bervariasi. Tugas akhir ini juga akan membahas kebisingan akan berpengaruh pada akurasi dan bagaimana penanganannya agar akurasi tidak menurun menggunakan *Discret Wavelet Transform* (DWT).

Sistem pengenalan suara sudah banyak dikembangkan di perangkat PC yang mempunyai RAM dan CPU yang besar, karena sistem ini membutuhkan resource yang cukup besar. Saat ini perangkat mobile sedang menjamur di dunia. Pengenalan suara mulai dikembangkan di perangkat *mobile*, yang mempunyai RAM dan CPU terbatas. Pada tugas akhir ini akan dibangun sistem pengenalan suara pada perangkat bergerak berbasis Android.

## 1.2 Perumusan Masalah

Berdasarkan latar belakang masalah yang telah dikemukakan sebelumnya, maka masalah yang dirumuskan adalah

1. Bagaimana mengimplementasikan *speech recognition* dengan metode *Hidden Markov Model* pada perangkat bergerak berbasis Android?
2. Bagaimana membangun *grammar* untuk pengenalan nama-nama tempat?
3. Bagaimana mengimplementasikan *speech recognition* ke dalam aplikasi android untuk *speech command*?
4. Bagaimana akurasi yang dihasilkan ketika *speech recognition* dilakukan pada lingkungan berderau?
5. Bagaimana mengimplementasikan *Wavelet Transform Diskret* pada *speech recognition*?
6. Bagaimana akurasi yang dihasilkan oleh penguji yang berbeda-beda berdasarkan usia?
7. Bagaimana kecepatan pengucapan berpengaruh pada akurasi pengenalan ucapan?

Adapun batasan masalahnya adalah :

1. Implementasi dilakukan pada perangkat *mobile* berbasis android.
2. *Speech recognition* diimplementasikan untuk *speech command* tempat-tempat di Bandung yang berjumlah berkisar 150 tempat.
3. Menggunakan *tools pocketsphinx*.
4. Bahasa yang digunakan adalah Bahasa Indonesia

## 1.3 Tujuan

Tujuan dari tugas akhir ini adalah :

1. Mengimplementasikan *speech recognition* pada perangkat mobile berbasis Android.
2. Membangun *speech recognition* untuk *speech command* pada perangkat bergerak berbasis Android.
3. Mengimplementasikan *Discrete Wavelet Transform* untuk reduksi derau pada *speech recognition*.
4. Menganalisa akurasi sistem *speech recognition* yang dibangun pada perangkat mobile berbasis android yang diuji pada lingkungan berderau, *speaker* berbeda usia dan kecepatan *speaker* berbicara.

## 1.4 Hipotesa

1. Kebisingan, kecepatan dan perbedaan usia *speaker* berpengaruh pada penurunan akurasi, tetapi masih bisa ditangani sehingga akurasi tidak menurun terlalu drastis.
2. Implementasi *speech recognition* dalam perangkat mobile dan PC berbeda, karena perangkat mikro mempunyai resources yang terbatas.

## 1.5 Metodologi Penyelesaian Masalah

Metodologi yang digunakan untuk penyelesaian tugas akhir ini adalah :

1. Identifikasi Masalah  
Pada tahap ini, akan dilakukan identifikasi mengenai permasalahan yang dibahas dalam tugas akhir ini, yaitu mengenai pengenalan suara pada perangkat bergerak, pengenalan suara berbasis fonem, *Hidden Markov Model*, *Wavelet Diskret Transform*, tools yang digunakan, dan pengaruh kecepatan, perbedaan usia serta lingkungan berderau mempengaruhi akurasi pengenalan suara.
2. Studi Literatur  
Pada tahap ini, akan dilakukan pencarian terhadap materi-materi guna mendukung penulisan tugas akhir ini. Referensi yang dicari antara lain adalah jurnal, artikel, maupun buku mengenai pengenalan suara berbasis fonem, Android, *Hidden Markov Model*, *Wavelet Diskret Transform*, dan bagaimana meningkatkan akurasi pada lingkungan berderau, pengguna dengan perbedaan usia, dan kecepatan pengucapan, agar akurasi tidak turun drastis.
3. Perancangan Sistem  
Pada tahap ini, akan dilakukan perancangan sebuah sistem perangkat lunak dengan menggambarkan modul-modul perangkat lunak.
4. Implementasi Sistem, Pengujian, dan Analisis Hasil  
Mengimplementasikan sistem pengenalan suara berbasis fonem dengan *Hidden Markov Model* pada perangkat bergerak berbasis android, kemudian mengimplementasikan pada aplikasi *speech command*. Kemudian mengimplementasikan *Wavelet Diskret Transform* untuk reduksi derau. Tahap selanjutnya adalah pengujian perangkat lunak yang dibangun dan menganalisis hasil dari pengujian.
5. Perumusan Kesimpulan dan Penyusunan Buku Tugas Akhir  
Pada tahap ini, akan dilakukan perumusan kesimpulan berdasarkan analisis dari hasil implementasi sistem yang telah dilakukan pada tahap sebelumnya. Setelah diambil kesimpulan, akan dilakukan penyusunan skripsi dan pengumpulan dokumentasi dengan mengikuti aturan-aturan yang telah ditetapkan oleh institusi.