

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Setiap individu manusia mempunyai suara yang berbeda-beda. Suara manusia ini dihasilkan dari getaran udara yang muncul dari perpaduan antara paru-paru, katup tenggorokan, dan pita suara. Karena ketebalan masing-masing pita suara unik dan berbeda-beda. Salah satu teknologi yang sedang dikembangkan dalam dunia teknologi dan informasi adalah biometrik. Aspek biologis yang dimanfaatkan dalam biometrik antara lain wajah, sidik jari, iris mata, DNA dan suara. Keunggulan biometrik dalam hal otentifikasi dibandingkan dengan *password* karakter adalah kemampuannya membedakan satu individu dengan individu lain secara akurat, sulit diduplikasi, dan tidak mudah hilang.

Pada tugas akhir ini, dilakukan perancangan sistem pengenalan penutur (*speaker recognition*) dengan menggunakan metode *Mel Frequency Cepstrum Coefficient* (MFCC) sebagai ekstraksi ciri dan *Hidden Markov Model* (HMM) sebagai klasifikasi. Pada penelitian ini dilakukan perekaman oleh sepuluh pembicara dengan total data untuk *training* 210 file suara dan saat proses pengambilan suara dibedakan berdasarkan kata dan jarak yang nantinya akan di jadikan input pada sistem. Pada saat proses perekaman suara dilakukan di dalam ruangan untuk meminimalisir noise.

Tingkat akurasi pengenalan speaker bergantung pada jumlah data dan jarak. Karena semakin banyak data yang akan di *training* dan jarak saat perekaman yang terlalu jauh maka tingkat akurasi sistem akan berkurang. Oleh karena itu penulis pada penelitian kali ini akan membahas lebih dalam lagi tentang *speaker recognition* dengan variabel penelitiannya yaitu pada jumlah data, proses pelatihan hanya bergantung kepada teks dan jarak pengambilan rekaman yang bervariasi.

Tabel 1. 1 Hasil penelitian sebelumnya

No	Ekstraksi Ciri	Klasifikasi	Tahun	Hasil Akurasi	Sumber
1	MFCC	HMM	2014	100%	[1]
2	LPCC	HMM	2006	98,9%	[2]
3	MFCC	HMM	2013	52,9%	[3]
4	MFCC	HMM	2010	10% - 99%	[4]
5	MFCC	VQ	2016	48% - 63%	[5]
	MFCC	HMM	2016	65% - 90%	
6	MFCC	LDA	2012	70,5%	[6]
	MFCC	MLP	2012	82,1%	
	MFCC	HMM	2012	100%	
	LPCC	LDA	2012	55,6%	
	LPCC	MLP	2012	75,1%	
	LPCC	HMM	2012	94%	

MFCC : *Mel-Frequency Cepstrum Coefficient*

MLP : *Multiplayer Perception*

LDA : *Linear Discriminant Analysis*

HMM : *Hidden Markov Model*

VQ : *Vector Quantization*

LPCC : *Linear Predictive Coding Coefficients*

Pada penelitian sebelumnya oleh Devi Handaya dengan judul “Perbandingan Metode Kuantisasi Vektor dan Model Markov Tersembunyi Pada Pengenalan Pembicara Berbahasa Indonesia”. Berdasarkan pengujian selama sepuluh kali dengan 4 orang pembicara. Tingkat akurasi dengan pengujian kata, kalimat sederhana, dan kalimat lengkap. Akurasi metode Kuantisasi Vektor pada pengujian kata “saya” 20% - 90% sedangkan metode HMM 80% - 90%. Akurasi pada kalimat sederhana “saya sedang belajar” metode Kuantisasi Vektor 0% - 90% sedangkan metode HMM 70% - 100%. Akurasi pada kalimat lengkap “saya berangkat menggunakan bus ke kampus” metode Kuantisasi Vektor 0% - 90% sedangkan metode HMM 70% - 100%.

Penelitian oleh Abdallah S dengan judul “*Text-Independent Speaker Identification Using Hidden Markov Model*” dengan tingkat performasi untuk pengenalan speaker 80% - 100% dengan jumlah speaker dari 3 - 40.

Penelitian oleh Agus Buono dengan judul “Sistem Identifikasi Pembicara Berbasis Power Spektrum Menggunakan Hidden Markov Model” dapat mencapai tingkat keberhasilan 99% dengan data tanpa penambahan noise sedangkan untuk data yang ditambah dengan noise mencapai tingkat keberhasilan 10%.

Kemudian penelitian oleh Maryati Gultom dengan judul “Rancang Bangun Aplikasi Pengenal Penutur Menggunakan Metode Hidden Markov Model” secara *realtime* dapat mencapai tingkat keberhasilan 30% dan *file* rekaman tingkat keberhasilan 100%.

Hasil dari beberapa penelitian tersebut dapat diketahui bahwa semakin banyak jumlah speaker maka akurasi sistem semakin menurun dan panjang dataset berpengaruh terhadap akurasi sistem karena fitur informasi dari pembicara semakin detail. Penambahan noise pada nilai tertentu dapat menurunkan tingkat keberhasilan secara drastis.

## 1.2 Rumusan Masalah

Berdasarkan deskripsi latar belakang, beberapa hal yang dapat dirumuskan yaitu:

1. Bagaimana cara merancang sistem yang mampu mengekstrasi ciri dari masukan sinyal suara subjek menggunakan metode MFCC (*Mel Frequency Cepstrum Coefficient*) dengan metode klasifikasi HMM (*Hidden Markov Model*) ?
2. Bagaimana menganalisis performansi sistem berdasarkan hasil akurasi ?
3. Parameter – parameter apa saja yang memberikan akurasi terbaik ?

## 1.3 Tujuan Penelitian

Tujuan dari penelitian tugas akhir ini antara lain:

1. Mengekstrasi ciri masukan sinyal suara subjek menggunakan metode MFCC dan di klasifikasikan dengan metode HMM.
2. Mencari parameter yang sesuai supaya tingkat akurasi sistem lebih akurat.
3. Pengujian jarak perekaman suara terhadap akurasi sistem.

## 1.4 Batasan Masalah

Berikut batasan yang digunakan dalam penelitian tugas akhir ini:

1. Sinyal bicara diambil dari 10 orang dari usia 21 – 30 tahun.
2. *Sample rate* yang digunakan adalah 48000 Hz.
3. Jarak pengambilan rekaman (10cm, 30cm dan 50cm).

## 1.5 Metodologi Penelitian

Berikut metode penelitian yang digunakan dalam penelitian tugas akhir ini:

### 1. Studi literatur

Metode ini bertujuan untuk mempelajari dan memahami berbagai referensi yang terkait dengan tugas akhir ini, antara lain meliputi studi pustaka tentang:

- a) Speaker Recognition
- b) MFCC (*Mel Frequency Cepstrum Coefficient*)
- c) HMM (*Hidden Markov Model*)

### 2. Perancangan model sistem

Metode ini bertujuan untuk merancang model sistem *speaker recognition* secara teori antara lain menyusun diagram alir yang tepat dan menyesuaikan format keluaran data di setiap proses.

### 3. Implementasi perangkat lunak

Metode ini bertujuan untuk menerapkan rancangan model sistem ke dalam software yaitu Matlab R2015a.

### 4. Pengambilan data

Pengambilan data adalah kegiatan mengumpulkan data suara yang terdiri dari 5 orang perempuan dan 5 orang laki-laki dengan usia berbeda yang masing-masing satu orangnya mempunyai 4 *files* kata dari total kata tujuh dengan jarak pengambilan yang berbeda.

### 5. Pengujian dan analisis performansi sistem

Metode ini bertujuan untuk menguji sistem *speaker recognition* dengan memberi input sinyal bicara dan dianalisis akurasi kesesuaian antara data latih dan data uji serta dianalisis juga kecepatan komputasi sistem.

### 6. Perbaikan sistem

Metode ini bertujuan untuk memperbaiki model sistem *speaker recognition* baik secara teori maupun implementasinya pada Matlab R2015a apabila ditemukan ketidaksesuaian dengan hasil yang diinginkan, contohnya noise yang terlalu besar dan akurasi dibawah 50%.

## 7. Mengambil kesimpulan

Metode terakhir yang dilakukan dalam tugas akhir ini adalah pengambilan kesimpulan atas penelitian yang telah dilakukan.