

**BAB I****PENDAHULUAN****1.1 Latar Belakang**

Pengenalan ucapan atau speech recognition merupakan teknologi yang memungkinkan komputer untuk mengenali ucapan manusia. Teknologi ini telah mengalami kemajuan yang signifikan dalam beberapa dekade terakhir, terutama dalam mengenali bahasa-bahasa yang umum digunakan seperti Bahasa Inggris. Namun, bahasa-bahasa daerah atau lokal seringkali tidak mendapatkan perhatian yang sama dalam pengembangan teknologi speech recognition. Bahasa daerah memiliki kekhasan sendiri dalam pengucapannya, seperti vokal dan konsonan yang unik, intonasi khas, dan kosa kata yang berbeda dari bahasa-bahasa yang lebih umum digunakan.

Implementasi speech recognition yang mampu mengenali bahasa daerah memiliki manfaat yang signifikan. Pertama, ini memungkinkan akses yang lebih mudah dan inklusif terhadap teknologi komunikasi untuk komunitas yang menggunakan bahasa daerah sebagai bahasa ibu. Hal ini akan memungkinkan pengguna bahasa daerah untuk berkomunikasi dengan komputer dan perangkat elektronik secara lebih alami dan efisien. Selain itu, pengenalan bahasa daerah dalam speech recognition juga dapat mendukung pelestarian dan pengembangan bahasa dan budaya lokal. Bahasa daerah seringkali terancam punah karena kurangnya dukungan teknologi dan kurangnya sumber daya untuk pengajaran dan penggunaan bahasa tersebut. Dengan adanya speech recognition untuk bahasa daerah, system dapat mengenali ucapan yang diucapkan oleh user dan system akan mengenali aksen daerah apa yang diucapkan oleh user. Sehingga, memungkinkan penelitian ini dapat dijadikan sebagai media pembelajaran.

Oleh karena itu, peneliti berinisiatif ingin mengembangkan aplikasi ini yang berjudul *“Implementasi Speech Recognition Pada Pengucapan Bahasa Inggris Oleh Pembicara Non-Native Dengan Menggunakan Mel–Frequency Cepstral Coefficients*

*Berbantuan Recurrent Neural Network*". Walaupun pada peneliti sebelumnya sudah banyak mengambil tentang penelitian ini. Salah satunya oleh, *Teguh Puji Laksono* membuat penelitian tentang Speech To Text Untuk Bahasa Indonesia menggunakan Metode MFCC. Pada penelitian sebelumnya, hanya didapatkan Akurasi yang dihasilkan speech to text pada data Bahasa Indonesia adalah 65% sedangkan, untuk data Bahasa Jawa adalah 57%. Oleh karena itu, pada penelitian ini penulis berinisiatif ingin memfokuskan pada bagaimana proses suara bisa diekstraksi ke MFCC, bagaimana cara mengklasifikasikan model dari masing-masing class, dan bagaimana cara proses mengenali pola suaranya.

## **1.2 Identifikasi Masalah**

Berdasarkan yang telah dijelaskan, maka masalah pada uraian di atas adalah sebagai berikut:

1. Banyaknya masyarakat tidak memahami pengucapan yang diucapkan oleh seorang non-native dalam berbahasa Inggris.
2. Banyaknya masyarakat tidak mengetahui aksent dari berbagai suku bangsa yang diucapkan oleh seorang non-native.

## **1.3 Ruang Lingkup Masalah**

Dari uraian latar belakang masalah di atas terdapat permasalahan, yaitu class yang diberikan sebagai objek penelitian masih sedikit, yaitu hanya 5 daerah saja pada algoritma klasifikasi sistem speech recognition dengan kasus untuk mengenali kata Daerah dalam Bahasa Inggris yaitu, Betawi, Jawa Barat, Jawa Tengah, Maluku, Palembang.

## **1.4 Rumusan Masalah**

Berdasarkan latar belakang masalah yang telah diuraikan di atas, maka terdapat beberapa rumusan masalah yang ditemukan dalam penelitian ini, yaitu:

1. Bagaimana cara mengenali aksent daerah pada pengucapan Bahasa Inggris dari seorang non-native?
2. Bagaimana cara metode Mel-Frequency Cepstral Coefficient (MFCC) dapat

mengenali pengucapan Bahasa Inggris dari seorang non-native dengan menggunakan algoritma Recurrent Neural Network (RNN) ?

### **1.5 Tujuan Penelitian**

1. Pembuatan model dalam pola pronunciation bahasa Inggris oleh beberapa suku bangsa di Indonesia.
2. Menerapkan metode Mel-Frequency Cepstral Coefficient (MFCC) dengan menggunakan algoritma Recurrent neural network (RNN) untuk mengenali aksent daerah yang diucapkan oleh beberapa suku bangsa di Indonesia dalam berbahasa Inggris.

### **1.6 Batasan Masalah**

1. Dataset yang digunakan adalah rekaman suara dalam Bahasa Inggris yang diambil dari 5 suku yaitu, Betawi, Sumatera Selatan, Maluku, Jawa Timur, dan Jawa Barat.
2. Model yang dihasilkan adalah pola dari Speech Recognition adalah 5 suku bangsa. Apabila pengujian dilakukan selain dari 5 suku bangsa tersebut. Maka, hasilnya tidak dapat dikenali.
3. Total jumlah dataset dalam penelitian ini adalah 2.722 sample suara.
4. Total dataset di atas dibagi menjadi 2 label, yaitu data train dan data test. Pada data train terdapat 2.178 sample suara, sedangkan data test terdapat 544 sample suara.

### **1.7 Kerangka Berfikir**

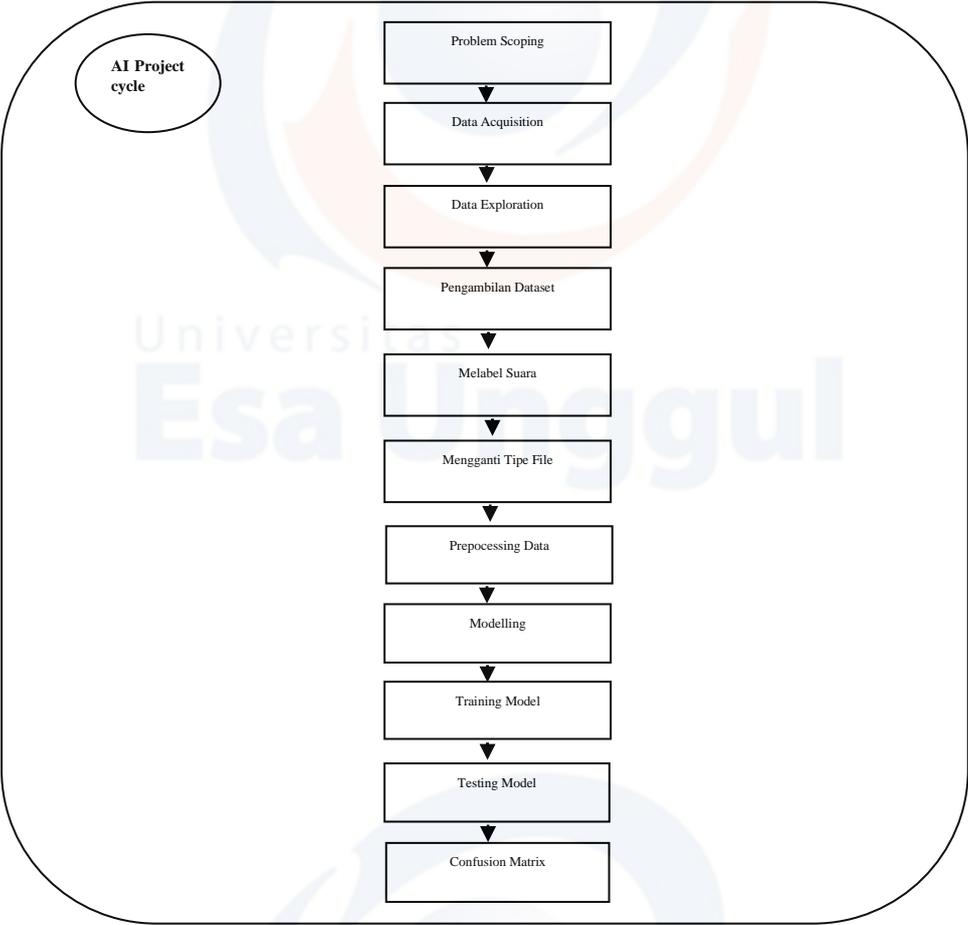
Kerangka berfikir adalah gambaran sebuah penelitian yang digambarkan dari awal penelitian hingga akhir. Berikut merupakan kerangka berfikir pada penelitian yang dijelaskan pada gambar sebagai berikut:

Topic: *“Implementasi Speech Recognition Pada Pengucapan Berbahasa Inggris Oleh Seorang Non-Native Dengan Menggunakan Mel-Frequency Cepstral Coefficients Berbantuan Recurrent Neural Network”*

- Identifikasi Masalah**
1. Banyaknya masyarakat tidak memahami pengucapan yang diucapkan oleh seorang non-native dalam berbahasa inggris.
  2. Banyaknya masyarakat tidak mengetahui aksen dari berbagai suku bangsa yang diucapkan oleh seorang non-native.

- Tujuan Penelitian**
1. Pembuatan model dalam pola pronunciation bahasa inggris oleh beberapa suku bangsa di Indonesia.
  2. Menerapkan metode Mel-Frequency Cepstral Coefficient (MFCC) dengan menggunakan algoritma Recurrent neural network (RNN) untuk mengenali aksen daerah yang diucapkan oleh beberapa suku bangsa di Indonesia dalam berbahasa inggris.

- Usulan Solusi**
1. Pengumpulan Data yang Representatif = Diperlukan dataset yang mencakup variasi luas dari suku bangsa di Indonesia, berbicara dalam bahasa inggris. Data ini harus mencakup variasi dialek, aksen, dan pola bicara yang mungkin muncul dari berbagai kelompok etnis.
  2. Pra-Pemrosesan Data = Sebelum mengaplikasikan MFCC, data audio perlu diproses untuk penghapusan noise, normalisasi, dan segmentasi. Ini membantu memastikan bahwa fitur yang diekstraksi berkualitas tinggi dan bebas dari gangguan.
  3. Ekstraksi Fitur MFCC = MFCC adalah fitur yang sangat efektif untuk merepresentasikan sinyal audio. Tetapi, Anda juga bisa mempertimbangkan untuk memasukan fitur lain seperti delta MFCC dan delta-delta MFCC untuk memperkaya informasi dalam fitur.
  4. Pembagian Data = Bagi dataset menjadi set pelatihan, validasi, dan pengujian. Ini penting untuk mengukur kinerja model dengan benar dan mencegah overfitting. Pastikan setiap kelompok etnis diwakili dengan baik di setiap bagian dataset.
  5. Validasi dan Evaluasi = Gunakan set validasi untuk mengoptimalkan model Anda. Jika model terbaik telah ditemukan, gunakan set pengujian yang terpisah untuk mengevaluasi kinerja akhir model.



## **1.8 Sistematika Penulisan Tugas Akhir**

Dalam penyusunan laporan tugas akhir ini, sistematika penulisan dibagi menjadi beberapa bab sebagai berikut:

### **BabI Pendahuluan**

Bab ini berisi tentang latar belakang permasalahan dari penelitian deteksi suara menggunakan RNN. Dari latar belakang tersebut kemudian dibuat rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penelitian.

### **BabII Studi Literatur**

Bab ini berisi tentang teori-teori dan konsep yang berhubungan dengan penelitian. Dengan adanya teori dan konsep digunakan untuk mendukung pemecahan permasalahan. Teori-teori dan konsep didapat dari sumber-sumber seperti, jurnal, buku, dan artikel.

### **BabIII Metodologi Penelitian**

Penelitian Bab ini berisi tahapan-tahapan yang dilakukan pada penelitian. Tahapan-tahapan tersebut yaitu pengumpulan data, membagi data, dan tahapan yang ada pada metode RNN seperti membangun arsitektur model, melatih model, dan melakukan pengujian.

### **BabIV Hasil & Pembahasan**

Bab ini berisi proses penelitian yang dijelaskan juga kode-kode yang digunakan. Dan terdapat hasil dari pengujian dari model yang telah dibangun.

### **BabV Kesimpulan & Saran**

Bab ini berisi kesimpulan yang diperoleh dari hasil penelitian dan analisa data yang telah dilakukan dan saran-saran yang dapat diterapkan dari hasil yang dapat menjadi masukan yang berguna kedepannya.