

BAB I

PENDAHULUAN

1.1 Latar Belakang

Bahasa sunda adalah bahasa daerah yang dipakai di Jawa Barat, sebagian Jawa Tengah sebelah barat, dan sebagian Daerah Ibu Kota Jakarta. Bahasa Sunda memiliki variasi Bahasa, atau undak-unsuk, yang terbentuk karena hubungan antara pembicara dan lawan bicara. Undak-usuk dalam bahasa Sunda (undak-usuk; tingkatan bahasa; unggah-ungguh; tata krama) ada 3 tingkatan yaitu basa kasar, basa loma (akrab), basa hormat/lemes (santun).

Dialek adalah varian dari sebuah bahasa menurut pemakai. Berbeda dengan ragam bahasa yaitu varian ini masih terdiri dari satu bahasa yang sama. Variasi ini berbeda satu sama lain, tetapi masih banyak menunjukkan kemiripan sehingga belum pantas disebut bahasa yang berbeda, pada bahasa sunda ini memiliki 6 dialek yaitu : Barat (Banten), Dialek Utara, Dialek Selatan (Priangan), Dialek Tengah Timur, Dialek Timur Laut (termasuk Bahasa Sunda Cirebon), Dialek Tenggara.

Walaupun bahasa sunda digunakan oleh banyak orang, riset *Automatic Speech Recognition* (ASR) untuk mengenali bahasa Sunda masih terbelang sedikit dibanding dengan ASR untuk mengenali bahasa lainnya. Pengenalan ASR bahasa Sunda memiliki beberapa kendala diantaranya : kurangnya data suara, banyaknya dialek bahasa Sunda, dan lainnya. Dalam hal ini, bagaimana suatu sistem ASR dapat memahami dan mengenal bahasa Sunda dialek tengah timur menggunakan metode *deep neural network*.

Pada umumnya, orang Indonesia yang berbeda daerah terkendala saat berkomunikasi dikarenakan ketidaktahuan dari bahasa daerah tersebut, sehingga dapat menimbulkan kesalahpahaman saat berkomunikasi. Dalam tugas akhir ini, metode *deep neural network* akan diterapkan pada ASR dengan menggunakan data ucapan bahasa sunda untuk memudahkan komunikasi antar daerah di Indonesia khususnya masyarakat Jawa Barat. Dengan demikian, sistem yang dibangun diharapkan dapat memiliki performansi yang baik.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, dapat dirumuskan permasalahan yang dihadapi adalah bagaimana membuat sistem *Automatic Speech Recognition* untuk pengenalan Bahasa Sunda dialek

Tengah Timur dengan metode *Deep Neural Network* dan memproses suatu masukan berupa suara ke suatu sistem agar dimengerti oleh sistem dan dapat mengklasifikasikan masukan kedalam Bahasa Sunda dialek Majalengka serta memberikan bobot nilai untuk masukan tersebut.

1.3 Tujuan

Tujuan yang ingin dicapai dari tugas akhir ini adalah :

1. Mengimplementasikan metode *deep neural network pada Automatic Speech Recognition* bahasa Sunda dialek tengah timur (Majalengka).
2. Mengukur akurasi kinerja *deep neural network*.
3. *Mengenal di dialek Majalengka dengan sistem automatic speech recognition.*

1.4 Batasan Masalah

Dengan merujuk pada rumusan masalah yang telah dijelaskan sebelumnya dan untuk lebih memfokuskan topik penelitian maka penulis membuat batasan masalah yang akan dibahas. Adapun batasan masalah tersebut adalah sebagai berikut :

1. Sampel yang diambil adalah kosa kata Bahasa Sunda.
2. Sampel akan diambil langsung dari orang Majalengka.
3. *Sampel kata yang digunakan hanya 6 kalimat atau 35 kata.*

1.5 Metodologi Penyelesaian Masalah

Metode yang digunakan pada penelitian ini memiliki lima tahapan, yaitu :

1. Studi Literatur
Mempelajari permasalahan yang ada pada *Automatic Speech Recognition* sistem dan mempelajari metode-metode yang dapat digunakan untuk menyelesaikan permasalahan tersebut.
2. Analisis dan Perancangan Sistem
Menganalisis kebutuhan dan perancangan arsitektur sistem yang digunakan dalam menyelesaikan permasalahan yang diangkat.
3. Pengambilan Data dan Pengujian Sistem
Pengambilan data dilakukan dengan mencari orang sunda di sekitar daerah Majalengka sebagai narasumber dan pengujian sistem dilakukan setelah dataset terkumpul.

4. Analisis Hasil Pengujian

Pada tahap ini dilakukan analisis dari hasil keluaran sistem tentang kinerja sistem dan tingkat akurasi dari metode yang digunakan..

5. Pembuatan Laporan

Tahap ini merupakan tahap akhir dari penelitian dimana dilakukan penyusunan laporan, dokumentasi dan kesimpulan berdasarkan analisis hasil pengujian.

1.6 Sistem Matika Penulisan

Sistematika penulisan pada laporan tugas akhir ini adalah sebagai berikut:

1. Pendahuluan

Menjelaskan latar belakang, perumusan masalah, batasan masalah, tujuan, metodologi penelitian, dan sistematika penulisan.

2. Landasan Teori

Menjelaskan dasar teori yang diambil dari beberapa kutipan buku, konsep dasar sistem, dan definisi lainnya yang berkaitan dengan permasalahan dan pembahasan dari *automatic speech recognition*, *deep neural network*, *mel-frequency ceptral coefficients*.

3. Perancangan Sistem

Menjelaskan perancangan dari sistem yang dibangun, dimulai dari preprocessing, ekstraksi ciri hingga tahapan klasifikasi dan rencana pengujian sistem yang dilakukan

4. Implementasi dan Pengujian

Membahas hasil pengujian sistem yang telah dilakukan dengan menganalisis skenario pengujian perubahan nilai parameter yang mempengaruhi performansi sistem yang dilakukan dalam mencapai tujuan akhir yang diinginkan pada tugas akhir ini.

5. Kesimpulan dan Saran

Membahas kesimpulan akhir yang didapat dari pengerjaan tugas akhir, serta saran yang diusulkan dalam mendukung pengembangan yang lebih lanjut.