BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Speech Recognition adalah proses menginterpretasikan ucapan manusia agar dikenali maknanya oleh komputer [1]. Dalam sebuah sistem *speech recognition* diperlukan sebuah algoritma klasifikasi untuk mengenali makna dari sinyal suara. Terdapat berbagai macam algoritma klasifikasi yang dapat digunakan pada sebuah *speech recognition*, salah satunya adalah *Elman Recurrent Neural Network* (ERNN).

Elman Recurrent Neural Network (ERNN) dapat bekerja dengan baik dalam *speech recognition* dikarenakan kemampuannya untuk mengenali pola yang berbeda berdasarkan waktu[2]. Terdapat beberapa penelitian sebelumnya yang menggunakan ERNN sebagai algoritma klasifikasi dalam sistem *speech recognition*. Diantaranya adalah penelitian yang dilakukan Robert B. Allen mengenai ERNN untuk mengenali ucapan berupa kata dalam Bahasa Inggris yang menghasilkan akurasi mencapai 87.5% [3]. Pada penelitian lainnya yang dilakukan oleh Yousef Alotaibi [4], ERNN menghasilkan akurasi 99.47% dan 96.46% untuk mengenali ucapan berupa angka dalam Bahasa Arab. Dengan bahasa yang sama namun kasus yang berbeda penelitian yang dilakukan oleh M. M. El Choubassi menyatakan bahwa ERNN mampu menghasilkan akurasi 90% pada sistem *Automatic Speech Recognition* (ASR) yang dapat mengenali 6 kata dalam Bahasa Arab [5]. Pada penelitian yang dilakukan Muhammad Fachrie [6], ERNN mampu menghasilkan akurasi paling tinggi sebesar 99.30% pada kasus pengenalan ucapan digit 0 sampai 9. Dari beberapa penelitian tersebut ERNN terlihat mampu memberikan hasil yang baik ketika digunakan untuk *speech recognition* dalam beberapa kasus yang berbeda, namun performa dari ERNN untuk kasus *speech recognition* kata dalam Bahasa Indonesia belum diketahui.

Berdasarkan uraian di atas, maka pada penelitian ini akan menggunakan ERNN sebagai algoritma klasifikasi pada speech recognition dengan kasus kata

dalam Bahasa indonesia. Harapan dari penelitian ini adalah untuk mengetahui bagaimana akurasi yang dihasilkan ERNN untuk mengenali ucapan berupa kata dalam Bahasa Indonesia.

1.2 Perumusan Masalah

Dari uraian latar belakang masalah di atas, terdapat permasalahan yaitu belum diketahuinya akurasi yang dihasilkan oleh metode *Elman Recurrent Neural Network* (ERNN) sebagai algoritma klasifikasi pada sistem *speech recognition* dengan kasus untuk mengenali kata dalam Bahasa Indonesia.

1.3 Maksud dan Tujuan

1.3.1 Maksud

Maksud dari penelitian ini adalah untuk mengimplementasikan *Elman Recurrent Neural Network* (ERNN) ke dalam sistem *speech recognition* kata dalam Bahasa Indonesia.

1.3.2 Tujuan

Adapun tujuan dari penelitian ini adalah untuk mengetahui akurasi dari metode *Elman Recurrent Neural Network* (ERNN) untuk mengenali ucapan berupa kata dalam Bahasa Indonesia.

1.4 Batasan Masalah

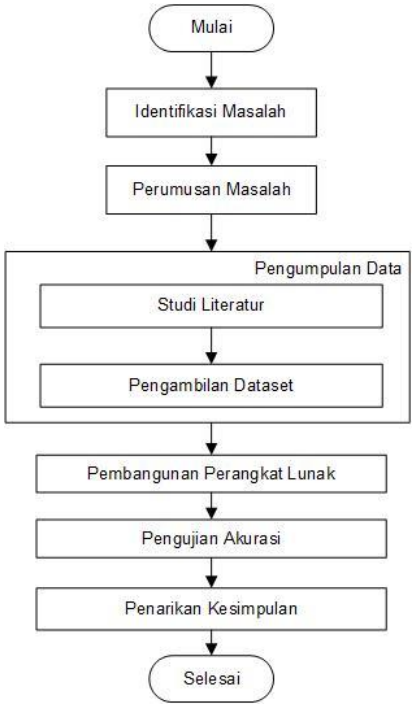
Agar penelitian yang dilakukan lebih terarah dan mencapai sasaran yang ditentukan, maka dapat dibuat batasan masalah sebagai berikut :

1. *Dataset* merupakan hasil rekaman suara 40 orang laki-laki. Setiap pembicara akan mengucapkan 10 kata sebanyak 4kali, sehingga akan dihasilkan 1600 sampel suara (40 orang x 10 kata x 4 kali ucapan).
2. Durasi file rekaman untuk setiap kata adalah 1 detik, dengan proses perekaman suara dilakukan menggunakan *software* audacity.
3. Format file yang digunakan pada *dataset* adalah .wav dengan frekuensi sampling 16000 Hz dan *channel* mono.

4. Pengambilan *dataset* dilakukan dengan cara merekam suara pada lingkungan yang tenang (*low noise environment*).
5. Metode ekstraksi yang akan digunakan adalah MFCC.
6. Program dapat mengenali 10 kata dalam Bahasa Indonesia (kata yang dikenali merupakan kata kerja dengan *rank* tertinggi pada *Leipzig Corpora Collection* [7] yaitu, “ada”, ”adalah”, ”beri”, ”bisa”, ”guna”, ”jadi”, ”laku”, ”milik”, ”rupa”, ”sebut”).

1.5 Metodologi Penelitian

Metode penelitian yang digunakan pada penelitian ini menggunakan metodologi kuantitatif. Penelitian ini dimulai dengan identifikasi masalah, kemudian masalah tersebut dirumuskan dan dibatasi. Setelah didapatkan rumusan masalah maka penelitian ini akan dilanjutkan dengan melakukan pengumpulan data yang terdiri dari studi literatur yang nantinya akan menjadi landasan teori dalam penelitian ini dan pengambilan *dataset* yang nantinya akan digunakan pada tahap pengujian. Setelah dilakukan pengumpulan data maka akan dilakukan pembangunan perangkat lunak menggunakan metodologi *sequential linear model*. Pada tahap pembangunan perangkat lunak akan dilakukan analisis mengenai algoritma yang dibutuhkan, kebutuhan fungsional dan non-fungsional hingga desain terhadap perangkat lunak yang akan dibangun. Setelah perangkat lunak dibangun maka akan dilakukan pengujian akurasi *Elman Recurrent Neural Network* (ERNN) dari perangkat lunak yang telah dibangun. Setelah dilakukan pengujian akurasi maka akan dilakukan penarikan kesimpulan terhadap penelitian yang telah dilakukan. Adapun tahapan penelitian akan dilakukan dengan alur seperti pada Gambar 1.1.



Gambar 1.1 Alur Tahapan Penelitian

1.5.1 Identifikasi Masalah

Pada tahap ini akan dilakukan pemaparan mengenai beberapa penelitian sebelumnya sehingga dapat dilakukan identifikasi masalah mengenai *speech recognition* dan *Elman Recurrent Neural Network* sebagai algoritma klasifikasi.

1.5.2 Perumusan Masalah

Pada tahap ini akan dilakukan perumusan masalah mengenai implementasi *Elman Recurrent Neural Network* pada *speech recognition*.

1.5.3 Pengumpulan Data

Tahap pengumpulan data terdiri atas studi literatur dan pengambilan *dataset*.

1.5.3.1 Studi Literatur

Pada tahap ini akan dilakukan pembelajaran untuk menambah pamahaman mengenai konsep-konsep *speech recognition*, *Mel Frequency Cepstrum Coefficient*, *Elman Recurrent Neural Network*, *Backpropagation Through Time* sebagai algoritma pelatihan. Studi literatur dilakukan dengan cara membaca referensi seperti buku, jurnal ilmiah, dan sumber lain yang relevan. Diskusi dengan

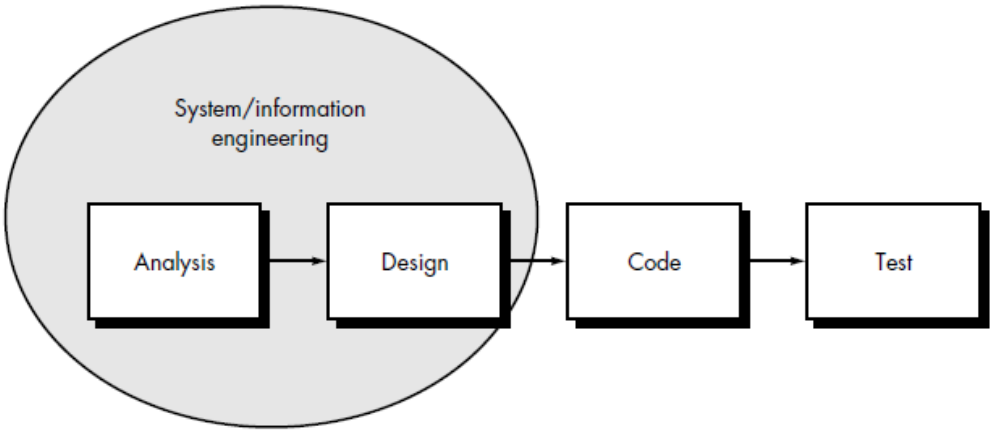
orang-orang yang berkompeten di bidang-bidang yang terkait juga akan dilakukan, seperti dengan pembimbing, rekan mahasiswa, ataupun orang lain yang berkompetensi di bidang *speech recognition*.

1.5.3.2 Pengambilan Dataset

Sebagai bahan untuk data training dan data uji maka akan dilakukan proses perekaman suara dari 10 orang *native speaker* (10 orang laki-laki) dimana setiap orang akan mengucapkan 10 kata yang diulang sebanyak 20 kali. Suara akan direkam dalam format .wav dengan panjang setiap rekaman 1 detik, pengambilan suara dilakukan menggunakan *microphone* pada *channel* mono dengan *sample rate* 16Khz.

1.5.4 Pembangunan Perangkat Lunak

Metode pembangunan perangkat lunak yang akan digunakan pada penelitian ini adalah *Sequential Linear Model*. Metode ini diterapkan karena kebutuhan perangkat lunak yang akan dibangun pada penelitian ini adalah untuk mensimulasikan sistem pengenalan ucapan menggunakan *Elman Recurrent Neural Network*. Adapun tahapan pada *Sequential Linear Model* dapat dilihat pada Gambar 1.2.



Gambar 1.2 *Sequential Linear Model*

Berikut tahapan pembangunan perangkat lunak yang akan dilakukan pada penelititan ini.

1. *Analysis* (Analisis)

Analisis yang akan dilakukan pada penelitian ini adalah analisis terhadap data masukkan, algoritma yang akan digunakan pada *speech recognition* (Ekstraksi Ciri MFCC dan ERNN), hingga analisis kebutuhan fungsional dan non fungsional pada perangkat lunak yang akan dibangun.

2. *Design* (Desain perangkat lunak)

   Pada tahap ini akan dilakukan perancangan terhadap keseluruhan sistem mulai dari struktur menu, antarmuka dan hal lain yang dibutuhkan dalam perangkat lunak.

3. *Code* (Implementasi)

   Analisis dan desain yang sebelumnya telah dilakukan akan diimplementasikan kedalam perangkat lunak agar dapat menghasilkan perangkat lunak untuk melakukan pengenalan ucapan.

4. *Test* (Pengujian perangkat lunak)

   Perangkat lunak yang telah dibangun akan diuji untuk mengetahui apakah perangkat lunak yang dibangun berjalan sesuai diharapkan.

1.5.5 Pengujian Akurasi

Pada tahap ini akurasi dari perangkat lunak *speech recognition* yang menggunakan *Elman Recurrent Neural Network* (ERNN) akan diuji, selain itu beberapa perubahan parameter pada algoritma yang terkait juga akan dilakukan untuk mengetahui bagaimana pengaruhnya terhadap akurasi yang dihasilkan oleh ERNN.

1.5.6 Penarikan Kesimpulan

Pada tahap ini akan dilakukan penarikan kesimpulan berdasarkan hasil dari pengujian akurasi pada *speech recognition* yang menggunakan *Elman Recurrent Neural Network* sebagai metode klasifikasinya.

1.6 Sistematika Penulisan

BAB 1 PENDAHULUAN

Bab ini menerangkan secara umum latar belakang masalah, rumusan masalah, maksud dan tujuan penelitian, batasan masalah, metodologi penelitian serta sistematika penulisan laporan penelitian ini.

BAB 2 LANDASAN TEORI

Bab ini membahas tentang teori atau tinjauan dari literatur pendukung yang akan digunakan pada penelitian ini. Adapun teori-teori yang akan dibahas pada bab ini meliputi teori proses produksi suara pada manusia, metode MFCC (*dc removal*, *pre emphasis*, *frame blocking*, *windowing*, *discrete fourier transform*, *mel frequency warping*, *discrete cosine transform*, *cepstral liftering*), metode *Elman Recurrent Neural Network* (ERNN), teori mengenai analisis dan perancangan aplikasi yang akan dibangun, dan perangkat pendukung yang digunakan dalam penelitian ini.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini berisi tentang analisis masalah, analisis sistem yang akan dibuat, analisis proses mulai dari analisis data masukkan, Ekstraksi Ciri, hingga analisis klasifikasi *Elman Recurrent Neural Network*. Bab ini juga akan membahas tentang perancangan antarmuka, dan perancangan jaringan semantik dalam perangkat lunak yang akan dibangun untuk mengenali ucapan kedalam bentuk teks.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini menampilkan implementasi dari hasil analisis dan perancangan yang telah dilakukan sebelumnya kedalam bentuk pada aplikasi *speech recognition.* yang telah dianalisis pada bab sebelumnya, dan pengujian terhadap fungsional yang ada pada *speech recognition* ini.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran dari hasil penelitian mengenai performa *speech recognition* menggunakan metode *Elman Recurrent Neural Network* yang telah dilakukan agar dapat digunakan dimasa mendatang sebagai referensi dalam penelitian mengenai *speech recognition*.