

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Teknologi saat ini sudah semakin maju seiring perkembangan zaman. Khususnya di bidang interaksi manusia dengan komputer. Manusia terus berusaha melakukan penelitian supaya komputer dapat merepresentasikan apa yang manusia inginkan. Misalkan dalam pengenalan emosi. Komputer dapat melakukan pengenalan emosi salah satunya berdasarkan suara. Suara manusia memiliki banyak informasi penting yang dapat diidentifikasi dan dikenali berupa ciri. Setiap orang memiliki ciri suara yang berbeda maka dari itu perlu dilakukan ekstraksi ciri untuk mendapat hasil ciri suara tersebut yang nantinya akan dilakukan klasifikasi untuk menentukan emosi apa yang sedang dirasakan oleh orang tersebut.

Sudah banyak penelitian yang berfokus pada pengenalan emosi berdasarkan suara dengan beragam metode ekstraksi ciri dan metode klasifikasi. Penelitian-penelitian tersebut menggunakan dataset ucapan yang tersedia sebagai standar seberapa baik sebuah sistem melakukan pengenalan emosi. Banyak *database* emosi yang disediakan oleh beberapa lembaga diluar negeri seperti SUSAS (*Speech Under Simulated and Actual Stress*), *Berlin database of emotional speech* (Emo-DB), *Surrey Audio-Visual Expressed Emotion (SAVEE) Database*, *The Ryerson Audio-Visual Database of Emotional Speech and Song* (RAVDESS) dan masih banyak lagi. Adapun beberapa penelitian yang dilakukan untuk mengenali emosi diantaranya oleh A. Iqbal & K. Barua pada tahun 2019 dengan judul *A Real-time Emotion Recognition from Speech using Gradient Boosting* dengan menggunakan dataset RAVDESS dan metode klasifikasi SVM, KNN dan *Gradient Boosting*. Hasil akurasi untuk data suara pria dengan metode SVM sebesar 65%, metode KNN sebesar 56% dan metode *Gradient Boosting* sebesar 70%. Lalu hasil akurasi untuk data suara wanita dengan metode SVM sebesar 45%, metode KNN sebesar 55% dan metode *Gradient Boosting* sebesar 62%. Kemudian penelitian serupa lainnya dilakukan oleh D. Issa et.al pada tahun 2020 dengan judul *Speech Emotion*

*Recognition With Deep Convolutional Neural Networks* dengan menggunakan dataset yang sama yaitu RAVDESS dan mencapai akurasi sebesar 71,61%.

Adapun metode ekstraksi ciri yang paling sering digunakan dalam pengolahan suara adalah *Mel Frequency Cepstral Coefficient* (MFCC) seperti penelitian yang telah dilakukan oleh R. Umar et.al pada tahun 2019 dengan judul Analisis Bentuk Pola Suara Menggunakan Ekstraksi Ciri *Mel-Frequency Cepstral Coefficients* (MFCC) dengan menggunakan data suara dari beberapa pengguna dengan pengucapan yang sama yaitu “login”. Hasilnya metode MFCC dapat melakukan ekstraksi ciri dengan baik sehingga memberikan hasil bentuk pola suara yang berbeda dari setiap penggunanya.

Lalu metode klasifikasi yang memiliki tingkat akurasi tinggi adalah Multi Layer Perceptron (MLP). Metode ini merupakan turunan dari Jaringan Syaraf Tiruan (JST). Akurasi tinggi dari MLP dalam proses klasifikasi dibuktikan dari penelitian yang telah dilakukan oleh N. Husain & N. Aji pada tahun 2019 dengan judul Klasifikasi Sinyal EEG Dengan *Power Spectra Density* Berbasis Metode *Welch* Dan MLP *Backpropagation* dengan menggunakan 5 kelas dataset (set A, set B, set C, set D dan set E) dari database yang dibuat oleh Dr. Ralph Andrzejak dari Pusat Epilepsi di Universitas Bonn, Jerman. Hasilnya MLP dapat mengklasifikasi sinyal EEG dengan nilai akurasi yang tinggi yaitu 99,68 %.

Dengan latarbelakangi beberapa referensi tersebut, penulis ingin berkontribusi dengan melakukan penelitian tentang pengenalan emosi dengan judul Model Klasifikasi Emosi Berdasarkan Suara Dengan Metode Multilayer Perceptron. Dataset yang digunakan merupakan data sekunder karena penelitian ini dilakukan pada masa pandemi Covid-19. Adapun dataset bersumber dari database *The Ryerson Audio-Visual Database of Emotional Speech and Song* (RAVDESS). Lalu metode ekstraksi ciri yang digunakan adalah *Mel Frequency Cepstral Coefficient* (MFCC) karena metode ini dapat merepresentasikan sinyal dengan baik sehingga dapat memberikan hasil bentuk pola suara yang berbeda dari setiap pengguna. Lalu Multi Layer Perceptron (MLP) dipilih untuk metode klasifikasi dalam penelitian ini karena mampu melakukan klasifikasi pada dataset multiclass dengan baik. Adapun

algoritma pembelajaran yang digunakan adalah *Backpropagation* untuk mengupdate nilai bobot.

## 1.2 Rumusan Masalah

Dalam penelitian ini, permasalahan yang dibahas adalah:

1. Bagaimana cara menghasilkan ciri dari suara manusia dengan menggunakan metode MFCC?
2. Berapa tingkat akurasi dari hasil klasifikasi emosi berdasarkan suara dengan menggunakan MLP *Backpropagation*?

## 1.3 Ruang Lingkup

Ruang lingkup pada penelitian ini adalah sebagai berikut:

1. Format file yang digunakan adalah wav.
2. Data yang digunakan merupakan data sekunder yang bersumber dari *Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)* <https://www.kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio> dengan spesifikasi data 16 bit, 48kHz
3. Data suara yang digunakan adalah 12 suara aktor laki-laki dan 12 suara aktor perempuan.
4. Jenis emosi yang digunakan adalah emosi netral, tenang, senang, sedih, marah, takut, jijik dan terkejut.
5. Proses ekstraksi ciri menggunakan metode Mel-Frequency Cepstral Coefficient (MFCC) dan proses klasifikasi menggunakan Multilayer Perceptron (MLP) *Backpropagation*.

#### 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah melakukan pengujian dan evaluasi dengan menggunakan metode MFCC sebagai metode ekstraksi ciri dan Multilayer Perceptron *Backpropagation* sebagai metode klasifikasi emosi pada dataset RAUDESS.

#### 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah dapat mengetahui tingkat akurasi dari hasil pengujian dan evaluasi dengan menggunakan metode MFCC dan Multilayer Perceptron dalam klasifikasi emosi.

#### 1.6 Luaran yang diharapkan

Luaran yang diharapkan dari penelitian ini adalah model yang dapat melakukan pengujian dan evaluasi dengan menggunakan MFCC sebagai metode ekstraksi ciri dan Multilayer Perceptron *Backpropagation* dalam klasifikasi emosi pada dataset RAUDESS.

#### 1.7 Sistematika Penulisan

Untuk mempermudah pemahaman ide dari penelitian, penulis menyusun sistematika penulisan yang terdiri dari tiga bab:

**BAB I: PENDAHULUAN** Bab ini berisi tentang latar belakang, rumusan masalah, ruang lingkup, tujuan penelitian, manfaat penelitian, luaran yang diharapkan dan sistematika penulisan.

**BAB II: TINJAUAN PUSTAKA** Bab ini berisi teori-teori yang digunakan dalam penelitian meliputi sinyal suara, emosi, Mel Frequency Cepstral Coefficient, Multilayer Perceptron, *underfitting overfitting*, *dropout*, regularisasi, fungsi aktivasi, fungsi optimasi, *confusion matrix* serta penelitian terkait.

**BAB III: METODOLOGI PENELITIAN** Bab ini berisi kerangka pikir yang digunakan untuk memecahkan masalah. alat bantu penelitian, tempat penelitian dan jadwal penelitian

**BAB IV: HASIL DAN PEMBAHASAN** Bab ini berisi persiapan data, praproses data, ekstraksi ciri, perancangan model, implementasi model dan evaluasi.

**BAB V: PENUTUP** Bab ini berisi kesimpulan dan saran.