

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Evolusi waktu kini tumbuh besar sehingga dengan cepat membentuk pertumbuhan penduduk yang semakin tinggi. Kemacetan, serta antrian telah sering kali terjadi diberbagai tempat seperti jalan raya, pertokoan, atau tempat pada umumnya. Ini pula bisa mengganggu emosional atau mental seseorang. Terburu-buru yang melipatgandakan masalah yang tiba silih berganti dalam berkomunikasi dengan seseorang tentu mempengaruhi emosi juga. Biasanya untuk bisa melihat emosi seseorang berdasarkan ekspresi wajahnya, Tetapi kini bisa mengenali keadaan emosi atau mental seseorang hanya mendengarkan suaranya. Suara manusia ialah contoh sinyal analog mengandung informasi. Bandwidth formant, pitch, energy suara, formant, serta durasi pengucapan nya merupakan karakter suara manusia. Karakteristik akustik ini pula bisa mengenali keadaan emosional seseorang apakah senang, marah, ataupun sedih (Magdalena 2014).

Emosi merupakan perasaan kuat yang diarahkan untuk seseorang atau sesuatu. Tidak hanya itu, emosi bisa dijelaskan menjadi reaksi yang muncul akibat tindakan seseorang ataupun peristiwa tersebut. Tipe emosional bisa diklasifikasikan semacam kecemasan, depresi, jijik, kebosanan, tidak puas, dominasi, frustrasi, ketakutan, kebahagiaan, ketidakpedulian, ironi, kegembiraan, netralitas, panik, kejutan, penghambatan, kelelahan, stres, kebingungan, kesedihan (Fridea 1993).

Emosi sering menyebabkan perubahan sikap seseorang yang mengakibatkan gangguan lingkungan. Mengenali emosi manusia merupakan masalah penting untuk hal yang membutuhkan komunikasi manusia serta teknologi. Contohnya, seorang *customer care* bisa memberikan akses yang tepat ke emosi pelanggan dengan mengenali emosi dalam ucapan pelanggan. Sistem informasi bisa memberikan informasi sesuai dengan suasana hati yang menggunakan sistem serta metode pencarian informasi pengguna, dan sebagainya. Pada sisi lain, mengenali emosi ialah permasalahan kompleks. Itu berasal dari adat serta dialek yang berbeda di mana-mana serta sulit untuk menilai secara subjektif sebab emosi adalah masalah

yang terjadi dibawah tingkat kesadaran manusia.

Mengenali emosi berdasarkan suara bertujuan untuk mengidentifikasi keadaan emosional melalui suaranya secara otomatis. Misalnya pada bidang psikolog adalah psikolog tidak bisa bertatap muka langsung dengan pasien maka dapat melakukan panggilan suara dan dengan aplikasi ini psikolog tersebut akan tau keadaan emosi pasiennya seperti apa. Sistem pengenalan suara terdiri dari empat topik utama yaitu emosi masukan, ekstraksi ciri, klasifikasi emosi dan emosi keluaran. Banyak peneliti telah mengusulkan ciri suara penting yang mengandung informasi mengenai emosi, seperti energi, frekuensi, frekuensi pitch, frekuensi formant, dan beberapa ciri spektrum seperti *Linear Predictive Coefficients* (LPC), *Mel-Frequency Cepstrum Coefficients* (MFCC), *Linear Predictive Cepstrum Coefficients* (LPCC). Selain itu, banyak peneliti telah mengeksplorasi beberapa metode klasifikasi, seperti Jaringan Syaraf, *Gaussian Mixture Model* (MLC), *Hidden Markov Model*, *K-Nearest Neighbor* (KNN), dan *Support Vector Machines* (SVM) (Pan 2012).

Pengenalan emosi menggunakan ucapan tidaklah merupakan suatu yang baru pada dunia teknologi. Beberapa peneliti sudah melakukan penelitian dengan menggunakan sistem cerdas dan metode tradisional. Penelitian-penelitian ini memakai data ucapan yang tersedia sebagai standar seberapa baik kinerja sistem pada pendeteksian emosi. Terdapat beberapa institusi luar negeri misalnya 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) menyediakan beberapa database emosi. Kedepannya diharapkan bisa melakukan segala hal dengan mendeteksi emosi manusia menggunakan perangkat atau benda mati (Ariska 2019).

Dalam penelitian lainnya, emosi sedih yang identik menggunakan kecepatan bicara yang lambat pada pitch rendah, sedangkan emosi marah menggunakan kecepatan bicara yang identik serta pitch yang tinggi. Peneliti terdahulu telah melakukan berbagai penelitian mengenai sinyal suara, diantaranya adalah "Pengenalan Emosi Berdasarkan Suara Menggunakan Algoritma HMM (Barlian Henryanu Prasetio, Wijaya Kurniawan, Mochammad Hannats Hanafi Ichsan 2017)". Penelitian ini menghasilkan tingkat keberhasilan sebesar 86.66% pada pengenalan suara emosi menggunakan ekstraksi ciri (*pitch*), energi, dan (*formant*)

dengan klasifikasi emosi dengan HMM, data latih yang dipakai dalam penelitian ini ialah data suara emosi marah, bahagia, dan netral. Adapun penelitian lainnya ialah pengenalan emosi melalui ucapan berdasarkan objek pengujian Emo-DB menggunakan fitur energi, ZCC, entropi, MFCC (Mel-Frequency Cepstral Coefficients), serta pitch menggunakan algoritma K-NN (*K-Nearest Neighbor*) mencapai akurasi 86.02%. Lalu penelitian serupa pula dilaksanakan buat mengembangkan sistem pendeteksian emosi dengan berbagai metode serta mencapai akurasi tertinggi sebesar 92% dengan GMM (*Gaussian Mixture Model*), serta 72% dengan K-NN pada emosional marah. Selain emosi marah, akurasi dihasilkan dibawah 80%. Dalam penelitian tersebut deteksi emosional dilakukan dengan algoritma HMM (*Hidden Markov Model*) menggunakan MFCC. Akurasi mean dari sistem ini adalah 86.66%. Banyak penelitian telah menunjukkan perlunya pengenalan emosi pada interaksi komputer dan manusia. Hasil penelitian lainnya memakai metode MFCC dan DTW-KNN antara lain yaitu pengukuran nilai keberhasilan dari metode MFCC dan DTW-KNN untuk pengenalan aksen etnik di Indonesia memberikan hasil paling efektif dengan taraf akurasi tinggi.

Untuk melakukan pendeteksian emosi manusia melalui ucapan maka diperlukan metode klasifikasi pada data latih yang dimiliki supaya bisa mengenali bahwa sinyal yang masuk diklasifikasikan sebagai emosi yang sudah ditentukan. Di bidang pengolahan suara, metode ekstraksi populer ialah MFCC dalam berbagai macam jenis metode klasifikasi yang di pergunakan dengan merekognisi sampel suara yang di pergunakan.

Penelitian ini dilakukan untuk dapat mendeteksi emosi berbasis ucapan menggunakan ekstraksi fitur ucapan manusia guna memperoleh nilai akurasi kecocokan diatas 90%. Ekstraksi ciri yang di pergunakan dalam penelitian ini ialah *Mel Frequency Cepstrum Coefficient* (MFCC) dan menggunakan metode klasifikasi *Dynamic Time Warping* dan *K Nearest Neighbour* sebab ialah metode ekstraksi fitur yang mendekati sistem pendengaran manusia. diharapkan dengan adanya pendeteksian emosi tersebut, sistem bisa mengidentifikasi emosi manusia berdasarkan ucapan.

Berdasarkan uraian latar belakang masalah tersebut, peneliti tertarik dalam meneliti mengenai **“Deteksi Emosi Manusia Berdasarkan Rekaman Suara menggunakan Python dengan Metode MFCC dan DTW-KNN”**.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka identifikasi masalah adalah Penelitian yang membahas tentang pengenalan emosi umumnya hanya menggunakan salah satu ciri, seperti melalui ekspresi wajah. Selain itu, penelitian pengenalan emosi yang dilakukan dengan menggunakan ekspresi wajah lebih banyak dibandingkan suara. Sehingga, penelitian ini akan mengembangkan sistem pengenalan emosi menggunakan suara secara real time dengan menggunakan metode MFCC dan DTW-KNN.

1.3 Ruang Lingkup

Ruang lingkup dari penelitian ini adalah menggunakan rekaman suara yang dapat mewakili beberapa emosi. Data suara ini diambil dari *Berlin Database of Emotional Speech*. Menggunakan sumber suara berupa rekaman yang telah ada dan memiliki ukuran yang hampir sama dengan durasi waktu 1-3 detik.

1.4 Rumusan Masalah

Berdasarkan latar belakang yang sudah dikemukakan, maka rumusan masalah pada penelitian ini sebagai berikut:

1. Bagaimana hasil ekstraksi ciri *mel frequency cepstrum coefficient*?
2. Bagaimana hasil deteksi emosi berdasarkan rekaman suara manusia dengan metode klasifikasi gabungan DTW-KNN?

1.5 Batasan Masalah

Agar penelitian lebih terarah serta topik pembahasan tidaklah meluas, maka peneliti menetapkan batasan masalah sebagai berikut:

1. Suara yang diperkenalkan yaitu suara dalam keadaan normal atau tidak dalam keadaan sakit.
2. Jenis emosi yang digunakan adalah netral, sedih, marah dan senang.
3. Metode pengenalan emosi berdasarkan rekaman suara manusia yang digunakan adalah metode MFCC dan DTW-KNN.
4. Data yang digunakan diambil dari *website database* yang berisi rekaman suara aktor yaitu *Berlin Database of Emotional Speech*

1.6 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka tujuan dari penelitian ini ialah sebagai berikut:

1. Menentukan hasil ekstraksi ciri *mel frequency cepstrum coefficient*
2. Mendapatkan hasil deteksi emosi berdasarkan rekaman suara manusia dengan metode klasifikasi gabungan MFCC dan DTW-KNN

1.7 Manfaat Penelitian

Melalui penelitian ini maka diharapkan dapat memberikan manfaat sebagai berikut:

1. Bagi peneliti, merupakan media pembelajaran dalam merancang aplikasi analisis deteksi emosi manusia dengan matlab menggunakan metode MFCC dan DTW-KNN dan memberikan sumbangan pemikiran berdasarkan disiplin ilmu yang diperoleh dibangku kuliah.
2. Bagi pembaca, memberikan informasi dan referensi tentang merancang aplikasi dalam analisis deteksi emosi manusia dengan matlab menggunakan metode MFCC dan DTW-KNN.

1.8 Definisi Operasional

Dalam peneliian ini variabel yang diteliti yaitu stabilitas emosi. Istilah stabilitas emosi menurut operasional adalah reaksi yang stabil terhadap seseorang atau kejadian. Pengertian stanilitas emosi adalah kemudahan seseorang dalam merespon emosinya terhadap situasi yang diberikan dan sesuai dengan besar kecilnya rangsangan situasi yang menumbuhkan reaksi emosional.