

BAB I

PENDAHULUAN

1.1 Latar Belakang

Emosi merupakan bagian tak terpisahkan dari kehidupan manusia, terutama dalam percakapan atau komunikasi, dimana emosi memiliki peran penting dalam memperkuat informasi yang disampaikan, melebihi kekuatan kata-kata saja. pengenalan emosi manusia dapat dilakukan melalui identifikasi berbagai sumber seperti teks, ucapan, ekspresi wajah, dan gerakan tubuh [1].

Gelombang atau sinyal yang dihasilkan oleh suara manusia dapat dianalisis untuk mendeteksi emosi yang dirasakan. Deteksi emosi suara manusia adalah salah satu topik penelitian yang sedang berkembang pesat dan banyak diminati, yang dikenal sebagai *Speech Emotion Recognition* (SER) [2]. SER dapat diterapkan di berbagai bidang, salah satunya adalah telekomunikasi, di mana teknologi ini membantu meningkatkan kualitas layanan untuk *customer service* [3]. Dalam hal ini, kepuasan pelanggan menjadi perhatian utama, terutama melalui emosi yang tersampaikan saat percakapan, karena kepuasan pelanggan sangat penting untuk keberhasilan perusahaan atau organisasi. Oleh karena itu, penting untuk memiliki sistem yang dapat secara otomatis mendeteksi emosi melalui suara manusia..

Klasifikasi dapat digunakan sebagai alternatif untuk mendeteksi emosi. Langkah pertama dalam proses ini adalah ekstraksi fitur. Ekstraksi fitur adalah proses pengambilan atribut yang relevan dari data audio suara manusia, yang kemudian dapat digunakan untuk tindakan seperti klasifikasi, pengenalan, dan deteksi emosi.. Metode ekstraksi fitur yang sering digunakan adalah MFCC (*Mel-Frequency Cepstrum Coefficient*). Metode MFCC ini sangat sensitif terhadap kebisingan, sehingga menghasilkan akurasi tinggi di lingkungan yang tidak berisik, tetapi mengalami penurunan akurasi yang signifikan di tempat yang bising. [4]. Metode lain yang dapat digunakan adalah *Root Mean Square Energy* (RMSE) atau amplitudo rata-rata, *Spectral Centroid* atau karakteristik frekuensi dominan, *Spectral Rolloff* atau frekuensi di mana *spektrum* mencapai titik yang ditentukan, *Zero Crossing Rate* atau tingkat perubahan sinyal dari positif ke negatif.

Sebuah penelitian menunjukkan deteksi emosi wicara menggunakan SVM (*Support Vector Machine*) dan LSTM (*Long Short-Term Memory*). Dataset yang digunakan adalah klip suara dari dataset publik HENLO (*Human Voice Natural Language from On-Demand Media*). Dalam dataset ini, terdapat empat kelas emosi dasar: takut, senang, sedih, dan marah. Hasil dari pelatihan pertama menunjukkan akurasi sebesar 30%, sedangkan pada pelatihan kelima akurasi meningkat menjadi 47,4%. Hal ini menunjukkan bahwa kombinasi kedua metode tersebut dapat meningkatkan akurasi sebesar 17,4% [5].

Dalam penelitian lain, dilakukan percobaan menggunakan metode *machine learning* untuk mengklasifikasikan musik berdasarkan genre dengan menggunakan *Random Forest*. Penelitian ini menggunakan dataset GTZAN yang diambil dari *Music Analysis, Retrieval, and Synthesis for Audio Signals* (MARSYAS), terdiri dari file audio berformat wav (*waveform audio file*) dengan 10 lagu. Genre musik yang dianalisis meliputi *disco*, *hiphop*, *jazz*, *pop*, dan *reggae*. Ekstraksi fitur yang digunakan adalah *Mel Frequency Cepstrum Coefficient* (MFCC), karena MFCC mampu mengadaptasi pendengaran manusia dalam bentuk digital. Metode *supervised learning* yang digunakan adalah *Random Forest*, karena metode ini lebih baik dalam mengklasifikasikan data dan *robust* terhadap *outliers* dan *noise*. Nilai performa klasifikasi yang dihasilkan untuk data *training* menunjukkan nilai akurasi dan *F-score* sebesar 99.69% dan 99.7%, sedangkan untuk data *testing* nilai akurasi dan *F-score* masing-masing sebesar 88.83% dan 88.82% [6].

Reduksi fitur dapat digunakan untuk mengoptimalkan fitur-fitur yang akan digunakan dalam proses klasifikasi. Dalam penelitian yang menggunakan metode *Random Forest* untuk deteksi *malware*, digunakan reduksi fitur *Principal Component Analysis* (PCA) yang bertujuan untuk optimasi *random forest* itu sendiri guna meningkatkan performanya. Fitur awal yang berjumlah 1084 direduksi menjadi 32 fitur yang menyebabkan performa *Random Forest* yang semula 98,3% meningkat sebesar 0,14% menjadi 98,44% [7].

Random Forest merupakan metode klasifikasi yang menggabungkan beberapa pohon keputusan untuk menghasilkan prediksi akhir yang lebih akurat dan stabil. Penelitian menunjukkan bahwa untuk mengklasifikasikan data suara manusia yang

terdiri dari 667 data menjadi 4 jenis emosi (senang, sedih, netral, dan marah), tiga teknik ekstraksi fitur digunakan, yaitu MFCC, *Chroma*, dan *Mel Scale*. Hasil yang diperoleh menggunakan *Random Forest* lebih baik daripada menggunakan *Multilayer Perceptron*, *Decision Tree*, maupun *K-Nearest Neighbor* (K-NN), dengan akurasi sebesar 75,5%. Hal ini disebabkan oleh kemampuan *Random Forest* dalam mengatasi *overfitting* yang sering terjadi pada *decision tree*, serta mengatasi masalah *high variance* yang biasa terjadi pada *Multilayer Perceptron* dan K-NN [3].

Berdasarkan hal tersebut, penelitian ini akan membangun sebuah model *Machine Learning Random Forest* dengan menggunakan reduksi fitur *Principal Component Analysis* sebagai sarana optimasi fitur yang digunakan. Model yang diperoleh akan digunakan untuk melakukan deteksi emosi pada suara manusia yang dapat mendeteksi emosi sebanyak 8, yaitu netral, tenang, senang, sedih, marah, takut, jijik, dan kaget.

1.2 Rumusan Masalah

Perumusan masalah yang dihasilkan berdasarkan latar belakang diatas adalah seperti berikut :

1.2.1 Permasalahan

Perkembangan teknologi mempengaruhi kebutuhan akan *customer service* yang dituntut untuk memberikan pelayanan terbaik. Kenyamanan yang diberikan oleh *customer service* dapat mempertahankan maupun menambah pelanggan agar tetap menggunakan produk perusahaan atau organisasi. Deteksi emosi dapat digunakan untuk memahami kondisi pelanggan saat percakapan berlangsung, sehingga perusahaan atau organisasi dapat melakukan analisis berdasarkan respon emosi yang dirasakan oleh pelanggan, guna meningkatkan tingkat kepuasan mereka.

Penelitian ini bertujuan untuk untuk menangani masalah klasifikasi emosi berdasarkan suara manusia menggunakan metode *Random Forest* pada dataset RAVDESS. Penelitian ini berfokus untuk mengidentifikasi model terbaik yang mampu memberikan akurasi tertinggi dalam klasifikasi emosi, yang dapat diaplikasikan dalam meningkatkan analisis pelayanan *customer service* suatu perusahaan atau organisasi.

1.2.2 Metode Usulan

Metode penelitian yang diusulkan dalam penelitian ini dimulai dari 5 metode *feature extraction* yaitu *Mel-Frequency Cepstrum Coefficient* (MFCC), *Zero Crossing Rate* (ZCR), *Root Mean Square Energy* (RMSE), *Spectral Centroid*, *Spectral Rolloff*. Selanjutnya dilakukan reduksi fitur menggunakan *Principal Component Analysis* (PCA). Kemudian dilanjutkan dengan klasifikasi menggunakan *Radom Forest*. Data yang digunakan adalah RAVDESS (*Ryerson Audio-Visual Database of Emotional Speech and Song*) *dateset* yang diambil dari *Kaggle* dengan jumlah 1440 data.

1.2.3 Pertanyaan Penelitian

Berapa nilai akurasi terbaik yang dihasilkan dari model *Random Forest* dalam melakukan klasifikasi untuk deteksi emosi data suara manusia dengan memakai *feature extraction* MFCC, ZCR, RMSE, *Spectral Centroid*, *Spectral Rolloff* dan reduksi fitur *Principal Component Analysis* pada RAVDESS *dataset*?

1.3 Tujuan dan Manfaat Penelitian

Tujuan dan manfaat yang dihasilkan dari penelitian tugas akhir ini adalah sebagai berikut :

1.3.1 Tujuan Penelitian

Mengetahui nilai akurasi terbaik yang dihasilkan oleh model klasifikasi menggunakan metode *Random Fores* dan reduksi fitur menggunakan *Principal Component Analysis* dalam mendeteksi emosi manusia pada RAVDESS *dataset*.

1.3.2 Manfaat Penelitian

1. Penelitian ini memiliki nilai sebagai referensi bagi penelitian lanjutan atau bagi pihak yang tertarik untuk mengeksplorasi deteksi emosi pada suara manusia menggunakan metode klasifikasi *Random Forest*. Selain itu, penelitian ini juga berperan sebagai literatur yang dapat memperkaya pengetahuan di bidang ini.
2. Membantu Perusahaan/Organisasi dalam membuat kebijakan guna meningkatkan kualitas pelayanan pada *customer service*.

1.4 Batasan Masalah

Ada batasan-batasan yang diterapkan pada penelitian tugas akhir ini, di antaranya yaitu :

1. Data audio yang digunakan berasal dari *dataset* RAVDESS yang tersedia di *Kaggle*, dengan total 1440 data. *Dataset* ini terdiri dari 60 rekaman yang diperankan oleh 24 aktor profesional (12 laki-laki dan 12 perempuan) yang menyuarakan pernyataan dalam aksen netral Amerika Serikat.
2. Total emosi yang akan diklasifikasi terdiri dari 8 emosi, yaitu netral, tenang, senang, sedih, marah, takut, jijik, dan kaget.
3. Jumlah Fitur yang digunakan sebanyak 24, 34, 44, dan 54 fitur berasal dari hasil ekstraksi fitur MFCC (20, 30, 40, dan 50 fitur), ZCR (1 fitur), RMSE (1 fitur), *Spectral Centroid* (1 fitur), dan *Spectral Rolloff* (1 fitur).

1.5 Sistematika Penulisan

Untuk sistematika penulisan skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang, rumusan masalah, solusi yang diusulkan untuk menyelesaikan masalah, tujuan dan manfaat penerapan solusi yang diusulkan, batasan masalah dalam penelitian, serta sistematisasi penulisan.

BAB II KAJIAN PUSTAKA

Bab ini menjelaskan mengenai landasan teori dalam penyusunan penelitian ini. Literatur yang dibahas relevan dengan topik deteksi emosi, metode ekstraksi fitur, reduksi dimensi *Principal Component Analysis*, model klasifikasi *Random Forest* beserta penelitian - penelitian terkait.

BAB III METODE USULAN

Bab ini menjelaskan mengenai metode penelitian yang digunakan, yaitu mengenai analisa rancangan sistem, dataset yang digunakan dan alur penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan mengenai hasil dan pembahasan kinerja model deteksi emosi menggunakan klasifikasi Random Forest. Dimulai dari lingkungan uji coba, analisis perbandingan hasil uji coba dari berbagai skenario percobaan, serta mengimplementasikan model yang diperoleh kedalam sistem dengan membuat antarmuka sistem.

BAB V PENUTUP

Bab ini menjelaskan mengenai rangkuman temuan utama dalam penelitian, jawaban terhadap permasalahan penelitian beserta saran untuk pengembangan penelitian.