

BAB I PENDAHULUAN

1.1 Latar Belakang

Pada saat ini kita berada di era di mana banyak informasi dihasilkan dengan sangat cepat. Hal ini tentunya tidak lepas dari era revolusi industri 4.0, era ini sebagian besar aktivitas manusia dilakukan secara *online* sehingga akan selalu ada data baru yang dicatat dan diolah oleh *database*. Salah satu cara dalam penggunaan data adalah dengan menerapkan teknik pembelajaran mesin *machine learning*. *Machine learning* merupakan salah satu teknologi baru yang menjadi bagian dari *Artificial Intelligence* atau kecerdasan buatan pada program komputer yang dirancang untuk meniru kecerdasan manusia baik dalam pengambilan keputusan, logika, dan karakteristik kecerdasan lainnya. Salah satu contoh penerapan *machine learning* adalah dengan mengklasifikasi dan mendeteksi emosi melalui suara manusia. Klasifikasi tersebut melakukan penyaringan emosi manusia dengan menggunakan media suara. Emosi mempunyai peran penting dalam menentukan kepuasan pengguna dan opini pelanggan serta merupakan bagian penting dari interaksi manusia [1]. Dalam penelitian pemrosesan sinyal audio digital saat ini, sistem pengenalan suara dan emosi *speech emotion recognition* (SER) yang cerdas merupakan persyaratan dasar. Banyak aplikasi yang melibatkan interaksi manusia-komputer atau *human-computer interaction* (HCI) sangat bergantung pada SER. Saat ini penting untuk meningkatkan kinerja prediksi sistem SER yang terancang sehingga dapat digunakan untuk aplikasi komersial secara *real-time* [1].

Dalam penelitian *speech emotion recognition* (SER) terdapat berbagai metode dan *framework* yang telah digunakan untuk mengembangkan sistem yang dapat mengenali emosi dalam ucapan. Ada beberapa metode dan *framework* yang terdapat dalam *machine learning* terutama dalam mengimplementasikan pengenalan emosi suara, salah satunya yaitu *Convolutional Neural Network* (CNN). *Convolutional Neural Network* (CNN) merupakan jenis jaringan saraf khusus yang dirancang untuk memproses data dengan struktur mirip grid. Nama "jaringan saraf konvolusional" menandakan penggunaan operasi

matematika yang disebut konvolusi. Konvolusi itu sendiri adalah operasi linier, sehingga jaringan konvolusional memanfaatkan konvolusi secara minimal dalam setiap lapisannya. [2].

Dalam penelitian ini, peneliti melakukan pendekatan terhadap *Speech Emotion Recognition* dengan menerapkan Normalisasi *Z-score* terhadap *Convolutional Neural Network* (CNN). Dengan menerapkan Normalisasi *Z-score* pada penelitian ini dapat memastikan bahwa model pembelajaran mesin dapat bekerja secara optimal dan menghasilkan hasil yang konsisten.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dijelaskan maka rumusan masalah dalam penelitian ini adalah bagaimana penggunaan CNN dalam *Speech Emotion Recognition* dapat ditingkatkan dengan menerapkan algoritma normalisasi *Z-score* menggunakan kode '*StandardScaler*' pada data audio?

Bagaimana pengaruh normalisasi *Z-score* terhadap kinerja model *Speech Emotion Recognition* (SER) menggunakan CNN?

1.3 Batasan Masalah

Pada penelitian ini memiliki beberapa batasan dari masalah di atas. Batasan masalah pada penelitian ini sebagai berikut:

- Data suara yang digunakan adalah CREMA-D *dataset* (publik)
- Terdapat delapan emosi yang dapat dikenali oleh model, yaitu netral, bahagia, sedih, marah, takut, dan jijik.
- Pada satu file suara terdapat satu emosi dengan satu kata
- Normalisasi yang digunakan adalah normalisasi *Z-score*.
- Metode yang digunakan adalah Convolutional Neural Network (CNN).

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengembangkan sebuah sistem *Speech Emotion Recognition* menggunakan metode Machine Learning, terutama CNN, dengan penerapan Normalisasi *Z-score*. Normalisasi ini diintegrasikan untuk

memastikan bahwa data audio yang digunakan oleh model memiliki skala yang seragam, memungkinkan CNN untuk belajar pola-pola yang kompleks dengan lebih baik.

1.5 Manfaat Penelitian

Manfaat penelitian dijabarkan menjadi 2 yaitu, bagi penulis: Mendapatkan gambaran umum tentang bidang deteksi emosi berbasis suara pada manusia. Penulis dapat mempelajari teori, metode dan langkah-langkah penyelesaian deteksi emosi melalui suara pada manusia. Penelitian ini melatih kemampuan analisis penulis. Dan bagi peneliti lain: Penelitian ini dapat dijadikan referensi untuk penelitian lebih lanjut dan lebih kompleks. Penelitian ini dapat dijadikan referensi dalam bidang ilmu komputer.

1.6 Sistematika Penulisan

Dalam penyusunan tugas akhir ini, sistematika penulisan dibagi menjadi beberapa bab sebagai berikut:

BAB I PENDAHULUAN, Berisi latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA, Berisi tinjauan pustaka, dasar-dasar teori yang digunakan dalam bab ini menggunakan jurnal, buku, dan artikel sebagai referensi dengan topik mulai dari masalah yang berkaitan dengan deteksi suara, emosi manusia, pembelajaran mesin, jaringan saraf konvolusional, dan beberapa alat yang digunakan.

BAB III METODE PENELITIAN, Didalamnya terdapat berbagai tahapan dan kebutuhan penelitian, sebagai acuan pemecahan masalah dalam penelitian ini. Bab ini meliputi pengumpulan data dan perancangan sistem menggunakan diagram alir.

BAB IV HASIL DAN PEMBAHASAN, Pada bab ini berisi hasil dan pembahasan setiap proses dalam sistem, pengujian kinerja sistem, kekuatan sistem, dan kelemahan. Pengujian kinerja sistem menggunakan uji pada data tes.

BAB V PENUTUP, Berisi kesimpulan dan saran. Kesimpulan tersebut memenuhi

tujuan penelitian atau tidak memenuhi rekomendasi penelitian, sehingga peneliti lain dapat melanjutkan penelitian melalui keterbatasan dan kekurangan penelitian.

