

BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Perancangan Model Voice Emotion Recognition Penelitian ini menggunakan perancangan model voice emotion recognition berbasis metode Convolutional Neural Network (CNN) yang mampu mengenali emosi utama, seperti marah, senang, sedih, frustrasi, dan netral. Model dirancang menggunakan tahapan preprocessing yang meliputi normalisasi, trimming, dan padding untuk memastikan data audio memiliki kualitas dan durasi yang seragam. Proses pelabelan dilakukan dengan encoding data kategorikal emosi menjadi representasi numerik. Dengan arsitektur CNN yang dirancang, model mampu mengekstraksi fitur penting dari data audio untuk mengidentifikasi emosi secara akurat.
2. Kinerja Model Berdasarkan evaluasi menggunakan metrik *precision*, *recall*, *f1-score*, dan *accuracy*, model CNN yang dirancang mencapai akurasi sebesar 84% pada data uji. Performanya konsisten untuk sebagian besar kategori emosi, dengan *precision*, *recall*, dan *f1-score* rata-rata masing-masing sebesar 84%. Kategori netral menunjukkan hasil terbaik dengan *recall* 95% dan *f1-score* 88%, menunjukkan bahwa model sangat baik dalam mengenali emosi ini. Namun, terdapat beberapa kelemahan pada kategori tertentu, seperti angry dan happy, yang memiliki *recall* lebih rendah, yaitu masing-masing 76% dan 77%. Hal ini menunjukkan bahwa meskipun model telah bekerja dengan baik, masih ada ruang untuk perbaikan, terutama pada kategori dengan pola fitur yang mirip.

Dengan demikian, penelitian ini berhasil menjawab rumusan masalah dengan merancang model *voice emotion recognition* berbasis CNN serta mengukur kinerjanya dalam mengenali emosi utama secara akurat.

5.2 Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan, berikut adalah saran untuk pengembangan lebih lanjut:

1. Peningkatan Dataset

Untuk meningkatkan akurasi dan generalisasi model, disarankan untuk menggunakan dataset yang lebih besar dan lebih beragam, mencakup variasi bahasa, intonasi, dan aksen dari berbagai pengguna.

2. Optimasi Model

Penelitian selanjutnya dapat mencoba arsitektur model yang lebih kompleks atau menggunakan pendekatan hybrid, seperti memadukan CNN dengan *Long Short-Term Memory* (LSTM) untuk menangkap pola temporal dalam data audio.

3. Penanganan Kategori yang Mirip

Untuk mengatasi kesalahan klasifikasi pada kategori emosi yang memiliki pola fitur serupa, seperti *happy* dan *surprised*, disarankan untuk menambahkan fitur tambahan, seperti analisis prosodi atau spektrum suara, yang dapat membantu membedakan emosi tersebut.

4. Pengujian pada Data Nyata

Model perlu diuji pada data suara nyata yang diambil dari lingkungan aplikasi sebenarnya untuk mengevaluasi performa model dalam menghadapi noise dan kondisi rekaman yang bervariasi.

5. Pengintegrasian dengan Sistem Lain

Disarankan untuk mengintegrasikan model ini dengan teknologi *speech recognition* atau *chatbot* untuk menciptakan aplikasi yang lebih interaktif dan mampu memahami konteks emosi pelanggan secara *real-time*.