

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan implementasi sistem penerjemah suara menjadi teks dan bahasa isyarat menggunakan arsitektur CNN–LSTM, dapat disimpulkan beberapa hal sebagai berikut:

1. Sistem yang dibangun berhasil mengintegrasikan proses pra-pemrosesan data, ekstraksi fitur *Mel-Spectrogram*, serta arsitektur CNN–LSTM untuk melakukan klasifikasi audio ke dalam sejumlah label tertentu dengan tingkat akurasi tinggi. Hasil pelatihan menunjukkan akurasi pelatihan sebesar 99,99% dan akurasi validasi sebesar 100%, yang mengindikasikan bahwa model dapat mengenali pola dengan sangat baik.
2. Hasil evaluasi melalui *confusion matrix* dan *classification report* memperlihatkan nilai *precision*, *recall*, dan $f1\text{-score} = 1.00$ (100%) pada seluruh kelas, menandakan bahwa prediksi model sepenuhnya sesuai dengan label sebenarnya pada dataset uji.
3. Implementasi sistem pada *backend* menggunakan Flask terbukti mampu mengelola alur pemrosesan mulai dari penerimaan *input* audio, konversi format, ekstraksi fitur, prediksi model, transkripsi teks, perhitungan *similarity score*, hingga pengiriman hasil ke *frontend*.
4. Antarmuka web (*frontend*) berhasil menampilkan hasil prediksi dalam bentuk teks dan gambar bahasa isyarat dengan desain sederhana dan mudah digunakan, sehingga memudahkan interaksi pengguna.
5. Berdasarkan skenario pengujian dengan tujuh label uji, sistem mampu memberikan prediksi dengan *confidence* tinggi, berkisar antara 90% hingga 100%, serta *similarity score* antara 99.90% hingga 100%. Hal ini menunjukkan konsistensi sistem dalam mengenali kata/frasa yang diucapkan meskipun terdapat sedikit variasi suara pada percobaan tertentu.

Secara keseluruhan, sistem penerjemah suara berbasis CNN–LSTM ini terbukti layak diimplementasikan untuk mendukung komunikasi, khususnya dalam mempertemukan pengguna suara dengan representasi bahasa isyarat.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat disampaikan guna memberikan masukan untuk pengembangan penelitian selanjutnya, yaitu:

1. Disarankan penggunaan dataset yang lebih besar, beragam, dan mencakup kondisi lingkungan nyata (misalnya variasi aksen, intonasi, serta latar belakang kebisingan) agar model CNN–LSTM yang dikembangkan memiliki daya generalisasi yang lebih baik.
2. Penelitian selanjutnya disarankan untuk mengembangkan sistem penerjemah suara–teks–bahasa isyarat ini dalam bentuk aplikasi lintas platform, khususnya berbasis mobile, agar pemanfaatannya dapat lebih luas dan mendukung komunikasi inklusif secara langsung di lapangan.
3. Jika keterbatasan waktu komputasi menjadi pertimbangan utama, penelitian selanjutnya dapat menerapkan metode optimasi *hyperparameter* yang lebih efisien atau menggunakan arsitektur model alternatif seperti *transformer* atau *attention-based models* untuk memperoleh performa optimal dengan waktu komputasi yang lebih singkat.