BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi emosi suara menggunakan metode Mel-Frequency Cepstral Coefficients (MFCC) dan Convolutional Neural Network (CNN), dapat disimpulkan beberapa hal sebagai berikut:

1. Proses preprocessing yang meliputi penghapusan bagian hening (silence removal), normalisasi, dan penyesuaian durasi sinyal berhasil menghasilkan data suara yang lebih konsisten dan siap untuk tahap ekstraksi fitur.
2. Ekstraksi fitur menggunakan MFCC yang dikombinasikan dengan fitur delta dan delta-delta mampu merepresentasikan karakteristik spektral dan dinamika temporal suara secara efektif sebagai input bagi model CNN.
3. Model CNN yang dirancang dalam penelitian ini mampu melakukan klasifikasi empat kategori emosi, yaitu angry, happy, neutral, dan sad, dengan tingkat akurasi sebesar 76% pada data uji.
4. Performa terbaik diperoleh pada emosi angry dan sad, sedangkan emosi neutral memiliki tingkat recall terendah, yang menunjukkan adanya tantangan dalam membedakan emosi dengan karakteristik suara yang relatif datar atau mirip dengan emosi lain.
5. Model yang dikembangkan telah berhasil diimplementasikan dalam bentuk aplikasi berbasis web menggunakan Streamlit, sehingga dapat digunakan untuk melakukan prediksi emosi suara secara langsung.

Secara keseluruhan, penelitian ini menunjukkan bahwa kombinasi MFCC dan CNN merupakan pendekatan yang cukup efektif untuk sistem Speech Emotion Recognition (SER).

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran untuk pengembangan lebih lanjut adalah sebagai berikut:

1. Menggunakan dataset dengan jumlah dan variasi data yang lebih besar untuk meningkatkan kemampuan generalisasi model.
2. Mengkombinasikan MFCC dengan fitur prosodi lain seperti pitch, energy, dan tempo bicara untuk meningkatkan akurasi klasifikasi.
3. Mencoba arsitektur model yang lebih kompleks, seperti kombinasi CNN-LSTM atau model berbasis transfer learning, untuk menangkap dinamika temporal yang lebih mendalam.
4. Melakukan tuning hyperparameter secara lebih sistematis, seperti variasi jumlah layer, ukuran kernel, learning rate, dan epoch.
5. Mengembangkan sistem agar dapat bekerja secara real-time dengan integrasi langsung terhadap input mikrofon.

Dengan pengembangan lebih lanjut, sistem deteksi emosi berbasis suara diharapkan dapat digunakan secara lebih luas dalam berbagai bidang seperti layanan pelanggan, asisten virtual, dan sistem pendukung kesehatan mental.