

## **BAB V**

### **PENUTUP**

#### **5.1. Kesimpulan**

Berdasarkan penelitian yang telah dilakukan oleh penulis guna mengklasifikasikan jenis emosi melalui ucapan dengan menggunakan metode *Convolutional neural network*, dapat disimpulkan beberapa hal sebagai berikut:

1. Implementasi algoritma *Convolutional neural network* dalam mengklasifikasikan emosi manusia melalui suara dilakukan melalui dua tahapan utama yaitu ekstraksi fitur suara serta pembuatan model menggunakan algoritma *machine learning*. Pada penelitian ini tahapan ekstraksi suara dilakukan dengan menggunakan beberapa metode ekstraksi antara lain MFCC (*Mel-Frequency Cepstral Coefficients*), *Energy*, *Pitch*, *Spectral Centroid*, *Spectral Flatness* dan *Spectral Roll-off*. Selanjutnya untuk tahapan klasifikasi dilakukan dengan melakukan pelatihan pada data *audio* percakapan dengan rasio 80% untuk data *training* dan 10% untuk data validasi dari total 1440 data suara yang telah melalui tahapan pra-proses. Pelatihan ini akan menghasilkan model CNN yang selanjutnya akan digunakan untuk pengujian menggunakan 10% dari total data suara sebagai data *testing*. Pengujian model menunjukkan bahwa model dapat mengklasifikasikan jenis emosi dalam suara menjadi beberapa kelas seperti netral, tenang, sedih, senang, takut, marah, jijik dan terkejut dengan tingkat akurasi tertentu.
2. Hasil pengujian yang dilakukan pada semua skenario yang ada didapatkan hasil bahwa skenario uji coba kedua yang menggunakan metode ekstraksi fitur suara *Mel-Frequency Cepstral Coefficients* (MFCC) mendapatkan akurasi tertinggi sebesar 70% diikuti dengan nilai rata-rata presisi dan *recall* masing-masing 68% dan 67%. Untuk kelas emosi yang paling sering ditebak dengan benar adalah emosi marah, terkejut, sedih dan tenang dengan rata-rata prediksi benar sebesar 77%.

## 5.2. Saran

Adapun beberapa saran yang dapat diberikan penulis untuk dapat menjadi acuan dalam pengembangan lebih lanjut adalah sebagai berikut:

1. *Data input* pada penelitian ini masih menggunakan suara dalam bahasa Inggris sehingga perlu dilakukan penelitian selanjutnya dengan menambahkan data suara yang menggunakan bahasa Indonesia.
2. Menambahkan *hyperparameter tuning* yang lebih banyak lagi pada saat penggunaan algoritma *Convolutional Neural Network* seperti *batch size*, *epochs*, *optimizers function*, dan lainnya. Hal ini bertujuan untuk mendapatkan model CNN yang menghasilkan performa terbaik dalam mengklasifikasikan kelas emosi.
3. Dapat menambahkan dataset suara yang digunakan sehingga pelatihan model mendapatkan sampel fitur yang lebih bervariasi atau juga dapat menggunakan metode *upsampling* seperti SMOTE untuk menaikkan jumlah dataset yang digunakan.
4. Dapat menggunakan metode ekstraksi fitur suara yang lebih banyak lagi serta berbeda untuk mengetahui metode yang paling baik untuk mengklasifikasikan sebuah emosi dari ucapan seseorang.
5. Dapat menggunakan algoritma machine learning yang berbeda namun relevan untuk kasus klasifikasi suara seperti *Support Vector Machine* (SVM), *Long Short Term Memory* (LSTM) dan lainnya.