

## BAB 5

### SIMPULAN DAN SARAN

#### 5.1 Simpulan

Perancangan dan evaluasi Perancangan dan Evaluasi Model Deteksi Deephoax Audio Berbasis Analisis Lapisan wav2vec 2.0 dan Klasifikasi Multilayer Perceptron (MLP) telah selesai. Penelitian ini menyajikan investigasi sistematis terhadap representasi fitur dan arsitektur pengklasifikasi untuk deteksi audio deep hoaks, dengan fokus pada pertimbangan implementasi praktis. Analisis *layer-wise ablation* mengungkapkan temuan yang bersifat berlawanan dengan intuisi namun signifikan: lapisan transformer awal (lapisan ke-4) dari *pre-trained model* wav2vec 2.0 menunjukkan performa yang lebih baik dibandingkan lapisan yang lebih dalam, dengan selisih akurasi sebesar 6,4 poin persentase (92,46% dibandingkan 86,04%). Hal ini menunjukkan bahwa artefak akustik tingkat rendah yang dipertahankan pada representasi awal lebih diskriminatif untuk deteksi audio hoaks dibandingkan fitur semantik tingkat tinggi. Temuan ini menantang asumsi konvensional mengenai abstraksi fitur pada model ucapan pra-latih serta memberikan panduan berharga bagi peneliti di bidang ini, sekaligus membuka potensi pengembangan arsitektur sistem deteksi audio deep hoaks yang lebih sederhana namun lebih efisien.

Perbandingan komprehensif terhadap berbagai arsitektur MLP selanjutnya mengidentifikasi keseimbangan optimal antara akurasi dan efisiensi komputasi. Meskipun MLP dengan dua lapisan tersembunyi menunjukkan performa yang kompetitif dengan presisi sempurna, konfigurasi empat lapisan dengan optimasi ambang batas berhasil mencapai akurasi keseluruhan tertinggi sebesar 94,93%, dengan tetap mempertahankan presisi 88,69% dan *recall* 90,81%. Yang terpenting, konfigurasi optimal ini mampu mempertahankan waktu inferensi di bawah satu milidetik, yaitu sekitar 28 mikrodetik per sampel, sehingga memungkinkan pemrosesan *real-time* lebih dari 35.000 sampel per detik pada perangkat CPU standar.

Secara keseluruhan, temuan ini menunjukkan bahwa sistem deteksi audio deephoax yang efektif tidak harus bergantung pada proses *fine-tuning* yang kompleks atau model yang mahal secara komputasi. Sebaliknya, pemilihan strategis fitur beku dari lapisan awal yang dikombinasikan dengan pengklasifikasi berukuran sesuai mampu menghasilkan performa setara dengan pendekatan mutakhir

sekaligus memenuhi kebutuhan untuk penerapan di dunia nyata. Penelitian ini memberikan kontribusi baik dari sisi metodologi pemilihan fitur dalam forensik audio maupun referensi panduan dalam membangun sistem deteksi deepfake yang efisien dan skalabel, khususnya untuk aplikasi keamanan siber yang menuntut akurasi dan kecepatan secara bersamaan.

## 5.2 Saran

Penelitian ini hanya dilakukan dengan satu dataset. Untuk melihat relevansi dari temuan dalam penelitian ini, diperlukan penggunaan *training data* yang lebih beragam. Untuk penelitian selanjutnya, dapat diberikan fokus pada dukungan multibahasa.

