BAB 5
PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan analisis yang telah dilakukan secara menyeluruh, serta mempertimbangkan berbagai aspek teknis dan implementatif dalam pengembangan model klasifikasi sertifikat berbasis NLP menggunakan BERT, dapat disimpulkan bahwa model BERT mampu mengklasifikasikan sertifikat ke dalam tujuh mata kuliah dengan tingkat akurasi yang tinggi. Dari dua konfigurasi yang diuji, BERT-base-uncased menunjukkan performa terbaik, dengan nilai akurasi pelatihan mencapai 98,73% dan akurasi validasi sebesar 95,54% ketika menggunakan skenario augmentasi Synonym Replacement. Sementara itu, konfigurasi BERT-multilingual-uncased juga menunjukkan kinerja yang baik, dengan akurasi validasi tertinggi sebesar 92,99% pada skenario yang sama. Temuan ini menegaskan bahwa pendekatan bidireksional BERT efektif dalam menangkap konteks teks sertifikat, dengan konfigurasi base-uncased memberikan generalisasi yang lebih optimal untuk dokumen berbahasa Indonesia dan Inggris.

Lebih lanjut, penggunaan teknik augmentasi data berpengaruh signifikan terhadap kinerja model. Di antara lima skenario dataset yang diuji, Synonym Replacement terbukti menjadi teknik augmentasi yang paling efektif dalam meningkatkan akurasi model, baik pada konfigurasi base-uncased maupun multilingual-uncased. Teknik ini tidak hanya menghasilkan akurasi pelatihan dan validasi tertinggi, tetapi juga mencatatkan nilai *loss* dan *validation loss* terendah. Sebaliknya, teknik augmentasi berbasis manipulasi karakter seperti *Character Insertion* dan *Character Deletion* cenderung menurunkan stabilitas model, dengan akurasi validasi yang lebih rendah dan kecenderungan terhadap kesalahan klasifikasi.

Analisis lebih lanjut melalui *confusion matrix* dan *classification report* memperkuat temuan tersebut. Penerapan *Synonym Replacement* mampu mengurangi kesalahan klasifikasi lintas kelas, khususnya pada kelas-kelas yang rawan tertukar seperti "Desain Antarmuka" dan "Pemrograman Web". Model dengan konfigurasi BERT-base-uncased yang menggunakan *Synonym Replacement* mencapai *precision*, *recall*, dan *F1-score* masing-masing sebesar (0,97), sementara

BERT-multilingual-uncased mencatatkan *F1-score* tertinggi sebesar 0,94. Sebaliknya, penggunaan augmentasi berbasis karakter memperburuk konsistensi prediksi model, menunjukkan pentingnya mempertahankan makna semantik dalam proses augmentasi data.

Secara keseluruhan, kombinasi model BERT-base-uncased dengan teknik augmentasi *Synonym Replacement* menghasilkan keseimbangan terbaik antara akurasi tinggi, stabilitas generalisasi, dan efisiensi pengolahan data. Selain itu, layanan sederhana berbasis Streamlit yang dikembangkan berhasil mengintegrasikan model ini ke dalam sistem praktis, sehingga berpotensi mendukung implementasi rekognisi pembelajaran lampau dengan cara yang lebih cepat, efisien, dan akurat.

5.2 Saran

Berdasarkan kesimpulan yang telah diuraikan, berikut adalah beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut guna meningkatkan kualitas dan cakupan penelitian ini:

1. Untuk meningkatkan performa model, disarankan untuk memperluas dataset dengan menambah jumlah sampel dari setiap kelas. Dataset yang lebih besar dan lebih bervariasi akan membantu model mempelajari pola yang lebih kompleks, sehingga dapat meningkatkan akurasi dan generalisasi model. Selain itu, pengumpulan data dari sumber yang lebih beragam, seperti sertifikat dari berbagai institusi atau platform pelatihan, juga dapat meningkatkan representasi data.
2. Selain tahapan praproses yang telah dilakukan, peneliti menyarankan untuk melakukan denoising pada hasil ekstraksi OCR guna menghilangkan karakter atau simbol yang tidak relevan. Dengan demikian, teks yang lebih bersih dan terstruktur akan memudahkan dalam tahap analisis dan klasifikasi selanjutnya, sehingga meningkatkan akurasi model yang digunakan.
3. Meskipun teknik *Synonym Replacement (SR)* menunjukkan hasil terbaik dalam penelitian ini, eksplorasi teknik augmentasi lainnya seperti *Random Swap, Random Deletion,* atau kombinasi beberapa teknik augmentasi dapat

dipertimbangkan. Hal ini bertujuan untuk mengevaluasi apakah teknik lain dapat memberikan hasil yang lebih baik atau melengkapi teknik yang sudah ada.

4. Selain model BERT, eksplorasi model NLP lainnya seperti RoBERTa, DistilBERT, atau bahkan model berbasis *transformer* generasi terbaru seperti Qwen atau Deepseek dapat dipertimbangkan. Model-model ini memiliki arsitektur yang lebih terkini dan mungkin dapat memberikan performa yang lebih baik dalam tugas klasifikasi teks. Selain itu, penerapan model multilingual lainnya juga dapat dieksplorasi untuk mendukung pengolahan teks dalam berbagai bahasa, sehingga aplikasi ini dapat digunakan dalam konteks internasional. Peneliti juga menyarankan untuk melakukan tuning hyperparameter untuk mengoptimalkan kinerja model lebih lanjut.

Penelitian ini diharapkan dapat menjadi landasan yang kokoh bagi pengembangan sistem klasifikasi sertifikat yang lebih adaptif, efisien, dan akurat. Ke depannya, hasil penelitian ini juga diharapkan mampu mendukung perluasan aplikasi sistem serupa di berbagai konteks, baik dalam lingkungan pendidikan formal maupun dalam sektor industri berbasis keterampilan.

Halaman ini sengaja dikosongkan