

BAB 5

KESIMPULAN DAN SARAN

Bab ini menyajikan kesimpulan yang ditarik dari hasil implementasi, pengujian, dan analisis sistem otomasi validasi biodata mahasiswa untuk pendaftaran wisuda sarjana. Selain itu, bab ini juga menguraikan sejumlah saran yang dapat dijadikan landasan bagi pengembangan dan penyempurnaan sistem pada penelitian di masa mendatang.

5.1 Kesimpulan

Berdasarkan keseluruhan tahapan penelitian, perancangan, dan pengujian yang telah dilakukan, dapat ditarik beberapa kesimpulan utama sebagai berikut:

1. **Keandalan Ekstraksi Teks Berbasis *Vision-Language*:** Pemanfaatan model komputasi awan (Gemini API) sebagai mesin utama untuk mendeduksi teks dari citra dokumen pindaian ijazah terbukti sangat presisi. Model mampu mencapai tingkat kesalahan yang sangat minim dengan nilai *Character Error Rate* (CER) sebesar 0.042. Kecerdasan multimodal pada komputasi awan ini juga terbukti tangguh dalam mengatasi anomali visual di lapangan, seperti teks tulisan tangan yang memudar akibat degradasi kualitas kertas.
2. **Optimalisasi Klasifikasi Entitas melalui *Fine-Tuning* IndoBERT:** Pelatihan lanjutan (*fine-tuning*) pada model bahasa IndoBERT-Base menggunakan dataset spesifik domain akademik berhasil mendongkrak performa *Named Entity Recognition* (NER) secara signifikan. Kemampuan model dalam mengekstraksi atribut Nama, Tempat Lahir, dan Tanggal Lahir beroperasi secara sangat optimal dengan rata-rata *F1-Score* agregat melesat dari 1% (sebelum *fine-tuning*) menjadi 90% hingga 97.9% pada evaluasi akhir, menjadikannya sangat relevan untuk memproses tata bahasa dokumen akademik Indonesia.
3. **Efektivitas Algoritma *Fuzzy String Matching*:** Penggabungan luaran klasifikasi AI dengan algoritma *Fuzzy String Matching* menggunakan metode *Token Set Ratio* sukses mengatasi masalah variasi penulisan dan

kesalahan ketik minor (seperti kelebihan tanda baca atau urutan kata yang terbalik). Sistem mampu mencapai tingkat kesuksesan validasi otomatis sebesar 97.8% saat dicocokkan dengan pangkalan data SIAMIK, yang secara langsung mereduksi lebih dari 80% beban kerja (*workload*) tinjauan manual oleh staf administrasi akademik.

4. **Keunggulan Arsitektur Sistem Hibrida (Awan dan Luring):** Implementasi arsitektur komputasi hibrida yang dibungkus dalam antarmuka *Streamlit* memberikan solusi infrastruktur yang ideal bagi operasional institusi pendidikan tinggi. Sistem mampu menekan waktu inferensi hingga 3.8 detik per berkas dengan membagi beban kerja secara terdistribusi: ekstraksi citra dieksekusi pada peladen awan, sementara pemahaman semantik (NLP) dieksekusi secara instan dan efisien menggunakan memori CPU lokal.

5.2 Saran

Meskipun sistem yang dikembangkan telah memenuhi tujuan penelitian dan menunjukkan kinerja yang sangat baik, terdapat beberapa ruang untuk penyempurnaan di masa mendatang. Beberapa saran yang direkomendasikan adalah sebagai berikut:

1. **Eksplorasi Model *Optical Character Recognition* (OCR) Alternatif:** Untuk mengurangi ketergantungan pada API komputasi awan pihak ketiga dan latensi internet, penelitian selanjutnya disarankan untuk mengeksplorasi dan menguji arsitektur OCR tingkat lanjut yang dapat berjalan sepenuhnya di peladen lokal (*on-premise*), seperti model **TrOCR** (*Transformer-based Optical Character Recognition*). Transisi ke TrOCR diharapkan dapat meningkatkan keamanan data tertutup sekaligus menjaga tingkat akurasi pada dokumen yang bervariasi.
2. **Ekspansi *Dataset* Anotasi CoNLL:** Performa pengenalan entitas, khususnya pada atribut Tempat Lahir (B-POB) yang masih rentan terhadap penulisan singkatan, dapat ditingkatkan dengan memperkaya korpus *dataset* latih. Peneliti selanjutnya disarankan untuk menambah sampel dokumen

dari berbagai jenjang dan institusi lain, serta memperbanyak variasi penulisan toponimi (nama daerah) agar model bahasa semakin adaptif.

3. **Integrasi Penuh melalui API Pangkalan Data (SIAMIK):** Pada prototipe ini, pencocokan *Fuzzy String* masih disimulasikan menggunakan data referensi lokal. Ke depannya, disarankan agar sistem web yang dibangun dapat diintegrasikan secara langsung dengan pangkalan data mahasiswa (SIAMIK) melalui *Application Programming Interface* (API) internal kampus. Hal ini akan memungkinkan proses validasi pendaftaran wisuda sarjana berjalan secara murni *real-time* dengan basis data yang terus diperbarui.