

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian, implementasi, dan pengujian yang telah dilakukan terhadap *Prototype* sistem web analisis dokumen Word berbasis *Machine learning* dengan *AI Chat Interface* untuk deteksi kesalahan SOP di Dinsos PPKB Purworejo, dapat disimpulkan hal-hal sebagai berikut. Secara keseluruhan, *Prototype* yang dikembangkan berhasil membuktikan konsep bahwa sistem berbasis AI dapat membantu pegawai dalam memvalidasi dokumen secara otomatis. Ketiga rumusan masalah yang diajukan pada awal penelitian telah terjawab melalui pengujian fungsional, evaluasi performa model, dan uji penerimaan pengguna.

Hasil observasi dan wawancara terkait proses pemeriksaan dokumen di Dinsos PPKB Purworejo mengonfirmasi bahwa validasi selama ini masih dilakukan secara manual dengan membuka file satu per satu. Metode ini rentan terhadap perbedaan penilaian antar pegawai dan memakan waktu operasional yang signifikan ketika volume surat tinggi. Elemen yang paling sering tidak lengkap dalam dokumen yang diujikan adalah Kop Surat, terutama pada bagian nama instansi pemerintah, alamat kantor, dan informasi kontak.

Selanjutnya, pengujian terhadap performa model *Machine Learning* dan *Natural Language Processing* (NLP) menunjukkan bahwa sistem berhasil diimplementasikan dengan mengintegrasikan tiga lapisan pemrosesan: ekstraksi fitur berbasis aturan menggunakan *Regular Expression* dan algoritma Jarak *Levenshtein* untuk menghasilkan vektor fitur biner dari delapan elemen SOP, klasifikasi menggunakan model *Support Vector Machine* (SVM) yang terpisah untuk Surat Undangan dan Surat Tugas, serta penyajian hasil melalui *AI Chat Interface* berbasis Google Gemini API. Pada data uji, model SVM mencapai *F1-Score* sebesar 86% untuk Surat Undangan dan 67% untuk Surat Tugas, keduanya melampaui target minimal yang ditetapkan sebesar 65%. Validasi silang *Stratified K-Fold* dengan nilai $K=2$ hingga $K=6$ menunjukkan bahwa model Surat Undangan stabil dengan rata-rata akurasi 0,75–0,92, sementara model Surat Tugas memiliki fluktuasi lebih besar (0,60–0,82) yang disebabkan oleh ukuran *dataset* yang terbatas.

Terakhir, evaluasi terhadap fitur rekomendasi membuktikan bahwa sistem mampu menyajikan arahan perbaikan yang spesifik dan kontekstual melalui dua mekanisme: rekomendasi berbasis aturan statis dari hasil analisis NLP yang selalu tersedia, dan rekomendasi naratif berbahasa Indonesia yang dihasilkan oleh Google Gemini API ketika kuota tersedia. Mekanisme *fallback* memastikan sistem tetap informatif meskipun layanan AI eksternal tidak tersedia. Seluruh delapan skenario pengujian *black-box* berhasil sesuai harapan.

Keseluruhan pengujian tersebut berhasil membuktikan konsep (proof of concept) bahwa sistem web berbasis *Machine learning* dengan *AI Chat Interface* dapat digunakan sebagai alat bantu validasi dokumen SOP secara digital di lingkungan instansi pemerintahan daerah, mendukung transformasi digital administrasi di Dinsos PPKB Purworejo.

5.2 Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan selama proses pengembangan dan pengujian *Prototype*, berikut disampaikan saran untuk pengembangan sistem lebih lanjut maupun untuk penelitian berikutnya. Saran-saran ini disusun berdasarkan prioritas dampak terhadap peningkatan kualitas sistem.

5.2.1 Perluasan dan Penyeimbangan *Dataset*

Prioritas utama untuk pengembangan selanjutnya adalah penambahan ukuran *dataset* secara signifikan. *Dataset* yang digunakan pada penelitian ini berjumlah 50 dokumen (40 latih, 10 uji) yang terbukti masih terlalu kecil untuk melatih model yang general, terutama untuk Surat Tugas yang hanya memiliki 20 data latih dan menghasilkan *Recall* 50%. Penelitian berikutnya disarankan mengumpulkan minimal 200–300 dokumen per jenis surat dengan variasi yang lebih beragam, mencakup dokumen dari berbagai periode, berbagai pegawai pengetik yang berbeda, serta berbagai kombinasi elemen SOP yang kurang. Keseimbangan kelas antara Sesuai SOP dan Tidak Sesuai SOP juga harus dijaga pada rasio minimal 50:50.

Dengan *dataset* yang lebih besar, teknik augmentasi data seperti sinonim kata atau perubahan minor pada format dapat diterapkan untuk memperkaya variasi pola yang dipelajari model.

5.2.2 Penggantian Algoritma Klasifikasi ke Model Berbasis Transformer dan IndoBERT

Penelitian ini menggunakan *Support Vector Machine* (SVM) dengan fitur biner hasil ekstraksi aturan, yang merupakan pendekatan yang sudah terbukti layak sebagai *proof of concept*. Namun, guna meningkatkan kemampuan pemahaman semantik dan kontekstual dokumen secara lebih mendalam, studi lanjutan sangat dianjurkan untuk mengeksplorasi model berbasis transformer, khususnya IndoBERT yang merupakan model BERT yang dilatih khusus menggunakan korpus bahasa Indonesia.

IndoBERT (indobenchmark/indobert base p1) mampu memahami konteks kalimat secara bidireksional sehingga dapat menangkap makna frasa seperti "Plt. Kepala Dinas" atau "Dasar Penugasan" dalam konteks kalimat yang lebih luas, bukan sekadar pencocokan pola. Selain IndoBERT, pengembangan sistem dapat membandingkan performa beberapa algoritma berikut: Random Forest yang mampu menangani fitur biner dengan baik dan memberikan informasi *feature importance* untuk interpretasi model Gradient Boosting (XGBoost atau LightGBM) yang dikenal lebih robust terhadap *overfitting* pada *dataset* kecil serta arsitektur transformer

ringan seperti DistilBERT-multilingual atau XLM RoBERTa yang lebih efisien secara komputasi dibandingkan IndoBERT penuh.

5.2.3 Penerapan Explainable AI (XAI) pada Sistem Analisis

Salah satu kelemahan model *Machine learning* dalam konteks administrasi pemerintahan adalah sulitnya menjelaskan mengapa sebuah dokumen diklasifikasikan sebagai Tidak Sesuai SOP kepada pengguna non teknis. Oleh karena itu, sangat penting untuk mengintegrasikan teknik Explainable AI (XAI) sehingga setiap prediksi model dapat dijelaskan secara transparan kepada pegawai.

Dua pendekatan XAI yang paling relevan untuk kasus ini adalah SHAP (SHapley Additive exPlanations) dan LIME (Local Interpretable Model agnostic Explanations). SHAP dapat menghitung kontribusi setiap fitur SOP (misalnya $Kop_Surat=0$ berkontribusi -0,42 terhadap prediksi) sehingga sistem dapat menampilkan visualisasi waterfall chart yang menunjukkan fitur mana yang paling berpengaruh terhadap keputusan klasifikasi. LIME bekerja dengan membuat model lokal yang dapat diinterpretasikan di sekitar satu prediksi, cocok untuk menjelaskan kasus per kasus kepada pengguna. Integrasi XAI pada *AI Chat Interface* akan memungkinkan *chatbot* tidak hanya menyebut "Kop Surat kurang" tetapi juga menjelaskan "Kop Surat berkontribusi 38% terhadap keputusan Tidak Sesuai SOP karena dua dari empat sub-komponennya tidak ditemukan.

5.2.4 Integrasi OpenCV untuk Deteksi Visual Kop Surat dan Logo Instansi

Pendekatan NLP berbasis teks pada sistem saat ini tidak mampu memvalidasi keberadaan elemen visual pada dokumen, seperti logo instansi (lambang garuda atau logo Kabupaten Purworejo) dan garis pembatas kop surat yang merupakan bagian dari standar tata naskah dinas. Untuk mengatasi celah ini, diperlukan integrasi Computer Vision melalui library OpenCV guna melakukan analisis visual pada halaman pertama dokumen.

Alur implementasi yang disarankan adalah mengonversi halaman pertama dokumen .docx menjadi gambar menggunakan pustaka pdf2image atau python docx2txt, kemudian menggunakan OpenCV untuk mendeteksi keberadaan logo menggunakan template matching atau kontur objek, mendeteksi garis horizontal pembatas kop surat menggunakan Hough Line Transform, serta memverifikasi posisi dan proporsi elemen visual sesuai standar. Kombinasi deteksi teks (NLP) dan deteksi visual (OpenCV) akan menghasilkan validasi SOP yang jauh lebih komprehensif dibandingkan pendekatan teks saja.

5.2.5 Perbaikan Dokumen Otomatis Berbasis *AI Chat Interface*

Keterbatasan utama *Prototype* saat ini adalah sistem hanya dapat mendeteksi dan merekomendasikan perbaikan, tetapi pengguna masih harus memperbaiki dokumen secara manual di Microsoft Word. Sebagai bentuk penyempurnaan, sistem perlu dibekali fitur perbaikan dokumen otomatis yang sepenuhnya terintegrasi dalam *AI Chat Interface*, sehingga pengguna cukup mengetik perintah seperti "tolong perbaiki dokumen ini" atau "perbaiki kop surat" di kolom chat, dan sistem akan secara otomatis menghasilkan dokumen yang sudah diperbaiki untuk diunduh.

Alur kerja fitur ini adalah sebagai berikut. Pengguna mengunggah dokumen dan sistem menampilkan hasil analisis beserta rekomendasi perbaikan seperti yang sudah berjalan saat ini. Ketika pengguna mengetik perintah perbaikan pada kolom chat, sistem akan mengirimkan prompt terstruktur kepada Google Gemini API yang berisi teks asli dokumen, daftar elemen yang kurang, dan instruksi perbaikan sesuai format SOP Dinsos PPKB. Gemini API kemudian menghasilkan teks dokumen yang sudah diperbaiki dalam format terstruktur, misalnya dalam bentuk JSON dengan key berupa nama elemen dan value berupa teks yang disarankan. Sistem Flask kemudian menggunakan pustaka python docx untuk menyuntikkan teks perbaikan ke dalam dokumen .docx asli, mempertahankan format dan layout yang ada, lalu menyediakannya untuk diunduh oleh pengguna.

5.2.6 Implementasi Berbasis Data Permanen dan Manajemen Pengguna

Prototype saat ini tidak memiliki basis data permanen sehingga seluruh riwayat analisis dan percakapan hilang setelah sesi berakhir. Tahap penyempurnaan berikutnya wajib mengimplementasikan basis data relasional . Penggunaan SQLAlchemy sebagai ORM (Object Relational Mapper) akan memudahkan migrasi antara SQLite (untuk pengembangan) dan MySQL atau PostgreSQL (untuk produksi) tanpa mengubah kode logika bisnis.

