

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan perumusan masalah, tujuan penelitian, serta hasil implementasi dan evaluasi yang telah dipaparkan pada bab-bab sebelumnya, maka dapat ditarik simpulan sebagai berikut:

1. Implementasi Algoritma: Penelitian ini telah berhasil mengimplementasikan algoritma *Conditional Random Fields* (CRF) untuk mendeteksi kesalahan penggunaan angka dan bilangan pada teks berita. Proses implementasi dilakukan melalui tahapan persiapan data (*scraping* dan ekstraksi kalimat), pelabelan data menggunakan format IOB (*Inside, Outside, Beginning*) untuk kategori kesalahan sesuai EYD, rekayasa fitur (*feature engineering*), pelatihan model, hingga evaluasi dan pengujian model. Model yang telah dilatih kemudian diintegrasikan ke dalam sistem *backend* menggunakan *framework* Flask, yang menyediakan API (*Application Programming Interface*) untuk menerima input teks dan memberikan respons berupa deteksi kesalahan beserta penjelasannya.
2. Kinerja Model: Evaluasi performa model CRF menunjukkan hasil yang sangat baik dalam mendeteksi kesalahan penggunaan angka dan bilangan. Berdasarkan pengujian menggunakan data uji, model menghasilkan nilai rata-rata F1-score, Precision, dan Recall (baik micro, macro, maupun weighted average) sebesar 0,94 atau 94%. Analisis confusion matrix juga menunjukkan bahwa model mampu membedakan antar-kategori kesalahan dengan tingkat kekeliruan yang sangat minim. Hal ini membuktikan bahwa metode CRF dengan fitur yang dirancang efektif untuk menangani permasalahan sekuensial dalam pengecekan tata bahasa Indonesia, khususnya terkait angka dan bilangan.

5.2 Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan selama proses pengembangan, terdapat beberapa saran yang dapat diajukan untuk pengembangan penelitian selanjutnya, yaitu:

1. Perluasan Cakupan Dataset: Meskipun augmentasi data telah dilakukan untuk menyeimbangkan kelas, penelitian selanjutnya disarankan untuk memperbanyak pengumpulan data riil dari artikel berita yang secara alami mengandung kesalahan penulisan angka dan bilangan. Hal ini bertujuan agar model dapat mempelajari variasi kesalahan yang lebih kompleks dan natural yang mungkin belum terwakili oleh data hasil augmentasi.
2. Penggunaan *Deep Learning*: Penelitian selanjutnya dapat mengeksplorasi penggunaan model *deep learning* untuk tidak hanya mendeteksi kesalahan, tetapi juga menghasilkan usulan koreksi kalimat secara otomatis. Pendekatan ini berpotensi memperluas fungsi sistem dari sekadar penanda kesalahan menjadi asisten penulisan yang mampu melakukan perbaikan langsung, sehingga memberi nilai tambah yang signifikan bagi proses penyuntingan teks berita.

