

## BAB 5 SIMPULAN DAN SARAN

### 5.1 Simpulan

Penelitian ini bertujuan untuk mengembangkan sistem deteksi kesalahan kapitalisasi sesuai dengan EYD menggunakan algoritma *Random Forest*. Hasil evaluasi menunjukkan bahwa model *Random Forest* dapat menghasilkan performa yang cukup baik dalam mendeteksi kesalahan kapitalisasi. Model ini mencapai akurasi 81% pada data uji, dan 82% setelah diterapkan pengaturan *class weight=balanced*, yang memberikan performa terbaik. Teknik *SMOTE* juga memberikan sedikit peningkatan, dengan akurasi mencapai 81,19%.

Metrik *precision*, *recall*, dan *f1-score* menunjukkan bahwa model ini cukup efektif, dengan keseimbangan yang baik antara *precision* dan *recall* untuk kedua kelas, meskipun masih terdapat beberapa kesalahan deteksi pada sampel minoritas (kelas 1). Secara keseluruhan, hasil penelitian ini memberikan kontribusi pada pengembangan sistem otomatis untuk mendeteksi kesalahan kapitalisasi dalam teks bahasa Indonesia sesuai dengan kaidah EYD.

Implementasi sistem ini dapat meningkatkan efisiensi dalam proses pemeriksaan teks dan memberikan solusi praktis untuk mendukung penulisan yang sesuai dengan aturan bahasa Indonesia.

### 5.2 Saran

Berdasarkan hasil penelitian ini, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut dalam sistem deteksi kesalahan kapitalisasi. Meskipun algoritma *Random Forest* menunjukkan kinerja yang cukup baik dan memberikan hasil yang memadai dalam sebagian besar kasus, namun hasil yang diperoleh tidak sepenuhnya optimal. Hal ini disebabkan oleh cakupan yang sangat luas dan kompleksitas dari permasalahan deteksi kesalahan kapitalisasi yang melibatkan berbagai aturan dan konteks yang lebih mendalam. Akibatnya, model terkadang kesulitan dalam menangkap pola yang lebih halus dan kontekstual dalam teks.

Sebagai langkah perbaikan, disarankan untuk mengeksplorasi penggunaan model berbasis *deep learning*, seperti *Long Short-Term Memory (LSTM)*, yang memiliki kemampuan lebih baik dalam memahami hubungan jangka panjang antar

kata serta konteks yang lebih kompleks dalam sebuah kalimat. Selain itu, untuk meningkatkan akurasi deteksi, pendekatan yang lebih terstruktur dalam penanganan aturan EYD dapat diterapkan. Salah satu solusi yang dapat dipertimbangkan adalah pembagian peran dalam penerapan deteksi kesalahan kapitalisasi, di mana masing-masing model fokus pada satu aturan EYD tertentu. Dengan membagi cakupan ke dalam bagian yang lebih kecil dan terfokus, setiap aturan EYD dapat ditangani lebih mendalam, sehingga memungkinkan penggabungan hasil deteksi dari setiap aturan tersebut untuk mencapai hasil yang lebih optimal dan akurat.

Penerapan pendekatan ini dapat membantu mengurangi kompleksitas dan memperbaiki akurasi deteksi kesalahan kapitalisasi, karena satu aturan saja sudah memiliki cakupan yang luas dan sering kali tidak cukup untuk mencakup semua aspek yang diperlukan dalam deteksi yang sempurna.

