

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan penelitian yang telah dilakukan, dapat ditarik kesimpulan bahwa model pendeteksi kesalahan tik menggunakan algoritma *random forest* telah berhasil dibangun. Dari hasil perhitungan *confusin matrix* yang dilakukan dengan menggunakan 100% *dataset* sebagai *data training* dan 20% dari *data training* sebagai data uji pada tahap evaluasi, didapat hasil dari *accuracy*, *precision*, *recall*, dan *F1 score* sebesar 100%. Hal ini menunjukkan bahwa model yang dibangun telah mampu mendeteksi kesalahan tik khususnya pada 50 artikel berita Tribunnews yang dimasukkan ketika melakukan pengujian. Namun perlu diingat bahwa hasil 100% ini adalah model mampu mendeteksi kata yang telah terdapat pada *dataset*. Berdasarkan hasil evaluasi yang telah dilakukan, dikarenakan kata yang dideteksi terdapat dalam *dataset* sehingga akurasi yang dikeluarkan adalah 100%. Namun jika tidak terdapat dalam *dataset*, seperti salah satu huruf tertukar atau terhapus, maka akan memungkinkan kata tersebut untuk dikenali oleh model yang dilakukan dengan *majority vote* dan akan memberikan koreksi kata benar dari kata tersebut dengan melihat tingkat kemiripan yang dilakukan oleh fungsi *get_close_matches* dari library *difflib*. Alasan dari penggunaan 20% *data training* sebagai data uji dikarenakan ketika melakukan perhitungan evaluasi, model menjadi sangat bergantung terhadap *dataset* sehingga jika pada perhitungan evaluasi terdapat kata yang tidak ada dalam *dataset*, maka akan menyebabkan model menjadi tidak dapat melakukan perhitungan evaluasi.

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, terdapat saran yang diberikan untuk pengembangan penelitian ke depannya, antara lain sebagai berikut.

1. Memperbanyak jenis kesalahan tik agar model kedepannya dapat lebih luas cakupannya dalam mendeteksi kesalahan tik yang bervariasi. Hal ini dikarenakan terdapat keterbatasan waktu dalam melakukan penelitian sehingga jenis kesalahan tik yang dimasukkan ke dalam model untuk saat ini adalah kesalahan-kesalahan yang dinilai umum terjadi pada 50 artikel berita yang digunakan.

2. Dapat memberikan hasil saran koreksi kata benar dari kata salah tik yang lebih baik dimana saran koreksi kata tidak hanya merupakan hasil dari *string matching* seperti yang dilakukan pada penelitian ini, akan tetapi dapat mengerti konteks dari kata salah tik sehingga hasil dari saran koreksi kata yang diberikan dapat lebih baik. Hal ini dapat dilakukan salah satunya dengan menambahkan algoritma probabilitas dalam menghasilkan saran koreksi kata.

