

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Implementasi algoritma *Symspell* dan *Bayesian network* dalam mendeteksi dan mengoreksi kata *typo* pada teks bacaan berbahasa Indonesia telah berhasil dilakukan. Implementasi dilakukan menggunakan bahasa pemrograman Python dengan bantuan perangkat lunak Jupyter Notebook sebagai *text editor* sekaligus *compiler*. Aplikasi *spell checker* dirancang dan dibentuk dalam bentuk aplikasi *desktop* dengan menggunakan *library* PyQt5 dalam pembuatan tampilan antarmuka aplikasi dan PyInstaller dalam pengonversian kode program ke dalam bentuk aplikasi *desktop*.

Parameter terbaik yang menjadi hasil akhir dari eksperimen sistem *spell checker* bahasa Indonesia adalah parameter dengan nilai *depth search* satu (1) yang merupakan jumlah iterasi pada algoritma *Symspell* dan nilai *suggestions limit* sepuluh (10) yang merupakan jumlah maksimal kata pada daftar kandidat kata algoritma *Symspell*. Rata-rata akurasi deteksi dan koreksi kata *typo* jika sistem menggunakan parameter terbaik adalah sebagai berikut.

1. Rata-rata akurasi deteksi sebesar 96,3020% tanpa atau dengan pengaplikasian algoritma *Bayesian network*.
2. Rata-rata akurasi koreksi sebesar 82,7056% dengan pengaplikasian algoritma *Bayesian network*.
3. Rata-rata akurasi koreksi sebesar 76,9712% tanpa pengaplikasian algoritma *Bayesian network*.

5.2 Saran

Saran untuk penelitian selanjutnya yang dapat dilakukan untuk meningkatkan akurasi yang didapat adalah sebagai berikut.

1. Mengkategorikan setiap kata pada unigram berdasarkan perannya dalam struktur kalimat (subjek, predikat, objek, dan keterangan) untuk menjaga relevansi pasangan kata yang digunakan pada bigram.
2. Mengatasi kasus di mana satu kata *typo* yang terdeteksi dapat dikoreksi menjadi beberapa kata yang sah maupun sebaliknya.
3. Memperlakukan kata-kata yang berada di dalam tanda kurung sebagai kalimat sendiri yang tidak memiliki hubungan dengan kalimat yang sedang diproses.