

## BAB V

### KESIMPULAN

#### 5.1 KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan pada analisis klasifikasi teks menggunakan metode yang sesuai pada *dataset* CPNS 2024, dapat disimpulkan bahwa:

##### 1. Pengembangan Model Klasifikasi Teks dengan Metode *Naïve Bayes*

Model klasifikasi teks berhasil dikembangkan dengan menerapkan metode *Naïve Bayes* pada *dataset* CPNS 2024. Model ini mampu mengelompokkan data berdasarkan *education\_level* dengan tingkat akurasi yang tinggi. Selain itu, kombinasi dengan metode TF-IDF membantu dalam meningkatkan kualitas representasi data teks, sehingga memperbaiki performa model secara keseluruhan.

##### 2. Konversi Data Teks Menggunakan TF-IDF

Data formasi CPNS yang berbentuk teks berhasil dikonversi menjadi representasi numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Teknik ini memungkinkan sistem untuk mengidentifikasi kata-kata yang paling signifikan dalam setiap kategori *education\_level* pada *dataset* CPNS 2024 dan meningkatkan akurasi model dalam mengklasifikasikan teks.

### 3. Hasil Evaluasi Model

Berdasarkan pengujian yang dilakukan, model klasifikasi berhasil mencapai tingkat akurasi 96% yang menunjukkan bahwa metode yang diterapkan efektif untuk mengolah *dataset* CPNS 2024 dengan total 9775 data. Evaluasi menggunakan metrik *precision*, *recall*, dan *F1-score* juga menunjukkan hasil yang memuaskan, yang menandakan bahwa metode ini dapat digunakan sebagai solusi klasifikasi analisis dataset CPNS 2024.

## 5.2 SARAN

Berdasarkan penelitian yang telah dilakukan, peneliti dapat memberikan rekomendasi untuk pengembangan penelitian sebagai berikut:

### 1. Pengembangan *Dataset* CPNS

Untuk meningkatkan kemampuan model dalam menggeneralisasi data, disarankan untuk menggunakan *dataset* yang mencakup berbagai periode seleksi CPNS. Selain itu, mempertimbangkan variasi kategori pekerjaan serta wilayah geografis yang berbeda dapat meningkatkan ketepatan model.

### 2. Peningkatan Akurasi Model dengan *Algoritma Hybrid*

Penelitian selanjutnya dapat menggabungkan metode lain atau pendekatan *hybrid* untuk meningkatkan akurasi dan performa model.

### 3. Penerapan *Deep Learning*

Penggunaan teknik *deep learning* seperti *Recurrent Neural Network* (RNN) atau model berbasis *Transformer* seperti BERT dapat membantu meningkatkan akurasi klasifikasi dengan memahami hubungan antar kata dalam teks dengan lebih baik

