

## **BAB V**

### **KESIMPULAN**

#### **5.1. Kesimpulan**

Berdasarkan hasil Analisa dan implementasi terhadap klasifikasi ulasan aplikasi Ruangguru pada playstore yang telah di lakukan, dapat disimpulkan sebagai berikut ini :

1. Pada penelitian ini terdapat beberapa tahapan yang dilakukan untuk membangun model analisis klasifikasi dan menemukan model terbaik untuk dapat di terapkan pada saat visualisasi. Beberapa tahapan diantaranya tahap Analisa Kebutuhan, tahap persiapan data (Pengambilan data, Pelabelan, Filtering,) kemudian tahapan sentiment analisis yang meliputi (text preprocessing, pembobotan TFIDF, Pembagian data, Ekstraksi Fitur, Klasifikasi, Pengukuran Akurasi), dan yang terakhir adalah tahapan visualisasi.
2. Pada penelitian ini untuk membangun suatu model klasifikasi dilakukan pembentukan model skenario eksperimen. Skenario Model pertama yaitu klasifikasi dengan *Gaussian Naïve Bayes*, Skenario Model kedua yaitu klasifikasi dengan *Gaussian Naïve Bayes* dengan menggunakan Metode *Principal Component Analysis* (PCA) sebagai metode Ekstraksi Fitur, Kemudian Skenario Model ketiga yaitu klasifikasi *Gaussian Naïve Bayes* dengan menggunakan Variabel Smoothing, dan Skenario Model terakhir yaitu *Gaussian Naïve Bayes* dengan menggunakan nilai *Var Smoothing* terbaik (0.5) dan *Principal Component Analysis*(PCA). Pada masi. Pada tiap skenario modelnya memiliki beberapa eksperimen yang mana nilai K pada tiap tiap eksperimennya adalah 10.
3. Skenario Model pertama pada penelitian ini yaitu diterapkan model *Gaussian naïve bayes* pada analisis sentiment klasifikasi ulasan aplikasi Ruangguru didapatkan hasil performansi dengan nilai akurasi 55.58%, Precision 89.29%, Recall 50.72%, dan F1 Score 64.7%.

4. Skenario Model kedua pada penelitian ini yaitu diterapkan model *Gaussian naïve bayes* pada analisis sentiment klasifikasi ulasan aplikasi Ruangguru dan ditambahkan metode *Principal Component Analysis* sebagai metode dari *feature selection* dengan pemotongan fitur optimum yaitu sebanyak 101 didapatkan hasil performansi dengan nilai akurasi 74.65%, Precision 83.71%, Recall 84.93%, dan F1 Score 84.32%. Dari penerapan pemotongan fitur sebesar 101 ini dapat diketahui bahwa adanya peningkatan hasil klasifikasi dan juga performance dari skenario pertama, dari sini dapat disimpulkan apabila penggunaan *Principal Component Analysis* dengan menerapkan jumlah pemotongan fitur yang tepat akan mempengaruhi peningkatan hasil klasifikasi dan performansi.
5. Skenario Model ketiga pada penelitian ini yaitu diterapkan model *Gaussian naïve bayes* pada analisis sentiment klasifikasi ulasan aplikasi Ruangguru dan ditambahkan optimasi parameter Variabel Smoothing, pada skenario ini dapat didapatkan hasil performansi dan klasifikasi dengan nilai optimum Ketika menerapkan parameter optimasi sebesar 0.5. Dari penerapan tersebut dihasilkan nilai akurasi 88.56%, Precision 94.2%, Recall 91.29%, dan F1 Score 85.58%. Dari penerapan parameter optimasi var smoothing = 0.5 pada klasifikasi *Gaussian naïve bayes* didapatkan hasil klasifikasi dan performansi yang lebih optimum dari skenario sebelumnya.
6. Skenario Model keempat pada penelitian ini yaitu diterapkan model *Gaussian naïve bayes* pada analisis sentiment klasifikasi ulasan aplikasi Ruangguru dan ditambahkan dengan parameter optimum *Var Smoothing* terbaik sebesar 0.5 dan *Principal Component Analysis*(PCA) sebesar 101 didapatkan hasil performansi dengan nilai akurasi 73.49%, Precision 82.35%, Recall 85.22%, dan F1 Score 83.76%. Dari penerapan parameter optimum dan juga pemotongan fitur yang dianggap optimum menghasilkan nilai klasifikasi, dari sini didapatkan nilai performansi dan nilai klasifikasi yang nilainya jauh lebih rendah dari pada skenario model sebelumnya.
7. Skenario Model Kelima pada penelitian ini yaitu diterapkannya model *Gaussian naïve bayes* pada analisis sentiment klasifikasi

ulasan aplikasi Ruangguru dan ditambahkan dengan metode untuk mengatasi imbalance data yaitu SMOTE, pada skenario eksperimen ini data yang dianggap tidak seimbang (kelas negative) akan di duplikasikan sehingga menjadi seimbang dengan data yang sudah ada. Dari sini digunakan teknik oversampling menggunakan metode SMOTE. Hasil dari penggunaan SMOTE ini

8. Pada beberapa skenario yang telah di jalankan dapat disimpulkan bahwasanya terdapat beberapa metode atau cara yang difungsikan untuk mengoptimalkan hasil klasifikasi dan juga performansi yang pertama menggunakan pemotongan fitur dengan metode Principal Component Analysis, dari penerapan metode ini diketahui nilai klasifikasi dan performansi meningkat dikarenakan ada beberapa fitur penting yang digunakan dan fitur yang dianggap tidak relevan di reduksi. Kemudian Penerapan parameter Var Smoothing dengan nilai 0.5 juga sangat berpengaruh dalam nilai klasifikasi dan performansi karena terbukti mengoptimalkan hasil. Kemudian yang terakhir penerapan SMOTE dari penerapan metode SMOTE ini mendapatkan nilai akurasi yang lebih akurat dikarenakan data positif dan negative dapat dikatakan seimbang. Hasil dari penerapan SMOTE ini tidak jauh lebih baik jika dibandingkan dengan hasil Ketika tidak menggunakan SMOTE. Pada kasus ini dapat disimpulkan bahwa penggunaan data asli jauh lebih baik dari pada menggunakan oversampling.
9. Visualisasi analisis sentiment terhadap ulasan aplikasi ruangguru dilakukan dengan menunjukkan hasil klasifikasi mulai dari table dan grafik distribusi data, wordcloud positif negatif, hasil validasi, hingga hasil pengujian, dan *Confussion Matrix*

## 5.2. Saran

Adapun beberapa saran yang dapat diberikan bagi pengembang penelitian ini lebih lanjut adalah sebagai beriku :

1. Dapat menambahkan kelas netral pada klasifikasi ulasan aplikasi ruangguru.

2. Dapat menambahkan suatu sistem yang mampu mendeteksi dan mengatasi kata singkatan ataupun typo pada pengolahan text.
3. Dapat menambahkan metode yang dapat menyeimbangkan data agar dapat menghasilkan performa yang lebih baik
4. Dapat menambahkan metode ekstraksi fitur tambahan seperti kernel PCA yang nantinya mungkin dapat menghasilkan performa yang lebih baik dan sebagai bahan perbandingan yang tepat.
5. Dapat menerapkan SMOTE sebagai salah satu metode imbalance data ke pada semua skenario dan dibuat suatu perbandingan anatara yang menggunakan SMOTE dan tanpa menggunakan SMOTE.

## **Lampiran**