

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan penelitian yang telah dilakukan, simpulan yang didapat pada analisis sentimen dengan algoritma *Naive Bayes* adalah sebagai berikut :

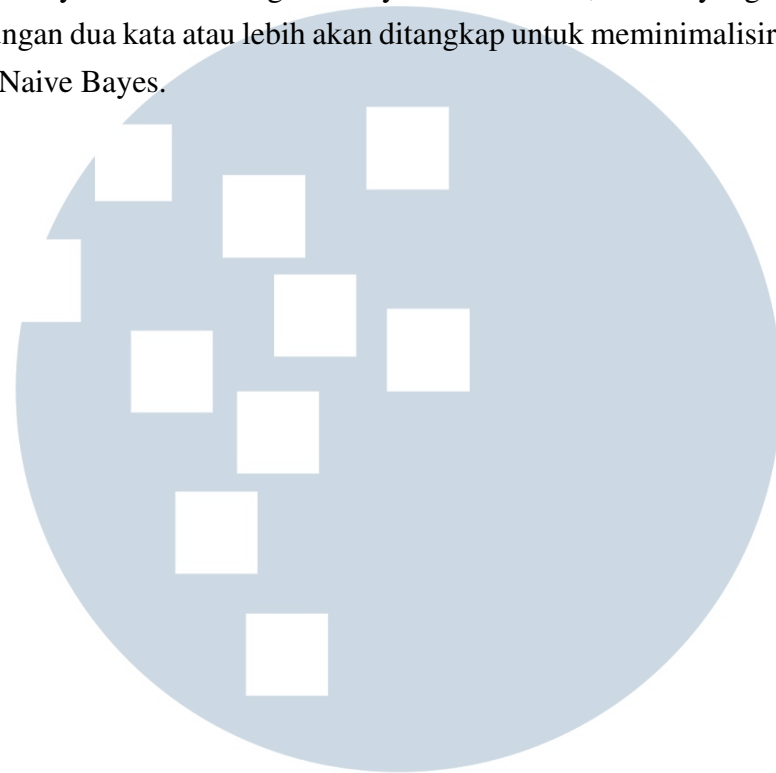
1. Analisis sentimen Naive Bayes telah berhasil dilakukan untuk menganalisa tanggapan mengenai resesi 2023 pada media sosial twitter. Data penelitian ini didapatkan melalui hasil scraping dari twitter dan setelah melalui proses preprocessing dan labeling, jumlah total tweets yang didapatkan adalah 6.685 dengan rincian 1.696 tweets positif, 4.631 tweets negatif, dan 358 tweets netral. Namun karena penelitian ini hanya mencakup tweets positif dan negatif, maka data tweets netral akan dibuang. Pada penelitian ini, terdapat perbandingan pembobotan antara penggunaan TF-IDF dan Count Vectorizer. Kemudian karena data tidak balance maka terdapat uji coba dengan menggunakan SMOTE dan Near Miss. Pada akhir penelitian juga diuji validasi dengan menggunakan Cross Validation.
2. Berdasarkan hasil uji coba yang telah dilakukan, hasil akurasi tertinggi yang didapatkan adalah dari penggunaan pembobotan Count Vectorizer dengan algoritma Multinomial Naive Bayes yaitu dengan akurasi sebesar 80,88%, *precision* 77%, *recall* 71%, dan *F1-score* sebesar 73%. dan tidak melalui proses *balancing data*. Uji coba yang dilakukan dengan menggunakan SMOTE dan Near Miss tidak secara pasti akan menaikkan hasil akurasi, namun dengan melakukan SMOTE dan Near Miss nilai dari *precision*, *recall*, dan *f1-score* akan menjadi lebih stabil.

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, berikut adalah saran yang dapat diberikan untuk pengembangan selanjutnya, antara lain :

1. Menggunakan modeling lain seperti Long Short Term Memory (LSTM) sebagai pembanding. Karena kelemahan dari algoritma Naive Bayes adalah tidak melihat hubungan antar kata sedangkan untuk Long Short Term Memory (LSTM) menangkap hubungan antar kata.

2. Mencoba menambahkan fitur N-Gram ketika ingin menggunakan algoritma Naive Bayes. Karena dengan adanya fitur N-Gram, makna yang didapat dari gabungan dua kata atau lebih akan ditangkap untuk meminimalisir kelemahan dari Naive Bayes.



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA