

## BAB V PENUTUP

### 5.1 Kesimpulan

Berdasarkan analisis dan pengujian yang dilakukan pada bab sebelumnya, maka kesimpulan yang dapat diambil adalah sebagai berikut:

1. Metode Naive Bayes Classifier berhasil diimplementasikan dengan bantuan beberapa metode seperti preprocessing, pelabelan vader lexicon, serta ekstraksi fitur TF-IDF, dari pengujian Multinomial *default* parameter. NBC mampu menganalisa klasifikasi sentimen dari komentar pengguna YouTube terhadap produk handphone keluaran terbaru dari brand itel di Indonesia dengan 1970 data bersih, dihasilkan 1125 sentimen netral, 656 sentimen positif, dan 205 sentimen negatif.
2. Hasil yang didapatkan menunjukkan bahwa banyak masyarakat di Indonesia yang cenderung biasa saja (netral) terhadap produk handphone keluaran terbaru dari brand ITEL dan kebanyakan masyarakat berkomentar buruk (negatif) dikarenakan nama brandnya sendiri yang memiliki makna jelek terkhusus pada daerah Jawa.
3. Performa algoritma NBC bekerja dengan baik dengan estimator terbaik menggunakan MultinomialNB default parameter, splitting data 70:30, dan TF-IDF dengan parameter N-Gram Range (1,2). Hasil evaluasi terbaik dari delapan percobaan yang dilakukan adalah nilai *accuracy* sebesar 74%.
4. Penelitian ini dilakukan dengan total delapan percobaan, yang dibagi menjadi dua bagian yaitu pembagian data 70:30 dengan empat percobaan dan pembagian data 80:20 juga dengan empat percobaan. Percobaan ini memberikan hasil yang tidak jauh berbeda yaitu, tingkat akurasi sebesar 73% pada pembagian data 80:20 dan 74% pada pembagian data 70:30. Hal ini menandakan perubahan pembagian data

tidak memberikan perbedaan yang signifikan untuk kasus pada dataset ini.

## 5.2 Saran

Untuk menambah referensi penelitian selanjutnya ada beberapa saran yang dikemukakan sebagai berikut:

1. Disarankan untuk mengembangkan kamus yang mencakup kata-kata tidak baku agar lebih banyak kata tidak baku yang dapat disaring, sehingga memungkinkan data menjadi lebih bersih dan terpilah dengan lebih baik.
2. Penelitian ini masih terbatas pada analisis teks dalam youtube. Namun, potensi pengembangan lebih lanjut dapat melibatkan faktor lain seperti penggunaan emoji/emotikon dan juga gambar dalam konteks analisis tersebut.
3. Disarankan untuk menggunakan dataset yang lebih luas agar hasil yang diperoleh dapat mencapai tingkat optimal yang lebih tinggi. Dengan menggunakan dataset yang lebih besar, akan ada keuntungan dalam mendapatkan hasil yang lebih baik dan lebih representatif.
4. Melakukan pengembangan lebih lanjut dari penelitian ini dengan mengubah beberapa hal seperti, dataset yang dipakai untuk pengujian diambil dari sosial media lain, melakukan perbandingan dengan metode lain, mengubah parameter pengujian yang digunakan dan beberapa hal lainnya.