

BAB V PENUTUP

5.1 Kesimpulan

Pada penelitian ini dilakukan pengujian dataset menggunakan Naïve Bayes dan pembobotan TF-IDF dengan cara melakukan proses *mining data* pada aplikasi Twitter, kemudian akan memasuki *pre-processing* guna data dapat diolah dan digunakan, dataset akan dilabeli dengan positif dengan negatif. Untuk selanjutnya dilakukan proses training dan prediksi. Dari tahap penelitian ini nanti akan disajikan nilai akhir berupa *accuracy*, *recall*, *precision* dan *f1 score*. Dari setiap algoritma *k-fold(5) cross validation*. Dari pengujian Naïve Bayes didapatkan nilai 81,33% dengan pembobotan TF-IDF dan *recall* 74,29%, *precision* 83,87% dan *f1 score* 78,79%.

Sedangkan pengujian kedua dengan dataset baru, memiliki hasil yang berbeda yaitu untuk nilai *accuracy* 91,67%. Pada pengujian kedua ini nilai *accuracy* lebih tinggi, tetapi nilai *recall*, *precision*, dan *f1 score* 0, dikarenakan label pada dataset tidak seimbang, karena dominan negatif. Untuk itu diperlukan pengujian kembali. Dari hasil penelitian ini diharapkan calon pelanggan dapat lebih bijak sebelum memutuskan dalam menentukan provider yang akan di gunakan. Walaupun pengujian terhadap testing memberikan hasil memuaskan, untuk perlu diperhatikan bahwa presisi atas suatu klasifikasi diperlukan evaluasi lagi terhadap prediksi yang telah didapat.

5.2 Saran

Berdasarkan kesimpulan diatas, penulis memberikan saran dari hasil penelitian sebagai berikut.

1. Untuk penelitian menggunakan metode naïve bayes, dapat menggunakan metode algoritma lainnya atau menggabungkan dengan beberapa metode, seperti dengan *support vector machine*.
2. Menggunakan pembobotan TF – IDF, perlu dikombinasikan dengan metode lainnya.
3. Dikarenakan dataset yang digunakan seimbang, perlu untuk selanjutnya ditingkatkan jumlah dataset agar dapat menghasilkan nilai akurasi yang lebih

tepat dan baik.

4. Pentingnya waktu atau momen pengambilan data, sesuai dengan studi kasus, akan menghasilkan dataset yang berkualitas.
5. Pentingnya kualitas dataset juga berpengaruh terhadap hasil akhir akurasi, untuk itu disarankan data pengujian memiliki sumber yang berkualitas dan data berlabel dengan baik.

