

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian sudah berhasil menerapkan model klasifikasi teks data *tweet* dari pengguna pada media sosial *twitter* terhadap pengguna dengan adanya metode pembayaran dengan *e-wallet* Ovo, Gopay dan Dana agar mendapatkan kelas sentimen dengan kamus *lexicon based*, memberikan visualisasi kata dengan *wordcloud* dan mendapatkan hasil *confusion matriks* dengan algoritma *Naïve Bayes Classifier*. Berdasarkan penelitian yang telah dilakukan dapat ditarik kesimpulan diantaranya yaitu:

1. Pada para pengguna data *tweet* yang diambil melalui metode *lexicon based* dengan total data bersih dengan total dari *e-wallet* Dana sebanyak yang terkumpul data bersih sebanyak 3980 dengan kelas positif sebanyak 1780 *tweet* dan kelas negatif sebanyak 2200 *tweet*. Sedangkan *e-wallet* OVO mendapatkan data bersih sebanyak 3976 *tweet* dengan kelas positif sebanyak 953 *tweet* dan kelas negatif sebanyak 3023 *tweet*. Dan yang terakhir *e-wallet* LinkAja mendapatkan data bersih sebanyak 3972 *tweet* dengan kelas positif sebanyak 1429 *tweet* dan kelas negatif sebanyak 2543 *tweet* dari penelitian ini.
2. Pada tahap hasil pengujian kinerja dengan menggunakan algoritma *Naïve Bayes Classifier* dalam menganalisis sentimen *tweet* dengan adanya *e-wallet* dengan *model split data* dengan 4 rasio menunjukkan bahwa hasil

akurasi terbaik dari pembagian data tersebut pada ketiga aplikasi *e-wallet* tersebut yang mendapatkan akurasi terbesar didapat pada aplikasi *e-wallet* LinkAja dengan akurasi sebesar 85.17%, sedangkan yang kedua didapat pada aplikasi *e-wallet* OVO dengan akurasi sebesar 81.97%, sedangkan yang terakhir didapat pada aplikasi *e-wallet* Dana dengan akurasi sebesar 79.39%. Hal ini menunjukkan bahwa pengguna *tweet* dengan adanya metode pembayaran *e-wallet* yang benar diprediksi berkomentar positif maupun yang berkomentar negatif dari keseluruhan pengguna *tweet* memiliki *accuracy* dengan nilai sebesar 85.17%, sedangkan tingkat keakuratan pengguna *tweet* terhadap metode pembayaran *e-wallet* yang bersentimen positif dari keseluruhan pengguna yang diprediksi bersentimen positif memiliki *precision* dengan nilai sebesar 87.50%, sedangkan nilai keberhasilan pengguna *tweet* terhadap metode pembayaran *e-wallet* yang diprediksi berkomentar positif dibandingkan keseluruhan pengguna *tweet* yang sebenarnya bersentimen positif maka memiliki *recall* dengan nilai sebesar 68.53%, dan hasil perbandingan rata-rata dari tingkat akuratan dan keberhasilan pengguna *tweet e-wallet* memiliki *f1-score* dengan nilai sebesar 76.86%.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, peneliti memiliki beberapa saran yang dapat dijadikan dan dipertimbangkan untuk penelitian selanjutnya sebagai berikut:

1. Untuk penelitian ini selanjutnya dapat diterapkan dengan algoritma

selanjutnya untuk *Random Forest*, *K-Nearest Neighbor* (KNN), Algoritma *C4.5* dan lain sebagainya. Karena pada *Random Forest* salah satu algoritma terbaik untuk klasifikasi data yang besar atau memiliki banyak fitur dengan akurasi tinggi.

2. Berdasarkan hasil analisis yang diperoleh, penelitian ini terbatas dan mengalami kendala salah satunya di tahap *text preprocessing*, dimana pengolahan data pada penelitian ini terbatas pada tahapan normalisasi kata yang terdapat pada kamus normalisasi Bahasa Indonesia saja. Sedangkan pada media sosial *twitter*, banyak ditemukan bahasa yang tidak baku yang sering digunakan sehari-hari seperti “kamu”, “yaha”, “eug”, “kalem”, “yuhu” dan banyak kata yang menggunakan huruf ganda “cangggihhhhh”, “perhatiannnn”, “mantapppttsss” yang kemudian kata-kata tersebut tidak ternormalisasi. Selain itu, beberapa data banyak menggunakan bahasa daerah yang tidak mengalami normalisasi.
3. Penggunaan kamus *Lexicon Inset* dengan tujuan untuk memberikan label positif dan negatif yang berisi kata-kata dengan *polarity score* -5 sampai dengan +5 masih dikatakan terbatas. Untuk penelitian selanjutnya dapat menggunakan metode seperti *TextBlob* atau bekerja sama dengan seorang ahli dalam bidang bahasa.
4. Diperlukannya kajian lebih lanjut pada proses *text processing* di bagian *library* yang dipakai untuk menseleksi data *crawling* yang terutama pada tahap visualisasi *wordcloud* kenapa kata netral masih muncul di gambar visualisasi.