

BAB 4

HASIL PENELITIAN

4.1 RINGKASAN HASIL PENELITIAN

Hasil penelitian perbandingan metode Naïve Bayes Classifier dan Support Vector Machine untuk analisis sentimen *review* pengguna Shopee di media Google Play Store. Pada penelitian ini data yang diambil adalah sebanyak 5000 data mentah menggunakan Google Colab. Data mentah tersebut dilakukan *preprocessing* sehingga mendapatkan sebanyak 3309 data bersih, dimana 600 data digunakan untuk *training* dengan label positif sebanyak 200 data, negatif 200 data, dan netral 200 data. Data hasil *preprocessing* juga digunakan untuk proses *testing* yaitu sebanyak 600 data dengan 200 data positif, 200 data negatif, dan 200 data netral. Analisis sentimen ini dibangun untuk mengolah data *review* pengguna Shopee pada media Google Play Store menggunakan Bahasa pemrograman Python dengan perbandingan metode Naïve Bayes Classifier dan Support Vector Machine untuk menghasilkan perhitungan sentimen menggunakan *framework* Flask.

4.2 AKURASI METODE

Pada penelitian ini dilakukan menggunakan dua metode yaitu Naïve Bayes Classifier dan Support Vector Machine, dimana menggunakan data sebanyak 600 data sebagai *training* yang masing-masing memiliki jumlah data yang sama yaitu 200 data positif, 200 data netral dan 200 data negatif. Sedangkan untuk data *testing* yang digunakan yaitu sebanyak 450 data dengan masing-masing data positif 150, netral 150 dan negatif yaitu 150 data. Hasil dan pembahasan mengenai pengujian yang telah dilakukan menggunakan metode NBC dan SVM dapat dijelaskan sebagai berikut:

4.3 HASIL EVALUASI KLASIFIKASI

Confusion matrix merupakan perhitungan yang digunakan untuk hasil evaluasi klasifikasi data *testing* yang menghitung nilai akurasi untuk menentukan perbedaan antara data *training* dan data *testing*. Data *testing* menggunakan data

sebanyak 450 data yang telah diberikan label positif, negatif dan netral diujikan menggunakan model yang sudah dibangun sebelumnya. Hasil perhitungan *multiclass confusion matrix* data *testing* NBC ditunjukkan pada Tabel 4.1.

Tabel 4.1 Hasil Perhitungan Confusion Matrix testing NBC

Kelas Aktual	Kelas Prediksi		
	Positif	Negatif	Netral
Positif	145	2	4
Negatif	1	114	1
Netral	8	6	169

Tabel 4.1 menjelaskan bahwa hasil dari perhitungan untuk data *Testing* menggunakan *confusion matrix* yaitu TPos = 145, FPostNeg = 2, FPostNet = 4, FNegPost = 1, TNeg = 114, FNegNet = 1 dan FNetPos = 8, FNegNet = 6, TNet = 169. Hasil klasifikasi data *testing* NBC dapat dilihat pada Tabel 4.2.

Tabel 4.2 Hasil Klasifikasi *Testing* NBC

Accuracy	0,75
Precision	0,76
Recall	0,75
F1 Score	0,75

Kemudian terdapat hasil perhitungan *confusion matrix* SVM yang dapat dilihat pada Tabel 4.3.

Tabel 4.3 Hasil Perhitungan *Confusion Matrix* SVM

Kelas Aktual	Kelas Prediksi		
	Positif	Negatif	Netral
Positif	147	0	4
Negatif	2	107	7
Netral	1	2	180

Tabel 4.3 menjelaskan bahwa hasil dari perhitungan untuk data *Testing* menggunakan *confusion matrix* yaitu $T_{Pos} = 147$, $F_{PostNeg} = 0$, $F_{PostNet} = 4$, $F_{NegPost} = 2$, $T_{Neg} = 107$, $F_{NegNet} = 7$ dan $F_{NetPos} = 1$, $F_{NegNet} = 2$, $T_{Net} = 180$. Hasil klasifikasi data *testing* SVM dapat dilihat pada Tabel 4.4.

Tabel 4.4 Hasil Klasifikasi *Testing* SVM

Acurracy	0,80%
Precision	0,80%
Recall	0,80%
F1 Score	0,79%

Dari hasil perhitungan klasifikasi NBC dengan SVM yang dapat dilihat pada table diatas dapat dibandingkan bahwa hasil klasifikasi SVM cukup lebih tinggi dibandingkan dengan hasil klasifikasi menggunakan NBC, dimana SVM menghasilkan akurasi data *testing* 0,80% sedangkan NBC menghasilkan 0,75%.

4.4 PERBANDINGAN AKURASI NBC DAN SVM

Pada penelitian ini dihasilkan perbandingan dari metode NBC dan SVM untuk data yang telah dilakukan pengujian. Hasil perbandingan tingkat akurasi dari kedua metode tersebut dapat dilihat pada Tabel 4.5.

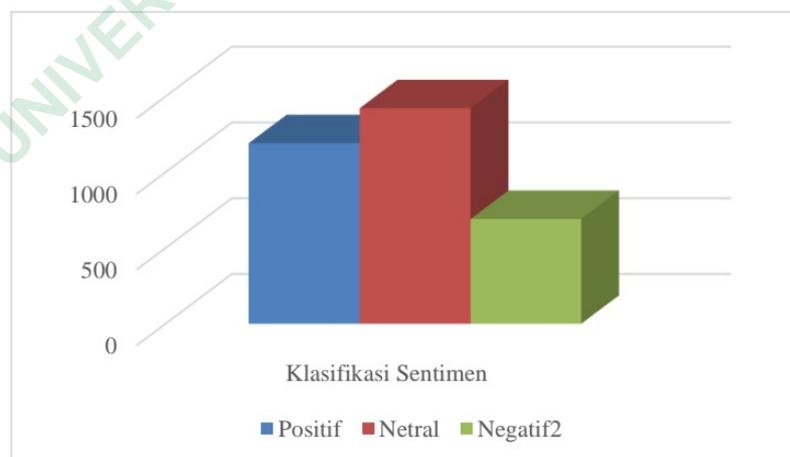
Tabel 4.5 Perbandingan Akurasi

Metode	NBC	SVM
Akurasi	75%	80%

Tingkat akurasi yang didapatkan untuk perbandingan yaitu hasil dari data *testing* dimana data yang digunakan yaitu sebanyak 400 data dengan masing-masing jumlah data positif yaitu 150, data netral 150 dan data negatif 150. Dari hasil tingkat akurasi tersebut dapat disimpulkan bahwa metode SVM memiliki nilai tingkat akurasi sedikit lebih tinggi dibandingkan dengan NBC.

4.5 HASIL KLASIFIKASI

Berdasarkan hasil akurasi tertinggi menggunakan algoritma SVM maka dihasilkan klasifikasi pada penelitian ini yang didapatkan dari data keseluruhan yaitu sebanyak 3309 data, dengan rincian 600 data tweet yang telah dilabeli secara manual dengan masing-masing nilai yang sama dalam proses *training* kemudian untuk proses *testing* dengan data sebanyak 450 data *review* yang dimana telah diduplikasinya nilai akurasi yang baik dari tahap *training* dan *testing*, hasil prediksi pada data *review* keseluruhan yang telah diketahui positif, netral dan negatif dapat dilihat pada Gambar 4.1.

**Gambar 4.1** Grafik Klasifikasi Sentimen

Pada Gambar 4.1 menunjukkan bahwa menghasilkan 1192 data positif, 1423 data netral dan 692 data negatif. Pada sentiment positif banyak masyarakat berkomentar mengenai gratis ongkir yang diberikan oleh Shopee, barang yang dijual sesuai dan *seller* yang ramah. Berikut contoh beberapa *review* positif dapat dilihat pada Tabel 4.6.

Tabel 4.6 Data *Review* Positif

No	Data <i>Review</i>
1.	banyak gratis ongkir promo kak
2.	belanja enak gratis ongkir ketimbang belanja lazada toko dia mending belanja disini murah aman
3.	bagus segi kirim dan kurir nya ramah
4.	harga murah lengkap barang kualitas sesuai tidak kecewa
5.	mudah cari barang selektif pilih barang sesuai pesanan

Sedangkan untuk sentimen negatif pada data review banyak yang membahas mengenai aplikais yang semakin hari lemot, terlalu banyak video iklan yang tidak jelas dan terlalu banyak iklan yang mengganggu. Contoh dari beberapa sentimen negatif dari masyarakat dapat dilihat pada Tabel 4.7.

Tabel 4.7 Data *Review* Negatif

No	Data <i>Review</i>
1.	iklan aplikasi nya mohonn pliss ganggu
2.	aplikasi lemottt banget emosi
3.	lemotttt bgttt pliss jaring bagus buka trss load nya lamaaa
4.	sopi skrng lot banget video reel nya apus sopi tolong baik ngambil enak doang
5.	kesini parah kaya lot hapus chace lakuin tetep lot gara gara nama shope video shope live belanja susah hapus belanja lihat video lihat video mending lihat youtube ramai minat

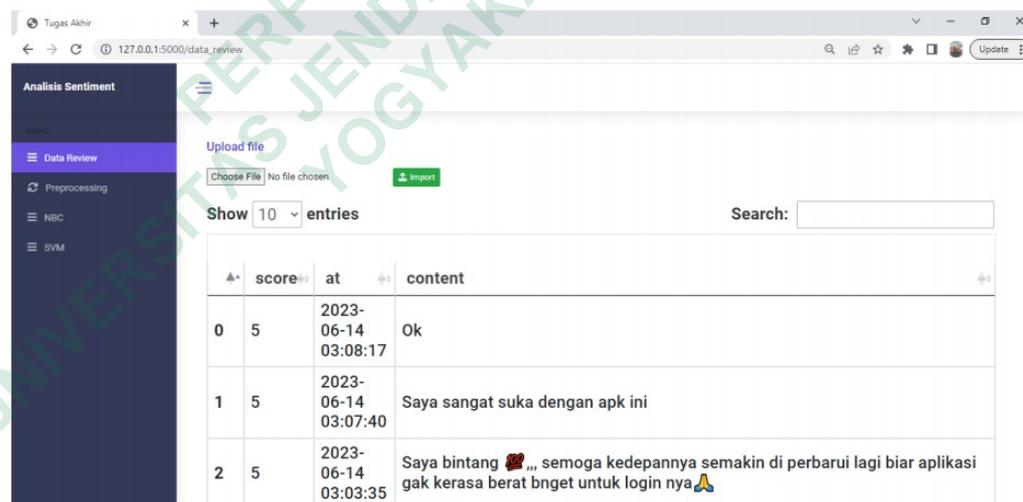
Berdasarkan hasil klasifikasi sentimen pada data *review* banyak yang diklasifikasikan ke negatif dikarenakan permasalahan aplikasi yang lambat, iklan yang terlalu mengganggu pengguna, video yang tidak jelas.

4.6 IMPLEMENTASI DESAIN *INTERFACE*

Implementasi *desain interface* pada penelitian ini yaitu menrealisasikan desain yang sudah dibuat kedalam sistem yang nyata dengan menggunakan bahasa pemrograman python. Implementasi desain *interface* kedalam sistem *web* ini digunakan untuk analisis sentimen pengguna Shopee pada media Google Play Store.

4.6.1 HALAMAN *UPLOAD DATA REVIEW*

Pada halaman *upload data review* digunakan untuk mengupload file CSV. Pada halaman ini akan ditampilkan data *review* dari file yang sudah diimport kedalam sistem. Implementasi halaman *upload data review* terdapat pada Gambar 4.2.



	score	at	content
0	5	2023-06-14 03:08:17	Ok
1	5	2023-06-14 03:07:40	Saya sangat suka dengan apk ini
2	5	2023-06-14 03:03:35	Saya bintang 🌟,,, semoga kedepannya semakin di perbarui lagi biar aplikasi gak kerasa berat bnget untuk login nya 🙏

Gambar 4.2 Implementasi Halaman *Data Review*

Data yang ditampilkan pada halaman *data review* merupakan data yang belum dilakukan proses *preprocessing* yang masih terdapat score dan juga tanggal, sedangkan *content* merupakan kalimat ulasan dari aplikasi Shopee pada media Google Play Store.

4.6.2 HALAMAN *PREPROCESSING*

Pada halaman *preprocessing* dilakukan pembersihan data yang sudah diupload pada halaman *upload data review*. Pada halaman ini akan ditampilkan data yang sudah melalui proses *preprocessing*, dan hasil data *preprocessing* akan digunakan untuk data pelabelan manual, *training*, *testing* dan juga klasifikasi. Data *preprocessing* juga dapat disimpan dengan format *xlsx*. Berikut ini implementasi halaman *preprocessing* pada Gambar 4.3.

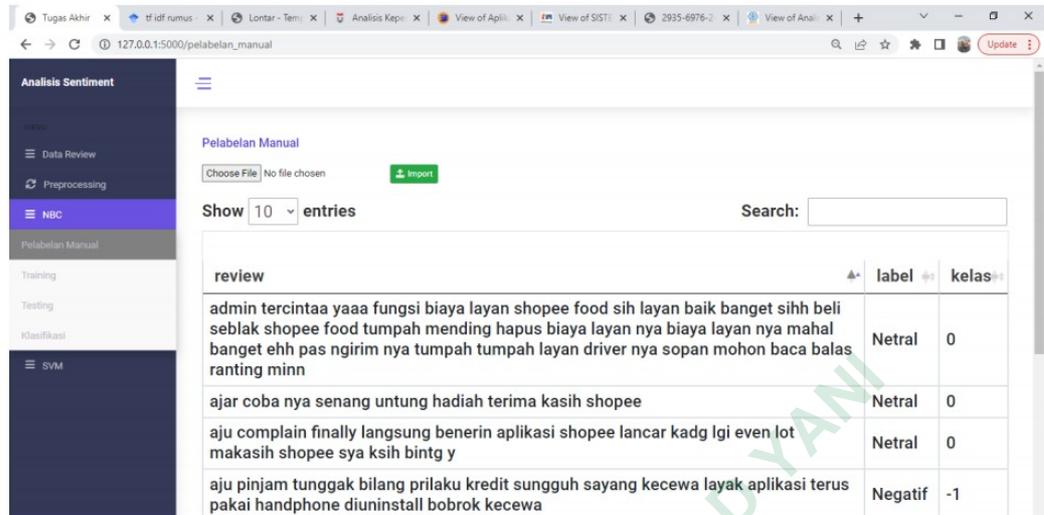


Gambar 4.3 Implementasi Halaman *Preprocessing*

Halaman *preprocessing* terdapat *button preprocessing* yang digunakan untuk memulai proses *preprocessing* yang kemudian hasil dari *preprocessing* akan menampilkan *content*.

4.6.3 HALAMAN PELABELAN MANUAL

Pada halaman pelabelan manual digunakan untuk mengupload data yang telah diberi label yang nantinya akan digunakan untuk proses *training* data. Pada halaman ini juga ditampilkan data pelabelan manual yang telah diupload. Tampilan halaman pelabelan manual terdapat pada Gambar 4.4.

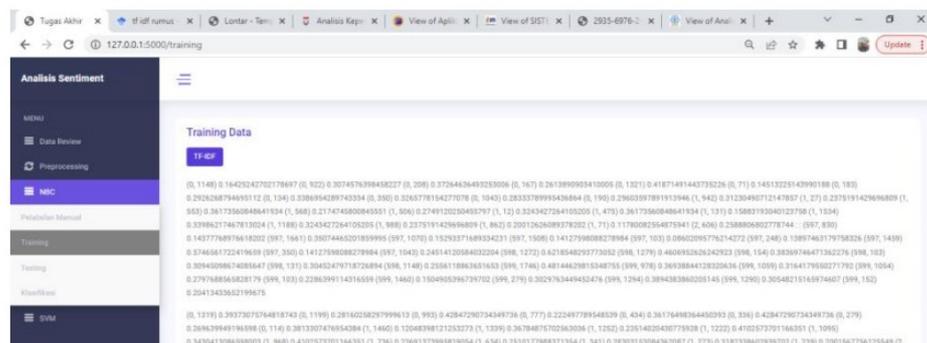


Gambar 4.4 Implementasi Halaman Pelabelan Manual

Pada halaman pelabelan manual terdapat fitur choose file yang digunakan untuk import file data yang sudah dilakukan pelabelan pada Microsoft Excel. Pada pelabelan manual ditampilkan *review* yang merupakan kalimat-kalimat ulasan sedangkan label dan kelas merupakan hasil pelabelan manual.

4.6.4 HALAMAN TRAINING

Pada halaman *training* digunakan untuk membuat model klasifikasi menggunakan metode NBC dan SVM dengan fitur ekstraksi menggunakan *term frequency invers document frequency* (TF-IDF) untuk menganalisis sentimen secara otomatis. Tampilan halaman *training* NBC terdapat pada Gambar 4.5.



Gambar 4.5 Implementasi Halaman Training

Pada halaman training terdapat *button* yang digunakan untuk membuat fitur ekstraksi menggunakan TF-IDF yang kemudian hasilnya akan ditampilkan pada sistem.

4.6.5 HALAMAN TESTING

Halaman *testing* ini digunakan untuk menampilkan *accuracy*, *precision*, *recall* dan *f1-score* dari data yang telah diprediksi kelas dan labelnya secara otomatis. Proses *testing* digunakan untuk menentukan akurat model *training* data yang sudah dibangun. Tampilan halaman *testing* terdapat pada Gambar 4.6.

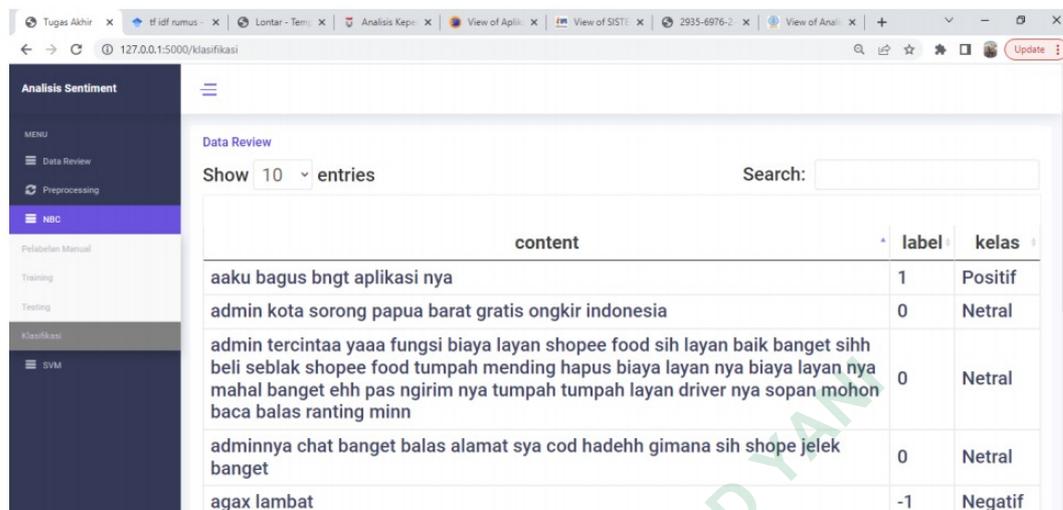
review	label	kelas
agax lambat	-1	Negatif
aju complain finally langsung benerin aplikasi shopee lancar kadg lgi even lot makasih shopee sya ksh bintang y	0	Netral
aktif spaylater gagal id da data isi lengkap kecewa	-1	Negatif
akun member gold turun silver belanja juta an	0	Netral

Gambar 4.6 Implementasi Halaman *Testing*

Pada halaman testing menampilkan hasil dari proses testing yaitu *accuracy*, *precision*, *recall* dan *f1-score* yang berasal dari perhitungan 450 data yang sudah diberi label.

4.6.6 HALAMAN KLASIFIKASI

Pada halaman klasifikasi digunakan untuk menampilkan data *review* yang telah diberi kelas dan label secara otomatis dari pembuatan model pada proses *training*. Berikut implementasi dari halaman klasifikasi terdapat pada Gambar 4.7.



content	label	kelas
aaku bagus bngt aplikasi nya	1	Positif
admin kota sorong papua barat gratis ongkir indonesia	0	Netral
admin tercintaa yaaa fungsi biaya layanan shopee food sih layanan baik banget sihh beli seblak shopee food tumpah mending hapus biaya layanan nya biaya layanan nya mahal banget ehh pas ngirim nya tumpah tumpah layanan driver nya sopan mohon baca balas ranting minn	0	Netral
adminnya chat banget balas alamat saya cod hadehh gimana sih shope jelek banget	0	Netral
agax lambat	-1	Negatif

Gambar 4.7 Implementasi Halaman Klasifikasi

Gambar 4.7 menampilkan hasil klasifikasi dimana content merupakan data review yang telah dilakukan proses pada tahapan sebelumnya, kemudian terdapat label dan kelas yang merupakan hasil klasifikasi.

4.7 PEMBAHASAN

Sistem analisis sentimen pengguna Shopee pada media Google Play Store dibangun menggunakan bahasa pemrograman Python dengan menggunakan *framework* Flask. Analisis sentimen pada penelitian ini memiliki beberapa *menu* yaitu data *review* yang digunakan untuk mengupload data yang telah diambil berbentuk CSV, selanjutnya terdapat *menu preprocessing* yang digunakan untuk membersihkan data sehingga dapat menjadi data yang terstruktur, selanjutnya terdapat *menu* pelabelan manual untuk metode NBC dan juga SVM yang digunakan untuk menampilkan data *review* yang telah dilakukan *labelling* secara manual. Kemudian terdapat *menu training* untuk metode NBC dan SVM yang digunakan untuk membangun model klasifikasi sentimen secara otomatis dengan fungsi ekstraksi menggunakan *term frequency invers document frequency* (TF-IDF) pada data teks, *menu testing* yang digunakan untuk menampilkan hasil akurasi dari data yang diprediksi secara otomatis untuk mengetahui seberapa akurat model klasifikasi yang telah dibangun menggunakan metode NBC dan

SVM. Kemudian terdapat menu klasifikasi yang digunakan untuk menampilkan hasil klasifikasi dari metode NBC dan SVM.

Analisis sentimen pada penelitian ini yang mengangkat topik Shopee merupakan analisis yang digunakan untuk perbandingan metode Naïve Bayes Classifier (NBC) dengan Support Vector Machine (SVM) dengan menggunakan data *review* Shopee di media Google Play Store. Data yang digunakan untuk analisis ini yaitu 3000 data yang sudah dilakukan *preprocessing*. Data *review* sebanyak 600 data dengan masing-masing 20 data positif, 200 data negatif dan 200 data netral yang menghasilkan model klasifikasi yang kemudian dilakukan evaluasi dengan 450 data uji yang sudah terlabeli secara manual. Hasil evaluasi menunjukkan bahwa tingkat akurasi mencapai 75% untuk algoritma NBC sedangkan untuk algoritma SVM mencapai 80%. Dari hasil akurasi yang didapat tersebut dapat ditentukan bahwa SVM merupakan algoritma terbaik untuk diterapkan pada analisis sentimen *review* pengguna Shopee. Selanjutnya dilakukan klasifikasi sentimen dengan menggunakan 3309 data yang menghasilkan sentiment positif sebanyak 1194, sentimen netral 1423 dan 692 sentimen negatif.