

BAB 4

HASIL PENELITIAN

4.1 RINGKASAN HASIL PENELITIAN

Penelitian ini melakukan analisis sentimen terhadap layanan internet Biznet berdasarkan data Twitter dengan menggunakan metode Naive Bayes Classifier. Data dikumpulkan dari 1 Januari hingga 26 Juni 2024 menggunakan *tools tweet-harvest*, menghasilkan 5.545 tweet dalam format CSV. Setelah *preprocessing*, dilakukan pelabelan manual yang menghasilkan 720 data, kemudian data dibagi menjadi data *training* dengan jumlah 576 dan data *testing* berjumlah 144 data. Hasil pengujian model menunjukkan *accuracy* 84,03%, *precision* 84,16%, *recall* 84,03%, dan *f1-score* 83,77%. Dengan menggunakan model yang sudah dibangun, dapat dilakukan klasifikasi dari keseluruhan data dengan total 5.545 data *tweet* dan model berhasil mengklasifikasikan 2.483 *tweet* sebagai sentimen negatif, 1.604 sebagai positif, dan 1.458 sebagai netral.

4.2 HASIL EVALUASI MODEL

Untuk mengukur kinerja dan akurasi dari suatu model dalam memprediksi atau mengklasifikasikan data dilakukan proses evaluasi terhadap model. Tujuan utamanya adalah untuk mengetahui seberapa baik model tersebut dapat melakukan prediksi yang akurat terhadap data baru yang belum pernah diketahui sebelumnya. Evaluasi model ini penting karena membantu dalam memahami seberapa efektif model dalam menangkap pola yang ada dalam sebuah data.

Pada penelitian ini, evaluasi model menggunakan *confusion matrix* adalah cara yang efektif untuk memahami kinerja model dalam mengklasifikasikan data. *Confusion matrix* merupakan sebuah table yang menggambarkan performa model klasifikasi dengan membandingkan prediksi yang dilakukan oleh model terhadap label sebenarnya dari data uji (data *testing*). Berikut merupakan hasil pengujian terhadap data *testing* ditunjukkan pada Tabel 4.1.

Tabel 4.1 Hasil pengujian data *testing*.

<i>Accuracy</i>	84,03%
<i>Precision</i>	84,16%
<i>Recall</i>	84,03%
<i>F1-Score</i>	83,77%

Tabel 4.1 menunjukkan bahwa model memiliki performa yang cukup baik berdasarkan *accuracy* yang diperoleh. Selanjutnya gambaran tentang bagaimana prediksi model dibandingkan dengan hasil sebenarnya dapat disajikan dalam *confusion matrix* seperti pada Tabel 4.2.

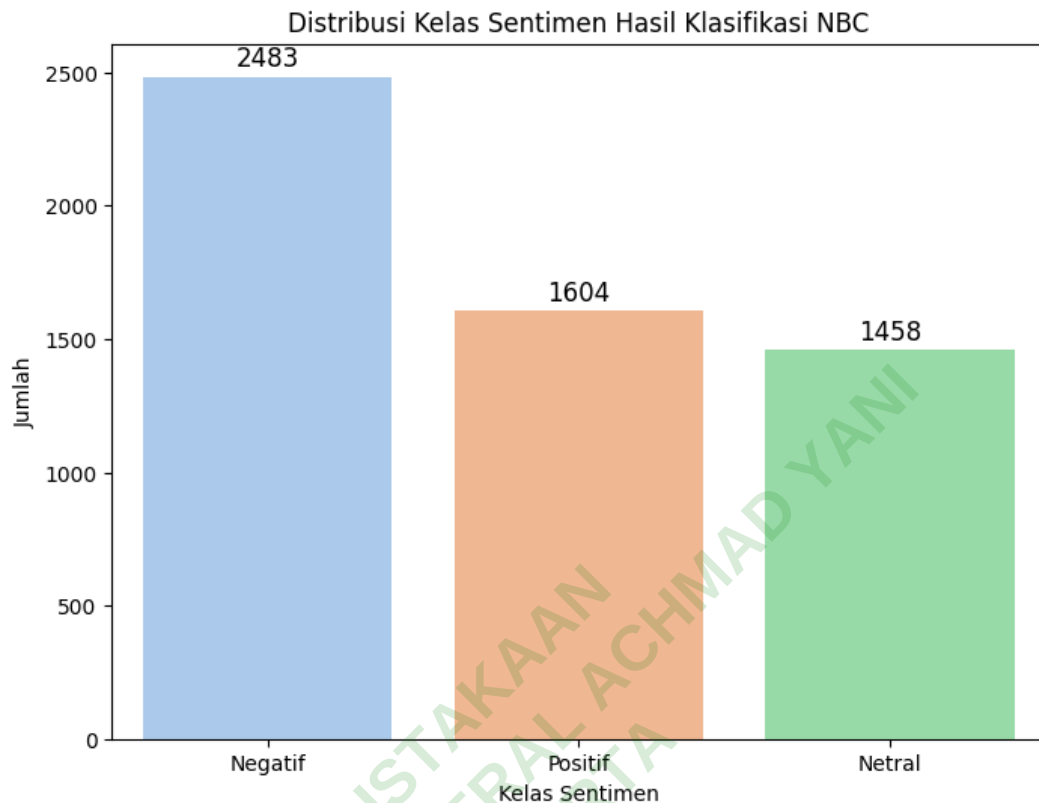
Tabel 4.2 Hasil Confusion Matrix Testing

<i>Actual</i>	<i>Prediction</i>		
	Positif	Negatif	Netral
Positif	54	0	2
Negatif	8	35	5
Netral	5	3	32

Tabel 4.2 menunjukan bahwa hasil *confusion matrik* dengan nilai TP = 54, FPN = 0, FPN_{Net} = 2, FNP = 8, TN = 35, FN_{Net} = 5, FNetP = 5, FN_{Net} = 3, TN = 32. *Confusion matrix* pada data *testing* berukuran 3x3 karena pada penelitian ini menggunakan 3 kelas untuk pengklasifikasian sentimen yaitu kategori kelas positif, negatif serta netral. Beberapa istilah yang digunakan pada *confusion matrix* ini diantaranya adalah TP (*True Positif*), FPN (*False Positif Negatif*), FPN_{Net} (*False Positif Netral*), FNP (*False Negatif Positif*), TN (*True Negatif*), FN_{Net} (*False Negatif Netral*), FNetP (*False Netral Positif*), FNetN (*False Netral Negatif*), Tnet (*True Netral*).

4.3 HASIL KLASIFIKASI

Klasifikasi dilakukan terhadap keseluruhan data *tweet* dengan jumlah 5.545 data yang sebelumnya belum diketahui kelas dan labelnya menggunakan model yang sudah dibuat sebelumnya. Berikut merupakan hasil klasifikasi yang divisualisasikan dengan menggunakan barplot ditunjukkan pada Gambar 4.1.



Gambar 4.1 Barplot Hasil Klasifikasi

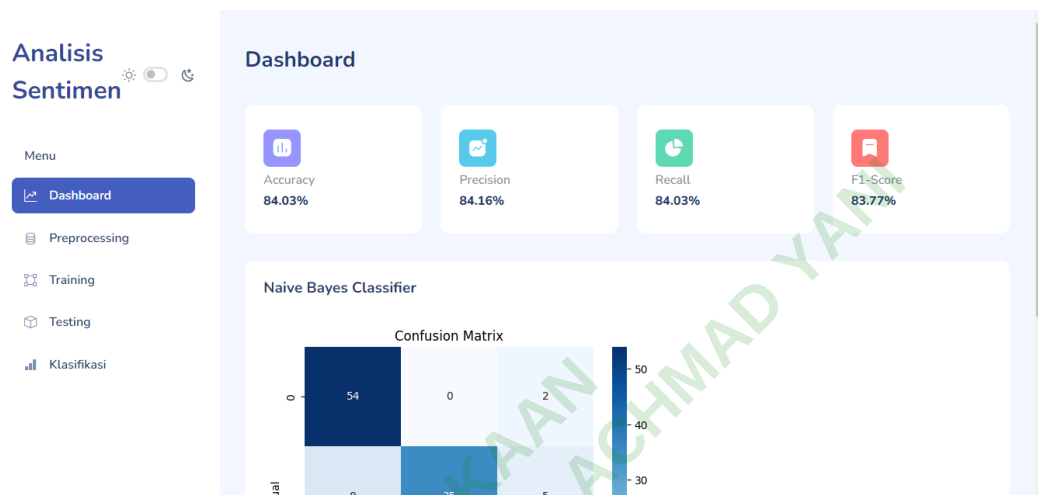
Berdasarkan Gambar 4.1 dapat diketahui bahwa hasil klasifikasi menunjukkan 2483 data memiliki sentimen negatif, 1604 data memiliki sentimen positif dan 1458 data memiliki sentimen netral. Berdasarkan klasifikasi tersebut artinya sentimen pengguna Twitter terhadap layanan internet Biznet memiliki hasil sentimen negatif. Sebagian besar sentimen negatif menunjukkan adanya keluhan signifikan dari pengguna terhadap layanan internet Biznet, terutama terkait gangguan dan kestabilan koneksi.

4.4 IMPLEMENTASI DASHBOARD

Hasil analisis yang dilakukan selanjutnya ditampilkan dalam sebuah dashboard yang dibangun dengan menggunakan Bahasa *Pemrograman Python* serta *Framework Flask*. Berikut merupakan beberapa halaman yang merupakan implementasi dalam hasil analisis sentimen ini.

4.4.1 Halaman Dashboard

Pada halaman dashboard berisi tentang informasi nilai dari *Accuracy*, nilai *Precision*, nilai *Recall* dan nilai *F1-Score* serta hasil *confusion matrix* dari testing. Tampilan halaman dashboard ditunjukkan pada Gambar 4.2.



Gambar 4.2 Halaman Dashboard

4.4.2 Halaman Preprocessing

Halaman *preprocessing* digunakan untuk menampilkan data yang sudah dilakukan preprocessing sebelumnya, tampilan halaman *preprocessing* ditunjukkan pada Gambar 4.3.

Unnamed: 0	full_text
0	kenapa provider yang speed nya oke kayak myrepublic sama biznet tidak pasang daerah coy butuh bgtt
1	biznet paling mending roket ungu kalau daerah waras kenceng cuman kalau alas jangkau pakai startlink ae
2	sepertinya sutet wifi biznet kamu dien
3	memang deh wifi biznet rumah juara banget cepat ketimbang yang kostan
4	biznet padahal sudah paling oke turut coba iconnet thur
5	aman tah pip gawe biznet ping e tidak stabil anyinggg
6	provider wifi oke biznet apa kawan
7	lah biznet kenapa dah
8	ganti biznet palagan tercover koyoke
9	kenapa biznet slow banget kayaknya Biznet memang lambat banget

Gambar 4.3 Halaman Preprocessing

4.4.3 Halaman Training

Halaman *training* menampilkan data yang digunakan untuk *training* dan sudah diberi kelas dan label sebelumnya. Tampilan halaman *training* ditunjukkan pada Gambar 4.4.



full_text	kelas	label
min bintaro jakarta selatan internet disconnect lalu connect mei ganggu kali juni ganggu lagiburuk sekali kualitas layanan biznet baru sy pakai tahun lebih bulan bikin emosi layanan internet biznetkita bayar min klu free putus nyambung wajar gratis	Negatif	-1
biznet bukan era kamu	Negatif	-1
sumpah biznet sih	Negatif	-1
biznet mki mulai tantrumnya	Negatif	-1
kak biznet ganggu apa gimana koneksi kamu s	Negatif	-1
oi biznet lelet kali kau punya jaring	Negatif	-1
iya kayak biznet lama	Negatif	-1
emosi lama lama main bareng sama orang pakai biznet	Negatif	-1

Gambar 4.4 Halaman Training

4.4.4 Halaman Testing

Halaman *testing* menampilkan data *testing* yang digunakan dalam proses analisis sentimen. Tampilan dari halaman *testing* ditunjukkan pada Gambar 4.5.



full_text	label
seperti biznet	0
info dong ada layanan biznet hari wireless maupun wired	0
pake biznet tidak sumpah sinyal jelek banget bikin emosi	-1
biznet down semua discord tidak akses sama sekali	-1
biznet oke oxygen juga oke	0
lah kenapa jadi ngurusin tugas customer care indihome oxygen biznet	-1
min kemarin down minggu tidak kompensasi biznet tagih nya sama masa aktif nya tidak tambah	-1
duh kontrak saya mas sampe akhir juni sehat sehat mental saya sudah sama ini provider asa indihome tidak parah bahkan daerah saya tidak pernah down salah fatal w milih biznet	-1
nyata biznet	0

Gambar 4.5 Halaman Testing

4.4.5 Halaman Klasifikasi

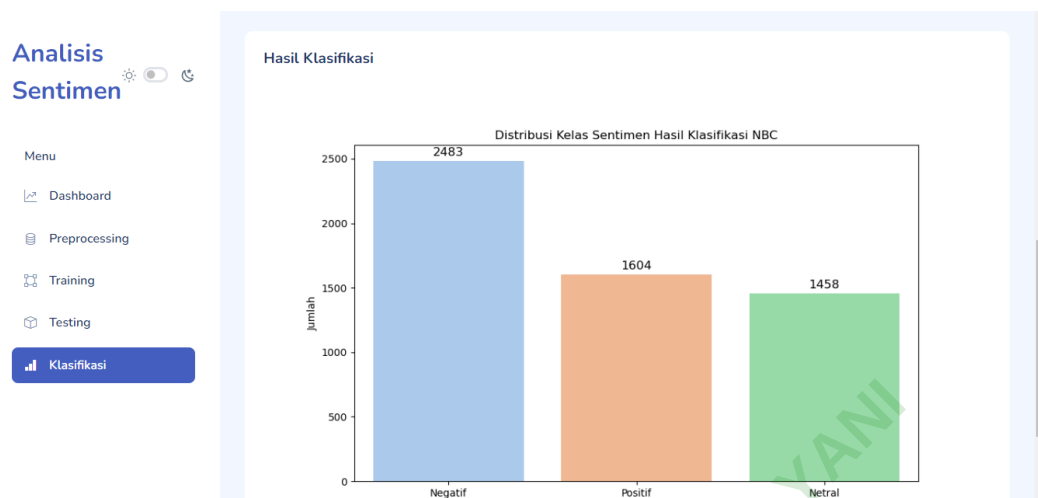
Pada halaman klasifikasi digunakan untuk melakukan proses klasifikasi data secara otomatis oleh sistem dengan cara melakukan *upload* data maka sistem akan menampilkan contoh hasil data klasifikasi dan barplot berdasarkan data yang di *upload*. Tampilan halaman klasifikasi ditunjukkan pada Gambar 4.6.



Gambar 4.6 Halaman Upload Data Klasifikasi



Gambar 4.7 Halaman Contoh Data Hasil Klasifikasi



Gambar 4.8 Halaman Visualisasi Data Hasil Klasifikasi

4.5 PEMBAHASAN

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap layanan internet Biznet berdasarkan data dari Twitter. Dengan menerapkan metode Naive Bayes Classifier, dilakukan penelitian untuk mengelompokkan data *tweet* ke dalam kategori sentimen positif, negatif, dan netral. Proses pengumpulan data dilakukan menggunakan *tools tweet-harvest* pada *Google Colaboratory*, dengan kata kunci "Biznet" dan "Internet Biznet". Data yang diambil mencakup *tweet* dan komentar dari 1 Januari 2024 hingga 26 Juni 2024. Data yang dihasilkan berupa file dengan format CSV dengan jumlah data yang diperoleh adalah 5545 data *tweet*. Pengolahan data atau *preprocessing* melibatkan beberapa tahapan yang penting untuk mempersiapkan serta membersihkan data sebelum dilakukan analisis. Tahapan ini meliputi *cleaning*, *case folding*, *number removal*, *punctuation removal*, *whitespaces removal*, *tokenizing*, *stopword removal (filtering)*, *stemming*, dan *normalization*. Hasil dari setiap tahapan *preprocessing* menunjukkan bahwa data teks menjadi lebih terstruktur dan siap untuk digunakan dalam analisis sentimen.

Setelah melalui tahap *preprocessing*, dilakukan pelabelan manual dan menghasilkan data yang berjumlah 720 data. Data tersebut kemudian dibagi menjadi data *training* dan data *test* dengan proporsi 80:20 sehingga data *training* berjumlah 576 dan data *test* yang berjumlah 144 data. Selanjutnya, dilakukan ekstraksi fitur menggunakan *TfidfVectorizer* untuk mengubah teks menjadi

representasi *numerik* agar dapat diolah. Selanjutnya model klasifikasi dibangun dengan menerapkan metode Naive Bayes Classifier berdasarkan data *training* yang sudah diberi kelas dan label sebelumnya. Tahap pengujian dilakukan dengan mengukur nilai *accuracy*, nilai *precision*, nilai *recall*, dan nilai *f1-score* dari model yang telah dibangun. Hasil dari pengujian menunjukkan bahwa model memiliki *accuracy* sebesar 84,03%, *precision* sebesar 84,16%, *recall* sebesar 84,03%, dan *f1-score* sebesar 83,77%. Dari 5545 data *tweet* yang dikumpulkan, model yang dibangun berhasil mengklasifikasikan 2483 data *tweet* sebagai sentimen negatif, 1604 data *tweet* sebagai sentimen positif, dan 1458 data *tweet* sebagai sentimen netral. Hasil sentimen menunjukkan bahwa sebagian besar pengguna memberikan tanggapan negatif terhadap layanan internet Biznet.