

BAB 4

HASIL PENELITIAN

4.1 RINGKASAN HASIL PENELITIAN

Ringkasan hasil penelitian ini yaitu pengambilan data *tweet* dari Twitter menggunakan Jupyter Notebook dengan keyword *bullying* menggunakan library *snsrscrape* selama Mei 2023 hingga Juni 2023 yang berjumlah jumlah sebanyak 13.000 data. Menggunakan data mentah sebanyak 966 data dari 13000 data mentah yang diperoleh. Setelah data mentah diperoleh dilanjutkan dengan tahap Preprocessing untuk membersihkan data dan dilabeli secara manual. Data yang digunakan pada proses training sebanyak 772 data *tweet*. Sedangkan untuk data testing yang digunakan sebanyak 194 *tweet* dengan rincian 395 data positif, dan 571 data negatif. Dari data training ini dilakukan pengujian, dan menghasilkan nilai akurasi sebesar 59,79%, sedangkan akurasi tahap testing sebesar 68,55%. Tahapan dalam penelitian ini adalah pengambilan data, preprocessing, training data, TF-IDF, testing data, dan klasifikasi.

1.2 AKURASI ALGORITMA

Pada penelitian ini menggunakan data training sebanyak 772 data, dengan rincian 469 data berlabel negatif dan 303 data berlabel positif. Hasil pada perhitungan *confusion matrix* pada data training dapat dilihat pada Tabel 4.1

Tabel 4.1 Confusion Matrix Data Training

Kelas Pediksi	Kelas Aktual	
	Positif	Negatif
Positif	192	194
Negatif	190	196

Pada Tabel 4.1 diatas menampilkan nilai TP sebesar 192, nilai FN sebesar 194, nilai FP sebesar 190, dan nilai TN sebesar 196. Hasil *accuracy* dan *f1-score* pada data *training* dapat dilihat pada Tabel 4.2

Tabel 4.2 Akurasi Data Training

Akurasi Data <i>Training</i>	Hasil
<i>Accuracy</i>	59,79%
<i>F1-Score</i>	59,94%

Pada tabel 4.2 diatas menampilkan tabel hasil perhitungan *confusion matrix* data *training* yang menghasilkan nilai *accuracy* sebesar 59,79% dan nilai *f-measure* bernilai 59,94%.

Tahap selanjutnya yaitu tahap testing, dengan menggunakan data *testing* sebanyak 194 data *tweet* dengan rincian 82 data berlabel positif, dan 112 data berlabel negatif. Berikut merupakan hasil dari *confusion matrix* data testing dapat dilihat pada Tabel 4.3

Tabel 4.3 *Confusion Matrix Data Testing*

Kelas Pediksi	Kelas Aktual	
	Positif	Negatif
Positif	92	17
Negatif	25	58

Pada tabel 4.3 diatas menghasilkan nilai TP sebesar 92, nilai FN sebesar 17, nilai FP sebesar 25, nilai TN sebesar 58. Nilai akurasi dari data testing dapat dilihat pada Tabel 4.4

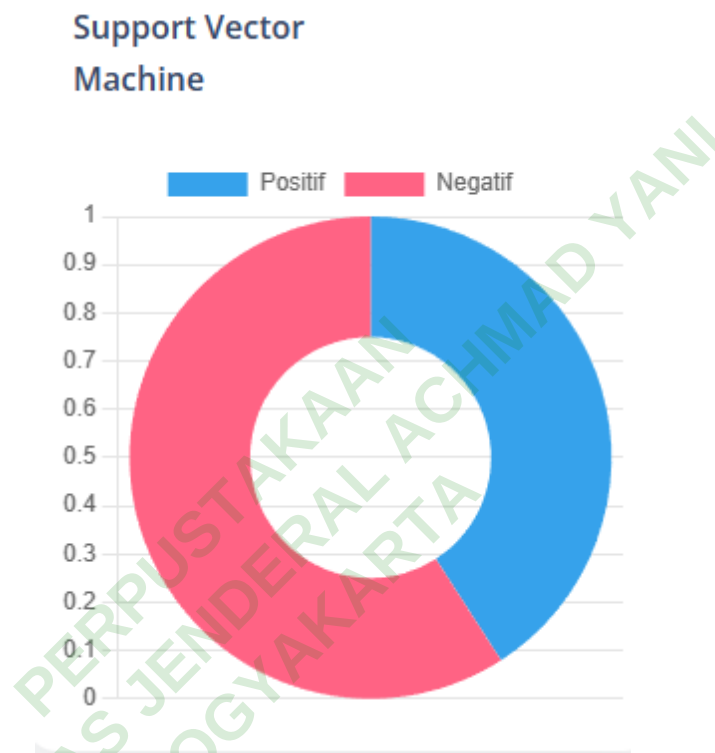
Tabel 4.4 Akurasi Data Testing

Akurasi Data Testing	Hasil
<i>Accuracy</i>	68,55%
<i>F1-Score</i>	67,96%

Pada tabel 4.4 diatas menampilkan Tabel akurasi perhitungan pada data testing, menghasilkan nilai *accuracy* sebesar 68,55% dan nilai *f-measure* sebesar 67,96%.

4.3 EVALUASI DAN KLASIFIKASI

Hasil dari klasifikasi ini dibagi menjadi dua sentimen, yaitu sentiment positif dan sentimen negatif. Grafik hasil dari keseluruhan data dapat dilihat pada grafik lingkaran pada Gambar 4.1



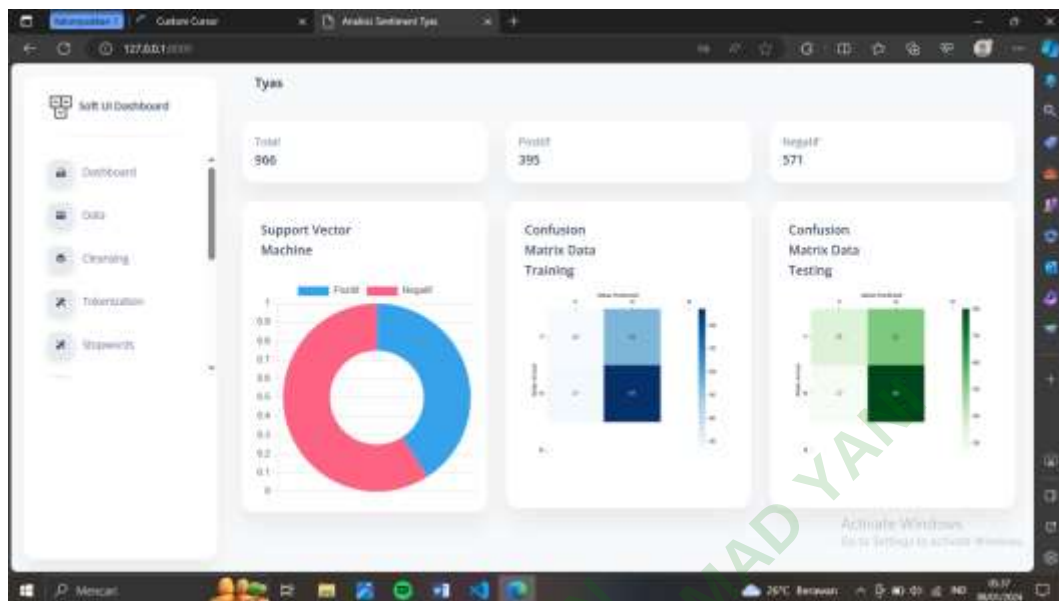
Gambar 4.1 Grafik klasifikasi sentimen

Pada Gambar 4.1 diatas menampilkan grafik hasil klasifikasi sentimen dengan hasil data *tweet* Negatif berjumlah 571, dan jumlah data sentimen Positif adalah 395.

4.4 FITUR-FITUR SISTEM

4.4.1 Fitur Dashboard

Fitur Dashboard terdapat halaman Dasboard yang berisi Support Vector Machine, Confusion Matrix Data Training dan Confusion Matrix Data Testing. Tampilan halaman ini dapat dilihat pada Gambar 4.2



Gambar 4.2 Tampilan halaman Dashboard

Pada Gambar 4.2 diatas menampilkan halaman dashboard yang menampilkan total data sentimen sebesar 986, jumlah data sentimen positif 588, dan jumlah data sentimen negatif adalah 398.

4.4.2 Fitur Data

Fitur Data terdapat halaman yang berisi sebuah data – data yang telah diambil dari twitter yang dapat dilihat pada Gambar 4.3

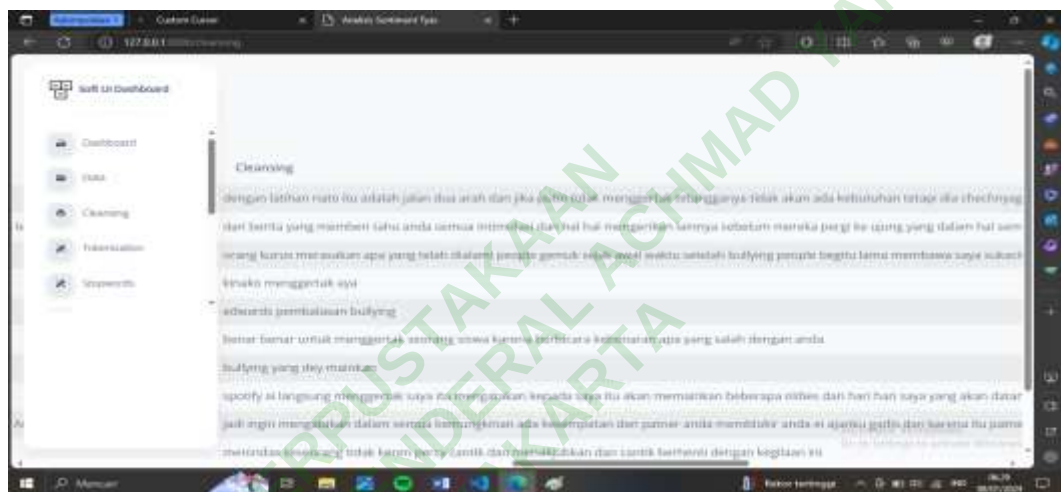
No.	Label	Content
1	1	Dengan latihan NATO itu adalah jalan dua arah dan jika Putin tidak menggentak entanggganya tidak akan ada ke
2	-1	Dari berita yang memberi tahu Anda semua informasi dan hal-hal mengenai latanya satupun mereka pergi b
3	-1	Orang kurus melakukan apa yang telah dialami ppl gimak sejak awal waktu setelah bullying ppl begitu lama m
4	-1	Rinako menggentak Apa
5	-1	Edward: Pembulatan bullying
6	1	benar-benar untuk menggentak seorang siswa karena kebencian kewanaran apa yang kalah dengan Anda
7	-1	Bullying yang day marikan
8	1	spotify akan langsung menggentak saya itu mengatakan kepada saya itu akan melakukan beberapa siklus dari ha
9	1	jadi ingin mengatakan dalam semua kemungkinan ada kesempatan dan patner Anda membantu Anda di agensi
10	1	Meminta seseorang tidak keren, Percy cantik dan menakutkan dan cantik, bermain dengan kegiatan m

Gambar 4.3 Tampilan halaman Data

Pada Gambar 4.3 diatas terdapat Pengambilan 966 data dari Twitter namun pada gambar di atas hanya dapat menampilkan 10 data saja.

4.4.3 Fitur Cleansing

Fitur Cleansing sebuah halaman yang melakukan pembersihan data yang telah diunggah lebih dari satu kali. Tampilan Cleansing dapat dilihat pada Gambar 4.4

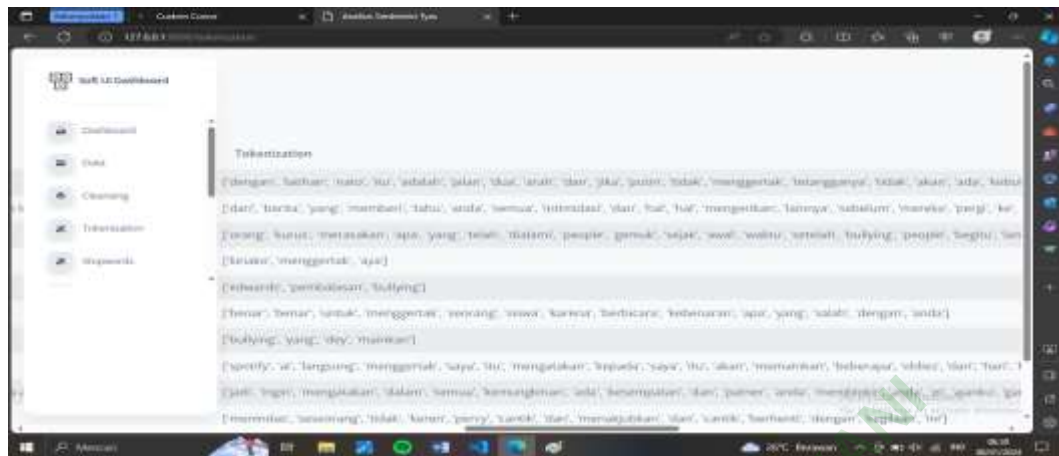


Gambar 4.4 Tampilan halaman Cleansing

Pada Gambar 4.4 menampilkan halaman Cleansing yang berisikan data yang akan digunakan untuk tahap Tokenizing.

4.4.4 Fitur Tokenizing

Fitur Tokenizing Tahap tokenizing berfungsi untuk menampilkan data berupa kata yang telah pecahkan atau melakukan *root* yang dapat dilihat pada Gambar 4.5

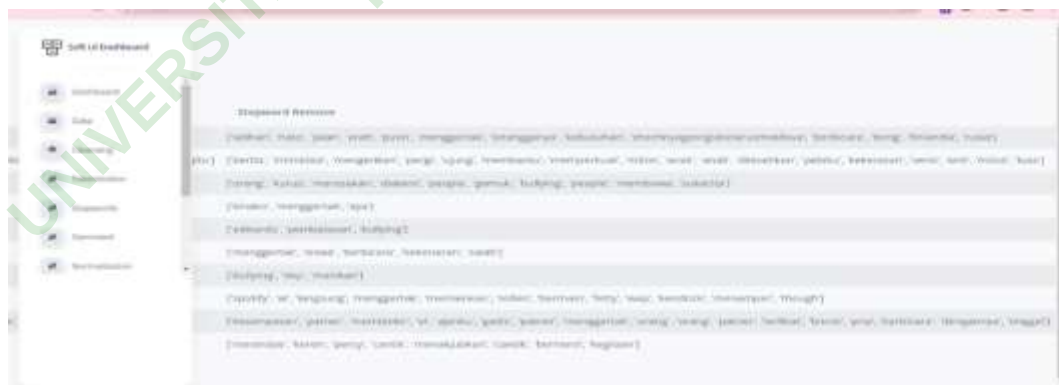


Gambar 4.5 Tampilan halaman Tokenizing

Pada Gambar 4.5 atas terdapat halaman Tokenizing yang menampilkan data tweet atau sebuah kalimat yang telah dipecahkan menjadi sebuah kata.

4.4.5 Fitur Stopword Removal

Fitur Stopword Removal fitur ini berfungsi untuk penghilang kata henti menghapus kata-kata yang tidak memiliki arti. Tampilan halaman ini dapat dilihat pada Gambar 4.6

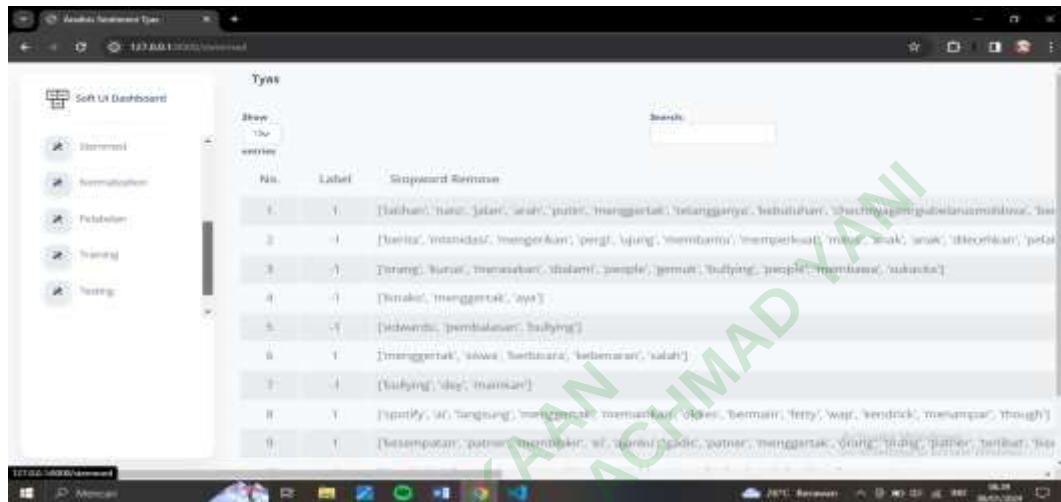


Gambar 4.6 Tampilan halaman Stopword Removal

Pada Gambar 4.6 menampilkan halaman Stopword Removal yang berisi data yang tidak memiliki arti akan dihapus secara otomatis.

4.4.6 Fitur Stemming

Fitur Stemming pada tahap ini berfungsi untuk menampilkan data yang telah di ubah menjadi kata dasar.



Gambar 4.7 Tampilan halaman Stemming

Gambar 4.8 menampilkan data hasil proses *stemming*, setelah itu data akan dilabeli lalu kemudian dibagi menjadi data training dan data testing.

4.4.7 Fitur Pelabelan

Fitur Pelabelan pada halaman ini fitur pelabelan berfungsi untuk melakukan pelabelan pada data yang dapat di lihat pada Gambar 4.8

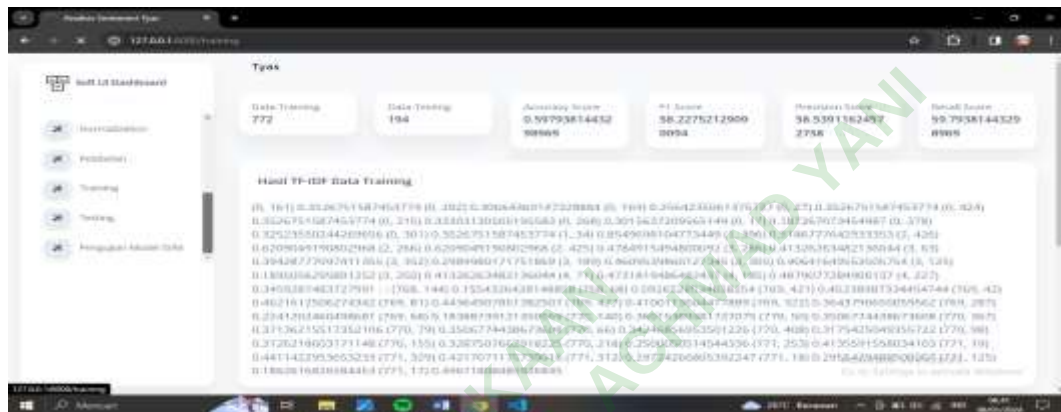


Gambar 4.8 Tampilan halaman Pelabelan

Pada Gambar 4.8 diatas terdapat sebuah halaman yang menampilkan halaman untuk melakukan pelabelan data yang akan digunakan proses data training dan testing.

4.4.8 Fitur Training Data

Fitur Training data adalah halaman yang berisikan hasil data TF-IDF yang telah diambil. Tampilan Training Data dapat di lihat pada Gambar 4.9

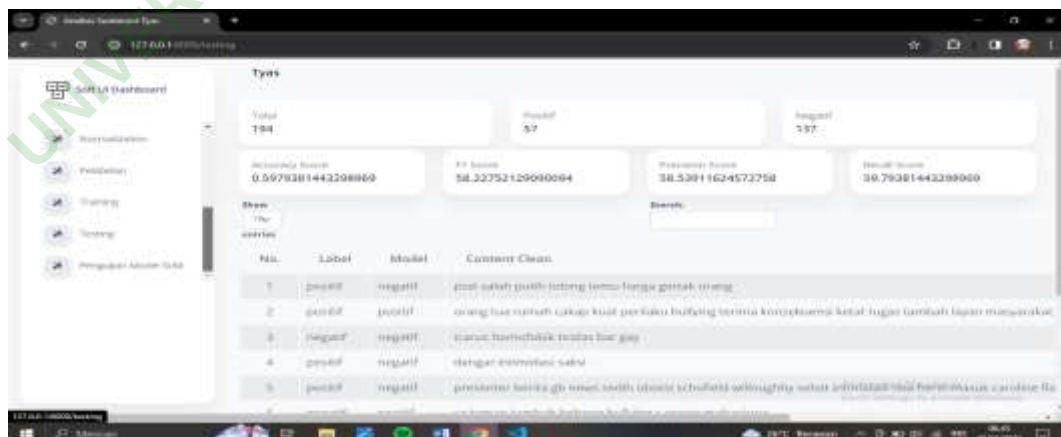


Gambar 4.9 Tampilan halaman Training Data

Pada Gambar 4.9 diatas terdapat sebuah tampilan hasil dari training data yang meliputi hasil data trainingnya 772.

4.4.9 Fitur Testing Data

Fitur Testing data yaitu sebuah halaman yang menampilkan halaman hasil TF-IDF data testing. Tampilan testing data dapat dilihat pada Gambar 4.10

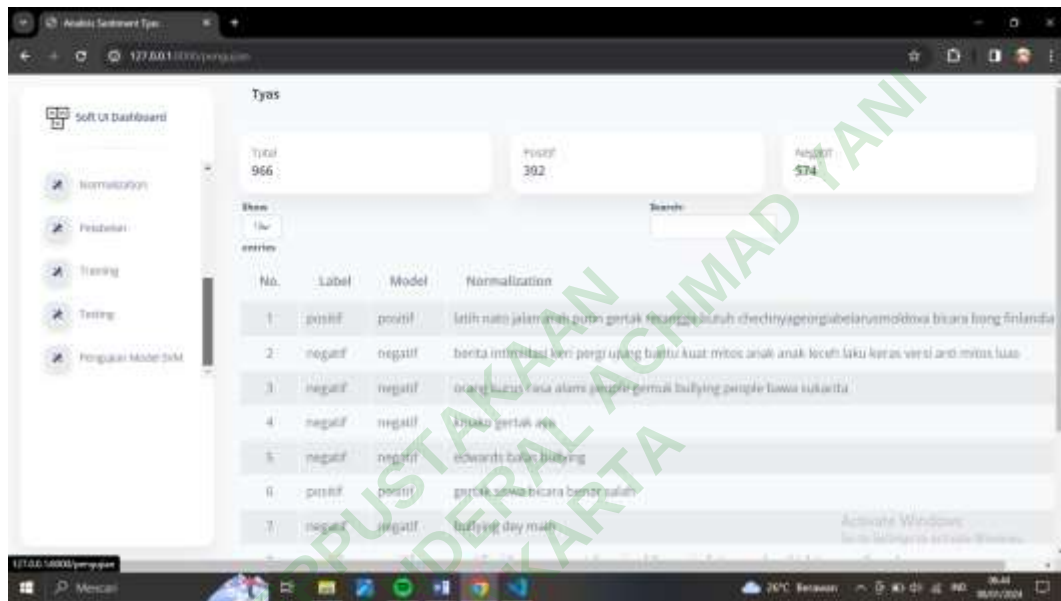


Gambar 4.10 Tampilan Hasil Testing Data

Pada Gambar 4.10 berisi data testing berjumlah 194 data label, dengan 137 sentimen Negatif dan 57 sentimen Positif.

4.4.10 Pengujian SVM Model

Fitur hasil menampilkan diagram lingkaran dari hasil perhitungan menggunakan SVM. Tampilan fitur hasil dapat dilihat pada gambar 4.11



Gambar 4.11 Tampilan Hasil Pengujian Support Verctor Machine

Pada gambar 4.11 di atas menampilkan hasil perhitungan algoritma SVM beserta akurasi data training serta data testing.

