

BAB 4

HASIL PENELITIAN

4.1 RINGKASAN HASIL PENELITIAN

Hasil dari penelitian penerapan model klasifikasi menggunakan metode Naive Bayes dan Decision Tree untuk melakukan analisis terhadap sentimen di media sosial Twitter dengan kata kunci “PLN” menggunakan data yang diambil dari tanggal 1 Mei 2022 sampai dengan 4 Juni 2022 dengan jumlah total data 40.745, baik *tweet* maupun *re-tweet*. Dalam proses *training* digunakan sebanyak 900 data dengan 300 data berlabel positif, 300 data berlabel negatif dan 300 data berlabel netral. Sedangkan untuk proses *testing* menggunakan 300 data dengan rincian 100 data berlabel positif, 100 data berlabel negatif dan 100 data berlabel netral. Hasil dari analisis ini ditampilkan kedalam sebuah *dashboard* dengan menggunakan *framework Flask*. Penelitian ini diharapkan dapat memberikan gambaran mengenai respon masyarakat mengenai pelayanan yang telah diberikan oleh PT.PLN dan juga untuk membandingkan antara metode Naive Bayes dan Decision Tree.

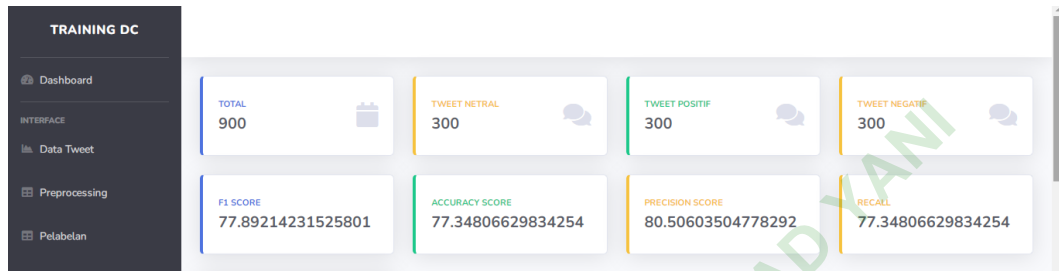
4.2 AKURASI MODEL DECISION TREE

Evaluasi model Decision Tree dilakukan dengan menggunakan data dengan jumlah 900 data dengan rincian 300 data dengan label positif, 300 data berlabel negatif dan 300 data berlabel netral, kemudian dilakukan pengujian *confusion matrix* untuk mendapatkan nilai aktual dan prediksi pada data, hasil dari *confusion matrix* dapat dilihat pada Tabel 4.1.

Tabel 4.1 *Confusion Matrix Training Decision Tree*

Kelas Aktual	Kelas Predikasi		
	positif	negatif	Netral
positif	49	1	14
negatif	1	42	16
netral	3	6	49

Dari percobaan yang telah dilakukan didapatkan bahwa akurasi dari model yang menggunakan metode Decision Tree sebesar 77 % kemudian mendapatkan nilai *precision* sebesar 80%, *recall* 77 % dan *f1-score* 77%, seperti pada Gambar 4.1.



Gambar 4.1 Akurasi *Training* Decision Tree

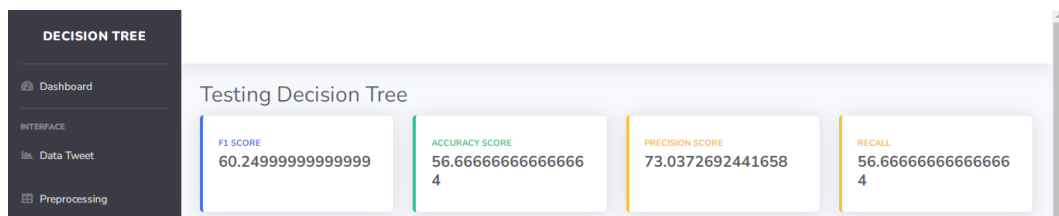
4.3 TESTING DECISION TREE

Testing yang dilakukan dengan menggunakan 300 data *testing* dibagi menjadi 100 data netral, 100 data negatif dan 100 data positif kemudian dilakukan penghitungan *confusion matrix* dan menghasilkan data seperti pada Tabel 4.2.

Tabel 4.2 *Confusion Matrix Testing* Decision Tree

Kelas Aktual	Kelas Predikasi		
	positif	negatif	Netral
positif	8	3	0
negatif	9	5	1
netral	12	1	21

menghasilkan akurasi sebesar 56 %, *precision* 73 %, *recall* 56% dan *f1-score* 60% seperti pada Gambar 4.2.



Gambar 4.2 *Testing* Decision Tree

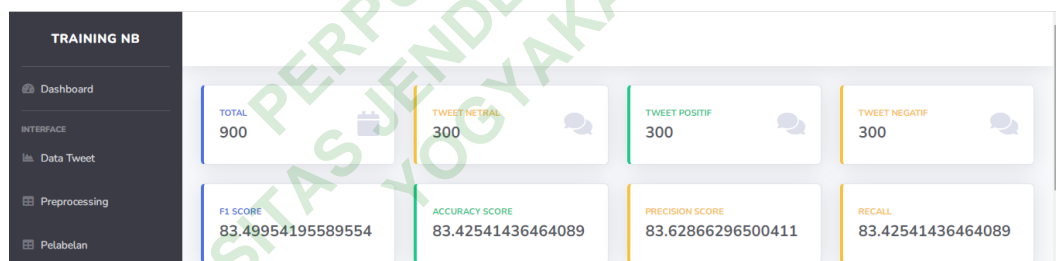
4.4 AKURASI MODEL NAÏVE BAYES CLASSIFIER

Evaluasi model dilakukan dengan menggunakan data dengan jumlah 900 data dengan rincian 300 data dengan label positif, 300 data berlabel negatif dan 300 data berlabel netral, kemudian dilakukan pengujian *confusion matrix* untuk mendapatkan nilai aktual dan prediksi pada data, hasil dari *confusion matrix* dapat dilihat pada Tabel 4.3

Tabel 4.3 *Confusion Matrix Training Naive Bayes*

Kelas Aktual	Kelas Predikasi		
	positif	negatif	Netral
positif	56	2	6
negatif	3	49	7
netral	2	10	46

Dari percobaan kedua yang telah dilakukan didapatkan bahwa akurasi dari model yang menggunakan metode Naive Bayes sebesar 83 %, *f1-score* 83 %, *Precision* 83% dan *recall* 83% seperti pada Gambar 4.3.



Gambar 4.3 Akurasi *Trainig* Naive Bayes

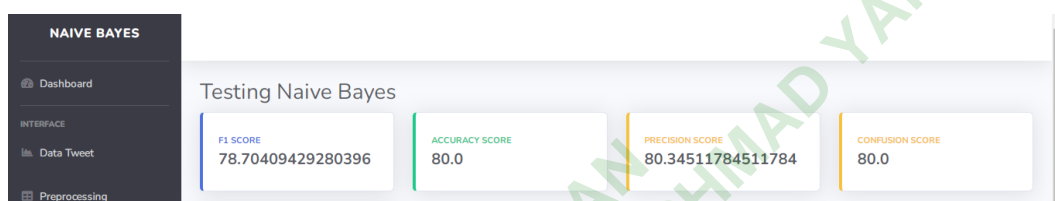
4.5 TESTING NAÏVE BAYES CLASSIFIER

Testing yang dilakukan dengan menggunakan 300 data *testing* dibagi menjadi 100 data netral, 100 data negatif dan 100 data positif dan dilakukan penghitungan *confusion matrix* dan menghasilkan data seperti pada Tabel 4.4.

Tabel 4.4 *Confusion Matrix Testing Naive Bayes*

Kelas Aktual	Kelas Predikasi		
	positif	negatif	Netral
positif	7	3	5
negatif	0	13	0
netral	2	2	28

Kemudian menghasilkan akurasi sebesar 80 %, *f1-score* 78 %, precision 80% dan recall 80% seperti pada Gambar 4.4.

**Gambar 4.4** *Testing Naive Bayes*

4.6 PERBANDINGAN HASIL AKURASI DECISION TREE DAN NAÏVE BAYES CLASSIFIER

Berdasarkan analisis yang dilakukan didapatkan bahwa akurasi metode Naive Bayes lebih tinggi dibandingkan dengan akurasi metode Decision Tree didapatkan bahwa nilai dari *training* Naive Bayes berada angka 83 % sedangkan untuk Decision Tree berada pada angka 77% dengan jumlah data *training* sebanyak 900 data dan menghasilkan perbandingan seperti pada Tabel 4.5.

Tabel 4.5 Hasil Akurasi *Training*

	Naive Bayes	Decision Tree
<i>Accuracy</i>	83%	77%

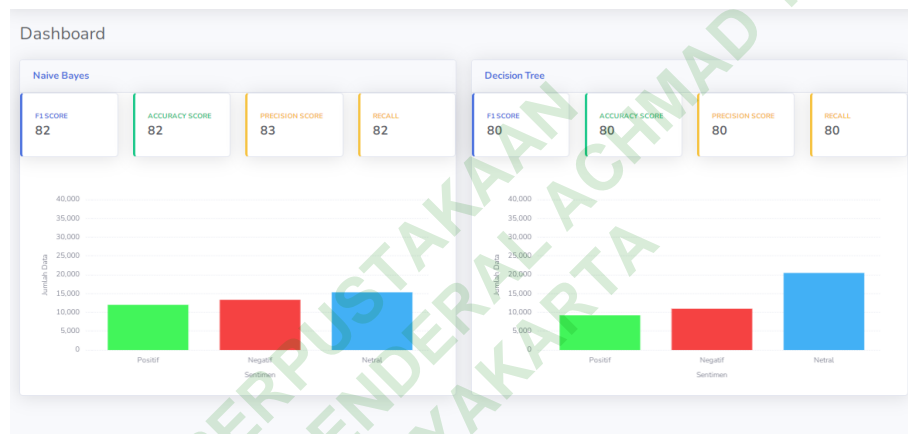
Kemudian untuk data *testing* sebanyak 300 data *tweet* dengan label positif sebanyak 100 data, negatif 100 data dan netral sebanyak 100 data sehingga menghasilkan perbandingan akurasi seperti pada Tabel 4.6

Tabel 4.6 Hasil Akurasi *Testing*

	Naive Bayes	Decision Tree
<i>Accuracy</i>	80%	56%

4.7 HASIL KLASIFIKASI TWEET

Klasifikasi dilakukan dengan menggunakan metode dengan akurasi *training* dan *testing* yang memiliki akurasi paling tinggi yaitu menggunakan metode Naive Bayes, Seperti pada Gambar 4.5.

**Gambar 4.5** Klasifikasi

Dari hasil klasifikasi dengan menggunakan 40.475 data tweet untuk metode Naive Bayes didapatkan jumlah sentimen positif sebanyak 12.029 dengan *tweet* berisi tentang apresiasi masyarakat mengenai kesiapan dalam menjaga pasokan tenaga listrik, untuk sentimen positif metode Decision Tree mendapatkan sebanyak 9.189 data *tweet* dengan pembahasan yang sama, sentimen negatif pada metode Naive Bayes mendapat 13.363 dengan pembahasan mengenai lamanya penanganan ketika ada masalah jaringan kelistrikan dan pemadaman listrik secara tiba-tiba sedangkan untuk metode Decision Tree mendapatkan sebanyak 11.097 data *tweet* dengan pembahasan yang sama, untuk data netral metode Naive Bayes mendapatkan 15.352 data *tweet* dan metode Decision Tree mendapat 20.458 data *tweet* dengan pembahasan mengenai berita dan pertanyaan pelanggan kepada admin PLN. Untuk rincian data sentimen seperti pada Tabel 4.7

Tabel 4.7 Jumlah Total Sentimen

Sentimen	Naive Bayes	Decision Tree
Positif	13.403	9.364
Negatif	14.811	16.649
Netral	12.530	14.731

Dari hasil analisis yang dilakukan sentimen positif berisi tentang pujian terhadap pelayanan pln yang berisi tentang siap siaga dalam menjaga pasokan listrik, pemberdayaan sumber tenaga yang baik, seperti pada Tabel 4.8.

Tabel 4.8 Tweet Positif

Tweet
ayo semangat memperbaiki pln pln realisasikan aturan plts
alhamdulillah terimakasih pln
hidup bbrp menit terimakasih admin respon twt
pln kucurkan dana rp miliar jamin ketersediaan listrik g bali terimakasih pln bantuannya
admin jaga kesehatan jamu bugur dm admin

Kemudian untuk sentimen negatif yang didapat membahas tentang lamananya pln dalam merespon laporan masyarakat, pemadaman listrik dalam jangka waktu yang lama tanpa adanya pemberitahuan, susahnya dalam mengurus keluhan, berikut beberapa tweet negatif yang terdapat pada Tabel 4.9.

Tabel 4.9 Tweet Negatif

Tweet
sumpahhhhhhhhhh ngeselin banget allah jam pagi smp jam set malam nyala jugaaaa
admin padam listrik area desa ciseeng kabupaten bogor tolong diperbaiki gerah anjing pln tai emg
bogor demen banget mati lampu
apaansih ajjing pln sore mati listrik mulu gada fakkkk saya emosi

Untuk Tweet yang bersifat netral berisi tentang berita dan pertanyaan pelanggan kepada admin PLN, seperti pada Tabel 4.10.

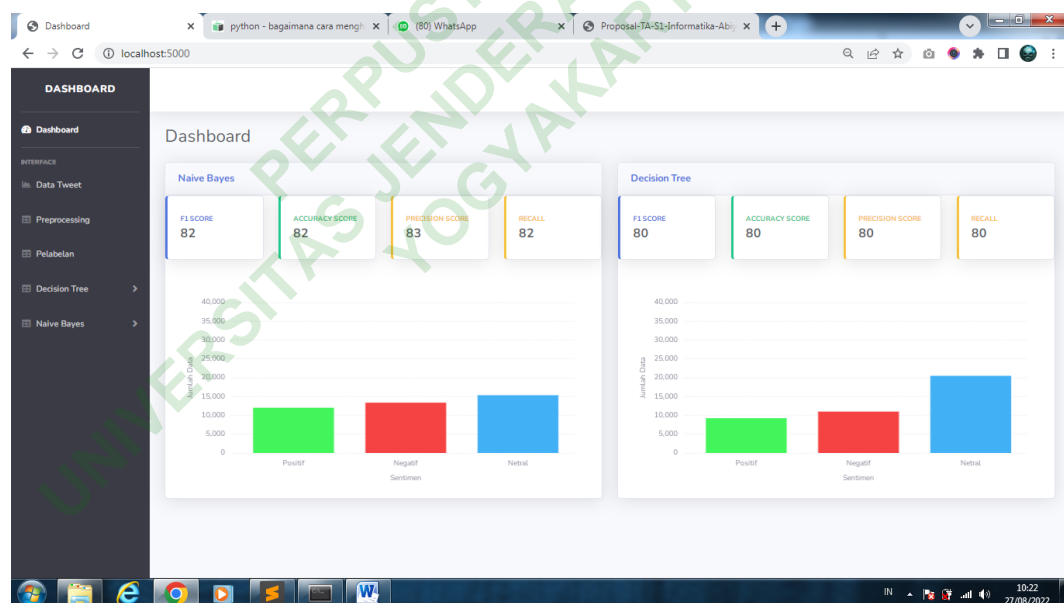
Tabel 4.10 Tweet Netral

Tweet
pln sabet penghargaan bumh entrepreneurial marketing award
hi kak admin informasikan xauntuk pelanggan layanan paskabayar nomor id pelanggan p
menteri bumh erick thohir seiring penambahan anggaran subsidi pln pertamin
manfaatkan layanan ratusan siswa guru sma air pura install pln mobile
hai kak mursyid mohon maaf ketidaknyamanannya mohon mengirimkan kendala kelis

4.8 IMPLEMENTASI APLIKASI

4.8.1 Dashboard

Dashboad merupakan tampilan awal dalam sistem dimana didalam halamanya akan menampilkan grafik perbandingan antara klasifikasi dengan menggunakan metode Naive Bayes dan Decision Tree seperti pada Gambar 4.7.

**Gambar 4.6** Halaman Dashboard

Dashboard adalah tampilan awal ketika halaman pertama kali di akses, pada halaman ini terdapat dua barchart yang mempresentasikan jumlah sentimen dari masing-masing analisis, data dipanggil langsung dari folder upload. Dengan menggunakan *library* pandas maka data bisa terbaca ke dalam sistem, kemudian

data dihitung berdasarkan kelas. Pada tampilan dashboard juga terdapat akurasi dari hasil klasifikasi atas keseluruhan data sehingga mempermudah dalam melakukan pengambilan keputusan.

4.8.2 Data Tweet

Data Tweet merupakan halaman yang berfungsi untuk meng-*upload* data yang akan di preprocessing dan menampilkan data dengan format csv. Data akan tampil di dalam tabel. Seperti pada Gambar 4.7.

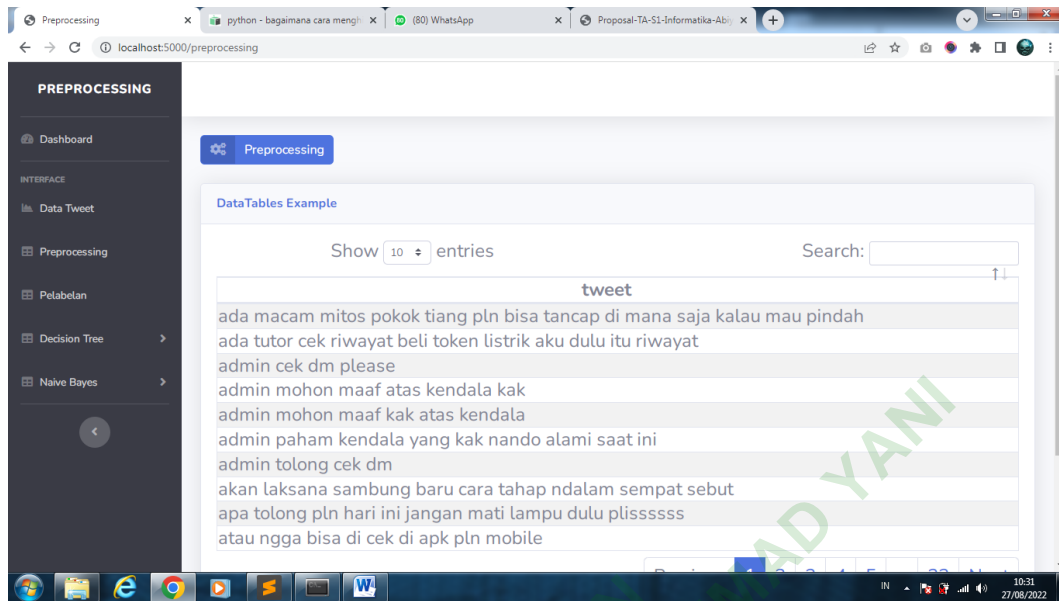


Gambar 4.7 Halaman Data Tweet

Halaman diatas juga bisa menampilkan data yang baru saja di-*upload*. Data yang di *upload* nantinya akan tersimpan kedalam folder upload yang ada didalam folder project. File yang di *upload* adalah format csv kemudian akan dibuka dan ditampilkan kedalam halaman html.

4.8.3 Preprocessing

Preprocessing merupakan halaman yang berfungsi untuk membersihkan mengubah data tweet menjadi struktur data yang baku sehingga mudah dalam melakukan analisis. Berikut merupakan tampilan dari halaman Preprocessing seperti pada Gambar 4.8.

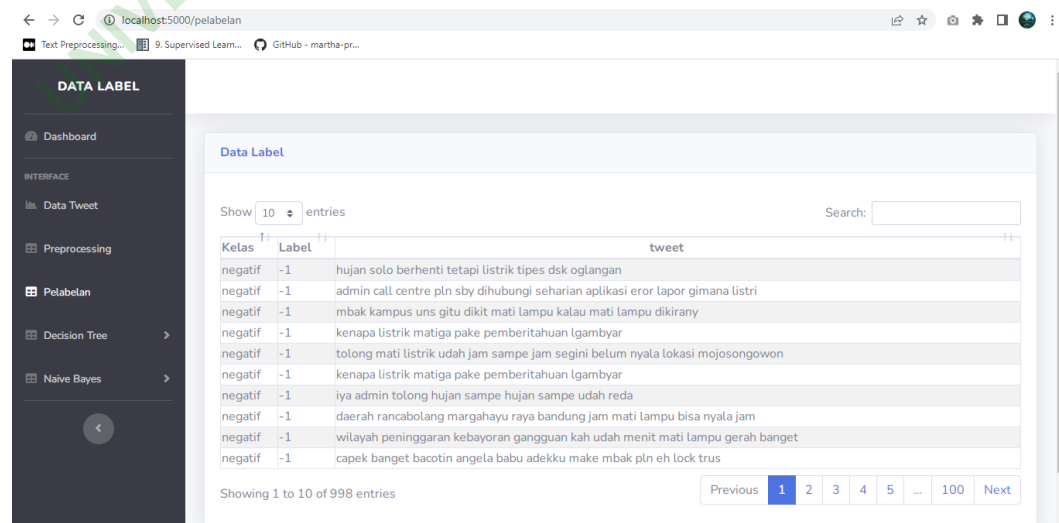


Gambar 4.8 Halaman Preprocessing

Hasil dari proses preprocessing adalah data yang telah bersih dan disimpan ke dalam file berformat .xlsx yang tersimpan ke dalam folder upload.

4.8.4 Pelabelan

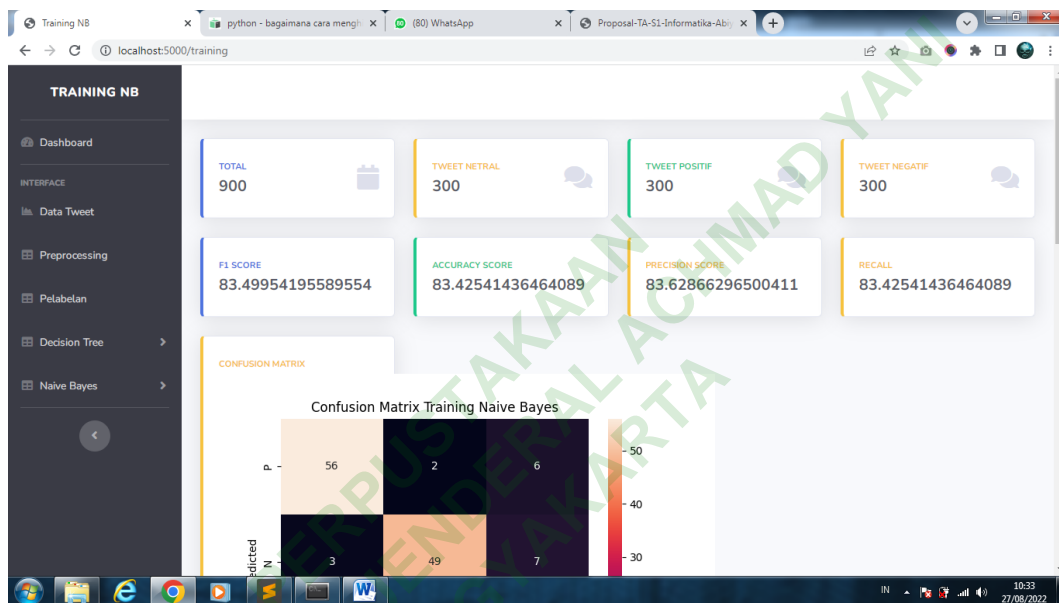
Pelabelan merupakan halaman yang menampilkan data yang telah melalui proses *preprocessing* dan telah memiliki label setimen, penambahan label dilakukan dengan cara manual menggunakan Microsoft Excel, data ini nantinya akan digunakan sebagai data *training* dalam *tesing* dan klasifikasi nantinya. Seperti pada Gambar 4.9.



Gambar 4.9 Halaman Pelabelan Manual

4.8.5 Training

Training merupakan halaman yang berfungsi untuk melatih data yang akan digunakan untuk melakukan klasifikasi, pada halaman *trainig* juga dilakukan proses *tf – idf* . Proses training dilakukan dengan menggunakan dua metode, yaitu metode Naive Bayes dan Decision Tree. Dapat dilihat pada Gambar 4.10.

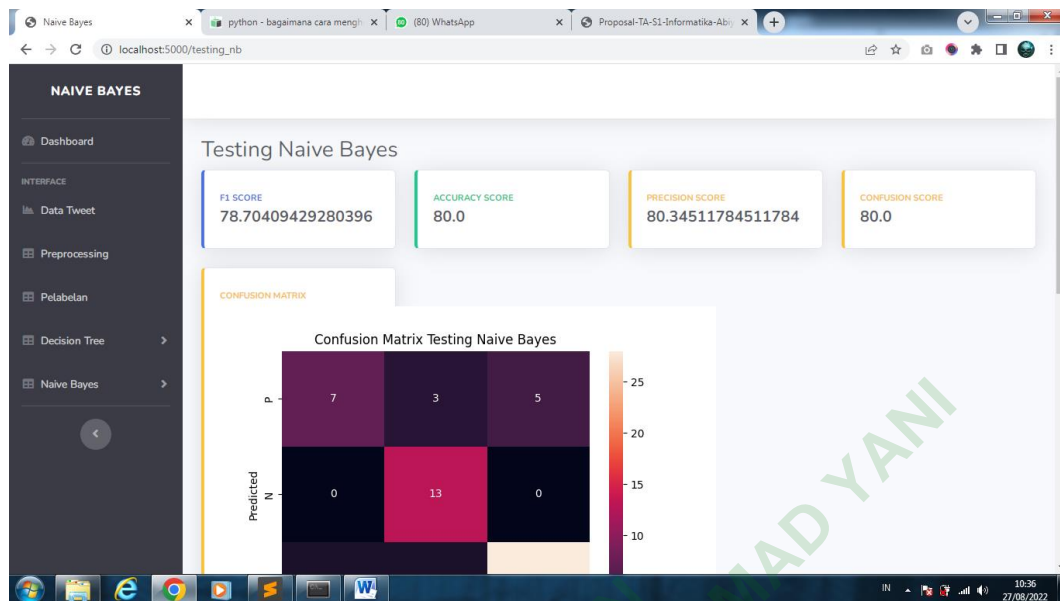


Gambar 4.10 Halaman *Training*

Data yang digunakan untuk *training* berjumlah 900 data dan dibagi menjadi 3 sentimen yaitu positif, negatif dan netral. Didalam halaman ini juga terdapat perhitungan akurasi untuk data *trainig*.. Didalam halaman *training* terdapat tampilah hasil *tf-idf* pada data training dan *confusion matrix*nya.

4.8.6 Testing

Halaman *testing* merupakan halaman uji coba untuk model yang telah dibuat, dalam halaman *testing* digunakan sebanyak 300 data testing, seperti pada Gambar 4.11.

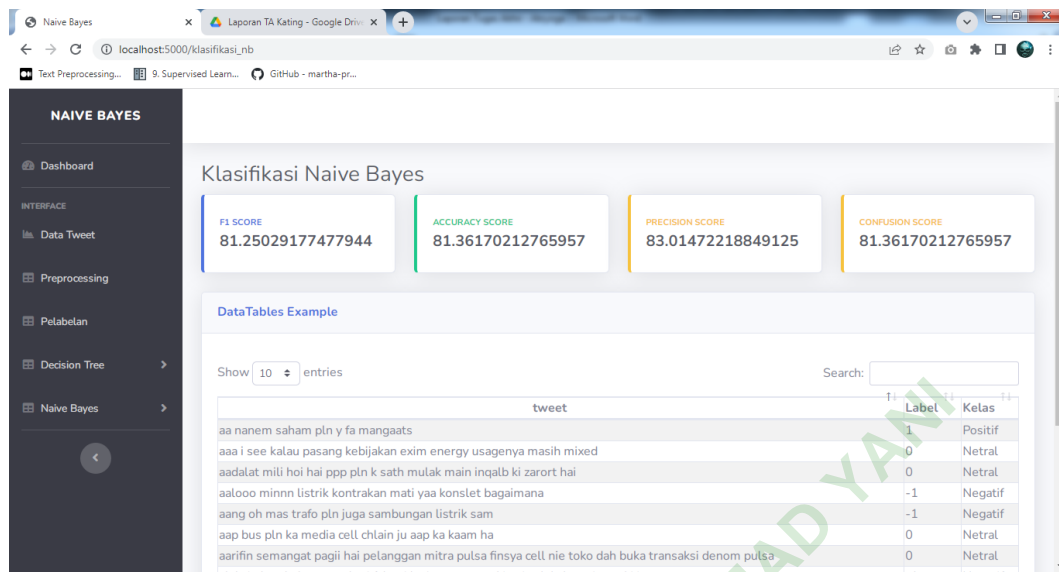


Gambar 4.11 Halaman *Testing*

Halaman *testing* menggunakan model pelatihan yang telah dibuat dihalaman *training* sebelumnya, sehingga bisa memprediksi kelas dan label yang terdapat pada 300 data yang ada dalam halaman *testing*.

4.8.7 Klasifikasi

Halaman klasifikasi merupakan halaman yang digunakan untuk melakukan proses pelabelan secara otomatis dengan menggunakan data *training* yang telah dilakukan sebelumnya. Dalam halaman klasifikasi menggunakan semua data yang telah diambil dengan jumlah total datanya 40.745. Kemudian menampilkan hasil akurasi seperti pada Gambar 4.12.



Gambar 4.12 Halaman Klasifikasi

Pada halaman klasifikasi pemberian label dan kelas dilakukan secara otomatis dengan menggunakan model yang sebelumnya dibuat pada halaman *training*.