

**LAPORAN AKHIR  
STUDI INDEPENDEN BERSERTIFIKAT  
AI Mastery Program  
Di Orbit Future Academy**

Diajukan untuk memenuhi persyaratan kelulusan  
Program MSIB MBKM

oleh :  
Dyah Ayu Wulandari / 195150201111062



**TEKNIK INFORMATIKA UNIVERSITAS BRAWIJAYA  
2022**

**Lembar Pengesahan Teknik Informatika Universitas Brawijaya**

**noHate : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis  
Website Menggunakan Kombinasi Model *Machine Learning***

**Di Orbit Future Academy**

oleh :

Dyah Ayu Wulandari / 195150201111062

disetujui dan disahkan sebagai

Laporan Magang atau Studi Independen Bersertifikat Kampus Merdeka

Malang, 17 Juni 2022

Pembimbing Studi Independen Teknik Informatika Universitas Brawijaya



Edy Santoso, S.Si., M.Kom.

NIP. 19740414 200312 1 004

**Lembar Pengesahan**

**noHate : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model *Machine Learning***

**Di Orbit Future Academy**

oleh :

Dyah Ayu Wulandari / 195150201111062

disetujui dan disahkan sebagai  
Laporan Magang atau Studi Independen Bersertifikat Kampus Merdeka

Jakarta, 17 Juni 2022

AI Coach



Eko Purbo Wahyono

NIP: 2201069

## **Abstraksi**

Program Studi Independen adalah salah satu program Merdeka Belajar Kampus Merdeka untuk menyiapkan mahasiswa dalam menghadapi perubahan sosial, budaya, dunia kerja dan kemajuan teknologi yang pesat. Salah satu mitra yang menyediakan program studi independen adalah PT.Orbit Ventura Indonesia. Terdapat 2 macam kegiatan studi independen yang dilaksanakan Orbit, yaitu *Foundations of AI and Life Skills for Gen-Z* dan *Artificial Intelligence Mastery Program*. *Artificial Intelligence Mastery Program* adalah program pelatihan *Artificial Intelligence* (AI) secara daring untuk pelajar yang bertujuan mengangkat perangkat AI, sehingga bisa membuat sesuatu yang menciptakan dampak sosial. Untuk mencapai tujuan tersebut, maka salah satu output dari kegiatan ini adalah dengan membuat *project* akhir yang berfokus ke salah satu domain, salah satunya yaitu *Natural Language Processing*.

*Project* yang dibuat yaitu “**noHate** : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model *Machine Learning*” menggunakan Bahasa Indonesia dengan metode *Logistic Regression*, *Support Vector Machine*, *Naive Bayes (Bernoulli)*, *Decision Tree*, dan *Stacking*. Proses pembuatan *project* akhir mulai dari mencari dataset, *import* dataset, *preprocessing* (*normalize* alay, *stopword removal*, *remove emoticon*, *casefolding*, *stemming*), *wordcloud*, *vectorization (TF-IDF)*, *modelling (Logistic Regression, Support Vector Machine, Naive Bayes, Decision Tree, dan Stacking)*, *testing*, evaluasi model, *front end*, *back end* dan *deployment*.

Tujuan dibuatnya aplikasi tersebut agar masyarakat awam akan dapat membedakan kalimat yang termasuk ujaran kebencian atau tidak dalam kebebasan berpendapat sehingga tidak menimbulkan diskriminasi, pengucilan, dan lain-lain dalam suatu individu atau golongan serta dapat lebih bijak dalam berpendapat dan tidak mudah terpengaruh oleh pendapat orang lain.

**Kata Kunci :** *Studi Independen, Artificial Intelligence, Ujaran Kebencian*

## Kata Pengantar

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa. Atas rahmat dan karunia-Nya, penulis dapat menyelesaikan laporan akhir Studi Independen Batch 2 dengan judul “**noHate** : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model *Machine Learning*”. Penyusunan laporan ini dilakukan untuk memenuhi persyaratan akhir Program Studi Independen *Batch 2*.

Selama proses program studi independen yang dilaksanakan secara *online (daring)* serta proses penyusunan laporan akhir ini tentu tak lepas dari bantuan, arahan, masukan, serta bimbingan dari berbagai pihak. Untuk itu, penulis mengucapkan terima kasih kepada :

1. Prof Widodo, S.Si.,M.Si.,Ph.D.Med.Sc selaku Rektor Universitas Brawijaya
2. Adhitya Bhawiyuga, S.Kom., M.Sc. selaku Ketua Program Studi Teknik Informatika Universitas Brawijaya.
3. Edy Santoso, S.Si., M.Kom. selaku Dosen Pembimbing Lapangan Program Studi Independen.
4. Eko Purbo Wahyono, Mukhlizar Nirwan, Ely Sudarsono, Hafid Galih, dan Wijaya Yudha Atmaja selaku Mentor Program Studi Independen.
5. Serta berbagai pihak yang tidak mungkin dapat penulis sebutkan satu per satu.

Meski demikian, penulis menyadari masih banyak kekurangan dalam penulisan laporan program studi independen ini, sehingga penulis secara terbuka menerima saran dan kritik positif dari pembaca. Semoga laporan program studi independen ini dapat bermanfaat dan dapat menjadi referensi yang baik bagi pembaca khususnya mahasiswa yang hendak melaksanakan studi independen baik di instansi yang sama maupun instansi yang berbeda. Terima kasih.

Ngawi, 17 Juni 2022

Penulis

Dyah Ayu Wulandari

## **Daftar Isi**

Bab I	Pendahuluan	1
I.1	Latar belakang	1
I.2	Lingkup	2
I.3	Tujuan	3
Bab II	Orbit Future Academy	1
II.1	Struktur Organisasi	1
II.2	Lingkup Pekerjaan	2
II.3	Deskripsi Pekerjaan	3
II.4	Jadwal Kerja	5
Bab III	noHate : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model <i>Machine Learning</i>	1
III.1	Latar Belakang Proyek Akhir	1
III.2	Proses Pelaksanaan Proyek Akhir	3
III.3	Hasil Proyek Akhir	6
Bab IV	Penutup	1
IV.1	Kesimpulan	1
IV.2	Saran	1
Bab V	Referensi	vii
Bab VI	Lampiran A. TOR	1
Bab VII	Lampiran B. Log Activity	1
Bab VIII	Lampiran C. Dokumen Teknik	1

## **Daftar Tabel**

Tabel 2.1 Agenda Kelas

5

## Daftar Gambar

Gambar 2.1 Logo Orbit Future Academy	1
Gambar 2.2 Struktur Organisasi OFA	2
Gambar 2.3 <i>Wordcloud</i> ujaran kebencian	4
Gambar 2.4 <i>Wordcloud</i> bukan ujaran kebencian	4
Gambar 2.5 <i>Landing page</i> Aplikasi <b>noHate</b>	5
Gambar 2.6 <i>Dashboard</i> Aplikasi <b>noHate</b>	5
Gambar 3.1 Data Statistik Pengguna Media Sosial di Indonesia	1
Gambar 3.2 <i>Wordcloud</i> ujaran kebencian	3
Gambar 3.3 <i>Wordcloud</i> bukan ujaran kebencian	3
Gambar 3.4 <i>Landing page</i> Aplikasi <b>noHate</b>	5
Gambar 3.5 Tampilan <i>Dashboard</i>	5
Gambar 3.6 Tampilan Fitur <i>Prediction</i>	6
Gambar 3.7 Tampilan Gambar <i>Dashboard</i> Hasil Prediksi Input Suatu Kalimat	7
Gambar 3.8 Tampilan Gambar <i>Dashboard</i> Hasil Prediksi Suatu File (.csv)	7
Gambar 3.9 Tampilan Halaman <i>Preprocessing Casefolding</i>	8
Gambar 3.10 Tampilan Halaman <i>Preprocessing Normalize Alay</i>	8
Gambar 3.11 Tampilan Halaman <i>Preprocessing Stopword removal</i>	9
Gambar 3.12 Tampilan Halaman <i>Preprocessing Remove emoticon</i>	9
Gambar 3.13 Tampilan Tampilan Halaman <i>Preprocessing Stemming</i>	10
Gambar 3.14 Tampilan Halaman Hasil <i>Preprocessing</i> dari semua proses ( <i>casefolding</i> , <i>normalize alay</i> , <i>stopword removal</i> , <i>remove emoticon</i> byte, dan <i>stemming</i> )	10
Gambar 3.15 Tampilan Halaman Visualisasi dan Akurasi	11
Gambar 3.16 Tampilan Halaman <i>Contact us</i>	11

Gambar 3.17 Tampilan Halaman <i>Home</i>	12
Gambar 3.18 Tampilan Halaman <i>About us</i>	12
Gambar 3.19 Tampilan Halaman <i>Services</i>	13
Gambar 3.20 Tampilan Halaman <i>Our team</i>	13
Gambar 3.21 <i>Classification Report Logistic Regression</i>	13
Gambar 3.22 <i>Confusion matrix Logistic Regression</i>	14
Gambar 3.23 <i>Classification Report Support Vector Machine</i>	14
Gambar 3.24 <i>Confusion matrix Support Vector Machine</i>	14
Gambar 3.25 <i>Classification Report Naïve Bayes (Bernoulli)</i>	15
Gambar 3.26 <i>Confusion matrix Naïve Bayes (Bernoulli)</i>	15
Gambar 3.27 <i>Classification Report Decision Tree</i>	16
Gambar 3.28 <i>Confusion matrix Decision Tree</i>	16
Gambar 3.29 <i>Classification Report Stacking</i>	16
Gambar 3.30 <i>Confusion matrix Stacking</i>	17
Gambar 8.1 Perbandingan Dataset Ujaran Kebencian (label 1) dan Bukan Ujaran Kebencian (label 0)	3
Gambar 8.2 <i>Modelling</i>	3
Gambar 8.3 <i>Modelling Decision Tree</i>	5
Gambar 8.4 <i>Modelling Stacking</i>	6
Gambar 8.5 <i>Classification Report Logistic Regression</i>	7
Gambar 8.6 <i>Confusion matrix Logistic Regression</i>	7
Gambar 8.7 <i>Classification Report Support Vector Machine</i>	8
Gambar 8.8 <i>Confusion matrix Support Vector Machine</i>	8
Gambar 8.9 <i>Classification Report Naïve Bayes (Bernoulli)</i>	8
Gambar 8.10 <i>Confusion matrix Naïve Bayes (Bernoulli)</i>	9

Gambar 8.11 <i>Classification Report Decision Tree</i>	9
Gambar 8.12 <i>Confusion matrix Decision Tree</i>	9
Gambar 8.13 <i>Classification Report Stacking</i>	10
Gambar 8.14 <i>Confusion matrix Stacking</i>	10
Gambar 8.15 Tampilan <i>Landing page</i> Aplikasi <b>noHate</b>	11
Gambar 8.16 Logo Aplikasi <b>noHate</b>	13
Gambar 8.17 Tampilan <i>Dashboard</i> Hasil Prediksi	14
Gambar 8.18 Tampilan Fitur <i>Prediction</i>	15
Gambar 8.19 Tampilan Gambar <i>Dashboard</i> Hasil Prediksi Input Suatu Kalimat	15
Gambar 8.20 Tampilan Gambar <i>Dashboard</i> Hasil Prediksi Suatu File (.csv)	16
Gambar 8.21 Tampilan Halaman <i>Preprocessing Casefolding</i>	16
Gambar 8.22 Tampilan Halaman <i>Preprocessing Normalize Alay</i>	17
Gambar 8.23 Tampilan Halaman <i>Preprocessing Stopword removal</i>	17
Gambar 8.24 Tampilan Halaman <i>Preprocessing Remove emoticon</i>	18
Gambar 8.25 Tampilan Tampilan Halaman <i>Preprocessing Stemming</i>	18
Gambar 8.26 Tampilan Halaman Hasil <i>Preprocessing</i> dari semua proses ( <i>casefolding</i> , <i>normalize alay</i> , <i>stopword removal</i> , <i>remove emoticon</i> byte, dan <i>stemming</i> )	19
Gambar 8.27 Tampilan Halaman Visualisasi dan Akurasi	20
Gambar 8.28 Kodingan <i>Wordcloud</i>	20
Gambar 8.29 Tampilan Halaman <i>Contact us</i>	21
Gambar 8.30 Tampilan Halaman <i>Home</i>	21
Gambar 8.31 Tampilan Halaman <i>About us</i>	22
Gambar 8.32 Tampilan Halaman <i>Services</i>	22
Gambar 8.33 Tampilan Halaman <i>Our team</i>	23

Gambar 8.34 Kodingan index.html	23
Gambar 8.35 Kodingan <i>dashboard.html</i> dan <i>casefolding.html</i>	24
Gambar 8.36 Kodingan <i>normalize.html</i> dan <i>stopword-removal.html</i>	24
Gambar 8.37 Kodingan <i>removeemoticon.html</i> dan <i>stemming.html</i>	25
Gambar 8.38 Kodingan <i>hasilpreprocess.html</i> dan <i>visualisasi.html</i>	25

## Bab I Pendahuluan

### I.1 Latar belakang

Dalam rangka menyiapkan mahasiswa dalam menghadapi perubahan sosial, budaya, dunia kerja dan kemajuan teknologi yang pesat, kompetensi mahasiswa harus disiapkan untuk lebih mengikuti dengan kebutuhan zaman. Kebijakan Merdeka Belajar - Kampus Merdeka diharapkan dapat menyiapkan mahasiswa dalam menghadapi perubahan tersebut. Kampus Merdeka merupakan wujud pembelajaran di perguruan tinggi yang otonom dan fleksibel sehingga tercipta kultur belajar yang inovatif, tidak mengekang, dan sesuai dengan kebutuhan mahasiswa. Salah satu bentuk kegiatan belajar di luar perguruan tinggi yaitu program MSIB (Magang dan Studi Independen). Terdapat banyak mitra yang bekerja sama dalam program MSIB, salah satunya yaitu PT. Orbit Ventura Indonesia.

PT. Orbit Ventura Indonesia menyediakan program studi independen untuk mahasiswa yang ingin mempelajari *Artificial Intelligence*. Terdapat 2 macam kegiatan studi independen yang dilaksanakan Orbit, yaitu *Foundations of AI and Life Skills for Gen-Z* dan *Artificial Intelligence Mastery Program*. *Artificial Intelligence Mastery Program* adalah program pelatihan *Artificial Intelligence* (AI) secara daring untuk pelajar yang bertujuan bukan hanya untuk memperkenalkan teknologi AI ke pelajar, tetapi juga untuk memungkinkan bisa mengangkat perangkat AI, sehingga bisa membuat sesuatu yang menciptakan dampak sosial. Program ini berfokus pada komponen utama AI seperti *Data Science*, *Natural Language Processing*, *Computer Vision*, dan *Reinforcement Learning*. Untuk mencapai salah satu tujuan tersebut yaitu mengangkat perangkat AI, maka salah satu output dari kegiatan ini adalah dengan membuat *project* akhir yang berfokus ke salah satu domain. *Project* akhir penulis berfokus pada *domain Natural Language Processing*. *Natural Language Processing* (NLP) adalah cabang dari *Artificial Intelligence* yang berhubungan dengan interaksi antara mesin dan manusia menggunakan bahasa natural [1].

Oleh karena itu, selama program studi independen diharapkan mahasiswa akan mendapatkan pengalaman baik berupa *hard skills* (keterampilan, *complex problem solving*, *analytical skills*, dsb.), maupun *soft skills* (etika profesi/kerja, komunikasi, kerjasama, dsb.). Melalui program studi independen yang dirancang dan diimplementasikan dengan baik, maka *hard skills* dan *soft skills* mahasiswa akan terbentuk dengan kuat. Selain itu, diharapkan program ini juga sesuai dengan visi dan misi dari orbit yaitu memberikan pembelajaran berbasis keterampilan transformatif dan melatih pemuda dengan keterampilan kewirausahaan serta mencocokkan mereka dengan peluang masa depan yang muncul di berbagai industri sebab program ini berpusat pada membangun *AI Readiness* / kesiapan AI dengan cara mempersiapkan pelajar dengan keterampilan yang dibutuhkan dalam pekerjaan masa depan dan sebagai fundamental dalam pengembangan mindset dan skillset untuk sukses dalam dunia yang sudah dan akan terus menjadi *AI dependant* / memiliki ketergantungan dengan AI.

Program Studi Independen diharapkan dapat menjawab tantangan Perguruan Tinggi untuk menghasilkan lulusan yang sesuai perkembangan zaman, kemajuan IPTEK, tuntutan dunia usaha dan dunia industri, maupun dinamika masyarakat.

## I.2 Lingkup

Pelaksanaan pembuatan *project* akhir berfokus pada domain *Natural Language Processing* dengan membuat aplikasi untuk mendeteksi ujaran kebencian dengan cara memberi inputan dari pengguna atau mengunggah file (.csv). Dalam *project* ini juga ditampilkan hasil setiap *preprocessing* yang dilakukan. Pelaksanaan *project* akhir ini, tugas dibagi dalam 3 bagian antara lain yaitu :

1. Tugas awal yaitu mencari dataset, *import* dataset, *preprocessing* (*normalize alay*, *stopword removal*, *remove emoticon*, *casefolding*, *stemming*), dan *wordcloud*.
2. Tugas tengah yaitu *vectorization (TF-IDF)*, *modelling (Logistic Regression, Support Vector Machine, Naive Bayes, Decision Tree, dan Stacking)*, dan membuat *testing* dengan kalimat masukan di setiap model

3. Tugas akhir yaitu evaluasi model, demonstrasi di colab, *front end*, *back end* dan *deployment*.

### I.3 Tujuan

1. Mampu memahami apa itu AI, penerapan dan pemanfaatannya.
2. Mampu memahami pentingnya data dalam AI.
3. Mampu membuat *project AI* yang berdampak sosial.
4. Mampu menulis kode dengan bahasa pemrograman Python.
5. Mampu memahami operasi dan logika sederhana pada Python.
6. Mampu membuat *project Python*.
7. Mampu membuat *repository* di akun Git/Github.
8. Mampu membuat portfolio dengan Git/Github.
9. Mampu memahami terkait *Data Science*, *Computer Vision*, *Natural Language Processing*, dan *Reinforcement Learning*.
10. Mampu membuat aplikasi deteksi ujaran kebencian
11. Mampu membuat model *Machine Learning*
12. Mampu membuat kode program untuk pengujian model *Machine Learning*.
13. Mampu melakukan pengujian model dan analisis.
14. Mampu men-deploy model menggunakan Heroku dan atau menggunakan *Machine Learning as a service* (MLaaS).

## Bab II Orbit Future Academy

### II.1 Struktur Organisasi



Gambar 2.1 Logo Orbit Future Academy

Orbit Future Academy (OFA) didirikan pada tahun 2016 dengan tujuan untuk meningkatkan kualitas hidup melalui inovasi, edukasi, dan pelatihan keterampilan. Label atau *brand* Orbit merupakan kelanjutan dari warisan mendiang Prof. Dr. Ing. B. J. Habibie (presiden Republik Indonesia ke-3) danistrinya, Dr. Hasri Ainun Habibie. Mereka berdua telah menjadi penggerak dalam mendukung perkembangan inovasi dan teknologi pendidikan di Indonesia. OFA mengurasi dan melokalkan program/kursus internasional untuk *upskilling* atau *reskilling* pemuda dan tenaga kerja menuju pekerjaan masa depan. Hal ini sesuai dengan slogan OFA, yakni “*Skills-for-Future-Jobs*”.

#### Visi:

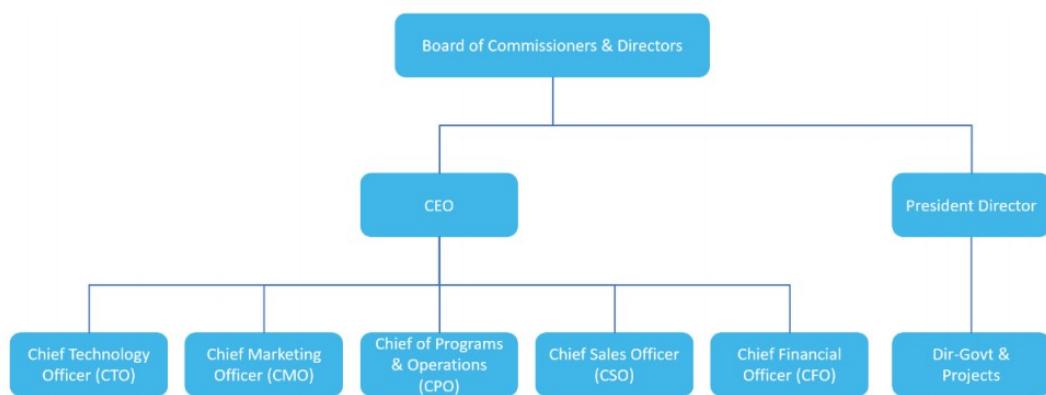
Memberikan pembelajaran berbasis keterampilan transformatif terbaik untuk para pencari kerja & pencipta lapangan kerja.

#### Misi:

1. Membangun jaringan Orbit Transformation Center (OTC) secara nasional untuk menyampaikan kurikulum keterampilan masa depan berbasis sertifikasi melalui *Platform* Konten Digital.

2. Secara proaktif bekerja dengan pemerintah & organisasi dengan mengubah tenaga kerja mereka agar sesuai dengan perubahan pekerjaan yang terjadi karena Industri 4.0.
3. Melatih pemuda dengan keterampilan kewirausahaan & mencocokkan mereka dengan peluang masa depan yang muncul di berbagai industri.
4. Menghubungkan jaringan inkubator dan akselerator yang dikurasi ke industri, investor, dan ekosistem start-up global.

Struktur organisasi OFA dapat dilihat pada Gambar 2.2.



Gambar 2.2 Struktur Organisasi OFA

## II.2 Lingkup Pekerjaan

Seorang fasilitator akan mendampingi kurang lebih 40 peserta MSIB (student) dalam satu kelas. Terdapat dua jenis fasilitator, yakni:

### a. ***Homeroom Coach***

*Homeroom coach* bertugas menyampaikan materi tentang dasar-dasar AI, memberikan penilaian pada student, dan mengarahkan *student* saat pengerjaan Proyek Akhir (PA).

### b. Domain Coach

Domain coach bertugas menyampaikan materi tentang domain AI dan memberikan penilaian pada student.

Lingkup pekerjaan student adalah mengikuti kelas bersama *homeroom* atau domain coach, sesuai agenda kelas, hingga program selesai.

## II.3 Deskripsi Pekerjaan

Berikut adalah deskripsi pekerjaan student sebelum penggerjaan PA:

- a. Mengikuti pre-test.
- b. Mengikuti kelas sesi pagi pada pukul 08.00 hingga 11.30 WIB.
- c. Mengikuti kelas sesi siang pada pukul 13.00 hingga 16.30 WIB.
- d. Mengulang materi yang telah disampaikan di kelas sesi pagi dan siang, setelah kelas sesi siang, selama 1 jam (*self-study*).
- e. Mengerjakan latihan individu atau kelompok yang diberikan oleh *homeroom* atau domain coach saat kelas berlangsung.
- f. Mengerjakan tugas yang diberikan *homeroom* atau domain coach hingga batas waktu tertentu.
- g. Mengerjakan *mini project* yang diberikan *homeroom* atau domain coach hingga batas waktu tertentu
- h. Mengikuti post-test.

Student memiliki peran membuat *wordcloud*, membuat *modelling* (*Decision Tree* dan *Stacking*), dan membuat *front end* aplikasi **noHate** selama penggerjaan PA, dengan deskripsi pekerjaan sebagai berikut:

- a. Membuat *Wordcloud*. *Wordcloud* adalah gambar yang menunjukkan kata-kata yang terdapat dalam suatu kalimat dan semakin banyak kata sering digunakan maka semakin besar ukuran kata tersebut dalam gambar. Berikut adalah *wordcloud* ujaran kebencian (label 1) :



Gambar 2.3 *Wordcloud* ujaran kebencian

Sedangkan, di bawah ini adalah *wordcloud* bukan ujaran kebencian (label 0) :



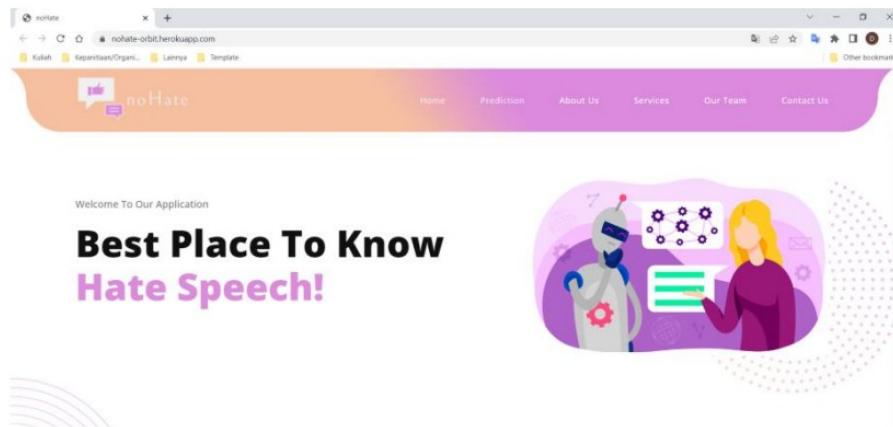
Gambar 2.4 *Wordcloud* bukan ujaran kebencian

b. Membuat *Modelling Decision Tree* dan *Stacking*

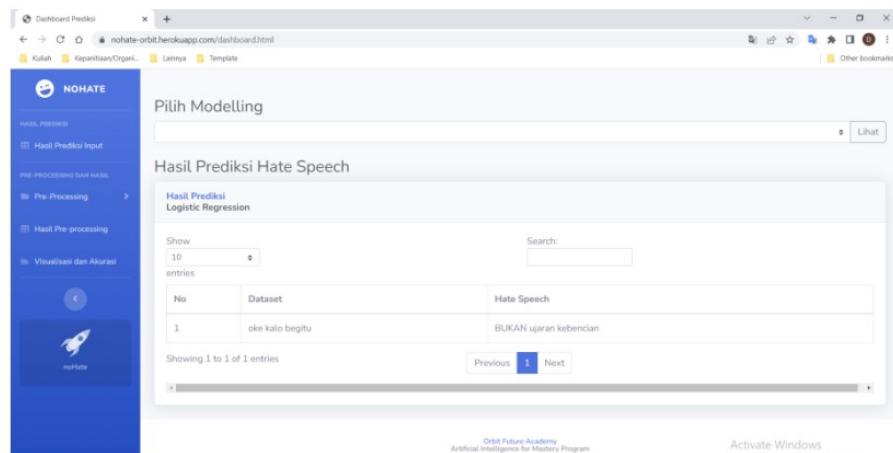
Setelah membuat model, maka dapat diketahui hasil akurasi dari evaluasi *modelling Decision Tree* adalah 0.78 dan *Stacking* adalah 0.85.

c. Membuat *front end* aplikasi **noHate**

Berikut adalah tampilan *front end* aplikasi **noHate** :



Gambar 2.5 *Landing page Aplikasi noHate*



Gambar 2.6 *Dashboard Aplikasi noHate*

## II.4 Jadwal Kerja

Program ini berlangsung setiap hari kerja (Senin sampai dengan Jumat) selama 8 jam per harinya, dengan rincian sebagai berikut:

Tabel 2.1 Agenda Kelas

Pukul (WIB)	Durasi (jam)	Aktivitas
08.00 s.d. 11.30	3.5	Kelas Sesi Pagi
13.00 s.d. 16.30	3.5	Kelas Sesi Siang
16.30 s.d. 17.30	1	<i>Self-Study</i>

Program ini berlangsung dari bulan Februari 2022 sampai dengan bulan Juli 2022.

### Bab III noHate : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model Machine Learning

#### III.1 Latar Belakang Proyek Akhir

Seiring dengan perkembangan zaman, teknologi berkembang semakin pesat. Perkembangan teknologi yang semakin cepat mengakibatkan maraknya penggunaan media sosial. Berdasarkan TiNewss.Com, terdapat 191,4 juta pengguna media sosial di Indonesia pada Januari 2022 dan jumlah pengguna media sosial di Indonesia pada awal tahun 2022 setara dengan 68,9 persen dari total populasi [2]. Menurut We Are Social, jumlah pengguna media sosial di Indonesia meningkat setiap tahunnya. Berikut adalah data statistik pengguna media sosial di Indonesia.



Gambar 3.1 Data Statistik Pengguna Media Sosial di Indonesia

Pengguna media sosial yang semakin banyak, maka semakin banyak pula kebebasan berekspresi masyarakat dalam menyampaikan pendapat/argumennya dalam kehidupan sehari-hari sehingga terjadi peningkatan kebebasan diskusi dan interaksi dalam berbagai bidang seperti politik, ekonomi, budaya, sosial, kesehatan, dan lain-lain. Kebebasan berpendapat tersebut dapat disampaikan dalam bentuk ekspresi positif seperti pujian, atau dapat juga disampaikan dalam ekspresi negatif seperti ujaran kebencian (*hate speech*). *Hate speech* (ujaran kebencian) adalah

tindakan komunikasi yang dilakukan oleh suatu individu atau kelompok dalam bentuk provokasi, hasutan, ataupun hinaan kepada individu atau kelompok yang lain dalam berbagai aspek seperti ras, warna kulit, etnis, gender, cacat, orientasi seksual, kewarganegaraan, agama, dan lain-lain [3]. Dalam bidang politik, *hate speech* sering menjadi alat untuk menghalangi proses demokrasi dan bahkan menjatuhkan pihak lawan. Berdasarkan bawaslu.go.id, anggota Bawaslu yaitu Mochammad Afifuddin mengatakan sebagai pengawas pemilu menyadari perkembangan demokrasi di Indonesia berkembang secara dinamis sehingga di berbagai tingkat perkembangannya terbukti rentan terhadap ujaran kebencian (*hate speech*) dan efek negatifnya serta. Dalam kampanye Pemilu menurutnya menyediakan lahan subur untuk ujaran kebencian dan hasutan kebencian, baik pejabat terpilih, partai politik, kandidat atau calon, masyarakat sipil termasuk penyebar ujaran kebencian serta efek menguatkan yang dimiliki oleh media massa, terutama media sosial [4]. Selain itu, pada detik.com menjelaskan bahwa anggota Badan Pengawas Pemilu (Bawaslu) RI Mochammad Afifuddin menilai definisi ‘ujaran kebencian’ masih memunculkan perbedaan dalam perspektif masyarakat dan masih multitafsir sehingga menjadi tantangan untuk menyatukan pemahaman agar tercipta batasan dalam ujaran kebencian [5]. Berdasarkan hal tersebut, penulis membuat “**noHate** : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model *Machine Learning*” yang menggunakan Bahasa Indonesia dengan metode *Logistic Regression*, *Support Vector Machine*, *Naive Bayes (Bernoulli)*, *Decision Tree*, dan *Stacking* yang akan digunakan untuk membangun sistem yang dapat memberikan informasi terhadap pendekripsi pernyataan atau argumen atau file dalam format csv di media sosial yang mengarah pada *hate speech*. Dengan aplikasi tersebut, maka masyarakat awam akan dapat membedakan kalimat yang termasuk ujaran kebencian atau tidak dalam kebebasan berpendapat sehingga tidak menimbulkan diskriminasi, pengucilan, dan lain-lain dalam suatu individu atau golongan. Selain itu, dengan aplikasi tersebut harapannya masyarakat dapat lebih bijak dalam berpendapat dan tidak mudah terpengaruh oleh pendapat yang menjatuhkan lawan dalam pemilu nanti dan dapat bijak dalam memilih calon pemilu.

### **III.2 Proses Pelaksanaan Proyek Akhir**

Proses dalam pelaksanaan *project* akhir aplikasi **noHate** yang saya lakukan antara lain yaitu :

### 1. Membuat *wordcloud*

*Wordcloud* adalah gambar yang menunjukkan kata-kata yang terdapat dalam suatu kalimat dan semakin banyak kata sering digunakan maka semakin besar ukuran kata tersebut dalam gambar.

Di bawah ini adalah gambar *wordcloud* ujaran kebencian (label 1) :



Gambar 3.2 *Wordcloud* ujaran kebencian

Sedangkan, di bawah ini adalah *wordcloud* bukan ujaran kebencian (label 0) :



Gambar 3.3 Wordcloud bukan ujaran kebencian

## 2. Membuat *Modelling Decision Tree* dan *Stacking*

### A. Decision Tree

*Decision Tree* adalah salah satu metode klasifikasi untuk mem-break down proses pengambilan keputusan yang kompleks menjadi lebih simple, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan. Jadi, dengan model ini bisa mendapatkan solusi apakah suatu kalimat termasuk ujaran kebencian atau tidak. Pembuatan *Decision Tree* sendiri menggunakan metode *supervised Machine Learning* yaitu proses pembelajaran dimana data baru diklasifikasikan berdasarkan *training samples* yang sudah ada. Terdapat 3 jenis node yang terdapat pada Decision Tree, yaitu:

- a. Root node, merupakan node yang terletak paling atas. Node ini tidak ada input dan juga bisa tidak mempunyai output ataupun mempunyai output lebih dari satu
- b. Internal Node/node percabangan, pada node ini terdapat satu input dan memiliki output mininamal dua.
- c. Leaf node/terminal node merupakan node akhir. Node ini terdapat satu input dan tidak mempunyai output.

Hasil akurasi *decision Tree* adalah 0.7880539499036608

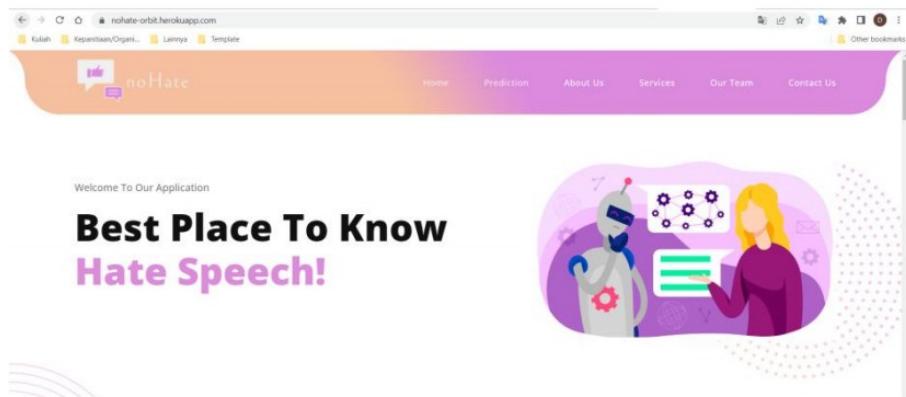
### B. Stacking

*Stacking* melibatkan pemasangan banyak jenis model yang berbeda (dalam *project* ini model *Logistic Regression*, *Support Vector Machine*, *naïve bayes (Bernoulli)*, dan *Decision Tree*) pada data yang sama dan menggunakan model lain untuk mempelajari cara terbaik menggabungkan prediksi. Dalam proses stacking set data pelatihan tidak berubah, algoritma ML yang berbeda untuk setiap anggota ensemble (ensemble member), dan model ML digunakan untuk mempelajari cara terbaik menggabungkan model prediksi. Hasil akurasi *Stacking* adalah 0.855009633911368

### 3. Membuat *Front end* aplikasi **noHate**

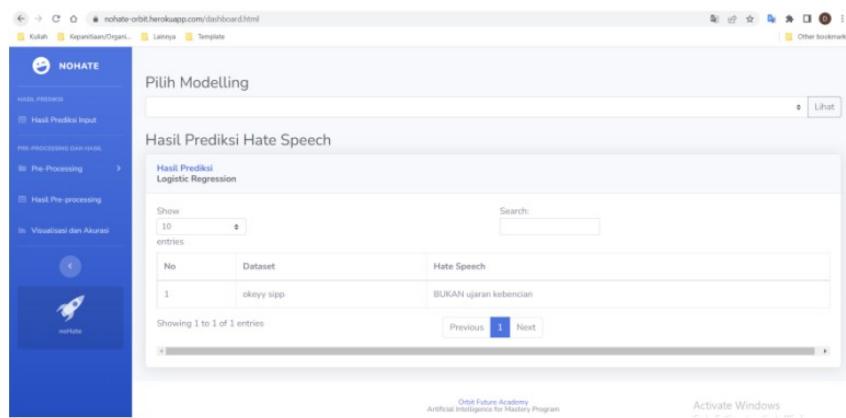
Untuk membuat *front end* aplikasi **noHate** perlu menentukan apa saja fitur yang akan ditampilkan. Terdapat beberapa fitur di aplikasi **noHate**, antara lain yaitu *home*, *prediction*, *about us*, *services*, *our team*, *contact us*, hasil prediksi

input, proses *preprocessing* (*casefolding*, *normalize* alay, *stopword removal*, *remove emoticon* byte, *stemming*), hasil *preprocessing*, visualisasi dan akurasi. Pertama kali yang dibuat yaitu halaman index.html untuk *landing page* yang berisi *home*, *prediction*, *about us*, *services*, *our team*, dan *contact us*. Berikut adalah tampilan *landing page* :



Gambar 3.4 *Landing page* Aplikasi noHate

Setelah *landing page* selesai dibuat, maka membuat *dashboard* untuk menampilkan hasil prediksi dari inputan beserta proses *preprocessing*, hasil *preprocessing* serta visualisasi dan akurasi. Berikut adalah tampilan *dashboard* yang telah dibuat :



Gambar 3.5 Tampilan *Dashboard*

Untuk membuat *dashboard* tersebut, terdapat 8 file yang harus dibuat antara lain yaitu *dashboard.html*, *casefolding.html*, *normalize.html*, *stopword-removal.html*, *removeemoticon.html*, *stemming.html*, *hasilpreprocess.html*, *visualisasi.html*.

### III.3 Hasil Proyek Akhir

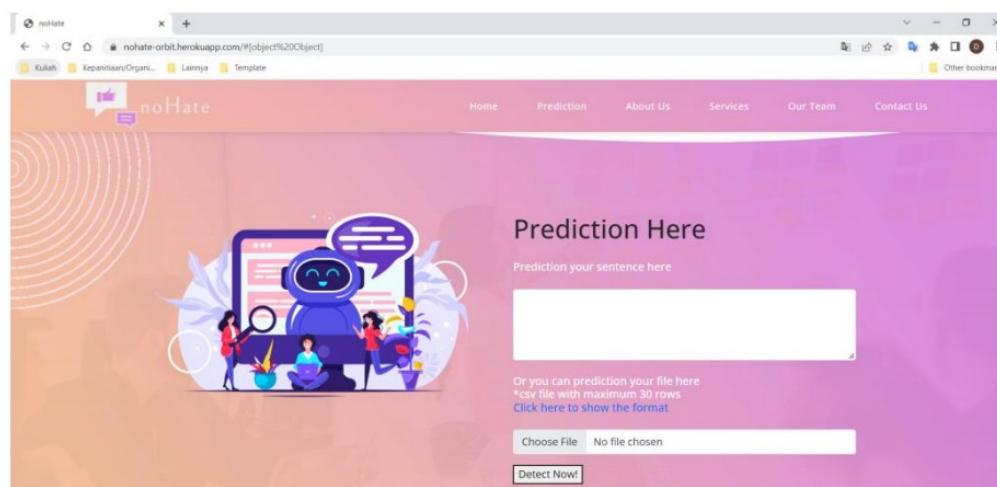
Hasil dari *project* akhir yaitu membuat aplikasi untuk mendeteksi ujaran kebencian yang diberi nama “**noHate**”. **noHate** adalah aplikasi untuk mendeteksi suatu kalimat mengandung ujaran kebencian (*hate speech*) atau tidak yang ditujukan kepada seseorang. Ujaran kebencian (*hate speech*) bisa berarti tindakan komunikasi yang dilakukan oleh suatu individu atau kelompok dalam bentuk provokasi, hasutan, ataupun hinaan kepada individu atau kelompok yang lain. Target user dari aplikasi **noHate** adalah pengguna media sosial baik dari kalangan anak-anak, remaja dan dewasa. Oleh karena itu, harapannya aplikasi **noHate** dapat bermanfaat untuk semua orang sehingga semua orang dapat bijak dalam mengutarakan pendapat atau menulis sesuatu di *platform* digital.

**noHate** terdiri atas 2 fungsi utama yaitu mendeteksi suatu kalimat yang diinput langsung oleh user apakah mengandung ujaran kebencian (*hate speech*) atau tidak, dan mendeteksi suatu file csv mengandung *hate speech* atau tidak serta terdapat visualisasi *Wordcloud* ujaran kebencian dan bukan ujaran kebencian, menampilkan akurasi, presisi, *Recall*, *F1-Score* dari tiap model, dan hasil *preprocessing*.

Aplikasi **noHate** memiliki beberapa fitur, antara lain yaitu :

#### 1. Fitur *Prediction*

Dalam fitur ini, pengguna dapat memprediksi kalimat baik berupa inputan kalimat atau file dengan format .csv yang diunggah.



Gambar 3.6 Tampilan Fitur *Prediction*

Berikut adalah tampilan setelah memasukkan kalimat “kamu jelek sekali” lalu menekan tombol “Detect Now!” dan diarahkan ke halaman *dashboard*. Setelah itu, memilih model yang digunakan dan menekan “Lihat”.

The screenshot shows the NOHATE dashboard interface. On the left sidebar, there are several sections: HASIL PREDIKSI (Hasil Prediksi Input), PRE-PROCESSING DAN HASIL (Pre-Processing & Hasil), and VISUALISASI DAN AKURASI (Visualisasi dan Akurasi). The main content area is titled "Pilih Modelling" and shows "Hasil Prediksi Hate Speech" using "Hasil Prediksi Logistic Regression". A table displays the result for the input "kamu jelek sekali", which is categorized as "Ujaran kebencian".

No	Dataset	Hate Speech
1	kamu jelek sekali	Ujaran kebencian

Gambar 3.7 Tampilan Gambar *Dashboard* Hasil Prediksi Input Suatu Kalimat

Berikut adalah tampilan setelah memasukkan file dengan format .csv lalu menekan tombol “Detect Now!” dan diarahkan ke halaman *dashboard*. Setelah itu, memilih model yang digunakan dan menekan “Lihat”.

The screenshot shows the NOHATE dashboard interface. The left sidebar includes sections for HASIL PREDIKSI, PRE-PROCESSING DAN HASIL, and VISUALISASI DAN AKURASI. The main area displays the results of a CSV file prediction using "Hasil Prediksi Logistic Regression". The table lists 10 entries from the CSV file, each with its original text and the predicted hate speech category. The categories include various types of hate speech such as "BUKAN ujaran kebencian" and "Ujaran kebencian".

No	Dataset	Hate Speech
1	RT @yayenberpr: Fadli Zon Minta Mendagri Segera Memvonis Atau Jadi Gubernur DKI <a href="https://t.co/kHtsvRvhfO">https://t.co/kHtsvRvhfO</a>	BUKAN ujaran kebencian
2	RT @bagusindonesian: Membaca teman melukai atau dalam rangka memperjajakan Atau atau Atuk gagal dalam Ptkada	BUKAN ujaran kebencian
3	Syahrir bagaimana gubernur instakukan kekerasan perempuan! Buktinya banyak Bo2 meo foto bareng #DebatFinalPilkadaKet	BUKAN ujaran kebencian
4	Ahmad Dhani Tak Puas Detek Pilkada, Masalah Jalan di Bekasi Belum Sempat Terungkap <a href="https://t.co/lnqkpPLUDu2026">https://t.co/lnqkpPLUDu2026</a>	BUKAN ujaran kebencian
5	RT @Bintangary28: Waspada KTP palsu... Awas!! PTKADA <a href="https://t.co/vCDnRQv4AM">https://t.co/vCDnRQv4AM</a>	Ujaran kebencian
6	@Hiloksuwe match Impak hat ageng yg bisa ditahan, tempa ke pilkada tenus, apa lingkar kemaruan anda cuma setetas pilkada????	BUKAN ujaran kebencian
7	(@wondroid) apa makna pernyataan. Ambil Apa anda mengenali kriminalitas Anies? Hanif dg mudahmu pak! <a href="https://t.co/H4DzVdYt8m">https://t.co/H4DzVdYt8m</a>	BUKAN ujaran kebencian
8	Ahok Akan Integrasikan Transportasi Publik dengan Sistem Single Ticketingku2026 <a href="https://t.co/cZPhv7pgfVb">https://t.co/cZPhv7pgfVb</a> AhokTriboen	Ujaran kebencian
9	Ahok mampu melawan korupsi di daukota, bangun infrastruktur, merencana memperbaiki dengan berbagai insinuasi dan works well	BUKAN ujaran kebencian
10	Awas! Kecurangan Pilkada Agus - Syahrir Pakai Aplikasi Ini <a href="https://t.co/wZSw4nufUJ">https://t.co/wZSw4nufUJ</a>	BUKAN ujaran kebencian

Gambar 3.8 Tampilan Gambar *Dashboard* Hasil Prediksi Suatu File (.csv)

## 2. Fitur *Preprocessing*

Untuk fitur *preprocessing*, terdapat berbagai macam *preprocessing* antara lain yaitu :

### a) *CASEFOLDING*

Pada halaman *casefolding*, menampilkan dataset yang sebelum dilakukan *casefolding* dan sesudah dilakukan *casefolding* seperti gambar di bawah ini.

The screenshot shows a web interface titled 'Casefolding'. On the left, there's a sidebar with navigation links: Home, Pre-Processing, and Visualisasi dan Analisis. The main area has two tables: 'Sebelum Casefolding' and 'Sesudah Casefolding'. Both tables have columns for 'No', 'Text', and 'Action'. The 'Text' column contains 10 entries of Indonesian text. The 'Sesudah Casefolding' column shows the same text converted to lowercase. At the bottom, there are buttons for 'Previous', 'Next', and 'Activate Windows'.

No	Text (Sebelum Casefolding)	Text (Sesudah Casefolding)
1	RT (@pembalas: Fadli Zon Mintu Mendagi Segara Memnaktikan Ahok Jadi Gubemur DKI https://VA.co/0XfHvQwPdG)	spandeksi fadli zon mintu mendagi segara memnaktikan ahok jadi gubernur dki https://VA.co/0XfHvQwPdG
2	RT (@baguscondongwiso: Menika turu mletak akoi dalam rangka memperjakan Ahok atau Ahok gagal dalam Pilikada)	baguscondongwiso menika turu mletak akoi dalam rangka memperjakan ahok atau ahok gagal dalam pilkada
3	Sylvi Isagamane gubernur melakukan kekerasan perempuan? Buktinya banyak ibu masi foto bareng #DibebatPilkadaAKT	sylvi isagamane gubernur melakukan kekerasan perempuan? buktinya banyak ibu masi foto bareng #dibebatpilkadakt
4	Ahmad Dhani Tak Puan Dibat Pilikada, Masalah Jalan di Bekasi Belum Sempat Terungkap https://VA.co/0XfHvQwPd2E	ahmad dhani ta puas dibat pilkada, masalah jalan di bekasi belum sempat terungkap https://VA.co/0XfHvQwPd2E
5	RT (@Indulayu28: Waspada KTP palsu... kewal PILKADA https://VA.co/0OuERQV4SM	indulayu waspada ktp palsu... kewal pilkada https://VA.co/0OuERQV4SM
6	@Indulayu masih banyak hal agama yg bisa ditutup, kemoga ke pilkada turus, apa tingkat kemarahan anda cuma setetas pilkada????	@Indulayu masih banyak hal agama yg bisa ditutup, kemoga ke pilkada turus, apa tingkat kemarahan anda cuma setetas pilkada????
7	@nevel01 apakah maklumat pernyataan anda? Apa anda mengemukakan kriminalisasi Ahok?? Hati2 zg dg mulutmu pahl https://VA.co/0XfHvQwPdM	@nevel01 apakah maklumat pernyataan anda? Apa anda mengemukakan kriminalisasi ahok?? hati2 zg dg mulutmu pahl https://VA.co/0XfHvQwPdM
8	Ahok Akan Integrasi Transportasi Publik dengan Sistem Single TicketingJu2026 https://VA.co/42Ph7yqD0	Ahok akan integrasi transponasi publik dengan sistem single ticketingju2026 https://VA.co/42Ph7yqD0
9	Ahok mampu melawan korupsi di Buleleng, bangun infrastruktur, mencoba memperbaiki dengan berbagai tindakan dan works well.	Ahok mampu melawan korupsi di Buleleng, bangun infrastruktur, mencoba memperbaiki dengan berbagai tindakan dan works well.
10	Awas Kecurangan Pilikada, Agus - Sylvi Pakai Aplikasi Iri https://VA.co/0XfHvQwPdUfJU	Awas Kecurangan pilkada, agus - sylvi pakai aplikasi iri https://VA.co/0XfHvQwPdUfJU

Gambar 3.9 Tampilan Halaman *Preprocessing Casefolding*

### b) *Normalize alay*

Pada halaman *normalize alay*, menampilkan dataset yang sebelum dilakukan *normalize* dan sesudah dilakukan *normalize* seperti gambar di bawah ini.

The screenshot shows a web interface titled 'Normalize'. The layout is identical to the 'Casefolding' page, with a sidebar and two tables: 'Sebelum Normalize' and 'Sesudah Normalize'. The 'Text' column contains 10 entries of Indonesian text, and the 'Sesudah Normalize' column shows the text normalized to standard Indonesian. The bottom features 'Previous', 'Next', and 'Activate Windows' buttons.

No	Text (Sebelum Normalize)	Text (Sesudah Normalize)
1	RT (@pembalas: Fadli Zon Mintu Mendagi Segara Memnaktikan Ahok Jadi Gubemur DKI https://VA.co/0XfHvQwPdG)	RT (@pembalas: Fadli Zon Mintu Mendagi Segara Memnaktikan Ahok Jadi Gubemur DKI https://VA.co/0XfHvQwPdG)
2	RT (@baguscondongwiso: Menika turu mletak akoi dalam rangka memperjakan Ahok atau Ahok gagal dalam Pilikada)	RT (@baguscondongwiso: Menika turu mletak akoi dalam rangka memperjakan Ahok atau Ahok gagal dalam pilkada)
3	Sylvi Isagamane gubernur melakukan kekerasan perempuan? Buktinya banyak ibu2 mso foto bareng #DibebatPilkadaAKT	Sylvi Isagamane gubernur melakukan kekerasan perempuan? Buktinya banyak ibu masi foto bareng #dibebatpilkadakt
4	Ahmad Dhani Tak Puan Dibat Pilikada, Masalah Jalan di Bekasi Belum Sempat Terungkap https://VA.co/0XfHvQwPd2E	Ahmad Dhani Tak Puan Dibat Pilkada, Masalah Jalan di Bekasi Belum Sempat Terungkap https://VA.co/0XfHvQwPd2E
5	RT (@Indulayu28: Waspada KTP palsu... kewal PILKADA https://VA.co/0OuERQV4SM	RT (@Indulayu28: Waspada KTP palsu... kewal PILKADA https://VA.co/0OuERQV4SM
6	@Indulayu masih banyak hal agama yg bisa ditutup, kemoga ke pilkada turus, apa tingkat kemarahan anda cuma setetas pilkada????	@Indulayu masih banyak hal agama yg bisa ditutup, kemoga ke pilkada turus, apa tingkat kemarahan anda cuma setetas pilkada????
7	@nevel01 apakah maklumat pernyataan anda? Apa anda mengemukakan kriminalisasi Ahok?? Hati2 zg dg mulutmu pahl https://VA.co/0XfHvQwPdM	@nevel01 apakah maklumat pernyataan anda? Apa anda mengemukakan kriminalisasi ahok?? hati2 zg dg mulutmu pahl https://VA.co/0XfHvQwPdM
8	Ahok Akan Integrasi Transportasi Publik dengan Sistem Single TicketingJu2026 https://VA.co/42Ph7yqD0	Ahok akan integrasi transponasi publik dengan sistem single ticketingju2026 https://VA.co/42Ph7yqD0
9	Ahok mampu melawan korupsi di Buleleng, bangun infrastruktur, mencoba memperbaiki dengan berbagai tindakan dan works well.	Ahok mampu melawan korupsi di Buleleng, bangun infrastruktur, mencoba memperbaiki dengan berbagai tindakan dan works well.
10	Awas Kecurangan Pilikada, Agus - Sylvi Pakai Aplikasi Iri https://VA.co/0XfHvQwPdUfJU	Awas Kecurangan pilkada, agus - sylvi pakai aplikasi iri https://VA.co/0XfHvQwPdUfJU

Gambar 3.10 Tampilan Halaman *Preprocessing Normalize Alay*

### c) Stopword removal

Pada halaman *stopword removal*, menampilkan dataset yang sebelum dilakukan *stopword removal* dan sesudah dilakukan *stopword removal* seperti gambar di bawah ini.

The screenshot shows a web browser window titled 'Dashboard Preproses'. On the left sidebar, there are several tabs: 'Home', 'Predict', 'Hasil Prediksi Input', 'Hasil Prediksi Output', 'Pre-Processing', 'Hasil Pre-processing', and 'Visualisasi dan Analisis'. The main content area is titled 'Stopword Removal' and contains a table with 30 entries. The first few entries are:

- No. 1: RT (@pendeksy: Fadli Zon Mintu Mengabdi Segara Memerintahkan Ahok Jadi Gubernur DKI https://t.co/oXHsuRnID)
- No. 2: RT (@tagusconomicomew: Munkka lorus, muluk akid datang rangka memperparah Ahok atau Ahok gagal dalam Pilakada)
- No. 3: Syuh! Inggerisme gubernur melakukan kekerasan pemerasan? Buktinya banyak lbs2 misi foto barang #DobelFinalPilkadaKT
- No. 4: Ahmad Dhani Tak Puas Debat Pilakada, Masalah Jalan di Bekasi Belum Sempat Terungkap https://t.co/nHkqspPLQDwz2026
- No. 5: RT (@Indahsy28\_Waspada\_KTP\_pilaku...\_kawal\_PILKADA https://t.co/vQOoDRQV4SM)

Each entry has a detailed description in the right column, such as 'Sesudah Stopword Removal' and 'Search' bar at the top right.

Gambar 3.11 Tampilan Halaman *Preprocessing Stopword removal*

### d) Remove emoticon Byte

Pada halaman *remove emoticon*, menampilkan dataset yang sebelum dilakukan *remove emoticon* dan sesudah dilakukan *remove emoticon* seperti gambar di bawah ini.

The screenshot shows a web browser window titled 'Dashboard Preproses'. On the left sidebar, there are several tabs: 'Home', 'Predict', 'Hasil Prediksi Input', 'Hasil Prediksi Output', 'Pre-Processing', 'Hasil Pre-processing', and 'Visualisasi dan Analisis'. The main content area is titled 'Remove Emoticon' and contains a table with 30 entries. The first few entries are:

- No. 1: Selbelum Remove Emoticon
- No. 2: RT (@pendeksy: Fadli Zon Mintu Mengabdi Segara Memerintahkan Ahok Jadi Gubernur DKI https://t.co/oXHsuRnID)
- No. 3: RT (@tagusconomicomew: Munkka lorus, muluk akid datang rangka memperparah Ahok atau Ahok gagal dalam Pilakada)
- No. 4: Syuh! Inggerisme gubernur melakukan kekerasan pemerasan? Buktinya banyak lbs2 misi foto barang #DobelFinalPilkadaKT
- No. 5: Ahmad Dhani Tak Puas Debat Pilakada, Masalah Jalan di Bekasi Belum Sempat Terungkap https://t.co/nHkqspPLQDwz2026
- No. 6: RT (@Indahsy28\_Waspada\_KTP\_pilaku...\_kawal\_PILKADA https://t.co/vQOoDRQV4SM)

Each entry has a detailed description in the right column, such as 'Sesudah Remove Emoticon' and 'Search' bar at the top right.

Gambar 3.12 Tampilan Halaman *Preprocessing Remove emoticon*

### e) Stemming

Pada halaman *stemming*, menampilkan dataset yang sebelum dilakukan *stemming* dan sesudah dilakukan *stemming* seperti gambar di bawah ini.

No.	Sebelum Stemming	Sesudah Stemming
1	RT @spandeyc: Fadli Zon Minta Mendagri Segara Menonaktifkan Ahok Jadi Gubernur DKI https://t.co/kHbvRwPdO	rt spandeyc fadli zon minta mendagi segara nonaktif ahok jadi gubernur dk https t co khbvrwpd
2	RT @lagucondongmewo: Mereka terus melukuk akhir dalam rangka memperpanjang Ahok atau Ahok gagal dalam Pilpres	rt lagucondongmewo mereka terus luka akhir dalam rangka perpanjang ahok atau ahok gagal dalam pilkada
3	Syulf bagaimana gubernur melakukan kekerasan pemerasan? Buktinya banyak ibu2 mase foto boeng #DobutuInisiPilkadaKT	syulf bagaimana gubernur laku keras pemerasan buktinya banyak ibu2 mase foto boeng dobantuiniplikada
4	Ahmad Dhani Tak Pernah Dilehat Pilpres, Mewasiat Jalan di Bekasi Belum Sempat Terungkap https://t.co/lnqapLUDQn2020	ahmad dhani tak pernah dilihat pilkada mewasiat jalan di bekasi belum sempat ungkap https t co lnqapludqn2020
5	RT @Indonesiay28: Wikipedia KTP polisi... kawal PILKADA https://t.co/DOheRQv4SM	rt indonesiay28 wikipedia ktp polisi kawal pilkada https t co doherqv4sm
6	@Neliscause masih banyak hal agresif yg bisa diwasih, kenapa ke pilkada terus, apa tingkat keimanan anda cuma setiap pilkada??????	neliscause masih banyak hal agresif yg bisa diwasih kenapa ke pilkada terus apa tingkat iman anda cuma batas pilkada
7	@saedap! apa makna pernyataan, Anda? Apa anda menganggap kriminalisasi Ahok?? Hal2 yg dg mudah pakai? https://t.co/H4CQVnTfM4	saedap! apa makna pernyataan anda apa anda menganggap kriminalisasi ahok hal2 yg dg mudah pakai https t co h4cqvnfm4
8	Alok Akan Integrasi Transportasi Publik dengan Sistem Single Ticketinglu2020 https://t.co/4Z9nU7poy0B #WeTibus	alok akan integrasi transportasi publik dengan sistem single ticketinglu2020 https t co 4z9nu7poy0b infotibus
9	Alok mampu melawan korupsi di ikutlak, bangun infrastruktur, mencoba mempersiapkan dengan berbagi trobosan dan works well	alok mampu lawan korupsi di ikutlak bangun infrastruktur coba baik dengan berbagi trobosan dan works well
10	Awas Kecurangan Pilkada Agus - Syulf Pakai Aplikasi Iri https://t.co/WZWh4nUfRJ	awas curang pilkada agus - syulf pakai aplikasi iri https t co wzwh4nufrj

Gambar 3.13 Tampilan Tampilan Halaman *Preprocessing Stemming*

### 3. Fitur Hasil *Preprocessing*

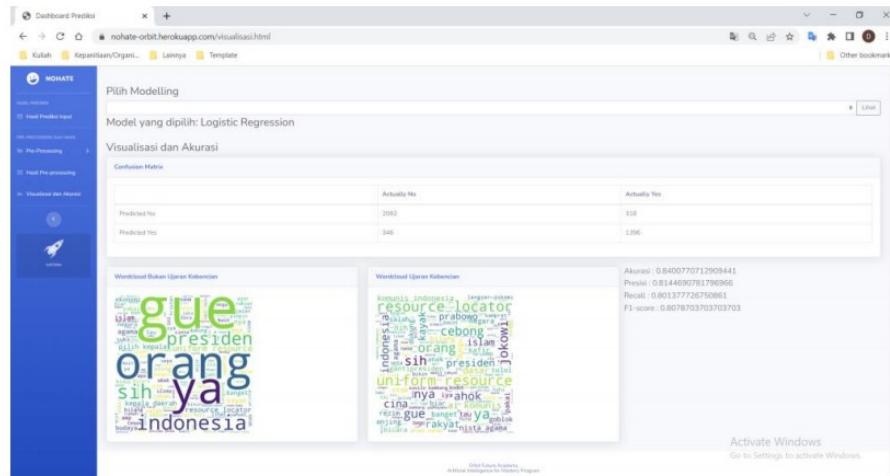
Pada halaman hasil *preprocessing*, menampilkan dataset yang sebelum dilakukan *preprocessing* dan sesudah dilakukan *preprocessing* seperti gambar di bawah ini. Hasil sesudah dilakukan *preprocessing* yang ditampilkan yaitu sesudah dilakukan berbagai macam *preprocessing* antara lain *casefolding*, *normalize alay*, *stopword removal*, *remove emoticon byte*, dan *stemming*.

No.	Data Sebelum Preprocessing	Data Sesudah Preprocessing
1	RT @spandeyc: Fadli Zon Minta Mendagri Segara Menonaktifkan Ahok Jadi Gubernur DKI https://t.co/kHbvRwPdO	spandeyc fadli zon minta mendagi segara nonaktif ahok jadi gubernur dk https t co khbvrwpd
2	RT @lagucondongmewo: Mereka terus melukuk akhir dalam rangka memperpanjang Ahok atau Ahok gagal dalam Pilpres	lagucondongmewo mereka terus luka akhir dalam rangka perpanjang ahok atau ahok gagal dalam pilkada
3	Syulf bagaimana gubernur melakukan kekerasan pemerasan? Buktinya banyak ibu2 mase foto boeng #DobutuInisiPilkadaKT	syulf bagaimana gubernur laku keras pemerasan buktinya banyak ibu2 mase foto boeng dobantuiniplikada
4	Ahmad Dhani Tak Pernah Dilehat Pilpres, Mewasiat Jalan di Bekasi Belum Sempat Terungkap https://t.co/lnqapLUDQn2020	ahmad dhani tak pernah dilihat pilkada mewasiat jalan di bekasi belum sempat ungkap https t co lnqapludqn2020
5	RT @Indonesiay28: Wikipedia KTP polisi... kawal PILKADA https://t.co/DOheRQv4SM	indonesiay28 wikipedia ktp polisi kawal pilkada https t co doherqv4sm
6	@Neliscause masih banyak hal agresif yg bisa diwasih, kenapa ke pilkada terus, apa tingkat keimanan anda cuma setiap pilkada??????	neliscause masih banyak hal agresif yg bisa diwasih kenapa ke pilkada terus apa tingkat iman anda cuma batas pilkada
7	@saedap! apa makna pernyataan, Anda? Apa anda menganggap kriminalisasi Ahok?? Hal2 yg dg mudah pakai? https://t.co/H4CQVnTfM4	saedap! apa makna pernyataan anda apa anda menganggap kriminalisasi ahok hal2 yg dg mudah pakai https t co h4cqvnfm4
8	Alok Akan Integrasi Transportasi Publik dengan Sistem Single Ticketinglu2020 https://t.co/4Z9nU7poy0B #WeTibus	alok akan integrasi transportasi publik dengan sistem single ticketinglu2020 https t co 4z9nu7poy0b infotibus
9	Alok mampu melawan korupsi di ikutlak, bangun infrastruktur, mencoba mempersiapkan dengan berbagi trobosan dan works well	alok mampu lawan korupsi di ikutlak bangun infrastruktur coba baik dengan berbagi trobosan dan works well
10	Awas Kecurangan Pilkada Agus - Syulf Pakai Aplikasi Iri https://t.co/WZWh4nUfRJ	awas curang pilkada agus - syulf pakai aplikasi iri https t co wzwh4nufrj

Gambar 3.14 Tampilan Halaman Hasil *Preprocessing* dari semua proses (*casefolding*, *normalize alay*, *stopword removal*, *remove emoticon byte*, dan *stemming*)

#### 4. Fitur Visualisasi dan Akurasi

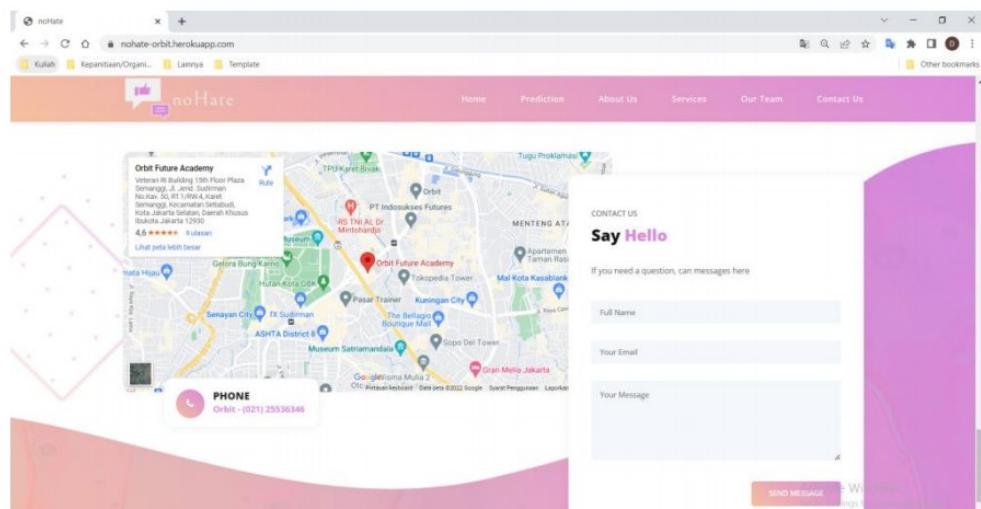
Menampilkan *confusion matrix*, dan *metrics* (akurasi, presisi, *Recall*, *F1-Score*) dari setiap *modelling* yang dipilih. Serta menampilkan *wordcloud* dari ujaran kebencian dan bukan ujaran kebencian seperti gambar di bawah ini.



Gambar 3.15 Tampilan Halaman Visualisasi dan Akurasi

#### 5. Fitur *Contact us*

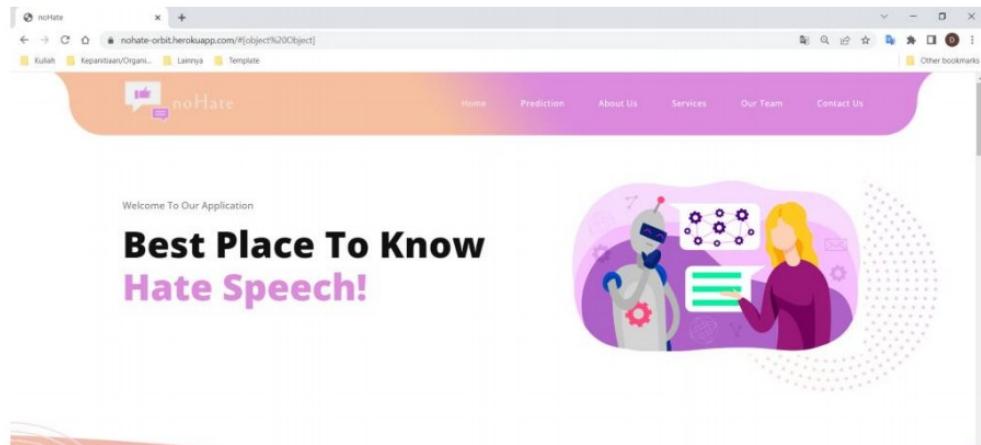
Fitur ini digunakan apabila pengguna ingin bertanya atau memberikan saran dapat mengetik pesan beserta nama dan email. Berikut adalah tampilan *contact us* aplikasi noHate.



Gambar 3.16 Tampilan Halaman *Contact us*

## 6. Home

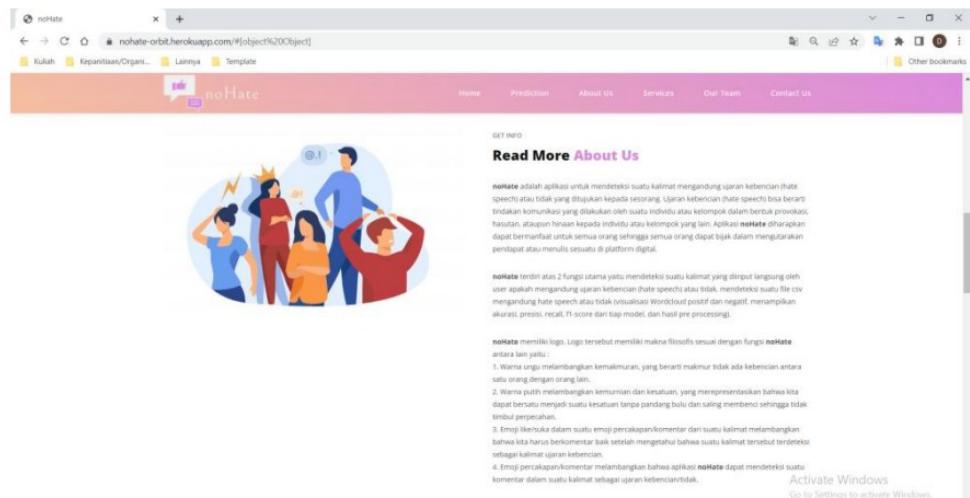
Halaman awal pada aplikasi **noHate**



Gambar 3.17 Tampilan Halaman *Home*

## 7. About us

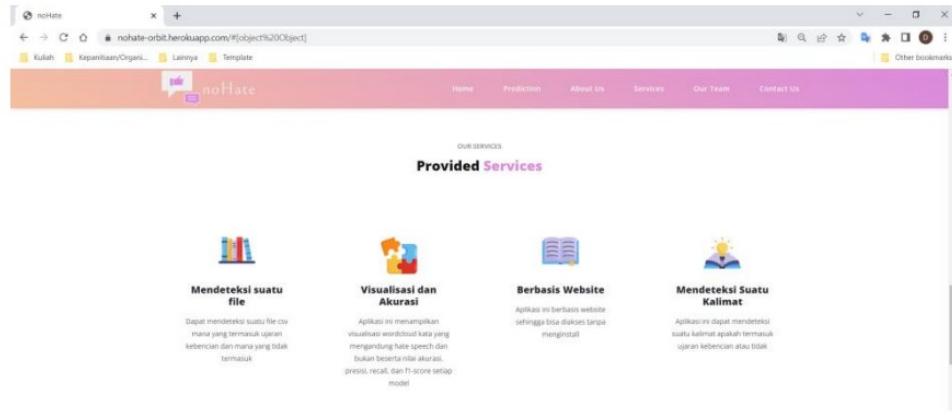
Menampilkan informasi terkait aplikasi **noHate** mulai dari apa itu aplikasi **noHate**, harapan aplikasi **noHate**, fungsi aplikasi **noHate**, dan makna filosofis dari logo aplikasi **noHate**.



Gambar 3.18 Tampilan Halaman *About us*

## 8. Services

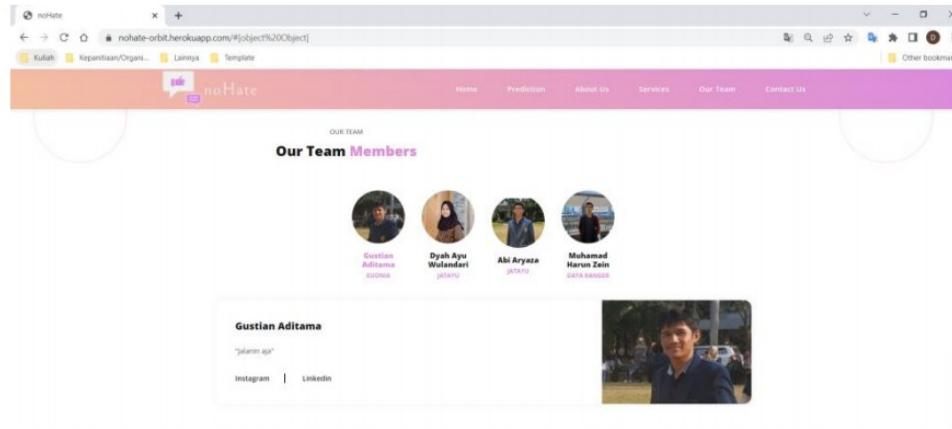
Menampilkan apa saaja *services/layanan* yang disediakan oleh aplikasi **noHate**



Gambar 3.19 Tampilan Halaman *Services*

## 9. Our team

Menampilkan anggota tim yang membuat aplikasi **noHate** mulai dari foto, nama, *quotes* dan media sosial (Instagram dan linkedIn).



Gambar 3.20 Tampilan Halaman *Our team*

Berikut adalah hasil evaluasi dari beberapa model :

### 1. Logistic Regression

#### a. Classification Report

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.85	0.87	0.86	2426
1	0.81	0.78	0.80	1726
accuracy			0.83	4152
macro avg	0.83	0.83	0.83	4152
weighted avg	0.83	0.83	0.83	4152

Gambar 3.21 Classification Report Logistic Regression

b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[2114 312]
 [ 381 1345]]
True Positif : 1345
True Negatif : 2114
False Positif : 312
False Negatif : 381
```

Gambar 3.22 *Confusion matrix Logistic Regression*

c. *Metrics* :

- AUC : 0.9100836677591827
- Akurasi : 0.833092485549133
- Presisi : 0.8117079058539529
- Recall : 0.7792584009269988
- F1-Score: 0.7951522317469701

2. *Support Vector Machine*

a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.84	0.91	0.87	2426
1	0.85	0.76	0.80	1726
accuracy			0.85	4152
macro avg	0.85	0.83	0.84	4152
weighted avg	0.85	0.85	0.84	4152

Gambar 3.23 *Classification Report Support Vector Machine*

b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[2199 227]
 [ 414 1312]]
True Positif : 1312
True Negatif : 2199
False Positif : 227
False Negatif : 414
```

Gambar 3.24 *Confusion matrix Support Vector Machine*

c. *Metrics* :

- AUC : 0.9193512679842457

- Akurasi : 0.845616570327553
- Presisi : 0.852501624431449
- *Recall* : 0.7601390498261877
- *F1-Score*: 0.8036753445635528

### 3. *Naive Bayes (Bernoulli)*

#### a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.88	0.81	0.84	2426
1	0.76	0.84	0.80	1726
accuracy			0.82	4152
macro avg	0.82	0.82	0.82	4152
weighted avg	0.83	0.82	0.82	4152

Gambar 3.25 *Classification Report* Naïve Bayes (*Bernoulli*)

#### b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[1966 460]
 [ 278 1448]]
```

```
True Positif : 1448
True Negatif : 1966
False Positif : 460
False Negatif : 278
```

Gambar 3.26 *Confusion matrix* Naïve Bayes (*Bernoulli*)

#### c. *Metrics* :

- AUC : 0.9085597653462538
- Akurasi : 0.8222543352601156
- Presisi : 0.7589098532494759
- *Recall* : 0.8389339513325609
- *F1-Score*: 0.7969179966978537

### 4. *Decision Tree*

#### a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.83	0.80	0.81	2426
1	0.73	0.77	0.75	1726
accuracy			0.79	4152
macro avg	0.78	0.79	0.78	4152
weighted avg	0.79	0.79	0.79	4152

Gambar 3.27 Classification Report Decision Tree

#### b. Confusion matrix

Dengan Confusion Matrix:

```
[[1935 491]
 [ 389 1337]]

True Positif : 1337
True Negatif : 1935
False Positif : 491
False Negatif : 389
```

Gambar 3.28 Confusion matrix Decision Tree

#### c. Metrics :

- AUC : 0.7860112397654228
- Akurasi : 0.7880539499036608
- Presisi : 0.7314004376367614
- Recall : 0.7746234067207416
- F1-Score: 0.7523916713562183

### 5. Stacking

#### a. Classification Report

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.87	0.88	0.88	2426
1	0.83	0.82	0.82	1726
accuracy			0.86	4152
macro avg	0.85	0.85	0.85	4152
weighted avg	0.85	0.86	0.85	4152

Gambar 3.29 Classification Report Stacking

#### b. Confusion matrix

Dengan Confusion Matrix:

```
[[2131 295]
 [ 307 1419]]
```

```
True Positif : 1419
True Negatif : 2131
False Positif : 295
False Negatif : 307
```

Gambar 3.30 *Confusion matrix Stacking*

c. *Metrics* :

- AUC : 0.9274677618575895
- Akurasi : 0.855009633911368
- Presisi : 0.8278879813302217
- *Recall* : 0.8221320973348784
- *F1-Score*: 0.825

Kelebihan aplikasi **noHate** adalah :

1. Dapat memprediksi suatu kalimat apakah termasuk ujaran kebencian atau tidak baik dari inputan kalimat maupun file yang diunggah oleh pengguna dengan berbagai model (5 model).
2. Dapat melihat hasil *preprocessing* (*casefolding*, *normalize* alay, *stopword removal*, *remove emoticon byte*, dan *stemming*) serta dapat melihat visualisasi dan akurasi (*wordcloud*, *confusion matrix*, akurasi, presisi, *Recall*, *F1-Score*) dengan berbagai model (5 model).
3. Dapat diakses dengan mudah karena tidak perlu menginstall aplikasi
4. Bebas *platform* karena penggunaan aplikasi web tidak terbatas pada satu atau dua sistem operasi saja sehingga semua *device* dan semua web browser bisa digunakan untuk mengakses aplikasi **noHate**.

Kekurangan aplikasi **noHate** adalah :

1. Sumber daya yang terbatas sehingga setelah aplikasi *dideployment* hanya bisa mendeteksi 30 baris dan hanya bisa mendeteksi dalam format .csv
2. Dapat mendeteksi suatu kalimat termasuk ujaran kebencian atau tidak hanya dengan menggunakan Bahasa Indonesia.

Rencana pengembangan aplikasi **noHate** di masa depan adalah :

1. Dapat mendeteksi lebih dari 1 bahasa/bermacam-macam Bahasa
2. Dapat melakukan pembobotan setiap kata sehingga dalam suatu kalimat dapat dilakukan pembobotan apakah termasuk ujaran kebencian atau tidak atau bahkan hanya berupa sindiran sehingga dapat memahami suatu makna kata dengan benar.
3. Dapat menambahkan algoritma klasifikasi *deep learning*

## Bab IV      Penutup

### IV.1 Kesimpulan

Dalam pelaksanaan Studi Independen di PT. Orbit Ventura Indonesia dengan topik *Artificial Intelligence Mastery Program* membahas beberapa domain yaitu *Data Science*, *Natural Language Processing*, *Computer Vision*, dan *Reinforcement Learning*. Setelah pembahasan materi selesai maka dilanjutkan pembuatan *project* akhir, berdasarkan hasil *project* yang dilakukan dapat diambil kesimpulan yaitu sistem dapat mendeteksi suatu kalimat termasuk ujaran kebencian atau bukan dengan menggunakan inputan kalimat atau file dengan format .csv yang diunggah, dapat memilih *modelling* yang akan digunakan untuk mendeteksi suatu kalimat ujaran kebencian atau bukan, dapat menampilkan hasil *preprocessing*, serta dapat menampilkan visualisasi dan akurasi. Aplikasi **noHate** dapat diakses dengan mudah dan bebas *platform* karena tidak terbatas pada satu atau dua sistem operasi saja sehingga semua *device* dan semua web browser bisa digunakan untuk mengakses aplikasi **noHate**. Dari hasil evaluasi diketahui bahwa hasil akurasi menggunakan model *Logistic Regression* adalah 0.833092485549133, *Support Vector Machine* adalah 0.845616570327553, *Naive Bayes (Bernoulli)* adalah 0.8222543352601156, *Decision Tree* adalah 0.7880539499036608, dan *Stacking* adalah 0.855009633911368. Selain itu, dilakukan pengujian dan dapat diketahui bahwa fungsi-fungsi aplikasi **noHate** dapat berjalan dengan benar dan sesuai.

### IV.2 Saran

Berdasarkan hasil kesimpulan yang didapatkan terdapat beberapa saran yang ditunjukkan untuk penelitian selanjutnya. Diharapkan penelitian selanjutnya untuk dapat mendeteksi lebih dari 1 bahasa/bermacam-macam Bahasa, dapat melakukan pembobotan setiap kata sehingga dalam suatu kalimat dapat dilakukan pembobotan apakah termasuk ujaran kebencian atau tidak atau bahkan hanya berupa sindiran sehingga dapat memahami suatu makna kata dengan benar. Implementasi pembelajaran ini sesuai dengan tujuan PT. Orbit Ventura Indonesia yaitu bukan

hanya untuk memperkenalkan teknologi AI ke pelajar, tetapi juga untuk memungkinkan bisa mengangkat perangkat AI, sehingga bisa membuat sesuatu yang menciptakan dampak sosial, *project* ini diharapkan menjadi referensi untuk penelitian selanjutnya dalam perkembangan digital saat ini. Penulis menyarankan agar persiapan dalam pelaksanaan studi independen ini dapat dimaksimalkan untuk hasil yang lebih maksimal.

## Bab V Referensi

- [ W. Kwan, “APA ITU *NATURAL LANGUAGE PROCESSING?*,” 16 Februari 1 2021. [Online]. Available: <https://algorit.ma/blog/natural-language-processing/>. ]
- [ R. Nuryama, “TiNewss.Com,” 14 June 2022. [Online]. Available: 2 <https://www.tinewss.com/indonesia-news/pr-1853617810/jumlah-pengguna->] media-sosial-di-indonesia-pada-tahun-2022#:~:text=TiNewss.Com%20%2D%20Ada%20191%2C,9%20persen%20dari%20total%20populasi.. [Accessed 12 June 2022].
- [ A. prokomsetda, “Pemerintah Kabupaten Buleleng Bagian Protokol dan Komunikasi Pimpinan,” 16 Januari 2017. [Online]. Available: ] [https://prokomsetda.bulelengkab.go.id/informasi/detail/artikel/hate-speech-36#:~:text=Hate%20Speech%20\(Ucapan%20Penghinaan%2Fatau,cacat%2C%20orientasi%20seksual%2Ckewarganegaraan%2C.](https://prokomsetda.bulelengkab.go.id/informasi/detail/artikel/hate-speech-36#:~:text=Hate%20Speech%20(Ucapan%20Penghinaan%2Fatau,cacat%2C%20orientasi%20seksual%2Ckewarganegaraan%2C.)
- [ nurisman, “Melawan Ujaran Kebencian dalam Pemilu,” 12 Januari 2018. 4 [Online]. Available: <https://www.bawaslu.go.id/id/berita/melawan-ujaran->] kebencian-dalam-pemilu.
- [ E. Safitri, “Bawaslu RI: Makna Ujaran Kebencian Masih Multitafsir,” 15 Februari 2019. [Online]. Available: <https://news.detik.com/berita/d-4429440/bawaslu-ri-makna-ujaran-kebencian-masih-multitafsir.>

## **Bab VI            Lampiran A. TOR**

**TERM OF REFERENCE (TOR)**  
**STUDI INDEPENDEN BERSERTIFIKAT**  
**AI MASTERY PROGRAM**  
**DI ORBIT FUTURE ACADEMY**

### **A. Rincian Program**

AI *Mastery Program* adalah program pelatihan *Artificial Intelligence* (AI) daring yang bertujuan untuk memperkenalkan teknologi dan perangkat AI kepada pelajar, sehingga diharapkan mereka dapat mengembangkan produk AI yang memiliki dampak sosial. Program ini berfokus pada komponen utama AI, seperti *Data Science* (DS), *Natural Language Processing* (NLP), *Computer Vision* (CV), dan *Reinforcement Learning* (RL).

### **B. Tujuan Program**

Tujuan yang diharapkan setelah peserta mengikuti program ini:

1. Mampu memahami apa itu AI, penerapan dan pemanfaatannya.
2. Mampu memahami terkait tiga domain utama AI (DS, NLP, dan CV).
3. Mampu mengelaborasi kemampuan terkait AI dengan bidang lain.
4. Mampu memahami pentingnya data dalam AI.
5. Mampu membuat *project* AI yang berdampak sosial.
6. Mampu menulis kode dengan bahasa pemrograman Python.
7. Mampu memahami operasi dan logika sederhana pada Python.
8. Mampu membuat *project* Python.
9. Mampu melakukan kolaborasi secara interaktif dengan Git/Github.
10. Mampu membuat *repository* di akun Git/Github.
11. Mampu membuat portfolio dengan Git/Github.
12. Mampu menganalisis algoritma *Machine Learning* (ML) yang paling sesuai.
13. Mampu membuat model ML.

14. Mampu memahami dan menerapkan algoritma ML untuk membantu kehidupan.
15. Mampu membuat model *Deep learning* (DL).
16. Mampu membuat kode program untuk pengujian model *Data Science*.
17. Mampu melakukan pengujian model dan analisis.
18. Mampu membuat ramalan dan prediksi berdasarkan data.
19. Mampu mengolah data yang besar untuk membuat keputusan.
20. Mampu men-clustering untuk memetakan pola.
21. Mampu membuat dokumentasi hasil pengujian model DS.
22. Mampu Membuat model DS dengan ML & DL.
23. Memahami NLP.
24. Mampu membuat model pengenalan suara.
25. Mampu membuat chatbot.
26. Mampu membuat *project* terkait dengan RL.
27. Mampu mengkombinasikan dan membuat *project* terkait AI, IoT, dan sensor.
28. Mampu mengaplikasikan konsep RL dan diterapkan bersama domain AI lain.
29. Mampu memahami dan membuat *project* terkait CV.
30. Mampu menggunakan teknologi terkait *Computer Vision*.
31. Mampu mengembangkan *project* CV untuk kepentingan sosial.
32. Mampu membuat model ML dan DL untuk berbagai kasus.
33. Mampu men-deploy model menggunakan Heroku dan atau menggunakan layan *Machine Learning as a service* (MLaaS).

### C. Jadwal dan Tempat Pelaksanaan

Jadwal pelaksanaan tertera dalam tabel berikut:

Pukul (WIB)	Durasi (jam)	Aktivitas
08.00 s.d. 11.30	3.5	Kelas Sesi Pagi
13.00 s.d. 16.30	3.5	Kelas Sesi Siang
16.30 s.d. 17.30	1	Self-Study

Kelas akan diselenggarakan secara daring melalui aplikasi *video conference*.

### D. Peserta

Peserta program ini adalah mahasiswa yang berasal dari Perguruan Tinggi di bawah Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

### E. Uraian Tugas Peserta

Selama mengikuti program ini, peserta diharuskan:

1. Mengikuti program dari awal hingga selesai.
2. Mematuhi aturan program.
3. Mematuhi aturan kelas yang dibuat bersama *homeroom* atau *domain coach*.
4. Mengikuti kelas dengan presensi minimal 85%.
5. Membuat laporan harian dan mingguan di *website* Kampus Merdeka.
6. Menyelesaikan Proyek Akhir (PA) beserta laporannya.

*Homeroom Coach,*



Eko Purbo Wahyono  
NIP 2201069

Ngawi, 21 Februari 2022

Peserta Program,



Dyah Ayu Wulandari  
195150201111062

## Bab VII      Lampiran B. Log Activity

Minggu/Tgl	Kegiatan	Hasil
9 Mei 2022	Mengumpulkan asset yang diperlukan untuk <i>front end</i> dan mencari referensi untuk tambahan fitur pada aplikasi	Asset untuk <i>front end</i> sudah terkumpul dan referensi sudah didapatkan
10 Mei 2022	Mencari referensi untuk membuat model	Referensi untuk membuat model sudah didapatkan
11 Mei 2022	Membuat power point untuk bimbingan dan melakukan bimbingan kepada coach domain <i>Natural Language Processing</i> (NLP) mengenai ide dan saran tambahan fitur untuk <i>project akhir</i>	Mendapat masukan saran dan contoh aplikasi <i>Natural Language Processing</i>
12 Mei 2022	Membuat <i>front end</i> pada tampilan <i>landing page</i>	Tampilan <i>front end landing page</i> bagian <i>home, prediction, dan about us</i> telah berhasil dibuat
13 Mei 2022	Membuat <i>front end</i> pada tampilan <i>landing page</i>	Tampilan <i>front end landing page</i> bagian <i>services, our team, contact us</i> telah berhasil dibuat
16 Mei 2022	Libur Hari Raya Waisak	Libur Hari Raya Waisak
17 Mei 2022	Membuat <i>wordcloud</i> ujaran kebencian dan bukan ujaran kebencian	<i>Wordcloud</i> sudah dibuat

18 Mei 2022	Mencari Referensi Model yang tepat untuk kasus ujaran kebencian	Referensi tambahan telah didapatkan
19 Mei 2022	Membuat model <i>Decision Tree</i>	Model <i>Decision Tree</i> sudah berhasil dibuat
20 Mei 2022	Membuat tampilan kasar front-end <i>dashboard</i> yang berisi nama-nama fitur dari hasil referensi yang telah didapat serta membuat logo	Tampilan kasar <i>front end dashboard</i> sudah dibuat dan logo sudah dibuat
23 Mei 2022	Melakukan presentasi progress ke coach domain dan mengadakan rapat kelompok untuk membahas mengenai fitur-fitur dalam aplikasi	Presentasi progress telah dilakukan dan sudah mendapatkan pandangan terkait fitur yang akan diexplore
24 Mei 2022	Membuat tampilan front-end <i>dashboard</i> hasil prediksi input ( <i>dashboard.html</i> )	Tampilan <i>front end dashboard</i> hasil prediksi input ( <i>dashboard.html</i> ) telah berhasil dibuat
25 Mei 2022	Membuat tampilan front-end <i>dashboard</i> bagian <i>casefolding, normalize, stopword removal</i> ( <i>casefolding.html, normalize.html, stopword-removal.html</i> )	Tampilan <i>front end dashboard</i> bagian <i>casefolding, normalize, stopword removal</i> ( <i>casefolding.html, normalize.html, stopword-removal.html</i> )
26 Mei 2022	Libur Kenaikan Isa al Masih	Libur Kenaikan Isa al Masih
27 Mei 2022	Membuat tampilan front-end <i>dashboard</i> bagian <i>remove emoticon byte, stemming, hasil preprocessing, visualisasi dan akurasi</i>	Tampilan front-end <i>dashboard</i> bagian <i>remove emoticon byte, stemming, hasil</i>

	(removeemoticon.html, <i>stemming</i> .html, hasilpreprocess.html, visualisasi.html)	<i>preprocessing</i> , visualisasi dan akurasi removeemoticon.html, <i>stemming</i> .html, hasilpreprocess.html, visualisasi.html)
30 Mei 2022	Menambahkan kodingan ensemble-method yaitu <i>Stacking</i>	Model <i>Stacking</i> telah berhasil dibuat
31 Mei 2022	Melakukan presentasi progressan	Presentasi sudah dilakukan dan mendapatkan masukan
1 Juni 2022	Libur Hari Pancasila	Libur Hari Pancasila
2 Juni 2022	Mengadakan rapat kelompok untuk mencari solusi dari permasalahan backend	Rapat kelompok telah diadakan dan masih mencari solusi lebih lanjut
3 Juni 2022	Mencari solusi dari suatu permasalahan <i>back end</i>	Solusi sudah ditemukan
6 Juni 2022	Membuat laporan <i>project</i> akhir	Beberapa bagian laporan sudah dibuat
7 Juni 2022	Mengadakan rapat kelompok membahas mengenai permasalahan tampilan dan beberapa fitur pada website	Rapat sudah dilakukan dan beberapa fitur pada <i>dashboard</i> sudah berfungsi
8 Juni 2022	Melakukan konsultasi ke coach domain NLP terkait evaluasi model yang digunakan (overfitting/underfitting/tidak) dan terkait visualisasi data	Mendapatkan saran dan jawaban atas evaluasi model dan visualisasi data

9 Juni 2022	Melanjutkan pembuatan website	Beberapa fitur sudah berhasil dibuat
10 Juni 2022	Membuat laporan <i>project</i> akhir	Beberapa bagian laporan <i>project</i> akhir telah dibuat
13 Juni 2022	Melakukan presentasi progressan terkait aplikasi yang sudah berfungsi dengan sesuai	Menerima masukan terkait aplikasi yang sudah dibuat
14 Juni 2022	Mengadakan rapat kelompok untuk finalisasi terkait website dan dilakukan <i>deployment</i>	Website telah berhasil dibuat dan sudah dilakukan <i>deployment</i>
15 Juni 2022	Membuat laporan <i>project</i> akhir	Beberapa bagian laporan <i>project</i> akhir telah dibuat
16 Juni 2022	Melakukan presentasi kelompok terkait <i>project</i> akhir aplikasi <b>noHate</b> dan membuat laporan <i>project</i> akhir	Presentasi kelompok telah dilakukan dan laporan <i>project</i> akhir telah selesai
17 Juni 2022	Meminta tanda tangan untuk laporan <i>project</i> akhir dan mengumpulkan laporan <i>project</i> akhir	Laporan <i>project</i> akhir telah dibuat dan dikumpulkan

## Bab VIII      Lampiran C. Dokumen Teknik

### 1. AI Project Cycle

#### a. Problem Scoping

Jumlah pengguna media sosial di Indonesia pada awal tahun 2022 setara dengan 68,9 persen dari total populasi sehingga semakin banyak kebebasan berekspresi masyarakat dalam menyampaikan pendapat/argumennya dalam berbagai bidang seperti politik, ekonomi, budaya, sosial, kesehatan, dan lain-lain [2]. Kebebasan berpendapat tersebut dapat disampaikan dalam bentuk ekspresi positif seperti puji, atau dapat juga disampaikan dalam ekspresi negatif seperti ujaran kebencian (*hate speech*). Dalam bidang politik, *hate speech* sering menjadi alat untuk menghalangi proses demokrasi bahkan menjatuhkan pihak lawan. Berdasarkan bawaslu.go.id, anggota Bawaslu yaitu Mochammad Afifuddin mengatakan dalam kampanye Pemilu dapat menyediakan lahan subur untuk ujaran kebencian baik ujaran kebencian yang ditujukan kepada pejabat terpilih, partai politik, kandidat atau calon, masyarakat sipil [4]. Penyebar kebencian bisa dari berbagai kelompok, mulai dari anak-anak, remaja dan dewasa yang sudah dapat menggunakan sosial media untuk berpendapat.

Selain itu, pada detik.com menjelaskan bahwa anggota Badan Pengawas Pemilu (Bawaslu) RI Mochammad Afifuddin menilai definisi ‘ujaran kebencian’ masih memunculkan perbedaan dalam perspektif masyarakat dan masih multitafsir sehingga menjadi tantangan untuk menyatukan pemahaman agar tercipta batasan dalam ujaran kebencian (Safitri, 2019). Berdasarkan hal tersebut, penulis membuat “**noHate** : Aplikasi Deteksi Kalimat Ujaran Kebencian Berbasis Website Menggunakan Kombinasi Model *Machine Learning*” yang menggunakan Bahasa Indonesia dengan metode *Logistic Regression*, *Support Vector Machine*, *Naive Bayes (Bernoulli)*, *Decision Tree*, dan *Stacking*. Dengan aplikasi tersebut, maka masyarakat awam akan dapat membedakan kalimat yang termasuk ujaran kebencian atau tidak dalam

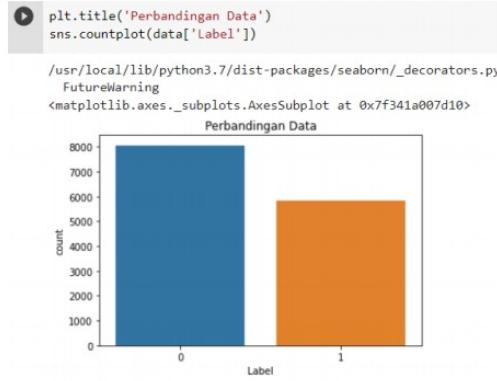
kebebasan berpendapat sehingga tidak menimbulkan diskriminasi, pengucilan, dan lain-lain dalam suatu individu atau golongan.

### **b. Data Acquisition**

Sumber data yang digunakan berasal dari github (<https://github.com/ialfina/id-hatespeech-detection>) dan Kaggle (<https://www.kaggle.com/datasets/ilhamfp31-indonesian-abusive-and-hate-speech-twitter-text>). Dataset ini berupa komentar/ tweet yang menggunakan Bahasa Indonesia. Dataset pada github berisi 713 baris dan 2 kolom yaitu kolom label (Non\_HS untuk non-hate-speech yang akan diberi label 0 dan HS untuk hate-speech yang akan diberi label 1) dan tweet. Sedangkan, dataset pada kaggle berisi 13.169 baris dan 13 kolom tetapi yang diperlukan hanya 2 kolom yaitu kolom Tweet dan HS (*hate speech* label). Label yang terdapat pada dataset yaitu 0 (bukan ujaran kebencian) dan 1 (ujaran kebencian). Dataset dari github dan kaggle tersebut didownload dan dilakukan *preprocessing*, *modelling*, dan evaluasi di google colaboratory. Sebelum dilakukan *preprocessing*, karena dataset dari github berformat .txt sehingga perlu diubah/dikonversi ke format .xlsx terlebih dahulu, sedangkan dataset dari Kaggle berformat .csv. Berikut link google colaboratory yang digunakan (<https://colab.research.google.com/drive/1fItgmAnFIIn9HIJtSo7LFxmYBo94L43o?usp=sharing>).

### **c. Data Exploration**

Setelah data digabung, maka berikut adalah visualisasi perbandingan banyaknya data yang merupakan ujaran kebencian (label 1) dan bukan ujaran kebencian (label 0). Dari gambar di bawah ini, dapat diketahui bahwa dataset yang merupakan ujaran kebencian (label 1) memiliki jumlah 5812 baris, sedangkan dataset yang merupakan bukan ujaran kebencian (label 0) memiliki jumlah 8025 baris.



Gambar 8.1 Perbandingan Dataset Ujaran Kebencian (label 1) dan Bukan Ujaran Kebencian (label 0)

#### d. *Modelling*

Yang pertama dilakukan yaitu membuat fungsi untuk menampilkan hasil evaluasi dan menyimpan model dengan parameter data text yang sudah dilakukan *vectorization*, label, dan model yang sudah dilakukan *training*. Lalu, dilakukan prediksi model dengan data text yang sudah dilakukan *vectorization* serta dilakukan prediksi probabilitas, lalu mencari *confusion matrix* (TP, TN, FP, FN), AUC, akurasi, presisi, *Recall*, dan *F1-Score*, serta dilakukan penyimpanan model dengan nama file yang diinginkan. Setelah itu, model, akurasi, presisi, *Recall*, dan *F1-Score* dilakukan dump joblib. Format kodingan *modelling* di bawah ini akan dipanggil pada setiap *modelling* untuk mengetahui hasil dari evaluasi masing-masing model.

```
• Modelling
[ ] def finalize(X,y,model): #menampilkan hasil evaluasi dan menyimpan model (optional)
    #X = data text yang sudah dilakukan vectorization
    #y = label
    #model = Objek model yang sudah dilakukan training
    y_predict = model.predict(X)
    y_prob = model.predict_proba(X)[:,1]
    fpr, tpr, thresholds = roc_curve(y, y_prob)
    roc_auc = auc(fpr, tpr)
    tn, fp, fn, tp = confusion_matrix(y, y_predict)
    acc = accuracy_score(y, y_predict)
    prec = precision_score(y, y_predict)
    rec = recall_score(y, y_predict)
    f1 = f1_score(y, y_predict)
    cm = confusion_matrix(y, y_predict)
    tn, fp, fn, tp = cm.ravel()

    print('*****\n*****\n*****\n*****\n*****\n*****\n*****')
    print('Dengan Hasil Model:\n*****')
    print('Dengan Classification Report:\n', classification_report(y,y_predict))
    print('Dengan Confusion Matrix:\n',cm)
    print('True Positif : ', tp)
    print('True Negatif : ', tn)
    print('False Positif : ', fp)
    print('False Negatif : ', fn)
    print('*****\n*****\n*****\n*****\n*****\n*****')

    print('Dengan Beberapa Metrik:')
    print('Akurasi : ', acc)
    print('AUC : ', roc_auc)
    print('Presisi : ', prec)
    print('Recall : ', rec)
    print('F1-score : ', f1)

    ans = input('\nIngin menyimpan model? [y/n]')
    if ans.lower() == 'y':
        print('\nMenyimpan model dan beberapa metrik. Masukkan nama file yang diinginkan:')
        name = input()
        joblib.dump((model, acc, prec, rec, f1), filename=name)
        print('\nModel tersimpan dengan nama file', name)

    print('*****\n*****\n*****\n*****\n*****\n*****\n*****')
```

Gambar 8.2 *Modelling*

Aplikasi **noHate** menggunakan 5 *modelling*, antara lain yaitu :

1. *Logistic Regression*

Model *Logistic Regression* yang digunakan yaitu binary *Logistic Regression*, yaitu digunakan untuk mengklasifikasikan suatu objek. Binary *Logistic Regression* hanya menyediakan dua kemungkinan jawaban. Konsep ini biasanya direpresentasikan sebagai 0 atau 1 dalam pengkodean. Dalam *project* deteksi ujaran kebencian ini untuk menilai apakah termasuk ujaran kebencian yang bernilai 1 dan bukan ujaran kebencian yang bernilai 0.

2. *Support Vector Machine*

*Support Vector Machine* (SVM) merupakan salah satu metode dalam *supervised learning* yang biasanya digunakan untuk mengklasifikasikan sebuah kelas (seperti Support Vector Classification). SVM digunakan untuk mencari hyperplane terbaik dengan memaksimalkan jarak antar kelas. Jadi, pada *project* ini dengan menggunakan model SVM dapat mengklasifikasikan tweet yang termasuk ujaran kebencian dan bukan ujaran kebencian.

3. *Naive Bayes (Bernoulli)*

*Bernoulli Naïve Bayes* di mana fitur diasumsikan biner (0s dan 1s). *Bernoulli Naïve Bayes* dalam klasifikasinya lebih berfokus pada hasil ya/tidak. Prediktor yang di-input adalah variabel Boolean sehingga satunya nilai yang ada adalah benar atau salah. Dalam *project* ini, dapat memprediksi atas sebuah kalimat termasuk ujaran kebencian atau tidak.

4. *Decision Tree*

*Decision Tree* adalah salah satu metode klasifikasi untuk mem-break down proses pengambilan keputusan yang kompleks menjadi lebih simple, sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan. Jadi, dengan model ini bisa mendapatkan solusi apakah suatu kalimat termasuk ujaran kebencian atau tidak. Pembuatan *Decision Tree* sendiri menggunakan metode *supervised Machine Learning* yaitu proses pembelajaran dimana data baru diklasifikasikan berdasarkan

*training* samples yang sudah ada. Untuk melakukan pelatihan dan evaluasi, dilakukan dengan cara memanggil model *Decision Tree Classifier* dan dilakukan fit dengan data *train* yang sudah divectorization tf-idf dan data *y\_train*. Setelah itu, dilakukan memanggil fungsi *finalize* yang sudah dibuat pada *modelling* sebelumnya, dengan parameter data *test* yang sudah dilakukan *vectorization* tf-idf, *y\_test* dan model *Decision Tree* seperti gambar di bawah ini.

```

    ▾ Decision Tree
    Q.
    [x] ▾ Pelatihan dan Evaluasi
    □
    [ ] dt = tree.DecisionTreeClassifier()
        dt.fit(X_train_vectors_tfidf, y_train)
        DecisionTreeClassifier()

    [ ] *finalize(X_test_vectors_tfidf,y_test,dt)
    =====
    Evaluasi Hasil Model
    =====

    Dengan Classification Report:
    precision    recall    f1-score   support
    0            0.83     0.80      0.81     2426
    1            0.73     0.77      0.75     1726
    accuracy                           0.79     4152
    macro avg       0.78     0.79      0.78     4152
    weighted avg    0.79     0.79      0.79     4152
    =====

    Dengan Confusion Matrix:
    [[1337 491]
     [ 389 1337]]
    True Positif : 1337
    True Negatif : 1337
    False Positif : 491
    False Negatif : 389
    =====

    Dengan Beberapa Metrik:
    AUC : 0.7860112397054228
    Akurasi : 0.7860112397054228
    Presisi : 0.7746244673676114
    Recall : 0.7746244673676114
    F1-score : 0.7523916713562183
    Ingin menyimpan model? [y/n/y]
    Penyimpanan model dan beberapa metrik. Masukkan nama file yang diinginkan:
    dt
    Model tersimpan dengan nama file dt.
    =====
    Selesai
    =====

```

Gambar 8.3 *Modelling Decision Tree*

## 5. Stacking

*Stacking* melibatkan pemasangan banyak jenis model yang berbeda (dalam *project* ini model *Logistic Regression*, *Support Vector Machine*, naïve bayes (*Bernoulli*), dan *Decision Tree*) pada data yang sama dan menggunakan model lain untuk mempelajari cara terbaik menggabungkan prediksi. Untuk melakukan pelatihan dan evaluasi, dilakukan membuat variable *base\_model* yang terdiri 4 macam model (*Logistic Regression*, *Support Vector Machine*, *naïve bayes* (*Bernoulli*), dan *Decision Tree*) serta dilakukan *Stacking* dengan memanggil model *Stacking Classifier* dengan

parameter estimator, final estimator dan cv. Lalu, dilakukan fit dengan data *train* yang sudah *vectorization* tf-idf dan data *y\_train*. Setelah itu, dilakukan memanggil fungsi *finalize* yang sudah dibuat pada *modelling* sebelumnya, dengan parameter data *test* yang sudah dilakukan *vectorization* tf-idf, *y\_test* dan model *stacked* seperti gambar di bawah ini.

```
[ ] base_models = [
    ('Logistic Regression', lr),
    ('Support Vector Machine', svm),
    ('Naive Bayes', nb),
    ('Decision Tree', dt),
]

stacked = StackingClassifier(
    estimators = base_models,
    final_estimator = lr,
    cv = 5)

stacked.fit(X_train_vectors_tfidf, y_train)

StackingClassifier(cv=5,
    estimators=[('Logistic Regression',
        LogisticRegression(C=10, solver='liblinear')),
        ('Support Vector Machine',
        SVC(C=1, probability=True)),
        ('Naive Bayes', BernoulliNB()),
        ('Decision Tree', DecisionTreeClassifier())),
    final_estimator=LogisticRegression(C=10, solver='liblinear'))
```

```
[ ] finalize(X_test_vectors_tfidf,y_test,stacked)
=====
Evaluasi Hasil Model
=====

Dengan Classification Report:

```

	precision	recall	f1-score	support
0	0.87	0.88	0.88	2426
1	0.83	0.82	0.82	1726

```
accuracy           0.86
macro avg       0.85
weighted avg    0.85
```

```
=====
Dengan Confusion Matrix:
[[2131 295]
 [ 307 1419]]
```

```
True Positif : 1419
True Negatif : 2131
False Positif : 295
False Negatif : 307
```

```
=====
Dengan Beberapa Metrik:

```

AUC	Akurasi	Presisi	Recall	F1-score
: 0.9274677618575895	: 0.852005613911168	: 0.8879813302117	: 0.8211320973348784	: 0.825

```
Ingin menyimpan model? [y/n]y
Menyimpan model dan beberapa metrik. Masukkan nama file yang diinginkan:
stacking
Model tersimpan dengan nama file stacking
=====
Selesai
=====
```

Gambar 8.4 *Modelling Stacking*

#### e. Evaluation

Dalam evaluasi beberapa model, terdapat *confusion matrix* yang ditampilkan, *confusion matrix* merepresentasikan prediksi dan kondisi sebenarnya(aktual) dari data yang dihasilkan oleh algoritma ML. Berdasarkan *Confusion matrix*, kita bisa menentukan *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Akurasi adalah rasio prediksi benar (positif dan

negatif) dengan keseluruhan data atau tingkat kedekatan antara nilai prediksi dengan nilai aktual. Presisi adalah rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. *Recall* adalah rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. *F1-Score* merupakan perbandingan rata-rata presisi dan *Recall* yang dibobotkan. Selain itu, terdapat AUC (*Area Under the Curve*). Semakin tinggi AUC semakin bagus model tersebut.

Berikut adalah hasil evaluasi dari beberapa model :

### 1. *Logistic Regression*

#### a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.85	0.87	0.86	2426
1	0.81	0.78	0.80	1726
accuracy			0.83	4152
macro avg	0.83	0.83	0.83	4152
weighted avg	0.83	0.83	0.83	4152

Gambar 8.5 *Classification Report Logistic Regression*

#### b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[2114  312]
 [ 381 1345]]
```

```
True Positif : 1345
True Negatif : 2114
False Positif : 312
False Negatif : 381
```

Gambar 8.6 *Confusion matrix Logistic Regression*

#### c. *Metrics* :

- ✓ AUC : 0.9100836677591827
- ✓ Akurasi : 0.833092485549133
- ✓ Presisi : 0.8117079058539529
- ✓ Recall : 0.7792584009269988
- ✓ F1-Score: 0.7951522317469701

### 2. *Support Vector Machine*

a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.84	0.91	0.87	2426
1	0.85	0.76	0.80	1726
accuracy			0.85	4152
macro avg	0.85	0.83	0.84	4152
weighted avg	0.85	0.85	0.84	4152

Gambar 8.7 Classification Report Support Vector Machine

b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[2199 227]
 [ 414 1312]]
```

```
True Positif : 1312
True Negatif : 2199
False Positif : 227
False Negatif : 414
```

Gambar 8.8 Confusion matrix Support Vector Machine

c. *Metrics* :

- ✓ AUC : 0.9193512679842457
- ✓ Akurasi : 0.845616570327553
- ✓ Presisi : 0.852501624431449
- ✓ Recall : 0.7601390498261877
- ✓ F1-Score: 0.8036753445635528

3. *Naive Bayes (Bernoulli)*

a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.88	0.81	0.84	2426
1	0.76	0.84	0.80	1726
accuracy			0.82	4152
macro avg	0.82	0.82	0.82	4152
weighted avg	0.83	0.82	0.82	4152

Gambar 8.9 Classification Report Naïve Bayes (Bernoulli)

b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[1966 460]
 [ 278 1448]]
True Positif : 1448
True Negatif : 1966
False Positif : 460
False Negatif : 278
```

Gambar 8.10 *Confusion matrix Naïve Bayes (Bernoulli)*

c. *Metrics* :

- ✓ AUC : 0.9085597653462538
- ✓ Akurasi : 0.8222543352601156
- ✓ Presisi : 0.7589098532494759
- ✓ *Recall* : 0.8389339513325609
- ✓ *F1-Score*: 0.7969179966978537

#### 4. *Decision Tree*

a. *Classification Report*

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.83	0.80	0.81	2426
1	0.73	0.77	0.75	1726
accuracy			0.79	4152
macro avg	0.78	0.79	0.78	4152
weighted avg	0.79	0.79	0.79	4152

Gambar 8.11 *Classification Report Decision Tree*

b. *Confusion matrix*

Dengan Confusion Matrix:

```
[[1935 491]
 [ 389 1337]]
True Positif : 1337
True Negatif : 1935
False Positif : 491
False Negatif : 389
```

Gambar 8.12 *Confusion matrix Decision Tree*

c. *Metrics* :

- ✓ AUC : 0.7860112397654228
- ✓ Akurasi : 0.7880539499036608

- ✓ Presisi : 0.7314004376367614
- ✓ Recall : 0.7746234067207416
- ✓ F1-Score: 0.7523916713562183

## 5. Stacking

### a. Classification Report

Dengan Classification Report:

	precision	recall	f1-score	support
0	0.87	0.88	0.88	2426
1	0.83	0.82	0.82	1726
accuracy			0.86	4152
macro avg	0.85	0.85	0.85	4152
weighted avg	0.85	0.86	0.85	4152

Gambar 8.13 Classification Report Stacking

### b. Confusion matrix

Dengan Confusion Matrix:

```
[[2131 295]
 [ 307 1419]]
True Positif : 1419
True Negatif : 2131
False Positif : 295
False Negatif : 307
```

Gambar 8.14 Confusion matrix Stacking

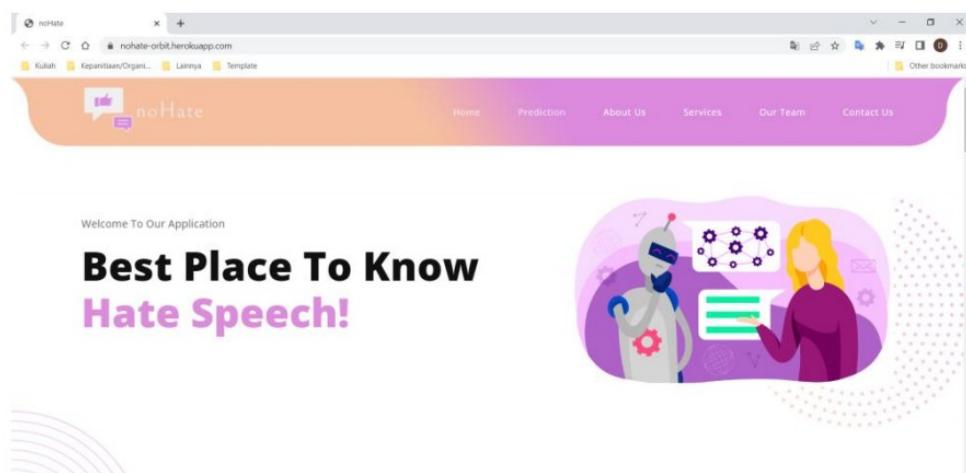
### c. Metrics :

- ✓ AUC : 0.9274677618575895
- ✓ Akurasi : 0.855009633911368
- ✓ Presisi : 0.8278879813302217
- ✓ Recall : 0.8221320973348784
- ✓ F1-Score: 0.825

### f. Deployment

*Deployment* aplikasi merupakan proses menyalin, mengkonfigurasi, dan mengaktifkan aplikasi tertentu ke URL dasar tertentu di sebuah server atau di cluster sehingga tujuan *deployment* setelah proses penerapan selesai maka aplikasi tersebut menjadi dapat diakses publik di URL dasar. Berikut adalah

link aplikasi **noHate** : <https://noHate-orbit.herokuapp.com/>. Deployment project ini menggunakan Heroku. Proses deployment aplikasi terdiri dari beberapa Langkah-langkah, antara lain yaitu : Registrasi Akun di Heroku, Membuat Aplikasi Baru, Menginstall heroku CLI, melakukan clone Repository dari Github, Login ke Heroku, melakukan push Local Repository ke Heroku, lalu membuka Aplikasi yang Berhasil di-Deploy. Berikut adalah gambar aplikasi **noHate** yang berhasil dideployment.



Gambar 8.15 Tampilan *Landing page* Aplikasi **noHate**

## 2. Profil Tim dan Deskripsi Pembagian Tugas

Tim *project* akhir aplikasi **noHate** beranggotakan 4 orang, antara lain yaitu :

- Gustian Aditama dari kelas Euonia
- Dyah Ayu Wulandari dari kelas Jatayu
- Abi Aryaza dari kelas Jatayu
- Muhamad Harun Zein dari kelas Data Ranger

Semua anggota tim kami berasal dari Universitas Brawijaya. Berikut adalah pembagian peran dalam *project* akhir aplikasi **noHate** :

- Tugas Awal :
  - Gustian : Mencari dataset dan melakukan import dataset
  - Abi : Melakukan *preprocessing* (*normalize alay, stopword removal, remove emoticon*)
  - Harun : Melakukan *preprocessing* (*Casefolding, Stemming*)

- Dyah : Membuat *wordcloud* dari label 0 (bukan ujaran kebencian) dan label 1 (ujaran kebencian)
- b. Tugas Tengah :
- Gustian : Membuat *vectorization (TF-IDF)*
  - Abi : Membuat *modelling (Logistic Regression, Support Vector Machine, Naive Bayes)*
  - Dyah : Membuat *modelling (Decision Tree, Stacking)*
  - Harun : Membuat *testing* dengan kalimat masukan di setiap model
- c. Tugas Akhir :
- Gustian : Melakukan evaluasi model dan demonstrasi di colab
  - Dyah : Membuat *front end* dari aplikasi **noHate**
  - Abi : Membuat *back end*
  - Harun : Membuat *back end* dan melakukan *deployment*

### 3. Deskripsi Aplikasi

#### a. Nama dan Fungsi Aplikasi

**noHate** adalah aplikasi untuk mendeteksi suatu kalimat mengandung ujaran kebencian (*hate speech*) atau tidak yang ditujukan kepada seseorang. Ujaran kebencian (*hate speech*) bisa berarti tindakan komunikasi yang dilakukan oleh suatu individu atau kelompok dalam bentuk provokasi, hasutan, ataupun hinaan kepada individu atau kelompok yang lain. Target user dari aplikasi **noHate** adalah pengguna media sosial baik dari kalangan anak-anak, remaja dan dewasa. Oleh karena itu, harapannya aplikasi **noHate** dapat bermanfaat untuk semua orang sehingga semua orang dapat bijak dalam mengutarakan pendapat atau menulis sesuatu di *platform* digital.

**noHate** terdiri atas 2 fungsi utama yaitu mendeteksi suatu kalimat yang diinput langsung oleh user apakah mengandung ujaran kebencian (*hate speech*) atau tidak, dan mendeteksi suatu file csv mengandung *hate speech* atau tidak serta terdapat visualisasi *Wordcloud* ujaran kebencian dan bukan

ujaran kebencian, menampilkan akurasi, presisi, *Recall*, *F1-Score* dari tiap model, dan hasil pre processing.



Gambar 8.16 Logo Aplikasi **noHate**

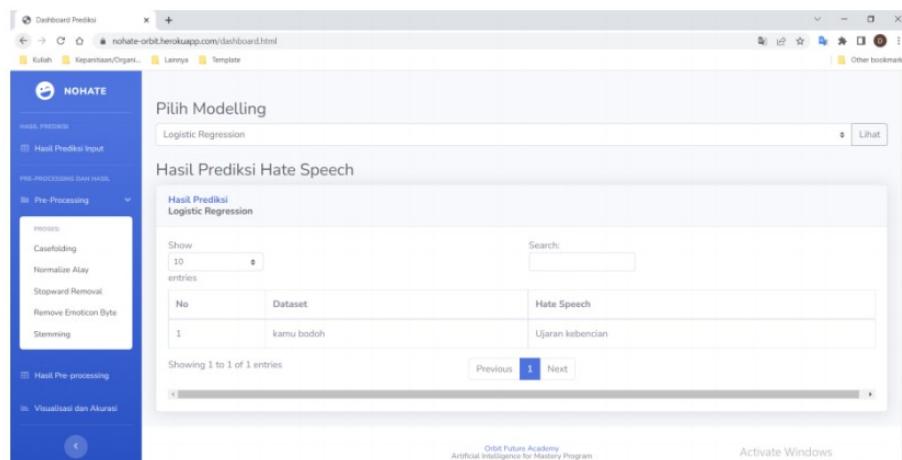
Aplikasi **noHate** memiliki logo, gambar logo aplikasi seperti gambar di atas. Logo tersebut memiliki makna filosofis sesuai dengan fungsi **noHate** antara lain yaitu :

1. Warna ungu melambangkan kemakmuran, yang berarti makmur tidak ada kebencian antara satu orang dengan orang lain.
2. Warna putih melambangkan kemurnian dan kesatuan, yang merepresentasikan bahwa kita dapat bersatu menjadi suatu kesatuan tanpa pandang bulu dan saling membenci sehingga tidak timbul perpecahan.
3. Emoji like/suka dalam suatu emoji percakapan/komentar dari suatu kalimat melambangkan bahwa kita harus berkomentar baik setelah mengetahui bahwa suatu kalimat tersebut terdeteksi sebagai kalimat ujaran kebencian.
4. Emoji percakapan/komentar melambangkan bahwa aplikasi **noHate** dapat mendeteksi suatu komentar dalam suatu kalimat sebagai ujaran kebencian/tidak.

Berikut adalah cara kerja aplikasi **noHate** :

- a. Masukkan input untuk mendeteksi apakah termasuk ujaran kebencian atau tidak. Apabila input yang dimasukkan adalah suatu kalimat dan file maka akan muncul kalimat “Choose one method!!”. Apabila tidak menginput apapun maka akan muncul kalimat “Please input text or file!!”.
- b. Setelah memasukkan input di *landing page* maka akan diarahkan ke *dashboard* setelah proses deteksi selesai.

- c. Selanjutnya, dapat memilih *modelling* yang akan digunakan untuk deteksi, di aplikasi **noHate** terdapat 5 model yaitu *Logistic Regression*, *SVM*, *Naïve Bayes*, *Decision Tree*, dan *Stacking*.
- d. Setelah memilih model maka dapat menekan tombol “Lihat” untuk melihat hasil yang ingin diprediksi.
- e. Lalu, hasil prediksi dapat dilihat. Berikut di bawah ini adalah tampilan prediksi. Selain itu, terdapat beberapa fitur di *dashboard* yang dapat dilihat seperti *preprocessing* (*casefolding*, *normalize alay*, *stopword removal*, *remove emoticon byte*, *stemming*), hasil pre-processing, dan visualisasi dan akurasi.



Gambar 8.17 Tampilan *Dashboard* Hasil Prediksi

### b. Jenis Aplikasi dan *Specific Requirement*

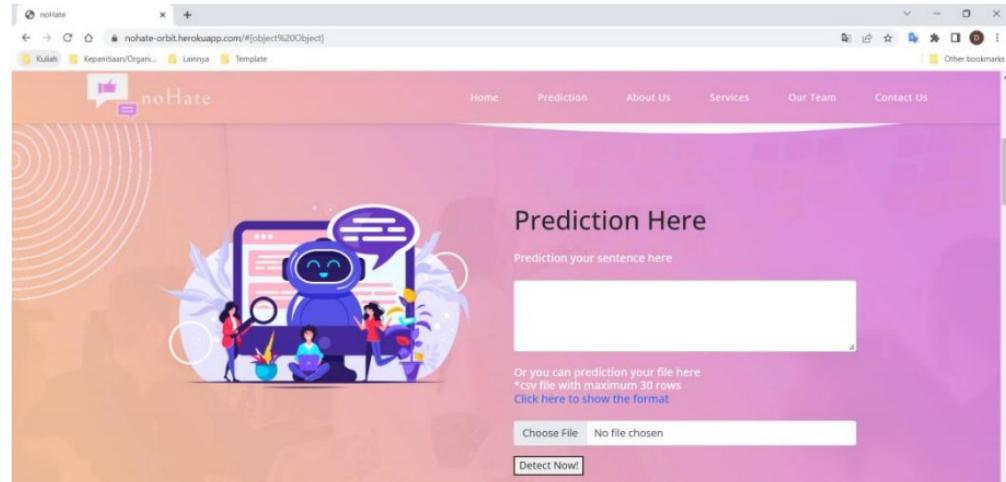
Aplikasi **noHate** adalah aplikasi berbasis website sehingga dapat diakses di handphone maupun di laptop dengan menggunakan browser. Aplikasi ini membutuhkan akses internet yang memadai serta untuk memprediksi file yang berisi komentar ujaran kebencian atau bukan hanya dapat diproses jika file yang diunggah dengan format .csv.

### c. *User Interface*

Aplikasi **noHate** memiliki beberapa fitur, antara lain yaitu :

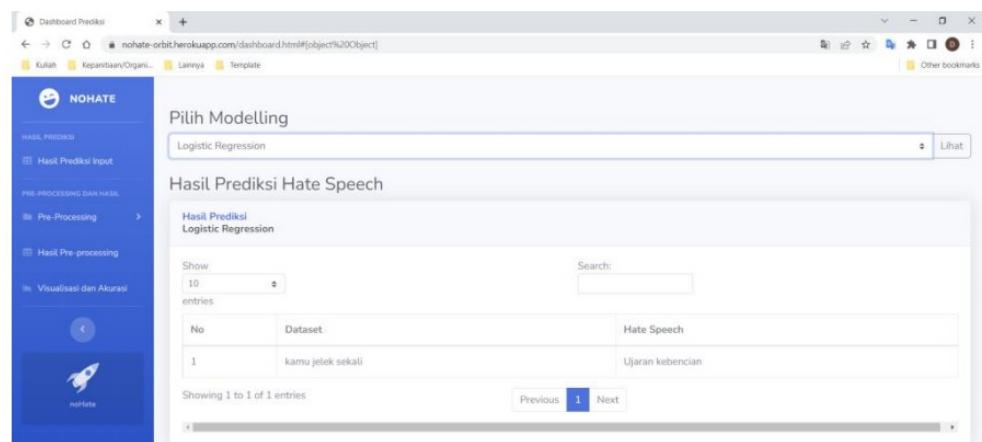
1. Fitur *Prediction*

Dalam fitur ini, pengguna dapat memprediksi kalimat baik berupa inputan kalimat atau file dengan format .csv yang diunggah.



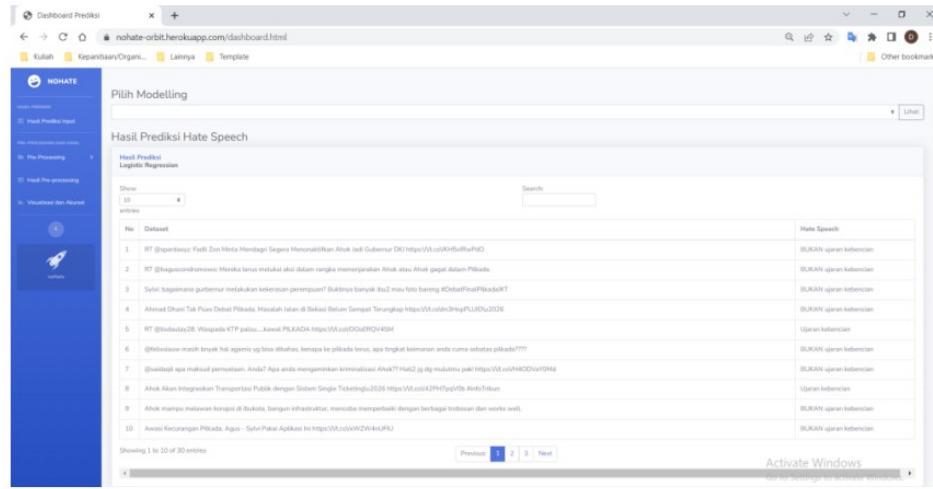
Gambar 8.18 Tampilan Fitur *Prediction*

Berikut adalah tampilan setelah memasukkan kalimat “kamu jelek sekali” lalu menekan tombol “Detect Now!” dan diarahkan ke halaman *dashboard*. Setelah itu, memilih model yang digunakan dan menekan “Lihat”.



Gambar 8.19 Tampilan Gambar *Dashboard* Hasil Prediksi Input Suatu Kalimat

Berikut adalah tampilan setelah memasukkan file dengan format .csv lalu menekan tombol “Detect Now!” dan diarahkan ke halaman *dashboard*. Setelah itu, memilih model yang digunakan dan menekan “Lihat”.



Gambar 8.20 Tampilan Gambar *Dashboard Hasil Prediksi Suatu File (.csv)*

## 2. Fitur *Preprocessing*

Untuk fitur *preprocessing*, terdapat berbagai macam *preprocessing* antara lain yaitu :

### a. *CASEFOLDING*

Pada halaman *casefolding*, menampilkan dataset yang sebelum dilakukan *casefolding* dan sesudah dilakukan *casefolding* seperti gambar di bawah ini.

Casefolding	
Nos. 1 : Sebelum Casefolding	Sesudah Casefolding
1. RT @spesialis_Fadli: Zon Minta Mendangi Segara Memerlukan Ahok Jadi Gubernur DKI https://t.co/kHfS0hvePhD	spesialis fadli zon minta mendangi segara memerlukan ahok jadi gubernur dk https://t.co/kHfS0hvePhD
2. RT @tagarcondongkuning: Mereka terus merilaku akhir dalam rangka memerlukan Ahok atau Ahok gagal dalam Pilkada	tagarcondongkuning mereka terus merilaku akhir dalam rangka memerlukan ahok atau ahok gagal dalam pilkada
3. Suhu bagaimana gubernur melakukan kelecekan perempuan? Buktinya banyak dsb2 mau foto bareng #DewiFirdiaPilkadaDKT	suhu bagaimana gubernur melakukan kelecekan perempuan? buktinya banyak dsb2 mau foto bareng #DewiFirdiaPilkadaDKT
4. Ahmad Dhani Tak Puas Dapat Pilkada, Haruslah Jalan di Belakang Belum Sempat Terungkap https://t.co/3mEqPfUDtw2026	ahmad dhani tak puas dapat pilkada, haruslah jalan di belakang belum sempat terungkap https://t.co/3mEqPfUDtw2026
5. RT @Indonesiatv28: Waspada KTP palsu... Jurnal PIALADA https://t.co/4QODDvVq9M	indonesiatv28 waspada ktp palsu... jurnal pilkada https://t.co/4QODDvVq9M
6. @Hilmanz masih banyak hal agresif yg bisa ditulis, kerapa ke pilkada terso, apa tingkat keimanan anda cuma setebal pilkada????	hilmanz masih banyak hal agresif yg bisa ditulis, kerapa ke pilkada terso, apa tingkat keimanan anda cuma setebal pilkada????
7. @calonjaya apa maklum perempuan? Andi? Apa anda mengemukakan kriminasi Ahok?? Hati2 jg dg mulutmu pake https://t.co/4QODDvVq9M	calonjaya apa maklum perempuan? andi? apa anda mengemukakan kriminasi ahok?? hati2 jg dg mulutmu pake https://t.co/4QODDvVq9M
8. Ahok Akan Integrasi Transportasi Publik dengan Sistem Single Ticketinghttps://t.co/col42PNFay@Bh_RhezaHibuu	ahok akan integrasi transportasi publik dengan sistem single ticketinghttps://t.co/col42PNFay@Bh_RhezaHibuu
9. Ahok mampu melawan korupsi di Balotela, bangun infrastruktur, mencipta memperbaiki dengan berbagai bidang dan works well.	ahok mampu melawan korupsi di balotela, bangun infrastruktur, mencipta memperbaiki dengan berbagai bidang dan works well.
10. Awas Keciranngan Pilkada, Agus - Syah Pakar Aplikasi ke https://t.co/nZWh4uLPUJ	awas keciranngan pilkada, agus - syah pakar aplikasi ke https://t.co/nZWh4uLPUJ

Gambar 8.21 Tampilan Halaman *Preprocessing Casefolding*

### b. *Normalize alay*

Pada halaman *normalize alay*, menampilkan dataset yang sebelum dilakukan *normalize* dan sesudah dilakukan *normalize* seperti gambar di bawah ini.

The screenshot shows a web application interface titled 'Dashboard Prediksi'. On the left, there's a sidebar with navigation items: 'Home', 'Model', 'Host Predict Input', 'Pre Processing', 'Host Pre-processing', and 'Visualisasi dan Analisis'. The main content area has a title 'Normalize' and a sub-section 'Data Normalizing'. It displays two tables: 'Sebelum Normalize' and 'Sesudah Normalize'. The 'Sebelum Normalize' table lists 10 tweets in their original, unprocessed form. The 'Sesudah Normalize' table lists the same 10 tweets, where each tweet has been processed to remove punctuation, stop words, and emoticons. Below the tables, there are sections for 'Activate Windows' and 'Show 1 to 10 of 30 entries'.

No.	Sebelum Normalize	Sesudah Normalize
1	RT @spesialistik: Fadli Zon Mintu Mendagri Segara Memerlukan Ahok Jadi Gubernur DKI <a href="https://t.co/XHf5llwvH0">https://t.co/XHf5llwvH0</a>	RT @spesialistik: Fadli Zon Mintu Mendagri Segera Memerlukan Ahok Jadi Gubernur DKI <a href="https://t.co/XHf5llwvH0">https://t.co/XHf5llwvH0</a>
2	RT @BogorGubernormous: Mereka terus merusak akik dalam rangka memperparah Ahok atau Ahok gagal dalam Pilkada	RT @BogorGubernormous: Mereka terus merusak akik dalam rangka memperparah Ahok atau Ahok gagal dalam Pilkada
3	Sylvi Isappramee gubernur melakukan kekerasan pemimpin? Buktinya banyak bls2 atau foto bierung #Detail#FinalPilkadaKT	Sylvi Isappramee gubernur melakukan kekerasan pemimpin? Buktinya banyak bls2 atau foto bierung #Detail#FinalPilkadaKT
4	Ahmad Dhani Tak Penuh Dukat Pilkada, Masalah Jalan di Bekasi Belum Sempat Terungkap <a href="https://t.co/PLUDzu2206">https://t.co/PLUDzu2206</a>	Ahmad Dhani Tak Penuh Dukat Pilkada, Masalah Jalan di Bekasi Belum Sempat Terungkap <a href="https://t.co/PLUDzu2206">https://t.co/PLUDzu2206</a>
5	RT @Endahay28: Waspada KTP polisi... kawal PILKADA <a href="https://t.co/ERQVq45H">https://t.co/ERQVq45H</a>	RT @Endahay28: Waspada KTP polisi... kawal PILKADA <a href="https://t.co/ERQVq45H">https://t.co/ERQVq45H</a>
6	@hendriksw masih banyak hal aheng yg bisa dibahas, kerapa ke pilkada tonas, apa tingkat komarman anda cuma sebutas pilkada????	@hendriksw masih banyak hal aheng yg bisa dibahas, kerapa ke pilkada tonas, apa tingkat komarman anda cuma sebutas pilkada????
7	@hendriksw apa makna pernyataan. Anda? Apa anda mengaminkan kriminalisasi Ahok?? Hsl2 zg dg mulutmu pasti <a href="https://t.co/H4QD9VY0M4">https://t.co/H4QD9VY0M4</a>	@hendriksw apa makna pernyataan. Anda? Apa anda mengaminkan kriminalisasi Ahok?? Hsl2 zg dg mulutmu pasti <a href="https://t.co/H4QD9VY0M4">https://t.co/H4QD9VY0M4</a>
8	Ahok Akan Integrasi Transportasi Publik dengan Sistem Single Ticketinglu2026 <a href="https://t.co/42Ph7jeV0n">https://t.co/42Ph7jeV0n</a>	Ahok Akan Integrasi Transportasi Publik dengan Sistem Single Ticketinglu2026 <a href="https://t.co/42Ph7jeV0n">https://t.co/42Ph7jeV0n</a>
9	Ahok insompi melawan kompi di biukita, biangk infrastruktur, merencana memperbaiki dengan berbagai trodosen dan works well.	Ahok insompi melawan kompi di biukita, biangk infrastruktur, merencana memperbaiki dengan berbagai trodosen dan works well.
10	Awas Kocurangan Pilkada, Agus - Sylvi Pakai Aplikasi ini <a href="https://t.co/oWZWh4HJFU">https://t.co/oWZWh4HJFU</a>	Awas Kocurangan Pilkada, Agus - Sylvi Pakai Aplikasi ini <a href="https://t.co/oWZWh4HJFU">https://t.co/oWZWh4HJFU</a>

Gambar 8.22 Tampilan Halaman *Preprocessing Normalize Alay*

### c. Stopword removal

Pada halaman *stopword removal*, menampilkan dataset yang sebelum dilakukan *stopword removal* dan sesudah dilakukan *stopword removal* seperti gambar di bawah ini.

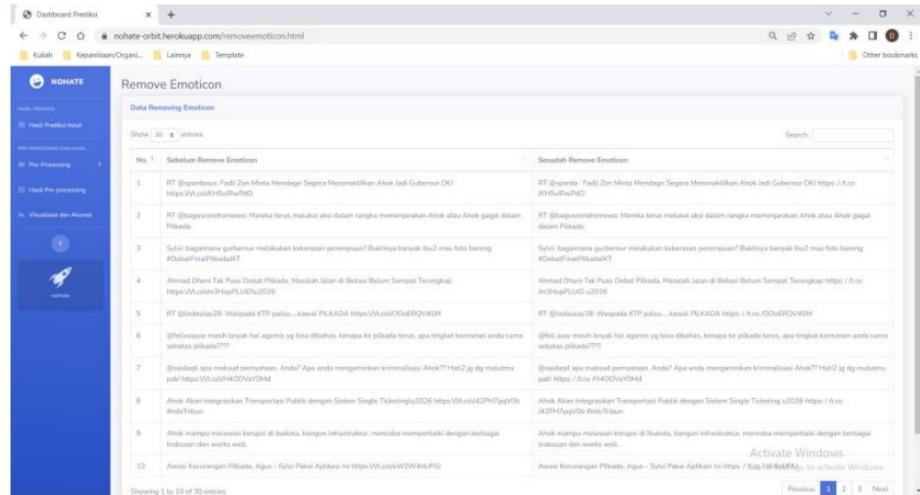
The screenshot shows a web application interface titled 'Dashboard Prediksi'. On the left, there's a sidebar with navigation items: 'Home', 'Model', 'Host Predict Input', 'Pre Processing', 'Host Pre-processing', and 'Visualisasi dan Analisis'. The main content area has a title 'Stopword Removal' and a sub-section 'Data Stopword Removal'. It displays two tables: 'Sebelum Stopword Removal' and 'Sesudah Stopword Removal'. The 'Sebelum Stopword Removal' table lists 10 tweets in their original, unprocessed form. The 'Sesudah Stopword Removal' table lists the same 10 tweets, where stop words have been removed from the text. Below the tables, there are sections for 'Activate Windows' and 'Show 1 to 10 of 30 entries'.

No.	Sebelum Stopword Removal	Sesudah Stopword Removal
1	RT @spesialistik: Fadli Zon Mintu Mendagri Segara Memerlukan Ahok Jadi Gubernur DKI <a href="https://t.co/XHf5llwvH0">https://t.co/XHf5llwvH0</a>	RT @spesialistik: Fadli Zon Mintu Mendagri Segara Memerlukan Ahok Jadi Gubernur DKI <a href="https://t.co/XHf5llwvH0">https://t.co/XHf5llwvH0</a>
2	RT @BogorGubernormous: Mereka terus merusak akik dalam rangka memperparah Ahok atau Ahok gagal dalam Pilkada	RT @BogorGubernormous: Mereka merusak akik dalam rangka memperparah Ahok atau Ahok gagal dalam Pilkada
3	Sylvi Isappramee gubernur melakukan kekerasan pemimpin? Buktinya banyak bls2 atau foto bierung #Detail#FinalPilkadaKT	Sylvi: gubernur kekerasan pemimpin? Buktinya bls2 foto bierung #Detail#FinalPilkadaKT
4	Ahmad Dhani Tak Penuh Dukat Pilkada, Masalah Jalan di Bekasi Belum Sempat Terungkap <a href="https://t.co/PLUDzu2206">https://t.co/PLUDzu2206</a>	Ahmad Dhani Tak Penuh Dukat Pilkada, Masalah Jalan di Bekasi Belum Sempat Terungkap <a href="https://t.co/PLUDzu2206">https://t.co/PLUDzu2206</a>
5	RT @Endahay28: Waspada KTP polisi... kawal PILKADA <a href="https://t.co/ERQVq45H">https://t.co/ERQVq45H</a>	RT @Endahay28: Waspada KTP polisi... kawal PILKADA <a href="https://t.co/ERQVq45H">https://t.co/ERQVq45H</a>
6	@hendriksw masih banyak hal aheng yg bisa dibahas, kerapa ke pilkada tonas, apa tingkat komarman anda cuma sebutas pilkada????	@hendriksw masih banyak hal aheng yg bisa dibahas, pilkada tonas, tingkat komarman sebutas pilkada????
7	@hendriksw apa makna pernyataan. Anda? Apa anda mengaminkan kriminalisasi Ahok?? Hsl2 zg dg mulutmu pasti <a href="https://t.co/H4QD9VY0M4">https://t.co/H4QD9VY0M4</a>	@hendriksw makna pernyataan. Anda? Apa anda mengaminkan kriminalisasi Ahok?? Hsl2 zg dg mulutmu pasti <a href="https://t.co/H4QD9VY0M4">https://t.co/H4QD9VY0M4</a>
8	Ahok Akan Integrasi Transportasi Publik dengan Sistem Single Ticketinglu2026 <a href="https://t.co/42Ph7jeV0n">https://t.co/42Ph7jeV0n</a>	Ahok Akan Integrasi Transportasi Publik Sistem Single Ticketinglu2026 <a href="https://t.co/42Ph7jeV0n">https://t.co/42Ph7jeV0n</a>
9	Ahok insompi melawan kompi di biukita, biangk infrastruktur, merencana memperbaiki dengan berbagai trodosen dan works well.	Ahok melawan kompi biukita, biangk infrastruktur, merencana memperbaiki trodosen works well.
10	Awas Kocurangan Pilkada, Agus - Sylvi Pakai Aplikasi ini <a href="https://t.co/oWZWh4HJFU">https://t.co/oWZWh4HJFU</a>	Awas Kocurangan Pilkada, Agus - Sylvi Pakai Aplikasi ini <a href="https://t.co/oWZWh4HJFU">https://t.co/oWZWh4HJFU</a>

Gambar 8.23 Tampilan Halaman *Preprocessing Stopword removal*

### d. Remove emoticon Byte

Pada halaman *remove emoticon*, menampilkan dataset yang sebelum dilakukan *remove emoticon* dan sesudah dilakukan *remove emoticon* seperti gambar di bawah ini.

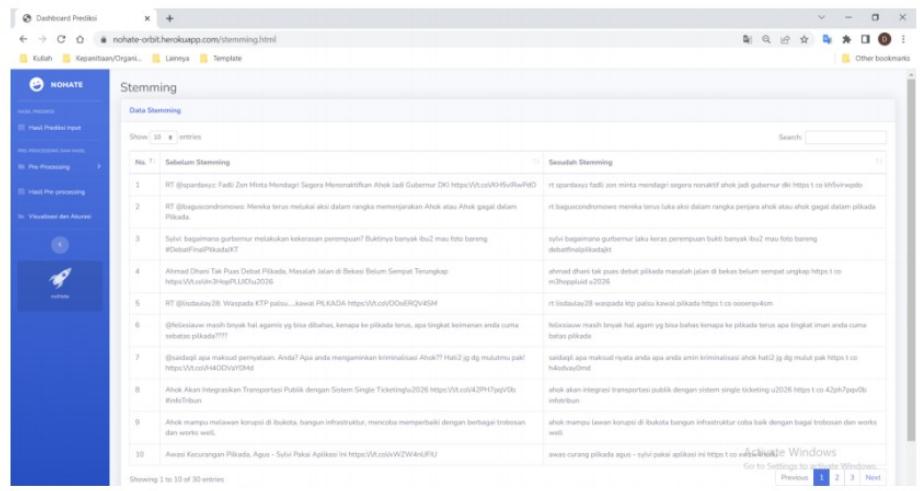


The screenshot shows a web application interface titled 'Dashboard Prediksi'. On the left, there's a sidebar with navigation items: 'Home', 'Model', 'Head Prediksi Input', 'Pre processing', 'Head Pre-processing', and 'Visualisasi dan Analisis'. The main content area has a title 'Remove Emoticon' and a sub-section 'Data Removing Emotion'. It displays two tables: 'Sebelum Remove Emotion' and 'Setelah Remove Emotion'. The 'Sebelum Remove Emotion' table contains 10 rows of raw text entries. The 'Setelah Remove Emotion' table contains the same 10 rows, with each entry having its original text followed by a cleaned version. For example, the first row 'RT (@bahusondromos: Fadi Zon Mintu Mendagri Segara Memerlukan Ahok Jadi Gubernur DKI https://t.co/kHfBvSwPhD)' is followed by 'RT (@bahusondromos: Menska terus melaui akai dalam rangka memperpanjang Ahok atau Ahok gagal dalam PilkaD)'.

Gambar 8.24 Tampilan Halaman *Preprocessing Remove emoticon*

#### e. *Stemming*

Pada halaman *stemming*, menampilkan dataset yang sebelum dilakukan *stemming* dan sesudah dilakukan *stemming* seperti gambar di bawah ini.

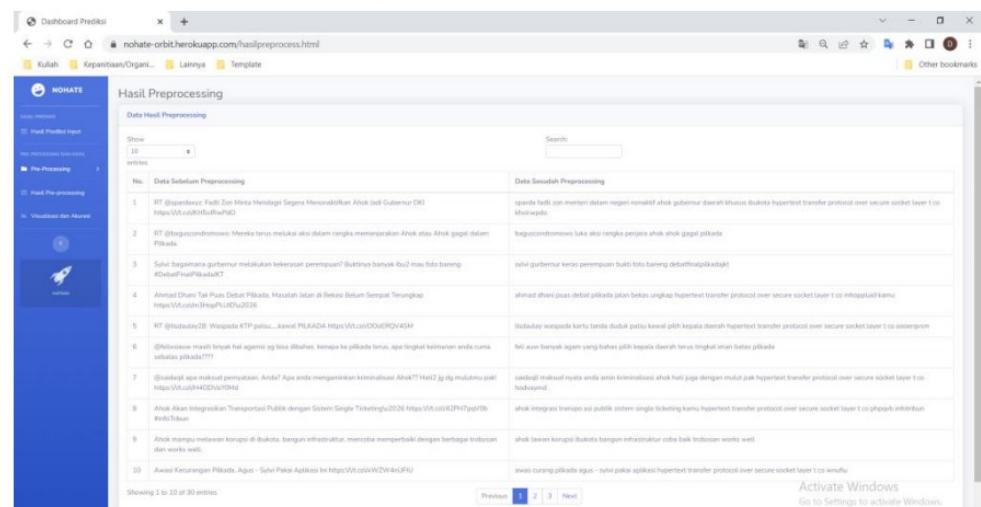


The screenshot shows the same 'Dashboard Prediksi' interface. The main content area has a title 'Stemming' and a sub-section 'Data Stemming'. It displays two tables: 'Sebelum Stemming' and 'Setelah Stemming'. The 'Sebelum Stemming' table contains 10 rows of raw text entries. The 'Setelah Stemming' table contains the same 10 rows, with each entry having its original text followed by a stemmed version. For example, the first row 'RT (@bahusondromos: Fadi Zon Mintu Mendagri Segara Memerlukan Ahok Jadi Gubernur DKI https://t.co/kHfBvSwPhD)' is followed by 'rt spandane fadi zon mintu mendagi segera nonakif ahok jadi gubernur dki https://t.co/kHfBvSwPhD'.

Gambar 8.25 Tampilan Tampilan Halaman *Preprocessing Stemming*

### 3. Fitur Hasil *Preprocessing*

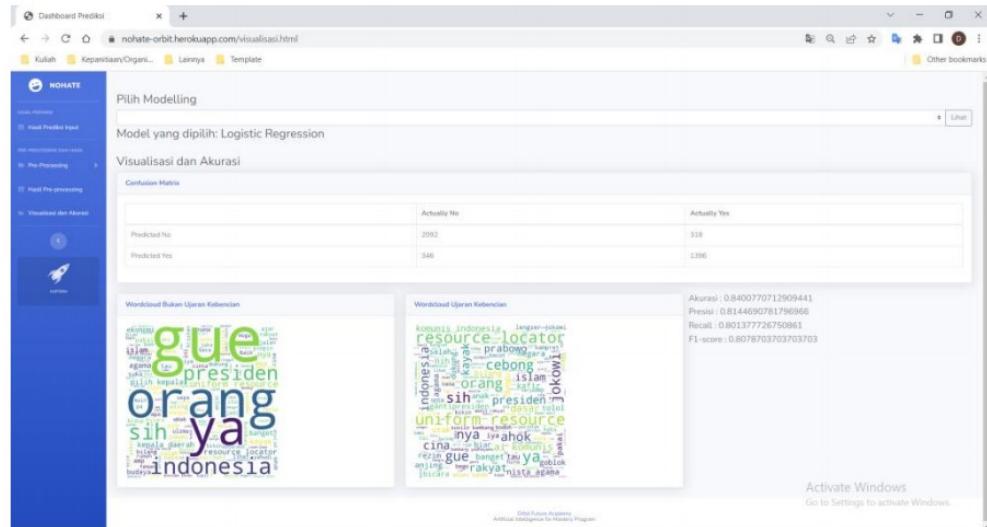
Pada halaman hasil *preprocessing*, menampilkan dataset yang sebelum dilakukan *preprocessing* dan sesudah dilakukan *preprocessing* seperti gambar di bawah ini. Hasil sesudah dilakukan *preprocessing* yang ditampilkan yaitu sesudah dilakukan berbagai macam *preprocessing* antara lain *casefolding*, *normalize alay*, *stopword removal*, *remove emoticon byte*, dan *stemming*.



Gambar 8.26 Tampilan Halaman Hasil *Preprocessing* dari semua proses (*casefolding*, *normalize alay*, *stopword removal*, *remove emoticon byte*, dan *stemming*)

#### 4. Fitur Visualisasi dan Akurasi

Menampilkan *confusion matrix*, dan *metrics* (akurasi, presisi, *Recall*, *F1-Score*) dari setiap *modelling* yang dipilih. Serta menampilkan *wordcloud* dari ujaran kebencian dan bukan ujaran kebencian seperti gambar di bawah ini.



Gambar 8.27 Tampilan Halaman Visualisasi dan Akurasi

Dalam menampilkan Wordcloud perlu dilakukan koding terlebih dahulu di python, berikut untuk kodingan *wordcloud* :

#### ▼ WordCloud

```
[ ] comment_0 = ''
comment_1 = ''

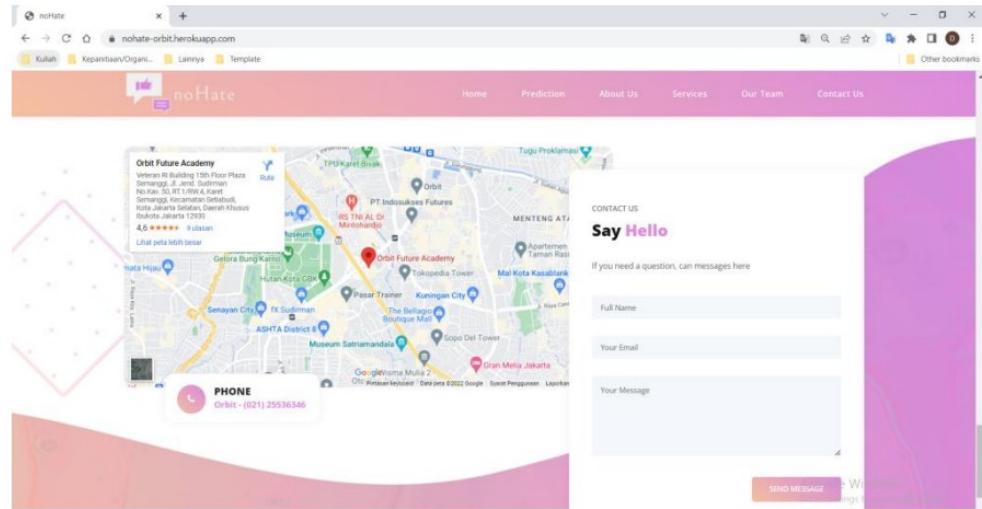
for val in data['Tweet'].loc[data['Label']==0]:
    val = str(val)
    tokens = val.split()
    for i in range(len(tokens)):
        tokens[i] = tokens[i].lower()
    comment_0 += " ".join(tokens)+" "

for val in data['Tweet'].loc[data['Label']==1]:
    val = str(val)
    tokens = val.split()
    for i in range(len(tokens)):
        tokens[i] = tokens[i].lower()
    comment_1 += " ".join(tokens)+" "
```

Gambar 8.28 Kodingan *Wordcloud*

## 5. Fitur *Contact us*

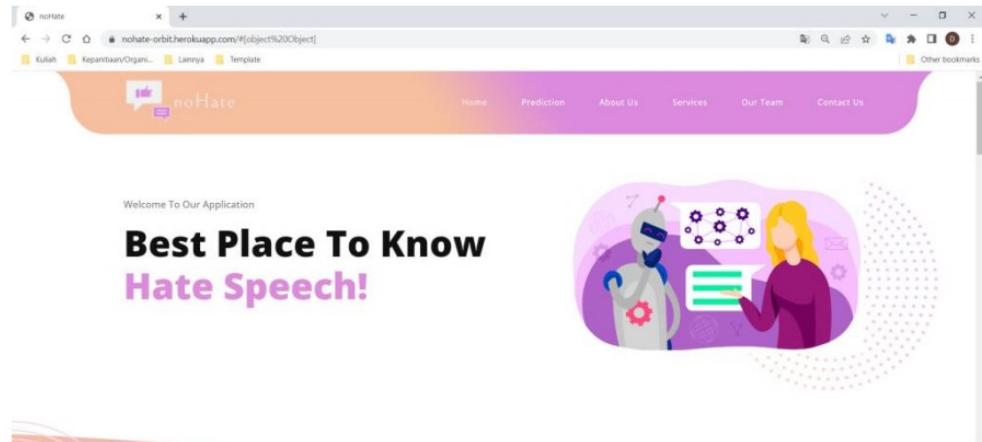
Fitur ini digunakan apabila pengguna ingin bertanya atau memberikan saran dapat mengetik pesan beserta nama dan email. Berikut adalah tampilan *contact us* aplikasi **noHate**.



Gambar 8.29 Tampilan Halaman *Contact us*

## 6. Home

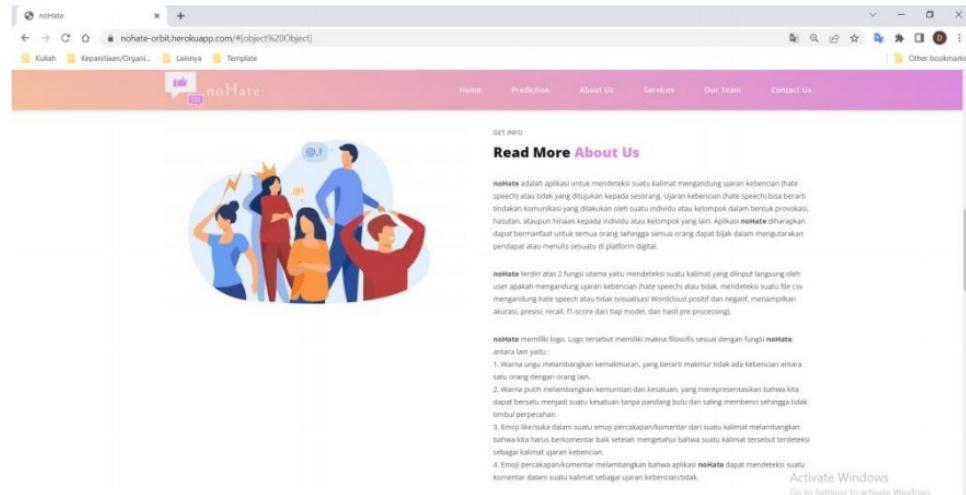
Halaman awal pada aplikasi **noHate**



Gambar 8.30 Tampilan Halaman *Home*

## 7. About us

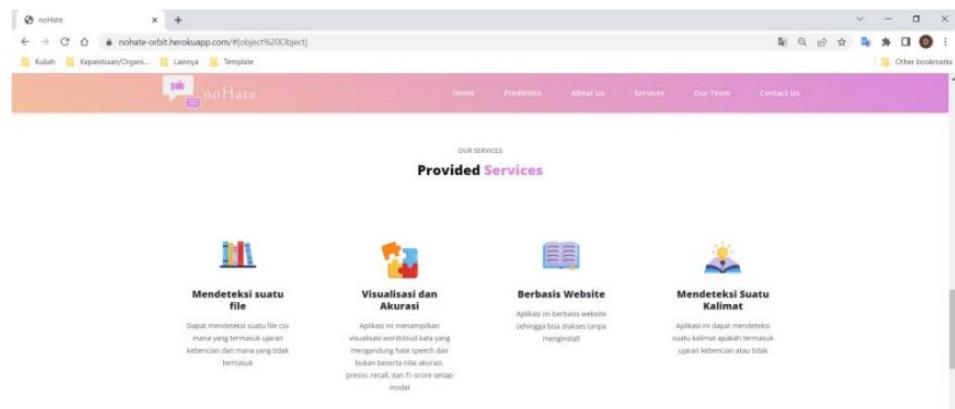
Menampilkan informasi terkait aplikasi **noHate** mulai dari apa itu aplikasi **noHate**, harapan aplikasi **noHate**, fungsi aplikasi **noHate**, dan makna filosofis dari logo aplikasi **noHate**.



Gambar 8.31 Tampilan Halaman *About us*

## 8. Services

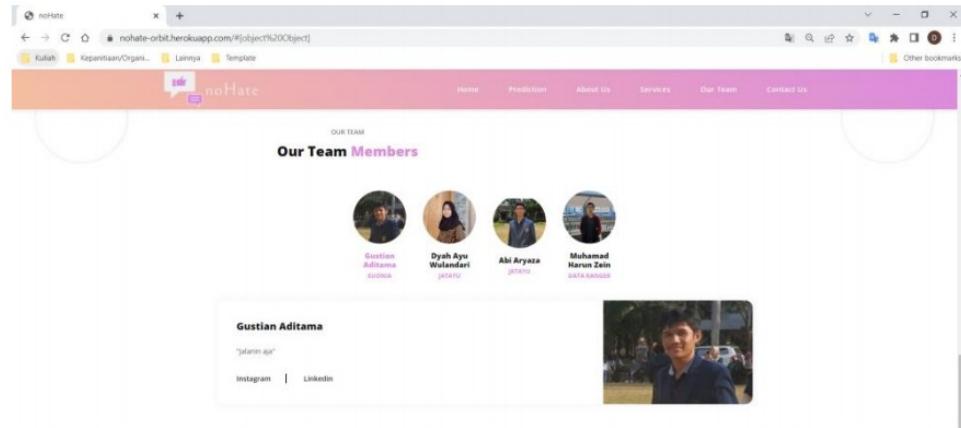
Menampilkan apa saaja *services/layanan* yang disediakan oleh aplikasi **noHate**



Gambar 8.32 Tampilan Halaman *Services*

## 9. Our team

Menampilkan anggota tim yang membuat aplikasi **noHate** mulai dari foto, nama, quotes dan media sosial (Instagram dan linkedIn).



Gambar 8.33 Tampilan Halaman *Our team*

Berikut adalah dokumentasi gambar kodingan front end aplikasi noHate :

1. Landing page (index.html) yang telah dibuat :

```
index.html
C:\Users\WIN 10>OneDrive>Dokumen>Kuliah>Orbit>Progress Project Akhir>Template>1>projek akhir orbit>templates>index.html>...
1 <!DOCTYPE html>
2 <html lang="en">
3
4   <head>
5
6     <meta charset="utf-8">
7     <meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-fit=no">
8     <meta name="description" content="">
9     <meta name="author" content="">
10    <link href="https://fonts.googleapis.com/css?family=Open+Sans:100,200,300,400,500,600,700,800,900" rel="stylesheet">
11
12    <title>noHate</title>
13
14    <!-- Bootstrap core CSS -->
15    <link href="{{url_for('static', filename='vendor/bootstrap/css/bootstrap.min.css')}}" rel="stylesheet">
16
17
18    <!-- Additional CSS Files -->
19    <link rel="stylesheet" href="{{url_for('static', filename='assets/css/fontawesome.css')}}">
20    <link rel="stylesheet" href="{{url_for('static', filename='assets/css/templatemo-eduwell-style.css')}}">
21    <link rel="stylesheet" href="{{url_for('static', filename='assets/css/owl.css')}}">
22    <link rel="stylesheet" href="{{url_for('static', filename='assets/css/lightbox.css')}}">
23
24  </head>
25
26  <body>
27
28    <!-- ***** Header Area Start ***** -->
29    <header>...</header>
```

Gambar 8.34 Kodingan index.html

2. Setelah *landing page* selesai dibuat, maka membuat *dashboard* untuk menampilkan hasil prediksi dari inputan beserta proses *preprocessing*, hasil *preprocessing* serta visualisasi dan akurasi. Untuk membuat *dashboard* tersebut, terdapat 8 file yang harus dibuat antara lain yaitu *dashboard.html*, *casefolding.html*, *normalize.html*, *stopword-removal.html*, *removeemoticon.html*, *stemming.html*, *hasilpreprocess.html*, *visualisasi.html*.

A) Berikut di bawah ini adalah gambar kodingan *dashboard.html* yang berisi hasil prediksi dari input yang dimasukkan pengguna, dan *casefolding.html*.

```
dashboard.html x casefolding.html normalize.html stopward-re  
Drive > Dokumen > Kulliah > Orbit > Progress Project Akhir > Template > 1 > projek akhir orbit  
  
</form>  
    <!-- Page Heading -->  
    <br><h1 class="h3 mb-2 text-gray-800">Hasil Prediksi Hate  
  
<!-- DataTables Example -->  
<div class="card shadow mb-4">  
    <div class="card-header py-3">  
        <h6 class="m-0 font-weight-bold text-primary">Hasil  
        <% if model is defined%>  
        <h6 class="m-0 font-weight-bold text-gray-800">{(%  
        <% endif %>  
    </div>  
    <div class="card-body">  
        <div class="table-responsive">  
            <table class="table table-bordered" id="dataTable"  
                <thead>  
                    <tr>  
                        <th>No.</th>  
                        <th>Dataset</th>  
                        <th>Rate Speech</th>  
                    </tr>  
                </thead>  
                <tbody>  
                    <% if df is defined and count is defined %>  
                    <% for i in range(count) %>  
                    <tr>  
                        <td>{{i+1}}</td>  
                        <td>{{tweet[i]}}</td>  
                        <td>{{hasil[i]}}</td>  
                    </tr>  
                <% endif %>  
            </tbody>  
        </table>  
    </div>  
    <div class="card shadow mb-4">  
        <div class="card-header py-3">  
            <h6 class="m-0 font-weight-bold text-primary">Data Cas  
        </div>  
        <div class="card-body">  
            <div class="table-responsive">  
                <table class="table table-bordered" id="dataTable"  
                    <thead>  
                        <tr>  
                            <th>No.</th>  
                            <th>Sebelum Casefolding</th>  
                            <th>Sesudah Casefolding</th>  
                        </tr>  
                    </thead>  
                    <tbody>  
                        <% if df is defined and count is defined %>  
                        <% for i in range(count) %>  
                        <tr>  
                            <td>{{i+1}}</td>  
                            <td>{{df['weet'][i]}}</td>  
                            <td>{{df['Casefolding'][i]}}</td>  
                        </tr>  
                    <% endif %>  
                </tbody>  
            </table>  
        </div>  
    </div>  
</div>
```

Gambar 8.35 Kodingan *dashboard.html* dan *casefolding.html*

B) Berikut di bawah ini adalah gambar kodingan *normalize.html*, dan *stopword-removal.html*.

```
men + Kullan > Orbit 3 Progress Project Akhir > Template > 1 > projek akhir orbit > templates
  men + Heading ->
  131 ass=<h3 class="mb-2 text-gray-800">Normalize</h3>
  132
  133 ales Example -->
  134 <card shadow="4px">
  135   lass="card-header py-3">
  136     6 class="m-0 font-weight-bold text-primary">Data Normalizing</
  137
  138   lass="card-body">
  139     iv class="table-responsive">
  140       table class="table table-bordered" id="dataTable" width="100%" border="1">
  141         thead
  142           tr
  143             th>No.
  144             th>Sebelum Normalize
  145             th>Sesudah Normalize
  146           /tr
  147         /thead
  148         tbody
  149           [% if df is defined and count is defined %]
  150             [% for i in range(count) %]
  151               tr
  152                 td>{{i+1}}</td>
  153                 td>{{df['Tweet'][i]}}</td>
  154                 td>{{df['Normalize'][i]}}</td>
  155               /tr
  156             [% endfor %]
  157           [% endif %]
  158         /tbody
  159       /table>
  160
  161       v>
  162       v>
  163       v>
  164       v>
```

Gambar 8.36 Kodingan *normalize.html* dan *stopword-removal.html*

C) Berikut di bawah ini adalah gambar kodingan `removeemoticon.html`, dan `stemming.html`.

```
removeemotionicon.html  stemming.html  haslispreshape.html
in > Kuliah > Orbit > Progress Project Akhir > Template > 1 > projek akhir orbit > templates
In -> Jaring ->
131 <h3 mb-2 text-gray-800>Remove Emoticon</h3>
132
133 : Example -->
134 <div shadow mb-4>
135 <div card-header py-3>
136 <ass m-0 font-weight-bold text-primary>Data Removing Emoticon</ass>
137
138 <div card-body>
139 <table class="table-responsive">
140 <table class="table table-bordered" id="dataTable" width="100%">
141 <thead>
142 <tr>
143 <th>No.</th>
144 <th>Sebelum Remove Emoticon</th>
145 <th>Sesudah Remove Emoticon</th>
146 </tr>
147 </thead>
148 <tbody>
149 (% if df is defined and count is defined %)
150 (% for i in range(count) %)
151 <tr>
152 <td>{{i+1}}</td>
153 <td>{{df['Tweet'][i]}}</td>
154 <td>{{df['RemoveEmoticon'][i]}}</td>
155 </tr>
156 (% endfor %)
157 (% endif %)
158 </tbody>
159 </table>
0 0 0 0 0 0 Open In Browser In 1, Col 1 Spaces: 4 UTF-8 LF HTML Go
removeemotionicon.html  stemming.html  haslispreshape.html
n > Kuliah > Orbit > Progress Project Akhir > Template > 1 > projek akhir orbit > templates > 0
100 -- Page Heading -->
101 <h1 class="mb-2 text-gray-800">Stemming</h1>
102
103 : DataTables Example -->
104 <div class="card shadow mb-4">
105 <div class="card-header py-3">
106 <ass m-0 font-weight-bold text-primary>Data Stemming</ass>
107 </div>
108 <div class="card-body">
109 <table class="table-responsive">
110 <table class="table table-bordered" id="dataTable" width="100%">
111 <thead>
112 <tr>
113 <th>No.</th>
114 <th>Sebelum Stemming</th>
115 <th>Sesudah Stemming</th>
116 </tr>
117 </thead>
118 <tbody>
119 (% if df is defined and count is defined %)
120 (% for i in range(count) %)
121 <tr>
122 <td>{{i+1}}</td>
123 <td>{{df['Tweet'][i]}}</td>
124 <td>{{df['Stemming'][i]}}</td>
125 </tr>
126 (% endfor %)
127 (% endif %)
128 </tr>
129 </tbody>
0 0 0 0 0 0 Open In Browser In 1, Col 1 Spaces: 4 UTF-8 LF HTML Go
```

Gambar 8.37 Kodingan removeemoticon.html dan stemming.html

D) Berikut di bawah ini adalah gambar kodingan hasil preprocess.html, dan visualisasi.html.

Gambar 8.38 Kodingan hasil preprocess.html dan visualisasi.html

Setelah semua dibuat maka tampilan front-end sudah jadi dan dilanjutkan untuk membuat *back end* aplikasi.

**d. Keterangan Lainnya**

- Kelebihan aplikasi **noHate** adalah :
    1. Dapat memprediksi suatu kalimat apakah termasuk ujaran kebencian atau tidak baik dari inputan kalimat maupun file yang diunggah oleh pengguna dengan berbagai model (5 model).

2. Dapat melihat hasil pre processing (*casefolding, normalize* alay, *stopword removal, remove emoticon* byte, dan *stemming*) serta dapat melihat visualisasi dan akurasi (*wordcloud, confusion matrix*, akurasi, presisi, *Recall, F1-Score*) dengan berbagai model (5 model).
  3. Dapat diakses dengan mudah karena tidak perlu menginstall aplikasi
  4. Bebas *platform* karena penggunaan aplikasi web tidak terbatas pada satu atau dua sistem operasi saja sehingga semua *device* dan semua web browser bisa digunakan untuk mengakses aplikasi **noHate**.
- Kekurangan aplikasi **noHate** adalah :
    1. Sumber daya yang terbatas sehingga setelah aplikasi *dideployment* hanya bisa mendeteksi 30 baris dan hanya bisa mendeteksi dalam format .csv
    2. Dapat mendeteksi suatu kalimat termasuk ujaran kebencian atau tidak hanya dengan menggunakan Bahasa Indonesia.
  - Harapan pengembangan aplikasi **noHate** di masa depan adalah :
    1. Dapat mendeteksi lebih dari 1 bahasa/bermacam-macam Bahasa
    2. Dapat melakukan pembobotan setiap kata sehingga dalam suatu kalimat dapat dilakukan pembobotan apakah termasuk ujaran kebencian atau tidak atau bahkan hanya berupa sindiran sehingga dapat memahami suatu makna kata dengan benar.
    3. Dapat menambahkan algoritma klasifikasi *deep learning*