



SKRIPSI

**OPTIMASI ALGORITMA *NAÏVE BAYES*,
RANDOM FOREST, DAN *LOGISTIC
REGRESSION* MENGGUNAKAN METODE
MAJORITY VOTE UNTUK ANALISIS SENTIMEN
PADA ULASAN APLIKASI SATUSEHAT**

YOHANES DIMAS WISNU WIRATAMA
NPM 20081010112

DOSEN PEMBIMBING
Budi Nugroho, S.Kom., M.Kom
Eva Yulia Puspaningrum, S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2024**



SKRIPSI

**OPTIMASI ALGORITMA *NAÏVE BAYES*,
RANDOM FOREST, DAN *LOGISTIC
REGRESSION* MENGGUNAKAN METODE
MAJORITY VOTE UNTUK ANALISIS
SENTIMEN PADA ULASAN APLIKASI
SATUSEHAT**

YOHANES DIMAS WISNU WIRATAMA
NPM 20081010112

DOSEN PEMBIMBING

Budi Nugroho, S.Kom., M.Kom

Eva Yulia Puspaningrum, S.Kom., M.Kom

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2024**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

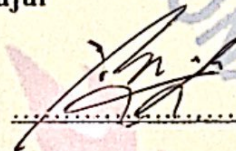
OPTIMASI ALGORITMA *NAÏVE BAYES*, *RANDOM FOREST*, DAN *LOGISTIC REGRESSION* MENGGUNAKAN METODE *MAJORITY VOTE* UNTUK ANALISIS SENTIMEN PADA ULASAN APLIKASI SATUSEHAT

Oleh :
YOHANES DIMAS WISNU WIRATAMA
NPM. 20081010112

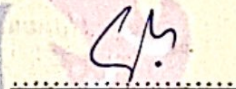
Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 30 Agustus 2024.

Menyetujui


Budi Nugroho, S.Kom., M.Kom
NIP. 19800907 2021211 005


..... (Pembimbing I)

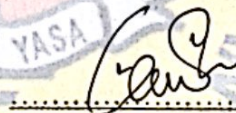
Eva Yulia Puspaningrum, S.Kom., M.Kom
NIP. 19890705 2021212 002


..... (Pembimbing II)

Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT
NPT. 222198 60 816400


..... (Ketua Penguji)

Agung Mustika Rizki, S.Kom., M.Kom.
NIP. 19930725 202203 1008


..... (Anggota Penguji)

Mengetahui,
Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**OPTIMASI ALGORITMA *NAÏVE BAYES*, *RANDOM FOREST*, DAN
LOGISTIC REGRESSION MENGGUNAKAN METODE *MAJORITY VOTE*
UNTUK ANALISIS SENTIMEN PADA ULASAN APLIKASI SATUSEHAT**

Oleh:

YOHANES DIMAS WISNU WIRATAMA

NPM. 20081010112



**Koordinator Program Studi Informatika
Fakultas Ilmu Komputer**

Fetty Tri Anggraeny, S.Kom., M.Kom.

NIP. 19820211 2021212 005

SURAT PERNYATAAN ORISINALITAS

Yang bertandatangan di bawah ini:

Nama Mahasiswa : YOHANES DIMAS WISNU WIRATAMA

Program Studi : Informatika

Dosen Pembimbing : Budi Nugroho, S.Kom., M.Kom

dengan ini menyatakan bahwa isi sebagian maupun keseluruhan skripsi dengan judul:

OPTIMASI ALGORITMA *NAÏVE BAYES*, *RANDOM FOREST*, DAN *LOGISTIC REGRESSION* MENGGUNAKAN METODE *MAJORITY VOTE* UNTUK ANALISIS SENTIMEN PADA ULASAN APLIKASI SATUSEHAT

adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diizinkan dan bukan merupakan karya pihak lain yang saya akui sebagai karya sendiri. Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka. Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.



Surabaya, 13 September 2024
Yang Membuat Pernyataan,



YOHANES DIMAS WISNU WIRATAMA
NPM. 20081010112

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Yohanes Dimas Wisnu Wiratama / 20081010112
Judul Skripsi : Optimasi Algoritma *Naïve Bayes*, *Random Forest*,
Dan *Logistic Regression* Menggunakan Metode
Majority Vote Untuk Analisis Sentimen Pada
Ulasan Aplikasi SATUSEHAT
Dosen Pembimbing : 1. Budi Nugroho, S.Kom., M.Kom
2. Eva Yulia Puspaningrum, S.Kom., M.Kom

Aplikasi SATUSEHAT yang merupakan pembaharuan dari aplikasi Peduli Lindungi, telah menjadi aplikasi Kesehatan secara umum bagi Masyarakat Indonesia. Dalam pengembangannya, diperlukan berbagai masukan dari pengguna untuk meningkatkan layanannya di masa depan. Masukan tersebut dapat ditemui pada ulasan di aplikasi *Google Play Store*. Untuk memudahkan pengelompokan ulasan tersebut di masa depan, analisis sentimen telah digunakan pada beberapa penelitian terdahulu dan memberikan hasil yang cukup baik. Dalam melakukan analisis sentimen, terdapat beragam algoritma klasifikasi yang dapat digunakan. Dalam penelitian ini, peneliti memilih menggunakan algoritma *Multinomial Naïve Bayes*, *Random Forest*, dan *Logistic Regression* yang pada penelitian sebelumnya terbukti ampuh dalam beberapa kasus analisis sentimen. Peneliti juga akan menggabungkan hasil dari masing-masing algoritma dengan metode *Majority Vote* untuk menghasilkan analisis sentiment dengan akurasi lebih tinggi. Hasil penelitian menunjukkan bahwa metode *Majority Vote* berhasil memberikan hasil uji dengan akurasi tertinggi dibanding 3 algoritma secara Tunggal. Hasil akurasi tertinggi yang didapat sebesar 97,6% dengan presisi 98%, recall 98%, dan F-1 score sebesar 98%.

Kata kunci : *Analisis Sentimen, Majority Vote, SATUSEHAT*

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Yohanes Dimas Wisnu Wiratama / 20081010112
Thesis Title : Optimasi Algoritma *Naïve Bayes*, *Random Forest*,
Dan *Logistic Regression* Menggunakan Metode
Majority Vote Untuk Analisis Sentimen Pada
Ulasan Aplikasi SATUSEHAT
Advisor : 1. Budi Nugroho, S.Kom., M.Kom
2. Eva Yulia Puspaningrum, S.Kom., M.Kom

The SATUSEHAT application, which is an update of the Peduli Lindungi app, has become a general health application for the Indonesian public. In its development, various user feedback is needed to improve its services in the future. This feedback can be found in the reviews on the Google Play Store. To facilitate the categorization of these reviews in the future, sentiment analysis has been used in several previous studies and has yielded quite good results. Various classification algorithms can be used for sentiment analysis. In this study, the researcher chose to use the Multinomial Naïve Bayes, Random Forest, and Logistic Regression algorithms, which have proven effective in several sentiment analysis cases in prior research. The researcher will also combine the results of each algorithm using the Majority Vote method to produce sentiment analysis with higher accuracy. The results show that the Majority Vote method successfully provided the highest test accuracy compared to the three individual algorithms. The highest accuracy achieved was 97.6%, with a precision of 98%, recall of 98%, and an F1-score of 98%.

Keywords: *Sentiment Analysis, Majority Vote, SATUSEHAT*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas berkat dan kebaikannya, peneliti dapat menyelesaikan skripsi berjudul “Optimasi *Algoritma Naïve Bayes*, *Random Forest*, dan *Logistic Regression* Menggunakan Metode *Majority Vote* Untuk Analisis Sentimen Pada Ulasan Aplikasi Satusehat” Berbagai dukungan juga telah diterima peneliti sehingga dapat menyelesaikan skripsi ini.

Peneliti telah mendapatkan berbagai bantuan dari berbagai pihak, baik dukungan moral, spiritual, juga dalam bentuk materi. Oleh karena itu, peneliti berterima kasih secara khusus kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
4. Bapak Fawwaz Ali Akbar, S.Kom., M.Kom. selaku dosen wali peneliti yang telah memberikan bantuan peneliti dalam perwalian sejak memulai kuliah hingga menyelesaikannya.
5. Bapak Budi Nugroho, S.Kom., M.Kom. selaku dosen pembimbing ke-1 peneliti yang telah membimbing dan memberi arahan dengan sangat baik sehingga peneliti dapat menuntaskan tugas akhir skripsi ini dengan baik.
6. Ibu Eva Yulia Puspaningrum, S.Kom., M.Kom selaku dosen pembimbing kedua saya yang telah membimbing, meluangkan waktu, dan memberikan berbagai arahan positif sehingga saya dapat menyelesaikan tugas akhir skripsi ini dengan maksimal.
7. Seluruh Dosen dan Staff Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan ilmu dan bantuan yang bermanfaat selama perkuliahan.

8. Ibu peneliti yang telah memberikan doa dan dukungan secara materi maupun non materi sehingga penulis dapat menyelesaikan perkuliahan dari awal hingga menyelesaikan tugas akhir skripsi ini untuk mengakhiri perkuliahan.
9. Mendiang Bapak peneliti yang telah mendoakan, memberi dukungan penuh, dan memotivasi untuk dengan bertanggung jawab menempuh dan menyelesaikan kuliah dengan baik.
10. Seluruh teman serta kakak tingkat pengurus HIMATIFA UPN “Veteran” Jawa Timur periode 2021/2022 yang telah membimbing dan Bersama berproses dalam meningkatkan ilmu *soft skill* peneliti di jenjang perkuliahan.
11. Seluruh anggota Keluarga Mahasiswa Katolik UPN “Veteran” Jawa Timur KMK St. Patrisius dan komunitas Lakik yang telah memberikan banyak pembelajaran secara *soft skill* maupun *hard skill* dalam perkuliahan serta mendukung peneliti untuk menyelesaikan tugas akhir skripsi ini.
12. Seluruh teman anggota PSM Gita Widya Giri UPN “Veteran” Jawa Timur yang memberi pengalaman bernyanyi serta *soft skill* di jenjang perkuliahan.
13. Seluruh teman-teman angkatan 2020 program studi Informatika yang sudah berjuang bersama dari awal memasuki perkuliahan.
14. Anastasia Anggia Paramitha selaku teman nomor 1 yang selalu mendukung, memberikan bantuan dan menemani peneliti dalam menyelesaikan skripsi ini.
15. Ibu Nurul dan Bapak Kris sekeluarga yang mendukung dan membantu peneliti selama ini.
16. Bagas, Avie, Kak Bima, Mbak Dinda, dan Nenek yang telah menemani, mendoakan dan memberi support.
17. Seluruh teman-teman Alumni Paduan Suara Sixteen Choir Group Angkatan 9 yang telah mendukung peneliti selama masa perkuliahan.
18. Seluruh anggota Komunitas Asade yang telah menjadi rekan dalam berbagi cerita dan melakukan kegiatan positif.
19. Komunitas MU Kalah yang telah berproses bersama dan memberi support.
20. Semua pihak yang penulis tidak sebutkan satu persatu.

Peneliti berharap tugas akhir skripsi ini dapat bermanfaat bagi orang banyak dan semoga peneliti dapat memberikan kembali manfaat untuk membalas

segala kebaikan dan dukungan yang diterima selama masa perkuliahan ini. Dengan penuh kesadaran, penulis memahami bahwa penyusunan skripsi ini masih memiliki banyak kekurangan. Oleh karena itu, peneliti mengharapkan kritik maupun saran yang membangun untuk mengembangkan skripsi ini. Akhir kata, semoga penelitian yang belum sempurna ini dapat bermanfaat bagi peneliti dan masyarakat umum.

Surabaya, 13 September 2024

Yohanes Dimas Wisnu Wiratama

NPM. 20081010112

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	iii
SURAT PERNYATAAN ORISINALITAS	v
ABSTRAK.....	vii
<i>ABSTRACT</i>	ix
KATA PENGANTAR	xi
DAFTAR ISI.....	xv
DAFTAR TABEL.....	xix
DAFTAR GAMBAR	xxi
DAFTAR KODE PROGRAM.....	xxiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Tujuan.....	5
1.4. Manfaat Penelitian.....	5
1.5. Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1. Penelitian Terdahulu	7
2.2. <i>Google Play Store</i>	9
2.3. Aplikasi SATUSEHAT	10
2.4. Analisis Sentimen.....	10
2.5. <i>Text Mining</i>	11
2.6. <i>Preprocessing</i>	11
2.7. <i>TF-IDF</i>	13
2.8. <i>SMOTE</i>	14
2.9. <i>Naïve Bayes Classifier</i>	14

2.10.	<i>Multinomial Naïve Bayes Classifier</i>	15
2.11.	<i>Random Forest Classifier</i>	16
2.12.	<i>Logistic Regression Classifier</i>	18
2.13.	Metode <i>Majority Vote</i>	19
2.14.	<i>Confusion Matrix</i>	19
BAB III METODOLOGI PENELITIAN		23
3.1.	Alur Penelitian.....	23
3.2.	Pengumpulan Data	23
3.3.	<i>Preprocessing</i>	24
3.3.1.	<i>Cleaning</i>	25
3.3.2.	<i>Case Folding</i>	25
3.3.3.	<i>Tokenization</i>	25
3.3.4.	<i>Normalization</i>	26
3.3.5.	<i>Stopword Removal</i>	26
3.3.6.	<i>Stemming</i>	26
3.4.	<i>Labelling Data</i>	27
3.5.	Ekstraksi Fitur <i>TF-IDF</i>	27
3.6.	<i>SMOTE</i>	29
3.7.	Pembagian Data <i>Training</i> dan Data <i>Testing</i>	30
3.8.	<i>Multinomial Naïve Bayes</i>	30
3.9.	<i>Random Forest</i>	34
3.10.	<i>Logistic Regression</i>	39
3.11.	<i>Majority Vote</i>	43
3.12.	Skenario Pengujian Hasil	45
3.12.1.	Skenario Pengujian Hasil <i>Multinomial Naïve Bayes</i>	45
3.12.2.	Skenario Pengujian Hasil <i>Random Forest</i>	47

3.12.3.	Skenario Pengujian Hasil <i>Logistic Regression</i>	49
3.12.4.	Skenario Pengujian Hasil Metode <i>Majority Vote</i>	52
BAB IV HASIL DAN PEMBAHASAN		55
4.1.	Lingkup Perangkat Penelitian	55
4.2.	Pengumpulan Data (<i>Scraping</i>)	55
4.3.	<i>Preprocessing</i>	58
4.3.1.	<i>Cleaning</i>	59
4.3.2.	<i>Case Folding</i>	60
4.3.3.	<i>Tokenization</i>	61
4.3.4.	<i>Normalization</i>	63
4.3.5.	<i>Stopwords Removal</i>	64
4.3.6.	<i>Stemming</i>	65
4.4.	<i>Labelling Data</i>	66
4.5.	<i>Feature Extraction TF-IDF</i>	68
4.6.	<i>SMOTE</i>	70
4.7.	Pembagian Data <i>Training</i> dan Data <i>Testing</i>	71
4.8.	Klasifikasi Algoritma <i>Multinomial Naïve Bayes</i>	71
4.9.	Klasifikasi Algoritma <i>Random Forest</i>	72
4.10.	Klasifikasi Algoritma <i>Logistic Regression</i>	74
4.11.	Metode <i>Majority Vote</i>	74
4.12.	Hasil Pengujian	76
4.12.1.	Hasil Pengujian <i>Multinomial Naïve Bayes</i>	76
4.12.2.	Hasil Pengujian <i>Random Forest</i>	85
4.12.3.	Hasil Pengujian <i>Logistic Regression</i>	94
4.12.4.	Hasil Pengujian <i>Majority Vote</i>	103
4.12.5.	Perbandingan Hasil Pengujian	106

BAB V KESIMPULAN DAN SARAN.....	107
5.1. Kesimpulan.....	107
5.2. Saran.....	108
DAFTAR PUSTAKA	109

DAFTAR TABEL

Tabel 2. 1 <i>Confussion Matrix</i>	20
Tabel 3. 1 <i>Cleaning</i>	25
Tabel 3. 2 <i>Case Folding</i>	25
Tabel 3. 3 <i>Tokenization</i>	26
Tabel 3. 4 <i>Normalization</i>	26
Tabel 3. 5 <i>Stopword Removal</i>	26
Tabel 3. 6 <i>Stemming</i>	27
Tabel 3. 7 <i>Labelling Data</i>	27
Tabel 3. 8 Data Sampel Acak.....	28
Tabel 3. 9 Perhitungan <i>IDF</i>	28
Tabel 3. 10 <i>TF-IDF</i>	29
Tabel 3. 11 Data Setelah <i>SMOTE</i>	30
Tabel 3. 12 Perhitungan Probabilitas Kata <i>MNB</i>	33
Tabel 3. 13 Data Sampel Acak.....	35
Tabel 3. 14 Hasil <i>Entropy</i> dan <i>Information Gain</i>	37
Tabel 3. 15 Data Sampel Acak Mengandung Kata Tolong.....	38
Tabel 3. 16 Hasil <i>Entropy</i> dan <i>Information Gain</i> 2.....	38
Tabel 3. 17 Hasil <i>Entropy</i> dan <i>Information Gain</i> 3.....	39
Tabel 3. 18 <i>Majority Vote</i>	44
Tabel 3. 19 Skenario Pengujian <i>Splitting Data Multinomial Naïve Bayes</i>	45
Tabel 3. 20 Skenario Pengujian Parameter <i>alpha Multinomial Naïve Bayes</i>	46
Tabel 3. 21 Skenario Pengujian Parameter <i>fit_prior Multinomial Naïve Bayes</i>	46
Tabel 3. 22 Skenario Pengujian <i>Splitting Data Random Forest</i>	47
Tabel 3. 23 Skenario Pengujian Parameter <i>random_state Random Forest</i>	48
Tabel 3. 24 Skenario Pengujian Parameter <i>random_state Random Forest</i>	49
Tabel 3. 25 Skenario Pengujian <i>Splitting Data Logistic Regression</i>	50
Tabel 3. 26 Skenario Pengujian Parameter <i>C Logistic Regression</i>	50
Tabel 3. 27 Skenario Pengujian Parameter <i>tol Logistic Regression</i>	51
Tabel 3. 28 Skenario Pengujian Metode <i>Majority Vote</i>	52
Tabel 4. 1 Spesifikasi Perangkat Lunak Penelitian.....	55

Tabel 4. 2 Hasil <i>Cleaning</i>	60
Tabel 4. 3 Hasil <i>Case Folding</i>	61
Tabel 4. 4 Hasil <i>Tokenization</i>	62
Tabel 4. 5 Hasil <i>Normalization</i>	64
Tabel 4. 6 Hasil <i>Stopword Removal</i>	65
Tabel 4. 7 Hasil <i>Stemming</i>	66
Tabel 4. 8 Pengujian Hasil <i>Multinomial Naïve Bayes</i> Skenario <i>Splitting Data</i>	79
Tabel 4. 9 Pengujian Hasil <i>Multinomial Naïve Bayes</i> Skenario Parameter <i>alpha</i> ..	82
Tabel 4. 10 Pengujian Hasil <i>Multinomial Naïve Bayes</i> Skenario <i>fit_prior</i>	84
Tabel 4. 11 Pengujian Hasil <i>Random Forest</i> Skenario <i>Splitting Data</i>	88
Tabel 4. 12 Pengujian Hasil <i>Random Forest</i> Skenario Parameter <i>random_state</i> ..	91
Tabel 4. 13 Pengujian Hasil <i>Random Forest</i> Skenario Parameter <i>n_estimators</i> ...	93
Tabel 4. 14 Pengujian Hasil <i>Logistic Regression</i> Skenario <i>Splitting Data</i>	97
Tabel 4. 15 Pengujian Hasil <i>Logistic Regression</i> Skenario Parameter <i>C</i>	100
Tabel 4. 16 Pengujian Hasil <i>Logistic Regression</i> Skenario Parameter <i>C</i>	102
Tabel 4. 17 Pengujian Hasil Metode <i>Majority Vote</i>	105
Tabel 4. 18 Perbandingan Hasil Pengujian Terbaik.....	106

DAFTAR GAMBAR

Gambar 2. 1 Ilustrasi 1 <i>Random Forest</i>	17
Gambar 2. 2 Ilustrasi 2 <i>Random Forest</i>	18
Gambar 3. 1 Alur Penelitian	23
Gambar 3. 2 Alur <i>Preprocessing</i>	24
Gambar 3. 3 Alur Analisis Sentimen <i>Multinomial Naïve Bayes</i>	31
Gambar 3. 4 Alur Analisis Sentimen <i>Random Forest</i>	35
Gambar 3. 5 Alur Analisis Sentimen <i>Logistic Regression</i>	40
Gambar 3. 6 Alur Analisis Sentimen <i>Majority Vote</i>	44
Gambar 4. 1 <i>Screenshot Program Scraping</i>	56
Gambar 4. 2 <i>Screenshot rawsatusehat0.csv</i>	57
Gambar 4. 3 Kumpulan Hasil <i>Scraping</i>	57
Gambar 4. 4 <i>Screenshot rawsatusehat0</i>	58
Gambar 4. 5 <i>Screenshot rawsatusehata</i>	58
Gambar 4. 6 <i>Screenshot Cleaning</i>	60
Gambar 4. 7 <i>Screenshot Case Folding</i>	61
Gambar 4. 8 <i>Screenshot Tokenization</i>	62
Gambar 4. 9 <i>Screenshot Normalization</i>	63
Gambar 4. 10 <i>Screenshot Stopwords Removal</i>	64
Gambar 4. 11 <i>Screenshot Stemming</i>	66
Gambar 4. 12 <i>Screenshot Labelling Data</i>	68
Gambar 4. 13 <i>Screenshot TF-IDF</i>	69
Gambar 4. 14 <i>Before After TF-IDF</i>	69
Gambar 4. 15 Data setelah <i>SMOTE</i>	71
Gambar 4. 16 <i>Confusion Matrix Multinomial Naïve Bayes 70:30</i>	77
Gambar 4. 17 <i>Confusion Matrix Multinomial Naïve Bayes 80:20</i>	77
Gambar 4. 18 <i>Confusion Matrix Multinomial Naïve Bayes 90:10</i>	78
Gambar 4. 19 <i>Confusion Matrix Multinomial Naïve Bayes alpha=1.0</i>	80
Gambar 4. 20 <i>Confusion Matrix Multinomial Naïve Bayes alpha=0.01</i>	80
Gambar 4. 21 <i>Confusion Matrix Multinomial Naïve Bayes alpha=1e-10</i>	81
Gambar 4. 22 <i>Confusion Matrix Multinomial Naïve Bayes fit_prior=True</i>	83

Gambar 4. 23	<i>Confusion Matrix Multinomial Naïve Bayes fit_prior=False</i>	84
Gambar 4. 24	<i>Confusion Matrix Random Forest 70:30</i>	86
Gambar 4. 25	<i>Confusion Matrix Random Forest 80:20</i>	87
Gambar 4. 26	<i>Confusion Matrix Random Forest 90:10</i>	88
Gambar 4. 27	<i>Confusion Matrix Random Forest random_state=42</i>	90
Gambar 4. 28	<i>Confusion Matrix Random Forest random_state=78</i>	90
Gambar 4. 29	<i>Confusion Matrix Random Forest n_estimators=150</i>	92
Gambar 4. 30	<i>Confusion Matrix Random Forest n_estimators=200</i>	93
Gambar 4. 31	<i>Confusion Matrix Logistic Regression 70:30</i>	95
Gambar 4. 32	<i>Confusion Matrix Logistic Regression 80:20</i>	96
Gambar 4. 33	<i>Confusion Matrix Logistic Regression 90:10</i>	97
Gambar 4. 34	<i>Confusion Matrix Logistic Regression C=1.0</i>	98
Gambar 4. 35	<i>Confusion Matrix Logistic Regression C=100</i>	99
Gambar 4. 36	<i>Confusion Matrix Logistic Regression tol=0.1</i>	101
Gambar 4. 37	<i>Confusion Matrix Logistic Regression tol=1e-4</i>	102
Gambar 4. 38	<i>Confusion Matrix Majority Vote tanpa parameter</i>	103
Gambar 4. 39	<i>Confusion Matrix Majority Vote dengan parameter</i>	105

DAFTAR KODE PROGRAM

Kode Program 4. 1 <i>Scraping Data</i>	55
Kode Program 4. 2 <i>Cleaning</i>	59
Kode Program 4. 3 <i>Case Folding</i>	60
Kode Program 4. 4 <i>Tokenization</i>	62
Kode Program 4. 5 <i>Normalization</i>	63
Kode Program 4. 6 <i>Stopword Removal</i>	64
Kode Program 4. 7 <i>Stemming</i>	65
Kode Program 4. 8 <i>Labelling Data</i>	67
Kode Program 4. 9 <i>TF-IDF</i>	68
Kode Program 4. 10 Menampilkan 1 Kalimat <i>TF-IDF</i>	69
Kode Program 4. 11 <i>SMOTE</i>	70
Kode Program 4. 12 Pembagian Data <i>Training</i> dan <i>Testing</i>	71
Kode Program 4. 13 Klasifikasi Algoritma <i>Multinomial Naïve Bayes</i>	72
Kode Program 4. 14 <i>Confusion Matrix Multinomial Naïve Bayes</i>	72
Kode Program 4. 15 Klasifikasi Algoritma <i>Random Forest</i>	73
Kode Program 4. 16 <i>Confusion Matrix Random Forest</i>	73
Kode Program 4. 17 Klasifikasi Algoritma <i>Logistic Regression</i>	74
Kode Program 4. 18 <i>Confusion Matrix Logistic Regression</i>	74
Kode Program 4. 19 Metode <i>Majority Vote</i>	75
Kode Program 4. 20 <i>Confusion Matrix Majority Vote</i>	75

Halaman ini sengaja dikosongkan