

**ANALISIS SENTIMEN TERHADAP REBRANDING
TWITTER MENJADI X DENGAN NAÏVE
BAYES, SVM, DAN ENSEMBLE
LEARNING**



SKRIPSI

**HANSEN
00000060325**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISIS SENTIMEN TERHADAP REBRANDING
TWITTER MENJADI X DENGAN NAÏVE
BAYES, SVM, DAN ENSEMBLE
LEARNING**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

**HANSEN
00000060325**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Hansen
Nomor Induk Mahasiswa : 00000060325
Program Studi : Informatika

Skripsi dengan judul:

Analisis Sentimen terhadap Rebranding Twitter menjadi X dengan Naïve Bayes, SVM, dan Ensemble Learning

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

UMM

Tangerang, 1 Juli 2025



UNIVERSITAS (Hansen)
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISIS SENTIMEN TERHADAP REBRANDING TWITTER MENJADI X DENGAN NAÏVE BAYES, SVM, DAN ENSEMBLE LEARNING

Oleh

Nama : Hansen

NIM : 00000060325

Program Studi : Informatika

Fakultas : Fakultas Teknik dan Informatika

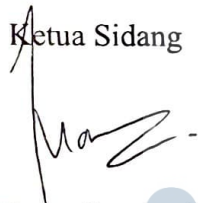
Telah diujikan pada Rabu, 9 Juli 2024

Pukul 15.00 s.d 17.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang



Marlinda Vasty Overbeek, S.Kom,
M. Kom.

NIDN: 0818038501

Penguji



Moeljono Widjaja, B.Sc., M.Sc., Ph.D.
NIDN: 0311106903

Pembimbing



David Agustriawan, S.Kom., M.Sc., Ph.D.

NIDN: 0525088601

Ketua Program Studi Informatika

UNIVERSITAS
MULTIMEDIA
NUSANTARA

Arya Wicaksana, S.Kom., M.Eng.Sc. (OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Hansen
NIM : 00000060325
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Analisis Sentimen terhadap
Rebranding Twitter menjadi X
dengan Naïve Bayes, SVM, dan
Ensemble Learning

Menyatakan dengan sesungguhnya bahwa saya bersedia:

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 1 Juli 2025

Yang menyatakan



Hansen

HALAMAN PERSEMBAHAN / MOTTO

"A good name is to be more desired than great wealth, Favor is better than silver and gold."

Proverbs 22:1 (NASB)



KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan Skripsi, ini dengan judul: Analisis Sentimen terhadap Rebranding Twitter menjadi X dengan Naïve Bayes, SVM, dan Ensemble Learning dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai penyusunan laporan skripsi ini, sulit bagi saya untuk menyelesaikan laporan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak David Agustriawan, S.Kom., M.Sc., Ph.D., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.

Semoga karya ilmiah ini dapat bermanfaat dan menjadi referensi bagi peneliti yang sedang melakukan penelitian dengan topik Analisis Sentimen atau topik lainnya.

Tangerang, 1 Juli 2025



Hansen

ANALISIS SENTIMEN TERHADAP REBRANDING TWITTER MENJADI X DENGAN NAÏVE BAYES, SVM, DAN ENSEMBLE LEARNING

Hansen

ABSTRAK

Rebranding Twitter menjadi X merupakan salah satu topik hangat pada pertengahan tahun 2023. Logo burung putih dengan latar biru dan nama Twitter resmi diubah menjadi huruf "X" dengan latar hitam. Perubahan ini memicu berbagai respon dari pengguna X yang akan dianalisis menggunakan pendekatan *machine learning*, yaitu algoritma Naïve Bayes, SVM, dan *Ensemble Learning*. Model terbaik diperoleh dari model SVM dengan TF-IDF yang dilakukan *hyperparameter tuning*. Model ini menghasilkan nilai *accuracy* sebesar 92%, *precision* 91%, *recall* 89%, dan serta *f1-score* 90%. Berdasarkan penelitian yang dilakukan, ditemukan bahwa mayoritas sentimen dari topik ini adalah netral terhadap *rebranding* yang dilakukan.

Kata kunci: Analisis Sentimen, *Ensemble Learning*, *Machine Learning*, Rebranding Twitter.



**SENTIMENT ANALYSIS OF TWITTER REBRANDING TO X USING NAIVE
BAYES, SVM, AND ENSEMBLE LEARNING**

Hansen

ABSTRACT

The rebranding of Twitter to X became one of the most widely discussed topics in mid-2023. The iconic white bird logo on a blue background and the name "TWitter" were officially replaced with the letter "X" on a black background. This transformation triggered various responses from X users, which will be analyzed using a machine learning approach involving the Naïve Bayes, SVM, and Ensemble Learning algorithms. The best-performing model was obtained using SVM with TF-IDF and optimized through hyperparameter tuning. This model achieved an accuracy of 92%, 91% precision, 89% recall, and 90% f1-score. Based on the analysis, it was found that the majority of sentiments regarding the rebranding were neutral.

Keywords: *Ensemble Learning, Machine Learning, Sentiment Analysis, Twitter Rebranding.*



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	5
2.1 Analisis Sentimen	5
2.2 Twitter	5
2.3 Data Preprocessing	5
2.4 CountVectorizer	6
2.5 TF-IDF	7
2.5.1 TF (Term Frequency)	7
2.5.2 IDF (Inverse Document Frequency)	7
2.6 Naïve Bayes	8
2.7 Support Vector Machine (SVM)	9
2.8 Logistic Regression	13
2.9 SMOTE	14
2.10 Hyperparameter Tuning	15
2.11 Ensemble Learning	15
2.12 Cross Validation	16
2.13 Confusion Matrix	16
BAB 3 METODOLOGI PENELITIAN	19
3.1 Studi Literatur	19
3.2 Pengumpulan Data	20
3.3 Perancangan Sistem	21
3.3.1 Import Libraries	22
3.3.2 Import and Merge Data	22
3.3.3 Remove Duplicate Data	22
3.3.4 Data Pre-processing	22
3.3.5 Labeling	23
3.3.6 Data Split	23
3.3.7 Feature Extraction	23

3.3.8	Data Balancing	24
3.4	Pembentukan Model	24
3.5	Pengujian	30
3.6	Validasi dan Evaluasi	31
3.7	Dokumentasi	31
3.8	Spesifikasi Sistem	31
3.8.1	Hardware	31
3.8.2	Software	32
BAB 4	HASIL DAN DISKUSI	33
4.1	Pengumpulan Data	33
4.2	Perancangan Sistem	34
4.2.1	Import Libraries	34
4.2.2	Import and Merge Data	35
4.2.3	Remove Duplicate Data	35
4.2.4	Data Pre-processing	36
4.2.5	Labeling	45
4.2.6	Data Split	47
4.2.7	Feature Extraction	48
4.2.8	Data Balancing	49
4.3	Pembentukan Model	50
4.3.1	Model tanpa Hyperparameter Tuning	50
4.3.2	Model dengan Hyperparameter Tuning	51
4.3.3	Model Ensemble Learning: Stacking	53
4.4	Pengujian	53
4.5	Validasi dan Evaluasi	54
4.5.1	Validasi	54
4.5.2	Evaluasi	54
4.5.3	Diskusi	56
BAB 5	SIMPULAN DAN SARAN	58
5.1	Simpulan	58
5.2	Saran	59
DAFTAR PUSTAKA		60

DAFTAR TABEL

Tabel 2.1	Confusion matrix	16
Tabel 3.1	Parameter CountVectorizer()	26
Tabel 3.2	Parameter TfidfVectorizer()	27
Tabel 3.3	Parameter MultinomialNB()	28
Tabel 3.4	Parameter SVC()	29
Tabel 4.1	Hasil <i>cleaning</i> data	37
Tabel 4.2	Hasil <i>case folding</i>	38
Tabel 4.3	Hasil <i>slangword replace</i>	39
Tabel 4.4	Hasil <i>translate</i>	40
Tabel 4.5	Hasil <i>tokenization</i>	41
Tabel 4.6	Hasil <i>stopword removal</i>	43
Tabel 4.7	Hasil <i>stemming</i>	44
Tabel 4.8	Hasil <i>labeling</i> menggunakan TextBlob	46
Tabel 4.9	Hasil <i>cross-validation</i> dengan <i>Stratified K-fold</i>	54
Tabel A1	<i>Research gap</i>	82



DAFTAR GAMBAR

Gambar 2.1	Grafik SVM dengan <i>margin</i> , <i>hyperplane</i> , dan keterangan kelas	10
Gambar 2.2	Ilustrasi <i>multi-class classification</i> SVM	13
Gambar 3.1	<i>Flowchart</i> metodologi penelitian	19
Gambar 3.2	<i>Flowchart</i> pengumpulan data	21
Gambar 3.3	<i>Flowchart</i> perancangan sistem	21
Gambar 3.4	<i>Flowchart</i> perancangan sistem	23
Gambar 3.5	<i>Flowchart</i> pembentukan model tanpa <i>hyperparameter tuning</i>	25
Gambar 3.6	<i>Flowchart</i> dengan <i>hyperparameter tuning</i>	25
Gambar 3.7	<i>Flowchart</i> pengujian dengan <i>ensemble learning</i>	30
Gambar 4.1	Confusion matrix dari model terbaik	55
Gambar 4.2	Classification report model terbaik	56



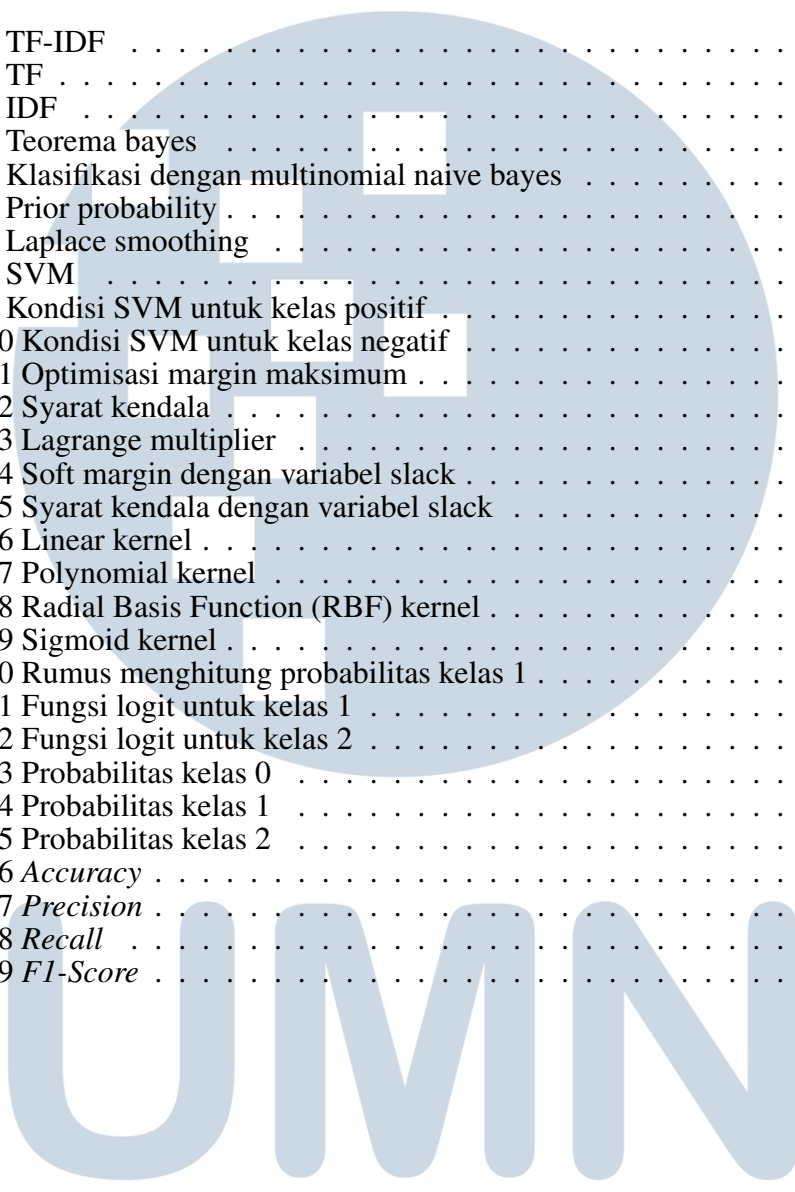
DAFTAR KODE

Kode 4.1	Instalasi <i>library</i>	33
Kode 4.2	<i>Crawl</i> data	34
Kode 4.3	<i>Merge</i> data	35
Kode 4.4	Remove duplicate data	36
Kode 4.5	<i>Cleaning</i> data	36
Kode 4.6	<i>Case folding</i>	37
Kode 4.7	<i>Slangword replace</i>	38
Kode 4.8	<i>Translate</i> data ke bahasa Inggris	39
Kode 4.9	<i>Tokenization</i>	41
Kode 4.10	<i>Stopword removal</i>	42
Kode 4.11	<i>Stemming</i>	43
Kode 4.12	<i>Labeling</i> dengan TextBlob	45
Kode 4.13	<i>Split</i> data	47
Kode 4.14	<i>Feature extraction</i> CountVectorizer	48
Kode 4.15	<i>Feature extraction</i> TF-IDF	49
Kode 4.16	Data balancing dengan SMOTE	49
Kode 4.17	Pembentukan model naive bayes	50
Kode 4.18	Pembentukan model support vector machine (SVM)	51
Kode 4.19	Pembentukan model hyperparameter tuning dengan random search	51
Kode 4.20	Pembentukan model dengan teknik stacking	53
Kode 4.21	Penerapan fungsi untuk prediksi	54
Kode 4.22	Menampilkan <i>confusion matrix</i>	54
Kode A1	<i>Import libraries</i>	81



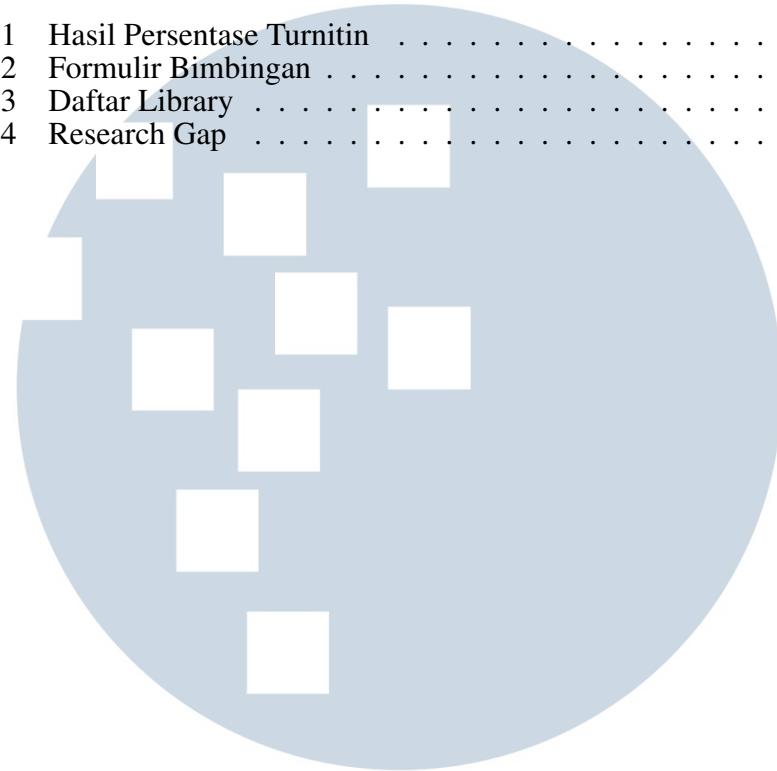
DAFTAR RUMUS

Rumus 2.1 TF-IDF	7
Rumus 2.2 TF	7
Rumus 2.3 IDF	7
Rumus 2.4 Teorema bayes	8
Rumus 2.5 Klasifikasi dengan multinomial naive bayes	9
Rumus 2.6 Prior probability	9
Rumus 2.7 Laplace smoothing	9
Rumus 2.8 SVM	10
Rumus 2.9 Kondisi SVM untuk kelas positif	10
Rumus 2.10 Kondisi SVM untuk kelas negatif	10
Rumus 2.11 Optimisasi margin maksimum	11
Rumus 2.12 Syarat kendala	11
Rumus 2.13 Lagrange multiplier	11
Rumus 2.14 Soft margin dengan variabel slack	11
Rumus 2.15 Syarat kendala dengan variabel slack	11
Rumus 2.16 Linear kernel	12
Rumus 2.17 Polynomial kernel	12
Rumus 2.18 Radial Basis Function (RBF) kernel	12
Rumus 2.19 Sigmoid kernel	12
Rumus 2.20 Rumus menghitung probabilitas kelas 1	13
Rumus 2.21 Fungsi logit untuk kelas 1	14
Rumus 2.22 Fungsi logit untuk kelas 2	14
Rumus 2.23 Probabilitas kelas 0	14
Rumus 2.24 Probabilitas kelas 1	14
Rumus 2.25 Probabilitas kelas 2	14
Rumus 2.26 <i>Accuracy</i>	17
Rumus 2.27 <i>Precision</i>	17
Rumus 2.28 <i>Recall</i>	17
Rumus 2.29 <i>F1-Score</i>	18


 UNIVERSITAS
 MULTIMEDIA
 NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	64
Lampiran 2	Formulir Bimbingan	79
Lampiran 3	Daftar Library	81
Lampiran 4	Research Gap	82



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA