

**IMPLEMENTASI ALGORITMA NAIVE BAYES SEBAGAI FILTER
UJARAN KEBENCIAN PADA BOT AUTOMENFESS TWITTER**

SKRIPSI



Disusun oleh:

Asrori Akhsan

19.21.1353

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**IMPLEMENTASI ALGORITMA NAIVE BAYES SEBAGAI FILTER
UJARAN KEBENCIAN PADA BOT AUTOMENFESS TWITTER**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



Disusun oleh:

Asrori Akhsan

19.21.1353

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

IMPLEMENTASI ALGORITMA NAIVE BAYES SEBAGAI FILTER UJARAN KEBENCIAN PADA BOT AUTOMENFESS TWITTER

yang dipersiapkan dan disusun oleh

Asrori Akhsan

19.21.1353

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 09 Februari 2021

Dosen Pembimbing,

Arif Akbarul Huda, S.Si, M.Eng.

NIK. 190302287

PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA NAIVE BAYES SEBAGAI FILTER
UJARAN KEBENCIAN PADA BOT AUTOMENFESS TWITTER**

yang dipersiapkan dan disusun oleh

Asrori Akhsan
19.21.1353

telah dipertahankan di depan Dewan Penguji
pada tanggal 16 Februari 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ika Nur Fajri, M.Kom.
NIK. 190302268

Ferry Wahyu Wibowo, S.Si, M.Cs.
NIK. 190302235

Arif Akbarul Huda, S.Si, M.Eng.
NIK. 190302287

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 16 Februari 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom.
NIK. 190302096

PERNYATAAN

Saya yang bertanda tangan di bawah ini menyatakan bahwa, Skripsi ini merupakan karya penulis (ASLI), dan isi dalam Skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan penulis juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggung jawab penulis pribadi.

Yogyakarta, 03 Februari 2021



Asrori Akhsan

NIM. 19.21.1353

MOTTO

“Stay hungry, stay foolish.”

(Steve Jobs)

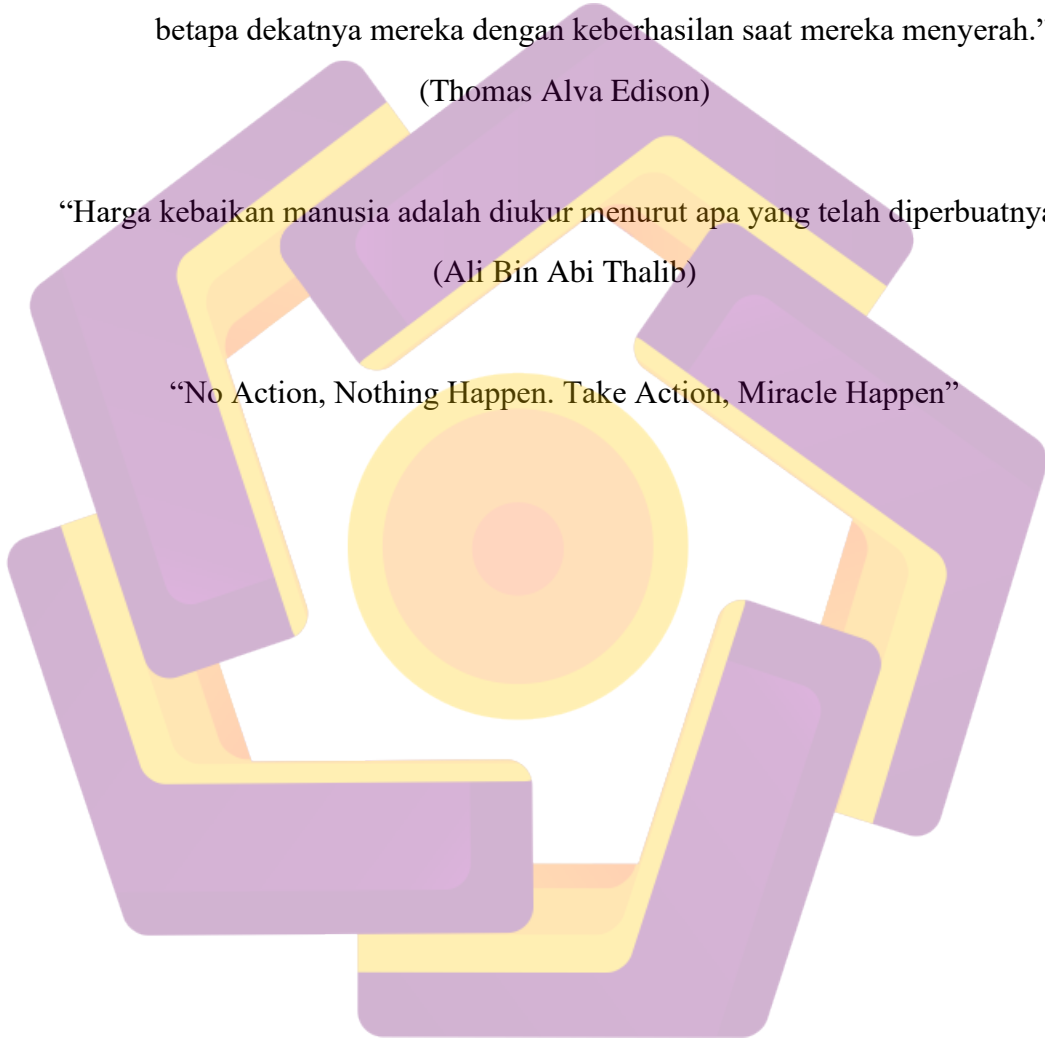
“Banyak kegagalan dalam hidup ini dikarenakan orang-orang tidak menyadari betapa dekatnya mereka dengan keberhasilan saat mereka menyerah.”

(Thomas Alva Edison)

“Harga kebaikan manusia adalah diukur menurut apa yang telah diperbuatnya”

(Ali Bin Abi Thalib)

“No Action, Nothing Happen. Take Action, Miracle Happen”



PERSEMBAHAN

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan rahmat dan hidayah serta karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan sebaik-baiknya dan mendapatkan hasil yang maksimal.

Tidak lepas dari bantuan berbagai pihak, oleh karena itu penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT yang telah memberikan kelancaran kepada penulis dalam menyelesaikan Tugas Akhir ini.
2. Ibu, kakak, dan saudara-saudara yang selalu memberikan doa terbaiknya.
3. Bapak Ibu Dosen Universitas Amikom Yogyakarta yang telah ikhlas dan selalu sabar memberikan ilmu mulai dari semester awal hingga akhir.
4. Bapak Arif Akbarul Huda, S.Si, M.Eng sebagai dosen pembimbing yang telah mendampingi dalam penyusunan Tugas Akhir ini.
5. Teman-teman 16-D3TI-03 yang telah memberi doa dan dukungannya. Serta kawan-kawan saya di manapun kalian berada yang tidak henti memberikan doa dan dukungan kepada saya.
6. Teman-teman 19-IF Transfer yang sudah menemani dalam suka dan duka dunia perkuliahan.
7. Kontrakan Gabut, Lesehan Pak Din, dan Angkringan Pak Samd atas apa yang sudah kalian berikan.
8. Binar Academy yang telah memberikan ilmu dan fasilitas yang nyaman untuk menyelesaikan berbagai permasalahan selama ini.

Asrori Akhsan

KATA PENGANTAR

Puji dan syukur penulis persembahkan untuk Allah SWT yang telah memberikan rahmat, hidayah dan kekuatan sehingga penulis dapat menyelesaikan Skripsi ini sesuai dengan waktu yang diinginkan penulis. Tidak lupa sholawat serta salam penulis haturkan pada junjungan Nabi besar kita yakni Muhammad SAW.

Skripsi ini disusun sebagai salah satu syarat kelulusan bagi setiap mahasiswa Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang Strata-I dan untuk memperoleh gelar Sarjana Komputer.

Dengan selesainya Skripsi ini, maka penulis tidak lupa mengucapkan terima kasih kepada :

1. Universitas AMIKOM Yogyakarta selaku pihak yang telah mengizinkan penulis melakukan penelitian.
2. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas AMIKOM Yogyakarta.
3. Bapak Hanif Al Fatta, M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
4. Ibu Windha Mega Pradnya D, M.Kom selaku Ketua Program Studi S1 Informatika Universitas AMIKOM Yogyakarta.
5. Bapak Arif Akbarul Huda, S.Si, M.Eng selaku dosen pembimbing yang telah memberikan petunjuk, bimbingan dan nasihatnya dalam proses penulisan Skripsi ini.
6. Orang tua dan keluarga besar yang memberikan motivasi, doa dan semangat.

DAFTAR ISI

PERSETUJUAN.....	iii
PENGESAHAN.....	iv
PERNYATAAN.....	v
MOTTO.....	vi
PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
INTISARI.....	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	2
1.5. Metode.....	3
1.5.1. Metode Pengumpulan data.....	3
1.5.2. Metode Analisis.....	3
1.5.3. Metode Pengembangan.....	3
1.5.4. Metode Pengujian.....	3
1.6. Sistematika Penulisan.....	4
1.6.1. BAB I PENDAHULUAN.....	4
1.6.2. BAB II LANDASAN TEORI.....	4
1.6.3. BAB III METODOLOGI PENELITIAN.....	4
1.6.4. BAB IV HASIL DAN PEMBAHASAN.....	4
1.6.5. BAB V PENUTUP.....	4

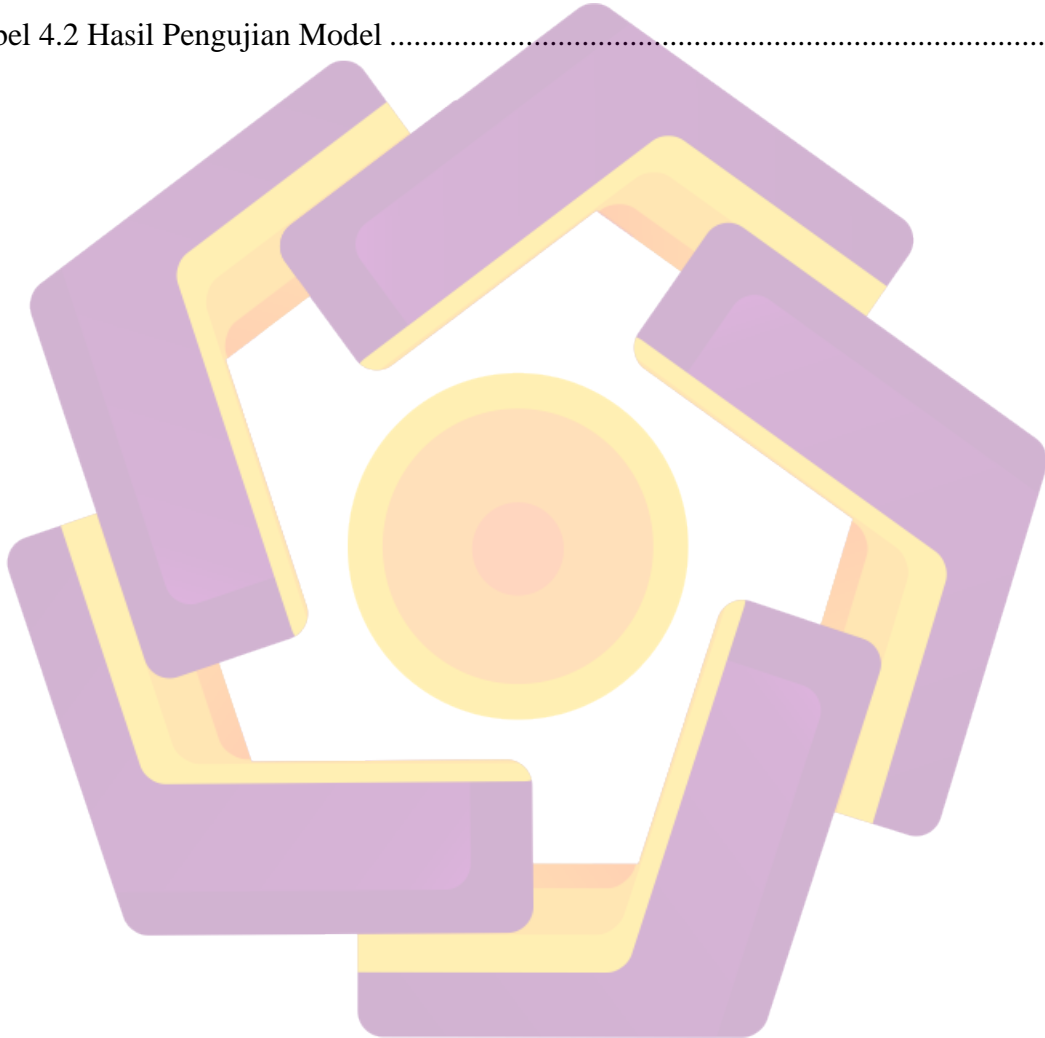
BAB II LANDASAN TEORI.....	5
2.1. Tinjauan Pustaka	5
2.2. Dasar Teori.....	7
2.2.1. Twitter	7
2.2.2. Hate Speech.....	7
2.2.3. Analisis Sentimen	8
2.2.4. Data mining.....	8
2.2.5. Text Mining.....	10
2.2.6. Tahap Preprocessing	11
2.2.7. Naïve Bayes Classifier	13
2.2.8. Cross Validation.....	13
2.2.9. Performance Evaluation Measure	15
2.2.10. Version Control System.....	16
2.2.11. Continuous Integration dan Continuous Deployment (CI/CD).....	16
BAB III METODOLOGI PENELITIAN	18
3.1. Populasi dan Sample Penelitian	18
3.2. Metode pengumpulan Data	18
3.3. Variabel Penelitian	18
3.4. Metode Analisis Data.....	19
3.5. Tahapan Penelitian.....	20
3.6. Rancangan Arsitektur Sistem.....	21
BAB IV IMPLEMENTASI DAN PEMBAHASAN	23
4.1. Pengambilan Data	23
4.2. Tahap Preprocessing	25
4.2.1. Data Cleaning	25
4.2.2. Correction Slang Word	26
4.2.3. Tokenizing.....	26
4.2.4. Stemming	26
4.3. Klasifikasi dengan Metode Naïve Bayes Classifier	28
4.4. Uji dan Evaluasi Model.....	29
4.5. Implementasi pada Bot Twitter.....	31

4.6. Deployment Aplikasi Bot Twitter.....	34
BAB V PENUTUP	35
5.1. Kesimpulan	35
5.2. Saran.....	35
DAFTAR PUSTAKA.....	36



DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian	6
Tabel 2.2 Tabel Confusion Matrix.....	16
Tabel 3.1 Variabel Penelitian	18
Tabel 4.1 Hasil Preprocessing Data.....	27
Tabel 4.2 Hasil Pengujian Model	30



DAFTAR GAMBAR

Gambar 2.1 Data Mining	10
Gambar 2.2 Tahapan Preprocessing	11
Gambar 2.3 Tahapan Case Folding	12
Gambar 2.4 Tahapan Tokenizing.....	12
Gambar 2.5 Tahapan Stemming	13
Gambar 2.6 Persamaan Teorema Bayes	13
Gambar 2.7 Rumus Precision	15
Gambar 2.8 Rumus Recall.....	15
Gambar 2.9 Rumus Accuracy.....	15
Gambar 2.10 Tahapan CI/CD.....	17
Gambar 3.1 Diagram Alir Penelitian	20
Gambar 3.2 Rancangan Arsitektur Sistem.....	21
Gambar 4.1 Credentials API Twitter	23
Gambar 4.2 Potongan Script Crawling Twitter	24
Gambar 4.3 Isi Dataset	24
Gambar 4.4 Potongan Script Data Cleaning.....	25
Gambar 4.5 Potongan Script Correction Slang Word	26
Gambar 4.6 Potongan Script Import Library	28
Gambar 4.7 Potongan Script Data Split.....	28
Gambar 4.8 Potongan Script pengujian model.....	29
Gambar 4.9 Hasil Pengujian Model.....	29
Gambar 4.10 Direktori Aplikasi BOT	31
Gambar 4.11 Potongan Script app.py	31
Gambar 4.12 Potongan Script twitter.py	32
Gambar 4.13 Potongan Script db_mongo.py.....	32
Gambar 4.14 Potongan Script constant.py	33
Gambar 4.15 Potongan Script nbc.py	33

Gambar 4.16 Potongan Script gitlab-ci.yml 34



INTISARI

Twitter merupakan media sosial yang memungkinkan untuk pengguna mengirimkan cuitan dengan menggunakan teks, gambar, dan video. Akun twitter biasanya digunakan secara personal, namun terdapat juga akun berjenis automenfess. Akun automenfess ini sengaja dibuat untuk dijadikan bot yang memiliki kemampuan untuk mengirimkan cuitan yang dikirimkan oleh pengguna lain melalui fitur pesan atau direct message.

Akun berjenis automenfess rawan terkena suspend karena melanggar Rules aplikasi Twitter yaitu spam dan ujaran kebencian. Untuk mengetahui teks yang mengandung unsur ujaran kebencian dapat dilakukan dengan cara melakukan analisis sentimen. Analisis sentimen merupakan metode yang digunakan untuk memahami, mengekstraksi, dan mengolah data kalimat opini tekstual yang kemudian diklasifikasikan menjadi opini positif dan opini negatif.

Dari permasalahan tersebut, penulis melakukan penelitian dengan judul "Implementasi Algoritma Naive Bayes Sebagai Filter Ujaran Kebencian Pada Bot Automenfess Twitter". Pada penelitian ini penulis mengintegrasikan bot automenfess dengan algoritma Naive Bayes sebagai metode klasifikasi data teks. Hasil yang didapatkan dari penelitian ini adalah bot mampu melakukan filter ujaran kebencian dengan Nilai Akurasi rata-rata sebesar 86.44%, F1-Score sebesar 86.03%, Nilai Precision sebesar 88.48%, dan Nilai Recall sebesar 83.93%.

Kata Kunci: Akurasi, Analisis Sentimen, Automenfess, Bot, Klasifikasi, Naive Bayes Classifier, Twitter.

ABSTRACT

Twitter is a social media that allows users to send tweets using text, images, and videos. twitter accounts are usually used personally, but there are also automenfess-type accounts. This automenfess account was deliberately created to be a bot that has the ability to send tweets sent by other users via direct message.

Automenfess accounts are prone to being suspended for violating Twitter's rules of spam and hate speech. To find out the text that contains elements of hate speech can be done by doing a sentiment analysis. Sentiment analysis is a method used to understand, extract, and process textual opinion sentence data that is then classified into positive opinions and negative opinions.

From these problems, the author conducted a study with the title "Implementation of the Naive Bayes Algorithm as a Filter for Hate Speech on Twitter Automenfess Bot". In this study, the authors integrated the automenfess bot with the Naive Bayes algorithm as a text data classification method. The results obtained from this study are bots are able to filter hate speech with an average accuracy rate 86.44%, F1-Score 86.03%, Precision Score 88.48%, and Recall Score 83.93%.

Keywords: *Accuracy, Sentiment Analysis, Automenfess, Bots, Classification, Naive Bayes Classifier, Twitter.*