

**PERBANDINGAN ALGORITMA NAÏVE BAYES DAN
DECISION TREE PADA ANALISIS SENTIMEN INVASI
RUSIA KE UKRAINA**

TUGAS AKHIR

Disusun oleh:
Geraldo Saroinsong
(18013017)



PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS UNIKA DE LA SALLE
MANADO
2022

**PERBANDINGAN ALGORITMA NAÏVE BAYES DAN
DECISION TREE PADA ANALISIS SENTIMEN INVASI
RUSIA KE UKRAINA**

TUGAS AKHIR

Ditulis untuk memenuhi sebagai persyaratan untuk mendapatkan Gelar Sarjana
Teknik pada Program Studi Teknik Informatika

Disusun oleh:

Geraldo Saroinsong

(18013017)



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS UNIKA DE LA SALLE
MANADO
2022**

LEMBAR PERNYATAAN

Yang bertandatangan di bawah ini :

Nama : Geraldo Saroinsong
NIM : 18013017
Tempat dan Tanggal Lahir : Manado, 25 Agustus 2000
Fakultas/Program Studi : Teknik/ Teknik Informatika

Menyatakan bahwa Karya Ilmiah atau Tugas Akhir dan atau Aplikasi atau Program yang berjudul "**Perbandingan Algoritma Naïve Bayes dan Decision tree Pada Analisis Sentimen Invasi Rusia Ke Ukraina**" yang saya buat merupakan benar hasil karya saya dan bukan karya tulis orang lain, baik sebagian ataupun seluruhnya kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar, maka saya bersedia menerima sanksi akademis sesuai dengan yang ditetapkan oleh Fakultas, berupa pembatalan Tugas Akhir dan hasilnya.

Manado, 26 Juli 2022

Yang Menyatakan,



Geraldo Saroinsong

Menyetujui,

Dosen Pembimbing I

Apriandy Angdresey, S.T., M.Sc.

Dosen Pembimbing II

Michael G. Sumampouw, S.T., M.T.

Mengetahui,

Ketua Program Studi

Vivie Deyby Kumenap, S.T., M.Cs.

Dekan Fakultas Teknik



Ronald Albert Rachmadi, S.T., MT.

Vivie Deyby Kumenap, S.T., M.Cs. Ronald Albert Rachmadi, S.T., MT.



**UNIVERSITAS KATOLIK DE LA SALLE
MANADO - INDONESIA**

Nama : Geraldo Saroinsong
NIM : 18013017
Fakultas : Teknik
Program Studi : Teknik Informatika
Judul Tugas Akhir : Perbandingan Algoritma Naïve Bayes Dan *Decision tree*
Pada Analisis Sentimen Invasi Rusia Ke Ukraina
Pembimbing I : Apriandy Angdresey, ST., M.Sc.
Pembimbing II : Michael George Sumampouw, S.T., M.T.

Manado, 26 Juli 2022

Menyetujui,

Dosen Pembimbing I

Apriandy Angdresey, S.T., M.Sc.

Dosen Pembimbing II

Michael G. Sumampouw, S.T., M.T.

Mengetahui,

Ketua Program Studi

Vivie Deyby Kumenap, S.T., M.Cs.

Dekan Fakultas Teknik



Ronald Albert Rachmadi, S.T., MT.

KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa atas berkat dan penyertaan-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir ini yang berjudul “Perbandingan Algoritma *Naïve Bayes* dan *Decision tree* Pada Analisis Sentimen Invasi Rusia Ke Ukraina”. Tujuan dari penulisan laporan tugas akhir ini yaitu untuk memenuhi sebagian persyaratan akademik untuk mendapatkan gelar Sarjana Teknik pada Program Studi Teknik Informatika Universitas Katolik De La Salle Manado.

Dalam penggeraan Tugas Akhir ini yaitu laporan dan aplikasi, penulis memperoleh banyak bantuan dari beberapa pihak, baik secara langsung maupun tidak langsung. Oleh sebab itu, penulis mengucapkan banyak terima kasih kepada pihak-pihak yang telah membantu dalam Tugas Akhir ini baik pada penyusunan laporan maupun dalam pembuatan aplikasi, beberapa di antaranya:

1. Prof. Dr. Johanis Ohoitimur, selaku Rektor Universitas Katolik De La Salle Manado.
2. Bapak Ronald A. Rachmadi, S.T., M.T, selaku Dekan Fakultas Teknik Universitas Katolik De La Salle Manado.
3. Ibu Vivie D. Kumenap, S.T., M.Cs., sebagai Ketua Program Studi Teknik Informatika Universitas Katolik De La Salle Manado.
4. Bapak Apriandy Angdresey, S.T., M.Sc., selaku Dosen Pembimbing I Tugas Akhir yang telah membimbing dan mengarahkan dalam penyusunan laporan dan pembangunan aplikasi
5. Bapak Michael George Sumampouw, S.T., M.T., sebagai Dosen Pembimbing II Tugas Akhir yang telah membantu dalam Tugas Akhir.
6. Bapak Steven Pandelaki, S.T., M.Sc. sebagai Dosen Pembimbing Akademik yang selalu memberi bimbingan dan nasihat.
7. Keluarga, khususnya buat Ayah, Ibu, Kakak yang selalu mendoakan dan memberi dukungan baik secara moral maupun materi.
8. Saudari Fiorella Sangkay yang selalu mendoakan dan memberi dukungan selama pelaksanaan Tugas Akhir.
9. Teman-teman Teknik Informatika Angkatan 2018, khususnya teman-teman seperjuangan Tugas Akhir selalu memberikan dukungan.
10. Pihak-pihak lain yang penulis tidak dapat menuliskannya satu persatu.

Selama penulisan laporan ini, penulis menyadari bahwa masih terdapat kekeliruan dan kekurangan. Maka itu, saran dan kritik yang membangun guna untuk penyempurnaan laporan Tugas Akhir ini sangat diharapkan. Akhirnya, semoga laporan Tugas Akhir ini dapat bermanfaat. Tuhan Yesus memberkati.

Manado, Juli 2022

Penulis

ABSTRACT

Twitter is one of the social media that is most often used by the public to express various opinions openly on a topic which is usually marked with a hashtag. Russia's invasion of Ukraine is one of the hot or trending topics since the beginning of 2022. This has certainly attracted the attention of the world, so that many give opinions about the issue. To obtain the polarity sentiment of these opinions, a sentiment analysis application was built using the Naïve Bayes and Decision tree algorithms, as well as to compare both algorithms. This application works by crawling data from Twitter that goes through 3 main phases, namely pre-processing, processing and post-processing with confusion matrix.

This research builds an application that can determine the polarity of sentiment from tweet data using the naive bayes and decision tree algorithms. Based on the tests carried out with 1069 dataset, the accuracy results by the naïve bayes in the partition 80% training data compared to 20% testing data are 88.10% which is better than using the decision tree algorithm that is 85.61%. Meanwhile, the execution time of this application to analyze the sentiment is around ≤ 0.57 seconds.

Keywords: *Text Mining, Sentiment Analysis, Algorithm Comparison, Naïve Bayes, Decision tree.*

ABSTRAK

Twitter merupakan salah satu media sosial yang paling sering digunakan masyarakat untuk menyampaikan berbagai opini secara terbuka mengenai suatu topik yang biasanya ditandai dengan sebuah *hashtag*. Invasi yang dilakukan oleh Rusia terhadap Ukraina adalah salah satu topik yang sedang panas sejak awal tahun 2022. Hal ini tentunya menarik perhatian masyarakat sehingga banyak yang memberi opini tentang masalah tersebut. Untuk mengetahui tingkat sentimen dari opini-opini tersebut, maka dibangun aplikasi analisis sentimen dengan menggunakan algoritma *Naïve Bayes* dan *Decision tree*. Cara kerja dari aplikasi ini, yaitu data yang ditarik dari *Twitter* melewati 3 tahapan utama, yaitu *pre-processing*, *processing* dan *post-processing* dengan menggunakan *confusion matrix*.

Penelitian ini menghasilkan aplikasi yang dapat menentukan tingkat sentimen dari data *tweet* dengan menggunakan algoritma *naïve bayes* dan *decision tree*, serta membandingkan kedua algoritma tersebut. Dari pengujian yang telah dilakukan dengan 1069 data diperoleh hasil akurasi pada partisi 80% data latih 20% data uji dengan menggunakan *naïve bayes* diperoleh 88.10% lebih unggul daripada algoritma *decision tree* yaitu sebesar 85.61%. Sedangkan untuk waktu eksekusi dari aplikasi dalam melakukan analisis sentimen yaitu sekitar ≤ 0.57 detik.

Kata kunci: *Text Mining*, Analisis Sentimen, Perbandingan Algoritma, *Naïve Bayes*, *Decision tree*.

DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERNYATAAN	ii
LEMBAR PENGESAHAN	iv
KATA PENGANTAR	v
<i>ABSTRACT</i>	vi
ABSTRAK.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
DAFTAR FORMULA	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I: PENDAHULUAN.....	Error! Bookmark not defined.
1.1 Latar Belakang	Error! Bookmark not defined.
1.2 Rumusan Masalah	Error! Bookmark not defined.
1.3 Tujuan Tugas Akhir.....	Error! Bookmark not defined.
1.4 Manfaat Tugas Akhir.....	Error! Bookmark not defined.
1.5 Batasan Masalah.....	Error! Bookmark not defined.
1.6 Metode Penelitian.....	Error! Bookmark not defined.
1.7 Sistematika Penulisan.....	Error! Bookmark not defined.
BAB II: STUDI PUSTAKA.....	Error! Bookmark not defined.
2.1 Analisis Sentimen.....	Error! Bookmark not defined.
2.2 Invasi Rusia Ke Ukraina.....	Error! Bookmark not defined.
2.3 Twitter	Error! Bookmark not defined.
2.4 Text Mining.....	Error! Bookmark not defined.
2.4.1 <i>Pre-Processing</i>	Error! Bookmark not defined.
2.4.2 <i>Processing</i>	Error! Bookmark not defined.
2.4.3 <i>Confusion Matrix</i>	Error! Bookmark not defined.
2.5 Bahasa Pemrograman	Error! Bookmark not defined.
2.5.1 HTML	Error! Bookmark not defined.
2.5.2 Python	Error! Bookmark not defined.
2.5.3 CSS.....	Error! Bookmark not defined.
2.6 Metodologi Penelitian	Error! Bookmark not defined.
2.7 Kakas Pemodelan	Error! Bookmark not defined.
2.7.1 <i>Flowchart</i>	Error! Bookmark not defined.
2.7.2 DFD.....	Error! Bookmark not defined.
2.7.3 <i>Entity relationship diagram (ERD)</i> ..	Error! Bookmark not defined.
2.8 Penelitian Terkait	Error! Bookmark not defined.
2.9 Metode Pengumpulan Data	Error! Bookmark not defined.
BAB III: ANALISIS	Error! Bookmark not defined.
3.1 <i>Create Project Vision</i>	Error! Bookmark not defined.
3.1.1 Latar Belakang	Error! Bookmark not defined.
3.1.2 Tujuan Proyek	Error! Bookmark not defined.

3.1.3	Lingkup Proyek	Error! Bookmark not defined.
3.1.4	Analisis Aplikasi Serupa	Error! Bookmark not defined.
3.1.5	Persyaratan dan Preferensi Sistem Berjalan	Error! Bookmark not defined.
3.1.6	Rencana Pengendalian	Error! Bookmark not defined.
3.1.7	Manajemen Risiko	Error! Bookmark not defined.
3.1.8	Analisis Teknologi	Error! Bookmark not defined.
3.2	<i>Identification Scrum Master and Stakeholders</i>	Error! Bookmark not defined.
3.2.1	<i>Scrum Master</i>	Error! Bookmark not defined.
3.2.2	Spesifikasi Pengguna	Error! Bookmark not defined.
3.2.3	Definisi Sistem	Error! Bookmark not defined.
3.3	<i>Form Scrum Team</i>	Error! Bookmark not defined.
3.4	<i>Develop Epic</i>	Error! Bookmark not defined.
3.4.1	<i>Content Outline</i>	Error! Bookmark not defined.
3.4.2	<i>Analyze the System</i>	Error! Bookmark not defined.
3.4.3	<i>Flowchart</i> Aplikasi yang Dibangun	Error! Bookmark not defined.
3.5	<i>Create Prioritized Product Backlog</i>	Error! Bookmark not defined.
3.6	<i>Conduct Release Planning</i>	Error! Bookmark not defined.
 BAB IV: PERANCANGAN		
4.1	<i>Create User Stories</i>	Error! Bookmark not defined.
4.2	Rancangan Modul Program	Error! Bookmark not defined.
4.3	<i>Approve, Estimate and Commit User Stories</i>	Error! Bookmark not defined.
4.4	<i>Estimate Tasks</i>	Error! Bookmark not defined.
4.5	<i>Create Sprint Backlog</i>	Error! Bookmark not defined.
4.5.1	<i>Data Flow Diagram</i>	Error! Bookmark not defined.
4.5.2	<i>Entity Relationship Diagram</i>	Error! Bookmark not defined.
 BAB V: IMPLEMENTASI.....		
5.1	<i>Create Deliverables</i>	Error! Bookmark not defined.
5.1.1	Implementasi Antarmuka	Error! Bookmark not defined.
5.1.2	Implementasi Basis Data	Error! Bookmark not defined.
5.2	<i>Conduct Daily Stand Up</i>	Error! Bookmark not defined.
5.3	<i>Groom Prioritized Project Backlog</i>	Error! Bookmark not defined.
5.3.1	Simulasi Model Analisis Sentimen ..	Error! Bookmark not defined.
5.3.2	Melakukan Pemrograman	Error! Bookmark not defined.
 BAB VI: PENGUJIAN		
6.1	<i>Convence Scrum of Scrum</i>	Error! Bookmark not defined.
6.1.1	Kasus Uji	Error! Bookmark not defined.
6.1.2	Tujuan Pengujian	Error! Bookmark not defined.
6.1.3	Kriteria Pengujian	Error! Bookmark not defined.
6.2	<i>Demonstrate and Validate Sprint</i>	Error! Bookmark not defined.
6.3	<i>Retrospect Sprint</i>	Error! Bookmark not defined.
 BAB VII: KESIMPULAN DAN SARAN		

7.1	Kesimpulan.....	Error! Bookmark not defined.
7.2	Saran	Error! Bookmark not defined.

DAFTAR PUSTAKAError! Bookmark not defined.
LAMPIRAN

DAFTAR TABEL

Tabel 2.1 Simbol-simbol <i>Flowchart</i> [17].....	Error! Bookmark not defined.
Tabel 2.2 Simbol-simbol DFD [18]	Error! Bookmark not defined.
Tabel 2.3 Simbol-simbol ERD [19]	Error! Bookmark not defined.
Tabel 3.1 Perbandingan Aplikasi Serupa.....	Error! Bookmark not defined.
Tabel 3.2 Tabel Rencana Pengendalian	Error! Bookmark not defined.
Tabel 3.3 Tabel Manajemen Risiko	Error! Bookmark not defined.
Tabel 3.4 Perangkat Lunak	Error! Bookmark not defined.
Tabel 3.5 Perangkat Keras	Error! Bookmark not defined.
Tabel 3.6 Spesifikasi Pengguna	Error! Bookmark not defined.
Tabel 3.7 Spesifikasi Kebutuhan Sistem.....	Error! Bookmark not defined.
Tabel 3.8 Contoh Hasil <i>Crawling</i>	Error! Bookmark not defined.
Tabel 3.9 Contoh Hasil <i>Cleaning</i>	Error! Bookmark not defined.
Tabel 3.10 Contoh Hasil <i>Casefolding</i>	Error! Bookmark not defined.
Tabel 3.11 Contoh Hasil <i>Stopword Filtering</i>	Error! Bookmark not defined.
Tabel 3.12 Contoh Hasil <i>Stemming</i>	Error! Bookmark not defined.
Tabel 3.13 Contoh TF dan DF	Error! Bookmark not defined.
Tabel 3.14 Contoh Hasil <i>Entropy</i> dan <i>Gain</i>	Error! Bookmark not defined.
Tabel 3.15 Probabilitas Kelas	Error! Bookmark not defined.
Tabel 3.16 Probabilitas Kata Dalam Kelas	Error! Bookmark not defined.
Tabel 3.17 Struktur Organisasi Pembangunan Aplikasi	Error! Bookmark not defined.
Tabel 3.18 <i>Prioritized Product Backlog</i>	Error! Bookmark not defined.
Tabel 3.19 Penjadwalan	Error! Bookmark not defined.
Tabel 4.1 Rancangan Modul Program	Error! Bookmark not defined.
Tabel 4.2 <i>Approve and Estimate</i>	Error! Bookmark not defined.
Tabel 4.3 Kamus Data.....	Error! Bookmark not defined.
Tabel 6.1 Kasus Uji.....	Error! Bookmark not defined.
Tabel 6.2 Pengujian.....	Error! Bookmark not defined.
Tabel 6.3 Pengujian 80%-20%	Error! Bookmark not defined.
Tabel 6.4 Pengujian 70%-30%	Error! Bookmark not defined.
Tabel 6.5 Pengujian 60%-40%	Error! Bookmark not defined.

DAFTAR GAMBAR

Gambar 2.1 Contoh <i>Decision tree</i> [3]	Error! Bookmark not defined.
Gambar 2.2 Penggambaran Tabel <i>Confusion Matrix</i> [12]	Error! Bookmark not defined.
Gambar 3.1 Pohon Keputusan.....	Error! Bookmark not defined.
Gambar 3.2 <i>Content Outline</i>	Error! Bookmark not defined.
Gambar 3.3 <i>Flowchart</i> Aplikasi.....	Error! Bookmark not defined.
Gambar 4.1 Rancangan Tampilan <i>Crawling</i>	Error! Bookmark not defined.
Gambar 4.2 Rancangan Halaman <i>Pre-processing</i> .	Error! Bookmark not defined.
Gambar 4.3 Rancangan Halaman <i>Processing</i>	Error! Bookmark not defined.
Gambar 4.4 Rancangan Halaman Rata-rata	Error! Bookmark not defined.
Gambar 4.5 Diagram konteks	Error! Bookmark not defined.
Gambar 4.6 Diagram <i>Zero</i>	Error! Bookmark not defined.
Gambar 4.7 DFD Level 1 <i>Pre-processing</i>	Error! Bookmark not defined.
Gambar 4.8 DFD Level 1 <i>Post-processing</i>	Error! Bookmark not defined.
Gambar 4.9 ERD.....	Error! Bookmark not defined.
Gambar 5.1 Halaman <i>Crawling</i>	Error! Bookmark not defined.
Gambar 5.2 Halaman <i>Pre-processing</i>	Error! Bookmark not defined.
Gambar 5.3 Halaman <i>Processing</i>	Error! Bookmark not defined.
Gambar 5.4 Halaman <i>Testing</i>	Error! Bookmark not defined.
Gambar 5.5 Halaman Rata-rata.....	Error! Bookmark not defined.
Gambar 5.6 Tabel <i>dataset</i>	Error! Bookmark not defined.
Gambar 5.7 Tabel prepos	Error! Bookmark not defined.
Gambar 5.8 Tabel <i>processing</i>	Error! Bookmark not defined.
Gambar 5.9 Tabel datamentah	Error! Bookmark not defined.
Gambar 5.10 Tabel validasi	Error! Bookmark not defined.
Gambar 5.11 <i>Script Menarik Data</i>	Error! Bookmark not defined.
Gambar 5.12 <i>Script Cleaning</i>	Error! Bookmark not defined.
Gambar 5.13 <i>Script Tokenizing</i>	Error! Bookmark not defined.
Gambar 5.14 <i>Script Stopword Filtering</i>	Error! Bookmark not defined.
Gambar 5.15 <i>Script Stemming</i>	Error! Bookmark not defined.
Gambar 5.16 <i>Script Confusion Matrix</i>	Error! Bookmark not defined.
Gambar 6.1 Penarikan Data	Error! Bookmark not defined.
Gambar 6.2 <i>Preprocessing</i>	Error! Bookmark not defined.
Gambar 6.3 <i>Testing Web Browser Chrome</i>	Error! Bookmark not defined.
Gambar 6.4 <i>Testing Web Browser Mozilla Firefox</i>	Error! Bookmark not defined.

DAFTAR FORMULA

Formula 2.1 TF-IDF	8
Formula 2.2 Nilai <i>Priori</i>	9
Formula 2.3 Probabilitas Kata	9
Formula 2.4 Nilai <i>Posteriori</i>	10
Formula 2.5 <i>Entropy</i>	11
Formula 2.6 <i>Gain</i>	11

DAFTAR LAMPIRAN

Lampiran A – *Dataset Tweet*A-1

