

**ANALISA SENTIMEN SOSIAL MEDIA *TWITTER* TERHADAP
PILPRES 2024 MENGGUNAKAN ALGORITMA *NAÏVE*
*BAYES***

SKRIPSI

OLEH

DENNY HERDIANSYAH CHAN

188160019



FAKULTAS TEKNIK

PROGRAM STUDI TEKNIK INFORMATIKA

UNIVERSITAS MEDAN AREA

2023

UNIVERSITAS MEDAN AREA

© Hak Cipta Di Lindungi Undang-Undang

i

Document Accepted 19/4/24

1. Dilarang Mengutip sebagian atau seluruh dokumen ini tanpa mencantumkan sumber
2. Pengutipan hanya untuk keperluan pendidikan, penelitian dan penulisan karya ilmiah
3. Dilarang memperbanyak sebagian atau seluruh karya ini dalam bentuk apapun tanpa izin Universitas Medan Area

Access From (repository.uma.ac.id)19/4/24

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

Judul Skripsi : Analisa Sentimen Sosial Media *Twitter* Terhadap Pilpres 2024
Menggunakan Metode *Naive Bayes*

Nama : Denny Herdiansyah Chan

NPM 188160019

Fakultas : Teknik

Prodi : Teknik Informatika

Disetujui Oleh :

Komisi Pembimbing

Nurul Khairina S.Kom, M.Kom
Pembimbing

Diketahui :

Dr. Eng. Supriatno, S.T, M.T
Dekan
Rizki Muliono, S. Kom, M.Kom
Prodi/WD I

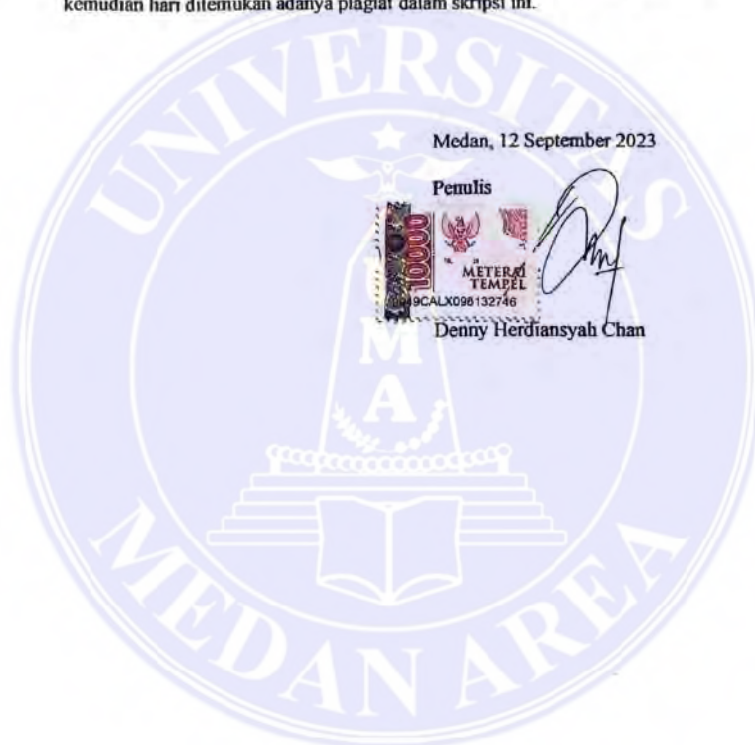
Tanggal Lulus : 12 September 2023

HALAMAN PERNYATAAN

HALAMAN PERNYATAAN

Saya menyatakan bahwa skripsi yang saya susun, sebagai syarat memperoleh gelar sarjana merupakan hasil karya tulis saya sendiri. Adapun bagian-bagian tertentu dalam penulisan skripsi ini saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi-sanksi lainnya dengan peraturan yang berlaku, apabila di kemudian hari ditemukan adanya plagiat dalam skripsi ini.



HALAMAM PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS

HALAMAM PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR/SKRIPSI/TESIS UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Medan Area, saya yang bertanda tangan dibawah ini :

Nama : Denny Herdiansyah Chan
NPM : 188160019
Program Studi : Teknik Informatika
Fakultas : Teknik
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Medan Area Hak Bebas Royalti Noneksklusif (Non-Exclusive Royalti-Free Right) atas karya ilmiah saya yang berjudul :

Analisa Sentimen Sosial Media *Twitter* Terhadap Pilpres 2024 Menggunakan Algoritma *Naive Bayes*

Beserta perangkat yang ada (jika diperlukan). Dengan hak bebas royalti noneksklusif ini Universitas Medan Area berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data(database), merawat, dan mempublikasikan tugas akhir/skripsi/tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Demikian pernyataan ini saya buat dengan sebenarnya

Dibuat di : Medan

Pada tanggal : 12 September 2023

Yang menyatakan

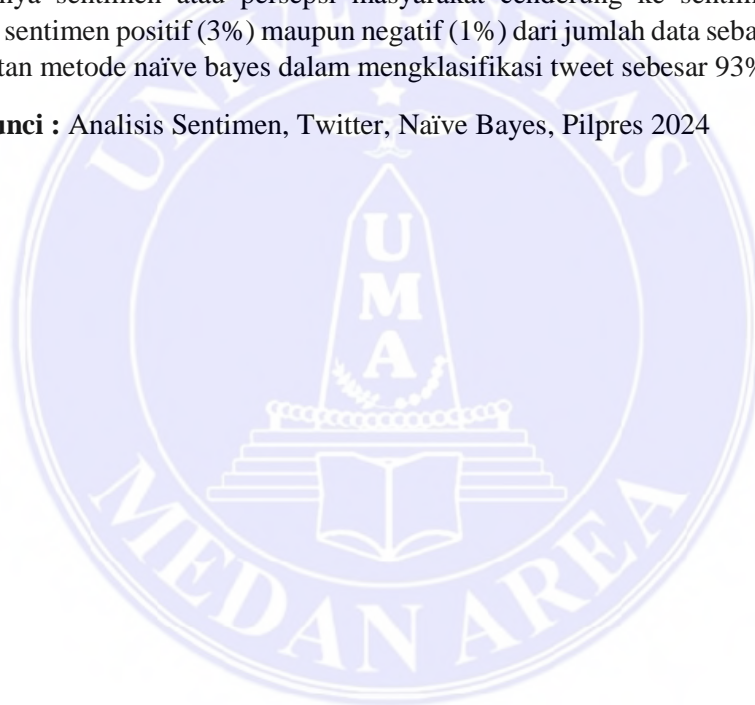


(Denny Herdiansyah Chan)

Abstrak

Twitter merupakan salah satu sosial media paling populer di kalangan pengguna internet saat ini. *Twitter* banyak digunakan sebagai platform untuk berbagi informasi terbaru, mempublikasikan opini, atau untuk berkomunikasi dengan orang lain. Isi *Twitter* yang disebut *tweet*, yang biasanya berisi informasi atau berita yang sedang viral. Pendapat yang diperoleh melalui *tweet* dapat digunakan oleh peneliti untuk menganalisis pendapat publik tentang tokoh politik yang mencalonkan diri sebagai calon presiden Indonesia tahun 2024. Untuk menerapkan teknik pembelajaran mesin dalam mengklasifikasikan polaritas pendapat publik dari jumlah sumber data yang sangat banyak. Teknik ini dapat dilakukan dengan menggunakan teks mining, salah satu fungsi dari teks mining adalah klasifikasi dokumen. Data diperoleh melalui scraping, yaitu dengan mengambil data langsung dari *tweet* menggunakan *Twitter API*. Data dikumpulkan dengan menggunakan *keyword* PILPRES 2024, sejumlah 11000 data diperoleh. Hasil yang didapatkan dari platform twitter bahwasanya sentimen atau persepsi masyarakat cenderung ke sentimen netral (93%) daripada sentimen positif (3%) maupun negatif (1%) dari jumlah data sebanyak 11000 data. Keakuratan metode naïve bayes dalam mengklasifikasi tweet sebesar 93%.

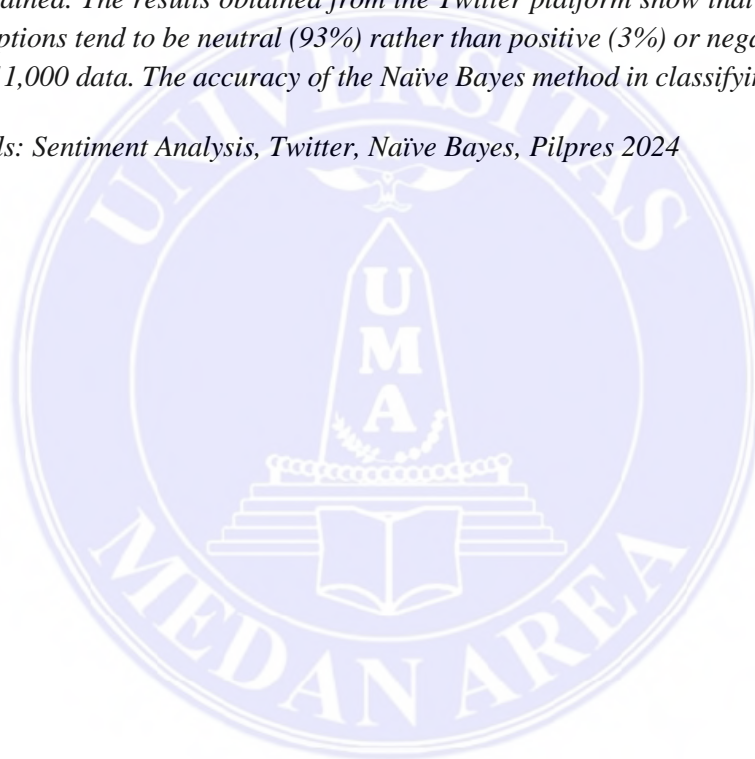
Kata Kunci : Analisis Sentimen, Twitter, Naïve Bayes, Pilpres 2024



Abstract

Twitter is one of the most popular social media among internet users today. Twitter is widely used as a platform to share the latest information, publish opinions, or to communicate with other people. Twitter content is called a tweet, which usually contains information or news that is currently viral. Opinions obtained through tweets can be used by researchers to analyze public opinion about political figures who are running for president of Indonesia in 2019. To apply machine learning techniques to classify the polarity of public opinion from a very large number of data sources. This technique can be done using text mining, one of the functions of text mining is document classification. the data is obtained through scraping, namely by taking data directly from tweets using the Twitter API. Data was collected using the 2024 PILPRES keyword, a total of 11,000 data were obtained. The results obtained from the Twitter platform show that public sentiment or perceptions tend to be neutral (93%) rather than positive (3%) or negative (1%) from a total of 11,000 data. The accuracy of the Naïve Bayes method in classifying tweets is 93%.

Keywords: Sentiment Analysis, Twitter, Naïve Bayes, Pilpres 2024

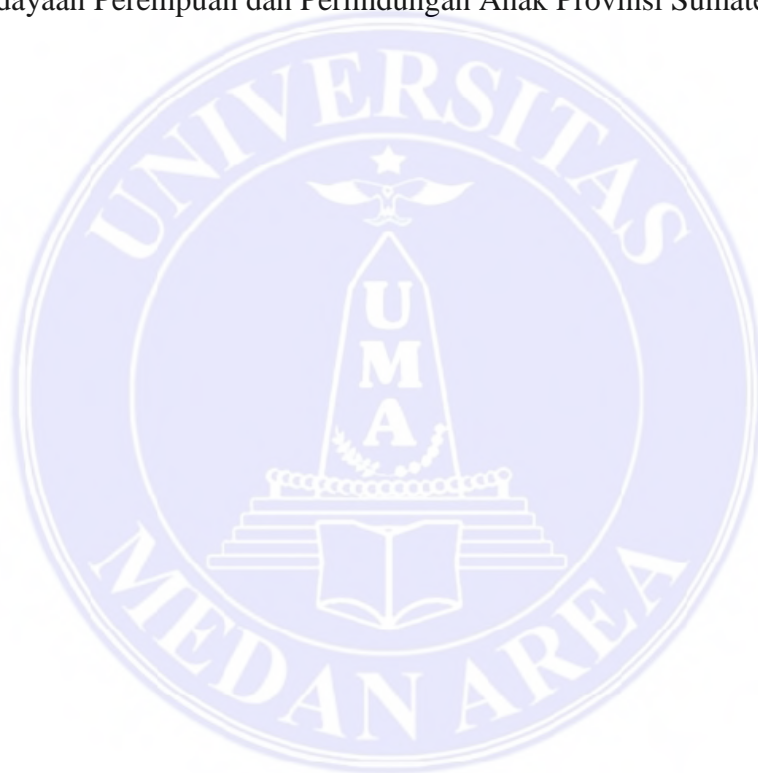


RIWAYAT HIDUP

Penulis merupakan putra atau anak pertama dari ayah Edy Sofyan dan Ibu Zuarnita yang dilahirkan di Medan, pada tanggal 11 September 2000.

Pada tahun 2018 Penulis lulus dari SMK Telkom Medan, lalu pada tahun yang sama penulis terdaftar sebagai mahasiswa Fakultas Teknik Prodi Teknik Informatika Universitas Medan Area.

Pada tahun 2021 Penulis melaksanakan Kerja Praktek (KP) di Dinas Pemberdayaan Perempuan dan Perlindungan Anak Provinsi Sumatera Utara



KATA PENGANTAR

KATA PENGANTAR

Segala puji dan syukur penulis ucapkan kepada Tuhan Yang Maha Esa oleh kasih karunia-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul "Analisa Sentimen Media Sosial *Twitter* Tentang Pilpres 2024 Menggunakan Metode *Naïve Bayes*". Tugas Akhir ini penulis susun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada program studi Teknik Informatika di Universitas Medan Area.

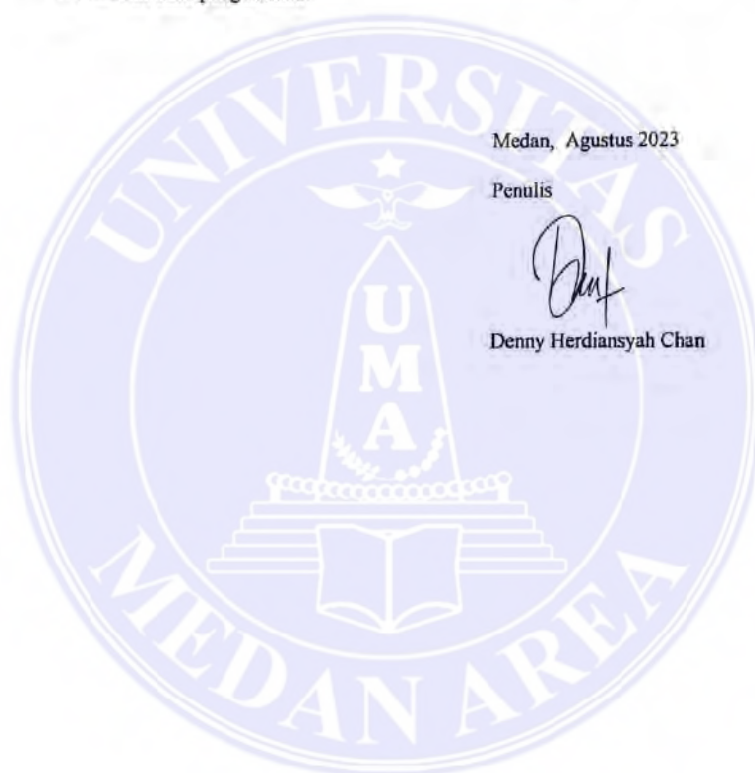
Dalam pengerjaan Tugas Akhir ini, penulis mendapatkan banyak bantuan dan dukungan dari berbagai pihak. Maka dari itu, pada kesempatan ini dengan segala kerendahan hati, penulis ingin menyampaikan terima kasih kepada:

1. Bapak Prof. Dadan Ramdan, M.Eng, M. Kom selaku Rektor Universitas Medan Area.
2. Bapak Dr. Rahmad Syah, S. Kom, M. Kom selaku Dekan Fakultas Teknik Universitas Medan Area.
3. Bapak Rizki Muliono, S. Kom, M. Kom selaku Ketua Program Studi Teknik Informatika sekaligus sebagai Sekretaris saya dalam ujian skripsi yang telah memberikan saran dalam penulisan tugas akhir ini.
4. Ibu Nurul Khairina, S. Kom, M. Kom selaku dosen pembimbing dan pembanding I yang telah meluangkan waktu dan dengan sabar membimbing saya dalam pembuatan Tugas Akhir ini sehingga dapat selesai.
5. Bapak Dr. Dian Noviandri, S.T, M.Kom sebagai ketua dalam ujian skripsi saya yang telah memberikan saran dalam penulisan Tugas Akhir ini.
6. Ibu Susilawati, S. Kom, M. Kom sebagai pembanding II saya pada ujian skripsi yang telah memberikan saran dalam penulisan Tugas Akhir ini.
7. Teristimewa kepada orang tua saya, ayahanda Edy Sofyan dan ibunda tersayang Zuarnita, yang selalu mendoakan, memberikan nasehat dan pengorbanannya, dari segi moril, dan materi kepada saya sehingga saya dapat menyelesaikan Tugas Akhir.
8. Teruntuk adik-adik saya yang memberikan dukungan penulis dalam

mengerjakan Tugas Akhir.

9. Kepada seluruh keluarga besar yang selalu memberikan semangat dan mengingatkan dalam mengerjakan Tugas Akhir.
10. Seluruh teman-teman Teknik Informatika angkatan 2018 dan seluruh teman-teman yang tidak dapat saya sebutkan satu persatu yang telah menemani, menghibur dan memotivasi selama mengerjakan Tugas Akhir ini.

Dalam penyusunan skripsi ini penulis menyadari sepenuhnya bahwa ini masih jauh dari kesempurnaan, oleh karena itu segala kritik dan saran yang bersifat membangun sangat penulis harapkan demi kesempurnaan skripsi ini dan untuk menambah ilmu pengetahuan.



Medan, Agustus 2023

Penulis

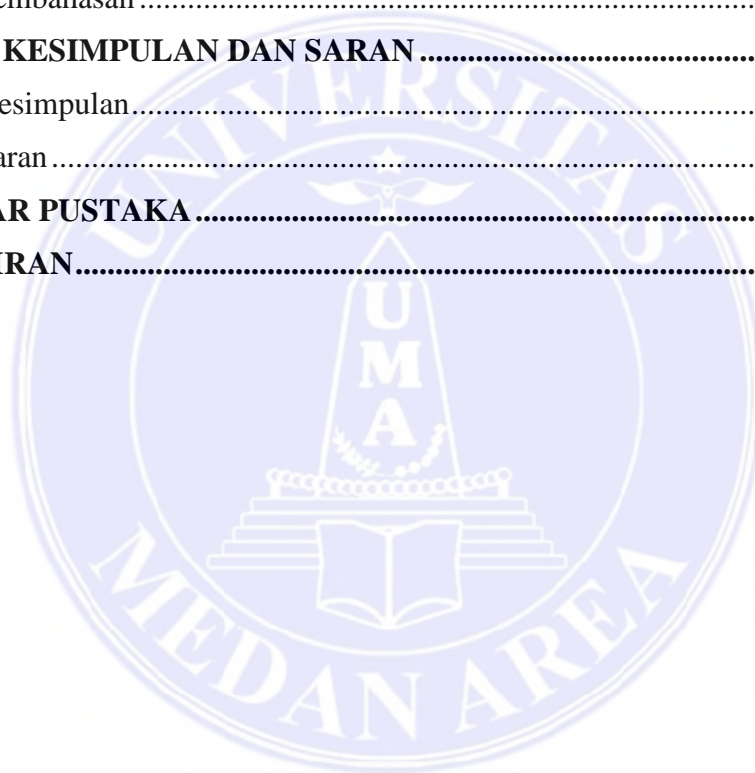
A handwritten signature in black ink, appearing to read 'Denny'.

Denny Herdiansyah Chan

DAFTAR ISI

HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
Abstrak	v
<i>Abstract</i>	vi
RIWAYAT HIDUP	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
BAB II LANDASAN TEORI	5
2.1 <i>Twitter</i>	5
2.2 <i>Data Mining</i>	5
2.3 TF – IDF.....	6
2.4 Analisa Sentimen.....	6
2.5 Pilpres / Pemilu	7
2.6 Klasifikasi <i>Naïve Bayes</i>	7
2.7 <i>Preprocessing</i>	8
2.8 <i>Python</i>	9
2.9 <i>Confusion Matrix</i>	10
2.10 <i>Flowchart</i>	11
2.11 Penelitian Terdahulu.....	13
BAB III METODOLOGI PENELITIAN	19
3.1 Metode Pengumpulan Data	19
3.2 Tahapan Penelitian	19
3.2.2 <i>Preprocessing</i>	20

3.2.3 Contoh Perhitungan Algoritma <i>Naïve Bayes</i>	29
BAB IV HASIL DAN PEMBAHASAN	32
4.1 Hasil.....	32
4.1.1 Deskripsi Data.....	32
4.1.2 <i>Text Preprocessing</i>	35
4.1.3 <i>TF-IDF</i>	38
4.1.4 <i>Naïve Bayes</i>	38
4.2 <i>Performance</i>	41
4.2.1 <i>Wordcloud</i>	41
4.3 Pembahasan	43
BAB V KESIMPULAN DAN SARAN	44
5.1 Kesimpulan.....	44
5.2 Saran	44
DAFTAR PUSTAKA	45
LAMPIRAN.....	47



DAFTAR TABEL

Tabel 3.1 Nilai dari TF-IDF	25
Tabel 3.3 Contoh data pemilihan umum Presiden Indonesia 2024.....	29
Tabel 3.4 Contoh perhitungan data negatif	30
Tabel 3.5 Contoh perhitungan data positif	31
Tabel 3.6 Keterangan mengenai table 3.2 dan 3.3	31



DAFTAR GAMBAR

Gambar 4.1 Scrapping Data	32
Gambar 4.2 Mendaftar pada Developer Twitter	33
Gambar 4.3 Kode Python Authentication Token	33
Gambar 4.4 Kode Python Crawling Data	34
Gambar 4.5 Tampilan Data	34
Gambar 4.6 Labelling Otomatis	35
Gambar 4.7 Proses Case folding	36
Gambar 4.8 Proses Tokenization	36
Gambar 4.9 Proses Stopwords	37
Gambar 4.10 Proses Stemming	37
Gambar 4.11 Tahap TF-IDF	38
Gambar 4.12 Tahap Pengujian	38
Gambar 4.13 Hasil Presisi dan Recall	39
Gambar 4.14 Confusion Matrix	39
Gambar 4. 15 Hasil Klasifikasi Menggunakan Naïve Bayes	40
Gambar 4.16 Wordcloud Positif	41
Gambar 4.17 Wordcloud negatif	42
Gambar 4.18 Wordcloud netral	42

BAB I PENDAHULUAN

1.1 Latar Belakang

Indonesia sebagai salah satu negara yang menganut sistem demokrasi, yang dibuktikan dengan adanya pemilihan umum untuk memilih pemimpin baru. Dalam semua pemilihan ini, setiap calon tokoh politik mempertimbangkan popularitasnya di mata publik. Perkembangan teknologi informasi yang sangat pesat telah mengubah kecenderungan netizen untuk mengungkapkan pendapatnya di media sosial, salah satunya adalah *Twitter*. *Twitter* merupakan salah satu sosial media paling populer di kalangan pengguna internet saat ini. *Twitter* banyak digunakan sebagai platform untuk berbagi informasi terbaru, mempublikasikan opini, atau untuk berkomunikasi dengan orang lain.(Silitonga & Sihotang, 2019)

Isi *Twitter* yang disebut *tweet*, yang biasanya berisi informasi atau berita yang sedang viral. *Tweet* banyak digunakan untuk memposting hal-hal tentang pengguna dan berbagi informasi. Isi *tweet* juga bisa mengungkapkan perasaan atau suasana hati penggunanya, yang merupakan pendapat yang subjektif. Pendapat yang diperoleh melalui *tweet* dapat digunakan oleh peneliti untuk menganalisis pendapat publik tentang tokoh politik yang mencalonkan diri sebagai calon presiden Indonesia tahun 2024. Analisis sentimen pada *tweet* dapat digunakan untuk mengetahui tingkat popularitas tokoh politik tersebut di mata publik, sehingga dapat membantu dalam pengambilan keputusan politik.(Silitonga & Sihotang, 2019)

Meskipun ada banyak penelitian tentang analisis sentimen media sosial terkait politik di berbagai negara, penelitian tentang sentimen masyarakat terhadap Pilpres 2024 di Indonesia masih terbatas. Oleh karena itu, penelitian ini akan mencoba mengisi celah tersebut dengan melakukan analisis sentimen terhadap percakapan di Twitter terkait Pilpres 2024.

Naïve Bayes atau multinomial *naïve bayes* adalah suatu teknik yang digunakan untuk mengklasifikasikan sebuah kumpulan dokumen. Algoritma ini memanfaatkan prinsip probabilitas dan statistik yang dikemukakan oleh Thomas Bayes, seorang ilmuwan asal Inggris. Metode *Naïve Bayes* terdiri dari dua tahap

dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap pengujian (klasifikasi).(Hidayat,2023) Algoritma ini didasarkan pada teori probabilitas dan memiliki kemampuan untuk mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral.

Penelitian ini bertujuan untuk melakukan analisis terhadap sentimen yang ada di *Twitter* terkait Pilpres 2024 di Indonesia dengan menggunakan algoritma Naive Bayes. Data yang akan digunakan dalam penelitian ini mencakup berbagai jenis teks seperti *tweet*, *retweet*, dan balasan, yang akan dianalisis untuk mengidentifikasi sentimen yang terkandung dalam teks-teks tersebut.

Diharapkan bahwa hasil penelitian ini akan memberikan pemahaman yang lebih mendalam mengenai bagaimana masyarakat Indonesia merespon dan memandang Pilpres 2024 melalui media sosial. Informasi yang diperoleh dari penelitian ini memiliki potensi untuk memberikan manfaat kepada peneliti politik, praktisi kampanye politik, dan masyarakat umum dalam memahami dinamika politik yang sedang berlangsung.

Selain itu, penelitian ini juga memiliki potensi untuk menjadi landasan dalam pengembangan sistem analisis sentimen otomatis yang dapat digunakan untuk memantau dan mengukur sentimen masyarakat secara langsung selama periode kampanye Pilpres 2024.

Oleh karena itu, penelitian ini memiliki relevansi yang kuat dalam bidang ilmu politik, teknologi informasi, dan media sosial, dan memiliki potensi untuk memberikan kontribusi yang positif dalam memahami serta melaksanakan proses pemilihan umum di Indonesia.

1.2 Rumusan Masalah

Adapun rumusan masalah pada penelitian ini adalah :

1. Bagaimana persepsi masyarakat terhadap calon-calon presiden yang akan bersaing di pemilu 2024 menurut *tweet* di media sosial *Twitter*?

1.3 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah :

1. Penelitian ini melakukan analisis sentimen terhadap *tweet* di media sosial *Twitter* yang berisi tentang Pilpres 2024.
2. Penelitian ini menganalisis *tweet* yang terkait dengan calon-calon presiden yang akan bersaing di pemilu 2024.
3. Penelitian ini menggunakan data kurang lebih sebanyak 11000 data dari *twitter*
4. Penelitian ini dilakukan dalam waktu bulan Mei-Juni 2023
5. Penelitian ini bertujuan untuk melakukan evaluasi dengan menggunakan *confusion matrix*

1.4 Tujuan Penelitian

Adapun tujuan penelitian pada penelitian ini adalah :

1. Untuk mengetahui persepsi masyarakat terhadap calon-calon presiden yang akan bersaing di pemilu 2024 menurut *tweet* di media sosial *Twitter*.
2. Untuk mengetahui akurasi, presisi, *recall* dan *F1 score* metode *naïve bayes* dalam mengklasifikasikan *tweet* menjadi positif, negatif, atau netral.
3. Untuk mengetahui apakah ada perbedaan sentimen terhadap calon-calon presiden yang akan bersaing di pemilu 2024 menurut *tweet* di media sosial *Twitter*.

1.5 Manfaat Penelitian

Adapun manfaat penelitian pada penelitian ini adalah :

1. Hasil dari penelitian ini dapat memberikan informasi mengenai persepsi masyarakat terhadap calon-calon presiden yang akan bersaing di pemilu 2024 menurut *tweet* di media sosial *Twitter*.

2. Penelitian ini dapat menjadi referensi bagi pihak-pihak yang ingin melakukan analisis sentimen terhadap *tweet* di media sosial menggunakan metode *Naive Bayes*.
3. Hasil dari penelitian ini dapat memberikan masukan bagi calon-calon presiden yang akan bertarung di pemilu 2024 mengenai bagaimana persepsi masyarakat terhadap mereka.
4. Penelitian ini dapat memberikan informasi bagi para pemilih dalam mengambil keputusan di pemilu 2024 nanti.



BAB II LANDASAN TEORI

2.1 Twitter

Twitter adalah platform media sosial yang bisa diakses melalui *smartphone*, *iOS*, *Windows Phone*, *Blackberry*, dan juga perangkat komputer atau laptop. Saat ini, *Twitter* menjadi salah satu situs jaringan sosial yang sedang tren. *Twitter* adalah salah satu platform jaringan sosial yang paling populer. Di *Twitter*, kamu dapat berbagi segala hal dalam bentuk pesan, yang biasanya disebut tweet. *Twitter* awalnya diperkenalkan oleh Jack Dorsey, seorang mahasiswa sarjana dari Universitas New York, dalam sebuah diskusi di acara yang diadakan oleh perusahaan podcast bernama Odeo. Jack Dorsey memberikan penjelasan singkat tentang cara penggunaannya. Pada tahap awal, platform ini dikenal sebagai "twtrr", yang kemudian dijelaskan oleh Evan Williams dan Noah Glass.

Aplikasi *Twitter* mulai populer pada tahun 2007 di konferensi South by Southwest Interactive. Selama acara tersebut berlangsung, pengguna *Twitter* meningkat dari 20 ribu tweet per hari menjadi 60 ribu tweet per hari. Pada tanggal 22 Januari 2010, astronot NASA bernama T. J. Creamer mengirimkan tweet pertama dari luar angkasa. (Dewi & Larasati, 2023)

2.2 Data Mining

Data mining adalah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin untuk mengekstraksi dan mengidentifikasi informasi yang berharga serta pengetahuan yang terkait dari berbagai database yang besar. Analisis asosiasi, atau *association rules mining*, merupakan teknik dalam data mining yang digunakan untuk menemukan aturan asosiasi antara kombinasi item yang berhubungan.

Data mining merupakan proses penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah besar data. Dalam konteks ini, data mining juga sering disebut sebagai *knowledge discovery in database* (KDD), yang mencakup pengumpulan dan penggunaan data historis untuk

menemukan keteraturan, pola, atau hubungan dalam kumpulan data yang besar.(Mawarni & Silegar, 2023)

2.3 TF – IDF

Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode yang digunakan untuk menentukan bobot kata berdasarkan frekuensi dokumen (DF). Jika suatu kata muncul lebih sering dalam dokumen, maka kepentingan kata tersebut dianggap rendah, sehingga bobot yang diberikan kecil. Sebaliknya, jika kemunculan kata tersebut jarang, bobot yang diberikan menjadi besar, menunjukkan bahwa kata tersebut memiliki kepentingan yang lebih tinggi.

Frekuensi kemunculan suatu kata dalam dokumen menggunakan semua kata-kata yang ada dalam teks, sementara fitur yang digunakan mencakup semua kata pada dataset. *Term Frequency* (TF) digunakan untuk menghitung bobot kata dalam sebuah dokumen. Jika suatu kata muncul lebih sering, bobotnya akan lebih besar. Sebaliknya, *Inverse Document Frequency* (IDF) mengukur sejauh mana kata tersebar dalam dokumen. Penyebaran kata yang tidak sesuai dapat memengaruhi hasil perhitungan bobot dokumen. (Sibaroni, 2023.)

2.4 Analisa Sentimen

Analisis sentimen, yaitu metode untuk memahami, mengekstrak, dan memproses data teks secara otomatis untuk mendapatkan sentimen yang terkandung dalam opini. Analisis sentimen merupakan salah satu aplikasi dari teknik data mining yang bertujuan untuk mengekstrak informasi yang berguna dari data teks. Dengan menggunakan analisis sentimen, kita dapat memahami sentimen yang terkandung dalam opini atau pendapat yang diberikan oleh pengguna, misalnya positif, negatif, atau netral. Analisis sentimen berguna untuk mengukur tingkat kepuasan pengguna terhadap suatu produk atau layanan, atau untuk mengukur opini masyarakat terhadap suatu peristiwa atau topik tertentu.(Sari dkk., 2020)

2.5 Pilpres / Pemilu

Pemilihan umum di Indonesia, sebagai manifestasi kedaulatan rakyat dan bagian dari sistem demokrasi, harus diadakan sesuai dengan prinsip langsung, umum, bersih, jujur, dan adil, sebagaimana diamanatkan oleh Pasal 22 E ayat (1) Undang Undang Dasar 1945. Namun, dalam pelaksanaan pemilu, munculnya beberapa masalah penegakan hukum disebabkan oleh faktor-faktor tertentu. Pertama, batasan mengenai terjadinya pelanggaran sering tidak pasti, sehingga menciptakan multitafsir dan asumsi publik yang dapat mengakibatkan kontroversi. Kedua, mekanisme dan prosedur penanganan pelanggaran tidak selalu jelas, membuat penanganannya menjadi sulit. Ketiga, sanksi hukum yang tercantum dalam Pasal 523 belum sepenuhnya efektif dalam menciptakan efek jera bagi pelaku praktik politik uang dalam pemilihan umum di Indonesia, sehingga para pelaku sering kali melakukan tindakan tersebut kembali. (A'dhim F, 2023)

2.6 Klasifikasi *Naïve Bayes*

Naïve Bayes Classifier adalah algoritma dalam data mining yang menggunakan teorema Bayes untuk melakukan klasifikasi. Metode ini adalah pengklasifikasi probabilitas sederhana yang menghitung peluang dengan menambahkan frekuensi dan kombinasi nilai dari dataset. *Naïve Bayes Classifier* didasarkan pada asumsi bahwa nilai-nilai variabel adalah saling bebas jika nilai outputnya diberikan. Persamaan teorema Bayes dituliskan sebagai berikut: (Yolanda Paramitha dkk, 2023)

$$P(X|Y) = P(Y|X) * P(X) / P(Y) \quad (2.1)$$

dimana:

$P(Y|X)$: Peluang terjadinya Y berdasarkan kondisi X (posteriori prob)

$P(Y)$: Peluang terjadinya Y (prior prob)

$P(X|Y)$: Peluang terjadinya X berdasarkan kondisi pada hipotesis Y

$P(X)$: Peluang terjadinya X

2.7 Preprocessing

Data *preprocessing*, yaitu teknik data mining yang bertujuan untuk mengubah data mentah menjadi format yang mudah dipahami. Data *preprocessing* merupakan langkah penting dalam proses data mining karena data mentah biasanya tidak dapat langsung digunakan untuk analisis. Data mentah biasanya terdiri dari data yang tidak terstruktur, bersifat redundan, atau memiliki nilai-nilai yang hilang. Dengan melakukan data *preprocessing*, masalah-masalah tersebut dapat diselesaikan sehingga data dapat digunakan untuk analisis. Langkah-langkah dalam data *preprocessing* meliputi *tokenisasi*, *case folding*, *stemming*, *filtering*, dan *labelling*. Dengan menjalankan semua langkah tersebut, data mentah akan dikonversi menjadi bentuk yang lebih mudah dipahami dan dapat digunakan untuk analisis. (Gunawan dkk., 2018)

1. Tokenisasi

Tokenization memiliki fungsi untuk membagi kalimat menjadi kata-kata yang memiliki makna, sehingga dapat meningkatkan akurasi data. Dalam proses ini, digunakan pustaka Python yang disebut `nlk.tokenize`. (Khatib Sulaiman dkk)

2. Case Folding

Case folding, Langkah pertama dalam tahapan preprocessing adalah melakukan case folding. Pada sebuah data, seringkali terdapat variasi penulisan huruf yang tidak konsisten, seperti adanya huruf kapital dan huruf kecil yang bercampur. Oleh karena itu, perlu dilakukan konversi agar semua huruf menjadi huruf kecil sehingga data memiliki format standar. Proses ini dikenal sebagai case folding, yang berguna untuk mengubah alfabet menjadi huruf kecil atau lower case. (Hidayati dkk., 2023)

3. Stemming

Stemming adalah proses untuk menyaring kata-kata yang termasuk kata sambung, kata ganti, atau kata depan, dengan tujuan menghasilkan kata dasar dengan menghilangkan awalan atau akhiran yang tidak penting. (Rizki Fakhri dkk, 2023)

4. *Filtering*

Filtering adalah langkah dalam tahapan preprocessing yang bertujuan untuk mengeliminasi kata-kata yang sering muncul, tidak memiliki makna deskriptif, atau tidak relevan dengan topik yang dibahas. Dengan melakukan *filtering*, kita dapat memperoleh hanya kata-kata yang penting dan relevan. (Febriyani & Februariyanti, 2023)

5. *Labelling*

Pada tahap labeling, dilakukan penandaan atau pemberian label (kelas) berdasarkan fitur atau karakteristik yang terdapat dalam sebuah dokumen atau kalimat. Pada fase ini, teks ulasan yang telah melalui tahap *preprocessing* dan menghasilkan teks yang bersih akan diberikan kategori positif atau negatif. Kelas positif digunakan untuk ulasan yang mengandung pernyataan yang baik dan mendukung, sedangkan kelas negatif digunakan untuk ulasan yang berisi koreksi, komentar, atau ekspresi kekecewaan terhadap pelayanan yang diterima. (Salim & Syafrullah, 2023)

6. *Stopword*

Stopword, yaitu kata-kata yang sering muncul dalam teks namun tidak memiliki signifikansi dalam konteks tertentu. *Stopword* umumnya tidak memberikan informasi berarti dalam proses pemodelan topik atau analisis lainnya, bahkan dapat menghambat proses analisis. Untuk menghilangkan *stopword*, kita dapat memanfaatkan pustaka seperti NLTK (*Natural Language Toolkit*) atau *spaCy* di *Python*. Dengan menggunakan pustaka tersebut, kita dapat dengan mudah mengakses daftar *stopword* yang telah disediakan dan menghapusnya dari teks. Meskipun demikian, perlu diingat bahwa dalam beberapa situasi, beberapa *stopword* mungkin memiliki makna yang penting dalam konteks tertentu, sehingga penghapusan *stopword* perlu dilakukan dengan hati-hati. (Sains & Teknologi, 2023)

2.8 *Python*

Python adalah bahasa pemrograman berbasis objek yang memungkinkan interaksi secara interaktif. Bahasa ini memiliki struktur data tingkat tinggi. Sebagai

bahasa pemrograman interpretatif, *Python* menonjol dengan berbagai fungsi dan didesain dengan fokus pada kejelasan serta kemudahan pemahaman kode. Python diakui sebagai bahasa yang menggabungkan kemampuan dan sintaks kode yang jelas. Dirancang khusus untuk memberikan kemudahan kepada para programmer dalam pembuatan program dengan efisiensi waktu, kemudahan pengembangan, dan kompatibilitas dengan berbagai sistem. Python dapat digunakan untuk membuat aplikasi mandiri maupun untuk pemrograman skrip. (Triono dkk, 2023)

2.9 Confusion Matrix

Confusion matrix adalah metode yang digunakan dalam data mining untuk membuat perhitungan yang akurat untuk konsep klasifikasi. Evaluasi dengan matriks kebingungan memberikan nilai presisi, akurasi, dan daya ingat. Akurasi klasifikasi adalah persentase akurasi dari rekaman yang terklasifikasi dengan benar setelah dilakukan pengujian terhadap hasil klasifikasi. Akurasi adalah proporsi kasus yang diprediksi secara positif yang juga benar-benar positif dalam data aktual. *Recall* atau sensitivitas adalah proporsi kasus positif sejati yang diprediksi dengan benar menjadi positif. Dengan menggunakan *confusion matrix*, kita dapat mengevaluasi performa dari suatu model klasifikasi dan mengidentifikasi kelemahan atau kekuatannya. (Mayadewi & Rosely)

Correct Classification	Classified as	
	+	-
+	True positives (A)	False negatives (B)
-	False positives (C)	True negatives (D)

Gambar 2.1 Rumus *Confusion Matrix*

Perhitungan akurasi dengan tabel *confusion matrix* adalah sebagai berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} * 100\% \quad (2.2)$$

Akurasi didefinisikan sebagai rasio elemen penting yang dipilih untuk semua elemen yang dipilih. Akurasi dapat diartikan sebagai kecocokan antara permintaan informasi dan tanggapan terhadap permintaan tersebut. Formula presisinya adalah:

$$Presicion = \frac{TP}{TP+FP} * 100\% \quad (2.3)$$

Recall didefinisikan sebagai rasio item signifikan yang dipilih terhadap jumlah total item signifikan yang tersedia. *Recall* dihitung menggunakan rumus berikut:

$$Recall = \frac{TP}{TP+FN} * 100\% \quad (2.4)$$




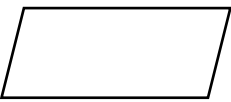
Precision dan *Recall* dapat ditentukan sebagai persentase (1-100%) atau sebagai nilai numerik antara 0-1. Sistem pemberi rekomendasi dianggap baik ketika skor presisi dan *recall* tinggi.

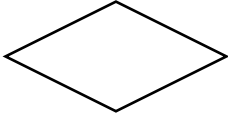
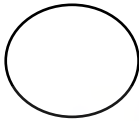





$$F1 \text{ Score} = \frac{2*(recall*precision)}{(recall+precision)} * 100\% \quad (2.5)$$


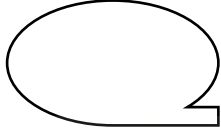
2.10 Flowchart

Diagram alir, yang juga dikenal dengan sebutan *flowchart*, adalah jenis diagram yang menunjukkan tahapan-tahapan yang harus dilakukan dalam suatu algoritma atau sistem. (Rosaly & Prasetyo, 2019)

Tabel 2.1 Simbol – Simbol *Flowchart*

Simbol	Nama	Fungsi
	<i>Terminator</i>	Untuk memulai atau mengakhiri suatu proses
	<i>Connecting Line</i>	Digunakan untuk menghubungkan simbol-simbol yang satu dengan yang lainnya dan menunjukkan langkah-langkah dalam suatu proses.
	<i>Processing</i>	Proses pengolahan akan dilakukan di dalam komputer
	<i>Input / Output</i>	Menyatakan bahwa input dan output akan terjadi

		tanpa memperhatikan jenisnya
	<i>Decision</i>	menyatakan bahwa proses yang sedang dilakukan akan dipilih berdasarkan kondisi tertentu.
	<i>Connector</i>	Bahwa proses yang sedang berlangsung akan terus berlanjut ke proses berikutnya di halaman yang sama.
	<i>Offline Connector</i>	Proses yang sedang berlangsung akan terus berlanjut ke proses berikutnya di halaman yang berbeda
	<i>Manual Operation</i>	Menyatakan bahwa proses pengolahan tidak dapat dilakukan oleh komputer.
	<i>Display</i>	Menyatakan bahwa keluaran akan ditampilkan melalui layar monitor.
	<i>Manual Input Symbol</i>	Menyatakan bahwa data akan dimasukkan secara manual menggunakan keyboard.
	<i>Predefined Process</i>	Menyatakan bahwa penyimpanan akan

		disiapkan dengan memberikan harga awal.
	<i>Document</i>	Menyatakan bahwa masukan dan keluaran akan berasal dari dokumen.
	<i>Magnetic Tape</i>	menyatakan bahwa masukan dan keluaran akan berasal dari pita magnetis.

(Sumber : (Rosaly & Prasetyo,2019.)

2.11 Penelitian Terdahulu

Penelitian terdahulu merupakan referensi dasar untuk melakukan penelitian, dimana membandingkan penelitian terdahulu dengan penelitian yang sedang dilakukan, dengan fungsi sumber informasi serta memperdalam teori yang sedang digunakan.

Berikut ini merupakan beberapa penelitian terdahulu yang membantu atau pendukung didalam penelitian yang peneliti lakukan saat ini yaitu akan ditampilkan dalam bentuk table berikut ini :

Tabel 2.2 Penelitian Terdahulu

No	Peneliti (Tahun)	Judul	Hasil Penelitian
1	(Aji Andika & Amalia Nur Azizah, 2019)	Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter	Model Naive Bayes yang dihasilkan memiliki tingkat akurasi sebesar 82,90% dengan $\alpha = 0,05$. Hasil klasifikasi menunjukkan bahwa 34,5% (471) tweet dianggap positif dan

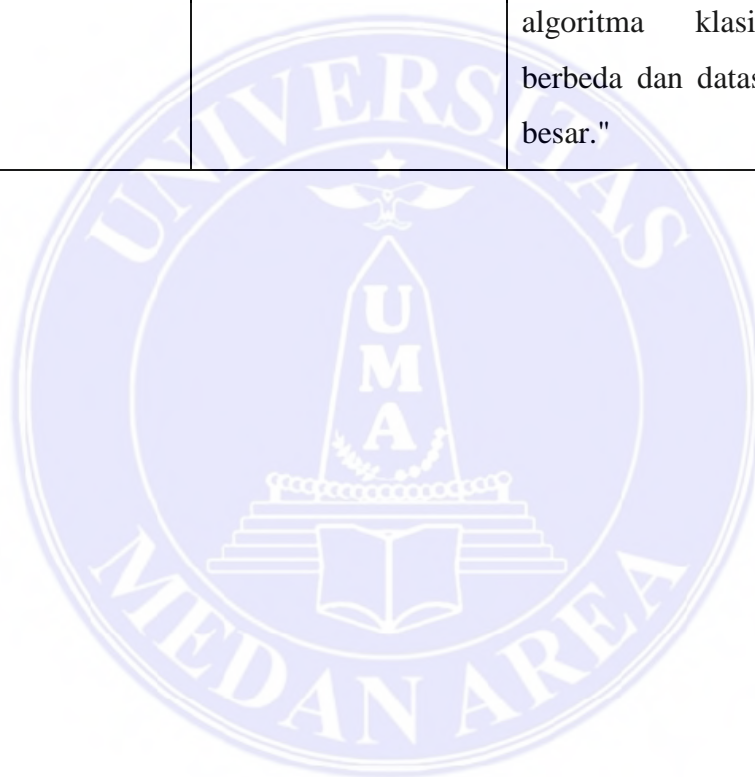
		Menggunakan Metode Naive Bayes Classifier	65,5% (895) tweet dianggap negatif terhadap hasil quick count.
2	(Dhimas Agusti dkk,2019)	Analisis Sentimen Isu PILPRES 2019 Dengan Menggunakan Algoritma Naive Bayes	Observasi terhadap sentimen publik pada pemilu presiden tahun 2019 menunjukkan peningkatan jumlah tweet di Twitter saat proses pemilu berlangsung. Antusiasme publik tampak meningkat, seperti pada saat debat dimulai. Dengan menganalisis tweet, dapat diperoleh informasi tentang seberapa antusias masyarakat terhadap pemilu, serta sentimen yang positif dan negatif terhadap proses pemilu, calon presiden dan wakil presiden. Hasil analisis polaritas sentimen terhadap proses pemilu dan calon presiden menunjukkan bahwa sekitar 80,8% sentimen positif dan 19,2% sentimen negatif. Dengan menggunakan metode Naive Bayes, hasil perhitungan sentimen analisis cukup baik.
3	(Silitonga & Sihotang, 2019)	Analisis Sentimen Pemilihan Presiden Indonesia Tahun 2019 Di Twitter Berdasarkan	Proses konfigurasi Naive Bayesian Classifier untuk pra-pemilihan presiden tahun 2019 dimulai dengan pengumpulan data manual dari Twitter. Data

		<p>Geolocation Menggunakan Metode Naive Bayesian Classification</p>	<p>tersebut kemudian diproses dan diklasifikasikan menggunakan aplikasi WEKA dengan perhitungan Naive Bayesian Classifier terhadap probabilitas kata. Hasil probabilitas tertinggi digunakan untuk mengategorikan kalimat menjadi sentimen yang sebenarnya. Sentimen tweet berdasarkan wilayah diperoleh melalui koordinat geografis yang terhubung dengan aplikasi Twitter menggunakan API Geolocation yang terdiri dari garis lintang dan garis bujur pada Google Maps yang telah ditentukan oleh pengguna. Perbandingan klasifikasi Naive Bayesian Classifier dengan algoritma Tree.J48 menunjukkan bahwa Naive Bayesian Classifier memiliki tingkat akurasi yang lebih tinggi (77,62%) dibandingkan Tree.J48 (70%).</p>
4	(Juanita, 2020)	<p>Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes</p>	<p>Data tweet dikumpulkan dari tanggal 16 April 2018 hingga 16 April 2019 dan menghasilkan 221 tweet. Kemudian, tweet tersebut diberi label oleh 2 orang ahli komunikasi untuk</p>

			<p>menentukan sentimen positif, negatif, dan netral. Data kemudian dibagi menjadi data latih (70%, atau 154 tweet) dan data uji (30%, atau 67 tweet). Setelah proses preprocessing dan transformasi teks, data set pola persepsi terbentuk untuk analisis sentimen menggunakan modeling klasifikasi Naive Bayes. Berdasarkan data set pola persepsi tersebut, ditemukan bahwa sentimen negatif pada data tweet adalah 52%, jauh lebih besar daripada sentimen positif (18%) dan netral (31%). Tingkat akurasi klasifikasi Naive Bayes terhadap data latih bernilai 81%, dan 76% pada data uji. Nilai precision rata-rata untuk sentimen positif adalah 86,65%, sentimen negatif 77,15%, dan sentimen netral 80,95%. Nilai recall rata-rata untuk sentimen positif adalah 36,8%, sentimen negatif 93,2%, dan sentimen netral 86,8%.</p>
5	(Heri prasetyo & Arif Senja Fitriani, 2023)	Analisa Sentimen Jelang Pilpres 2024 Menggunakan Naive Bayes Classifier	Hasilnya adalah 3.021 tweet yang berhasil diekstraksi dalam rentang waktu 1 Desember 2022 hingga 31 Januari 2023, dan data ini diubah menjadi format

		<p>Berdasarkan Opini Publik Di Twitter</p>	<p>dokumen dalam bentuk '.csv'. Setelah melakukan preprocessing menggunakan bahasa pemrograman Python, diperoleh 2.157 tweet yang telah dibersihkan. Data yang telah melalui proses preprocessing kemudian diberi label positif dan negatif, sehingga menjadi data yang memiliki sentimen. Analisis sentimen dilakukan dengan menggunakan algoritma Naïve Bayes Classifier dalam tiga percobaan berbeda dengan pembagian data uji dan data latih yang berbeda pada setiap percobaan. Hasil penelitian menunjukkan bahwa model Naïve Bayes terbaik ditemukan pada percobaan pertama, di mana data uji sebesar 10% dan data latih sebesar 90%, dengan tingkat akurasi sebesar 71%, presisi 93%, recall 66%, dan nilai F-measure sebesar 77%. Kesimpulan dari penelitian ini adalah bahwa elektabilitas pemilihan presiden 2024 memiliki dampak pada opini publik yang tercermin dalam tweet yang memiliki sentimen positif, negatif, dan</p>
--	--	--	--

			<p>perbincangan di Twitter. Sentimen positif memiliki persentase lebih tinggi, yaitu 71,5% (1.543 tweet), sementara sentimen negatif sebesar 28,5% (614 tweet). Penelitian selanjutnya diharapkan dapat menghasilkan informasi yang berbeda dengan menggunakan algoritma klasifikasi yang berbeda dan dataset yang lebih besar."</p>
--	--	--	--



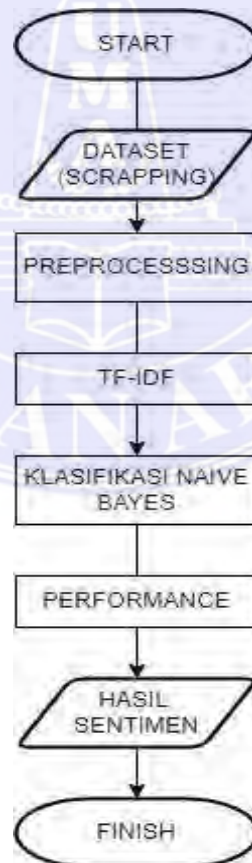
BAB III METODOLOGI PENELITIAN

3.1 Metode Pengumpulan Data

Untuk menyelesaikan penelitian ini, data diperoleh melalui scraping, yaitu dengan mengambil data langsung dari *tweet* menggunakan *Twitter API*. Data dikumpulkan dengan menggunakan *keyword* PILPRES 2024, sejumlah 11000 data diperoleh.

3.2 Tahapan Penelitian

Tahapan penelitian berikut ini dapat dilihat melalui *flowchart* berikut ini mulai dari pengumpulan data (*Scrapping*) hingga menghasilkan data sentimen.



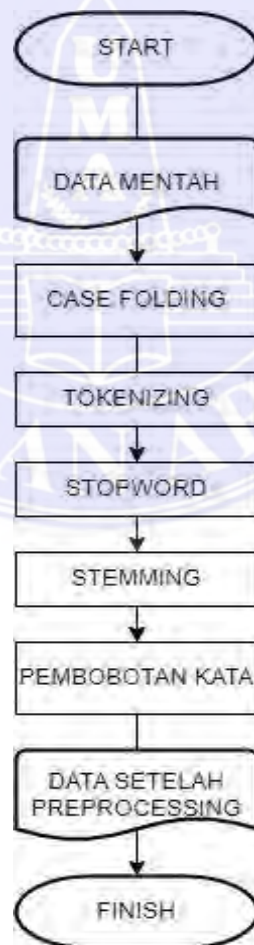
Gambar 3.1 Tahapan Penelitian

Pada gambar 3.1 terdapat tahapan-tahapan penelitian sebagai berikut yaitu :

1. Data set (*scraping*) : dibagian ini peneliti mengambil data dari *twitter*
2. *Preprocessing* : dibagian ini peneliti mengumpulkan data untuk diolah sebelum menganalisis
3. *Naive Bayes Classifier* : dibagian ini peneliti mengklasifikasi data menggunakan algoritma *naive bayes* untuk menentukan kelas dari data baru berdasarkan data latih yang telah diketahui.
4. Hasil sentimen : dibagian ini peneliti mendapatkan hasil sentimen yaitu berupa sentimen positif, negatif, ataupun Netral

3.2.2 Preprocessing

Dalam tahap *processing* berikut ini dapat dilihat melalui *flowchart* berikut ini mulai dari mencari data mentah sampai mendapatkan hasilnya.



Gambar 3.2 Preprocessing

Pada gambar 3.2 terdapat tahapan-tahapan untuk melakukan *preprocessing* yaitu :

No	Label	<i>Tweet</i>
1	-	1 Jokowi aja bisa tambah utang banyak banget gimana ditambah 1 lagi yg mirip die? Bisa2 rakyat mati kelaparan
2	+	Pemimpin yang merakyatnya dari hati.. cocok sih kalau mau nyapres di 2024 kayaknya bakal banyak yang dukung apalagi kaum perempuan @BantengGirl
3	-	Waspada! Anies Baswedan didukung kaum intoleran, kaum radikal, dan para pengasong
4	+	Menteri BUMN @erickthohir menjadi calon wakil presiden (cawapres) pilihan masyarakat di Pilpres 2024. Dan dinilai sebagai sosok yang cerdas dan memiliki kinerja sangat baik juga merupakan sosok yang layak menjadi panutan dan suri tauladan bangsa Indonesia.
5	-	Anis merupakan sosok yang tidak cocok untuk menjadi Pilpres 2024 dikarenakan "Tukang tipu" alia #TenggelamkanAniesNasdem Nusa Tenggara Barat

1. Data mentah : dibagian ini sistem mengambil data mentah yang ada di twitter
2. *Case folding* : dibagian ini sistem melakukan proses konversi teks menjadi satu kasus yang sama.

No	Label	<i>Case Folding</i>
1	-	jokowi aja bisa tambah utang banyak banget gimana ditambah lagi yg mirip die bisa rakyat mati kelaparan
2	+	pemimpin yang merakyatnya dari hati cocok sih kalau mau nyapres di kayaknya bakal banyak yang dukung apalagi kaum perempuan bantenggirl
3	-	waspada anies baswedan didukung kaum intoleran kaum radikal dan para pengasong
4	+	menteri bumn erickthohir menjadi calon wakil presiden cawapres pilihan masyarakat di pilpres dan dinilai sebagai sosok yang cerdas

		dan memiliki kinerja sangat baik juga merupakan sosok yang layak menjadi panutan dan suri tauladan bangsa indonesia
5	-	anis merupakan sosok yang tidak cocok untuk menjadi pilpres dikarenakan tukang tipu alia tenggelamkan anies nasdem nusa tenggara barat

3. *Tokenizing* : dibagian ini sistem melakukan proses pemecahan teks menjadi unit yang lebih kecil, yang disebut token.

No	Label	<i>Tokenization</i>
1	-	'jokowi', 'aja', 'bisa', 'tambah', 'utang', 'banyak', 'banget', 'gimana', 'ditambah', 'lagi', 'yg', 'mirip', 'die', 'bisa', 'rakyat', 'mati', 'kelaparan'
2	+	'pemimpin', 'yang', 'merakyatnya', 'dari', 'hati', 'cocok', 'sih', 'kalau', 'mau', 'nyapres', 'di', 'kayaknya', 'bakal', 'banyak', 'yang', 'dukung', 'apalagi', 'kaum', 'perempuan', 'bantenggirl'
3	-	'waspada', 'anies', 'baswedan', 'didukung', 'kaum', 'intoleran', 'kaum', 'radikal', 'dan', 'para', 'pengasong'
4	+	'menteri', 'bumn', 'erick', 'thohir', 'menjadi', 'calon', 'wakil', 'presiden', 'cawapres', 'pilihan', 'masyarakat', 'di', 'pilpres', 'dan', 'dinilai', 'sebagai', 'sosok', 'yang', 'cerdas', 'dan', 'memiliki', 'kinerja', 'sangat', 'baik', 'juga', 'merupakan', 'sosok', 'yang', 'layak', 'menjadi', 'panutan', 'dan', 'suri', 'tauladan', 'bangsa', 'indonesia'
5	-	'anis', 'merupakan', 'sosok', 'yang', 'tidak', 'cocok', 'untuk', 'menjadi', 'pilpres', 'dikarenakan', 'tukang', 'tipu', 'alia', 'tenggelamkan', 'anies', 'nasdem', 'nusa', 'tenggara', 'barat'

4. *Stopwords* : dibagian ini sistem melakukan pengecekan kata-kata yang dianggap tidak penting dalam analisis teks karena sering muncul dan tidak memberikan informasi yang signifikan.

No	Label	<i>Stopwords</i>
1	-	'jokowi', 'tambah', 'utang', 'banyak', 'banget', 'gimana', 'ditambah', 'mirip', 'rakyat', 'mati', 'kelaparan'
2	+	'pemimpin', 'merakyatnya', 'hati', 'cocok', 'nyapres', 'banyak', 'dukung', 'kaum', 'perempuan', 'bantenggirl'
3	-	'waspada', 'anies', 'baswedan', 'didukung', 'kaum', 'intoleran', 'kaum', 'radikal', 'pengasong'
4	+	'menteri', 'bumn', 'erick', 'thohir', 'calon', 'wakil', 'presiden', 'cawapres', 'pilihan', 'masyarakat', 'pilpres', 'dinilai', 'sosok', 'cerdas', 'kinerja', 'sosok', 'layak', 'panutan', 'suri', 'tauladan', 'bangsa', 'indonesia'
5	-	'anis', 'sosok', 'cocok', 'pilpres', 'tukang', 'tipu', 'alia', 'tenggelamkan', 'anies', 'nasdem', 'nusa', 'tenggara', 'barat'

5. *Stemming* : dibagian ini sistem melakukan proses perubahan kata yang berbeda-beda menjadi kata dasar (*root word*) atau kata induk. Ini digunakan dalam analisis teks untuk mengurangi variasi dalam kata yang sama, seperti kata kerja yang berbeda-beda (mengambil, diambil, mengambalikan) atau kata benda yang berbeda-beda (mobil, mobilnya, mobilku) menjadi kata dasarnya (ambil, mobil) sehingga dapat digunakan untuk analisis yang lebih efektif.

No	Label	<i>Stemming</i>
1	-	'jokowi', 'tambah', 'utang', 'banyak', 'banget', 'gimana', 'tambah', 'mirip', 'rakyat', 'mati', 'lapar'
2	+	'pemimpin', 'rakyat', 'hati', 'cocok', 'nyapres', 'banyak', 'dukung', 'kaum', 'perempuan', 'bantenggirl'
3	-	'waspada', 'anies', 'baswedan', 'dukung', 'kaum', 'intoleran', 'kaum', 'radikal', 'pengasong'
4	+	'menteri', 'bumn', 'erick', 'thohir', 'calon', 'wakil', 'presiden', 'cawapres', 'pilihan', 'masyarakat', 'pilpres', 'nilai', 'sosok', 'cerdas', 'kinerja', 'sosok', 'layak', 'panutan', 'suri', 'tauladan', 'bangsa', 'indonesia'

5	-	'anis', 'sosok', 'cocok', 'pilpres', 'tukang', 'tipu', 'alia', 'tenggelam', 'anies', 'nasdem', 'nusa', 'tenggara', 'barat'
---	---	--

6. Pembobotan kata TF-IDF: dibagian ini sistem memproses untuk memberikan bobot atau skor pada kata-kata dalam teks untuk mengukur tingkat keterkaitan atau pentingnya kata-kata tersebut dalam konteks analisis teks.



Tabel 3.1 Nilai dari TF-IDF

Term	D1	D2	D3	D4	D5	IDF	TF-IDF				
							D1	D2	D3	D4	D5
jokowi	1	0	0	0	0	0.698	0.698	0	0	0	0
tambah	2	0	0	0	0	0.698	0	0	0	0	0
utang	1	0	0	0	0	0.698	0.698	0	0	0	0
banyak	1	1	0	0	0	0.397	0.698	0.698	0	0	0
banget	1	0	0	0	0	0.698	0.698	0	0	0	0
gimana	1	0	0	0	0	0.698	0.698	0	0	0	0
mirip	1	0	0	0	0	0.698	0.698	0	0	0	0
rakyat	1	1	0	0	0	0.397	0.397	0.397	0	0	0
mati	1	0	0	0	0	0.698	0.698	0	0	0	0
lapar	1	0	0	0	0	0.698	0.698	0	0	0	0
pemimpin	0	1	0	0	0	0.698	0	0.698	0	0	0
hati	0	1	0	0	0	0.698	0	0.698	0	0	0
cocok	0	1	0	0	1	0.397	0	0.397	0	0	0.397
nyapres	0	1	0	0	0	0.698	0	0.698	0	0	0
dukung	0	1	1	0	0	0.397	0	0.698	0.698	0	0

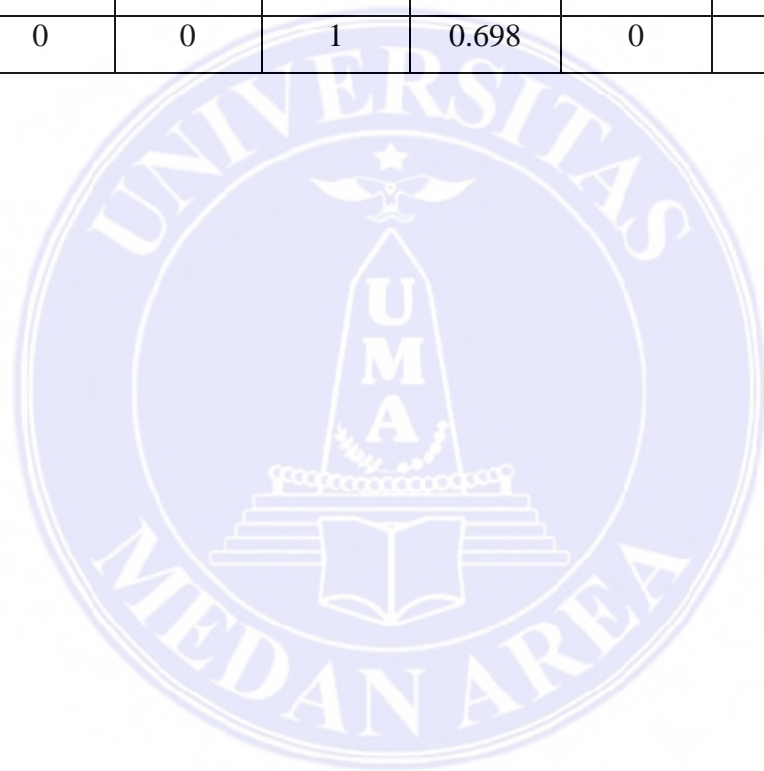
kaum	0	1	1	0	0	0.397	0	0.397	0.397	0	0
perempuan	0	1	0	0	0	0.698	0	0.698	0	0	0
bantenggirl	0	1	0	0	0	0.698	0	0.698	0	0	0
waspada	0	0	1	0	0	0.698	0	0	0.698	0	0
anies	0	0	1	0	1	0.397	0	0	0.397	0	0.397
baswedan	0	0	1	0	0	0.698	0	0	0.698	0	0
intoleran	0	0	1	0	0	0.698	0	0	0.698	0	0
radikal	0	0	1	0	0	0.698	0	0	0.698	0	0
pengasong	0	0	1	0	0	0.698	0	0	0.698	0	0
menteri	0	0	0	1	0	0.698	0	0	0	0.698	0
bumn	0	0	0	1	0	0.698	0	0	0	0.698	0
erick	0	0	0	1	0	0.698	0	0	0	0.698	0
thohir	0	0	0	1	0	0.698	0	0	0	0.698	0
calon	0	0	0	1	0	0.698	0	0	0	0.698	0
wakil	0	0	0	1	0	0.698	0	0	0	0.698	0
presiden	0	0	0	1	0	0.698	0	0	0	0.698	0
cawapres	0	0	0	1	0	0.698	0	0	0	0.698	0
pilihan	0	0	0	1	0	0.698	0	0	0	0.698	0

masyarakat	0	0	0	1	0	0.698	0	0	0	0.698	0
pilpres	0	0	0	1	1	0.397	0	0	0	0.397	0.397
nilai	0	0	0	1	0	0.698	0	0	0	0.698	0
cerdas	0	0	0	1	0	0.698	0	0	0	0.698	0
kinerja	0	0	0	1	0	0.698	0	0	0	0.698	0
sosok	0	0	0	1	1	0.397	0	0	0	0.397	0.397
layak	0	0	0	1	0	0.698	0	0	0	0.698	0
panutan	0	0	0	1	0	0.698	0	0	0	0.698	0
suri	0	0	0	1	0	0.698	0	0	0	0.698	0
tauladan	0	0	0	1	0	0.698	0	0	0	0.698	0
bangsa	0	0	0	1	0	0.698	0	0	0	0.698	0
indonesia	0	0	0	1	0	0.698	0	0	0	0.698	0
anis	0	0	0	0	1	0.698	0	0	0	0	0.698
tukang	0	0	0	0	1	0.698	0	0	0	0	0.698
tipu	0	0	0	0	1	0.698	0	0	0	0	0.698
alia	0	0	0	0	1	0.698	0	0	0	0	0.698
tenggelam	0	0	0	0	1	0.698	0	0	0	0	0.698
nasdem	0	0	0	0	1	0.698	0	0	0	0	0.698

nusa	0	0	0	0	1	0.698	0	0	0	0	0.698
tenggara	0	0	0	0	1	0.698	0	0	0	0	0.698
barat	0	0	0	0	1	0.698	0	0	0	0	0.698

Keterangan :

D=Data/Twitter



7. Data setelah *preprocessing* : dibagian ini sistem melakukan pembuatan data yang telah diolah atau dipersiapkan sebelum digunakan untuk analisis.

3.2.3 Contoh Perhitungan Algoritma *Naïve Bayes*

Ini adalah contoh perhitungan metode Naive Bayes yang digunakan pada data pemilihan Presiden Indonesia 2024:

Tabel 3.2 Contoh data pemilihan umum Presiden Indonesia 2024.

No	Label	<i>Tweet</i>
1	-	1 Jokowi aja bisa tambah utang banyak banget gimana ditambah 1 lagi yg mirip die? Bisa2 rakyat mati kelaparan
2	+	Pemimpin yang merakyatnya dari hati.. cocok sih kalau mau nyapres di 2024 kayaknya bakal banyak yang dukung apalagi kaum perempuan @BantengGirl
3	-	Waspada! Anies Baswedan didukung kaum intoleran, kaum radikal, dan para pengasong
4	+	Menteri BUMN @erickthohir menjadi calon wakil presiden (cawapres) pilihan masyarakat di Pilpres 2024. Dan dinilai sebagai sosok yang cerdas dan memiliki kinerja sangat baik juga merupakan sosok yang layak menjadi panutan dan suri tauladan bangsa Indonesia.
5	-	Anis merupakan sosok yang tidak cocok untuk menjadi Pilpres 2024 dikarenakan "Tukang tipu" alia #TenggelamkanAniesNasdem Nusa Tenggara Barat
6	?	Pilpres dapat berjalan damai

Keterangan :

- + = Data yang memiliki nilai positif
- - = Data yang memiliki nilai negative
- ? = Data yang belum diberi label/nilai.
- Data yang tercantum di atas memiliki jumlah data training sebanyak 5 tweet.

Data training yang berjumlah 5 dibagi menjadi dua kelompok sesuai dengan labelnya, yaitu positif dan negatif. Setelah dibagi, data tersebut dihitung menggunakan rumus metode *Naive Bayes*. Contoh perhitungan data negatif dan data positif dapat dilihat pada tabel 3.2 dan 3.4. Lebih lanjut tentang tabel 3.2 dan 3.4 dapat ditemukan pada tabel 3.5.

Dalam menyelesaikan perhitungan manual menggunakan algoritma *naive bayes* adalah sebagai berikut :

Terdapat rumus *naive bayes* yaitu contoh seperti perhitungan negatif $P(-).P(S|-)$. Setelah itu perhitungan manual pada tabel 3.1 untuk mengetahui banyak kalimat negatif dari tabel tersebut yaitu sebanyak 3 kalimat yang disimbolkan sebagai huruf P. setelah itu lakukan perhitungan manual yang ingin kita cari kata katanya, yaitu “Pilpres Dapat Berjalan Damai” setelah semua dihitung kemudian hitunglah dengan menggunakan rumus *naive bayes* tersebut dan begitu juga dengan mencari perhitungan manual kalimat positif. Setelah mendapatkan hasil dari $P(-).P(S|-)$ dan $P(+).P(S|+)$

$$P(-).P(S|-)=3.035 \times 10^{-9}$$

$$P(+).P(S|+)=1.572 \times 10^{-9}$$

Dibandingkan kedua hasil, yang lebih besar adalah yang berlabel positif. Oleh karena itu, kalimat dalam pengujian "pilpres dapat berjalan damai" diberi label positif.

Tabel 3.3 Contoh perhitungan data negatif

P(-)	(3/5)
Pilpres(-)	$(2 + 1)/(48 + 107) = 0.0193$
Dapat(-)	$(0 + 1)/(48 + 107) = 0.0064$
Berjalan(-)	$(0 + 1)/(48 + 107) = 0.0064$
Damai(-)	$(0 + 1)/(48 + 107) = 0.0064$
$P(-).P(S -)$	$(3/5) \times 0.0193 \times 0.0064 \times 0.0064 \times 0.0064$
	3.035×10^{-9}

Tabel 3.4 Contoh perhitungan data positif

P(+)	$(2/5)$
Pilpres(+)	$(2 + 1)/(57 + 107) = 0.0182$
Dapat(+)	$(0 + 1)/(57 + 107) = 0.0060$
Berjalan(+)	$(0 + 1)/(57 + 107) = 0.0060$
Damai(+)	$(0 + 1)/(57 + 107) = 0.0060$
P(+).P(S +)	$(2/5) \times 0.0182 \times 0.0060 \times 0.0060 \times 0.0060$
	1.572×10^{-9}

Tabel 3.5 Keterangan mengenai table 3.2 dan 3.3

Keterangan	
2	Jumlah kata "pilpres" dalam label negatif dapat ditemukan pada tabel.
1	Nilai bawaan dari rumus.
48	Jumlah dari semua kata dalam label negatif.
107	Jumlah dari semua kata dari label positif dan negatif.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Hasil prediksi menggunakan algoritma *Naïve Bayes* dengan pembagian data training 80% dan data testing 20%, mendapatkan hasil positif sebanyak 78 (3,55%), negatif sebanyak 88 (4%), dan netral sebanyak 2034 (92,45%).
2. Terdapat juga hasil dari *confusion matrix* berupa akurasi sebesar 93%, presisi sebesar 96%, *recall* sebesar 93% dan F1 Score sebesar 94%

5.2 Saran

Penelitian yang dilakukan ini sudah tentu tidak terlepas dari kekurangan dan kelemahan. Oleh karena itu, maka ada beberapa saran yang perlu dilakukan dalam penelitian berikutnya, diantaranya :

1. Penelitian selanjutnya diharapkan menggunakan dua atau lebih metode atau algoritma SVM dan lain lain sebagai pembandingan agar meningkatkan kekurangan sistem
2. Membuat *interface* berupa *desktop*, *web* maupun aplikasi *mobile* dari proses pengujian model dan visualisasi dari performa metode yang digunakan

DAFTAR PUSTAKA

- Aji Andika, L., & Amalia Nur Azizah, P. (2019). *Analisis Sentimen Masyarakat Terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 Pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier*.
- Dewi, O. :, & Larasati, P. (2023). "Meraih Keunggulan Bidang Ilmu Sosial Dan Politik Identitas Virtual Pada @Jjergaf Dalam Media Sosial Twitter. *Prosiding Konferensi Nasional Sosial Politik (Konaspol)*, 1.
- Dhimas Agusti, S., Yuslam Dahda, S., Hidayah, N., Topaz Tektonika, A., & Adi, S. (2019). *Analisis Sentiment Isu Pilpres 2019 Dengan Menggunakan Algoritma Naive Bayes*.
- Febriyani, E., & Februariyanti, H. (N.D.). *Analisis Sentimen Terhadap Program Kampus Merdeka Menggunakan Algoritma Naive Bayes Classifier Di Twitter*. 17(1).
- Gunawan, B., Sasty, H. P., & Esyudha, E. P. (2018). *Jepin (Jurnal Edukasi Dan Penelitian Informatika) Sistem Analisis Sentimen Pada Ulasan Produk Menggunakan Metode Naive Bayes*. 4(2), 17–29. www.femaledaily.com
- Hidayat Informatika, A. (2023). Klasifikasi Spam Email Menggunakan Metode Naive Bayes. In *Teknologipintar.Org* (Vol. 3, Issue 2).
- Hidayati, A. R., Fitriani, A. S., Rosid, M. A., Sains, F., & Teknologi, D. (2023). *Analisa Sentimen Pemilu 2019 Pada Judul Berita Online Menggunakan Metode Logistic Regression* (Vol. 4, Issue 2).
- Juanita, S. (2020). Analisis Sentimen Persepsi Masyarakat Terhadap Pemilu 2019 Pada Media Sosial Twitter Menggunakan Naive Bayes. *Jurnal Media Informatika Budidarma*, 4(3), 552. <https://doi.org/10.30865/mib.v4i3.2140>
- Khatib Sulaiman, J., Setiyawati, D., Cahyono, N., & Amikom Yogyakarta, U. (N.D.). Analisa Sentimen Pengguna Sosial Media Twitter Terhadap Perokok Di Indonesia. *Indonesian Journal Of Computer Science Attribution*, 12(1), 2023–2262.
- Mawarni, R., & Silegar, A. (2023). Pengolahan Data Mining Terhadap Penjualan Menggunakan Algoritma Apriori Pada Toko Alfamart Mulya Asri Tulang Bawang Barat. In *Jisn* (Vol. 04, Issue 01).
- Rizki Fakhri, M., Munandar, H., & Christian Chandra, J. (2023). *2 Nd Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (Senafiti) 21 Maret 2023-Jakarta*. 2(1).
- Rosaly, R., Prasetyo, A., & Kom, M. (2018). *Pengertian Flowchart Beserta Fungsi Dan Simbol-Simbol Flowchart Yang Paling Umum Digunakan*.

- Sains, H. F., & Teknologi, D. (2023). Pemodelan Topik Dalam Al-Qur'an Menggunakan Library Bertopic Pada Model Bahasa Bert. *Jurnal Simetris*, 14(2).
- Salim, E., & Syafrullah, M. (2023). *Jakarta Barat Menggunakan Algoritme K-Nearest Neighbor* (Vol. 20, Issue 1). https://kemsalim.space/ulasan_dukcapil/
- Sari, D. N., Adelia, F., Rosdiana, F., Butar, B. B., & Hariyanto, M. (2020). Analisa Sentimen Terhadap Review Produk Kecantikan Menggunakan Metode Naive Bayes Classifier. In *Jika: Vol. Issn*.
- Sentimen, A., Pilpres, J., Naïve, M., Classifier, B., Opini, B., Di, P., Prasetyo, T. H., & Fitriani, A. S. (2023). Seminar Nasional & Call Paper Fakultas Sains Dan Teknologi (Senasains 6 Th. In *Procedia Of Engineering And Life Science* (Vol. 4).
- Sibaroni, Y. (2023). *Perbandingan Pembobotan Fitur Tf-Idf Dan Tf-Abs Dalam Klasifikasi Berita Online Menggunakan Support Vector Machine (Svm) 2 Nd*.
- Silitonga, W. H., & Sihotang, J. I. (2019). *Analisis Sentimen Pemilihan Presiden Indonesia Tahun 2019 Di Twitter Berdasarkan Geolocation Menggunakan Metode Naïve Bayesian Classification*.
- Triono, A., Setia Budi, A., & Abdillah, R. (2023). Implementasi Peretasan Sandi Vigenere Chipper Menggunakan Bahasa Pemrograman Python. In *Jurnal Jocotis-Journal Science Informatica And Robotics E-Issn : Xxxx-Xxxx* (Vol. 1, Issue 1).
- Yolanda Paramitha, N., Nuryaman, A., Faisol, A., Setiawan, E., Dina Eka Nurvazly, Dan, Matematika, J., Mipa, F., Lampung Jl Soemantri Brojonegoro No, U., Lampung, B., Kunci, K., & Bayes, N. (2023). Klasifikasi Penyakit Stroke Menggunakan Metode Naïve Bayes. In *Jurnal Siger Matematika* (Vol. 04, Issue 01). <https://www.kaggle.com/datasets/zettrkalkpakbal/full-filled->

LAMPIRAN

1. Source code

1. Pengujian Data Pada Google Colaboratory

```
# Scraping data menggunakan API Twitter
import tweepy as tw
import pandas as pd

# Menggunakan token dari twitter
consumer_key= 'MN2zvRTQdqlcL0q03qYeEtNCF'
consumer_secret=
'vFMLcQiJCKOQ9NRnrMZXpnwEZlNvuv8TTx13o7gsUqvttbogH7'
access_token='1539625531591004163-
BZiYlAyjX89IzsmauFhmzJnrHdhSVs'
access_token_secret='cfdOKPeXQYsMIDfXhkoAL0HoTGyb1WcfU53grTeFe
LBg4'

auth = tw.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

api = tw.API(auth, wait_on_rate_limit=True)

# menentukan query
search_words = "pilpres 2024"
new_search = search_words + " -filter:retweets"

tweets = tw.Cursor(api.search_tweets,
                    q = new_search,
                    count = 11000,
                    include_entities = True).items(11000)

users_locs = [[tweet.created_at, tweet.id, tweet.text] for
tweet in tweets]

tweet_text = pd.DataFrame(data = users_locs,
                           columns = ["date", "id", "teks"])

tweet_text.to_csv('pilpres24.csv', index=False)
```



```
import pandas as pd

data = pd.read_csv('/content/pilpres24.csv')
data

# Cek jumlah baris dan kolom

data.shape

# Cek data terduplicat

data.duplicated().sum()

# Cek data missing value

data.isna().sum()

# Cek type data

data.info()

# Mengubah tipe data dari kolom date

data['date'] = pd.to_datetime(data['date'], format='%Y-%m-%d')

import nltk
nltk.download('vader_lexicon')
from nltk.sentiment.vader import SentimentIntensityAnalyzer

sentimen = SentimentIntensityAnalyzer()
polaritas = lambda x : sentimen.polarity_scores(x)
nilai_sentimen = data.teks.apply(polaritas)
nilai_sentimen.head()

sentiment_data = pd.DataFrame(data = list(nilai_sentimen))
sentiment_data.head()

# Menentukan Label positif, negatif, dan netral

label = lambda x : 'Netral' if x==0 else ('Positif' if x>0 else
'Negatif')
sentiment_data['Label'] = sentiment_data.compound.apply(label)
sentiment_data.head()

# Menggabungkan dataset dan label

dataset = data.join(sentiment_data.Label)
```

```
dataset.head()

# Menampilkan jumlah masing-masing kelas

dataset['Label'].value_counts()

import re

def cleaningulasan(teks):
    teks = re.sub(r'@[A-Za-a0-9]+', ' ', teks)
    teks = re.sub(r'#[A-Za-z0-9]+', ' ', teks)
    teks = re.sub(r"http\S+", ' ', teks)
    teks = re.sub(r'[0-9]+', ' ', teks)
    teks = re.sub(r"[-()\"#/@;:<>{}'+=~|.!?,_]", " ", teks)
    teks = teks.strip(' ')
    teks = teks.lower()
    return teks
dataset['case_folding'] = dataset['teks'].apply(cleaningulasan)
dataset.head()

import nltk
nltk.download('punkt')
from nltk.tokenize import word_tokenize

def tokenizingText(teks):
    teks = word_tokenize(teks)
    return teks
dataset['Tokenization'] =
dataset['case_folding'].apply(tokenizingText)
dataset.head()

from nltk.corpus import stopwords
nltk.download('stopwords')

daftar_stopword = stopwords.words('indonesian')
daftar_stopword.extend(["yg", "dg", "rt"])
daftar_stopword = set(daftar_stopword)

def stopwordText(words):
    return [word for word in words if word not in
daftar_stopword]

dataset['Stopword'] =
dataset['Tokenization'].apply(stopwordText)
dataset.head()

!pip install sastrawi
```

```

from Sastrawi.Stemmer.StemmerFactory import StemmerFactory

factory = StemmerFactory()
stemmer = factory.create_stemmer()

def porterstemmer(text):
    text = ' '.join(stemmer.stem(word) for word in text if word
in text)
    return text

dataset['Stemming'] = dataset['Stopword'].apply(porterstemmer)
dataset.head()

print(dataset.case_folding[0])
print(dataset.Tokenization[0])
print(dataset.Stopword[0])
print(dataset.Stemming[0])

from wordcloud import WordCloud, STOPWORDS
import matplotlib.pyplot as plt

Positif= " ".join(review for review in
dataset[dataset['Label'] == 'Positif'].Stemming)
Negatif= " ".join(review for review in
dataset[dataset['Label'] == 'Negatif'].Stemming)
Netral= " ".join(review for review in dataset[dataset['Label']
== 'Netral'].Stemming)
stopwords = set(STOPWORDS)

def plot_cloud(wordcloud):
    plt.figure(figsize=(12, 8))
    plt.imshow(wordcloud)
    plt.axis("off");

Positif = WordCloud(width = 3000, height = 2000,
random_state=3, background_color='black', colormap='Set2',
collocations=False, stopwords = STOPWORDS).generate(Positif)
plot_cloud(Positif)

Negatif = WordCloud(width = 3000, height = 2000,
random_state=3, background_color='black', colormap='Set2',
collocations=False, stopwords = STOPWORDS).generate(Negatif)
plot_cloud(Negatif)

Netral = WordCloud(width = 3000, height = 2000,
random_state=3, background_color='black', colormap='Set2',
collocations=False, stopwords = STOPWORDS).generate(Netral)

```

```
plot_cloud(Neutral)

stem = dataset['Stemming'].astype(str)

from sklearn.feature_extraction.text import TfidfVectorizer

vectorizer = TfidfVectorizer()
vector = vectorizer.fit_transform(stem).toarray()

print(vectorizer.vocabulary_)

print('teks', dataset.Stemming[0])
print('\n')
print(vector[0])

from sklearn.model_selection import train_test_split

# Tentukan y
y = dataset['Label']

# split training dan testing
X_train, X_test, y_train, y_test = train_test_split(vector, y,
test_size = 0.2, random_state=1)

y_test

from sklearn.naive_bayes import GaussianNB # library Naive
Bayes

nb = GaussianNB()
nb.fit(X_train, y_train)

y_pred = nb.predict(X_test)

# Menampilkan hasil akurasi

from sklearn.metrics import classification_report

print(classification_report(y_test, y_pred))

# Menampilkan Hasil Presisi dan Recall

from sklearn.metrics import precision_score, recall_score,
accuracy_score

acc = accuracy_score(y_test, y_pred)
```

```
ps = precision_score(y_test, y_pred, average='weighted')
rs = recall_score(y_test, y_pred, average='weighted')

print('Akurasi : ', acc)
print('Presisi : ', ps)
print('Recall : ', rs)

# Menampilkan hasil confusion matrix

from sklearn.metrics import confusion_matrix,
ConfusionMatrixDisplay

cm = confusion_matrix(y_test, y_pred, labels=nb.classes_)
disp = ConfusionMatrixDisplay(confusion_matrix=cm,
display_labels=nb.classes_)
disp.plot()

plt.show()

# Hasil klasifikasi menggunakan algoritma naive bayes

hasil_NB = pd.DataFrame({'label_NB':y_pred})
hasil_NB.value_counts()

import matplotlib.pyplot as plt

# Membuat jumlah visualisasi dari masing-masing klasifikasi

data = hasil_NB.label_NB.value_counts()
data.plot(kind='pie', autopct='%.2f%%')
plt.show()


hasil_NB.value_counts()

# label otomatis sebelum masuk ke tahapan machine learning

label_default = pd.DataFrame({'label_default':y_test})
label_default

label_default.value_counts()
```

2. Plagiasi



 **Similarity Report ID.** oid:29477:50816612

PAPER NAME	AUTHOR
188160019_Denny.docx	Denny Herdiansyah

WORD COUNT	CHARACTER COUNT
6393 Words	38648 Characters

PAGE COUNT	FILE SIZE
53 Pages	1.8MB

SUBMISSION DATE	REPORT DATE
Jan 30, 2024 10:59 PM GMT+7	Jan 30, 2024 11:00 PM GMT+7

- **25% Overall Similarity**
The combined total of all matches, including overlapping sources, for each database.
 - 20% Internet database
 - 8% Publications database
 - Crossref database
 - Crossref Posted Content database
 - 16% Submitted Works database
- **Excluded from Similarity Report**
 - Small Matches (Less than 10 words)

3. Surat Keterangan Pembimbing



UNIVERSITAS MEDAN AREA

FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax. (061) 7366998 Medan 20223
Kampus II : Jalan Seiabadi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225602, Fax. (061) 8226331 Medan 20122
Website www.teknik.uma.ac.id E-mail univ_medanarea@uma.ac.id

Nomor : 543/FT.6/01.10/VII/2023

18 Juli 2023

Lamp : -

H a l : Perpanjang SK Pembimbing Tugas Akhir

Yth. Pembimbing Tugas Akhir
Nurul Khairina, S.Kom, M.Kom
di
Tempat

Dengan hormat,
Sehubungan telah berakhirnya waktu masa berlaku SK pembimbing nomor 305/FT.6/01.10/XI/2022 tertanggal 9 Nopember 2022 maka perlu diterbitkan kembali SK Pembimbing Skripsi baru atas nama mahasiswa berikut :

N a m a : Denny Herdiansyah Chan
N P M : 188160019
Jurusan : Teknik Informatika

Oleh karena itu kami mengharapkan kesediaan saudara :

Nurul Khairina, S.Kom, M.Kom (Sebagai Pembimbing)

Adapun Tugas Akhir Skripsi berjudul :

"Analisis Sentimen Sosial Media Twitter terhadap Pilpres 2024 menggunakan Algoritma Naive Bayes"

SK Pembimbing ini berlaku selama enam bulan terhitung sejak SK ini diterbitkan. Jika proses pembimbing melebihi batas waktu yang telah ditetapkan, SK ini dapat ditinjau ulang.

Demikian kami sampaikan, atas kesediaan saudara diucapkan terima kasih.

Dekan.



Dr. Rahmad Syah, S. Kom, M. Kom

4. Surat Ijin Melakukan Riset dari Fakultas Teknik



UNIVERSITAS MEDAN AREA FAKULTAS TEKNIK

Kampus I : Jalan Kolam Nomor 1 Medan Estate/Jalan PBSI Nomor 1 ☎ (061) 7366878, 7360168, 7364348, 7366781, Fax (061) 7366998 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A, ☎ (061) 8225602, Fax (061) 8226331 Medan 20122
Website www.teknik.uma.ac.id E-mail univ_medanarea@uma.ac.id

Nomor : 305 /FT.6/01.10/IV/2023
Lamp : -
Hal : **Penelitian Dan Pengambilan Data Tugas Akhir**

11 April 2023

Yth. Wakil Rektor Bid. Pengembangan SDM & Adm. Keuangan
Jln. Kolam No.1
Di
Medan

Dengan hormat, kami mohon kesediaan ibu kiranya berkenan untuk memberikan izin dan kesempatan kepada mahasiswa kami tersebut dibawah ini :

NO	N A M A	N P M	PRODI
I	Denny Herdiansyah Chan	188160019	Teknik Informatika

Untuk melaksanakan Penelitian dan Pengambilan Data Tugas Akhir di **Laboratorium Komputer Program Studi Teknik Informatika Fakultas Teknik Universitas Medan Area.**

Perlu kami jelaskan bahwa Pengambilan Data tersebut adalah semata-mata untuk tujuan Ilmiah dan Skripsi, yang merupakan salah satu syarat bagi mahasiswa tersebut untuk mengikuti ujian sarjana pada Fakultas Teknik Universitas Medan Area dan tidak untuk dipublikasikan, dengan judul :

Analisis Sentimen Sosial Media Twitter terhadap Pilpres 2024 menggunakan Algoritma Naive Bayes.

Mohon kiranya tanggal Surat Izin Pengambilan Data Tugas Akhir agar disesuaikan dengan tanggal Terbitnya SK ini.

Atas perhatian dan kerja sama yang baik diucapkan terima kasih.


Dekan

Dr. Rahmadsyah, S. Kom, M. Kom

Tembusan :
1. Ka. BAMAI
2. Mahasiswa
3. File

5. Surat Riset Dari BATRI



UNIVERSITAS MEDAN AREA

Kampus I : Jalan Kolam Nomor 1 Medan Estate ☎ (061) 7360168, 7366878, 7364348 📠 (061) 7368012 Medan 20223
Kampus II : Jalan Setiabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A ☎ (061) 8225602 📠 (061) 8226331 Medan 20122
Website: www.uma.ac.id E-Mail: univ_medanarea@uma.ac.id

Nomor : 708 /UMA/B/01.7/V/2023
Lamp. : 1 (satu) lembar.
Hal : Penelitian dan Pengambilan Data Tugas Akhir

08 Mei 2023.

Kepada Yth.

Dekan Fakultas Teknik

Universitas Medan Area

di - M e d a n

Dengan hormat, sesuai dengan surat yang kami terima pada tanggal 06 Mei 2023 dari Dekan Fakultas Teknik Universitas Medan Area nomor 305/FT.6/01.10/IV/2023 tertanggal 11 April 2023, perihal Izin Penelitian dan Pengambilan Data Tugas Akhir di Laboratorium Komputer Program Studi Teknik Informatika Universitas Medan Area oleh mahasiswa sebagai berikut :

Nama : Denny Herdiansyah Chan
No. Pokok Mahasiswa : 188160019
Program Studi : Teknik Informatika
Fakultas : Teknik

Pada prinsipnya disetujui yang bersangkutan melaksanakan pengambilan data tersebut, untuk penunjang tugas akhir dengan judul Skripsi “Analisis Sentimen Sosial Media Twitter terhadap Pilpres 2024 menggunakan Algoritma Naive bayes”.

Wakil Rektor Bidang Pengembangan
SDM dan Administrasi Keuangan,

Dr. Ir. Suswati, MP

Tembusan :

1. Mahasiswa Ybs
2. File

6. Surat Selesai Riset



UNIVERSITAS MEDAN AREA

Kampus I : Jalan Kolam Nomor 1 Medan Estate ☎ (061) 7360168, 7366878, 7364348 ☎ (061) 7368012 Medan 20223
Kampus II : Jalan Setabudi Nomor 79 / Jalan Sei Serayu Nomor 70 A ☎ (061) 8225602 ☎ (061) 8226331 Medan 20122
Website: www.uma.ac.id E-Mail: univ_medanarea@uma.ac.id

SURAT KETERANGAN

Nomor : 1047/UMA/B/01.7/VI/2023

Rektor Universitas Medan Area dengan ini menerangkan bahwa :

Nama : Denny Herdiansyah Chan
No. Pokok Mahasiswa : 188160019
Program Studi : Teknik Informatika
Fakultas : Teknik

Benar telah selesai Pengambilan Data di Laboratorium Komputer Prodi Teknik Informatika Fakultas Teknik Universitas Medan Area dengan Judul “Analisis Sentimen Sosial Media Twitter Terhadap Pilpres 2024 Menggunakan Algoritma Naïve Bayes.”

Dan kami harapkan Data tersebut kiranya dapat membantu yang bersangkutan dalam penyusunan skripsi dan dapat bermanfaat bagi mahasiswa khususnya Fakultas Teknik.

Demikian surat ini diterbitkan untuk dapat digunakan seperlunya.

Medan, 14 Juni 2023.
An Rektor,
Wakil Rektor Bidang Pengembangan SDM dan
Administrasi Keuangan,


Suswati, MP

Tembusan :

1. Mahasiswa Ybs
2. File

