

**ANALISIS SENTIMEN TRENDING TOPIK DI MEDIA SOSIAL
TWITTER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER**

SKRIPSI
Program Studi Sistem Informasi
Jenjang Sarjana S1



Oleh

Ibnu Alana Muhaimin
NIM 09031381924138

FAKULTAS ILMU KOMPUTER
UNIVERSITAS SRIWIJAYA
JULI 2023

LEMBAR PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN TRENDING TOPIK DI MEDIA SOSIAL
TWITTER MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER**

Sebagai salah satu syarat untuk penyelesaian
studi di Program Studi Sistem Informasi

Oleh

Ibnu Alana Muhaimin
09031381924138

Palembang, 20 Juli 2023

Pembimbing,



Fathoni, MMSI.
NIP 197210182008121001

Mengetahui
Ketua Jurusan Sistem Informasi,



Endang Lestari Ruskan, M.T.
NIP 197811172006042001

HALAMAN PERNYATAAN

Yang Bertanda Tangan Di Bawah Ini

Nama : Ibnu Alana Muhaimin
Nim : 09031381924138
Program Studi : Sistem Informasi Bilingual
Judul Skripsi : Analisis Sentimen Trending Topik Di
Media Sosial Twitter Menggunakan
Metode Naïve Bayes Classifier

Hasil Pengecekan, Ithenticate/Turnitin : 4%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam lapoan skripsi, maka saya bersedia menerima sanksi akademis dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun



Palembang, 20 Juli 2023


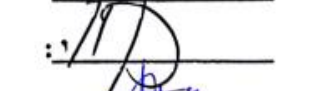


Ibnu Alana Muhaimin
NIM 09031381924138

HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Rabu
Tanggal : 12 Juli 2023
Nama : Ibnu Alana Muhaimin
NIM : 0903138124138
Program Studi : Sistem Informasi Bilingual
Judul Skripsi : Analisis Sentimen Trending Topik Di Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier

Tim Penguji :

1. Pembimbing	: Fathoni, MMSI.	:	
2. Ketua Penguji	: Dr. Ali Ibrahim, M.T.	:	
3. Penguji 1	: Ken Ditha Tania, M.Kom.	:	
4. Penguji 2	: Allsela Meiriza, M.T.	:	

Mengetahui,

Ketua Jurusan Sistem Informasi,



Endang Lestari Ruskan, M.T

NIP. 197811172006042001

HALAMAN PERSEMBAHAN

MOTTO :

**"ORANG YANG MAMPU BELAJAR DARI KESALAHAN
ADALAH ORANG YANG BERANI UNTUK SUKSES."**

Skripsi ini dipersembahkan kepada :

- ❖ Allah SWT
- ❖ Ayah dan Mama
- ❖ Adik adik
- ❖ Dosen Pembimbing dan Penguji
- ❖ Sahabat dan Teman Seperjuangan
- ❖ Almamaterku, Universitas Sriwijaya

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillah!, Puji syukur penulis haturkan kepada Allah SWT karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Analisis Sentimen Trending Topik Di Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier” dapat diselesaikan dengan baik.

Selama pengerjaan Tugas Akhir ini penulis memperoleh bimbingan, bantuan, dukungan, dan juga doa dari berbagai pihak sehingga tugas akhir dapat diselesaikan dengan baik. Oleh karena itu dalam kesempatan ini penulis ingin mengucapkan terima kasih kepada :

1. Sang Maha Pemberi Pertolongan Allah SWT, atas segala kasih sayang, hidayah ilmu, kekuatan, dan petunjuk sehingga penulis dapat merampungkan tugas akhir ini yang mana banyak sekali hikmah dalam setiap proses yang dilalui.
2. Ayah dan Mama selaku kedua orang tua, serta adik-adik penulis yang selama ini selalu mendukung di setiap kegiatan positif penulis, selalu mengupayakan yang terbaik, dan dukungan serta doa yang tak terhenti.
3. Bapak Jaidan Jauhari, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Ibu Endang Lestari Ruskan, M.T. selaku Ketua Jurusan Sistem Informasi.
5. Bapak Fathoni, MMSI. Selaku Dosen Pembimbing Skripsi penulis yang selalu sabar dalam membimbing penulis serta telah memberikan banyak bimbingan serta bantuan selama penyusunan tugas akhir ini hingga dapat diselesaikan dengan baik.

6. Seluruh Dosen Fakultas Ilmu Komputer Universitas Sriwijaya yang telah berjasa mengajarkan ilmunya, dan tiada bosan memberikan motivasi, dan nasihatnya untuk penulis selama perkuliahan.
7. Seluruh teman Jurusan Sistem Informasi angkatan 2019 yang tidak dapat penulis sebutkan satu persatu.

Penulis berharap semoga tugas akhir ini bermanfaat bagi kita semua terkhusus untuk mahasiswa Fakultas Ilmu Komputer Universitas Sriwijaya. Penulis juga menyadari bahwa tugas ini masih jauh dari kesempurnaan. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang bersifat membangun agar dapat lebih baik lagi di kemudian hari

Palembang, 20 Juli 2023

Penulis,



Ibnu Alana Muhaimin

NIM. 09031381294138

ANALISIS SENTIMEN TRENDING TOPIK DI MEDIA SOSIAL TWITTER MENGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Oleh

Ibnu Alana Muhaamin
09031381924138

ABSTRAK

Twitter merupakan media sosial yang di dalamnya memiliki informasi seperti berita terkini, biografi seseorang, dan tweet dari penggunaanya. Twitter memiliki suatu fitur yang dinamakan trending topik yang berfungsi untuk mengetahui informasi topik tertentu yang sedang populer. Namun, sering kali sulit untuk memahami sentimen apa yang ada pada trending topik tersebut. Oleh karena itu, penting untuk mengklasifikasikan sentimen pada trending topik yang bertujuan mengetahui tentang bagaimana masyarakat merespon dan memandang topik-topik yang sedang populer di Twitter. Metode yang digunakan untuk analisis sentimen pada penelitian ini menggunakan naïve bayes classifier dan dilakukan validasi data menggunakan k-fold cross validation dengan uji nilai fold = 2,3,4,5,6,7,8,9 dan 10 untuk mendapat model akurasi terbaik. Dari pengujian yang telah dilakukan menggunakan aplikasi Rapidminer, diperoleh hasil akurasi terbaik terdapat pada model fold = 7 dengan nilai akurasi sebesar 65,55%, nilai precision 59,71%, dan nilai recall 36,53%.

Kata Kunci : Analisis Sentimen, Twitter, Trending Topik, Naïve Bayes Classifier

SENTIMENT ANALYSIS OF TRENDING TOPICS ON THE TWITTER SOCIAL MEDIA USING THE NAIVE BAYES CLASSIFIER METHOD.

By

Ibnu Alana Muhaimin

09031381924138

ABSTRACT

Twitter is a social media platform that contains information such as current news, personal biographies, and tweets from its users. Twitter has a feature called trending topics, which functions to identify specific popular topics. However, it is often difficult to understand the sentiment associated with these trending topics. Therefore, it is important to classify the sentiment of trending topics with the aim of understanding how people respond to and perceive popular topics on Twitter. The method used for sentiment analysis in this study is the naive Bayes classifier, and data validation is performed using k-fold cross-validation with fold values of 2,3,4,5,6,7,8,9 and 10 to obtain the best accuracy model. The testing conducted using the RapidMiner application revealed that the best accuracy was achieved with the model at fold = 7, with an accuracy value of 65.55%, precision value of 59.71%, and recall value of 36.53%.

Keywords: *Sentiment Analysis, Twitter, Trending Topics, Naive Bayes Classifier.*

DAFTAR ISI

LEMBAR PENGESAHAN	ii
HALAMAN PERNYATAAN.....	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	6
BAB II TINJAUAN PUSTAKA	7
2.1 Landasan Teori	7
2.1.1 Analisis	7
2.1.2 Analisis Sentimen	7
2.1.3 Media Sosial	8
2.1.4 <i>Twitter</i>	9
2.1.5 <i>Trending Topic Twitter</i>	10
2.1.6 <i>Data Mining</i>	10
2.1.7 Fungsi Data Mining	11
2.1.8 Pengolahan Data	11
2.1.9 Sastrawi.....	13
2.1.10 <i>TextBlob</i>	14
2.1.11 <i>NLTK (Natural Language Tool Kit)</i>	14
2.1.12 <i>K- Fold Cross Validation</i>	14
2.1.13 <i>Naïve Bayes</i>	15
2.1.14 <i>Confusion Matrix</i>	17
2.1.15 Frekuensi Kata dan <i>Wordcloud</i>	19
2.1.16 Rapidminer	20
2.1.17 Penelitian Terdahulu	21
BAB III METODOLOGI PENELITIAN	24

3.1 Desain Penelitian	24
3.2 Implementasi Metode Naïve Bayes.....	28
3.3 Alat dan Bahan Penelitian	29
BAB IV PENGOLAHAN DATA	31
4.1 Proses Pengumpulan Data	31
4.2 Proses Preprocessing Data.....	33
4.2.1 <i>Case Folding</i> Data	33
4.2.2 Tokenizing	34
4.2.3 <i>Stopword Removal</i>	35
4.2.4 <i>Stemming</i> Data	36
4.3 Pelabelan	37
4.4 Implementasi <i>Algoritma Naïve Bayes</i>	37
4.5 Wordcloud	38
BAB V HASIL DAN PEMBAHASAN.....	40
5.1 Hasil Pengumpulan Data	40
5.2 Hasil Preprocessing	41
5.2.1 Hasil <i>Case Folding</i> Data	41
5.2.2 Hasil <i>Tokenize</i>	42
5.2.3 Hasil <i>Stopword</i>	42
5.2.4 Hasil Proses <i>Stemming</i>	43
5.2.5 Hasil <i>Preprocessing</i>	44
5.3 Pelabelan	45
5.4 Hasil Pembagian Data	45
5.5 Hasil Implementasi Metode <i>Naïve Bayes</i>	46
5.5.1 Hasil Pengujian <i>2-Fold Cross Validation</i>	47
5.5.2 Hasil Pengujian <i>3-Fold Cross Validation</i>	48
5.5.3 Hasil Pengujian <i>4-Fold Cross Validation</i>	50
5.5.4 Hasil Pengujian <i>5-Fold Cross Validation</i>	51
5.5.5 Hasil Pengujian <i>6-Fold Cross Validation</i>	52
5.5.6 Hasil Pengujian <i>7-Fold Cross Validation</i>	53
5.5.7 Hasil Pengujian <i>8-Fold Cross Validation</i>	54
5.5.8 Hasil Pengujian <i>9-Fold Cross Validation</i>	56
5.5.9 Hasil Pengujian <i>10-Fold Cross Validation</i>	57
5.6 Hasil Penelitian.....	58
5.7 Hasil Analisis <i>Wordcloud</i>	60
5.7.1 <i>Wordcloud</i> Sentimen Positif.....	60

5.7.2 <i>Wordcloud</i> Sentimen Negatif	62
BAB VI PENUTUP	64
6.1 Kesimpulan.....	64
6.2 Saran.....	65
DAFTAR PUSTAKA	66

DAFTAR GAMBAR

	Halaman
Gambar 2. 1 Proses Kerja K-Fold Cross Validation	15
Gambar 2. 2 Contoh Confusion Matrix 2x2.....	18
Gambar 2. 3 Frekuensi Kata	19
Gambar 2. 4 Wordcloud	19
Gambar 3. 1 Desain Penelitian.....	24
Gambar 4. 1 Operator Pengumpulan Data	31
Gambar 4. 2 Operator Didalam Subprocess.....	32
Gambar 4. 3 Proses Pemangilan Data	33
Gambar 4. 4 Data Sebelum Case Folding	34
Gambar 4. 5 Proses Case Folding	34
Gambar 4. 6 Proses Tokenize	35
Gambar 4. 7 Proses Stopword Removal	36
Gambar 4. 8 Proses Stemming	36
Gambar 4. 9 Proses Labelisasi	37
Gambar 4. 10 Implementasi Metode Naive Bayes.....	38
Gambar 4. 11 K-Fold Cross Validation	38
Gambar 4. 12 Operator Wordcloud.....	38
Gambar 4. 13 Process Document From Data.....	39
Gambar 4. 14 Operator Filters Token (By Length).....	39
Gambar 4. 15 Operator Filter Example Range.....	39
Gambar 5.2 Hasil Crawling Data.....	39
Gambar 5. 1 Hasil Pengumpulan Data.....	40
Gambar 5. 2 Data Sebelum Case Folding	41
Gambar 5. 3 Data Setelah Proses Case Folding	41
Gambar 5. 4 Data Sebelum Tokenize	42
Gambar 5. 5 Data Setelah Proses Tokenize	42
Gambar 5. 6 Data Sebelum Proses Stopword	43
Gambar 5. 7 Data Setelah Proses Stopword.....	43
Gambar 5. 8 Data Setelah Proses Stemming.....	44
Gambar 5. 9 Data Setelah Preprocessing	44
Gambar 5. 10 Data Setelah Preprocessing Diexcel.....	44
Gambar 5. 11 Data Setelah Proses Labelisasi Diexcel	45
Gambar 5. 12 Data Setelah Proses Labelisasi	45
Gambar 5. 13 Pembagian Data.....	46
Gambar 5. 14 Grafik Pembagian Data 7-Fold Cross Validation	59
Gambar 5. 15 Grafik Frekuensi Wordcloud Sentimen Positif	60
Gambar 5. 16 Wordcloud Sentimen Positif	61
Gambar 5. 17 10 Kata Teratas Sentimen Positif	61
Gambar 5. 18 Grafik Frekuensi Wordcloud Sentimen Negatif.....	62
Gambar 5. 19 Wordcloud Sentimen Negatif.....	63
Gambar 5. 20 5 Kata Teratas Sentimen Negative	63

DAFTAR TABEL

	Halaman
Tabel 2. 1 Case Folding	12
Tabel 2. 2 Tokenizing	12
Tabel 2. 3 Stopword Removal.....	13
Tabel 2. 4 Stemming	13
Tabel 2. 5 Penelitian Terdahulu	21
Tabel 3. 1 Format Pengujian K-Fold Cross Validation.....	28
Tabel 4.1 Pembagian Data	45
Tabel 5. 1 Jumlah Data.....	40
Tabel 5. 2 Sampel Hasil Pelabelan.....	45
Tabel 5. 3 Pembagian Data	46
Tabel 5. 4 Hasil Pengujian Metode Naive Bayes.....	47
Tabel 5. 5 Hasil Pengujian 2-Fold Cross Validation	47
Tabel 5. 6 Hasil Pengujian 3-Fold Cross Validation	49
Tabel 5. 7 Hasil Pengujian 4-Fold Cross Validation	50
Tabel 5. 8 Hasil Pengujian 5-Fold Cross Validation	51
Tabel 5. 9 Hasil Pengujian 6-Fold Cross Validation	52
Tabel 5. 10 Hasil Pengujian 7-Fold Cross Validation.....	53
Tabel 5. 11 Hasil Pengujian 8-Fold Cross Validation.....	55
Tabel 5. 12 Hasil Pengujian 9-Fold Cross Validation	56
Tabel 5. 13 Hasil Pengujian 10-Fold Cross Validation.....	57
Tabel 5. 14 Hasil 7-Fold Cross Validation.....	59
Tabel 5. 15 Tabel Frekuensi Wordcloud Sentimen Positif	60
Tabel 5. 16 Tabel Frekuensi Wordcloud Sentimen Negatif.....	62

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data set Awal.....	A-1
Lampiran 2 Hasil Preprocessing Text.....	B-1
Lampiran 3 Surat Pernyataan Similarity.....	C-1
Lampiran 4 Form Perbaikan Ujian Komprehensif.....	D-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan internet yang sangat pesat di Indonesia berdasarkan data dari Asosiasi Penyelenggara Internet Indonesia (APJII) menyatakan jumlah penduduk Indonesia yang telah terkoneksi dengan internet pada kurun 2021-2022 mencapai 210 juta orang dengan tingkat penetrasi internet sudah mencapai 77.02% dimana dengan tingkat penggunaan internet yang sudah mencapai 220 juta dan penetrasi internet mencapai 77,02% mempermudah masyarakat dalam melakukan komunikasi melalui media sosial (Dimas, 2022)

Menurut (Witro et al., 2020) media sosial adalah media yang berbasis internet yang memungkinkan pengguna untuk berinteraksi dan menunjukkan diri mereka, baik secara langsung atau tertunda, dengan audiens yang luas atau sempit, yang meningkatkan nilai dari konten yang dihasilkan pengguna dan persepsi interaksi dengan orang lain. Salah satu media sosial terbesar yaitu *twitter*. Pengguna *twitter* di Indonesia berdasarkan laporan dari *We Are Sosial dan Hootsuite* , memiliki 24 juta pengguna di Indonesia dari total 556 juta pengguna global (Annur, 2023).

Menurut (Khaira et al., 2020) *twitter* adalah layanan komunikasi daring yang menyediakan fasilitas untuk menuliskan teks yang singkat sehingga pengguna dapat menyampaikan opininya secara jelas, ringkas dan padat serta dapat dipublikasikan kepada semua orang atau kelompok tertentu. Kelebihan dari *twitter* diantaranya adalah komunikasi terjadi dengan sangat cepat, publikasinya luas dan membantu dalam penyebaran informasi dengan lebih cepat yang kemudian akan menjadi topik yang dibahas oleh para penggunanya

Menurut(Fikri et al., 2020) *tweets* yang merupakan hasil dari penyaluran opini dan komentar, merupakan *resource* yang dapat digunakan untuk menganalisis sentimen khalayak umum terhadap suatu instansi maupun perorangan. Hal tersebut disebabkan karena *tweets* tersebut mengandung sentimen yang dapat dijadikan sebagai tolak ukur pandangan khalayak umum yang dapat dijadikan sebagai bahan evaluasi ke depannya. Dalam menentukan sentimen suatu tweet, dapat dilakukan dengan mengklasifikasikannya ke dalam tiga kelas, yaitu positif, negatif, dan netral. *Tweet (twit)* atau kicauan pada setiap pengguna Twitter dapat berpengaruh dalam pembentukan citra suatu produk, program, atau kebijakan karena semakin banyak suatu isu tertentu dibahas dalam twit pengguna maka topik tersebut dapat menjadi *trending* topik.

Twitter merupakan media sosial yang didalamnya memiliki informasi seperti berita terkini, biografi seseorang, dan *tweet* dari penggunanya. Twitter memiliki suatu fitur yang dinamakan *trending* topik yang berfungsi untuk mengetahui informasi topik tertentu yang sedang populer. Kenyataannya seringkali sulit untuk memahami *trending* topik apa yang sedang terjadi. Oleh karena itu, diperlukan pengklasifikasian *trending* topik kedalam suatu kategori umum(Ramadhy & Sibaroni, 2022).

Analisis sentimen adalah sebuah bidang dalam *Natural Language Processing* (NLP) yang fokus pada identifikasi perasaan atau emosi yang terkandung dalam teks. Ini juga dikenal sebagai "*opinion mining*", yaitu proses otomatis memahami, mengekstrak, dan mengolah data tekstual untuk menentukan sentimen yang terkandung di dalamnya. Tujuan dari analisis sentimen adalah untuk mengidentifikasi pola dalam teks, kalimat, dan *twett*, dengan mendeteksi polaritas

dalam dokumen tersebut (Khaira et al., 2020) pada penelitian ini, analisis sentimen digunakan untuk mengelompokkan sentiment Positif dan negatif, analisis sentiment terhadap *trending* dapat memberikan wawasan secara *real-time* dan relevan tentang pandangan dan perasaan masyarakat terkait topik tersebut dan analisis sentiment pada *trending* topik dapat membantu memahami preferensi dan opini pengguna, mendeteksi tren, mengukur kepuasan pelanggan, serta mengidentifikasi masalah atau peluang yang muncul dalam waktu nyata.

Naïve Bayes adalah metode pengklasifikasian yang menggunakan probabilitas dan statistik yang ditemukan oleh Thomas Bayes, seorang ilmuwan Inggris. Ini memprediksi peluang di masa depan berdasarkan pengalaman masa lalu dan dikenal sebagai Teorema Bayes. Dalam hal ini, algoritma menghitung probabilitas bahwa kelas keputusan adalah benar dengan menggunakan informasi dari obyek yang bersangkutan, dengan asumsi bahwa atribut obyek independen. Probabilitas yang terlibat dalam menghasilkan perkiraan akhir dihitung berdasarkan frekuensi "master" tabel keputusan. *Naive Bayes Classifier* berperforma lebih baik daripada model klasifikasi lain (Pardede et al., 2020).

Menurut (Pebdika et al., 2023) keunggulan metode *naive bayes* adalah kemampuannya untuk menghasilkan hasil yang baik meskipun hanya menggunakan sedikit data latih. Tujuannya adalah memastikan parameter yang diharapkan pada langkah klasifikasi. Beberapa kelebihan dari metode ini antara lain: 1. Mengatasi masalah kuantitatif dan diskrit secara efektif, 2. Mampu mengatasi titik kebisingan terisolasi, seperti saat menghitung probabilitas bersyarat dari data, 3. Hanya memerlukan jumlah data latih yang kecil untuk memperkirakan parameter yang diperlukan dalam proses klasifikasi, seperti rata-rata dan variansi

dari variabel, 4.Mampu mengatasi nilai yang hilang karena mengabaikan kejadian selama estimasi kemungkinan peluang, 5.Prosesnya cepat dan menghemat waktu, 6. Tahan terhadap karakteristik yang tidak relevan. Penelitian tentang *trending* topik yang berjudul analisis *Trending* Topik Twitter dengan Fitur Ekspansi FastText Menggunakan Metode *Logistic Regression*, Pada penelitian ini, penerapan hyperparameter tuning pada model klasifikasi Logistic Regression terbukti dapat meningkatkan nilai akurasi sebesar 76,25% (+0,28) dan F1-Score sebesar 76,17% (+0,32) dengan skenario splitting data 90:10. Kemudian untuk implementasi fitur ekspansi FastText juga terbukti dapat meningkatkan nilai akurasi pada sistem, hasil terbaik didapatkan dengan top 1 feature menggunakan klasifikasi Logistic Regression dengan nilai akurasi sebesar 76,39% (Ramadhy & Sibaroni, 2022)

Pada penelitian yang berjudul Analisis Sentimen Twitter menggunakan Metode *Naive Bayes* dengan Seleksi Fitur Frekuensi Relevansi (Studi Kasus: Opini Masyarakat tentang Kebijakan New Normal). Data yang digunakan berasal dari platform media sosial Twitter. Hasil dari metode *Naive Bayes* menunjukkan tingkat akurasi tertinggi sebesar 73,3%, dengan presisi 75,8%, *recall* 73,3%, dan *f-measure* 73,4% saat diuji pada fold-4. Sementara itu, akurasi terendah ditemukan pada fold-5, dengan tingkat akurasi sebesar 55%, presisi 54%, *recall* 55%, dan *f-measure* 54,2%. (Verena et al., 2021).

Pada penelitian yang berjudul Analisis Sentimen terhadap *Cryptocurrency* dengan menggunakan *Python TextBlob* dan Algoritma *Naïve Bayes*. Analisis sentimen bertujuan untuk mengetahui bagaimana sentimen masyarakat global terhadap *hashtag* populer di media sosial *Twitter*, yaitu *#crypto*. Tujuan dari analisis ini adalah untuk memahami sejauh mana sentimen terkait *hashtag* tersebut bersifat

positif, netral, atau negatif. Setelah melakukan pelabelan otomatis menggunakan *Python TextBlob* pada total 1032 *tweet*, ditemukan bahwa sentimen positif mencapai 61.24%, sentimen netral sebesar 26.68%, dan sentimen negatif sebesar 10.07%. Hasil pengujian menggunakan algoritma *Naïve Bayes* dengan data testing dan data *training* masing-masing sebesar 0.2 dan 0.8 menunjukkan nilai akurasi sebesar 71.98%, presisi 83.04%, *recall* 60.88%, dan *f1_score* 71.98%. Dengan angka-angka tersebut, dapat disimpulkan bahwa analisis sentimen terhadap *#crypto* di media sosial *Twitter* menghasilkan hasil yang cukup baik dalam mengukur sentimen masyarakat terkait topik tersebut."

Dari latar belakang diatas, maka akan dibahas lebih lanjut dalam tugas akhir dengan judul "**Analisis Sentimen Trending Topik Di Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier**"

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang dapat dirumuskan masalah yang akan dibahas sebagai berikut :

1. Bagaimana penerapan metode *naïve bayes classifier* pada analisis sentiment *trending* topik yang ada pada media sosial *twitter*
2. Berapa tingkat keakuratan metode *naïve bayes classifier* dalam melakukan klasifikasi sentiment

1.3 Tujuan Penelitian

Berikut merupakan tujuan dari penelitian ini antara lain sebagai berikut:

1. Menilai kemampuan metode *naïve bayes classifier* dalam menganalisis sentiment yang ada pada *trending* topik di media sosial *twitter*

2. Mengetahui tingkat akurasi dalam metode *naïve bayes classifier* dalam analisis sentimen yang ada pada *trending* topik di media sosial *twitter*
3. Mengetahui tentang bagaimana masyarakat merespons dan memandang topik - topik yang sedang populer di *twitter*.

1.4 Manfaat Penelitian

1. hasil penelitian ini diharapkan dapat menjadi alat bantu memahami opini, pandangan masyarakat terkait dengan hal - hal yang sedang trend di media sosial *twitter* Dengan demikian, informasi ini dapat menjadi dasar untuk pengambilan keputusan yang lebih baik dalam berbagai konteks, seperti pemasaran, kebijakan publik, pengembangan produk.
2. hasil penelitian ini diharapkan dapat membantu dalam mengukur kepuasan, kepercayaan masyarakat terhadap publik figur atau suatu kebijakan

1.5 Batasan Masalah

Adapun batasan masalah yang mencangkup ruang lingkup dalam penelitian ini sebagai berikut:

1. Data penelitian yang digunakan berupa data *tweet* masyarakat yang ada pada *twitter*
2. Metode yang digunakan dalam analisis ini menggunakan *metode naïve bayes*
3. Tools yang digunakan untuk mengolah data adalah *rapidminer 10.1.003*
4. Data yang digunakan dari media sosial *twitter*
5. *Trending* topik *twitter* dalam penelitian ini adalah “Aldi Taher”, ”Ganjar”, ”Anies”, ”Pdip”, ”Jokowi” ,”Manchester United”,”Prabowo”,”Messi“,“dudung abdurachman”

DAFTAR PUSTAKA

- Alfikri, M. Z. (2020). Analisis Sentimen Twitter terhadap Kartu Prakerja di tengah Pandemi COVID-19 menggunakan Algoritma Pencocokan String dan library TextBlob. 1–6. <https://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2019-2020/Makalah/Makalah-Stima-2020-031.pdf>
- Anisa Putri, A. M. (2022). Analisis Sentimen Cyberbullying Kpop Di Media Sosial Twitter Menggunakan Metode Naive Bayes Anisa. *Jurnal Ilmiah Indonesia*, 7(8.5.2017), 2003–2005.
- Annur, cindy M. (n.d.). Pengguna Twitter di Indonesia Capai 24 Juta hingga Awal 2023, Peringkat Berapa di Dunia? Retrieved April 4, 2023, from <https://databoks.katadata.co.id/datapublish/2023/02/27/pengguna-twitter-di-indonesia-capai-24-juta-hingga-awal-2023-peringkat-berapa-di-dunia>
- Ashari, M. Y., Fauzan, A. C., & Muhamat, H. M. (2022). Perbandingan Kinerja Sistem Klasifikasi Berbasis K-Fold Cross Validation pada Algoritma Decision Tree ID3 dan C5.0. *Jurnal ICT : 1*, 21(1), 44 - 52. https://ejournal.ikmi.ac.id/index.php/jict_ikmi/article/view/450/235
- Cahyaningrum, N. I., Yoshida Fatima, D. W., Kusuma, W. A., Ramadhani, S. A., Destanto, M. R., & Nooraeni, R. (2020). Analysis of User Sentiment of Twitter to Draft KUHP. *Jurnal Matematika, Statistika Dan Komputasi*, 16(3), 273. <https://doi.org/10.20956/jmsk.v16i3.8239>
- Dimas, B. (2022, June). APJII: Pengguna Internet Indonesia Tembus 210 Juta pada 2022. <https://dataindonesia.id/digital/detail/apjii-pengguna-internet-indonesia-tembus-210-juta-pada-2022>
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *Smatika Jurnal*, 10(02), 71–76. <https://doi.org/10.32664/smatika.v10i02.455>
- Juditha, C. (2017). Sentimen Dan Imparsialitas Isi Berita Tentang Ahok Di Portal Berita Online. *Jurnal Penelitian Komunikasi Dan Pembangunan*, 18(1), 57. <https://doi.org/10.31346/jpkp.v18i1.839>
- Khaira, U., Johanda, R., Eko, P., Utomo, P., Suratno, T., & Info, A. (2020). Sentiment Analysis of Cyberbullying on Twitter Using SentiStrength. 3(1), 21–27.
- M Rizki, M Arhami, H. (2021). Perbaikan Algoritma Naive Bayes Classifier Menggunakan Teknik Laplacian Correction. *Jurnal Teknologi*, 21(1), 39–45.
- Mahmud, R., & Hartanto, A. (2020). Penerapan Data Mining Rekomendasi Laptop Menggunakan Algoritma Apriori. *Juisi*, 06(02), 21–30.
- Maulana., D., & Yahya, R. (2019). Implementasi Algoritma Naïve Bayes Untuk Klasifikasi Penderita Penyakit Jantung Di Indonesia Menggunakan Rapid Miner. 84(10), 1511–1518. <https://doi.org/10.1134/s0320972519100129>
- Muhdar, A. (2021). Analisis mengenai pengaruh politik terhadap kehidupan di kampung yeflio distrik mayamuk kabupaten sorong. 9, 67–72.
- Naraswati, N. P. G., Nooraeni, R., Rosmilda, D. C., Desinta, D., Khairi, F., & Damaiyanti, R. (2021). Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification. *Sistemasi*, 10(1), 222. <https://doi.org/10.32520/stmsi.v10i1.1179>

- Novianti, D. (2019). Implementasi Algoritma Naïve Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner. *Paradigma - Jurnal Komputer Dan Informatika*, 21(1), 49–54. <https://doi.org/10.31294/p.v21i1.4979>
- Nurul, S., Fitriyyah, J., Safriadi, N., Esyudha, E., & #3, P. (2019). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes. (*Jurnal Edukasi Dan Penelitian Informatika*), 5(3), 279–285. <http://dev.twitter.com>.
- Parasati, W., Abdurrachman Bachtiar, F., & Setiawan, N. Y. (2020). Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(4), 1090–1099. <http://j-ptiik.ub.ac.id>
- Pardede, J., Miftahuddin, Y., & Kahar, W. (2020). Deteksi Komentar Cyberbullying Pada Media Sosial Berbahasa Inggris Menggunakan Naïve Bayes Classification. 7(1), 46–54.
- Pebdika, A., Herdiana, R., & Solihudin, D. (2023). Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 452–458. <https://doi.org/10.36040/jati.v7i1.6303>
- Purnamasari, N. M. G. D., Fauzi, M. A., Indriarti, & Dewi, L. S. (2018). Identifikasi Tweet Cyberbullying pada Aplikasi Twitter menggunakan Metode Support Vector Machine (SVM) dan Information Gain (IG) sebagai Seleksi Fitur. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(11), 5326–5332.
- Rachman, F. F., & Pramana, S. (2020). Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. *Health Information Management Journal*, 8(2), 100–109. <https://inohim.esaunggul.ac.id/index.php/INO/article/view/223/175>
- Rahman ali, A. (2017, May). Pengenalan Natural Language Toolkit (NLTK). 3 Mei 2017. <https://code.tutsplus.com/id/tutorials/introducing-the-natural-language-toolkit-nltk--cms-28620>
- Ramadhy, I. F., & Sibaroni, Y. (2022). Analisis Trending Topik Twitter dengan Fitur Ekspansi FastText Menggunakan Metode Logistic Regression. *JURIKOM (Jurnal Riset Komputer)*, 9(1), 1. <https://doi.org/10.30865/jurikom.v9i1.3791>
- Rengkung, A. L., & Pitoy, C. (2022). Analisis Kesalahan Siswa Dalam Menyelesaikan Soal Pada Materi Operasi Bentuk Aljabar. 2(2), 273–281.
- Rezeki, S. R. I. (2020). Penggunaan sosial media twitter dalam komunikasi organisasi (studi kasus pemerintah provinsi dki jakarta dalam penanganan covid-19). *Journal of Islamic and Law Studies*, 04(02), 63–78.
- Ridwansyah, T. (2022). Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naïve Bayes Classifier. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 2(5), 178–185. <https://doi.org/10.30865/klik.v2i5.362>
- Rifano, E. J., Fauzan, A. C., Makhi, A., Nadya, E., Nasikin, Z., & Putra, F. N. (2020). Text Summarization Menggunakan Library Natural Language Toolkit

- (NLTK) Berbasis Pemrograman Python. *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 2(1), 8–17. <https://doi.org/10.28926/ilkomnika.v2i1.32>
- Rufaidha, N. F., & Irhandayaningsih, A. (2022). Perilaku Informasi Mahasiswa Fakultas Ilmu Budaya Universitas Diponegoro dalam Pemanfaatan Fitur Trending Topic Twitter Sebagai Pemenuhan Kebutuhan Informasi. *Anuva: Jurnal Kajian Budaya, Perpustakaan, Dan Informasi*, 6(4), 493–504. <https://doi.org/10.14710/anuva.6.4.493-504>
- Sabransyah, M., Nasution, Y. N., & Tisna, D. (2017). Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung Naive Bayes Method for a Heart Risk Disease Prediction Application. *Jurnal EKSPONENSIAL*, 8, 111–118.
- Setiawan, G., Palit, H. N., & Setyati, E. (2019). Aspect Based Sentiment Analysis pada Layanan Umpan Balik Universitas dengan Menggunakan Metode Naive Bayes dan Latent Semantic Analysis. *Jurnal Infra*, 7(1), 170–174.
- Siregar, H. (2022). Analisis Pemanfaatan Media Sosial Sebagai Sarana Sosialisasi Pancasila. *Pancasila: Jurnal Keindonesiaan*, 1, 71–82. <https://doi.org/10.52738/pjk.v2i1.102>
- Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner. *JBASE - Journal of Business and Audit Information Systems*, 4(1), 13–21. <https://doi.org/10.30813/jbase.v4i1.2729>
- Verena, K., Toy, S., Sari, Y. A., & Cholissodin, I. (2021). Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus : Opini Masyarakat mengenai Kebijakan New Normal). 5(11), 5068–5074.
- Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naive Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25. <https://doi.org/10.36787/jti.v13i1.78>
- Widowati, T. T., & Sadikin, M. (2021). Analisis Sentimen Twitter terhadap Tokoh Publik dengan Algoritma Naive Bayes dan Support Vector Machine. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 11(2), 626–636. <https://doi.org/10.24176/simet.v11i2.4568>
- Witro, D., Putri, L. A., & Oviensy, V. (2020). Kontribusi Media Sosial Terhadap Produktivitas Karyawan Generasi Milenial Pt Perkebunan Nusantara Vi Kayu Aro. *Ekonomi & Bisnis*, 18(2), 119–125. <https://doi.org/10.32722/eb.v18i2.2492>