

**ANALISIS SENTIMEN TREN 'KABUR AJA DULU'
PADA MEDIA SOSIAL X SEBAGAI DASAR
PERANCANGAN SISTEM PEMANTAUAN
SENTIMEN PUBLIK MENGGUNAKAN NAÏVE
BAYES DAN SVM**

SKRIPSI

Diajukan sebagai salah satu syarat kelulusan program Strata – 1

(S-1)

Program Studi : Sistem Informasi

Tahun Akademik 2025/2026



**STIKOM
CIPTA KARYA INFORMATIKA**

Disusun Oleh :

Nama : Fachrur Rozi

Nim : 21120110008

**SEKOLA TINGGI ILMU KOMPUTER CIPTA KARYA
INFORMATIKA (STIKOM CKI) JAKARTA**

2025



STIKOM
CIPTA KARYA INFORMATIKA

PENGESAHAN UJIAN

Skripsi ini diujikan pada tanggal 16 bulan 07 tahun 2025, dan dinyatakan

: ~~LULUS~~/TIDAK LULUS

Nama : Fachrur Rozi

Nim : 21120110008

Program Studi/Program : Sistem Infromasi/S-1

Judul SKRIPSI : Analisis Sentimen Tren 'Kabur Aja Dulu' Pada Media Sosial X Sebagai Dasar Perancangan Sistem Pemantauan Sentimen Publik Menggunakan Naive Bayes Dan Svm.

NAMA PENGUJI :

TANDA TANGAN

1. Veri Arinal, S.kom., M.kom

NIDN : 0328047101

2. RASIBAN, S.kom., M.kom

NIDN : 3080007006

Mengetahui,

Ketua Sidang SKRIPSI

(Dr. Mesra Betty Yel., MM., DBA., M.Kom)

NIDN : 0416057202



STIKOM
CIPTA KARYA INFORMATIKA

SEKOLAH TINGGI ILMU KOMPUTER CIPTA KARYA INFORMATIKA

SURAT KETERANGAN KEASLIAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, SKRIPSI ini merupakan karya saya sendiri (ASLI), dan isi dalam SKRIPSI ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain *karya tulis ini juga sudah di cek plagiarisme dengan aplikasi turnitin kurang dari 30% dan kami sertakan hasil pengecekan turnitin nya dalam lampiran*, dan menerangkan kembali bahwa saya :

NIM : 21120110008
Nama : Fachrur Rozi
Judul SKRIPSI : **Analisis Sentimen Tren 'Kabur Aja Dulu' Pada Media Sosial X Sebagai Dasar Perancangan Sistem Pemantauan Sentimen Publik Menggunakan Naive Bayes Dan Svm.**

Adalah benar karya Tulis saya dan segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi

Jakarta,



NIM : 21120110008

ABSTRAK

Media sosial X (Twitter) telah menjadi platform utama bagi masyarakat Indonesia dalam menyampaikan opini, termasuk terhadap tren kabur aja dulu. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat menggunakan metode Naïve Bayes dan Support Vector Machine (SVM), serta membandingkan akurasi keduanya dalam analisis sentimen. Data dikumpulkan melalui Twitter API dengan hashtag #kaburajadulu dan menghasilkan 2.067 tweet, yang setelah proses cleansing dan pelabelan manual, tersisa 385 data. Proses analisis mengikuti tahapan CRISP-DM yang mencakup pemahaman bisnis, data, persiapan data, pemodelan, evaluasi, dan deployment. Evaluasi model dilakukan menggunakan confusion matrix dengan metrik akurasi, presisi, dan recall. Hasil klasifikasi menunjukkan bahwa 82% tweet bersentimen positif dan 18% negatif. Algoritma Naïve Bayes menghasilkan akurasi sebesar 86,49%, sedikit lebih rendah dibandingkan SVM yang mencapai 88,05%. Kesimpulannya, Support Vector Machine lebih efektif dalam klasifikasi sentimen pada data opini publik. Penelitian ini memberikan kontribusi dalam pemetaan opini publik secara digital dan merekomendasikan pengembangan metode pelabelan otomatis serta eksplorasi algoritma lanjutan di masa mendatang.

Kata Kunci: Analisis Sentimen, Kabur Aja Dulu, Media Sosial X, Naive Bayes, SVM, Pemantauan Opini Publik.

ABSTRACT

Social media X (Twitter) has become the main platform for the Indonesian public to express opinions, including on the trend of 'kabur aja dulu' (let's just run away for a bit). This research aims to classify the sentiments of the public using the Naïve Bayes and Support Vector Machine (SVM) methods, and to compare the accuracy of both in sentiment analysis. Data was collected via the Twitter API with the hashtag #kaburajadulu, resulting in 2,067 tweets, which, after the cleansing process and manual labeling, left 385 data points. The analysis process followed the CRISP-DM stages, which include business understanding, data understanding, data preparation, modeling, evaluation, and deployment. Model evaluation was conducted using a confusion matrix with accuracy, precision, and recall metrics. The classification results show that 82% of tweets have a positive sentiment and 18% negative. The Naïve Bayes algorithm achieved an accuracy of 86.49%, slightly lower than SVM, which reached 88.05%. In conclusion, Support Vector Machine is more effective in sentiment classification on public opinion data. This research contributes to the digital mapping of public opinion and recommends the development of automatic labeling methods as well as the exploration of advanced algorithms in the future.

Keywords: *Sentiment Analysis, Kabur Aja Dulu, Social Media X, Naive Bayes, SVM, Public Opinion Monitoring.*

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh, Alhamdulillah Puji dan syukur saya panjatkan kehadiran Allah SWT, karena berkat rahmat dan kasih- Nya, saya sebagai penulis dapat menyelesaikan laporan skripsi dengan judul **“ANALISIS SENTIMEN TREN 'KABUR AJA DULU' PADA MEDIA SOSIAL X SEBAGAI DASAR PERANCANGAN SISTEM PEMANTAUAN SENTIMEN PUBLIK MENGGUNAKAN NAIVE BAYES DAN SVM”** guna untuk memenuhi sebagian persyaratan untuk memperoleh gelar Sarjana SI program studi Sistem Informasi di STIKOM CKI (Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika). Saya menyadari kelemahan serta keterbatasan yang ada sehingga dalam menyelesaikan Laporan Skripsi ini, banyak bantuan dari berbagai pihak, dalam kesempatan ini kami menyampaikan ucapan terimakasih kepada :

1. Bapak Muhammad Farel Ardhan, S.T. DBA Selaku Ketua Yayasan Cipta Karya Intelektual.
2. Ibu Dr. Mesra Betty Yel, MM., DBA., M.Kom selaku Ketua Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.
3. Bapak Yuma Akbar, M.Kom selaku Wakil Ketua I Bidang Akademik dan Kemahasiswaan Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.
4. Bapak M. Zaeni, S.E, M.Pd selaku wakil ketua II Bidang Keuangan dan SDM Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.
5. Bapak Veri Arinal, M.Kom selaku Ketua Program Studi Sistem Informasi Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.
6. Bapak Dr. Suhendi, S.T.,S.Kom., M.Msi selaku Ketua LPPM.
7. Bapak Sutisna, S.Kom., MM selaku Dosen Pembimbing yang telah banyak memberikan masukan, semangat, serta dorongan selama penyusunan skripsi.
8. Bapak Veri Arinal, S.Kom., M.Kom selaku Dosen Penguji 1
9. Bapak Rasiban, S.Kom., M.Kom selaku Dosen Penguji 2
10. Management dan Seluruh Dosen Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.

11. Orang Tua yang selalu mendukung dan menyemangati dalam pelaksanaan penelitian ini.
12. Untuk teman-teman Mahasiswa/i seperjuangan Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.
13. Terimakasih juga kepada semua pihak yang telah membantu dalam penyelesaian Laporan Skripsi.

Akhir kata Penulis mengucapkan Terima Kasih pada semua pihak yang telah membantu dalam penulisan Laporan Skripsi. penulis menyadari bahwa Laporan proposal skripsi ini masih banyak kekurangan baik isi maupun susunannya. Semoga apa yang tertulis dalam laporan ini dapat bermanfaat tidak hanya bagi penulis juga bagi para pembaca.

Jakarta 28 Juni 2025



Fachrur Rozi

DAFTAR ISI

PENGESAHAN UJIAN	i
SURAT KETERANGAN KEASLIAN.....	ii
ABSTRAK	iii
ABSTRACT	iv
KATA PENGANTAR.....	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah/RP.....	2
1.3 Rumusan Masalah/RQ.....	2
1.4 Tujuan Penelitian/RO.....	3
1.5 Pendekatan masalah.....	3
1.6 State Of The Art (SOTA)	4
1.7 Kontribusi Penelitian / Novelty	5
BAB II.....	6
TINJAUAN PUSTAKA.....	6
2.1 Systematic Literature Review (SLR).....	6
2.1.1 Survei Metodologi.....	7
2.1.2 Survei Review Protocol.....	8
2.1.3 Kajian Pustaka Utama.....	11
2.1.4 Matrik Perbandingan RP, RQ, RO.....	30
2.2 Definisi dan Pengertian.....	31
2.2.1 Text Mining.....	31
2.2.2 Data Mining	31
2.2.3 Analisa Sentimen	32
2.2.4 Twitter (X)	32
2.2.5 Application Programming Interface (API).....	32
2.2.6 Google Colab	33

2.2.7 Python	33
2.2.8 Klasifikasi	33
2.2.9 Naïve Bayes	34
2.2.10 Support Vector Machine (SVM).....	34
2.2.12 Dataset.....	35
2.2.13 Rapid Miner	35
2.2.14 Cross-Industry Process For Data Mining (CRISP-DM)	35
2.2.15 Flowchart	39
BAB III.....	42
METODOLOGI PENELITIAN	42
3.1 Data Penelitian	42
3.2 Penerapan Metodologi	43
3.2.1 Proses Pengumpulan Data.....	45
3.3 Perancangan Penelitian	46
3.3.1 Pemahaman Bisnis (<i>Business Understanding</i>).....	46
3.3.2 Pemahaman Data.....	47
3.3.3 Data Preparation.....	47
3.3.4 Modeling	49
3.3.5 Evaluation & Results	49
BAB IV	50
HASIL DAN PEMBAHASAN	50
4.1 Alat Penelitian.....	50
4.1.1 Perangkat Lunak (<i>Software</i>).....	50
4.1.2 Perangkat Keras (<i>Hardware</i>)	50
4.2 Implementasi dan Pengujian.....	51
4.2.1 Pemahaman Bisnis (<i>Business Understanding</i>).....	51
4.2.2 Pemahaman Data (<i>Data Understanding</i>)	51
4.2.3 Data Preparation.....	51
4.2.4 Modeling	57
4.2.5 Evaluation & Result.....	58
BAB V.....	63
PENUTUP.....	63
5.1 Kesimpulan	63
5.2 Saran	63

DAFTAR PUSTAKA.....	64
LAMPIRAN	68

DAFTAR TABEL

Tabel 2. 1 Review PICOC	7
Tabel 2. 2 Review Survei Protocol.....	8
Tabel 2. 3 Kajian Pustaka Utama	11
Tabel 2. 4 Matrix Perbandingan RP, RQ, RO.....	30
Tabel 2. 5 Simbol Flowchart	39
Tabel 3. 1 Deskripsi Nama Atribut	42
Tabel 4. 1 Spesifikasi Software	50
Tabel 4. 2 Spesifikasi Hardware.....	50

DAFTAR GAMBAR

Gambar 2. 1 Studies Selected Strategy.....	10
Gambar 2. 2 Fase CRISP-DM.....	36
Gambar 2. 3 Algoritma Permodelan.....	38
Gambar 3. 1 Tahapan Penerapan Metodologi	44
Gambar 3. 2 Flowchart Pengumpulan Data	45
Gambar 4. 1 Instalasi Package Python	52
Gambar 4. 2 Pengambilan Token Dari X.....	52
Gambar 4. 3 Memasukan Token	52
Gambar 4. 4 Proses Crawling Data	53
Gambar 4. 5 Proses Export Data	53
Gambar 4. 6 Jumlah Data yang Didapat.....	54
Gambar 4. 7 Model Process Cleansing.....	54
Gambar 4. 8 Hasil Cleansing Data	55
Gambar 4. 9 Data yang Sudah Dilabeli	55
Gambar 4. 10 Model Tokenizing Data	56
Gambar 4. 11 Model Tokenizing Naïve Bayes	56
Gambar 4. 12 Model Tokenizing Support Vector Machine	57
Gambar 4. 13 Model Proses Data Naïve Bayes.....	57
Gambar 4. 14 Model Proses Data Support Vector Machine	58
Gambar 4. 15 Hasil Akurasi Naïve bayes	58
Gambar 4. 16 Deskripsi Akurasi Naïve bayes.....	59
Gambar 4. 17 Plot View Naïve bayes.....	59
Gambar 4. 18 Grafik AUC (Optimistic) Naïve Bayes.....	60
Gambar 4. 19 Hasil Akurasi Support Vector Machine.....	61
Gambar 4. 20 Deskripsi Akurasi Support Vector Machine.....	61
Gambar 4. 21 Plot View Support Vector Machine.....	61
Gambar 4. 22 Grafik AUC (Optimistic) Support Vector Machine	62

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang pesat telah membawa perubahan signifikan dalam cara manusia berkomunikasi dan menyampaikan pendapat. Saat ini, penggunaan media sosial telah menjadi salah satu kegiatan sehari-hari yang tidak dapat dihindari. Media sosial dapat digunakan sebagai cara untuk berkomunikasi, berbagi informasi, dan mengemukakan pendapat. Penggunaan media sosial dapat memiliki efek positif maupun negatif pada kehidupan seseorang.[1]. Salah satu dampak dari perkembangan tersebut adalah meningkatnya penggunaan media sosial sebagai sarana utama masyarakat dalam mengekspresikan opini, perasaan, dan tanggapan terhadap berbagai peristiwa. Media sosial X (sebelumnya dikenal sebagai Twitter) merupakan salah satu platform yang paling banyak digunakan oleh masyarakat untuk menyampaikan pendapat secara terbuka, cepat, dan real-time.

Di bulan Februari 2025 sampai Juli 2025, muncul sebuah tren populer di media sosial X dengan tren "kabur aja dulu". Tren ini digunakan oleh banyak pengguna untuk mengekspresikan sikap atau respon terhadap situasi yang dianggap tidak menguntungkan, penuh tekanan, atau membingungkan. Tren ini mencerminkan fenomena sosial yang menarik, karena menunjukkan cara masyarakat dalam merespon dinamika sosial, politik, dan budaya yang sedang terjadi. Namun demikian, makna dari penggunaan tren ini sangat bervariasi, tergantung pada konteks dan nada emosional dari penggunaannya, sehingga diperlukan analisis yang lebih mendalam untuk memahami sentimen yang terkandung di dalamnya.

Analisis sentimen merupakan proses untuk mengidentifikasi dan mengklasifikasikan opini atau perasaan pengguna terhadap suatu topik ke dalam kategori sentimen positif, negatif, atau netral. Analisis Sentimen juga bisa diartikan teknik data tekstual yang menggunakan *Natural Language Processing* (NLP) dan *Machine Learning* (ML) untuk mengevaluasi teks

secara otomatis untuk perasaan yang dirasakan oleh penulis (positif, negatif, dan netral)[2]. Melalui analisis sentimen, dapat diperoleh gambaran mengenai persepsi publik terhadap suatu isu tertentu. Hal ini sangat bermanfaat, terutama bagi pemangku kebijakan, lembaga sosial, maupun pelaku usaha dalam merumuskan strategi dan mengambil keputusan berdasarkan data yang bersumber dari opini publik.

Untuk melakukan analisis sentimen terhadap data dalam jumlah besar dan tidak terstruktur seperti unggahan media sosial, diperlukan pendekatan yang sistematis dan berbasis teknologi.. Naive Bayes dikenal karena kesederhanaan dan efisiensinya dalam menangani data besar, sedangkan SVM memiliki keunggulan dalam menghasilkan klasifikasi yang akurat, terutama pada data yang kompleks dan tidak linear. Kombinasi algoritma machine learning seperti Naive Bayes (NB) dan Support Vector Machine (SVM) adalah pendekatan populer, terutama dikarenakan kemampuan mereka dalam mengklasifikasi teks informal, slang, dan emotikon[3].

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk melakukan analisis sentimen terhadap tren "kabur aja dulu" di media sosial X serta merancang sistem pemantauan sentimen publik menggunakan algoritma Naive Bayes dan SVM. Sistem yang dikembangkan diharapkan dapat digunakan sebagai alat bantu dalam memahami opini publik secara otomatis dan menyeluruh, serta sebagai dasar dalam pengambilan keputusan strategis di berbagai bidang.

1.2 Identifikasi Masalah/RP

Adapun identifikasi masalah berdasarkan latar belakang antara lain :

- a. Adanya pro & kontra masyarakat terkait tren "kabur aja dulu".
- b. Sentimen masyarakat tren "kabur aja dulu".

1.3 Rumusan Masalah/RQ

Sehubungan identifikasi masalah diatas maka peneliti merumuskan masalah sebagai berikut :

1. Bagaimana cara menerapkan text mining dengan metode *Naïve Bayes* dan *Support Vector Machine* untuk melakukan analisa sentimen masyarakat terhadap tren "kabur aja dulu"?
2. Berapakah perbandingan tingkat akurasi metode *Naïve Bayes* dan *Support Vector Machine* dalam analisa sentimen terhadap tren "kabur aja dulu"?

1.4 Tujuan Penelitian/RO

Adapun tujuan dari dilakukanya penelitian ini adalah sebagai berikut :

1. Menerapkan text mining menggunakan metode *Naïve Bayes* dan *Support Vector Machine* untuk mengetahui presentase sentimen positif dan negatif tentang tren "kabur aja dulu".
2. Agar publik mengetahui sentimen tentang tren "kabur aja dulu".

1.5 Pendekatan masalah

Pendekatan yang digunakan dalam penelitian ini terdiri atas beberapa tahap:

a. Pengumpulan Data

Data dikumpulkan dari media sosial X dengan menggunakan Teknik *web scraping* berdasarkan kata kunci yang relevan, seperti “Kabur aja dulu”, dan sejenisnya dalam rentan waktu tertentu.

b. Pra-pemrosesan Data

Data yang diperoleh dari media sosial umumnya tidak terstruktur, sehingga perlu dilakukan proses pembersihan (*cleaning*), seperti menghapus tanda baca, URL, emotikon, *stopword*, serta melakukan *tokenizing* dan *stemming* untuk menyiapkan data agar siap untuk dianalisis.

c. Labeling Sentimen

Setiap data teks label sentiment (positif, negatif) baik secara manual maupun dengan bantuan kamus sentimen (sentiment lexicon) sebagai tahap pelatihan model.

d. Implementasi Algoritma

1. *Naïve Bayes*: Model probabilistik berbasis teori Bayes, mengasumsikan fitur saling independen.

2. *Support Vector Machine* (SVM): Model mencari hyperplane atau bidang pemisah terbaik yang dapat memisahkan data ke dalam dua kelas atau lebih.

e. Evaluasi Model

Model dievaluasi menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score untuk mengetahui seberapa baik kinerja Naïve Bayes dan *Support Vector Machine* (SVM) dalam mengklasifikasikan sentiment masyarakat.

f. Implementasi Hasil

Hasil analisis sentimen diinterpretasikan untuk mengetahui kecenderungan opini masyarakat terhadap tren "kabur aja dulu" yang kemudian disajikan dalam bentuk visualisasi data dan analisis naratif.

1.6 State Of The Art (SOTA)

Analisis sentimen merupakan sub-bidang dari *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi opini atau perasaan (positif, negatif, atau netral) dari suatu teks. Media sosial seperti Twitter (sekarang X) menjadi sumber utama karena sifatnya yang real-time, publik, dan kaya opini.

Salah satu algoritma yang sering digunakan dalam analisis sentimen adalah Naïve Bayes yang bekerja berdasarkan prinsip probabilistic dan asumsi independensi antar fitur, sedangkan SVM bekerja dengan prinsip margin *maksimum* dan penggunaan kernel. SVM mampu menghasilkan model yang kuat dan generalisasi yang baik terhadap data baru.

Namun, belum banyak penelitian yang secara spesifik menganalisis tren "kabur aja dulu" menggunakan pendekatan supervised learning seperti Naïve Bayes dan SVM, yang membuat penelitian ini relevan, baru, dan penting sebagai kontribusi akademik maupun sosial. Secara umum, algoritma Naïve Bayes dan SVM tetap menjadi pendekatan yang valid dalam analisis sentimen media sosial.

Oleh karena itu, penelitian ini memposisikan dirinya untuk mengisi kekosongan riset terhadap opini masyarakat tentang tren "kabur aja dulu".

1.7 Kontribusi Penelitian / Novelty

Diharapkan dari hasil penelitian ini akan dapat memberikan kontribusi kepada pihak – pihak terkait sebagai berikut :

a. Penulis

Mampu menerapkan metode *Naïve Bayes* dan *Support Vector Machine* dalam pengelompokan sentimen positif dan negatif dari respon masyarakat terkait tren "kabur aja dulu". Penulis yakin dengan melakukan penelitian ini maka penulis memiliki wawasan yang baru tentang penerapan text mining, dan dapat untuk mengetahui sentimen publik tentang tren "kabur aja dulu".

b. Pembaca

Penulis berharap penelitian ini dapat memberikan kontribusi bagi para pembacanya khususnya mahasiswa Jurusan Teknik Informatika dan Sistem Informasi sehingga hasil penelitian ini dapat digunakan sebagai referensi dan perbandingan bagi para mahasiswa yang akan atau sedang melakukan penelitian yang sama dengan penulis.

BAB II

TINJAUAN PUSTAKA

2.1 Systematic Literature Review (SLR)

Systematic Literature Review (SLR) atau tinjauan pustaka sistematis adalah metode penelitian yang sistematis untuk mengumpulkan, mengevaluasi, dan menyintesis seluruh literatur yang relevan tentang suatu topik pada penelitian tertentu. Dalam SLR, peneliti mengikuti langkah-langkah yang terstruktur untuk melakukan pencarian karya tulis atau sumber informasi yang ada sebagai referensi, mengevaluasi kualitas metodologi studi yang ada, dan menyusun temuan-temuan tersebut secara rinci. Tujuan utama dari SLR itu sendiri ialah menyajikan pemahaman menyeluruh tentang topik penelitian, mengidentifikasi kesenjangan pengetahuan, dan menyediakan dasar yang kuat untuk pengambilan keputusan atau pengembangan penelitian selanjutnya. Tahapan pada SLR yang biasanya diikuti antara lain sebagai berikut :

1. *Planing*

Research Question (RQ) adalah bagian awal dan dasar dari SLR. RQ digunakan untuk menuntun proses pencarian dan ekstraksi literatur. Analisis dan sintesis data yang dihasilkan oleh SLR adalah jawaban dari RQ yang telah kita tentukan sebelumnya. RQ yang baik adalah yang bermanfaat dan dapat diukur yang mengarah pada pemahaman topik penelitian tentang penelitian kontemporer.

2. *Conduiting*

Tahapan *conduiting* merupakan tahapan yang dimulai dengan penentuan keyword pencarian literatur (*search string*) yang basisnya adalah dari PICOC (*Population, Intervention, Comparison, Outcome, & Context*) yang telah kita desain didepan. Pemahaman terhadap sinonim dan alternatif pengganti kata akan menentukan akurasi pencarian literatur penulis. Kemudian langkah berikutnya adalah penentuan sumber (*digital library*) dari pencarian literatur.

3. *Reporting*

Reporting adalah tahapan penulisan hasil dari SLR dalam bentuk tulisan, baik untuk dipublikasikan dalam bentuk paper ke jurnal ilmiah atau untuk menyusun Bab 2 tentang Literature Review dari skripsi/tesis/disertasi.

2.1.1 Survei Metodologi

Survei Metodologi adalah penelitian pada populasi besar maupun kecil; namun, data yang dipelajari berasal dari sampel populasi untuk menemukan hubungan antara variabel sosiologis dan psikologis.. Dalam Perumusan RQ tentang efektivitas dalam penelitian harus fokus pada 5 elemen yang dikenal sebagai *PICOC* :

- a. *Popularitas (P)* : kelompok sasaran untuk penyelidikan (misalnya orang, Perangkat lunak, dll)
- b. *Intervention (I)* : Menentukan aspek penyelidikan atau masalah yang menarik bagi peneliti.
- c. *Comparison (C)* : Aspek dari investigasi dimana *Intervention (I)* akan dibandingkan.
- d. *Outcomes (O)* : Efek dan hasil dari *Intervention*.
- e. *Context (C)* : Latar atau lingkungan penyelidikan.

Untuk *Review PICOC* digambarkan pada tabel 2.1 berikut :

Tabel 2. 1 Review PICOC

Judul Tema Survei : ANALISIS SENTIMEN TREN 'KABUR AJA DULU' PADA MEDIA SOSIAL X SEBAGAI DASAR PERANCANGAN SISTEM PEMANTAUAN SENTIMEN PUBLIK MENGGUNAKAN NAIVE BAYES DAN SVM	
Population	Analisa sentimen menggunakan metode <i>Naïve bayes dan Support Vector Machine (SVM)</i>

Intervention	Bagaimana cara peneliti dapat menentukan sentimen positif
	maupun negatif dari masyarakat menggunakan sosial media X
Comparison	Membandingkan sentimen positif dan negatif
Outcomes	Menerapkan konsep data mining menggunakan <i>Naïve Bayes</i> dan <i>Support Vector Machine</i> (KNN) untuk menentukan dan mengetahui sentimen masyarakat terkait kebijakan pemerintah
Context	Publik

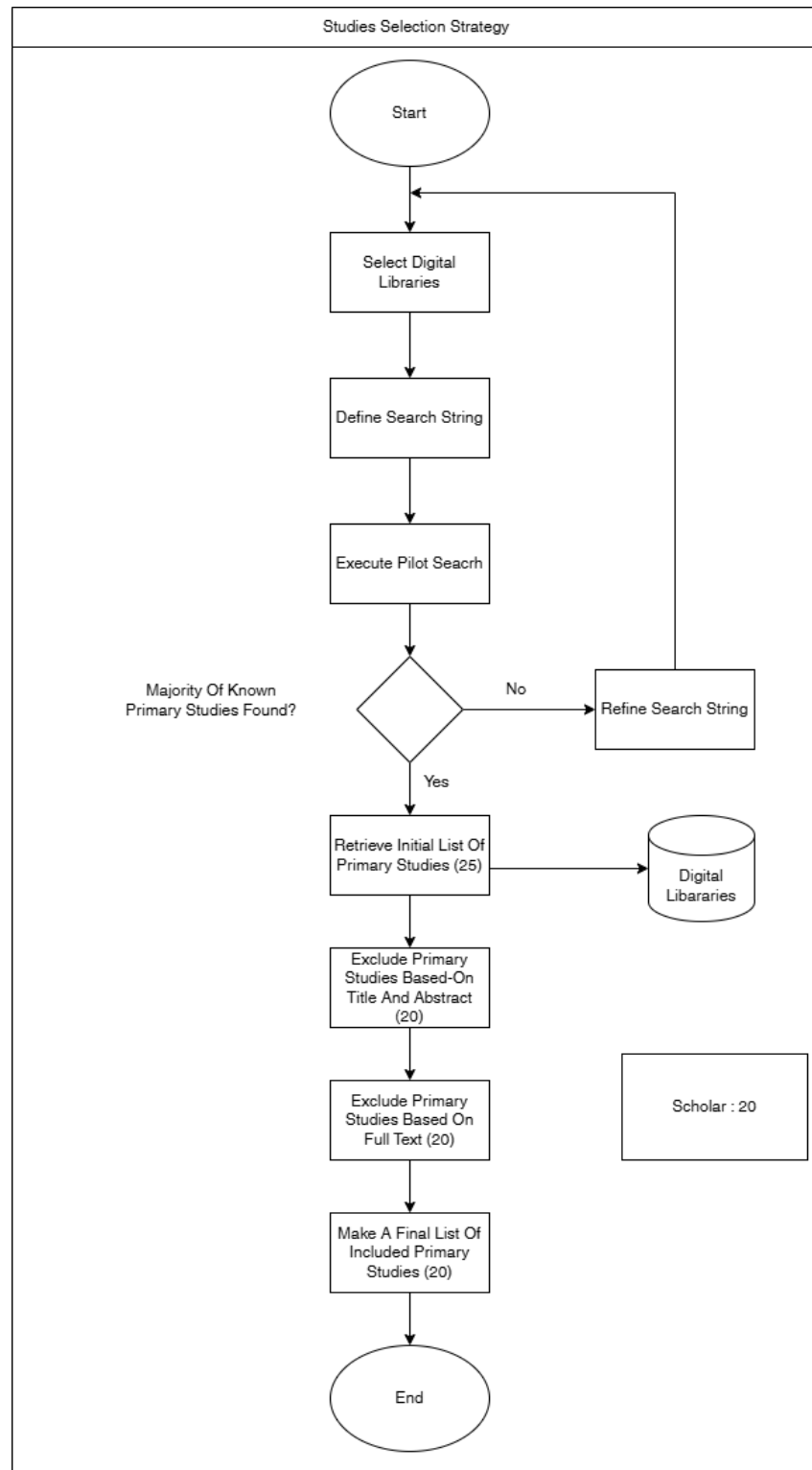
2.1.2 Survei Review Protocol

Penelitian survei adalah Penelitian ini dilakukan baik pada populasi besar maupun kecil, tetapi data yang dipelajari berasal dari sampel populasi untuk menemukan hubungan antara variabel sosiologis dan psikologis.. Proses melakukan metodologi survei dalam pengumpulan jurnal – jurnal ilmiah sesuai dengan judul skripsi dan yang di tulis dalam tabel 2.2 *Review Survei Protocol* ini yaitu :

Tabel 2. 2 Review Survei Protocol

<i>Review Survei Protocol</i>	
<i>Publication Year</i>	2022 s/d 2025
<i>Publication Type</i>	Jurnal, Buku
<i>Search String</i>	("All Metadata":Text Mining) OR ("All Metadata":Analisa sentimen) OR ("All Metadata":

	<i>Support Vector Machine , dan Naïve Bayes)</i>
<i>Final Selected</i>	25



Gambar 2. 1 Studies Selected Strategy

2.1.3 Kajian Pustaka Utama

Kajian pustaka merupakan data pendukung yang wajib dijadikan sebagai acuan yang didapat dari penelitian terdahulu yang berhubungan dengan penelitian ini. Berikut merupakan beberapa hasil penelitian yang berhubungan dengan penelitian ini antara lain seperti pada Tabel 2.3 sebagai berikut :

Tabel 2. 3 Kajian Pustaka Utama

KAJIAN PUSTAKA UTAMA PAPER			
JURNAL			
No.	Judul Paper	Nama Penulis	Nama Jurnal & No. ISSN / ISBN / DOI
	Analisis Sentimen Menggunakan Support Vector Machine Masyarakat Indonesia di Twitter terkait Bjorka	Adhitya Karel Maulaya, Junadhi	Jurnal Computer Science and Information Technology (CoSciTech) ISSN E: 2723-5661 ISSN P: 2723-567X
1	<p>Masalah yang ditemukan :</p> <p>Seorang hacker baru-baru ini membuat nama Bjorka menjadi kontroversi di media sosial. Karena gerakannya meretas banyak data pribadi, termasuk dokumen pemerintah, yang sering digunakan sebagai tujuan. Selain itu, sebagian besar dokumen diduga milik Presiden Indonesia Joko Widodo, dan gerakan hacker bersamarkan nama Bjorka membongkar data pribadi yang dimiliki pemerintah juga mendapatkan dukungan dari mayoritas netizen di media sosial. Untuk menghasilkan tahapan optimal dalam hal ini, penulis menggunakan metode Support Vector Machine.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Dibutuhkan sistem yang dapat mengkategorikan banyak cuitan</p>		

	<p>berbobot opini yang mengarah pada suatu kelas, tergolong positif, negatif, atau netral, karena media sosial yang berkomputasi dengan waktu nyata terhadap masyarakat seperti Twitter mampu mengirimkan berbagai ungkapan dan tanggapan terhadap tindakan Bjorka. Setelah itu, tim penulis melakukan pengujian dengan menggunakan 1000 crawling data tweet Twitter, yang menunjukkan bahwa metode Support Vector Machine dapat menghasilkan akurasi sebesar 62,33%[4].</p>		
	<p>ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGUNAKAN METODE NAIVE BAYES CLASSIFIER</p>	<p>Alfandi Safira, Firman Noor Hasan</p>	<p><i>ZONasi: Jurnal Sistem Informasi, Vol. 5 No. 1, Jan 2023</i> ISSN E: 2656-7407 ISSN P: 2656-7393</p>
2	<p>Masalah yang ditemukan :</p> <p>Karena mudah dan nyaman saat dilakukan, belanja online sangat disukai oleh masyarakat. Belanja online lebih mudah dengan metode pembayaran paylater. Namun, paylater juga menyebabkan perilaku yang tidak baik, seperti membeli secara impulsif. Peneliti melakukan penelitian ini karena tanggapan masyarakat terhadap paylater bervariasi.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Dalam penelitian ini, analisis sentimen dilakukan dengan menggunakan metode Naive Bayes Classifier dan TextBlob dari pustaka TetxBlob, yang digunakan dengan bahasa pemrograman Python. Dataset dari Twitter menghasilkan 405 data. Dengan menggunakan metode Naive Bayes Classifier, kami menemukan sentimen negatif sebesar 70,62%, atau 286 data, sentimen positif sebesar 22,72%, atau 92 data, dan sentimen netral sebesar 6,67%, atau 27 data. Dengan menggunakan metode TextBlob, kami menemukan sentimen negatif sebesar 55,8%, atau 226 data, sentimen positif sebesar 33,09%, atau 134 data, dan sentimen netral sebesar 11. Oleh karena itu, dapat disimpulkan bahwa masyarakat tidak</p>		

	menyukai penggunaan paylater. Dalam pengujian model dengan confusion matrix, algoritma Naive Bayes Classifier menunjukkan peningkatan akurat sebesar 91% dibandingkan TextBlob.yang hanya 61%[5].		
	Analisis Sentimen Masyarakat Terhadap Pilpres 2019 Berdasarkan Opini Dari Twitter Menggunakan Metode Naive Bayes Classifier.	Khoirul Zuhri, Nurul Adha Oktarini Saputri	Journal of Computer and Information Systems Ampera. ISSN: 2775-2496.
3	<p>Masalah yang ditemukan :</p> <p>Karena jumlah data yang besar, banyak komentar di media sosial Twitter tentang Pemilihan Presiden 2019 yang perlu diproses dan diperiksa secara otomatis. Klasifikasi sentimen positif dan negatif secara manual membutuhkan banyak waktu dan tenaga. Selain itu, mengklasifikasikan opini yang menggunakan kata-kata yang tidak baku dan beragam sulit.efektivitas.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Penerapan metode Naive Bayes Classifier untuk mengklasifikasikan tweet secara otomatis berdasarkan sentimen positif dan negatif. Metode ini cukup efektif, mencapai akurasi sekitar 71% dalam menentukan sentimen dari data tweet yang dikumpulkan[6].</p>		
4	A Comparison Support Vector Machine, Logistic Regression And Naïve Bayes For Classification Sentimen Analisis user Mobile App	Kiki Ahmad Baihaqi , Iwan Setyawan, Danny Manongga, Hendryanto Dwi Purnomo, Hendry , Ahmad Fauzi , Aprilia Hananto.	International Journal of Artificial Intelgence Research ISSN 2579-7298
	Masalah yang ditemukan :		

	<p>Data adalah hal yang paling penting, penggunaan data dapat berguna untuk mendapatkan evaluasi dari pengguna sistem atau aplikasi yang dibangun berdasarkan mobile. Tidak hanya itu, hasil penilaian atau penerimaan aplikasi mobile selama tahap uji coba dianggap penting, penilaian dan komentar dari pengguna langsung juga merupakan hal penting yang dapat menjadi masukan bagi pengembang aplikasi mobile. Penambangan data, atau yang dikenal dalam bahasa Inggris sebagai data mining, adalah jawaban untuk proses pengambilan data di media manapun. Dalam penelitian ini, penambangan data dilakukan pada media penyedia layanan unduh aplikasi mobile Google Playstore, yang menyediakan data dalam bentuk komentar dan penilaian..</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Setelah mengumpulkan data dan memperoleh parameter data terbaru yang ditentukan oleh 2000 komentar terakhir, data tersebut diproses terlebih dahulu dengan menghapus karakter ikon emot dan menghilangkan variabel yang tidak diperlukan sehingga data yang diperoleh dapat diproses ke tahap selanjutnya, yaitu klasifikasi berdasarkan penilaian dan komentar sentimen. Algoritma yang digunakan atau dibandingkan dalam penelitian ini adalah Support Vector Machine, regresi logistik, dan naïve bayes yang dikenal handal dalam pemrosesan data mining. Dalam penelitian ini, hasil akurasi adalah 88% untuk SVM, 90,5% untuk Regresi Logistik dan 91% untuk naïve bayes[7].</p>		
5	<p>Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM)</p>	<p>Hovi Sohibul Wafa, Asep Id Hadiana, Fajri Rakhmat Umbara</p>	<p>INFORMATICS AND DIGITAL EXPERT (INDEX) - VOL. 4 NO. 1 (2022) 40-45 ISSN- 2775-2208 EISSN - 2715-0453</p>
	<p>Masalah yang ditemukan :</p> <p>Diabetes mellitus adalah penyakit yang berbahaya, dan banyak orang di banyak negara, termasuk Indonesia. Meskipun penyebab utama diabetes</p>		

	<p>belum diketahui, banyak orang percaya bahwa gaya hidup dan gen dapat memainkan peran penting. Para peneliti bioinformatika berusaha mengatasi penyakit ini dan membuat sistem yang membantu memprediksi diabetes. Banyak penelitian yang ada menggunakan metode seperti C4.5, KNN, Naive Bayes, dan SVM Linier untuk membangun sistem. Namun, karena hasil yang dihasilkannya tidak cukup untuk digunakan pada sistem prediksi diabetes, metode SVM Radial Basis Function (RBF) jarang digunakan.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Penelitian ini menemukan bahwa metode algoritma SVM Radial Basis Function (RBF) dapat menghasilkan akurasi yang tinggi sebesar 91% dengan menggunakan matrika kekacauan dan peramalan kesalahan rata-rata kuadrat dengan kfold kelipatan 10. Dengan menggunakan teknik data mining dan klasifikasi menggunakan algoritma SVM Radial Basis Function Forward Selection, penelitian ini mencoba mengetahui apakah pasien atau penderita mungkin menderita diabetes.[8].</p>		
	<p>Analisa Sentimen Terhadap Kereta Cepat Jakarta – Bandung Menggunakan Algoritma Naïve Bayes Dan K-Nearest Neighbor</p>	<p>Zaky Rahman Hakim, Sugiyono</p>	<p>Jurnal Sains dan Teknologi, Volume 5 No. 3, Februari 2024, ISSN - 2714-8661</p>
6	<p>Masalah yang ditemukan :</p> <p>Banyak opini masyarakat Indonesia di media sosial tentang pembangunan proyek Kereta Cepat Jakarta – Bandung menunjukkan persepsi yang beragam, dengan sebagian besar bersikap negatif terhadap proyek tersebut. Hal ini menunjukkan perlunya analisis sentimen untuk memahami persepsi publik secara sistematis.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Penelitian ini menggunakan algoritma Naïve Bayes dan K-Nearest Neighbor untuk mengklasifikasikan tweet tentang kereta cepat menjadi dua kategori sentimen: positif dan negatif. Data tweet sebanyak 2000 dari</p>		

	<p>Twitter diolah dengan proses labeling otomatis, penggabungan data, dan preprocessing. Kemudian dilakukan pelatihan dan pengujian model, yang menunjukkan bahwa algoritma K-Nearest Neighbor mampu mengklasifikasi sentimen dengan akurasi tinggi sebesar 99,76%. Hasil ini menunjukkan bahwa algoritma tersebut efektif dalam mendeteksi persepsi masyarakat tentang proyek kereta cepat Jakarta – Bandung[9].</p>		
	<p>MULTI-ASPECT SENTIMENT ANALYSIS BASED ON EMOTICON CONVERSION WITH NAÏVE BAYES ALGORITHM FOR CULINARY TOURISM REVIEW ON TRIPADVISOR WEB</p>	<p>Sitti Aliyah Azzahra, Arief Wibowo</p>	<p>Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK) DOI: 10.25126/jtiik.20207 1907 p-ISSN: 2355-7699 e-ISSN: 2528-6579</p>
7	<p>Masalah yang ditemukan :</p> <p>Wisatawan sering menggunakan situs web seperti TripAdvisor untuk menemukan informasi tentang tempat wisata tertentu. Situs web ini memiliki fitur di mana pengguna terdaftar dapat menulis ulasan tentang objek wisata dalam kategori kuliner dari berbagai negara, dan ulasan tersebut dapat digunakan oleh wisatawan sebagai referensi sebelum melakukan perjalanan ke lokasi kuliner yang ingin mereka tuju. Ada kemungkinan untuk memeriksa ulasan atau komentar yang ada di situs TripAdvisor untuk mengetahui nilai sentimen dari objek wisata yang diulas. Analisis ini dapat membantu pengelola tempat wisata, pengusaha kuliner, dan wisatawan lainnya. Saat melakukan analisis sentimen pada kalimat ulasan yang mengandung ikon emosi atau emoticon, ada tantangan karena ulasan dapat mengandung arti sentimen yang berbeda antara kalimat dengan ekspresi emosi yang telah dilakukan. Studi ini menganalisis ulasan kuliner Bandung di TripAdvisor dan membagi</p>		

	<p>sentimen menjadi tiga kelas.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Penelitian ini menggunakan metode klasifikasi data mining menggunakan algoritme Naive Bayes, serta metode pelabelan multi aspek yang mengubah ikon emosi ke dalam teks ulasan. Selain itu, evaluasi dilakukan untuk menentukan nilai ulasan berdasarkan jumlah ulasan yang dibuat oleh pemberi ulasan di situs TripAdvisor. Hasil pengujian menunjukkan bahwa kombinasi metode ini dapat menghasilkan nilai akurasi sebesar 98,67% dalam proses klasifikasi sentimen.[10].</p>		
8	<p>Klasifikasi Kualitas Air Bersih Menggunakan Metode Naïve baiyes.</p>	<p>Sutisna, Mirsandi Nazar Yuniar</p>	<p>Jurnal Sains dan Teknologi, E-ISSN 2714-8661</p>
	<p>Masalah yang ditemukan :</p> <p>Air merupakan kebutuhan penting bagi manusia, namun tidak semua air layak di konsumsi. Sehingga penting untuk menentukan kualitas air yang di gunakan untuk konsumsi.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Peneliti melakukan pengujian dengan teknik analisis yang dapat di gunakan untuk membangun model dari sampel data dalam kelompok yang sama. Hasil yang di dapatkan menggunakan rapid miner dengan metode Naïve Baiyes menunjukan akurasi sebesar 97,35%[11].</p>		
9	<p>Analisis Sentimen Pengguna Twitter Terhadap BTS Menggunakan Algoritma Support Vector Machine</p>	<p>Tiara Safitri, Yuyun Umaidah, Iqbal Maulana</p>	<p>Journal of Applied Informatics and Computing (JAIC) e-ISSN: 2548-6861</p>
	<p>Masalah yang ditemukan :</p> <p>Saat ini, masih sedikit penelitian yang dilakukan tentang analisis sentimen terhadap publik figure sebagai influencer media sosial yang dapat membantu perusahaan dalam memilih publik figure sebagai BA untuk produk mereka. Untuk menyelesaikan masalah ini, penulis menggunakan metode Support Vector Machine (SVM) untuk menganalisis sentimen</p>		

	<p>pengguna Twitter terhadap grup musik BTS..</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Menurut analisis sentimen yang dilakukan pengguna Twitter terhadap grup musik BTS menggunakan algoritma Support Vector Machine (SVM) dengan kernel linear dan metodologi KDD dengan lima tahap: pemilihan data, preprocessing data, transformasi data, pembangunan teks, dan evaluasi. Dari 622 tweet yang telah dipilih dan diberi label oleh ahli Bahasa Indonesia, 546 di antaranya diberi nilai positif, dan 76 di antaranya diberi nilai negatif. Oleh karena itu, dapat disimpulkan bahwa pengguna Twitter lebih cenderung bersikap positif terhadap grup musik BTS[12].</p>		
	<p>Perbandingan Algoritma Naive Bayes dan K-Nearts Neighbor pada Analisis Sentimen Ulasan Aplikasi MyPertamina.</p>	<p>Husain Taufiqurrahman, Fetty Tri Anggraeny, Muhammad Muharrom Al Haromainy</p>	<p>Jurnal JATI (Jurnal Mahasiswa Teknik Informatika) Vol. 7 No. 6, Desember 2023.</p>
10	<p>Masalah yang ditemukan :</p> <p>Masalah yang ditemukan dalam penelitian ini terkait dengan ulasan pengguna aplikasi MyPertamina yang menunjukkan berbagai kendala dan ketidakpuasan, seperti kesulitan membuka aplikasi, keterbatasan metode pembayaran, kegagalan saat mendaftar, dan poin hadiah yang tidak bertambah meskipun transaksi berhasil. Kondisi ini menyebabkan munculnya sentimen negatif dan netral pada ulasan pengguna, serta mempengaruhi penilaian keseluruhan terhadap aplikasi.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Hasil solusi yang dilakukan adalah menerapkan dan membandingkan dua algoritma klasifikasi yaitu Naïve Bayes dan K-Nearest Neighbor (KNN) untuk analisis sentimen dari data ulasan tersebut. Setelah proses preprocessing dan pelabelan sentimen, terpilihah algoritma Naïve Bayes yang memperoleh tingkat akurasi sekitar 75%, sementara KNN berkisar antara 56% sampai 73%. Hasil ini menunjukkan bahwa Naïve Bayes lebih</p>		

	efektif untuk mengklasifikasi sentimen ulasan pengguna dalam penelitian ini[13].		
11	<p>ANALISA KEPUASAN MAHASISWA TERHADAP E-LEARNING MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE</p>	<p>Iwan Mahendro, Dhanan Abimanto</p>	<p>Jurnal Saintek Maritim, Volume 23 Nomor 1, September 2022 ISSN : 1412-6826 e-ISSN : 2623-2030</p>
	<p>Masalah yang ditemukan : Pembelajaran e-learning menggantikan pembelajaran tatap muka. Perguruan tinggi harus mengetahui apakah siswa puas atau tidak dengan pembelajaran e-learning karena hal ini tentunya membuat siswa menunjukkan berbagai reaksi.</p> <p>Hasil dan solusi yang telah di lakukan : Hasil analisis menggunakan algoritma SVM menunjukkan nilai akurasi sebesar 98,23%, yang menunjukkan bahwa metode algoritma SVM dapat memprediksi responden yang puas dan tidak puas dengan benar 98,23% dari total data. Ketepatan adalah rasio prediksi benar positif dibandingkan dengan hasil keseluruhan yang diprediksi positif, sehingga algoritma SVM ini memprediksi responden yang benar puas sebesar 95,65% dibandingkan dengan hasil keseluruhan yang diprediksi puas. Recall adalah rasio prediksi benar negatif dibandingkan dengan Algoritma svm menghasilkan nilai recall sebesar 95,65%, yang menunjukkan bahwa algoritma ini memperkirakan bahwa responden yang diprediksi akan lebih puas daripada responden yang sebenarnya merasa puas.[1].</p>		
12	<p>Analisis Sentimen Masyarakat di Twitter terhadap Pemerintahan Anies Baswedan Menggunakan Metode Naive Bayes Classifier</p>	<p>Febriyanti Rambu Bangu Kahi, Alfrian C. Talakua</p>	<p>Jurnal Minfo Polgan, p-ISSN: 2089-9424 e-ISSN: 2797-3298</p>

	<p>Masalah yang ditemukan :</p> <p>Penelitian ini dilatarbelakangi oleh meningkatnya opini publik di media sosial, khususnya Twitter, terhadap kinerja Anies Baswedan selama menjabat sebagai Gubernur DKI Jakarta. Dengan banyaknya informasi dan opini yang menyebar, masyarakat menjadi ingin mengetahui apakah kinerja tersebut mencerminkan bukti nyata dan layak dijadikan pertimbangan dalam pencalonan Anies sebagai Presiden pada Pemilu 2024.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Hasil Klasifikasi:</p> <ul style="list-style-type: none"> - Akurasi: 76,04% - Presisi: 71,43% - Recall: 81,63% - F1-score: 76,19% - Jumlah Sentimen Positif: 746 tweet (47,7%) - Jumlah Sentimen Negatif: 817 tweet (52,3%) <p>Kesimpulan:</p> <ul style="list-style-type: none"> - Meskipun ada respons positif, sentimen masyarakat terhadap pemerintahan Anies Baswedan lebih banyak yang negatif. - Hal ini menjadi masukan dalam menilai persepsi publik menjelang pencalonan Anies sebagai Presiden[14]. 		
13	<p>Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes.</p>	<p>Yuni Nurtikasari, Syariful Alam</p>	<p>INSOLOGI: Jurnal Sains dan Teknologi e-ISSN (Online): 2828-4984 p-ISSN (Cetak): 2828-4992</p>
	<p>Masalah yang ditemukan :</p> <p>Banyaknya opini masyarakat mengenai film Ngeri-Ngeri Sedap di platform Twitter menimbulkan variasi sentimen, baik positif, negatif, maupun netral. Opini ini penting untuk dipetakan karena menjadi acuan bagi publik untuk menilai kualitas film. Namun, tanpa klasifikasi sentimen</p>		

	<p>yang sistematis, sulit untuk mengetahui kecenderungan persepsi masyarakat secara akurat.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Data Collection: Tweet dengan kata kunci “Ngeri-nger Sedap” dikumpulkan dari Twitter. Preprocessing: Termasuk transformation, tokenization, dan filtering. TF-IDF: Untuk pembobotan kata. Algoritma: Naïve Bayes digunakan untuk klasifikasi sentimen (positif, negatif, netral).Evaluasi: Menggunakan confusion matrix melalui tools <i>Orange</i>.</p> <p>Hasil Evaluasi:</p> <ul style="list-style-type: none"> - Akurasi: 75% - Presisi: 80% - Recall: 79% - Kesimpulan: Mayoritas opini masyarakat terhadap film <i>Ngeri-Ngeri Sedap</i> tergolong netral[15]. 		
14	<p>Analisis Sentimen Tanggapan Masyarakat Aplikasi TikTok Menggunakan Metode Naïve Bayes dan Categorical Proportional Difference (CPD)</p>	<p>Junda Alfiah Zulqornain, Indriati</p>	<p>Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN: 2548-964X</p>
	<p>Masalah yang ditemukan :</p> <p>Aplikasi TikTok banyak digunakan oleh masyarakat dari berbagai usia, termasuk anak-anak di bawah umur. Namun, TikTok juga menampilkan konten yang bisa bersifat vulgar atau tidak layak bagi anak-anak. Selain itu, fitur komentar memungkinkan terjadinya bullying dan narsisme. Banyak ulasan negatif tentang aplikasi ini muncul di Google Play Store. Oleh karena itu, perlu dilakukan klasifikasi opini pengguna terhadap TikTok untuk membantu pengambilan keputusan, terutama bagi orang tua.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Penulis melakukan analisis sentimen terhadap ulasan TikTok menggunakan:</p>		

	<ul style="list-style-type: none"> - Metode Naïve Bayes untuk klasifikasi sentimen. - Categorical Propotional Difference (CPD) untuk seleksi fitur. - Data berupa 1000 ulasan TikTok dari Google Play Store, dilabeli sebagai positif atau negatif. - Pengujian dilakukan dengan 5-Fold Cross Validation. <p>Hasil Pengujian Terbaik (100% term):</p> <ul style="list-style-type: none"> - Akurasi: 72,99% - Presisi: 74,69% - Recall: 92,61% - F-Measure: 82,45% <p>Dengan penggunaan 100% term memberikan hasil terbaik dalam klasifikasi opini pengguna TikTok. Naïve Bayes dengan CPD efektif untuk memproses data ulasan dan memberikan wawasan mengenai kecenderungan opini masyarakat terhadap aplikasi ini[16].</p>		
	<p style="text-align: center;">Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis pada Media Sosial X Menggunakan Algoritma Naïve Bayes</p>	<p style="text-align: center;">Altolyto Sitanggung, Yuyun Umaidah</p>	<p style="text-align: center;">JITET Jurnal Informatika dan Teknik Elektro Terapan. p-ISSN: 2303-0577 e-ISSN: 2830-7062 DOI:10.23960/jitet. v12i3.4902</p>
15	<p>Masalah yang ditemukan :</p> <p>Program makan siang gratis yang diusulkan oleh pasangan calon presiden nomor urut 02 untuk anak dan ibu hamil memicu pro dan kontra di media sosial X. Penelitian ini bertujuan mengetahui bagaimana opini publik (sentimen) terhadap program ini, karena opini tersebut bisa mencerminkan dukungan atau penolakan terhadap kebijakan yang berdampak nasional.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Metode yang digunakan:</p> <ul style="list-style-type: none"> - Knowledge Discovery in Database (KDD) sebagai kerangka kerja. - Algoritma Naïve Bayes digunakan untuk klasifikasi sentimen. 		

	<ul style="list-style-type: none"> - Preprocessing data mencakup: cleaning, case folding, stopword removal, tokenizing, stemming. - TF-IDF digunakan sebagai metode pembobotan kata. - Evaluasi menggunakan confusion matrix dengan rasio pembagian data: 80:20, 70:30, dan 60:40. <p>Hasil terbaik diperoleh pada rasio 60:40, dengan: Akurasi: 72,2% Presisi: 63,2% Recall: 66,1% F1-Score: 64% [17].</p>		
	Analisis Sentimen dengan SVM, Naive Bayes dan KNN untuk Studi Tanggapan Masyarakat Indonesia Terhadap Pandemi Covid-19 pada Media Sosial Twitter	Fajar Sodik Pamungkas, Iqbal Kharisudin	PRISMA, Prosiding Seminar Nasional Matematika ISSN: 2613-9189
16	<p>Masalah yang ditemukan :</p> <p>Pandemi Covid-19 memicu berbagai tanggapan masyarakat di media sosial, khususnya Twitter. Dengan meningkatnya opini publik, penting untuk mengetahui sentimen masyarakat secara sistematis guna mengevaluasi persepsi terhadap pandemi dan kebijakan pemerintah. Belum diketahui secara pasti algoritma klasifikasi mana yang paling efektif dalam menganalisis sentimen tersebut.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>SVM adalah algoritma terbaik dalam klasifikasi sentimen tweet tentang Covid-19. Sentimen masyarakat lebih didominasi oleh tanggapan positif (6.128 dari 10.000). Wordcloud dan asosiasi kata menunjukkan bahwa tanggapan positif mencakup semangat, dukungan, dan kesehatan, sedangkan tanggapan negatif menyoroti kritik terhadap lambannya penanganan, terutama di sektor ekonomi [18].</p>		
17	Analisis Sentimen Terhadap Pemindahan Ibu Kota Negara Menggunakan Algoritma	Dedi Pramana, M. Afdal	Jurnal Media Informatika Budidarma e-ISSN: 2548-8368

	Naive Bayes Classifier dan K-Nearest Neighbors (KNN)		p-ISSN: 2614-5278
	<p>Masalah yang ditemukan :</p> <p>Penelitian ini membahas isu kontroversial mengenai pemindahan Ibu Kota Negara Indonesia. Topik ini menimbulkan berbagai reaksi publik di media sosial, yang terdiri dari opini positif maupun negatif. Tantangan utamanya adalah bagaimana mengklasifikasikan sentimen publik tersebut secara akurat untuk memberikan gambaran umum terhadap respons masyarakat.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Akurasi model Naive Bayes dan K-NN dibandingkan dalam mengklasifikasikan opini menjadi sentimen positif atau negatif. Hasil menunjukkan bahwa salah satu dari algoritma tersebut (belum disebutkan eksplisit dalam kutipan) lebih unggul dalam performa klasifikasi sentimen[19].</p>		
	Analisis Sentimen Masyarakat Terhadap Pemerintah di Era Kabinet Joko Widodo Berdasarkan Sosial Media X Menggunakan Naive Bayes dan K-Nearest Neighbor (KNN)	Fadhel Naufal Rahman, Sri Lestari	Journal of Information Technology and Computer Science (INTECOMS) e-ISSN: 2614-1574 p-ISSN: 2621-3249
18	<p>Masalah yang ditemukan :</p> <p>Selama masa pemerintahan Presiden Joko Widodo, banyak kebijakan yang menimbulkan pro dan kontra di kalangan masyarakat. Opini publik banyak dituangkan melalui media sosial (platform X/Twitter), namun belum tersedia sistem otomatis yang mampu mengklasifikasikan opini-opini ini ke dalam kategori sentimen (positif atau negatif). Tanpa klasifikasi ini, pemerintah kesulitan menilai tanggapan masyarakat terhadap kebijakan yang diambil.</p> <p>Hasil dan solusi yang telah di lakukan :</p>		

	<p>Mengambil data 5329 tweet tentang "Kabinet Jokowi" melakukan pembersihan (cleansing) data menjadi 3437 tweet, melakukan pelabelan manual terhadap sentimen tweet lalu menggunakan RapidMiner untuk memodelkan dan mengukur akurasi kedua algoritma juga menerapkan metode CRISP-DM untuk proses data mining. Algoritma K-Nearest Neighbor lebih baik dibandingkan Naïve Bayes dalam klasifikasi sentimen pada data yang digunakan. Sebagian besar opini masyarakat (98%) bersifat positif terhadap Kabinet Jokowi[20].</p>		
	<p>Analisis Sentimen Pengguna Twitter Mengenai Calon Presiden Indonesia Tahun 2024 Menggunakan Algoritma LSTM</p>	<p>Mohammed Hafizh Al-Areef, Kana Saputra S</p>	<p>Jurnal SAINTIKOM e-ISSN: 2615-3475 p-ISSN: 1978-6603</p>
19	<p>Masalah yang ditemukan : Masyarakat Indonesia aktif menyampaikan opini mengenai calon presiden 2024 melalui Twitter. Namun, banyaknya data tersebut belum diklasifikasikan ke dalam kategori sentimen (positif atau negatif). Masyarakat dan pihak berkepentingan belum memiliki gambaran umum tentang persepsi publik terhadap calon-calon presiden, seperti Ganjar Pranowo, Prabowo Subianto, dan Ridwan Kamil. Maka diperlukan analisis sentimen otomatis yang akurat.</p> <p>Hasil dan solusi yang telah di lakukan : Model LSTM mampu mengklasifikasikan sentimen masyarakat di Twitter terhadap calon presiden 2024 dengan tingkat akurasi yang tinggi. Ridwan Kamil memperoleh performa model terbaik dari sisi semua metrik evaluasi. Pendekatan ini bisa menjadi alat bantu dalam memetakan opini publik secara real-time terhadap dinamika politik nasional[21].</p>		
20	<p>Analisa Sentimen Masyarakat Indonesia terhadap Kinerja Prabowo Subianto sebagai Anggota</p>	<p>Syamsu Hidayat</p>	<p>JUNIF: Jurnal Nasional Informatika p-ISSN: 2774-3497</p>

	Kabinet Indonesia Maju di Platform Twitter Menggunakan R dengan Algoritma Naïve Bayes		e-ISSN: 2774-3500
	<p>Masalah yang ditemukan :</p> <p>Pemerintah Indonesia, khususnya Kementerian Pertahanan, memiliki tantangan serius dalam hal modernisasi alat utama sistem persenjataan (alutsista) serta efisiensi belanja pertahanan. Di tengah sorotan terhadap kebijakan pertahanan, kinerja Prabowo Subianto sebagai Menteri Pertahanan menimbulkan berbagai tanggapan di masyarakat. Namun, belum tersedia analisis yang sistematis terhadap opini masyarakat di media sosial, terutama Twitter.</p> <p>Hasil dan solusi yang telah di lakukan :</p> <p>Penelitian ini menerapkan pendekatan analisis sentimen dengan memanfaatkan:</p> <ul style="list-style-type: none"> - Platform R dan RStudio - Twitter API untuk melakukan crawling data - Algoritma Naïve Bayes untuk klasifikasi sentimen - Visualisasi hasil menggunakan R Shiny <p>Dari total tweet yang dianalisis:</p> <ul style="list-style-type: none"> - Sentimen Negatif: 1.537 tweet - Sentimen Positif: 1.084 tweet - Sentimen Netral: 142 tweet <p>Masyarakat Indonesia cenderung memberikan penilaian negatif terhadap kinerja Prabowo Subianto sebagai anggota Kabinet Indonesia Maju, berdasarkan data Twitter yang dianalisis[22].</p>		
21	Analisis Sentimen dan Penggalan Opini	Asrumi, Didik Suharijadi, Agustina Dewi Setiari, Diah Putri Wulanda	Eureka Media Aksara ISBN : 978-623-151-907-8
	Masalah yang ditemukan :		

	<p>Apa itu Analisis Sentimen?</p> <p>Hasil dan Solusi Yang ditemukan :</p> <p>Bidang Studi yang menganalisis opini, sentiment, evaluasi, penilaian, sikap, dan emosi masyarakat terhadap entitas seperti produk, layanan, organisasi, individu, isu, peristiwa, topik, dan atributnya[23].</p>		
	<p>PENGANTAR METODE ANALISIS SENTIMEN</p>	<p>Detty Purnamasari, Ananda Bayu Aji, Desy Wulandari A.P., Fanka Arie Reza, Milda Safrila O., Nafa Yanda, Ulfa Hidayati</p>	<p>Penerbit Gunadarma</p>
22	<p>Masalah yang ditemukan :</p> <p>Karena kemajuan teknologi dan internet, orang sekarang dapat dengan mudah mengakses berbagai hal melalui ponsel mereka. Salah satu contohnya adalah banyak media, yang memungkinkan orang untuk mengungkapkan pendapat mereka dan memberikan komentar tentang produk, kejadian, dan hal-hal lainnya.</p> <p>Hasil dan Solusi yang ditemukan :</p> <p>Psikolinguistik adalah bidang yang mempelajari cara seseorang memperoleh, menghasilkan, dan memahami bahasa lisan dan tulisan. Perkembangan teknologi dalam proses pengolahan bahasa natural tentunya tidak terlepas dari konsep psikolinguistik. Analisis sentimen dapat dilakukan dengan berbagai algoritma dan metode. Beberapa di antaranya adalah Naive Baye, Support Vector Machine (SVM), Berbasis Lexicon, dan Long Short Term Memory (LSTM).[2].</p>		
23	<p>Data Mining</p>	<p>Amna, Wahyuddin S, I Gede Iwan Sudipa, Tri Andi E. Putra, Ahmad Jurnaidi Wahidin, Wara Alfa Syukrilla,</p>	<p>PT GLOBAL EKSEKUTIF TEKNOLOGI Anggota IKAPI No. 033/SBA/2022 ISBN: 978-623-198-088- 5</p>

		Anindya Khrisna Wardhani, Nono Heryana, Tutuk Indriyani, Leo Willyanto Santoso,	
	<p>Masalah yang ditemukan : Apa itu Data Mining?</p> <p>Hasil dan Solusi yang ditemukan : Data mining juga dikenal sebagai penemuan pengetahuan, ekstraksi pengetahuan, analisis data/pola, pengumpulan informasi, dan lainnya. Proses data mining terdiri dari pengumpulan data, ekstraksi data, analisa data, dan statistik data.[24].</p>		
	Naïve Bayes pada Google Colabs	Ari Mulyoto, S.Pd, M.Si.	CV. EUREKA MEDIA AKSARA ISBN : 978-623-516-800- 5
24	<p>Masalah yang ditemukan :</p> <ol style="list-style-type: none"> Apa itu Naïve bayes? Apa itu Google Colab? <p>Hasil dan Solusi yang ditemukan :</p> <ol style="list-style-type: none"> Naive Bayes adalah kumpulan algoritma klasifikasi yang didasarkan pada Teorema Bayes, yang mengasumsikan bahwa fitur (atau variabel) yang digunakan dalam prediksi adalah independen satu sama lain. Meskipun asumsi ini mungkin tidak selalu benar dalam praktik, Naive Bayes sering kali memberikan hasil yang baik, terutama dalam kasus di mana ukuran dataset relatif kecil atau saat fitur memiliki hubungan yang kompleks. Google Colab, atau Google Colaboratory, adalah platform berbasis cloud yang memungkinkan pengguna untuk menulis dan menjalankan kode Python dalam lingkungan yang interaktif. Dikembangkan oleh Google, Colab menyediakan kemampuan untuk mengakses sumber daya komputasi yang kuat, termasuk 		

	<p>GPU (Graphics Processing Unit) dan TPU (Tensor Processing Unit), tanpa memerlukan instalasi perangkat lunak tambahan di komputer lokal. Colab sangat populer di kalangan peneliti, pengembang, dan mahasiswa karena kemudahan akses dan kemampuannya untuk mendukung proyek-proyek machine learning, data science, dan analisis data[25].</p>		
	<p>ALGORITMA SUPPORT VECTOR MACHINE UNTUK MENGANALISA KEPUASAN MAHASISWA TERHADAP PEMBELAJARAN E- LEARNING</p>	<p>Iwan Mahendro, S.Kom., M.Pd., Dhanan Abimanto, S.Hum., M.Pd</p>	<p>CV. Pustaka STIMART-AMNI Semarang</p>
25	<p>Masalah yang ditemukan :</p> <p>Selama pandemi, perguruan tinggi telah menggunakan e-learning. Pembelajaran online memiliki kelebihan dan kekurangan dibandingkan dengan pembelajaran tatap muka. Tujuan khusus penelitian ini adalah untuk menemukan dan menguji potensi penggunaan e-learning dengan algoritma svm untuk meningkatkan kinerja pembelajaran siswa.</p> <p>Hasil dan Solusi yang ditemukan :</p> <p>Hasil dari analisis yang dilakukan menggunakan algoritma Support Vector Machine, yaitu munculnya Matriks Konfussion. Dalam masalah klasifikasi pembelajaran mesin, di mana keluaran biasanya terdiri dari dua kelas atau lebih, metrik kekacauan digunakan untuk mengukur kinerja. Tabel Confusion Matrix terdiri dari empat gabungan nilai prediksi dan nilai aktual. True Positif, True Negatif, False Positif, dan False Negatif adalah empat hasil klasifikasi dari matriks confusion. Kami menemukan akurasi sebesar 98,23%.[26].</p>		

2.1.4 Matrik Perbandingan RP, RQ, RO

Masalah Penelitian, Identifikasi Masalah dan Tujuan Penelitian Dari hasil tinjauan pustaka terkait judul proposal yang di tulis dan di rangkum secara keseluruhan pada Tabel 2.4 sebagai berikut yaitu :

Tabel 2. 4 Matrix Perbandingan RP, RQ, RO

Identifikasi Masalah (<i>Research Problem / RP</i>)	Perumusan Masalah (<i>Research Question / RQ</i>)	Tujuan Penelitian (<i>Research Objective / RO</i>)
--	--	---

RP -1	Adanya pro & kontra masyarakat terkait Tren Kabur aja Dulu	RQ -1	Bagaimana cara menerapkan <i>text mining</i> dengan metode <i>Naïve Bayes</i> dan <i>Support Vector Machine</i> untuk melakukan analisa sentimen masyarakat terhadap Tren Kabur Aja Dulu?	RO -1	Menerapkan <i>text mining</i> menggunakan metode <i>Naïve Bayes</i> dan <i>Support Vector Machine</i> untuk mengetahui presentase sentimen positif dan negatif tentang Tren Kabur Aja Dulu.
----------	--	----------	---	----------	---

RP -2	Sentimen masyarakat terhadap Tren Kabur Aja Dulu.	RQ -2	Berapakah tingkat akurasi kinerja metode <i>Naïve Bayes</i> dan <i>K-Nearest Neighbor</i> dalam analisa sentimen Tren Kabur Aja Dulu?	RO -2	Agar publik mengetahui sentimen tentang Tren Kabur Aja Dulu.
----------	---	----------	---	----------	--

2.2 Definisi dan Pengertian

2.2.1 Text Mining

Text mining penggunaan teknik AI dan NLP untuk mengekstraksi informasi dan wawasan dari teks. Teknik-teknik ini mengubah data yang tidak terstruktur menjadi data terstruktur sehingga memudahkan dalam analisis data. Tujuan *text mining* adalah untuk pengkategorian data teks, baik itu untuk menemukan kategori yang sesuai dengan kelas yang ditentukan (*supervised learning*) maupun berdasarkan kesamaan karakteristik dan memberikan label pada kelas yang belum diketahui (*unsupervised learning*)[27].

Analisis sentimen adalah salah satu dari beberapa kategori teknik text mining; ini adalah proses memahami, mengekstrak, dan mengolah data teks secara otomatis; atau itu adalah studi komputasi pendapat, perasaan, dan emosi yang diungkapkan dalam teks.

2.2.2 Data Mining

Data mining juga dikenal sebagai penemuan pengetahuan, ekstraksi pengetahuan, analisis data/pola, pengumpulan informasi, dan lainnya.

Proses ini terdiri dari pengumpulan data, ekstraksi data, analisa data, dan statistik data.[24].

2.2.3 Analisa Sentimen

Analisis sentimen adalah teknik pemrosesan bahasa alami (NLP) dan pembelajaran mesin (ML) untuk mengevaluasi secara otomatis teks berdasarkan perasaan yang dirasakan penulis (positif, negatif, dan netral). Analisis sentimen juga merupakan area pemrosesan bahasa alami (NLP) yang berkembang karena interaksi manusia dengan komputer, ekstraksi informasi, dan penyulingan perasaan (distillation of feelings) dari data yang sudah ada. Ini mencakup memahami kata atau frasa yang menunjukkan sikap netral, positif, atau negatif dalam teks. Analisis sentimen biasanya digunakan untuk mengekstrak berbagai karakteristik dari data teks terstruktur atau tidak terstruktur dan menganalisisnya untuk mengekstrak pikiran, pendapat, dan perasaan dari data tersebut.[2].

2.2.4 Twitter (X)

Twitter adalah suatu media dan layanan sosial *microblogging* yang populer di kalangan pengguna internet. Penggunanya dapat mengekspresikan apa yang dipikirkan dalam pesan waktu nyata. Pengguna Twitter bebas menyampaikan pendapat atau ekspresi mereka tentang layanan, fasilitas atau isu politik atau hal-hal yang sedang dibahas.

2.2.5 Application Programming Interface (API)

Menurut Ardhi Wijayanto, Arif Rohmadi, dan Udhi Permana (2018), "API" adalah singkatan dari Interface Pembelajaran Aplikasi, dan merupakan kumpulan fungsi, perintah, dan protokol yang berfungsi untuk menghubungkan aplikasi satu sama lain.

Neil Madden (2020) menyatakan bahwa API adalah batas antara komponen sistem perangkat lunak. Ini menjelaskan kumpulan

operasi yang diberikan oleh satu komponen kepada bagian lain sistem untuk digunakan.

Peneliti sampai pada kesimpulan bahwa API adalah bagian dari sistem perangkat lunak yang mencakup kumpulan protokol, fungsi, dan perintah yang memungkinkan sistem komputer berinteraksi satu sama lain. [28].

2.2.6 Google Colab

Google Colab, atau *Google Colaboratory*, adalah platform berbasis cloud yang memungkinkan pengguna untuk menulis dan menjalankan kode Python dalam lingkungan yang interaktif. Dikembangkan oleh Google, Colab menyediakan kemampuan untuk mengakses sumber daya komputasi yang kuat, termasuk GPU (Graphics Processing Unit) dan TPU (Tensor Processing Unit), tanpa memerlukan instalasi perangkat lunak tambahan di komputer lokal. Colab sangat populer di kalangan peneliti, pengembang, dan mahasiswa karena kemudahan akses dan kemampuannya untuk mendukung proyek-proyek machine learning, data science, dan analisis data[25].

2.2.7 Python

Python adalah bahasa pemrograman interpretatif yang mudah dipelajari dan dapat dijalankan pada berbagai platform dengan fokus utama pada keterbacaan kode[5]. *Python* diklaim sebagai bahasa yang menggabungkan kapabilitas, kemampuan, dengan sintaksis kode yang sangat jelas, dan dilengkapi dengan fungsionalitas pustaka standar yang besar serta komprehensif.

2.2.8 Klasifikasi

Klasifikasi adalah proses pengelompokan data ke dalam sebuah kelas yang sudah ditentukan berdasarkan ciri, sifat, dan pola masing-masing kelas. Proses ini dimulai dengan penambahan kelompok atau sekumpulan data sebagai pedoman kriteria dalam sebuah kelas, dan

kemudian pedoman tersebut digunakan untuk memilah-milah setiap data yang memenuhi kriteria kelas tersebut. Klasifikasi dibagi menjadi dua fase dalam prosesnya, yaitu fase belajar dan pengujian (Pramudiono, 2003)[26].

Tujuan utama klasifikasi adalah untuk mengidentifikasi pola atau karakteristik yang membedakan setiap kelas, sehingga objek yang tidak dikenal dapat diklasifikasikan ke dalam kelas yang sesuai berdasarkan atribut yang diamati. Pada penelitian ini klasifikasi di bagi menjadi tiga kategori yaitu positif dan negatif.

2.2.9 Naïve Bayes

Kumpulan algoritma klasifikasi Naïve Bayes didasarkan pada Teorema Bayes, yang menganggap bahwa fitur atau variabel yang digunakan untuk prediksi tidak bergantung satu sama lain. Meskipun asumsi ini mungkin tidak selalu benar dalam kehidupan nyata, Naive Bayes sering kali memberikan hasil yang baik, terutama ketika datasetnya kecil atau ketika hubungan antara fiturnya kompleks. Berikut ini adalah penjelasan lebih lanjut tentang Naive Bayes.[25].

2.2.10 Support Vector Machine (SVM)

Untuk klasifikasi (seperti Support Vector Classification) dan regresi (seperti Support Vector Regression), Support Vector Machines (SVM) adalah metode pengajaran yang diawasi. Dibandingkan dengan metode klasifikasi lainnya, SVM memiliki konsep matematis yang lebih matang dalam pemodelan klasifikasi. Selain itu, SVM dapat menangani masalah regresi dan klasifikasi baik linear maupun non-linear. SVM dapat mengklasifikasikan data di mana kelas tidak dapat dipisahkan secara linear dengan memetakan data. Algoritma SVM memiliki keunggulan dibandingkan dengan algoritma lain.[26].

2.2.12 Dataset

Dataset adalah kumpulan data yang terorganisir secara sistematis untuk tujuan analisis. Dataset dapat berisi berbagai jenis data, seperti teks, gambar, audio, atau nilai numerik, dan biasanya digunakan dalam berbagai bidang, termasuk ilmu komputer, ilmu data, statistik, dan penelitian ilmiah. Dataset sering digunakan untuk melatih model kecerdasan buatan, memvalidasi hipotesis, membuat prediksi, atau mendapatkan wawasan yang berharga melalui analisis data.

2.2.13 Rapid Miner

Rapid Miner adalah platform analitik yang kuat dan serbaguna yang digunakan untuk memproses, menganalisis, dan memodelkan data. Platform ini menyediakan lingkungan visual yang intuitif yang memungkinkan pengguna untuk melakukan berbagai tugas analitik tanpa perlu menulis kode secara manual. *Rapid Miner* mendukung sejumlah besar algoritma pembelajaran mesin dan analisis statistik, serta menyediakan alat untuk pembersihan data, ekstraksi fitur, visualisasi data, dan evaluasi model.

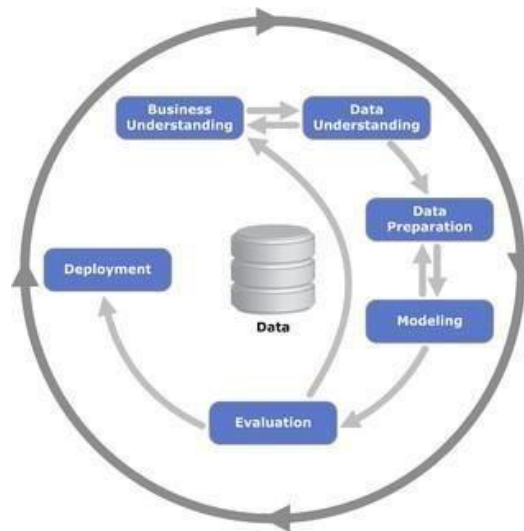
Dengan menggunakan *Rapid Miner*, pengguna dapat mengimpor data dari berbagai sumber, mulai dari file lokal hingga database terpusat, dan kemudian melakukan serangkaian operasi analitik untuk mendapatkan wawasan yang berharga dari data tersebut. Platform ini digunakan secara luas dalam berbagai industri, termasuk bisnis, keuangan, kesehatan, pemasaran, dan lainnya, untuk memecahkan berbagai masalah analitik dan mengambil keputusan yang didukung oleh data.

2.2.14 Cross-Industry Process For Data Mining (CRISP-DM)

CRISP-DM (Cross-Industry Standard Process for Data Mining) dibuat pada tahun 1996 oleh analisis industri seperti NCR, Daimler Chrysler, dan SPSS. CRISP-DM menawarkan standar proses data mining sebagai strategi pemecahan masalah umum untuk bisnis atau

unit penelitian. CRISP-DM juga menyediakan metodologi baku untuk data mining.[29].

CRISP-DM merupakan standar dalam data mining terdiri dari 6 tahap yang mana 3 tahap paling awal berdasarkan pengalaman penulis dapat *Non-Mutually Exclusive*. Berikut disajikan keenam tahapan CRISPDM tersebut dalam Gambar 2.2 dibawah ini :



Gambar 2. 2 Fase CRISP-DM

Pada gambar diatas terdapat 6 Fase pada CRISP-DM yang mana penjelasannya adalah sebagai berikut :

1. *Bussiness Understanding* (Pemahaman Bisnis)

Pada tahap pertama CRISP DM, pemahaman bisnis diperlukan untuk mendefinisikan masalah bisnis dengan sederhana dan tepat. Pada tahap ini membutuhkan pemahaman tentang dua aspek berbeda: pemahaman tentang proses bisnis termasuk regulasi yang mengaturnya, dan pemahaman tentang cara pengolahannya.

2. *Data Understanding* (Pemahaman Data)

Data Understanding merupakan tahap dimana data dikumpulkan dan dieksplorasi untuk mengenal karakteristiknya. Ini melibatkan kegiatan seperti pengumpulan data, deskripsi data, penilaian kualitas data, dan eksplorasi data awal.

3. *Data Preparation* (Persiapan Data)

Proses data preparasi merupakan proses data treatment menuju model berkualitas yang berguna. Tahapan ini adalah yang paling menguras resources dari tim analisis. Model yang baik dan akurat berawal dari data preparasi yang baik. Beberapa hal yang umum dilakukan pada tahapan ini seperti pada dibawah ini :

a. Melakukan pengecekan kembali kebenaran data

Dalam desain bertingkat, pengecekan data harus dilakukan untuk menjaga akuntabilitas dan untuk memastikan konsistensi input data. Sistem pengumpulan data yang baik, seperti default, dapat memastikan konsistensi data..

b. Mengelola data outlier

Data outlier dapat berupa *univariate* atau *multivariate*, dan dapat terjadi pada *variable dependent* atau *independent*. Untuk tujuan generalisasi, data mining harus dinetralisir karena berdampak pada data outlier. Sangat disarankan untuk melakukan evaluasi pengambilan dan pengisian data sebelum melakukan perawatan data outlier.

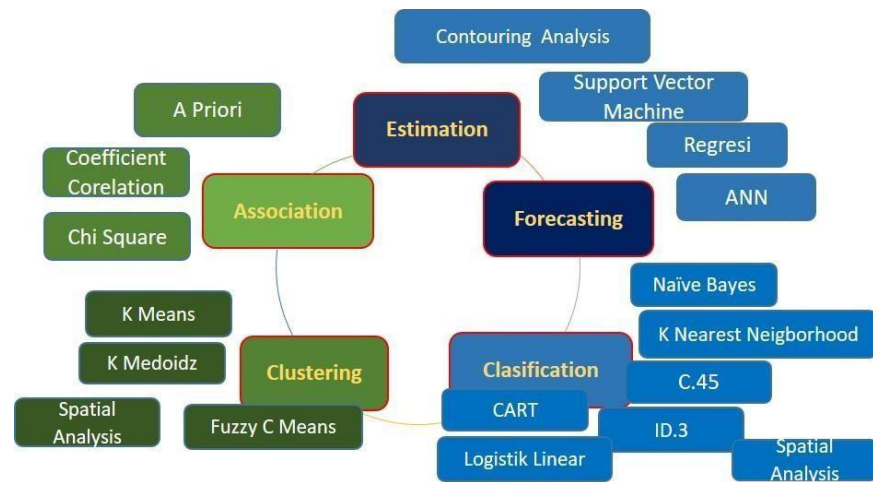
c. Memberlakukan *data missing* dan data inkonsistensi

Perilaku terhadap data yang tidak ada harus sesuai dengan tujuan dari pengolahan data. Misalnya, data yang tidak ada yang dipenuhi dengan rata-rata mungkin masih dapat digunakan untuk tujuan ramalan dan peramalan. Namun, mengklasifikasikan data yang tidak ada dapat mengarah pada kelompok yang kurang tepat..

4. *Modeling* (pemodelan)

Model adalah penjelasan yang berkualitas tinggi tentang pengetahuan yang dihasilkan oleh proses atau sistem melalui kalkulasi dan prediksi yang diterima. Dalam penelitian ekonomi, kata "sifat berterima" mengacu pada sedikitnya dua hal: teknikal tepat dan ekonomis tepat. Terdapat 5 (lima) tujuan utama dalam *data Mining* yaitu: Estimasi, Prediksi/Peramalan, Klasifikasi, Klastering, Asosiasi. Lebih dari satu

peran dapat dikombinasikan untuk mencapai tujuan yang lebih detail. Berdasarkan peran tersebut terdapat beberapa Algoritma yang dapat



digunakan sebagaimana disampaikan pada gambar 2.3 di bawah ini :

Gambar 2. 3 Algoritma Permodelan

Algoritma Gambar di atas biasanya membagi algoritma dalam dua kelompok. Kelompok biru (biru muda hingga biru tua) dikenal sebagai pembelajaran supervisi, sedangkan kelompok hijau (hijau muda hingga hijau tua) dikenal sebagai pembelajaran tidak supervisi. Pembelajaran supervisi memerlukan target atau *dependent variable*, sedangkan pembelajaran tidak supervisi tidak menggunakannya.

5. *Evaluation* (Evaluasi)

Melakukan interpretasi terhadap hasil dari data mining yang dihasilkan dalam proses pemodelan pada tahap sebelumnya. Evaluasi dilakukan terhadap model yang diterapkan pada tahap sebelumnya dengan tujuan agar model yang ditentukan dapat sesuai dengan tujuan yang ingin dicapai dalam tahap pertama.

6. *Deployment* (Pengembangan)

Tahap *deployment* atau rencana penggunaan model adalah tahap yang paling dihargai dari proses CRISP-DM. Perencanaan untuk *Deployment* dimulai selama *Business Understanding* dan harus menggabungkan tidak hanya bagaimana untuk menghasilkan nilai


model, tetapi juga bagaimana mengkonversi skor keputusan, dan bagaimana untuk menggabungkan keputusan dalam sistem operasional.


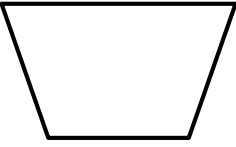
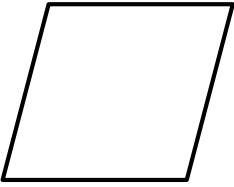


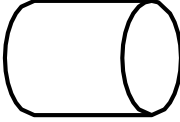

2.2.15 Flowchart

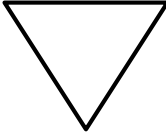
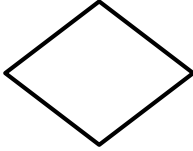
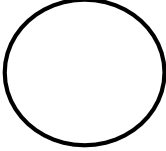
Flowchart adalah representasi visual dari alur kerja atau proses yang menunjukkan langkah-langkah yang harus diikuti untuk menyelesaikan tugas atau mencapai tujuan tertentu. Biasanya, flowchart menggunakan simbol-simbol geometris seperti persegi panjang, oval, dan panah untuk merepresentasikan langkah-langkah, keputusan, atau titik-titik penghubung dalam proses tersebut. Mereka sangat berguna dalam memvisualisasikan dan memahami aliran informasi atau keputusan dalam suatu sistem atau proses.

Dalam proyek pembuatan program yang melibatkan banyak orang sekaligus, flowchart sangat penting untuk menentukan langkah atau fungsionalitas. Selain itu, membuat diagram alur proses program akan membuatnya lebih mudah dipahami, lebih ringkas, dan mengurangi kemungkinan salah penafsiran. Dalam dunia pemrograman, flowchart adalah cara yang bagus untuk menghubungkan kebutuhan non-teknis dan teknis. Tujuan dari desain flowchart adalah untuk menunjukkan bagaimana proses klasifikasi tweet dilakukan menggunakan algoritma Naïve Bayes Classifier dan K-Nearest Neighbor, seperti yang ditunjukkan pada gambar Tabel 2.5 berikut:

Tabel 2. 5 Simbol Flowchart

Simbol	Nama	Keterangan
	Terminator	Simbol yang digunakan untuk menggambarkan awal atau akhir suatu proses dalam system.

	Proses	Simbol yang digunakan untuk menggambarkan proses komputerisasi dari suatu system.
	Proses Manual	Simbol yang digunakan untuk menggambarkan proses yang dikerjakan manual di dalam sistem.
	Catatan	Simbol yang digunakan untuk menggambarkan catatan data yang direkap sebelumnya didalam dokumen atau formular.
	Dokumen	Simbol ini menggambarkan sebuah dokumen input atau output yang terkait dalam suatu system.
	Multi Dokumen	Simbol ini menunjukkan beberapa dokumen lanjutan.
	Penyimpanan	Simbol ini digunakan media penyimpanan yang digunakan dalam suatu sistem.
	Garis Alir	Simbol yang digunakan untuk menjelaskan arah atau jalanya suatu proses dalam siste.

	Arsip	Simbol ini menunjukkan dokumen yang diarsipkan.
	Kondisi	Simbol ini digunakan untuk menunjukkan proses pengkondisian dari suatu proses kemudian keputusannya berupa Ya atau Tidak
	Penghubung	Simbol ini digunakan untuk penghubung ke halaman yang masih sama atau ke halaman lain

BAB III

METODOLOGI PENELITIAN

3.1 Data Penelitian

Kamus Besar Bahasa Indonesia (KBBI) mendefinisikan "data" sebagai suatu keterangan atau bahan nyata yang dapat digunakan sebagai dasar penelitian (analisis atau kesimpulan). Data juga disebut sebagai informasi yang dapat diproses oleh komputer, seperti representasi digital dari angka, teks, gambar grafis, atau suara.

Proses penelitian ilmiah bergantung pada data penelitian. Data penelitian harus akurat, dapat diandalkan, dan terkait dengan pertanyaan penelitian. Selain itu, data harus diorganisir, dianalisis, dan diinterpretasikan untuk menghasilkan informasi dan pengetahuan yang dapat digunakan untuk menjawab pertanyaan penelitian, menguji hipotesis, dan menemukan hubungan antara pola, tren, atau fenomena tertentu.

Data yang digunakan dalam penelitian merupakan data publik. Data mentah yang diperoleh berjumlah 5000 record dengan 2 atribut dapat dilihat pada tabel 3.1 sebagai berikut :

Tabel 3. 1 Deskripsi Nama Atribut

Atribut	Keterangan
Tweet	Komentar masyarakat/ <i>netizen</i> melalui media sosial twitter (X)
Sentimen	Sentimen positif dan negatif

Data tersebut didapatkan dari twitter menggunakan hashtag #KaburAjaDulu. Setelah tweet di temukan diperlukan API twitter untuk mendapatkan data tweets tersebut , lalu proses selanjut nya dibutuhkan aplikasi *rapidminer* untuk mendapatkan data tweet menggunakan *extention* API twitter yang ada di dalam aplikasi *rapidminer*, proses selanjutnya yang dilakukan pendeskripsian data, pemilihan atribut dan cleansing data. Setelah itu data di

export menjadi data csv agar memudahkan proses. Data yang digunakan dalam penelitian merupakan data kualitatif. Data kualitatif adalah data yang mencirikan atau menggambarkan sesuatu, data yang dapat diamati dan direkam. Data kualitatif dibagi menjadi 2 yaitu :

a. Data Primer

Data primer adalah data yang diperoleh langsung dari sumber aslinya sehingga peneliti bisa mengamati dan menulis jawaban langsung dari objek penelitian. Teknik pengumpulan data primer yaitu melakukan observasi langsung. Observasi adalah aktivitas yang dilakukan untuk mengamati secara langsung suatu objek tertentu dengan tujuan memperoleh sejumlah data dan informasi terkait objek tersebut. Dalam penelitian ini, penulis mengambil data berupa cuitan masyarakat dari sosial media (X).

b. Data Sekunder

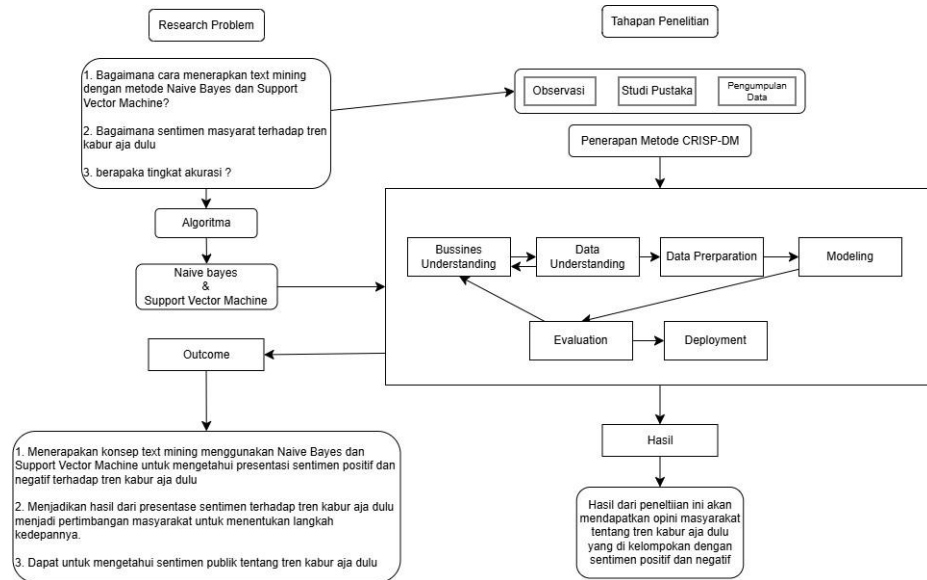
Data sekunder adalah data yang sudah ada atau data yang sudah diolah terlebih dahulu dan disusun berdasarkan pengaturan tertentu yang baru didapatkan peneliti dari sumber yang lain sebagai tambahan informasi. Teknik pengumpulan data sekunder terdiri dari 2 antara lain:

- 1) Studi pustaka adalah suatu metode pengumpulan data dari berbagai sumber tertulis seperti jurnal ilmiah, artikel, laporan, dan makalah konferensi.
- 2) Text book adalah suatu pengumpulan informasi yang merujuk pada sumber – sumber yang lebih umum seperti buku pelajaran, ensiklopedia, atau referensi standar yang digunakan untuk memperoleh pemahaman yang lebih luas tentang suatu objek.

3.2 Penerapan Metodologi

Pada penelitian ini metode yang digunakan adalah *Naïve Bayes*, dan *Support Vector Machine*. Data yang digunakan pada penelitian ini data tweet yang *search* menggunakan hastag #KaburAjaDulu. Pada tahapan penelitian ini akan dilakukan pengumpulan data awal maka proses selanjutnya akan dilakukan penarikan data, kemudian setelah itu dilakukan *Cleansing Data*, lalu pada tahap selanjutnya dilakukan Pemilihan Atribut, tahap berikutnya

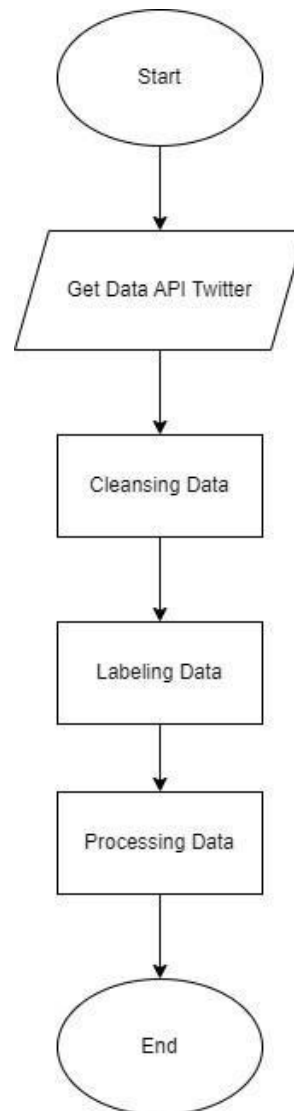
melakukan penerapan metode CRISP-DM yang pada proses penelitian ini mengacu pada enam tahap CRISP-DM. Berikut tahapan metodologi penelitian yang dapat dilihat pada Gambar 3.1 dibawah ini :



Gambar 3. 1 Tahapan Penerapan Metodologi

3.2.1 Proses Pengumpulan Data

Data tersebut diperoleh dari twitter (X) dengan menggunakan API twitter (X), Tweet yang diambil berupa tweet berbahasa Indonesia dengan kata kunci #KaburAjaDulu dengan limit 5000 dan berhasil di dapat berjumlah 3437 data Tweet yang telah selesai melalui tahap *text pre-processing*.



Gambar 3. 2 Flowchart Pengumpulan Data

Dalam gambar 3.1 diatas bisa dijelaskan penelitian ini meliputi 4 proses yang di lakukan yaitu :

- a. GET data API twitter

Objek pada deskripsi data penelitian ini adalah opini pengguna twitter (X) terhadap isu kebijakan pemerintah. Tweet yang diambil berupa tweet berbahasa Indonesia dengan kata kunci #KaburAJaDulu .

b. *Cleansing* data

Data tweet yang belum diolah melalui beberapa proses, yaitu: *cleaning, tokenisasi, transform cases, stopword, dan filtering*.

c. *Labeling* data

Data yang sudah melalui proses *cleansing* kemudian melakukan pelabelan secara manual. Pelabelan secara manual dilakukan pada seluruh tweet.

d. *Processing* data

Pada penelitian ini akan dilakukan pengujian menggunakan *confusion matrix* untuk melihat hasil pengujian data yang diperoleh dari tahapan *modelling* dengan menggunakan algoritma, *Naïve Bayes* dan *Support Vector Machine*. untuk menilai akurasi dari algoritma tersebut.

3.3 Perancangan Penelitian

Pada Pada tahap ini, metode Cross Industry Standard for Data Mining (CRISP-DM) digunakan. Ini adalah metode yang banyak digunakan para ahli untuk memecahkan masalah dan menggunakan model proses pengembangan data. Proses penelitian ini mencakup enam tahap CRISP-DM, yang terdiri dari:

3.3.1 Pemahaman Bisnis (*Business Understanding*)

Pemahaman bisnis dilakukan dengan menganalisa objek penelitian yaitu analisa sentimen publik terkait tren kabur aja dulu. Pada tahap ini, tujuan bisnis dari penelitian ini adalah untuk menentukan sentimen publik terkait tren kabur aja dulu. Untuk melakukan ini, API extension Twitter dari aplikasi rapid miner akan digunakan untuk mengambil data Twitter (X). Setelah mendapatkan data tersebut, fitur seperti mengganti RT, mengganti link, mengganti hastag, mengganti

mention, mengganti simbol, mengurangi duplikat, dan seterusnya. Setelah selesai, data tweet akan dihasilkan yang siap untuk digunakan untuk analisis klasifikasi menggunakan algoritme.

3.3.2 Pemahaman Data

Berdasarkan data yang dikumpulkan, langkah berikutnya dalam metode CRISP-DM adalah memahami kebutuhan data yang terkait dengan pencapaian tujuan untuk menentukan strategi yang efisien. Data yang diambil dari API Twitter berjumlah dua ribu tweet dengan dua atribut yang dipilih menjadi dua atribut tweet dan sentimen. Data dikumpulkan dari 10 Februari 2025 hingga 10 Juli 2025. Setelah data diperoleh, proses selanjutnya adalah mengganti dan membersihkan data.

3.3.3 Data Preparation

Semua proses yang diperlukan untuk membuat dataset yang akan digunakan dalam alat pemodelan—mulai dari data mentah yang berupa data tweet hingga proses data mining—termasuk persiapan data, yang merupakan tahap yang rumit dengan aktivitas pengolahan data. Tahap terakhir dari persiapan data adalah membuat dataset dari tweet publik, yang merupakan tahap persiapan data akhir untuk digunakan dalam pemodelan klasifikasi dalam proses data mining.. Ada 7 tahap yang akan dilakukan dalam mempersiapkan data tersebut yaitu sebagai berikut :

a. Pengumpulan Data

Dimulai dengan mengumpulkan data yang berasal dari twitter dengan kata kunci Kabur Aja Dulu.

b. Pembersihan data (*Cleansing data*)

Membersihkan data dari karakter khusus, tautan, tanda baca, dan elemen yang tidak diperlukan lainnya seperti menghapus stopwords

(kata-kata umum seperti “dan”, “atau”, “juga”) yang tidak memberikan banyak informasi tentang sentimen.

c. Tokenisasi

Pecah teks menjadi unit-unit yang lebih kecil, seperti kata atau frase. Ini disebut tokenisasi. Dalam bahasa Inggris, tokenisasi biasanya melibatkan pemisahan teks menggunakan spasi. Namun, dalam bahasa yang lebih kompleks, tokenisasi bisa lebih rumit.

d. *Stemming* atau lemmatisasi

Reduksi kata-kata menjadi bentuk dasarnya dapat membantu dalam analisis sentimen. Pilih antara stemming (pemangkasan akhiran kata) atau lemmatisasi (pengembalian kata ke bentuk dasarnya dengan mempertimbangkan konteks).

e. Menghilangkan *noise*

Hapus kata-kata yang tidak relevan, seperti angka, simbol, atau kata yang muncul secara berlebihan dalam teks.

f. Normalisasi

Normalisasi teks dapat melibatkan menghapus tanda baca atau karakter khusus tertentu, atau menggantikan kata kata atau tren umum dengan representasi standar (misalnya, mengganti “tidak bagus” dengan “buruk”).

g. *Labeling* sentimen

Labeling data dengan sentimen yang sesuai, seperti "positif" dan "negatif". Proses ini dapat dilakukan secara manual atau dengan menggunakan pendekatan otomatis, seperti menggunakan kamus sentimen atau model klasifikasi.

Setelah tahap data preparation selesai, proses selanjutnya yaitu dengan analisis sentimen menggunakan *Naïve Bayes* dan *Support Vector Machine*.

3.3.4 Modeling

Untuk tahap pemodelan, metode klasifikasi akan digunakan, yang menggunakan algoritma Naive Bayes dan Support Vector Machine. Metode klasifikasi akan membagi atribut menjadi dua kategori: tweet dan sentimen. Rapidminer adalah alat pemodelan yang digunakan.

3.3.5 Evaluation & Results

Pada titik ini, pengukuran ketepatan pemodelan akan dilakukan. Hasil evaluasi digunakan untuk menentukan apakah pemodelan yang digunakan tepat dan sesuai untuk penelitian ini, serta apakah sesuai dengan rencana awal penelitian. Berdasarkan hasil evaluasi, diputuskan apakah penelitian harus dilanjutkan atau diulang dari awal karena tidak sesuai dengan rencana awal.

.

.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Alat Penelitian

Dalam penelitian, membutuhkan alat untuk mendukung berjalannya penelitian. Alat penelitian yang digunakan yaitu berupa perangkat lunak (*software*) dan perangkat keras (*hardware*).

4.1.1 Perangkat Lunak (*Software*)

Berikut alat perangkat lunak (*Software*), versi, dan fungsinya dapat di lihat di table 4.1 :

Tabel 4. 1 Spesifikasi Software

No	Jenis	Versi	Fungsi
1	Microsoft Excel	2019	Untuk memfilter data dan menyimpan data.
2	Rapidminer	2025.1.0	Untuk melakukan Pengolahan Data

4.1.2 Perangkat Keras (*Hardware*)

Berikut alat perangkat keras (*Hardware*) dan spesifikasinya dapat di lihat di table 4.2 :

Tabel 4. 2 Spesifikasi Hardware

No	Jenis	Fungsi
1	Processor	Intel(R) Core(TM) i5- 8265U
2	RAM	12 GB
3	System Type	Windows
4	Operating System	Windows 11 Home Single

4.2 Implementasi dan Pengujian

Dalam penelitian ini, Naive Bayes dan Support Vector Machine (SVM) adalah algoritma yang digunakan. Data yang digunakan berasal dari tweet yang diambil dari aplikasi media sosial Twitter (X) menggunakan pemrograman python melalui Google Colab. Setelah data dikumpulkan, mereka dimasukkan ke dalam proses penerapan CRISP-DM, yang mencakup pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penyebaran. Untuk mendapatkan nilai akurasi klasifikasi, alat Rapidminer digunakan untuk menguji algoritma Naive Bayes dan Support Vector Machine (SVM). Metode Cross Industry Standard Process for Data Mining (CRISP-DM), yang terdiri dari 6 tahapan, digunakan untuk memproses data ini:

4.2.1 Pemahaman Bisnis (Business Understanding)

Dinamika sentiment masyarakat tentang tren kabur aja dulu dapat mempengaruhi berbagai aspek kehidupan masyarakat, termasuk industri, ekonomi, dan lain-lain. Meskipun umumnya dianggap sangat negatif, ada banyak orang yang melihat tren kabur aja dulu sebagai hal yang baik. Oleh karena itu, hanya tren kabur yang akan diperhatikan.

4.2.2 Pemahaman Data (Data Understanding)

Pada tahap ini, peneliti mencoba memahami data yang akan digunakan dari Twitter (X). Data ini telah diambil menggunakan pemrograman python, terdiri dari 2067 tweet dengan dua atribut, masing-masing sentimen dan tweet. Data dikumpulkan dari 10 Februari 2025 hingga 10 Juli 2025.

4.2.3 Data Preparation

Pada tahap data preparation dilakukan beberapa tahapan antara lain sebagai berikut :

- a. Pengumpulan Data

Mengambil data menggunakan pemrograman python dengan bantuan Google Colab. Prosesnya termasuk memulai instalasi paket python, mendapatkan token Twitter (X), mengcrawling data berdasarkan teks dan periode tanggal, dan kemudian menyimpan data. Semua ini ditunjukkan pada gambar 4.1 di bawah ini.

```
[ ] # Import required Python package
!pip install pandas

# Install Node.js (because tweet-harvest built using Node.js)
!sudo apt-get update
!sudo apt-get install -y ca-certificates curl gnupg
!sudo mkdir -p /etc/apt/keyrings
!curl -fsSL https://deb.nodesource.com/gpgkey/nodesource-repo.gpg.key | sudo gpg --dearmor -o /etc/apt/keyrings/nodesource.gpg

NODE_MAJOR=20 && echo "deb [signed-by=/etc/apt/keyrings/nodesource.gpg] https://deb.nodesource.com/node_$NODE_MAJOR.x nodistro main" | sudo tee /etc/apt/sources.list.d/nodesource.list

!sudo apt-get update
!sudo apt-get install nodejs -y

node -v
```

Gambar 4. 1 Instalasi Package Python

Setelah melalui package python dapat terinstall dengan baik, selanjutnya mengambil token dari akun twitter (X), lalu setelahnya memasukan token seperti gambar 4.2 dan gambar 4.3 dibawah ini

Name	Value	Domain	Path	Expires...	Size
Manifest					
Service workers					
Storage					
Local storage					
Session storage					
Extension storage					
IndexedDB					
Cookies					
https://x.com					
Private state tokens					
Interest groups					
Shared storage					
Cache storage					
Storage buckets					
Manifest					
Service workers					
Storage					
Local storage					
Session storage					
Extension storage					
IndexedDB					
Cookies					
https://x.com					
Private state tokens					
Interest groups					
Shared storage					
Cache storage					
Storage buckets					
Name	Value	Domain	Path	Expires...	Size
_cf_bm	MJHPidfrQ93uq\$SNA_9LkXrOBb7pKBsIAQJ.amH...	.x.com	/	2025-0...	177
_cuid	9e144d250e864e84a0d20bbb7dd56bef	.x.com	/	2026-0...	38
att	1-SPOvAsW04z2DlmlEuM9pmoL0JC2B3n9YbTd...	.x.com	/	2025-0...	45
auth_token	951d435738f00971895da70ef0cd62496da4a77c	.x.com	/	2026-0...	50
ct0	ddce7ef49cf15143229bf604d8a2ce639bc993c761...	.x.com	/	2026-0...	163
g_state	["i_*"]0	.x.com	/	2026-0...	16
gt	1952692745505821136	.x.com	/	2025-0...	21
guest_id	v1%3A175439320444410871	.x.com	/	2026-0...	31
guest_id_ads	v1%3A175439320444410871	.x.com	/	2026-0...	35
guest_id_marketing	v1%3A175439320444410871	.x.com	/	2026-0...	41
IDE	AHWqTUko9050gg2mtGy8yhiq6-eLiWt-rDnyYg...	.doubl...	/	2026-0...	70
kdt	9tjp45iZR7xCl2qfzkl13DZkZNRHFM81u2TO3EFx	.x.com	/	2026-0...	43
lang	id	.x.com	/	Session	6
personalization_id	"v1_20cF/RtSt3JE8nyW6Lfx0A=="	.x.com	/	2026-0...	47
twid	u%3D1242143978025136128	.x.com	/	2026-0...	27

Gambar 4. 2 Pengambilan Token Dari X

```
#@title Twitter Auth Token

twitter_auth_token = '951d435738f00971895da70ef0cd62496da4a77c' # change this auth token
```

Gambar 4. 3 Memasukan Token

Setelah memasukan token dan dapat terautentifikasi dengan baik maka selanjutnya tinggal memasukan keyword kaburajadulu, limit yang dicari yaitu 2000 data dengan bentuk text, dan periode yang dicari yaitu dari 10 Februari 2025 sampai 10 Juli 2025 seperti pada gambar 4.4 dibawah.

```
# Crawl Data
filename = 'kabur.csv'
search_keyword = 'kaburajadulu since:2025-02-10 until:2025-07-10 lang:id'
limit = 2000

lnpx -y tweet-harvest@2.6.1 -o '{filename}' -s '{search_keyword}' --tab 'LATEST' -l {limit} --token {twitter_auth_token}
```

Gambar 4. 4 Proses Crawling Data

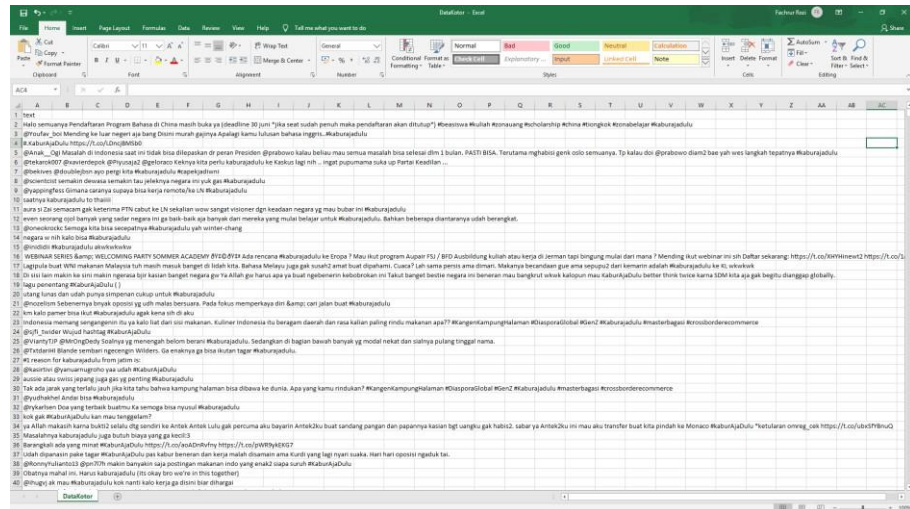
Setelah proses crawling data, file yang sudah siap tinggal di save dan di export sebagaimana yang ada pada gambar 4.5

```
import pandas as pd
# Specify the path to your CSV file
file_path = "tweets-data/{filename}"
# Read the CSV file into a pandas DataFrame
df = pd.read_csv(file_path, delimiter=',')
# Display the DataFrame
display(df)
```

	conversation_id_str	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_screen_name	lang	location	quote_count
0	1939652416976396524	Mon Jun 30 23:13:54 +0000 2025	0	@kegibgunfaedh Mentiri aja ga jatin jatin...	1939624748802046295	NaN	kegibgunfaedh	in	NaN	0
1	1939818301800292757	Mon Jun 30 22:48:17 +0000 2025	0	Nah yang mau #kaburajadulu di Italia.	1939818301800292757	NaN	NaN	in	NaN	0
2	1939733694501286063	Mon Jun 30 17:12:05 +0000 2025	0	soal pph 25 ni maah aja joo kewajiban pajak ...	1939733694501286063	NaN	NaN	in	NaN	0
3	1939649798601263852	Mon Jun 30 17:05:07 +0000 2025	1	@euljoo_zip 2026 KITA #kaburajadulu	1939731939201807894	NaN	euljoo_zip	in	NaN	0
4	193966726541765652	Mon Jun 30 13:53:16 +0000 2025	2	@tempodotco Mengajak mahasiswa ke luar negeri	1939713859172094175	NaN	tempodotco	in	NaN	0
...
107	1938666023240282456	Sat Jun 28 12:58:26 +0000 2025	1	@secqret_la saman logo	1938945122974278546	NaN	secqret_la	in	NaN	0

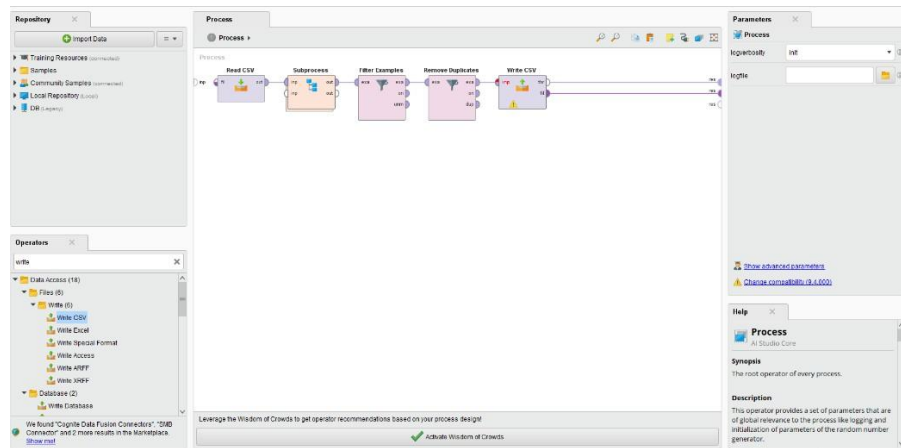
Gambar 4. 5 Proses Export Data

Setelah dapat di export data bisa didapatkan dengan jumlah 2067 data. Untuk data yang didapat masuk kedalam format csv atau Microsoft Excel seperti pada gambar 4.6.



Gambar 4. 6 Jumlah Data yang Didapat

b. **Cleansing Data:** Pada tahap ini, peneliti akan membersihkan data dengan menggunakan rapidminer sebelum memberi label pada sentimen. Ini akan memastikan bahwa dataset pada file csv terbaru tidak memiliki tanda dan data yang tidak diperlukan. Seperti yang ditunjukkan pada Gambar 4.7, proses ini akan menghilangkan berbagai gangguan, seperti menghilangkan URL, akun, retweet, digit angka, dan karakter..



Gambar 4. 7 Model Process Cleansing

Setelah selesai menjalankan proses Cleansing dengan menggunakan rapidminer. Awalnya terdapat 2067 data tweet pada file csv, kemudian setelah melewati proses Cleansing maka data tweet

berkurang menjadi 385 data seperti yang digambarkan pada gambar 4.8 dibawah ini.

Row No.	Tweets
1	Halo semuanya Pendaftaran Program Bahasa di China masih buka ya deadline 30 juni jika sead sudah penuh maka pendaftaran akan ditutup
2	Youfav_bcl Mending ke luar negeri aja bang Disini murah pajinya Apalagi kamu kulusan bahasa inggris
3	Anak_Ogi Masalah di Indonesia saat ini tidak bisa dilepaskan dr peran Presiden Prabowo kalau beliau mau semua masalah bisa selesai dim 1...
4	tekarok07 xaverdepok Piyusa2 gelaraco Kelnnya kita perlu kaburjadulu ke Kustus lagi nih ingat pupumama suka up Partai Keadilan
5	belkives doublejbn ayo pergi kita
6	scientist2 semakin dewasa semakin tau jeleknya negara ini yuk gas
7	yappingfess Gimana caranya supaya bisa kerja remoteke LN
8	saatnya kaburjadulu to thaili
9	aura si Zai semacam gak diterima PTN cabut ke LN sekalian wow sangat visioner dgn keadaan negara yg mau bubar ini
10	even seorang ojol banyak yang sadar negara ini ga baikbaik aja banyak dari mereka yang mulai belajar untuk
11	onekrocc: Semoga kita bisa secepatnya
12	negara w nih kalo bisa
13	imdidli
14	Lagipula buat WNI makanan Malaysia tuh masih masuk banget di lidah kita Bahasa Melayu juga gak susah2 amat buat dipahami Cucca Lah sa...
15	Di sisi lain makin ke sini makin ngerasa bjr kसान banget negara gw Ya Allah gw harus apa ya buat ngebenerin kebrokan ini Takut banget bes...
16	lagu penentang
17	utang lunas dan udah punya simpanan cukup untuk
18	km kalo pamer bisa ikut
19	Indonesia memang sengangenin itu ya kalo liat dari sisi makanan Kuliner Indonesia itu beragam daerah dan rasa kalian paling rindu makanan a...
20	sfl_widur Wujud hashtag
21	Vian7JP MrOngDedy Soalnya yg menengah belum berani
22	Tdkrt4t Rlbande samharhi nneraninn Wllfere Ca anakrus na Nics lrdfan banar

Gambar 4. 8 Hasil Cleansing Data

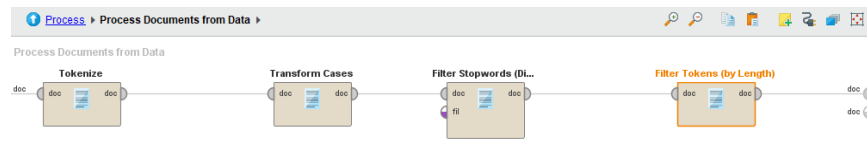
- c. Labeling Kemudian tahap selanjutnya adalah melakukan klasifikasi atau penentuan sentimen berdasarkan masing-masing tweet. Selama proses penentuan sentimen berlangsung, disini peneliti melakukan pemberian sentimen secara manual. Berikut ini data yang sudah di beri label seperti pada gambar 4.9.

Row	Tweet	Sentiment
1	1. Tweet	neutral
2	2. Mending ke luar negeri aja bang Disini murah pajinya Apalagi kamu kulusan bahasa inggris	positive
3	3. Masalah di Indonesia saat ini tidak bisa dilepaskan dr peran Presiden Prabowo kalau beliau mau semua masalah bisa selesai dim 1 bulan PARTI BISA Turutma ngubuhin gerak solo semuanya Tp kalau diri diam2 bisa yk wsi lahngah tepatna	positive
4	4. Kelnnya kita perlu kaburjadulu ke Kustus lagi nih ingat pupumama suka up Partai Keadilan	positive
5	5. ayo pergi kita	positive
6	6. belkives doublejbn ayo pergi kita	positive
7	7. scientist2 semakin dewasa semakin tau jeleknya negara ini yuk gas	positive
8	8. yappingfess Gimana caranya supaya bisa kerja remoteke LN	positive
9	9. saatnya kaburjadulu to thaili	positive
10	10. even seorang ojol banyak yang sadar negara ini ga baikbaik aja banyak dari mereka yang mulai belajar untuk	positive
11	11. onekrocc: Semoga kita bisa secepatnya	positive
12	12. negara w nih kalo bisa	positive
13	13. imdidli	positive
14	14. Lagipula buat WNI makanan Malaysia tuh masih masuk banget di lidah kita Bahasa Melayu juga gak susah2 amat buat dipahami Cucca Lah sa...	positive
15	15. Di sisi lain makin ke sini makin ngerasa bjr kसान banget negara gw Ya Allah gw harus apa ya buat ngebenerin kebrokan ini Takut banget bes...	positive
16	16. lagu penentang	positive
17	17. utang lunas dan udah punya simpanan cukup untuk	positive
18	18. km kalo pamer bisa ikut	positive
19	19. Indonesia memang sengangenin itu ya kalo liat dari sisi makanan Kuliner Indonesia itu beragam daerah dan rasa kalian paling rindu makanan a...	positive
20	20. sfl_widur Wujud hashtag	positive
21	21. Vian7JP MrOngDedy Soalnya yg menengah belum berani	positive
22	22. Tdkrt4t Rlbande samharhi nneraninn Wllfere Ca anakrus na Nics lrdfan banar	positive
23	23. ...	neutral
24	24. ...	neutral
25	25. ...	neutral
26	26. ...	neutral
27	27. ...	neutral
28	28. ...	neutral
29	29. ...	neutral
30	30. ...	neutral
31	31. ...	neutral
32	32. ...	neutral
33	33. ...	neutral
34	34. ...	neutral
35	35. ...	neutral
36	36. ...	neutral
37	37. ...	neutral
38	38. ...	neutral
39	39. ...	neutral
40	40. ...	neutral

Gambar 4. 9 Data yang Sudah Dilabeli

- d. Tokenizing Pada proses ini adalah untuk mengubah teks dari kalimat kalimat yang kompleks menjadi urutan token atau unit-unit yang

lebih sederhana. Tujuannya untuk mempersiapkan data teks agar dapat diproses oleh algoritma analisis sentimen dengan lebih efektif dan akurat. Untuk proses tokenizing, Hasil Tokenizing Naïve bayes dan hasil Tokenizing Support Vector Machine (SVM) dapat dilihat sebagaimana yang ada pada gambar 4.10, gambar 4.11 serta gambar 4.12 dibawah ini.



Gambar 4. 10 Model Tokenizing Data

Open in Turbo Prep Auto Model Interactive Analysis Filter (385 / 385 examples): all

Row No.	labeling	prediction[la...	confidence[...	confidence[...	aamin	abai	abis	abroad	absolutely	adilan
1	positif	positif	1	0	0	0	0	0	0	0
2	positif	positif	1	0	0	0	0	0	0	0
3	positif	positif	1	0	0	0	0	0	0	0
4	positif	negatif	0	1	0	0	0	0	0	0
5	positif	positif	1	0	0	0	0	0	0	0
6	positif	positif	1	0	0	0	0	0	0	0
7	positif	positif	1	0	0	0	0	0	0	0
8	positif	positif	1	0	0	0	0	0	0	0
9	positif	positif	1	0	0	0	0	0	0	0
10	negatif	negatif	0	1	0	0	0	0	0	0
11	positif	positif	1	0	0	0	0	0	0	0
12	negatif	negatif	0	1	0	0	0	0	0	0
13	positif	positif	1	0	0	0	0	0	0	0
14	negatif	negatif	0	1	0	0	0	0	0	0
15	negatif	negatif	0	1	0	0	0	0	0	0
16	positif	positif	1	0	0	0	0	0	0	0
17	positif	positif	1	0	0	0	0	0	0	0
18	positif	positif	1	0	0	0	0	0	0	0
19	positif	positif	1	0	0	0	0	0	0	0
20	positif	positif	1	0	0	0	0	0	0	0
21	positif	positif	1	0	0	0	0	0	0	0
22	positif	positif	1	0	0	0	0	0	0	0

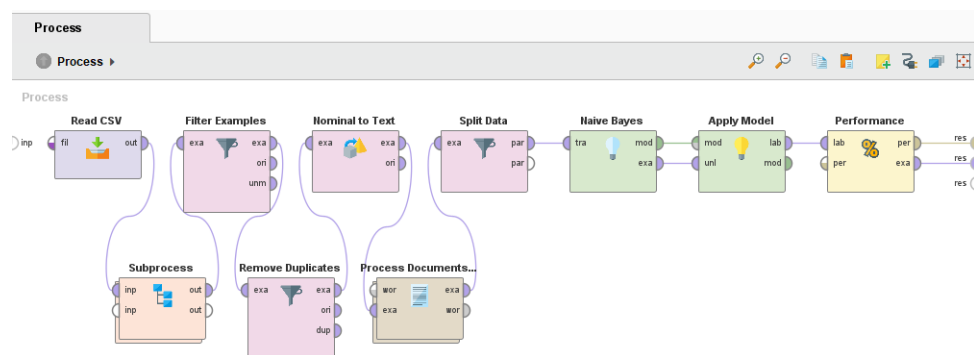
Gambar 4. 11 Model Tokenizing Naïve Bayes

Row No.	labeling	prediction(la...	confidence...	confidence...	aamiin	abai	abis	abroad	absolutely	adlan
1	positif	positif	0.749	0.251	0	0	0	0	0	0
2	positif	positif	0.749	0.251	0	0	0	0	0	0
3	positif	positif	0.750	0.250	0	0	0	0	0	0
4	positif	positif	0.749	0.251	0	0	0	0	0	0
5	positif	positif	0.749	0.251	0	0	0	0	0	0
6	positif	positif	0.749	0.251	0	0	0	0	0	0
7	positif	positif	0.749	0.251	0	0	0	0	0	0
8	positif	positif	0.749	0.251	0	0	0	0	0	0
9	positif	positif	0.750	0.250	0	0	0	0	0	0
10	negatif	positif	0.612	0.388	0	0	0	0	0	0
11	positif	positif	0.749	0.251	0	0	0	0	0	0
12	negatif	negatif	0.402	0.598	0	0	0	0	0	0
13	positif	positif	0.749	0.251	0	0	0	0	0	0
14	negatif	negatif	0.287	0.713	0	0	0	0	0	0
15	negatif	positif	0.676	0.324	0	0	0	0	0	0
16	positif	positif	0.748	0.252	0	0	0	0	0	0
17	positif	positif	0.749	0.251	0	0	0	0	0	0
18	positif	positif	0.749	0.251	0	0	0	0	0	0
19	positif	positif	0.749	0.251	0	0	0	0	0	0
20	positif	positif	0.749	0.251	0	0	0	0	0	0
21	positif	positif	0.749	0.251	0	0	0	0	0	0
22	positif	positif	0.749	0.251	0	0	0	0	0	0

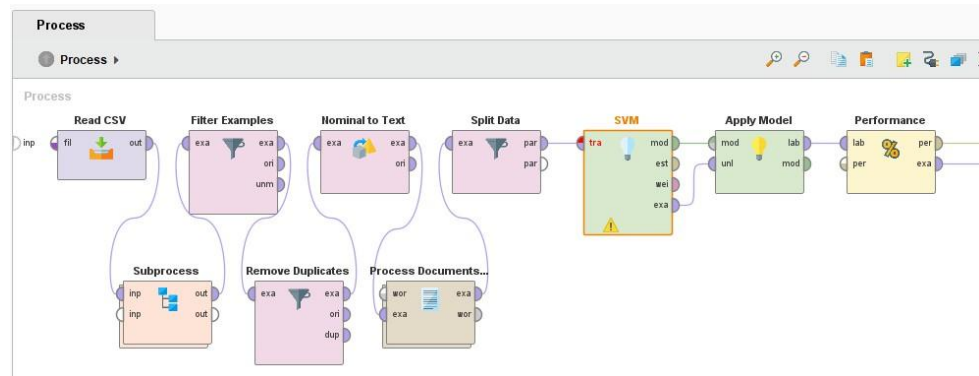
Gambar 4. 12 Model Tokenizing Support Vector Machine

4.2.4 Modeling

Pada tahap ini, peneliti akan melakukan preprocessing pada dataset untuk menyiapkan data yang bersih dan bebas suara. Selain itu, proses ini juga akan menghitung pembobotan kata untuk kebutuhan proses modelling di masa mendatang. Setelah melewati tahap persiapan teks, jumlah data tweet menjadi 385 data, yang merupakan data bersih untuk digunakan pada langkah berikutnya. Pada tahapan ini akan dilakukan pengukuran performa dua klasifikasi sekaligus dengan menggunakan algoritma Naïve bayes seperti pada gambar 4.13 dan Support Vector Machine (SVM) seperti pada gambar 4.14.



Gambar 4. 13 Model Proses Data Naïve Bayes



Gambar 4. 14 Model Proses Data Support Vector Machine

4.2.5 Evaluation & Result

Pada penelitian ini akan dilakukan pengujian menggunakan algoritma Naïve bayes dan K-Nearest Neighbors dengan hasil sebagai berikut :

a. Algoritma Naïve Bayes

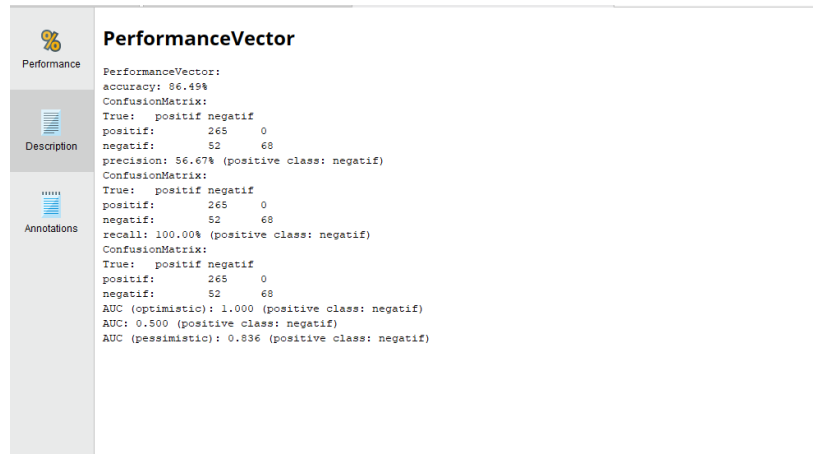
Jumlah data yang dikumpulkan adalah 385 data. Hasil dari tahapan model dengan menggunakan algoritma Naive Bayes berikut ini: Data yang benar diprediksi positif memiliki presentase ketepatan 100%, data dengan sentimen negatif memiliki presentase ketepatan 56,67%, data dengan sentimen positif memiliki presentase ketepatan (spesifisitas) 83,60%, dan data dengan sentimen negatif memiliki presentase ketepatan 56,67% .Nilai accuracy yang dihasilkan menggunakan model Naïve bayes adalah 86,49% yang dapat dilihat pada gambar 4.15 dibawah ini.

accuracy: 86.49%

	true positif	true negatif	class precision
pred. positif	265	0	100.00%
pred. negatif	52	68	56.67%
class recall	83.60%	100.00%	

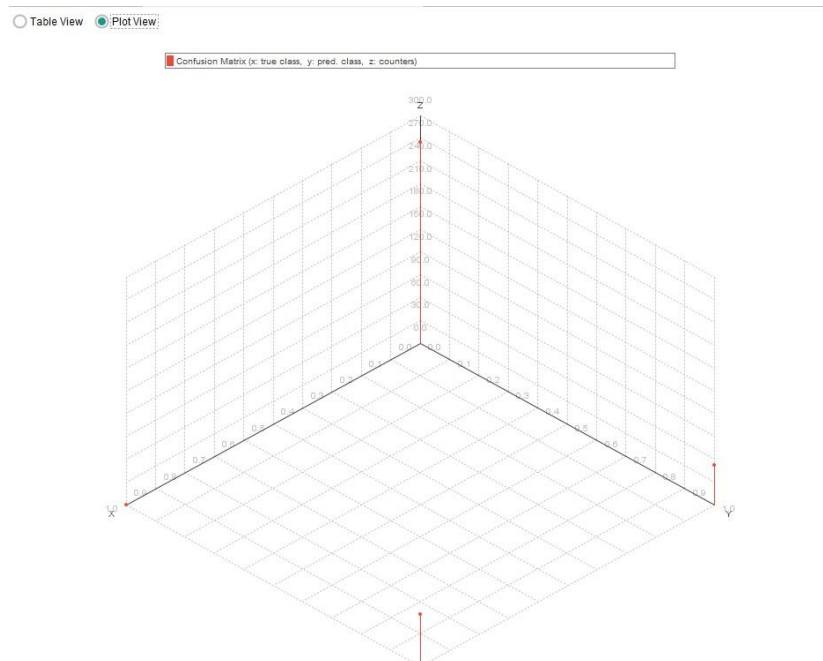
Gambar 4. 15 Hasil Akurasi Naïve bayes

Dari hasil berikut dideskripsikan olah aplikasi rapidminer seperti yang ada pada gambar 4.16.



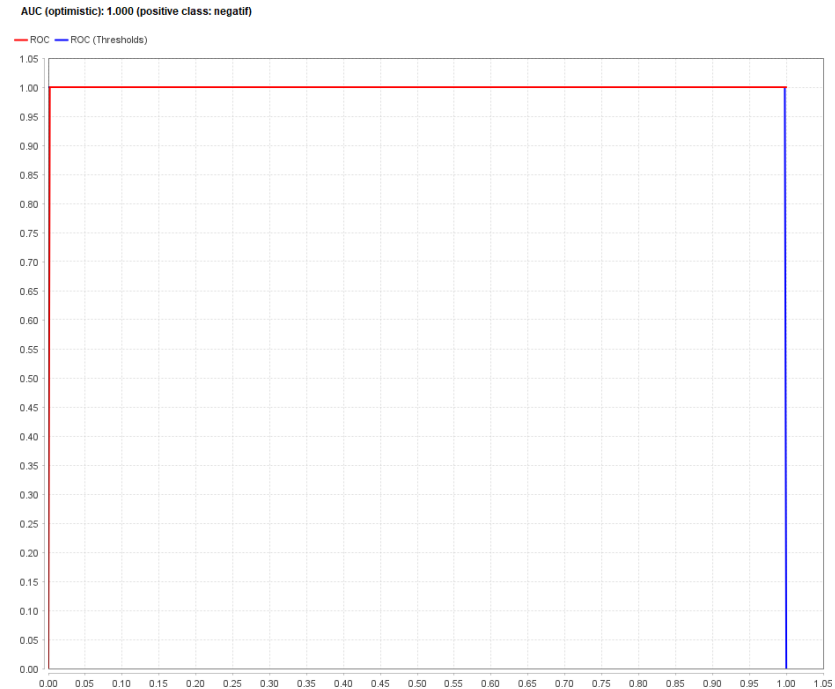
Gambar 4. 16 Deskripsi Akurasi Naïve bayes

Dari hasil tersebut tergambar plot view untuk algoritma Naïve bayes seperti yang digambarkan pada gambar 4.17 dibawah ini



Gambar 4. 17 Plot View Naïve bayes

Dari hasil berikut tergambar grafik AUC (Optimistic) untuk algoritma Naïve bayes seperti yang tergambar pada gambar 4.18 dibawah ini.



Gambar 4. 18 Grafik AUC (Optimistic) Naïve Bayes

Berdasarkan hasil diatas dapat disimpulkan bahwa, algoritma Naïve bayes dapat mengklasifikasi sentiment cukup menggunakan data tweet X kabur aja dulu.

b. Algoritma Support Vector Machine (SVM)

Total dataset yang dikumpulkan adalah 385 data. Berikut ini adalah hasil dari tahapan modelling dengan menggunakan algoritma Support Vector Machine dilihat hasil dari perhitungan hasil accuracy menunjukkan data yang benar diprediksi positif yang menghasilkan presentase ketepatannya adalah 87,33%, sedangkan untuk data bersentimen Negatif memiliki precision sebesar 100%. Untuk sentimen Positif memiliki recall (Specificity) adalah 100%, sedangkan pada sentimen negatif memiliki recall sebesar 32,36%. Nilai accuracy yang dihasilkan

menggunakan model Naïve bayes adalah 88,05% yang dapat dilihat pada gambar 4.19 dibawah ini.

accuracy: 88.05%

	true positif	true negatif	class precision
pred. positif	317	46	87.33%
pred. negatif	0	22	100.00%
class recall	100.00%	32.35%	

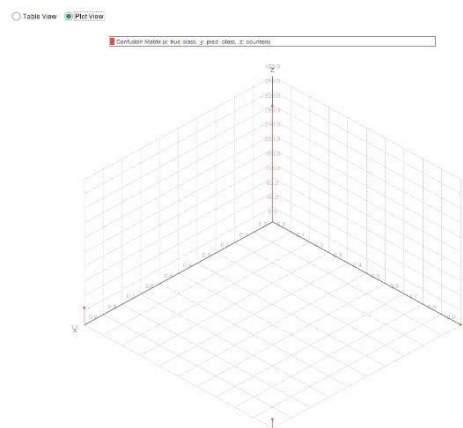
Gambar 4. 19 Hasil Akurasi Support Vector Machine

Dari hasil berikut dideskripsikan olah aplikasi rapidminer seperti yang ada pada gambar 4.20.



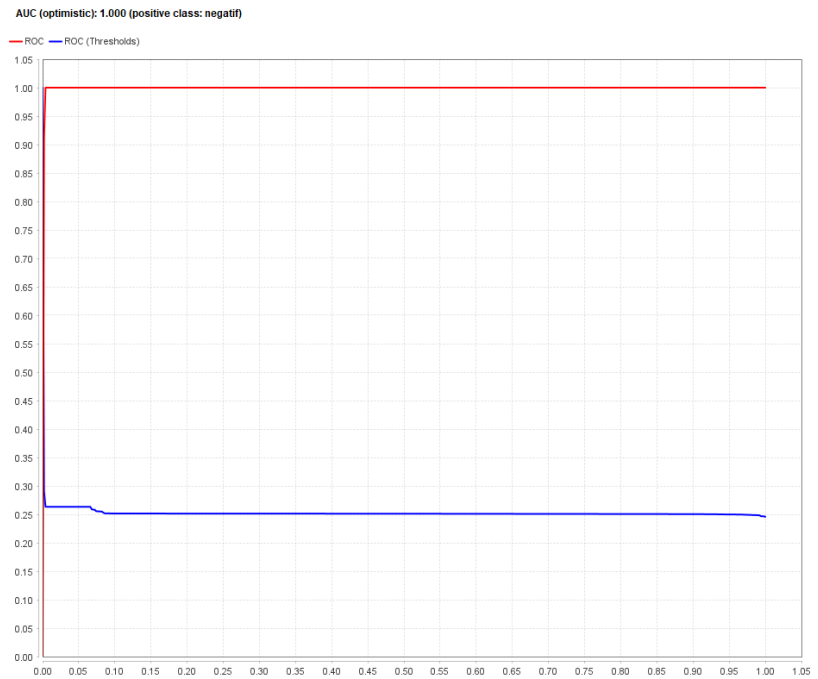
Gambar 4. 20 Deskripsi Akurasi Support Vector Machine

Dari hasil tersebut tergambar plot view untuk algoritma Naïve bayes seperti yang digambarkan pada gambar 4.21 dibawah ini



Gambar 4. 21 Plot View Support Vector Machine

Dari hasil berikut tergambar grafik AUC (Optimistic) untuk algoritma Naïve bayes seperti yang tergambar pada gambar 4.22 dibawah ini.



Gambar 4. 22 Grafik AUC (Optimistic) Support Vector Machine

Berdasarkan hasil diatas dapat disimpulkan bahwa, algoritma Support Vector Machine dapat mengklasifikasi sentiment cukup menggunakan data tweet X kabur aja dulu.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian dan penerapan algoritma Naïve Bayes dan Support Vector Machine dalam analisis sentimen terhadap Tren Kabur aja dulu dapat disimpulkan bahwa:

- a. Menurut penelitian yang telah dilakukan untuk mengklasifikasikan pendapat masyarakat Indonesia tentang tren kabur aja dulu, ditemukan bahwa dari 385 data, 317, atau 82 %, menunjukkan sentimen positif dan 68 data, atau 18 %, menunjukkan sentimen negatif.
- b. Hasil penelitian menunjukkan bahwa masyarakat Indonesia sebagian besar menunjukkan respon positif terhadap tren kabur aja dulu. Berdasarkan klasifikasi dataset tren kabur aja dulu dengan model algoritma Naive Bayes, tingkat akurasi 86,49% diperoleh, sedangkan tingkat akurasi 88,05% diperoleh dengan algoritma Support Vector Machine.
- c. Sehingga dapat dikatakan bahwa algoritma Support Vecotr Machine lebih baik dalam mengklasifikasikan data sentimen tentang tren kabur saja dulu daripada algoritma Naive Bayes.

5.2 Saran

Adapun saran-saran yang disampaikan berdasarkan hasil pengamatan dan analisa selama melakukan penelitian ini adalah:

- a. Untuk penelitian selanjutnya dapat mengembangkan penelitian ini dan menghubungkan dengan algoritma lain sehingga dapat menyentuh akurasi yang lebih bagus lagi.
- b. Dari hasil penelitian ini belum sangat lengkap dan efisien, diharapkan ada yang mengembangkan penelitian ini dengan menggunakan atau menambahkan algoritma lain agar dalam menentukan suatu objek lebih cepat dan akurat.
- c. Dari penelitian ini, diharapkan untuk bisa menentukan labelling sentiment secara otomatis agar dapat mempercepat menentukan sentimen.

DAFTAR PUSTAKA

- [1] Oryza Habibie Rahman, Gunawan Abdillah, and Agus Komarudin, “Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 17–23, Feb. 2021, doi: 10.29207/resti.v5i1.2700.
- [2] Purnamasari Detty *et al.*, *Pengantar Metode Analisis Sentimen*. Depok: Gunadarma, 2023.
- [3] R. F. P. Pratama and W. Maharani, “Comparative Analysis of Naive Bayes and SVM for Improved Emotion Classification on Social Media,” *Edumatic: Jurnal Pendidikan Informatika*, vol. 9, no. 1, pp. 11–20, Apr. 2025, doi: 10.29408/edumatic.v9i1.29087.
- [4] Adhitya Karel Maulaya and Junadhi, “Analisis Sentimen Menggunakan Support Vector Machine Masyarakat Indonesia Di Twitter Terkait Bjorka,” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 3, no. 3, pp. 495–500, Dec. 2022, doi: 10.37859/coscitech.v3i3.4358.
- [5] A. Safira, A. S. Masyarakat...v, and F. N. Hasan, “ANALISIS SENTIMEN MASYARAKAT TERHADAP PAYLATER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER,” *Jurnal Sistem Informasi*, vol. 5, no. 1, 2023.
- [6] K. Zuhri, N. Adha, and O. Saputri, “Analisis Sentimen Masyarakat Terhadap Pilpres 2019 Berdasarkan Opini Dari Twitter Menggunakan Metode Naive Bayes Classifier,” 2020. [Online]. Available: <https://journal-computing.org/index.php/journal-cisa/index>
- [7] K. A. Baihaqi *et al.*, “A Comparison Support Vector Machine, Logistic Regression And Naïve Bayes For Classification Sentimen Analisis user Mobile App,” *International Journal of Artificial Intelligence Research*, vol. 7, no. 1, p. 64, Jun. 2023, doi: 10.29099/ijair.v7i1.962.
- [8] I. And and D. Expert, “Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM) INFORMASI ARTIKEL ABSTRAK,” 2022. [Online]. Available: <https://e-journal.unper.ac.id/index.php/informatics>
- [9] Z. Rahman Hakim and S. Sugiyono, “Analisa Sentimen Terhadap Kereta Cepat Jakarta – Bandung Menggunakan Algoritma Naïve Bayes Dan K-Nearest Neighbor,” *Jurnal Sains dan Teknologi*, vol. 5, no. 3, pp. 939–945, Apr. 2024, doi: 10.55338/saintek.v5i3.1423.

- [10] S. A. Azzahra and A. Wibowo, “ANALISIS SENTIMEN MULTI-ASPEK BERBASIS KONVERSI IKON EMOSI DENGAN ALGORITME NAÏVE BAYES UNTUK ULASAN WISATA KULINER PADA WEB TRIPADVISOR,” vol. 7, no. 4, 2020, doi: 10.25126/jtiik.202071907.
- [11] M. Nazar Yuniar, “Klasifikasi Kualitas Air Bersih Menggunakan Metode Naïve baiyes,” *Jurnal Sains dan Teknologi*, vol. 5, no. 1, pp. 243–246, 2023, doi: 10.55338/saintek.v5i1.1383.
- [12] T. Safitri, Y. Umaidah, and I. Maulana, “Analisis Sentimen Pengguna Twitter Terhadap BTS Menggunakan Algoritma Support Vector Machine,” 2023. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [13] H. Taufiqqurrahman, F. Tri Anggraeny, and M. Muharrom Al Haromainy, “PERBANDINGAN ALGORITMA NAÏVE BAYES DAN K-NEAREST NEIGHBOR PADA ANALISIS SENTIMEN ULASAN APLIKASI MYPERTAMINA,” 2023.
- [14] F. R. B. Kahi, A. Talakua, and R. Reynaldi, “Analisis Sentimen Masyarakat Di Twitter Terhadap Pemerintahan Anies Baswedan Menggunakan Metode Naive Bayes Classifier,” *Jurnal Minfo Polgan*, vol. 13, no. 1, pp. 324–336, Apr. 2024, doi: 10.33395/jmp.v13i1.13636.
- [15] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, “Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes,” *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 4, pp. 411–423, Aug. 2022, doi: 10.55123/insologi.v1i4.770.
- [16] J. Alfiah Zulqornain and P. Pandu Adikara, “Analisis Sentimen Tanggapan Masyarakat Aplikasi Tiktok Menggunakan Metode Naïve Bayes dan Categorical Propotional Difference (CPD),” 2021. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [17] A. Sitanggang, Y. Umaidah, Y. Umaidah, R. I. Adam, and R. I. Adam, “ANALISIS SENTIMEN MASYARAKAT TERHADAP PROGRAM MAKAN SIANG GRATIS PADA MEDIA SOSIAL X MENGGUNAKAN ALGORITMA NAÏVE BAYES,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4902.
- [18] F. S. Pamungkas and I. Kharisudin, “Analisis Sentimen dengan SVM,” vol. 4, pp. 628–634, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [19] “ANALISIS SENTIMEN TERHADAP PEMINDAHAN IBU KOTA NEGARA MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER DAN K-NEAREST NEIGHTBORS.”

- [20] F. Naufal Rahman, S. Lestari, and S. Tinggi Ilmu Komputer Cipta Karya Informatika Jakarta, “ANALISIS SENTIMEN MASYARAKAT TERHADAP PEMERINTAH DI ERA KABINET JOKO WIDODO BERDASARKAN SOSIAL MEDIA X MENGGUNAKAN NAÏVE BAYES DAN K-NEAREST NEIGHBOR (KNN) ANALYSIS OF PUBLIC SENTIMENT TOWARDS THE GOVERNMENT IN THE JOKO WIDODO CABINET ERA BASED ON SOCIAL MEDIA X USING NAÏVE BAYES AND K-NEAREST NEIGHBOR (KNN),” *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 7, no. 5, 2024.
- [21] M. H. Al-Areef and K. Saputra, “Nomor 2,” *Agustus*, vol. 22, pp. 270–279, 2023, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jis/index>
- [22] P. Studi and T. Informatika, “ANALISA SENTIMEN MASYARAKAT INDONESIA TERHADAP KINERJA PRABOWO SUBIANTO SEBAGAI ANGGOTA KABINET INDONESIA MAJU DI PLATFORM TWITTER MENGGUNAKAN R DENGAN ALGORITMA NAÏVE BAYES PENULIS Syamsu Hidayat ABSTRAK,” vol. 2, no. 1, pp. 1–10, 2021.
- [23] Asrumi, Suharijadi Didik, Setiari Dewi Agustina, and Putri Wulanda Diah, *ANALISIS SENTIMEN DAN PENGALIAN OPINI*. Jawa Tengah: EUREKA MEDIA AKSARA, 2023.
- [24] Amna *et al.*, *DATA MINING*. Padang: PT GLOBAL EKSEKUTIF TEKNOLOGI , 2023.
- [25] S. P. M. Si. Ari Mulyoto, *NAIVE BAYES PADA GOOGLE COLABS*. Jawa Tengah: EUREKA MEDIA AKSARA, 2024.
- [26] S. Kom. , M. Pd. Iwan Mahendro and S. Hum. , M. P. Dhanan Abimanto, *ALGORITMA SUPPORT VECTOR MACHINE UNTUK MENGANALISA KEPUASAN MAHASISWA TERHADAP PEMBELAJARAN E-LEARNING*. Semarang: CV. Pustaka STIMART-AMNI , 2023.
- [27] M. Jelita, “Text Mining dengan Topic Modelling LDA dari Pertanyaan Gelar Wicara Literasi Perpustakaan Nasional RI,” *Media Pustakawan*, vol. 31, no. 3, pp. 253–265, Dec. 2023, doi: 10.37014/medpus.v31i3.5237.
- [28] G. M. G. Bororing and F. Gunawan, “Sistem Pemesanan dan Pembayaran Makanan berbasis Web Terintegrasi dengan Application Programming Interface (API),” *Jurnal Informatika dan Bisnis*, vol. 13, no. 1, pp. 37–48, Jun. 2024, doi: 10.46806/jib.v13i1.1147.
- [29] K. Zuhri, N. Adha, and O. Saputri, “Analisis Sentimen Masyarakat Terhadap Pilpres 2019 Berdasarkan Opini Dari Twitter Menggunakan

Metode Naive Bayes Classifier,” 2020. [Online]. Available: <https://journal-computing.org/index.php/journal-cisa/index>

LAMPIRAN

24% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text
- Small Matches (less than 12 words)

Top Sources

- 21%  Internet sources
- 14%  Publications
- 8%  Submitted works (Student Papers)



Fachrur Rozi

Jl.Cibubur III, No.11 Rt.009 🏠
Rw.001. Cibubur Jak-Tim
0857-7862-9502 📞
rozifachrur511@gmail.com ✉️

Saya seorang lulusan baru yang memiliki ketertarikan bekerja di bidang Design grafis, Fotografi dan juga Editing Foto/Video. Terbiasa menggunakan berbagai aplikasi Editing seperti Adobe Illustrator, Adobe Photoshop, Adobe Premier Pro, dan Adobe After Effect. Jujur, disiplin dan mampu bekerja dengan baik dalam tim. Pernah aktif dalam kegiatan ekstrakurikuler dan pengalaman project seperti Pembuatan Event di UKM Kampus.

Pengalaman Kerja

FEBRUARI 2019 – APRIL 2019

Design Grafis (Praktek Kerja Lapangan)_

Radio Silaturahmi

- Mengelola Content untuk keperluan Media Social seperti Feeds Instagram.
- Membuat Flyer untuk keperluan Event Perusahaan.
- Mengambil Foto Dokumentasi Event Perusahaan.
- Mengedit hasil Foto Maupun Video.

MEI 2020 – OKTOBER 2020

Layouter

Lettrb studio

- Memvisualisasikan Storyboard yang sudah di buat sesuai arahan.
- Mengembangkan Aset desain.

JULI 2022 – AGUSTUS 2022

Media Team Internship

STEGGO.id

- Mengelola Content untuk keperluan Media Social seperti Feeds Instagram.
- Membuat Keperluan Event seperti Mock Up Stand.
- Mengambil dan Mengedit foto Produk.
- Mengambil dan Mengedit Video Commercial.

FEBRUARI 2024 – JULI 2024

Labeler

Nexter AI

- Mengubah Foto menjadi Data berbentuk 2D dan 3D dengan teknologi LiDar

JULI 2024 – SEKARANG

Reviewer

Nexter AI

- Mempreview hasil pekerjaan labeler
- Memberikan feedback terhadap kerjaan labeler
- Memberikan pelatihan kepada labeler

Pengalaman Projek & Organisasi

FEBRUARY 2021 – MEI 2021

Ketua UKM/ UKM Desain ST.Ilmu Komputer CKI

- Memimpin penyelenggaraan acara UKM Desain ST.Ilmu Komputer CKI selama periode 2021 – 2022
- Mengkoordinir Pertemuan UKM Desain dengan frekuensi 2 kali sebulan dan bertindak sebagai narahubung pemateri.

- Menyusun rangkaian kegiatan UKM dan materi tiap pertemuan selama setahun agar program kerja terselenggara dengan baik.

Education

2017 - 2020

SMK Negeri 51 Jakarta / Jurusan Multimedia

2021 - 2025

St. Ilmu Komputer CKI / Jurusan S1 Sistem Informasi

Skills

Desain (Adobe Illustrator, Corel Draw, Adobe Photoshop), Editing Foto (Adobe Lightroom), dan Editing Video (Adobe Premier Pro dan Adobe After Effect), JavaScript, Codeigniter 4.