



**ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP INSTITUSI POLRI
PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE SUPPORT
VECTOR MACHINE DAN NAÏVE BAYES**

SKIRPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat untuk menyelesaikan
Program Studi Informatika (S1)

Disusun Oleh :

Ranis Lailatus Lisa

172410103028

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

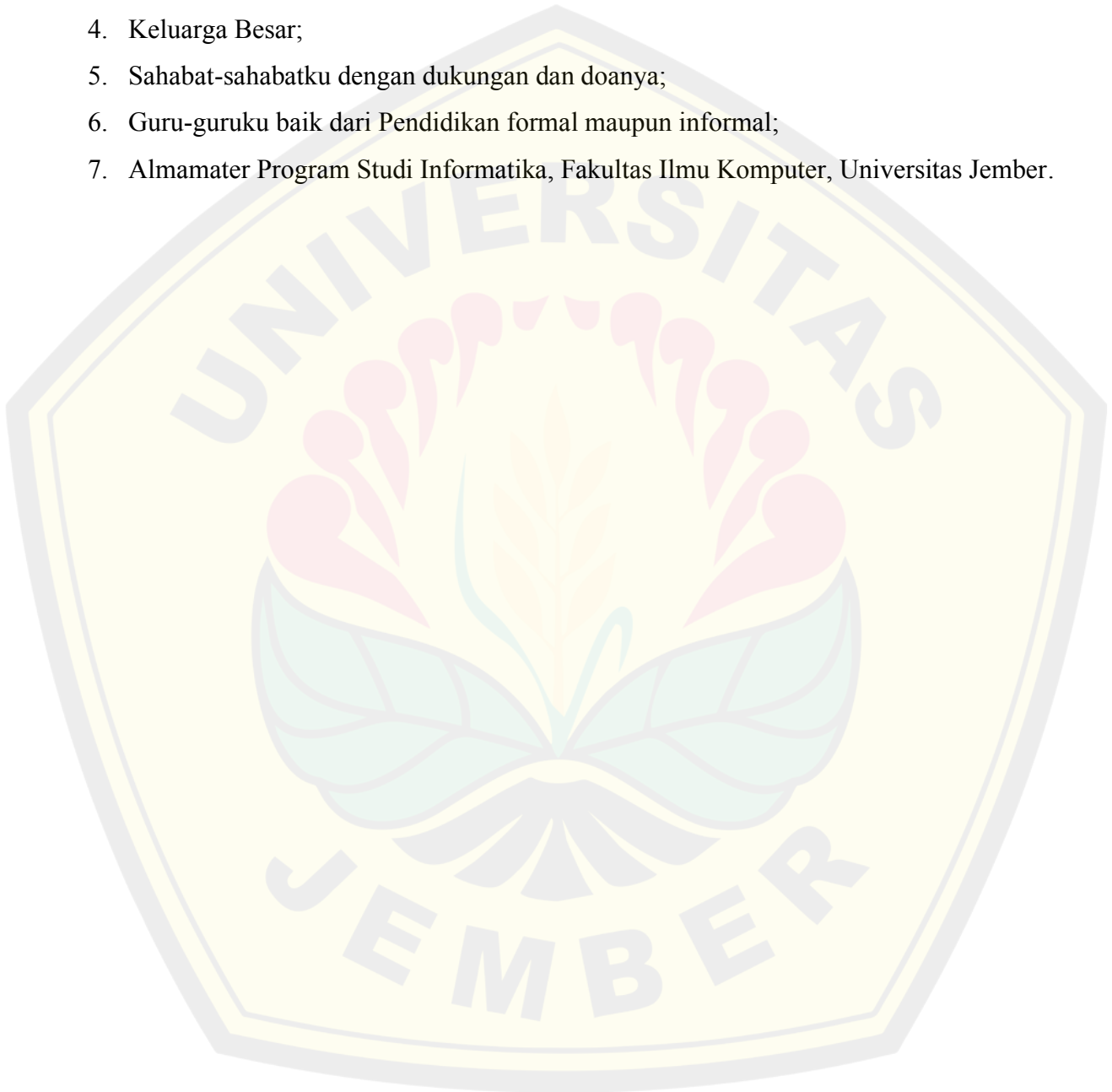
UNIVERSITAS JEMBER

2023

PERSEMBAHAN

Skripsi ini saya persembahkan untuk :

1. Allah SWT yang senantiasa memberikan rahmat dan hidayah-Nya untuk mempermudah dan melancarkan dalam mengerjakan skripsi;
2. Ibunda Siti Aisah;
3. Suami Tercinta Muhammad Prayudi;
4. Keluarga Besar;
5. Sahabat-sahabatku dengan dukungan dan doanya;
6. Guru-guruku baik dari Pendidikan formal maupun informal;
7. Almamater Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Jember.



MOTTO

“Karena sesungguhnya sesudah kesulitan ada kemudahan. Kemudahan ini Allah berikan untuk hamba-Nya sebagai rahmat yang tidak terhingga nilainya.”

-QS. Al Insyirah : 5-6-



PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Ranis Lailatus Lisa

NIM : 172410103028

Menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP INSTITUSI POLRI PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN NAÏVE BAYES ”, adalah benar-benar hasil karya sendiri, kecuali jika dalam pengutipan substansi disebutkan sumbernya, belum pernah diajukan pada institusi mana pun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa adanya tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika kemudian hari pernyataan ini tidak benar.

Jember, 14 Oktober 2022

Yang menyatakan,

Ranis Lailatus Lisa

SKRIPSI

**ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP INSTITUSI POLRI
PADA MEDIA SOSIAL TWITTER MENGGUNAKAN MTODE SUPPORT VECTOR
MACHINE DAN NAÏVE BAYES**

Oleh

Ranis Lailatus Lisa

172410103028

Pembimbing

Dosen Pembimbing Utama : Yanuar Nurdiansyah ST.,M.Cs

Dosen Pembimbing Pendamping : Priza Pandunata S.Kom.,M.Sc

PENGESAHAN PEMBIMBING

Skripsi berjudul "Analisis Sentimen Opini Masyarakat Terhadap Institusi Polri Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine dan Naïve Bayes" karya Ranis Lailatus Lisa Telah diuji dan disahkan pada:

hari, tanggal : Kamis, 19 Januari 2023

tempat : Fakultas Ilmu Komputer

Disetujui Oleh:

Pembimbing I,



Yanuar Nurdiansyah, ST., M.Cs

NIP 198201012010121004

Pembimbing II,



Priza Pandunata ST, M.Sc.

NIP 19830131201504001

PENGESAHAN PENGUJI

Skripsi berjudul “Analisis Sentimen Opini Masyarakat Terhadap Institusi Polri Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine dan Naïve Bayes” karya Ranis Lailatus Lisa Telah diuji dan disahkan pada:

hari, tanggal : Kamis, 19 Januari 2023

tempat : Fakultas Ilmu Komputer Universitas Jember

Disetujui Oleh:

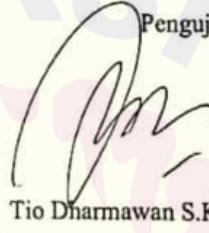
Penguji I,



Nelly Oktavia A, S.Si., MT.

NIP 198410242009122008

Penguji II,



Tio Dharmawan S.Kom., M.Kom

NIP 199111122022031000

Mengesahkan

Dekan Fakultas Ilmu Komputer



Drs. Antonius Cahya P., M.App.Sc., Ph.D

NIP 196909281993021001

RINGKASAN

Analisis Sentimen Opini Masyarakat Terhadap Institusi Polri Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine dan Naïve Bayes; Ranis Lailatus Lisa, 172410103028; 2022, 54 halaman; Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Jember.

Analisis sentimen merupakan proses mengestrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif dan negatif. Metode tersebut membantu untuk mempelajari sentimen dari berbagai macam konten-konten media sosial dalam hal berupa tweet tentang institusi kepolisian negara republik Indonesia.

Support Vector Machine merupakan teknik *supervised learning*, mempunyai tingkat akurasi dan kualitas yang baik. Akan tetapi, untuk implementasinya diperlukan tahap pelatihan *sequential training* dan harus melalui proses pengujian. Kelebihan metode Support Vector Machine dapat mengidentifikasi *hyperplane* yang terpisah sehingga bisa memaksimalkan margin dari kelas yang berbeda. Kekurangan dari metode ini yaitu pada masalah yang mempunyai fitur yang sama dapat memengaruhi tingkat akurasi secara signifikan. Dari dataset sejumlah 1100 didapatkan hasil akurasi antara 85% sampai dengan 86% dengan hasil nilai akurasi dari keempat percobaan pengujian sebesar 90%:10%, 80%:20%, 70%:30%, dan 60%:40%.

Naïve Bayes merupakan algoritma pengelompokan, yang sumbernya berasal dari teorema Bayes, yang prosesnya terdiri dari komputasi nilai probabilitas untuk data teks, dan yang dapat memproses data dalam jumlah besar dengan tingkat akurasi yang tinggi. Dari dataset sejumlah 1100 didapatkan hasil akurasi antara 79% sampai dengan 85% dengan hasil nilai akurasi dari keempat percobaan pengujian sebesar 90%:10%, 80%:20%, 70%:30%, dan 60%:40%.

PRAKATA

Puji syukur kehadiran Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul “Analisis Sentimen Opini Masyarakat Terhadap Institusi Polri Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine Dan Naïve Bayes”. Skripsi ini disusun untuk memenuhi salah satu syarat menyelesaikan pendidikan Strata Satu (S1) pada Program Studi Informatika Fakultas Ilmu Komputer Universitas Jember.

Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Allah SWT yang senantiasa memberikan rahmad dan hidayah-Nya untuk mempermudah dan melancarkan dalam mengerjakan skripsi;
2. Yanuar Nurdiansyah, ST., M.Cs. selaku Dosen Pembimbing Utama dan Priza Pandunata, S.Kom., M.Sc. selaku Dosen Pembimbing Pendamping yang telah meluangkan waktu, tenaga, dan perhatian dalam pembimbingan skripsi;
3. Seluruh Bapak dan Ibu dosen beserta staff karyawan di Fakultas Ilmu Komputer Universitas Jember.
4. Almarhumah Ibu Siti Aisah yang selalu memberikan kasih sayang serta dukungan dan motivasi selama ini;
5. Suami Tercinta Muhammad Prayudi yang selalu memberikan kasih sayang serta dukungan, bantuan dan membiayai segala kebutuhan penulis dalam hal apapun;
6. Teman-teman Informatika angkatan 2017 yang memberikan dukungan dan bantuan selama masa perkuliahan.

Penulis juga menyadari bahwa laporan penelitian skripsi ini masih jauh dari kata sempurna. Oleh karena itu, segala saran dan kritik akan sangat berarti bagi penulis. Penulis juga berharap dengan adanya penelitian skripsi ini dapat bermanfaat bagi semua pihak.

Jember, 29 November 2022

Penulis

DAFTAR ISI

PERSEMBAHAN	ii
MOTTO.....	iii
SKRIPSI.....	v
PENGESAHAN PEMBIMBING.....	Error! Bookmark not defined.
PENGESAHAN PENGUJI.....	Error! Bookmark not defined.
RINGKASAN.....	viii
PRAKATA	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiii
BAB 1. PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian.....	2
1.4 Manfaat Penelitian.....	3
1.5 Batasan Masalah.....	3
BAB 2. TINJAUAN PUSTAKA.....	4
2.1 Penelitian Terdahulu.....	4
2.2 Analisis Sentimen.....	5
2.3 Text mining.....	5
2.3.1 Case Folding.....	6
2.3.2 Cleansing.....	6
2.3.3 Stemming.....	6
2.3.4 Tokenizing.....	6
2.3.5 Stopword Removal.....	6
2.3.6 Normalization.....	6
2.4 Pembobotan Kata.....	6
2.5 Support Vector Machine.....	7
2.6 Naïve Bayes.....	7
BAB 3. METODELOGI PENELITIAN.....	10
3.1 Jenis Penelitian.....	10
3.2 Objek Penelitian.....	10
3.3 Tahapan Penelitian.....	10

3.3.1	Pengumpulan Data	11
3.3.2	Pelabelan Data.....	11
3.3.3	Pre-processing Data.....	11
3.3.4	Analisis Sentimen.....	12
BAB 4. HASIL DAN PEMBAHASAN		14
4.1	Hasil Dataset	14
4.2	Hasil Pelabelan.....	15
4.3	Hasil Preprocessing.....	15
4.4	Hasil Bag of Word	21
4.5	Hasil pembobotan TF-IDF	22
4.6	Hasil Pemodelan	26
4.6.1	Hasil Pemodelan Support Vector Machine.....	26
4.6.2	Hasil Pemodelan Naïve Bayes.....	27
4.7	Hasil Uji Evaluasi	27
4.8	Visualisasi Word Cloud	32
BAB 5. KESIMPULAN DAN SARAN		36
5.1	Kesimpulan	36
5.2	Saran	36
Daftar Pustaka		38

DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	10
Gambar 4.1 Kode Twitter API.....	14
Gambar 4.2 Kode Twitter API Key	14
Gambar 4.3 Kode Casefolding.....	15
Gambar 4.4 Kode Remove punctutation.....	16
Gambar 4.5 Kode Tokenizing.....	17
Gambar 4.6 Kode Normalisasi.....	18
Gambar 4.7 Kode Stopwords 1	19
Gambar 4.8 Kode Stopword 2	19
Gambar 4.9 Kode Stemming.....	20
Gambar 4.10 Kode Bag Of Words.....	21
Gambar 4.11 Kode TF-IDF.....	22
Gambar 4.12 Kode Split Data.....	26
Gambar 4.13 Kode SVM	27
Gambar 4.14 Kode Splitting Data.....	27
Gambar 4.15 Kode Naïve Bayes.....	27
Gambar 4.16 Kode Confusion Matrix	28
Gambar 4.17 Confusion Matrix SVM	28
Gambar 4.18 Confusion Matrix Naïve Bayes.....	29
Gambar 4.19 Kode Perhitungan Presisi, Recall, Fscore dan Akurasi.....	30
Gambar 4.20 Hasil Percobaan 1 SVM.....	30
Gambar 4.21 Hasil Percobaan 1 Naïve Bayes	31
Gambar 4.22 Kode program word cloud	32
Gambar 4.23 word cloud sentimen positif (naïve bayes)	33
Gambar 4.24 word cloud sentimen negatif (naïve bayes).....	33
Gambar 4.25 word cloud sentimen positif (SVM).....	34
Gambar 4.26 word cloud sentimen negatif (SVM).....	35

DAFTAR TABEL

Tabel 4.1 Hasil Crawling Data Twitter.....	15
Tabel 4.2 Hasil Pelabelan	15
Tabel 4.3 Hasil Casefolding.....	16
Tabel 4.4 Hasil Cleansing	17
Tabel 4.5 Hasil Tokenizing.....	17
Tabel 4.6 Hasil Normalisasi.....	18
Tabel 4.7 Hasil List Stopwords.....	19
Tabel 4.8 Hasil Stopwords.....	20
Tabel 4.9 Hasil Stemming.....	21
Tabel 4.10 Contoh Hasil Perhitungan TF	23
Tabel 4.11 Contoh Hasil Nilai DF dan IDF Setiap Kata	24
Tabel 4.12 Hasil TF-IDF Tiap Kata.....	25
Tabel 4.13 Hasil Data Uji Evaluasi SVM.....	30
Tabel 4.14 Hasil Data Uji Evaluasi Naïve Bayes	31
Tabel 4.15 Kata dan contoh kalimat positif dalam dataset	33
Tabel 4.16 Kata dan contoh kalimat negatif dalam dataset	34
Tabel 4.17 Kata dan contoh kalimat positif dalam dataset	35

BAB 1. PENDAHULUAN**1.1 Latar Belakang**

Kepolisian Negara Republik Indonesia yang disingkat POLRI merupakan kepolisian nasional di Indonesia, yang bertanggung jawab langsung di bawah Presiden. Sebelum organisasi ini bernama Badan Kepolisian Negara. Polisi Republik Indonesia memiliki moto Rastra Sewakotama yang merupakan abdi utama bagi Nusa Bangsa. Polisi Republik Indonesia juga mengemban tugas-tugas kepolisian diseluruh wilayah Indonesia yaitu memelihara keamanan serta ketertiban rakyat, tertib dan tegaknya hukum, terselenggaranya perlindungan, pengayoman serta pelayanan kepada masyarakat, dan terbinanya ketentraman rakyat dengan menjunjung tinggi hak azasi manusia (Wahyu Daeni, 2018).

Saat ini, dapat menyaksikan bagaimana media jejaring sosial dipakai oleh masyarakat umum dalam menggencarkan opini rakyat terhadap suatu proses politik atau kasus yang tengah terjadi di Indonesia. Contoh isu-isu sosial dan politik di Indonesia yang mendapat banyak respons melalui media sosial seperti Twitter antara lain viral aksi polisi polsek saat ingin diwawancara tidak beretika menyuruh wartawan untuk bicara dengan pohon sehingga menuai kontroversi di media sosial serta yang paling trending di media sosial twitter yaitu kasus pembunuhan Brigadir J dan Ferdi Sambo yang terlibat dalam kasus pembunuhan tersebut hingga akan dipecat secara tidak hormat oleh presiden dan dugaan pelecehan seksual yang sedang ramai diperbincangkan. Kasus-kasus lain yaitu oknum polisi berduaan dengan istri TNI di Jombang digrebek warga saat suami berduaan ke luar kota. Di luar Indonesia, banyak pihak yang juga memanfaatkan media sosial dalam berbagi pesan ke seluruh dunia yang mengakibatkan terbangunnya opini publik untuk mendukung tujuan tertentu.

Adanya sejumlah kasus yang melibatkan anggota kepolisian negara republik Indonesia menjadi perhatian publik dalam sebulan terakhir ini. Hal tersebut tentu berpengaruh terhadap masyarakat dan menimbulkan opini-opini yang dikemukakan di media sosial terutama pada media sosial twitter.

Sosial media merupakan tempat yang dikenal untuk mencurahkan segala hal oleh semua masyarakat. Segala bentuk pendapat dan pandangan yang dikemukakan oleh masyarakat disampaikan dengan bebas di sosial media dapat menimbulkan

dampak yang berbeda-beda di masyarakat tergantung bagaimana masing-masing individu menyikapi pendapat tersebut (Frizka, Ema and Hanif, 2021).

Analisis sentimen merupakan proses mengestrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif dan negatif. Metode tersebut membantu untuk mempelajari sentimen dari berbagai macam konten-konten media sosial dalam hal berupa tweet tentang institusi kepolisian negara republik indonesia. Penelitian ini berfokus pada twitter sebagai objek analisis sentimen karena pertumbuhan aktif twitter di Indonesia memiliki grafik yang baik dengan total 15,7 juta pengguna sehingga Indonesia menduduki peringkat ke-6 pada 2021 (Statista, 2021). Hal ini menjadikan twitter sebagai objek yang sesuai untuk analisis sentiment tweet opini masyarakat mengenai institusi polri. Penelitian ini menggunakan metode Support Vector Machine untuk proses analisis yang akan dimulai dengan mengubah data teks yang ada menjadi data vektor, kelebihan metode Support Vector Machine adalah kemampuannya mengidentifikasi *hyperplane* yang terpisah sehingga bisa memaksimalkan margin dari kelas yang berbeda (Brian Laurensz and Eko Sedyono, 2021). Metode yang ke 2 yaitu Naïve Bayes untuk garis dasar dalam tugas-tugas yang berhubungan dengan teks dan jumlah kumpulan data. Penggunaan kedua metode tersebut bertujuan untuk membandingkan performa dan tingkat akurasi yang dihasilkan dari analisis sentimen kedua metode tersebut. Penelitian ini diharapkan dapat memberikan kontribusi bagi pemerintah maupun masyarakat untuk mengetahui tentang dampak berbagai cuitan negatif atau positif yang bebas dikemukakan di media sosial twitter.

1.2 Rumusan Masalah

Berdasarkan permasalahan yang telah diuraikan pada latar belakang, maka dapat diambil rumusan masalah sebagai berikut :

1. Bagaimana proses analisis sentimen menggunakan metode *Support Vector Machine* dan *Naïve Bayes* terhadap opini masyarakat tentang institusi polri?
2. Berapa perbandingan akurasi yang didapatkan dari hasil penggunaan metode *Support Vector Machine* dan *Naïve Bayes* dalam menganalisis sentiment opini masyarakat terhadap institusi polri?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditetapkan, sehingga dapat ditetapkan tujuan penelitian untuk menjawab rumusan masalah sebagai berikut :

1. Mengetahui proses analisis sentimen menggunakan metode *Support Vector Machine* dan *Naïve Bayes* terhadap opini masyarakat tentang institusi polri.
2. Mengetahui nilai akurasi hasil penggunaa metode *Support Vector Machine* dan *Naïve Bayes* dalam menganalisis sentimen opini masyarakat terhadap institusi polri pada twitter.

1.4 Manfaat Penelitian

Manfaat yang disampaikan dari dampak yang ingin dihasilkan pada penelitian ini sebagai berikut :

1. Bagi pemerintah dan Pengguna Twitter
Penelitian ini dapat dijadikan sebagai sarana untuk mengetahui tentang konten dan akun yang bersifat negatif dan positif yang berhubungan dengan konten institusi kepolisian negara republik Indonesia.
2. Bagi Penulis
Penelitian ini dapat memberikan wawasan serta pengembangan keilmuan bagi penulis dalam bidang Metode Klasifikasi

1.5 Batasan Masalah

Terdapat beberapa batasan masalah yang dapat diangkat sebagai parameter dalam melakukan penelitian, sebagai berikut :

1. Penelitian berfokus pada tweet pengguna twitter berbahasa Indonesia dengan hashtag #institusipolri dan tanpa menganalisa gambar, video, dan pranala
2. Tweet yang terkumpul akan dikelompokkan kedalam dua analisis sentimen positif dan negatif
3. Data yang diambil menggunakan *Aplication Programming Interface(API)* *Twitter* dengan API key penulis
4. Melakukan penjadwalan untuk *Crawling Data* di twitter pada tanggal 29 Agustus 2022 dan data yang digunakan sebanyak 1100 tweet
5. Algoritma yang digunakan yaitu *support vector machine* dan *naïve bayes* serta membandingkan keduanya

BAB 2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Penulis mendapatkan referensi dari beberapa laporan penelitian yang sudah dilakukan sebelumnya oleh orang lain sebagai bahan referensi dalam melakukan penelitian ini dan sekaligus membandingkan hasilnya.

Taufik Kurniawan (2017) melakukan penelitian yang berjudul Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Media Mainstream Menggunakan Naïve Bayes Classifier Dan Support Vector Machine memaparkan bahwa proses pengklasifikasian sentiment analisis menggunakan metode Naïve Bayes diperoleh nilai akurasi yang berbeda, pada media tv one sebesar 95,8 % dan pada Kompas tv sebesar 97,8%, sedangkan hasil ketetapan dari klasifikasi yang menggunakan metode support vector machine menghasilkan akurasi yang berbeda pada media tv one sebesar 97,9% sedangkan pada media Kompas tv sebesar 99,3%. Secara keseluruhan perbandingan performa metode Naïve Bayes dan Support Vector Machine menunjukkan bahwa hasil dari performa Support Vector Machine lebih baik dan lebih akurat dalam mengklasifikasikan data dari pada metode Naïve Bayes.

Hennie Tuhuteru, Ade Iriani (2018) melakukan penelitian yang berjudul Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine Dan Naïve Bayes melakukan perbandingan performa kedua metode tersebut untuk mengetahui mana yang memiliki akurasi yang lebih baik. Hasil yang diperoleh pada penelitian ini menunjukkan rata-rata tingkat akurasi metode klasifikasi Support Vector Machine lebih baik dari pada metode Naïve Bayes Classifier, yaitu sebesar 76.42%.

Nurdiansyah, Bukhori dan Hidayat (Nurdiansyah et al., 2018) dalam judul “Sentiment Analysis System for Movie review in Bahasa Indonesia Using Naïve Bayes Classifier Method”. Penelitian ini dilakukan untuk membuat sebuah sistem yang dapat mengklasifikasikan sentimen dari dokumen review menjadi dua kelas sentimen positif dan sentimen negatif. Penelitian ini menggunakan naïve bayes classifier. Peneliti memilih movienthusiast, ulasan film dalam website Bahasa Indonesia sebagai sumber dokumen review. Hasil yang dikumpulkan 201 ulasan film: 783 ulasan positif dan 418 ulasan negatif yang digunakan dataset untuk machine learning classifier. Akurasi pengklasifikasian menghasilkan rata-rata 88,379% dari lima kali upaya pengukuran akurasi menggunakan kumpulan data yang disebutkan sebelumnya.

Berdasarkan beberapa penelitian tersebut dapat disimpulkan bahwa metode Support Vector Machine dapat diterapkan pada proses klasifikasi analisis sentiment dengan hasil akurasi yang cukup baik dibanding menggunakan metode Naïve Bayes. Adapun penelitian ini dilakukan untuk menetapkan polaritas opini masyarakat terhadap kasus-kasus institusi kepolisian negara republik Indonesia. Perbedaan yang ada antara penelitian terdahulu seperti telah dijelaskan sebelumnya dengan penelitian ini adalah proses untuk melakukan Normalization pada tahapan preprocessing data dengan menggunakan metode Support Vector Machine dan Naïve Bayes. Pengaplikasian klasifikasi menggunakan beberapa skenario data untuk mengetahui pengaruh berbagai skenario data terhadap akurasi dari model yang dihasilkan.

2.2 Analisis Sentimen

Analisis sentimen merupakan salah satu bidang studi atau metode yang fungsinya menganalisis opini masyarakat, sentiment, evaluasi, penilaian, perilaku dan emosi melalui entitas seperti produk, layanan umum, organisasi, individual, isu, kejadian dan topik (Akhmad and Arita, 2021). Analisis opini, atau disebut juga opinion mining, adalah bidang pertimbangan yang menganalisis kesimpulan, perkiraan, evaluasi, penilaian, sikap, dan perasaan orang terhadap substansi seperti item, administrasi, organisasi, orang, isu, peristiwa, topik, dan sifatnya (Liu, 2012). Analisis sentiment berfokus kepada opini yang mengandung pesan negatif dan positif. Analisis sentimen merupakan proses mengekstrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif. Metode tersebut membantu untuk mempelajari sentimen dari berbagai macam konten-konten media sosial dalam hal berupa tweet tentang institusi kepolisian negara republik indonesia.

2.3 Text mining

Text mining merupakan bagian dari data mining yang melakukan proses ekstraksi pengetahuan dan informasi dari pola-pola yang terdapat dalam sekumpulan dokumen text menggunakan alat analisis tertentu (Akhmad and Arita, 2021).

Perbedaan antara text mining dan data mining adalah pada penggunaan pola. Pola text mining diambil dari sekumpulan bahasa yang tidak terstruktur sedangkan pola data mining diambil dari database yang terstruktur (Taufik, 2017). Berikut ini tahapan text mining secara umum yaitu text preprocessing merupakan tahap awal untuk mempersiapkan data agar dapat digunakan pada proses selanjutnya. Text preprocessing memiliki beberapa tahapan :

2.3.1 Case Folding

Case folding yaitu proses pengubahan semua karakter huruf kapital menjadi huruf kecil.

2.3.2 Cleansing

Cleansing yaitu proses perubahan, perbaikan, atau penghapusan sejumlah data yang sudah tidak diperlukan seperti tanda baca, *hashtag*, karakter kosong, angka, dan link *URL* dengan tujuan menghasilkan *dataset* yang berkualitas (Veny Amilia Fitri et al., 2019).

2.3.3 Stemming

Stemming merupakan proses mengubah kata menjadi bentuk kata dasar untuk membantu menentukan kata yang paling sering digunakan pada sebuah kalimat .

2.3.4 Tokenizing

Tokenizing merupakan tahapan pemecah kalimat berdasarkan kata-kata yang penyusunannya dipisah oleh karakter *whitespace* atau spasi.

2.3.5 Stopword Removal

Stopword Removal merupakan proses menghilangkan kata-kata yang tidak berpengaruh pada makna maupun sentiment pada sebuah kalimat untuk mempermudah mendeteksi kata kunci disetiap kalimat (Octaviani et all., 2014).

2.3.6 Normalization

Normalization adalah proses perbaikan seluruh kata-kata yang tidak baku atau salah tulis yang ada didalam data. Dalam proses ini, penulis menggunakan file *excel* untuk pemilihan dan perbaikan kata-katanya.

2.4 Pembobotan Kata

Proses pembobotan kata menggunakan metode *Algoritma Term Frequency – Inverse Document Frequency* yang biasa disebut algoritma TF-IDF ini untuk menentukan seberapa jauh keterhubungan antar kata terhadap dokumen dengan memberikan bobot pada setiap kata (Aliandu, 2015).

$$W_{dt} = tf_{dt} \times idf$$

Keterangan :

W : bobot dokumen ke-d pada kata ke-t

d : kalimat ke-t

t : kata ke-t

tf : banyak kata yang dicari pada sebuah kalimat

idf : inversed document frequency

2.5 Support Vector Machine

Support Vector Machine merupakan teknik *supervised learning*, mempunyai tingkat akurasi dan kualitas yang baik. Akan tetapi, untuk implementasinya diperlukan tahap pelatihan *sequential training* dan harus melalui proses pengujian. Kelebihan metode Support Vector Machine dapat mengidentifikasi *hyperplane* yang terpisah sehingga bisa memaksimalkan margin dari kelas yang berbeda. Kekurangan dari metode ini yaitu pada masalah yang mempunyai fitur yang sama dapat memengaruhi tingkat akurasi secara signifikan (Brian Laurensz and Eko Sedyono, 2021). SVM adalah model pembelajaran yang diawasi dengan algoritma pembelajaran terkait yang menganalisis data yang digunakan untuk analisis regresi (Natt Leelawat et al., 2021).

Perhitungan Support Vector Machine menggunakan persamaan 2 :

$$f(x) = \sum_{i=1}^m a_i y_i K(X, X_i) + b \dots \dots \dots (2)$$

Keterangan :

n : banyaknya fitur

α : nilai bobot pada setiap titik data

m : jumlah support vector yang memiliki titik data $a_i > 0$

$K(X, X_i)$: fungsi kernel

Proses perhitungan untuk menentukan dua *hyperplane* memiliki ketentuan tidak ada poin data yang berada diantara dua *hyperplane* dan digunakan sebelum melakukan proses optimalisasi margin pada *hyperplane* tersebut sebagai berikut :

$$y_i(w \cdot X_i + b) \geq 1 \text{ untuk } 1 \leq i \leq n \dots \dots \dots (3)$$

Meminimalisir nilai $\|w\|$ dalam perhitungan margin sehingga proses optimalisasi dapat dilaksanakan pada persamaan 4 dengan bergantung pada persamaan 3 untuk setiap nilai i tersebut persamaannya sebagai berikut :

$$m = \frac{2}{\|w\|} \dots \dots \dots (4)$$

2.6 Naïve Bayes

Metode ini didasari oleh Thomas Bayes, seorang ilmuwan asal Inggris yang memperkenalkan Teorema Bayes yang digunakan untuk memprediksi peluang di masa

depan berdasarkan pengalaman pada masa lalu. Dalam penerapannya pada analisis sentimen, pengalaman dimasa lalu dianalogikan sebagai data training dan masa depan sebagai data testing. Naive Bayes merupakan sebuah metode pembelajaran probabilitas pada machine learning (Hennie Tuhuteru, Ade Iriani, 2018), dapat dirumuskan sebagai berikut :

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Keterangan :

$P(c|x)$: probabilitas hipotesis berdasarkan kondisi

x : data ke- x

c : hipotesis data pada suatu kelas yang spesifik

$P(x|c)$: probabilitas berdasarkan kondisi pada hipotesis

$P(c)$: probabilitas hipotesis

$P(x)$: probabilitas c

Tahap pertama dalam permodelan Naive Bayes yaitu menghitung prior probability. Prior probability merupakan perhitungan probabilitas dari total data.

Berikut persamaannya:

$$P(V_j) = \frac{D_j}{D} \dots\dots\dots(8)$$

Keterangan :

$P(V_j)$: prior probability

D : jumlah dokumen

D_j : jumlah dokumen (D) pada sebuah kelas

Untuk mendapatkan bobot term terhadap semua kelas prediksi atau conditional probability menggunakan persamaan (9) yaitu:

$$P(X_1|V_j) = \frac{w_{ij}+1}{N+N_j} \dots\dots\dots(9)$$

keterangan :

w_{ij} : bobot term

N : jumlah term pada sebuah dokumen

N_j : total term pada dokumen

Saat melakukan proses klasifikasi dokumen, Naive Bayes akan mencari nilai probabilitas tertinggi dari:

$$V_{map} = \underset{V_j \in V}{\operatorname{argmax}} \frac{P(X_1, X_2, X_3, \dots, X_n | V_j) P(V_j)}{P(X_1, X_2, X_3, \dots, X_n)} \dots\dots\dots(10)$$

Jika nilai dari $P(X_1, \dots, X_n)$ adalah permanen untuk semua kategori V_j maka persamaan (10) ini dapat ditulis sebagai berikut:

$$V_{MAP} = \underset{V_j \in V}{\operatorname{argmax}} P(X_1, X_2, X_3, \dots, X_n | V_j) P(V_j) \dots \dots \dots (11)$$

Sehingga dari persamaan (10) dapat ditulis sebagai:

$$V_{MAP} = \prod_{i=1}^n P(X_i | V_j) P(V_j) \dots \dots \dots (12)$$

2.1 Confution Matrix

Confution Matrix merupakan metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining. *Confidence* atau *presisi* merupakan proporsi kasus yang diprediksi positif serta memiliki hasil positif pada data yang sebenarnya. *Sensitivity* atau *recall* adalah proporsi kasus positif berdasarkan data sebenarnya dan diprediksi positif secara benar (Diaz Pilar et al., 2021). Berikut adalah istilah yang digunakan dalam confusion matrix:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

Keterangan :

- True Positive (TP) adalah kelas yang diprediksi positif dan ternyata faktanya benar oleh sistem
- True Negative (TN) adalah kelas yang diprediksi negatif dan ternyata faktanya benar oleh sistem
- False Positive (FP) adalah kelas yang diprediksi positif dan ternyata faktanya salah oleh sistem
- False Negative (FN) adalah kelas yang diprediksi negatif dan ternyata faktanya salah oleh system

BAB 3. METODELOGI PENELITIAN

3.1 Jenis Penelitian

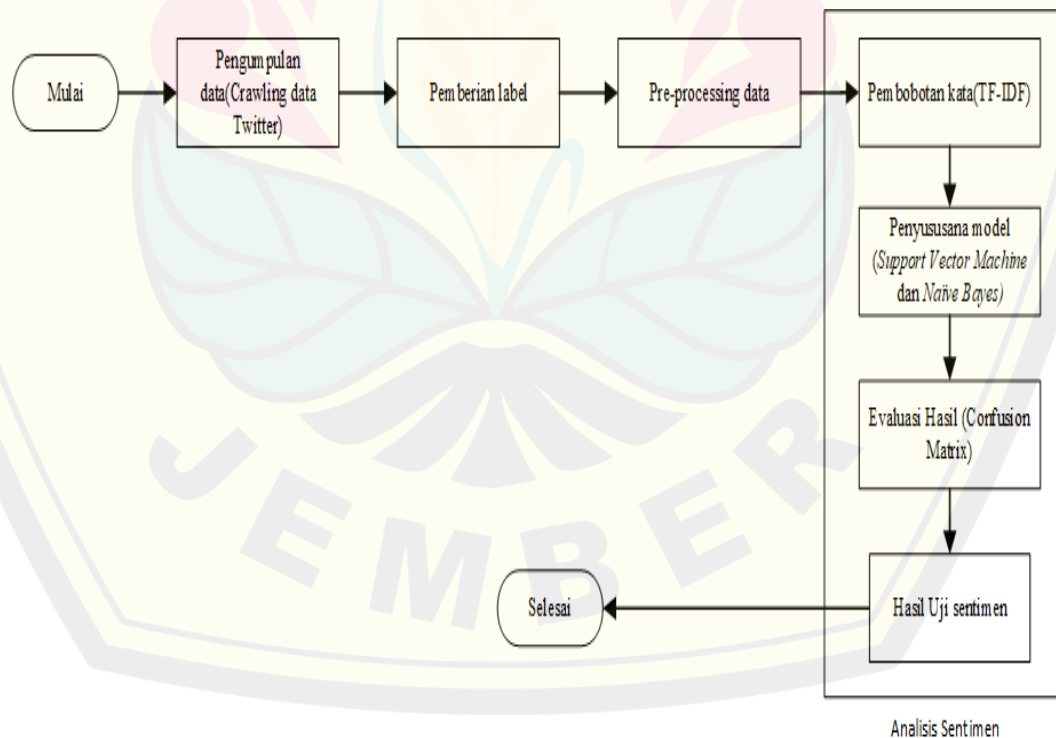
Penelitian yang dilakukan adalah penelitian kuantitatif. Penelitian kuantitatif merupakan suatu pendekatan penelitian yang bersifat objektif mencakup pengumpulan data dan analisis data kuantitatif menggunakan metode pengujian statistik (Hermawan, 2005). Penelitian ini akan fokus pada respon masyarakat terhadap kasus-kasus institusi kepolisian negara republik Indonesia.

3.2 Objek Penelitian

Objek penelitian yang digunakan adalah Tweet dari media sosial Twitter dengan topik opini masyarakat tentang institusi kepolisian negara republik Indonesia. Data yang dibutuhkan untuk pelaksanaan diperoleh dengan cara melakukan pencarian *tweet* menggunakan Teknik *crawling* dengan bantuan *Twitter API key* penulis dan dengan kata kunci yang berhubungan seperti *#institusipolri*. Kata kunci tersebut termasuk kedalam trending topik pada *Twitter*.

3.3 Tahapan Penelitian

Tahap penelitian yang dilakukan pada penelitian ini digambarkan seperti gambar 3.1 dibawah ini:



Gambar 3.1 Tahapan Penelitian

3.3.1 Pengumpulan Data

Tahap pengambilan data melakukan proses pengambilan postingan twitter menggunakan Twitter API. Proses pengambilan data tweet menggunakan library Tweepy yang tersedia pada bahasa pemrograman Python. Pengumpulan data dimulai pada bulan Agustus 2022. Tahap pertama pengambilan data melalui twitter diawali dengan mendaftarkan akun twitter menjadi akun developer, selanjutnya pihak twitter akan melakukan evaluasi terhadap permintaan sebagai akun developer twitter. Setelah mendapatkan akses sebagai developer twitter maka akan mendapatkan API key dan API secret dengan batas waktu selama 7 hari. Apabila crawling data melebihi 7 hari maka dapat dilakukan proses regenerate Api key dan api secret untuk melakukan crawling data lebih lanjut. Hasil *crawling data* yang didapatkan sebanyak 1100 *data tweet*.

3.3.2 Pelabelan Data

Pelabelan data adalah proses menandai data yang disimpan dalam file excel oleh ahli bahasa secara manual. Selama fase ini, ahli bahasa akan mengevaluasi input data ke dalam kelompok positif dan negatif. Kelompok positif adalah kumpulan data yang berisi kata-kata pujian, dukungan, dan motivasi. Kelompok negatif adalah kumpulan data yang berisi kata-kata ejekan, penyangkalan, dan kekerasan.

3.3.3 Pre-processing Data

Pre-processing data bertujuan untuk memproses data agar lebih terstruktur dan lebih mudah dipahami hingga dapat diproses pada tahapan selanjutnya. Preprocessing memiliki beberapa tahapan yaitu :

a. Case Folding

Case folding merupakan suatu proses pengubahan semua karakter huruf kapital menjadi huruf kecil.

b. Cleansing

Cleansing yaitu proses perubahan, perbaikan, atau penghapusan sejumlah data yang sudah tidak diperlukan seperti tanda baca, *hashtag*, karakter kosong, angka, dan link *URL* dengan tujuan menghasilkan *dataset* yang berkualitas.

c. Tokenizing

Tokenizing yaitu tahapan pemecah kalimat berdasarkan kata-kata yang penyusunannya dipisah oleh karakter whitespace atau spasi.

d. Stopword removal

Stopword removal merupakan proses menghilangkan kata-kata yang tidak berpengaruh pada makna maupun sentimen pada sebuah kalimat untuk mempermudah mendeteksi kata kata kunci disetiap kalimat.

e. Normalization

Normalization merupakan proses menyesuaikan kata yang masih belum sesuai dengan Kamus Besar Bahasa Indonesia. Proses normalization menggunakan library nltk dengan memanfaatkan kamus colloquial Indonesian lexicon. Kamus tersebut disimpan dalam google drive kemudian dipanggil menggunakan fungsi `pd.read_csv` agar sistem dapat membaca kamus tersebut untuk diidentifikasi dengan teks tweet.

f. Stemming

Stemming merupakan proses mengubah kata menjadi bentuk kata dasar untuk membantu menentukan kata yang paling sering digunakan pada sebuah kalimat.

3.3.4 Analisis Sentimen

Analisis sentimen merupakan proses mengestrak data teks untuk mendapatkan informasi tentang sentimen bernilai positif dan negatif. Metode tersebut membantu untuk mempelajari sentimen dari berbagai macam konten-konten media sosial dalam hal berupa tweet tentang institusi kepolisian negara republik indonesia. Berikut tahapan analisis sentimen :

a. Pembobotan Kata

Pembobotan kata yaitu *Term Frequency – Inversed Document Frequency* yang digunakan untuk menghitung bobot dari setiap kata (t) dalam sebuah dokumen (d). Metode ini merupakan penggabungan dari dua konsep yaitu *term frequency* yang kemunculan kata dalam kalimat yang menunjukkan seberapa penting sebuah kata dalam kalimat, dan *document frequency* merupakan frekuensi kalimat yang mengandung term atau kata yang menunjukkan seberapa umum kata tersebut. Kata yang semakin sering muncul pada suatu kalimat akan memberikan bobot

yang besar dan akan memberikan nilai bobot yang kecil apabila muncul dalam banyak dokumen.

b. Penyusunan Model

Penyusunan model ini bertujuan untuk dapat menguji seberapa akurat kemampuan prediksi sistem berdasarkan dataset bersih yang dirancang dari skenario pengujian dengan rasio 90:10, 80:20, 70:30, 60:40 untuk perbandingan data training dan data testing.

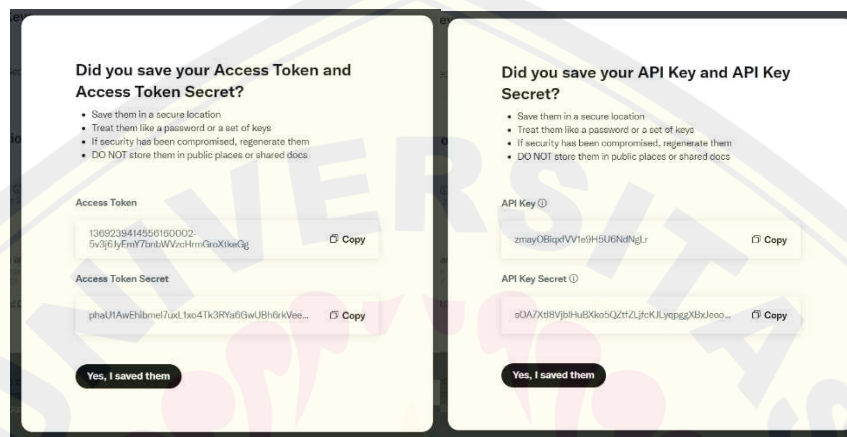
c. Evaluasi Hasil

Data yang sudah selesai diklasifikasikan divalidasi dengan menggunakan Confussion matrix yang merupakan metode yang digunakan untuk melakukan perhitungan akurasi terhadap sebuah sistem diklasifikasi sesuai dengan kelasnya(True Positive, True Negative) dan data yang diklasifikasi tidak sesuai dengan kelasnya(False Positive, False Negative). Data tersebut kemudian dijadikan acuan untuk menghitung akurasi, presisi dan recall.

BAB 4. HASIL DAN PEMBAHASAN

4.1 Hasil Dataset

Dataset yang digunakan dalam penelitian ini diambil dari proses *crawling* di *twitter*. Proses *crawling* data tweet dapat dilakukan dengan cara mendaftarkan akun *twitter* pada website developer portal. Akun *twitter* yang sudah terdaftar dalam mode developer akan mendapatkan *Consumer keys* berupa *API Key & API Key Secret* dan *Authentication Tokens* berupa *Acces Token & Acces Token Secret*. Hasil dari *Twitter API Key* bisa dilihat seperti pada gambar 4.1 berikut ini :



Gambar 4.1 Kode Twitter API

Proses *crawling* data menggunakan *Twitter API Key* yang telah didapatkan dari halaman website khusus developer *Twitter* untuk mendapatkan data tweet yang ada didalam media sosial *Twitter*. Kata kunci yang digunakan dalam proses *crawling* data adalah institusi polri. Kode-kode *Twitter API Key* pada tahap ini dapat dilihat pada Gambar 4.2 berikut ini :

```
api_key = "J79p1taFQsb2umk6Ihb2b7ag1"
api_key_secret = "NHu1FwNCRfQ56aQVXgIzTf6c8MirxYsPJk8E0q9XjFzWDpNDol1"
access_token = "1369239414556160002-ovTmaIboFd20zm1JxkqAZq9Tq4rWGA"
access_token_secret = "M2qmuZg0gnSf7VkeV15wh6iRz4Ufe7q5dDQd10E7Y2V1b"

search_key = "institusi polri"

csvFile = open(search_key+".csv","a+",newline="",encoding="utf-8")
csvWriter =csv.writer(csvFile)
u = []
t = []

for tweet in tweepy.Cursor(api.search,q=search_key,count=1000,lang="id").items():

    u.append(tweet.user.name)
    t.append(tweet.text.encode("utf-8"))
```

Gambar 4.2 Kode Twitter API Key

Hasil crawling data akan disimpan dalam bentuk excel. Contoh data yang dihasilkan dalam proses crawling data tweet dapat dilihat dalam Tabel 4.1 berikut :

Tabel 4.1 Hasil Crawling Data Twitter

No.	Tweet
1	Annashr Maulana RT @Prasetyo9Hh: @KompasTV PC itu Srikandinya Polri.... wajib di layani... seperti jaman Kerajaan kita dulu 😄😄😄
2	🐼 RT @arreneuz_s: @kang__L Pak @ListyoSigitP anda SBG Kapolri tidak menjalankan perintah UU dg baik.
3	RT @camilajenie: Terimakasih Joshua Hutabarat, Tragedi Kematianmu Telah Membuka Borokya Institusi Polri, Membuka Mata Publik Betapa Jahatny..

4.2 Hasil Pelabelan

Data yang didapat dari proses crawling data, kemudian diberi label oleh ahli bahasa. Pelabelan dilakukan secara manual oleh ahli bahasa yang memiliki keahlian dalam memahami makna sehingga dapat memberikan label secara tepat. Total data yang digunakan berjumlah 1100, diambil data yang representatif pada sentimen positif dan negatif sesuai dengan kata kunci yang ditetapkan. Tabel 4.2 merupakan contoh data yang telah diberi label oleh ahli bahasa.

Tabel 4.2 Hasil Pelabelan

No.	Tweet	Sentimen
1	Annashr Maulana RT @Prasetyo9Hh: @KompasTV PC itu Srikandinya Polri.... wajib di layani... seperti jaman Kerajaan kita dulu 😄😄😄	Negatif
2	🐼 RT @arreneuz_s: @kang__L Pak @ListyoSigitP anda SBG Kapolri tidak menjalankan perintah UU dg baik.	Negatif
3	RT @camilajenie: Terimakasih Joshua Hutabarat, Tragedi Kematianmu Telah Membuka Borokya Institusi Polri, Membuka Mata Publik Betapa Jahatny..	Negatif

4.3 Hasil Preprocessing

Implementasi preprocessing melalui beberapa tahapan berupa *casefolding*, *cleansing*, *tokenizing*, *stopword removal*, *normalization* dan *stemming*. *Casefolding* merupakan tahapan mengubah data *tweet* menjadi huruf kecil (*lowercase*). Berikut implementasi kode program *casefolding* dapat dilihat pada Gambar 4.3:

```
df['lower'] = df['text'].str.lower()
```

Gambar 4.3 Kode Casefolding

Kode program diatas berfungsi untuk mendefinisikan tabel teks yang diubah menjadi huruf kecil menggunakan perintah *lower()* dan ditampung dalam tabel *lower*. Kode program tersebut menghasilkan sampel data. Berikut beberapa contoh sampel data pada Tabel 4.3:

Tabel 4.3 Hasil Casefolding

No.	Tweet	Casefolding
1	Annashr Maulana RT @Prasetyo9Hh: @KompasTV PC itu Srikandinya Polri... wajib di layani... seperti jaman Kerajaan kita dulu 🤔🤔🤔	annashr maulana rt @prasetyo9hh: @kompastv pc itu srikandinya polri... wajib di layani... seperti jaman kerajaan kita dulu 🤔🤔🤔
2	🐼 RT @arreneuz_s: @kang__L Pak @ListyoSigitP anda SBG Kapolri tidak menjalankan perintah UU dg baik.	🐼 rt @arreneuz_s: @kang__l pak @listyosigitp anda sbg kapolri tidak menjalankan perintah uu dg baik.
3	RT @camilajenie: Terimakasih Joshua Hutabarat, Tragedi Kematianmu Telah Membuka Borokya Institusi Polri, Membuka Mata Publik Betapa Jahatny..	rt @camilajenie: terimakasih joshua hutabarat, tragedi kematianmu telah membuka borokya institusi polri, membuka mata publik betapa jahatny..

Setelah proses casefolding tahap selanjutnya adalah cleansing. Pada tahap ini menghapus berbagai komponen pada tweet yang tidak memiliki pengaruh seperti mention, hashtag, kode HTML, tweet dan menghapus angka, spaso serta tanda baca. Berikut implementasi kode program cleansing dapat dilihat pada Gambar 4.4:

```
# remove https* and www*
df['remove_punct'] = df['lower'].str.replace(r'((www\.[^\s]+)|(https?:\/\/[^\s]+))', '')
# remove username
df['remove_punct'] = df['remove_punct'].str.replace(r'@[^\s]+', '')
# remove additional white spaces
df['remove_punct'] = df['remove_punct'].str.replace(r'[\s]+', ' ')
# remove hashtag
df['remove_punct'] = df['remove_punct'].str.replace(r'#([^\s]+)', r'\1')
# remove RT
df['remove_punct'] = df['remove_punct'].str.replace(r'rt', ' ')
# remove punct
df['remove_punct'] = df['remove_punct'].str.replace(r'[\W\s]+', '')
# remove number
df['remove_punct'] = df['remove_punct'].str.replace(r'\d', '')
```

Gambar 4.4 Kode Remove punctutation

Tahapan *remove punctutation* data tweet pada gambar 4.4 memanfaatkan fungsi *library re* atau *regular expression (regex)*. Kode `(www\.[^\s]+)|(https?:\/\/[^\s]+)` berfungsi menghapus link seperti https dan www dan ditampung dalam tabel *remove_punct*. Kode merupakan kode program untuk menghapus username dengan mencocokkan `@([^\s]+)`. Berfungsi untuk menghapus spasi dengan kode `[\s]`. Kode yang berfungsi untuk menghapus hashtag dengan kode `#([^\s]+)`. Kode yang berfungsi untuk menghapus tanda baca dengan kode program `[\W\s]+`. Selanjutnya merupakan kode program menghapus angka dengan kode program `\d`. Berikut hasil dari kode program diatas sebagaimana Tabel 4.4:

Tabel 4.4 Hasil Cleansing

No.	tweet	remove_punct
1	annashr maulana rt @prasetyo9hh: @kompastv pc itu srikandinya polri... wajib di layani... seperti jaman kerajaan kita dulu 😊😊😊	annashr maulana pc itu srikandinya polri wajib di layani sepe i jaman kerajaan kita dulu
2	🐼 rt @arreneuz_s: @kang_l pak @listyosigitp anda sbg kapolri tidak menjalankan perintah uu dg baik.	pak anda sbg kapolri tidak menjalankan perintah uu dg baik
3	rt @camilajenie: terimakasih joshua hutabarat, tragedi kematianmu telah membuka borokya institusi polri, membuka mata publik betapa jahatny..	terimakasih joshua hutabarat tragedi kematianmu telah membuka borokya institusi polri membuka mata publik betapa jahatny

Tahapan selanjutnya yaitu tokenizing yang dilakukan untuk melakukan pemotongan terhadap kalimat menjadi kata-kata secara terpisah. Berikut kode program tokenizing dapat dilihat pada Gambar 4.5:

```
df['tokens'] = df['stemming'].apply(lambda x: word_tokenize(str(x)))
```

Gambar 4.5 Kode Tokenizing

Pada proses *tokenizing* menggunakan fungsi *word_tokenize* yang terdapat dalam library *nltk* sebagaimana pada kode program gambar 4.5. Tokenisasi yang dilakukan pada tabel *remove_punct* ditampung pada tabel *tokens*. Berikut hasil kode program dapat dilihat pada Tabel 4.5 dibawah ini:

Tabel 4.5 Hasil Tokenizing

No.	tweet	tokens
1	annashr maulana pc itu srikandinya polri wajib di layani seperti jaman kerajaan kita dulu	['annashr', 'maulana', 'pc', 'itu', 'srikandi', 'polri', 'wajib', 'di', 'layan', 'seperti', 'jaman', 'raja', 'kita', 'dulu']
2	pak anda sbg kapolri tidak menjalankan perintah uu dg baik	['pak', 'anda', 'sbg', 'kapolri', 'tidak', 'jalan', 'perintah', 'uu', 'dg', 'baik']
3	terimakasih joshua hutabarat tragedi kematianmu telah membuka borokya institusi polri membuka mata publik betapa jahatny	['terimakasih', 'joshua', 'hutabarat', 'tragedi', 'mati', 'telah', 'buka', 'borokya', 'institusi', 'polri', 'buka', 'mata', 'publik', 'betapa', 'jahatny']

Tahapan normalisasi merupakan proses mengubah kata yang tidak padu dan belum efektif mejadi kata yang standar sesuai dengan KBBI (kamus besar bahasa indonesia). Berdasarkan hasil *crawling* data tweet menunjukkan banyak penggunaar kalimat slang atau

bahasa gaul yang digunakan oleh pengguna, juga dilakukan proses mengubah kata yang disingkat menjadi frasa yang sesuai dengan bentuk aslinya. Berikut kode program normalisasi dapat dilihat pada Gambar 4.6:

```
normalized_word = pd.read_csv('drive/My Drive/BISMILLAH SKRIPSI/Sidang/colloquial-indonesian-lexicon.csv')
normalized_word_dict = {}

for index, row in normalized_word.iterrows():
    if row[0] not in normalized_word_dict:
        normalized_word_dict[row[0]] = row[1]

def normalized_term(document):
    return [normalized_word_dict[term] if term in normalized_word_dict else term for term in document]

df['normalized'] = df['tokens'].apply(normalized_term)
```

Gambar 4.6 Kode Normalisasi

Tahapan normalisais proses pertama membaca kamus *colloquial-indonesian-lexicon.csv* dalam *google drive*. Kode program *for index* merupakan proses perulangan untuk membaca dan mencocokkan kata yang kemudian didefinisikan ke dalam fungsi *normalized_term* dan dikembalikan nilainya pada *normalized_word_dict*. Hasil normalisasi kata ditampung dalam tabel *normalized* pada baris kode program 11. Berikut sampel hasil pada Tabel 4.6 dibawah ini:

Tabel 4.6 Hasil Normalisasi

No.	tweet	normalisasi
1	pak anda sbg kapolri tidak menjalankan perintah uu dg baik	['pak', 'anda', 'sebagai', 'kapolri', 'tidak', 'menjalankan', 'perintah', 'uu', 'dengan', 'baik']
2	gmni ini pak apa bnar rakyat percaya dgn institusi polri	['bagaimana', 'ini', 'pak', 'apa', 'bnar', 'rakyat', 'percaya', 'dengan', 'institusi', 'polri']
3	jk bener skrg saatnya pak institusi polri wajib tegas bersih dr oknum yg gak da tanggungjwbnya bagi nkri	['jika', 'benar', 'sekarang', 'saatnya', 'pak', 'institusi', 'polri', 'wajib', 'tegas', 'bersih', 'dari', 'oknum', 'yang', 'enggak', 'sudah', 'tanggungjwbnya', 'bagi', 'nkri']

Tahapan berikutnya yaitu stopwords atau pengambilan kata penting dalam sebuah kalimat. Proses ini menghapus kata yang tidak memiliki pengaruh terhadap makna kalimat seperti kata sambung, preposisi dan partikel. Berikut kode program stopwords dapat dilihat pada Gambar 4.7 dan Gambar 4.8 dibawah ini:

```
import nltk
from nltk.corpus import stopwords
list_stopwords = stopwords.words('indonesian')
list_stopwords.remove('tidak')
print(list_stopwords)

# convert list to dictionary
list_stopwords = set(list_stopwords)

#remove stopword pada list token
def stopwords_removal(words):
    return [word for word in words if word not in list_stopwords]
df['cleaned'] = df['normalized'].apply(stopwords_removal)
```

Gambar 4.7 Kode Stopwords 1

```
from nltk.tokenize.treebank import TreebankWordDetokenizer

df['clean_text']=df['cleaned'].apply(lambda x: TreebankWordDetokenizer().detokenize(x))
df.head()
```

Gambar 4.8 Kode Stopword 2

Kode program pada gambar 4.7 menunjukkan proses mengubah kata yang dianggap tidak penting berdasarkan list kata stopwords Indonesia pada *library nltk*. Kode program *list_stopwords.remove()* menunjukkan proses untuk menghapus kata tidak dalam *list stopword*. Kata tidak dihapus karena termasuk komponen penting dalam memberikan label positif dan negatif dalam analisis sentimen. Kode program *set(list_stopwords)* merupakan proses *convert list* stopword, kode selanjutnya merupakan proses stopword. Kemudian hasilnya ditampung dalam tabel *clean_text* sebagaimana kode program pada gambar 4.8. Hasil list stopwords yang telah dibersihkan dari kata tidak, dapat dilihat dalam Tabel 4.7:

Tabel 4.7 Hasil List Stopwords

['ada', 'adalah', 'adanya', 'adapun', 'agak', 'agaknya', 'agar', 'akan', 'akankah', 'akhir', 'akhiri', 'akhirnya', 'aku', 'akulah', 'amat', 'amatlah', 'anda', 'andalah', 'antar', 'antara', 'antaranya', 'apa', 'apaan', 'apabila', 'apakah', 'apalagi', 'apatah', 'artinya', 'asal', 'asalkan', 'atas', 'atau', 'ataukah', 'ataupun', 'awal', 'awalnya', 'bagai', 'bagaikan', 'bagaimana', 'bagaimanakah', 'bagaimanapun', 'bagi', 'bagian', 'bahkan', 'bahwa', 'bahwasanya', 'baik', 'bakal', 'bakalan', 'balik', 'banyak', 'bapak', 'baru', 'bawah', 'beberapa', 'begini', 'beginian', 'beginikah', 'beginilah', 'begitu', 'begitukah', 'begitulah', 'begitupun', 'bekerja', 'belakang', 'belakangan', 'belum', 'belumlah', 'benar', 'benarkah', 'benarlah', 'berada', 'berakhir', 'berakhirlah', 'berakhirnya', 'berapa', 'berapakah', 'berapalah', 'berapapun', 'berarti', 'berawal', 'berbagai', 'berdatangan', 'beri', 'berikan', 'berikut', 'berikutnya']

Kode program stopwords memberikan hasil sebagaimana sampel data pada Tabel 4.8 berikut ini:

Tabel 4.8 Hasil Stopwords

No.	tweet	clean_text
1	['pesan', 'kapolri', 'ke', 'seluruh', 'personel', 'polri', 'terus', 'semangat', 'karena', 'institusi', 'ini', 'sampai', 'kapan', 'pun', 'harus', 'kita', 'jaga', '']	pesan kapolri personel polri semangat institusi jaga
2	['kapolri', 'menyampaikan', 'terima', 'kasih', 'atas', 'dukungan', 'berbagai', 'pihak', 'terhadap', 'institusi', 'polri', 'yang', 'sedang', 'menghadapi', 'situasi', 'berat']	kapolri terima kasih dukungan institusi polri menghadapi situasi berat
3	['kasus', 'sambo', 'adalah', 'cerminan', 'parahnya', 'korupsi']	sambo cerminan parahnya korupsi

Stemming merupakan tahap pengambilan kata dasar atau root dengan cara membuang imbuhan, sisipan awalan, serta awala dan imbuhan. Berikut kode program stemming dapat dilihat pada Gambar 4.9 dibawah ini:

```
def mystem(cols):
    res = []
    cols = cols.apply(lambda x: word_tokenize(str(x)))
    """
    for item in cols:
        for word in item:
            print(word)
            s = stemmer.stem(word); print(s);
    """
    #for loop diatas buat cek fungsi stemmer.stem()

    cols = cols.apply(lambda x: [stemmer.stem(y) for y in x])
    return cols

stemmer = StemmerFactory().create_stemmer()
df['stemming'] = df['remove_punct']
df['stemming'] = mystem(df['stemming'])
```

Gambar 4.9 Kode Stemming

Proses stemming menggunakan library Sastrawi dengan memanfaatkan fungsi *StemmerFactory*.for loop pada gambar diatas untuk pengecekan fungsi *stemmer.stem()*. Kemudian disimpan dalam tabel *stemming*. Hasil dari kode program stemming diatas dapat dilihat pada Tabel 4.9:

Tabel 4.9 Hasil Stemming

No.	tweet	stemming
1	peringatan besar untuk kepolisian republik indonesia pak kapolri	['ingat', 'besar', 'untuk', 'polisi', 'republik', 'indonesia', 'pak', 'kapolri']
2	kabarnya lagi di aceh ada polisi tewas ditembak	['kabar', 'lagi', 'di', 'aceh', 'ada', 'polisi', 'tewas', 'tembak']
3	semoga institusi polri terus maju dan semakin mengedukasi masyarakat polri	['moga', 'institusi', 'polri', 'terus', 'maju', 'dan', 'makin', 'edukasi', 'masyarakat', 'polri']

4.4 Hasil Bag of Word

Bag of words merupakan sebuah bentuk representasi sederhana sebuah teks kedalam angka yang memanfaatkan *library sklearn* dengan fungsi *featur_extraction* kode program yang berfungsi untuk mendefinisikan variable *vectorize* dan hasil dari transformasi tersebut diubah kedalam bentuk *array*. Hasil bag of words ditampung ke dalam variable *y* pada kode program baris 5. Berikut kode program *bag of words* pada Gambar 4.9:

```
#Creating Bag of Words Model
from sklearn.feature_extraction.text import CountVectorizer
cv = CountVectorizer()
x = cv.fit_transform(df['clean_text']).toarray()
y = df.iloc[:, 1].values
```

Gambar 4.10 Kode Bag Of Words

Berikut contoh teks yang direpresentasikan sebagai *bag of word*:

- 1) contoh bejadnya akhlak di institusi polri
- 2) apa benar rakyat percaya dengan institusi polri

Berdasarkan kalimat diatas yang kemudian akan dibentuk list kata sebagai berikut:

“contoh”, “bejadnya”, “akhlak”, “di”, “institusi”, “polri”,
 “rakyat”, “tidak”, “percaya”, “dengan”, “institusi”, “polri”

Kemudian kata diatas dihitung frekuensinya dan dipetakan pada setiap kalimat pembentuknya dapat dilihat pada tabel 4.10 dibawah ini:

Tabel 4.10

teks	contoh	bejadnya	akhlak	di	institusi	polri	rakyat	tidak	percaya	dengan
1	1	1	1	1	1	1	0	0	0	0
2	0	0	0	0	1	1	1	1	1	1

4.5 Hasil Pembobotan TF-IDF

Pelaksanaan pembobotan kata pada penelitian ini dilakukan dengan mengimplementasikan algoritma term frequency-inverse document frequency (TF-IDF). TF-IDF merupakan suatu metode yang digunakan untuk memberikan pembobotan terhadap keterhubungan suatu kata dengan dokumen. Perhitungan term frequency kata yang sering muncul dianggap mewakili sebuah dokumen. Sedangkan proses term frequency inverse document frequency semakin sering kata muncul dalam banyak dokumen maka kata tersebut tidak dianggap mewakili sebuah dokumen karena bersifat umum. Proses ini dimulai dengan melakukan perhitungan term frequency (TF) kemunculan setiap kata pada setiap kalimat yang ada pada sebuah dokumen dibagi dengan jumlah total kata yang ada pada suatu kalimat. Kemudian menentukan nilai IDF dengan cara melakukan perhitungan nilai log dari hasil document frequency (DF) yang telah diproses oleh sistem. Tahap selanjutnya adalah melakukan perhitungan TF-IDF. Potongan kode program dapat dilihat pada gambar dibawah ini, TF-IDF ini memanfaatkan *library sklearn* dengan fungsi *feature_extraction* berfungsi mendefinisikan variabel *TfidfTransformer()* dengan nama variabel *tfidfconverter* dan hasil TF-IDF akan ditampung dalam variabel X. Berikut kode program TF-IDF pada Gambar 4.11:

```
#Creating TF-IDF
from sklearn.feature_extraction.text import TfidfTransformer
tfidfconverter = TfidfTransformer()
X= tfidfconverter.fit_transform(X).toarray()
```

Gambar 4.11 Kode TF-IDF

Berikut contoh perhitungan Term Frequency pada data yang telah menyelesaikan tahap preprocessing dapat dilihat pada Tabel 4.10 dibawah ini:

Tabel 4.10 Contoh Hasil Perhitungan TF

No	clean_text	Term Frequency (TF)	
1	['kasus', 'sambo', 'adalah', 'cermin', 'parah', 'korupsi']	Kata	Nilai TF
		kasus	1/6
		sambo	1/6
		adalah	1/6
		cermin	1/6
		parah	1/6
		korupsi	1/6
2	['moga', 'institusi', 'polri', 'makin', 'baik', 'depan']	Kata	Nilai TF
		moga	1/6
		institusi	1/6
		polri	1/6
		makin	1/6
		baik	1/6
		depan	1/6
3	['ingat', 'besar', 'untuk', 'polisi', 'republik', 'indonesia', 'pak', 'kaporli']	Kata	Nilai TF
		ingat	1/8
		besar	1/8
		untuk	1/8
		polisi	1/8
		republik	1/8
		indonesia	1/8
		pak	1/8
		kaporli	1/8

Proses selanjutnya melakukan perhitungan untuk menentukan document frequency (DF) dan invers document frequency (IDF). Pada tahap ini dilakukan perhitungan nilai DF

setiap kata terlebih dahulu. Setelah nilai DF telah berhasil ditentukan, kemudian dilakukan perhitungan nilai IDF setiap kata dengan cara menghitung nilai log dari nilai DF yang telah dihitung sebelumnya dengan menggunakan implementasi kode program pada Gambar 4.10. Berikut contoh hasil nilai DF dan IDF setiap kata yang ada di dalam data dapat dilihat pada Tabel 4.11 dibawah ini:

Tabel 4.11 Contoh Hasil Nilai DF dan IDF Setiap Kata

No.	Kata	d	df	idf
1	kasus	3	1	$\text{Log}(3/1) = 0.4771212547$
2	sambo	3	1	$\text{Log}(3/1) = 0.4771212547$
3	adalah	3	1	$\text{Log}(3/1) = 0.4771212547$
4	cermin	3	1	$\text{Log}(3/1) = 0.4771212547$
5	parah	3	1	$\text{Log}(3/1) = 0.4771212547$
6	korupsi	3	1	$\text{Log}(3/1) = 0.4771212547$
7	moga	3	1	$\text{Log}(3/1) = 0.4771212547$
8	institusi	3	1	$\text{Log}(3/1) = 0.4771212547$
9	polri	3	1	$\text{Log}(3/1) = 0.4771212547$
10	makin	3	1	$\text{Log}(3/1) = 0.4771212547$
11	baik	3	1	$\text{Log}(3/1) = 0.4771212547$
12	depan	3	1	$\text{Log}(3/1) = 0.4771212547$
13	ingat	3	1	$\text{Log}(3/1) = 0.4771212547$
14	besar	3	1	$\text{Log}(3/1) = 0.4771212547$
15	untuk	3	1	$\text{Log}(3/1) = 0.4771212547$
16	polisi	3	1	$\text{Log}(3/1) = 0.4771212547$
17	republic	3	1	$\text{Log}(3/1) = 0.4771212547$
18	indonesia	3	1	$\text{Log}(3/1) = 0.4771212547$
19	pak	3	1	$\text{Log}(3/1) = 0.4771212547$
20	kapolri	3	1	$\text{Log}(3/1) = 0.4771212547$

Tahap selanjutnya yaitu menghitung nilai term frequency-invers document frequency (TF-IDF) dengan cara melakukan perkalian antara Term frequency (TF) dan Invers document frequency (IDF). Berikut contoh hasil perhitungan nilai TF-IDF setiap kata dapat dilihat pada Tabel 4.12 dibawah ini:

Tabel 4.12 Hasil TF-IDF Tiap Kata

No.	Kata	d	idf	TF-IDF
1	kasus	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
2	sambo	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
3	adalah	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
4	cermin	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
5	parah	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
6	korupsi	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
7	moga	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
8	institusi	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
9	polri	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
10	makin	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
11	baik	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
12	depan	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
13	ingat	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
14	besar	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
15	untuk	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$

16	polisi	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
17	republic	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
18	indonesia	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
19	pak	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$
20	kapolri	3	0.4771212547	$3 * 0.4771212547$ $=1.4313637641$

4.6 Hasil Pemodelan

Hasil pemodelan pada penelitian ini menggunakan 2 metode yaitu support vector machine dan naïve bayes yang berfokus untuk membandingkan nilai akurasi dari kedua metode tersebut.

4.6.1 Hasil Pemodelan Support Vector Machine

Pemodelan *support vector machine* merupakan klasifikasi data yang mengharapkan output klasifikasi berupa kelas positif dan negatif. Split data memanfaatkan *library sklearn* pada potongan program *import train_test_split*. Kode selanjutnya berfungsi untuk melakukan splitting data testing dan data training. Berikut kode program split data pada Gambar 4.12 dibawah ini:

```
#Splitting dataset into training and testing
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.10, random_state = 0)
```

Gambar 4.12 Kode Split Data

Proses support vector machine memanfaatkan library sklearn dengan menggunakan jenis *support vector machine* (SVC). Kode program *import SVC* berfungsi untuk mengimpor modul SVM dan membuat objek *support vector classifier* dengan meneruskan kernel argument sebagai *random_state* dalam fungsi *SVC()* dengan nama classifier. Hasil pemodelan data train ditampung pada classifier menggunakan *fit()* dan melakukan prediksi pada data test menggunakan *predict()* dan menyimpan daftar label prediksi pada *y_pred*. berikut ptongan kode program dapat dilihat pada Gambar 4.12 dibawah ini:

```
#Import svm model
from sklearn.svm import SVC

#Create a svm Classifier
classifier = SVC(random_state=0)
#Train the model using the training sets
classifier.fit(X_train, y_train)
#Predictiong the test set results
y_pred = classifier.predict(X_test)
```

Gambar 4.13 Kode SVM

4.6.2 Hasil Pemodelan Naïve Bayes

Pemodelan naïve bayes merupakan klasifikasi data yang mengharapkan output klasifikasi berupa kelas positif dan negatif. Penerapan naïve bayes memanfaatkan library sklearn. Berikut potongan kode program yang berfungsi untuk melakukan splitting data testing dan training pada Gambar 4.13 dibawah ini:

```
#Splitting dataset into training and testing
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.10, random_state = 0)
```

Gambar 4.14 Kode Splitting Data

Proses naïve bayes memanfaatkan library sklearn dengan menggunakan jenis multinominal naïve bayes. Kode program *classifier = multinomialNB()* berfungsi untuk mendefinisikan variabel *multinomialNB()* dengan nama classifier. Hasil pemodelan naïve bayes ditampung dalam variabel classifier. Berikut potongan kode program dapat dilihat pada Gambar 4.14 dibawah ini:

```
from sklearn.naive_bayes import MultinomialNB
# from sklearn.svm import SVC
classifier = MultinomialNB()
classifier = classifier.fit(X_train, y_train)
```

Gambar 4.15 Kode Naïve Bayes

4.7 Hasil Uji Evaluasi

Tahapan hasil uji evaluasi pada penelitian ini menggunakan proses confusion matrik yang memanfaatkan library sklearn dengan menggunakan sklearn.metrics dengan tujuan untuk mengetahui seberapa baik tingkat akurasi pada proses klasifikasi yang telah diproses oleh sistem. Pada penelitian ini, tahapan ini dilakukan sebanyak empat kali percobaan dengan menggunakan data training dan data testing masing-masing sebanyak 90%:10%, 80%:20%, 70%:30%, dan 60%:40%. Dalam pengimplementasiannya, perhitungan akurasi, presisi, recall, dan f-score juga digunakan dalam penelitian ini dengan menggunakan rangkaian kode program. Pada kode program Gambar 4.16 *classification_report()* berfungsi menghasilkan presisi, recall, fscore untuk setiap kelas target serta memiliki beberapa nilai tambahan yaitu macro average dan weight average. Fungsi *sklearn*

`confusion_matrices()` mengembalikan nilai *confusion matrix*, *output* sedikit berbeda, dibutuhkan garis sebagai nilai actual dan kolom sebagai nilai yang diprediksi. Berikut kode program dapat dilihat pada Gambar 4.16:

```
#Confusion matrix
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import ConfusionMatrixDisplay

confusion_matrix(y_test, y_pred)

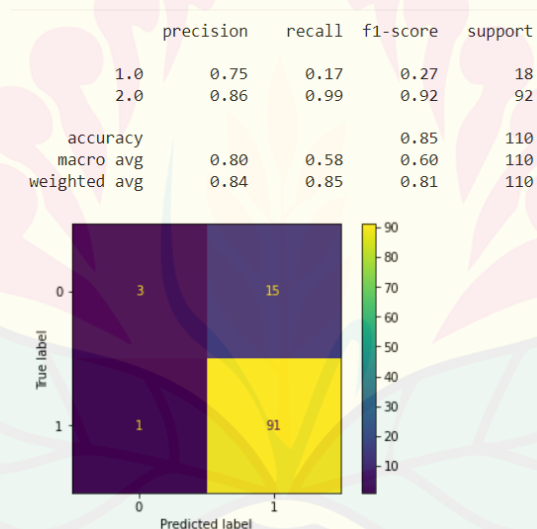
cm = confusion_matrix(y_test, y_pred)

cm_display = ConfusionMatrixDisplay(cm).plot()

print(classification_report(y_test, y_pred))
```

Gambar 4.16 Kode Confusion Matrix

Model yang dipasang berfungsi untuk menghitung prediksi model pada kumpulan data uji. Prediksi ini digunakan untuk menghitung *confusion matrix* yang diplot dengan *ConfusionMatrixDisplay*. Hasil dari *confusion matrix* untuk *naïve bayes* dan *support vector machine* yang divisualisasikan dalam Gambar 4.17 dan Gambar 4.18 dibawah ini:



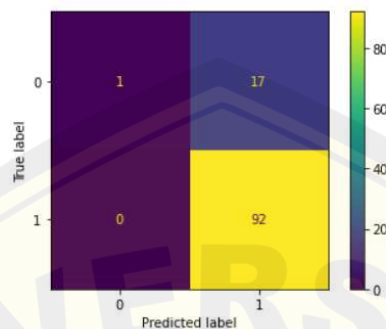
Gambar 4.17 Confusion Matrix SVM

Confusion matrix (Support Vector Machine) yang dibuat memiliki empat kuadran yang berbeda yaitu:

- True Negative (kuadran atas-kiri) = 3 yang artinya 3 titik data kelas negative diklasifikasikan dengan benar oleh model.
- False Negative (kuadran bawah-kiri) = 1 yang artinya 1 titik data kelas positif salah diklasifikasikan sebagai milik kelas negatif oleh model.
- False Positive (kuadran atas-kanan) = 15 yang artinya 15 titik data kelas negatif salah diklasifikasikan sebagai kelas positif oleh model.

- d. True Positive (kuadran bawah-kanan) = 91 yang artinya 91 titik data kelas positif diklasifikasikan dengan benar oleh model.

	precision	recall	f1-score	support
1.0	1.00	0.06	0.11	18
2.0	0.84	1.00	0.92	92
accuracy			0.85	110
macro avg	0.92	0.53	0.51	110
weighted avg	0.87	0.85	0.78	110



Gambar 4.18 Confusion Matrix Naïve Bayes

Confusion matrix (Naïve Bayes) yang dibuat memiliki empat kuadran yang berbeda yaitu:

- True Negative (kuadran atas-kiri) = 1 yang artinya 1 titik data kelas negative diklasifikasikan dengan benar oleh model.
- False Negative (kuadran bawah-kiri) = 0 yang artinya 0 titik data kelas positif salah diklasifikasikan sebagai milik kelas negatif oleh model.
- False Positive (kuadran atas-kanan) = 17 yang artinya 17 titik data kelas negatif salah diklasifikasikan sebagai kelas positif oleh model.
- True Positive (kuadran bawah-kanan) = 92 yang artinya 92 titik data kelas positif diklasifikasikan dengan benar oleh model.

Dalam pengimplementasian perhitungan akurasi, presisi, recall, dan fscore juga digunakan dalam penelitian ini dengan menggunakan rangkaian kode program yang dapat dilihat pada Gambar 4.19:

```

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from sklearn.metrics import precision_recall_fscore_support as score

precision,recall,fscore,support=score(y_test,y_pred,average='macro')
print('Precision : ',format(precision))
print('Recall    : ',format(recall))
print('F-score   : ',format(fscore))
print('Accuracy  : ',accuracy_score(y_test, y_pred))

```

Gambar 4.19 Kode Perhitungan Presisi, Recall, Fscore dan Akurasi

Pada Gambar 4.19 merupakan pengujian model yang berfungsi untuk menghitung tingkat akurasi model yang sebenarnya. Pada penelitian ini tahap *confusion matrix* memanfaatkan *library sklearn*. Proses perhitungan *confusion matrix* nilai yang digunakan adalah *macro average*. Kode program *print()* berfungsi untuk mencetak presisi, recall, fscore dan akurasi.

Hasil data uji evaluasi pada tahapan ini dapat dilihat pada Tabel 4.13 dan Gambar 4.20 untuk metode Support Vector Machine sebagai berikut:

Tabel 4.13 Hasil Data Uji Evaluasi SVM

Keterangan	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4
Data Training	90%	80%	70%	60%
Data Testing	10%	20%	30%	40%
True Positive	91	174	261	346
True Negative	3	13	24	32
False Positive	15	31	43	60
False Negative	1	2	2	2
Precision	81%	86%	89%	89%
Recall	57%	64%	67%	67%
F-score	59%	67%	72%	72%
Accuracy	85%	85%	86%	86%

```

Precision : 0.804245283018868
Recall    : 0.5778985507246377
F-score   : 0.595959595959596
Accuracy  : 0.8545454545454545

```

Gambar 4.20 Hasil Percobaan 1 SVM

Pada Tabel 4.13 dan Gambar 4.20 menunjukkan nilai precision, recall, fscore, dan accuracy yang didapatkan dari algoritma SVM dengan nilai average macro. Pada

penelitian ini dilakukan empat kali percobaan, berikut hasil yang didapat dari penelitian ini:

- Hasil percobaan 1 dihasilkan dari split data 90% : 10%, menghasilkan nilai *precision* sebesar 81%, *recall* sebesar 57%, *fscore* 59%, dan *accuracy* 85%.
- Hasil percobaan 2 dihasilkan dari split data 80% : 20%, menghasilkan nilai *precision* sebesar 86%, *recall* sebesar 64%, *fscore* 67%, dan *accuracy* 85%.
- Hasil percobaan 3 dihasilkan dari split data 70% : 30%, menghasilkan nilai *precision* sebesar 89%, *recall* sebesar 67%, *fscore* 72%, dan *accuracy* 86%.
- Hasil percobaan 4 dihasilkan dari split data 60% : 40%, menghasilkan nilai *precision* sebesar 81%, *recall* sebesar 67%, *fscore* 72%, dan *accuracy* 86%.

Hasil data uji evaluasi pada tahapan ini dapat dilihat pada Tabel 4.14 dan Gambar 4.21 untuk metode Naïve Bayes sebagai berikut:

Tabel 4.14 Hasil Data Uji Evaluasi Naïve Bayes

Keterangan	Percobaan 1	Percobaan 2	Percobaan 3	Percobaan 4
Data Training	90%	80%	70%	60%
Data Testing	10%	20%	30%	40%
True Positive	92	176	263	348
True Negative	1	1	1	1
False Positive	17	43	66	91
False Negative	0	0	0	0
Precision	92%	91%	89%	89%
Recall	53%	51%	51%	51%
F-score	51%	47%	46%	45%
Accuracy	85%	81%	80%	79%

Precision : 0.9220183486238532
 Recall : 0.5277777777777778
 F-score : 0.5103430217334382
 Accuracy : 0.8454545454545455

Gambar 4.21 Hasil Percobaan 1 Naïve Bayes

Pada Tabel 4.14 dan Gambar 4.21 menunjukkan nilai *precision*, *recall*, *fscore*, dan *accuracy* yang didapatkan dari algoritma Naïve Bayes dengan nilai *average macro*. Pada

penelitian ini dilakukan empat kali percobaan, berikut hasil yang didapat dari penelitian ini:

- Hasil percobaan 1 dihasilkan dari split data 90% : 10%, menghasilkan nilai *precison* sebesar 92%, *recall* sebesar 53%, *fscore* 51%, dan *accuracy* 85%.
- Hasil percobaan 2 dihasilkan dari split data 80% : 20%, menghasilkan nilai *precison* sebesar 91%, *recall* sebesar 51%, *fscore* 47%, dan *accuracy* 81%.
- Hasil percobaan 3 dihasilkan dari split data 70% : 30%, menghasilkan nilai *precison* sebesar 89%, *recall* sebesar 51%, *fscore* 46%, dan *accuracy* 80%.
- Hasil percobaan 4 dihasilkan dari split data 60% : 40%, menghasilkan nilai *precison* sebesar 89%, *recall* sebesar 51%, *fscore* 45%, dan *accuracy* 79%.

Hasil dari seluruh keempat percobaan pada kedua metode yang telah dilakukan dengan hasil yang telah dirata-rata dari empat kali percobaan metode support vector machine nilai akurasinya sebesar 85,5% sedangkan naïve bayes nilai akurasinya sebesar 81,25% dapat ditarik kesimpulan bahwa metode support vector machine lebih baik dalam mengklasifikasikan data. Dari segi waktu empat kali percobaan pada kedua metode menggunakan google colab tidak menunjukkan perbedaan waktu yang signifikan, kedua metode melakukan prediksi data dengan cepat.

4.8 Visualisasi Word Cloud

Visualisasi data tes yang sudah melalui tahapan preprocessing menggunakan *word cloud* berfungsi untuk mengetahui kata-kata yang paling sering muncul pada data. Pada penelitian ini, *word cloud* digunakan hanya bertujuan untuk visualisasi tweet berdasarkan kategori sentiment positif dan negatif sehingga dapat diketahui kata yang sering muncul pada setiap sentimennya. Berikut potongan kode program *word cloud* dapat dilihat pada Gambar 4.22:

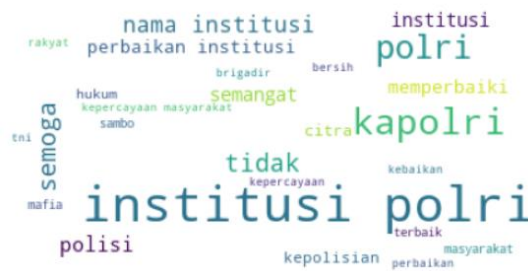
```
wordcloud = WordCloud(
    background_color = 'white',
    stopwords = stopwords,
    min_font_size = 8).generate(comment_words)

# plot the WordCloud image
plt.figure(figsize = (8, 8), facecolor = None)
plt.imshow(wordcloud)
plt.axis("off")

plt.show()
```

Gambar 4.22 Kode program word cloud

Gambar *word cloud* yang menunjukkan kata yang sering muncul di dalam seluruh data yang bersentimen positif pada metode naïve bayes dapat dilihat pada Gambar 4.23 berikut ini:



Gambar 4.23 Word cloud sentimen positif (naïve bayes)

Gambar diatas ini dapat mewakili kata-kata yang sering muncul didalam *tweet* masyarakat yang memiliki sentimen positif mengenai institusi kepolisian republik Indonesia. Berbagai macam ungkapan kata yang ada di antaranya seperti masyarakat yang memberi semangat dan do'a untuk memperbaiki citra institusi kepolisian republik Indonesia. Berikut ini pada Tabel 4.15 dapat dilihat beberapa contoh dari kalimat positif didalam dataset berdasarkan kata yang ada didalam *word cloud*:

Tabel 4.15 Kata dan contoh kalimat positif dalam dataset

No.	Kata	Kalimat di dalam dataset
1.	Semangat	semangat semangat rakyat pastinya mendukung perbaikan institusi
2.	Semoga	semoga pulih kekacauan institusi polri
3.	Kepercayaan	tujuannya untuk mengembalikan tingkat kepercayaan masyarakat seindonesia kepada institusi polri

Berikut gambar word cloud yang menunjukkan kata yang sering muncul di dalam seluruh data yang bersentimen negatif pada metode naïve bayes dapat dilihat pada Gambar 4.24:



Gambar 4.24 Word cloud sentimen negatif (naïve bayes)

Gambar diatas ini menunjukkan kata-kata yang sering kali muncul didalam *tweet* masyarakat yang memiliki sentimen negatif terhadap institusi kepolisian republik Indonesia. Berbagai kalimat atau ungkapan kata ejekan dari berbagai tragedi yang terjadi pada institusi polri seperti pada kata akhlak, jahatnya, boroknya dan oknum merupakan contoh kata yang menunjukkan sentimen negatif. Berikut ini pada Tabel 4.16 dapat dilihat beberapa contoh dari kalimat negatif didalam dataset berdasarkan kata yang ada didalam *word cloud*:

Tabel 4.16 Kata dan contoh kalimat negatif dalam dataset

No	Kata	Kalimat di dalam dataset
1	Akhlak	contoh bejadnya akhlak institusi polri ini
2	Kerusakan	sambosi pembawa kerusakan pada institusi polri
3	Oknum	gila gila institusi polri mulai harus berbenah ini atau tidak mau berbenah karena oknum serupa sambodoh banyak

Gambar *word cloud* yang menunjukkan kata yang sering muncul di dalam seluruh data yang bersetimena positif pada metode SVM dapat dilihat pada Gambar 4.25 berikut ini:



Gambar 4.25 Word cloud sentimen positif (SVM)

Gambar diatas ini dapat mewakili kata-kata yang sering muncul didalam *tweet* masyarakat yang memiliki sentimen positif mengenai institusi kepolisian rebuplik Indonesia. Berbagai macam ungkapan kata yang ada di antaranya seperti masyarakat yang memberi dukungan dan do'a untuk memperbaiki citra institusi kepolisian republik Indonesia. Berikut ini pada Tabel 4.17 dapat dilihat beberapa contoh dari kalimat positif didalam dataset berdasarkan kata yang ada didalam *word cloud*:

Tabel 4.17 Kata dan contoh kalimat positif dalam dataset

No	Kata	Kalimat di dalam dataset
1	Kepercayaan	mantap meningkatkan kepercayaan
2	Semoga	semoga membantu mendorong penyelidikan dugaan rekayasa keadilan nama institusi polri
3	Semangat	semangat jenderal badai menerpa jaga institusi polri

Berikut gambar word cloud yang menunjukkan kata yang sering muncul di dalam seluruh data yang bersentimen negatif pada metode SVM dapat dilihat pada Gambar 4.26:



Gambar 4.26 Word cloud sentimen negatif (SVM)

Gambar diatas ini menunjukan kata-kata yang sering kali muncul didalam *tweet* masyarakat yang memiliki sentimen negatif terhadap institusi kepolisian republik Indonesia. Berbagai kalimat atau ungkapan kata ejekan dari berbagai tragedi yang terjadi pada institusi polri. Berikut ini pada Tabel 4.18 dapat dilihat beberapa contoh dari kalimat positif didalam dataset berdasarkan kata yang ada didalam *word cloud*:

No	Kata	Kalimat di dalam dataset
1	Publik	publik hilang kepercayaan institusi kepolisian
2	Boroknya	orang hebat mengguncang institusi polri membuka boroknya tubuh polri rekayasa pembunuhan
3	Gila	gila institusi isinya sambo jago rekayasa

BAB 5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut:

1. Proses klasifikasi diawali dengan pelabelan manual oleh ahli bahasa, dilanjutkan preprocessing untuk melakukan pemrosesan awal terhadap teks tweet, dilanjutkan dengan ekstraksi fitur dibagi menjadi 2 yaitu bag of words yang berfungsi untuk menghitung frekuensi kemunculan kata dan TF-IDF yang berfungsi untuk mengetahui seberapa penting kata dalam sebuah dokumen, tahap berikutnya adalah pemodelan naive bayes classifier yang berfungsi mengklasifikasikan dataset kedalam kelas positif dan negatif. Kemudian dilakukan pengujian performansi dari model yang telah dibentuk dengan menggunakan confusion matrix. Langkah terakhir adalah dibuat visualisasi word cloud untuk mengetahui kata-kata yang sering kali muncul pada sentimen positif maupun negatif.
2. Pada penelitian ini data yang ada dibagi menjadi dua jenis data yaitu data training dan data testing dan digunakan dengan rancangan empat kali percobaan dengan porsi perbandingan data training dan data testing yang berbeda yaitu sebesar 90%:10%, 80%:20%, 70%:30%, dan 60%:40% dengan hasil nilai akurasi dari keempat percobaan pengujian menggunakan metode support vector machine sebesar 85,5% sedangkan hasil nilai akurasi dari keempat percobaan pengujian menggunakan metode naïve bayes sebesar 81,25%. Secara keseluruhan perbandingan performa metode support vector machine dan naïve bayes menunjukkan hasil bahwa performa dari metode support vector machine lebih baik dalam mengklasifikasikan data.

5.2 Saran

Saran yang dapat diberikan oleh penulis setelah melakukan serangkaian proses penelitian dan bisa digunakan pada pengembangan penelitian lanjutan yang akan dilakukan dimasa mendatang yaitu:

1. Pada penelitian ini, dataset yang didapatkan tidak merata pada setiap kelas sentimen. Data tweet yang bersentimen negatif jauh lebih banyak daripada data yang bersentimen positif. Diharapkan pada penelitian selanjutnya data yang dimiliki lebih merata ke setiap kelas dan mengembangkan jumlah kelas dengan menambah kelas netral. Sehingga penelitian selanjutnya dapat meningkatkan nilai akurasi dan mendapatkan hasil uji sentimen yang lebih baik.

2. Pada penelitian selanjutnya, diharapkan peneliti untuk menggunakan sumber data pada platform sosial media yang lainnya seperti Facebook, Instagram, Youtube, Tiktok dan semacamnya supaya bisa mendapatkan jenis data yang berbeda dan dapat dibandingkan dengan penelitian ini dan diharapkan peneliti dapat mengembangkan data dengan topik yang berbeda.



Daftar Pustaka

- Wahyu Daeni. 2018. Perlindungan Hukum Bagi Penyidik Polri Dalam Melaksanakan Tugas Dan Fungsinya Guna Terwujud Tegaknya Hukum dan Ketertiban Dalam Perspektif Hak Azasi Manusia
- Frizka Fitriana, Ema Utami, and Hanif Al Fatta, 2021. Analisis Sentimen Opini Terhadap Vaksin Covid-19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes
- Akhmad Muzaki and Arita Witanti. 2021. Sentiment Analysis Of The Community In The Twitter To The 2020 Election In Pandemic Covid-19 By Method Naive Bayes Classifier
- Taufik Kurniawan. 2017. Analisis Sentiment Pengguna Twitter Terhadap Media Mainstream Menggunakan Naïve Bayes Classifier Dan Support Vector Machine
- Hennie Tuhuteru, Ade Iriani. 2018. Analisis Sentimen Perusahaan Listrik Negara Cabang Ambon Menggunakan Metode Support Vector Machine Dan Naïve Bayes Classifier
- Brian Laurensz and Eko Sedyono. 2021. Analysis of Public Sentiment on Vaccination in Efforts to Overcome the Covid-19 Pandemic
- Nurdiansyah, Yanuar., Bukhori, Saiful., Hidayat, Rahmad. 2018. Sentimen Analysis System For Movie Review In Bahasa Indonesia Using Naive Bayes Classifier Method. Journal of Physics: Conference Series, 1008(), 012011–. doi:10.1088/1742-6596/1008/1/012011
- Hermawan. 2005 "Pengertian Penelitian Kuantitatif, Karakteristik dan Jenisnya" <https://katadata.co.id/iftitah/ekonopedia/6295749c7fdd7/pengertian-penelitian-kuantitatif-karakteristik-dan-jenisnya>. Penulis: Iftitah Nurul Laily [Diakses pada 23 Agustus 2022]
- Natt Leelawat, Sirawit Jariyapongpaiboonz, Arnon Promjun , Samit Boonyarak, Kumpol Saengtabtim, Ampan Laosunthara, Alfian Kurnia Yudha, Jing Tang. 2021. Twitter Data Sentiment Analysis Of Tourism In Thailand During The COVID-19 Pandemic Using Machine Learning. <https://doi.org/10.1016/j.heliyon.2022.e10894>.
- Dewi Hernikawati. 2021. Kecenderungan Tanggapan Masyarakat Terhadap Vaksin Sinovac Berdasarkan Lexicon Based Sentiment Analysis.

- Saiful Anwar Assaidi., Fatkhul Amin. 2022. Analisis Sentimen Evaluasi Pembelajaran Tatap Muka 100 Persen pada Pengguna Twitter menggunakan Metode Logistic Regression
- Pang, B., & Lee, L.(2008). *Opinion mining and sentiment analysis*. Foundations and Trends in Information Retrieval. <http://dx.doi.org/10.1561/15000000011>
- B., Liu. (2012). *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publisher.
- T., H., Susilo, and S., Rochimah. (2013). "Pengklasifikasian Topik dan Analisis Sentimen Dalam Media Sosial," SNASTI, 2013.
- I., Dilrukshi, and K., d., Zoysa. 2014. "A Feature Selection Method for Twitter News Classification," *International Journal of Machine Learning and Computing*, vol. 4, no. 4.
- P., Koncz and J., Paralic. 2011. "An approach to feature selection for sentiment analysis," *IEEE International Conference*, p. 357–362.
- G., Diaz Pilar., Sanchez., B., Isabel., P., Martin Diego., G., Avila., Jos'e Luis. 2022. A novel flexible feature extraction algorithm for Spanish tweet sentiment analysis based on the context of words. <https://doi.org/10.1016/j.eswa.2022.118817>.
- Veny Amilia Fitri et al. 2019. Sentiment Analysis of Social Media Twitter with Case of AntiLGBT Campaign in Indonesia using Naïve Bayes, Decision Tree, and Random Forest Algorithm. *Procedia Computer Science* 161 (2019) 765–772.
- Aliandu., P. 2015. Sentiment Analysis to determine Accommodation, Shopping and Culinary Location on Foursquare in Kupang City.
- Octaviani. P. A., Wilandari. Y., & Ispriyanti. D. 2014. "Penerapan Metode Klasifikasi Support Vector Machine (SVM) Pada Data Akreditasi Sekolah Dasar (SD) Di Kabupaten Magelang," *Jurnal Gaussian*, vol. 3, pp. 811-820.

LAMPIRAN 1 Surat keterangan validasi data oleh Ahli Bahasa

SURAT KETERANGAN

Yang bertanda tangan di bawah ini, menerangkan bahwa :

Nama : Ranis Lailatus Lisa
Nomer Induk Mahasiswa : 172410103028
Program Studi : Informartika
Fakultas : Ilmu Komputer
Instansi : Universitas Jember
Judul Skripsi : Analisis Sentimen Opini Masyarakat Terhadap Institusi Kepolisian Republik Indonesia Pada Media Sosial Twitter Menggunakan Metode Support Vector Machine dan Naïve Bayes

Telah melakukan tahap validasi data dalam penyusunan skripsi untuk keperluan penggunaan kata Bahasa Indonesia yang baik dan benar pada penelitian yang akan dikerjakan.

Profil validator data sebagai ahli bahasa :

Nama : Nisrina Nur Amalina Windari, S.Pd
Instansi : Alumnus FKIP Bahasa dan Sastra Indonesia UNEJ

Jember, 10 Oktober 2022

Ahli Bahasa,



Nisrina Nur Amalina W., S.Pd

Peneliti,



Ranis Lailatus Lisa

Lampiran 2 Contoh Hasil validasi data oleh Ahli Bahasa

tweet	sentimen
Annashr Maulana RT @Prasetyo9Hh: @KompasTV PC itu Srikandinya Polri... wajib di layani... seperti jaman Kerajaan kita dulu 😄😄😄	Negatif
🐼 RT @arreneuz_s: @kang__L Pak @ListyoSigitP anda SBG Kapolri tidak menjalankan perintah UU dg baik.	Negatif
RT @camilajenie: Terimakasih Joshua Hutabarat, Tragedi Kematianmu Telah Membuka Boroknya Institusi Polri, Membuka Mata Publik Betapa Jahatnya..	Negatif
Pesan Kapolri ke Seluruh Personel Polri: Terus Semangat, karena Institusi Ini Sampai Kapan Pun Harus kita Jaga.. https://t.co/fonzz6ZynJ	Positif
RT @barubuton: Sudah tibalah saatnya Institusi POLRI dilakukan refoasi hh'	Positif
@Pejuang01733866 @jokowi @mohmahfudmd @ListyoSigitP Kerusakan moral warisan zaman orba masih tersisa.\nBegitu sulit. https://t.co/xG1ieDDYoH	Negatif
RT @ResNgawi: Kapolri menyampaikan terima kasih atas dukungan berbagai pihak terhadap institusi Polri yang sedang menghadapi situasi berat.	Positif
@yashadipura Gua percaya ibu ini. Gimana mau profesional institusi polri bila ada oknum yg spt itu.	Negatif
RT @camilajenie: Pak Listyo Sigit Prabowo Tolong Tangkap Idham Azis dan Tito Karnavian Mereka Berdua Juga Banyak Terlibat Di kasus Pembunuhan.	Positif
@DivHumas Polri Sok ngomong SDM, benahi dulu isi institusi kalian pak.'	Negatif
RT @NgkongRoses: Pak @ListyoSigitP jika benar banyak kebobrokan, segera bersihkan institusi Polri.'	Negatif
RT @masguuusss_: @KompasTV Peringatan besar untuk Kepolisian Republik Indonesia Pak Kapolri	Positif
RT @OfficialiNewsTV: Presiden Joko Widodo (Jokowi) akan mencopot langsung Irjen Ferdy Sambo dari Institusi Polri jika putusan banding ditolak'	Positif
@yashadipura Nah Lho... Mulai RAME EX POLISI/POLWAN YANG BERANI CUAP CUAP UNGKAP KASUS2 BERMASALAH DITUBUH POLRI... https://t.co/A7MXGLAeHY	Positif
Ferdy Sambo resmi dipecat dari institusi Polri. Pemecatan atau Pemberhentian Tidak Dengan Hormat (PTDH) terhadap Sambo.	Positif
@ResNgawi Semoga institusi polri, trkhususnya Polres Ngawi tetap Solid, Tegas, berwibawa dalam menjalani Tugas'	Positif
@AimanWitjaksono Gila gila. Institusi polri mulai harus berbenah ini mahhh. Atau gamau berbenah krna oknum serupa sambodoh banyak?'	Negatif
@himawansugeha95 :Mungkin ini salah satu penyebab banyaknya oknum yang tidak bertanggungjawab.. https://t.co/yAR1TyZHB	Negatif
RT @Adi16343002: Sambo...si pembawa kerusakan pd institusi POLRI'	Negatif
dari kasus sambo, ada pertanyaan yg tidak terjawab sampai sekarang dan mungkin selamanya	Negatif
Betul, umumnya begitu. Tpi, kita tetap mempunyai harapan, institusi polri ini akan tetap memproses hukumnya. y kita lihat'	Positif
RT @jannotama: @hipohan @Alie_Fahnoor @ListyoSigitP Jadi ragu sama niat pak kapolri mau benahi institusi polri	Positif
@Dennysiregar7 ckckckckck\nGuna dan manfaat polisi itu sangat banyak... Bhkan oorg2 yg triak tak percaya atau benci	Positif

PelibRaelibatan polisi militer dalam penyidikan anggota polri didasarkan pada sisi stabilitas keamanan	Positif
RT @NenkMonica: Sudah terluka disiram air cuka, Setelah Ferdy Sambo kini Denny Siregar melecehkan Institusi Polri.	Negatif
@OposisiCerdas ___dapat di PASTIKAN para Penyidik Roy Suryo Adalah Jaringan Ferdy Sambo plus Penjilatnya Fadil Imran	Negatif
@ridho_suyono @CNNIndonesia selama sistem penerimaan nya blm bersih, kita ga bisa berharap lebih dr institusi polri.'	Negatif
Walau sy bukan pengamat sejak Polri dipimpin Tito .. institusi ini jadi rusak.	Negatif
RT @ImamWah25258448: @PresidenBingung Sambo cs adalah kerajaan dalam kerajaan polri..Dan Polri menuju kerajaan dalam kerajaan NKRI...	Negatif
@LOVE_AG4EVER @Dennysiregar7 Harus diingat bahwa oknum or geng yg jadi mafia di Polri akan menjadi batu sandungan..	Negatif
RT @KopipagiEncer: @Miss_warawiri Kalau memang benar, berarti institusi polri sudah menjadi mafia. Bukan hanya oknum lagi.'	Negatif
@ardhani83 @inong911 Sangat mustahil unt berubah, bahkan sampai kiamatpun, polisi tetap begitu.	Negatif
Ferdy Sambo secara resmi telah dipecat dari institusi Polri berdasarkan hasil sidang etik.	Negatif
RT @ecosocrights: Kasus Sambo adalah cerminan parahnya korupsi.	Negatif
@kr1t1kp3d45_pro @DonAdam68 @msaid_didu Mekanismenya memang bgitu...Nklw pengangkatan dan pemberhentian seorang Jendral.	Positif
@geloraco setuju banget bang @RamliRizal nkrn si IBLIS SAMBO dkk nya yg berada di INSTITUSI POLRI sdh jd bekingnya.	Negatif
RT @TTuwuh: INILAH YANG SEBENARNYA SUARA WAKIL KITA....SELAMAT BERTUGAS MENGEMBALIKAN MARWAH INSTITUSI POLRI PAK @ListyoSigitP'	Positif
RT @And_Hongg: Penyebaran hoax adalah kriminal , dan sbg negara hukum institusi penegakan hukum ,yakni Polri tdk boleh ada pembiarkan kriminal.	Positif
@OposisiCerdas Hukum nya sudah bagus @mohmahfudmd penegak hukumnyalah yg banyak bermasalah..masih banyak Sambo	Negatif
RT @KompasTV: Tidak Terima Dipecat dengan Tidak Hormat dari Institusi Polri, Ferdy Sambo Ajukan Banding!	Negatif
@ganarmdhn @aniesbaswedan @PT_Transjakarta Wah ada stiker institusi Bapak @ListyoSigitP @DivHumas Polri niihh...	Positif
contoh bejadnya akhlak di institusi polri saat ini	Negatif
@KompasTV Ferdy sambo memang beking judi\n\nKenapa warga Pekanbaru malah ditangkap. Polisi bodoh semua.	Negatif
@Dh4n1eL33 @korantempo @RamliRizal Sambo ini mafia besar terutama judi online yg bersembunyi dibalik institusi polri'	Negatif
@tvOneNews Azabmudidunia penghianat institusi polri.....!'	Negatif
@antdjn1: Sambo & PC msh dibeking terus ya ama babe?????? Awas jgn hrs A43dikawal & diawasi ketat tuh!!!! Terlalu byk sambo2 lain Di institusi polri, apalg menjelang thn 2024!!!!	Negatif
@susipudjiastuti Tuhan pasti menolong negara kita untuk membersihkan nama institusi polri yg di coreng oleh segelintir oknum	Negatif
@DivHumas_Polri perbaiki institusi...masyrkt tdk bodoh'	Positif

@Dennysiregar7 Cocok ni utk film komedi bung ds. Judulnya sambo ngeprank institusi polri'	Negatif
@keepithink_crab @Zulfanbinjai @DivHumas_Polri You know lah bg yg jual senjata kemereka dari institusi yg mana'	Negatif
@RamliRizal Gmn ini pak @ListyoSigitP. apa bnar rakyat percaya dgn institusi Polri'	Negatif
TERUNGKAP Kelemahan Terbesar Polri, Sosok Ini Berani Bongkar Institusi Imbas dari Kasus Sambo	Negatif
@DivHumas_Polri Yang aksi bentak\xc2\xb2 wartawan di menit keberapa ya? dan dia siapa ya ? Publish dong nama dan profil.	Negatif
@CNNIndonesia nah lo..., sebedjat inikah institusi POLRI sekarang..., ngeri maszeh	Negatif
@KompasTV Pantas saja institusi polri gak dipercaya krn pangkat jabatan dipakai kepentingan tertentu termasuk dukung mendukung pilpres lalu.'	Negatif
Pak @ListyoSigitP anda harus membuka kasus pemecatan ex polwan ini dan latar belakang pemecatannya, apa yg disampaikan!	Positif
@detikcom Rusak parah ini institusi polri.. Cc @DivHumas Polri @ListyoSigitP'	Negatif
@faddelm @DivHumas_Polri @jokowi @ListyoSigitP @Kemenkumham_RI Kabarnya lagi di Aceh ada polisi tewas ditembak..	Positif
@setan__alas @DivHumas_Polri Keknya bagus kepolisian ame Tni disatukan lg dlm satu institusi'	Positif
RT @IndonesiaBaikId: Hai SohIB! Dalam institusi Polri ada berbagai pangkat mulai dari tamtama hingga perwira. Seperti apa urutannya?	Positif
@MNW_MNW_MNW Percuma aja di blow up konsorsium lah, kaisar lah tetep modus yg dipake Kapolri soal "itu"	Negatif
@Dennysiregar7 Semoga institusi POLRI terus maju dan semakin mengedukasi masyarakat@Save POLRI'	Positif
@beritajokowi Semoga institusi polri semakin baik kedepannya	Positif
@CNNIndonesia Saya percaya dengan pernyataan mantan POLWAN yang dipecat oleh okmun POLRI. Ini buktik kebobrokan..	Negatif
@cobeh2021 Lah tapi bukan nya sudah dipecat secara tal hormat oleh institusi polri..kok masih ada lagi pemecatan.. https://t.co/XYM3rY9IZx '	Negatif
RT @syafii_ahmad_: @hipohan @ListyoSigitP @ListyoSigitP institusi polri marwah nya sudh jelleek. Nmbah jelek karena kapolri nya tidk taat'	Negatif
@restyca_yah @ListyoSigitP @DivHumas_Polri Dlm kasus Josua, gak ada yg hrs dibela, wong mrk kmplotan pnyamun.	Negatif
Pak @ListyoSigitP jika benar banyak kebobrokan, segera bersihkan institusi Polri. https://t.co/JAmiVTQsW0 '	Negatif
@zarazettirazr Gimana inj jendral!! PRESISI nya di mana??? Ada apa di tubuh polri??	Negatif
@HabibMu50362599 DPR dan PEMERINTAH ambil alih Kepemimpinan Institusi Polri, Demi Citra POLRI Kedepan..	Positif
RT @Zahrah922: Dulu, Di Jaman kepemimpinan Pak Hoegeng , institusi polri sangat di segani dan di hormati oleh rakyat dan luar negeri	Positif
@yashadipura Kalau memang kita cinta institusi Polri , kawal Polri sampai ke Pelosok. Potong palanya kalau ekornya tidak dipotong.'	Positif
@CNNIndonesia kalau dibubarkan aja institusi polri aja gimana, terus dibakilin lagi seperti dlu ABRI,'	Negatif

RT @Fajar_al_fakir: @androsalman @yashadipura beliau bisa dikenakan UU ITE dgn tuduhan mencemarkan nama baik institusi polri'	Positif
RT @YUDHI_Laksana: Jika Ini Benar, Maka Penyematan Slogan "Presisi" Institusi Polri Tidak Patut.\n\nPenegakan HukumTebangPilih.	Negatif
RT @Indones50981375: Ferdy Sambo sudah diberhentikan secara tidak hormat atau PTDH dari institusi Polri.\nFerdy Sambo bakal dipimpin langsung'	Negatif
@yashadipura @maspiyuaja Sebobrok itukah pak institusi ini @DivHumas_Polri @ListyoSigitP'	Negatif
Ingin turut serta membenahi institusi polri tapi di luar lagi banya angin :('	Negatif
RT @BERKARYA_NET: Padahal 17 Agustus gini, jabatan Kadiv Propam itu, gi gagah-gagahnya pas upacara di Mabes, dgn segala atribut penting nya	Positif
@susipudjiastuti Bagus kelompok itu dihukum mati semua, Insya Allah Institusi POLRI akan lebih baik'	Positif
RT @Aryaadi67273603: Pelecehan luar biasa ini terhadap institusi Polri. Kita tunggu aja ketegasan pihak Kepolisian ...'	Negatif
Mahfud MD Sebut Ada Kerajaan Ferdy Sambo di Dalam Institusi Polri	Negatif
RT @Raja_niel: Bongkar terus pantang mundur hancurkan institusi Polri yg bobrok'	Negatif
@KompasTV Kapolri mah nggak punya nyali mengungkap fakta , memang bagusny musnahkan institusi polri dan buat yang baru ,'	Negatif
@DivHumas_Polri saya ga percaya selama institusi polri tdk di reformasi total'	Negatif
RT @ekojhones7: Ayo Polri berani gak nih densi jadi tersangka kayak mas Roy Suryo, apalagi bawa institusi Polri tuh bajunya mosok Kapolri	Negatif
@CNNIndonesia Mosok institusi POLRI kalah ma genk sambo???'	Negatif
Kasus Ferdi Sambo ini sudah melebar kemana-mana opini dan analisisnya.\nSudah mengarah pada Trial by the press.	Negatif
RT @bigwinstoday: Ikut batisan atau bubar. Memang Kapolri harus tegas untuk memperbaiki Institusi Polri.	Positif
RT @NayDonuts: SURAT FS, permintaan maaf kok cuma ke institusi Polri? Masih ngarep empati dr institusi? Ngaca wouy!	Negatif
Ketua Setara Institute Hendardi menyebut kepercayaan publik terhadap institusi Polri sudah mulai kembali pulih.	Positif
Mulai nampak upaya-upaya melemahkan institusi POLRI lewat kasus Brigadir J, oleh pihak-pihak intoleran.	Negatif
Jika Ini Benar, Maka Penyematan Slogan "Presisi" Institusi Polri Tidak Patut.\n\nPenegakan Hukum Tebang.	Negatif
@CNNIndonesia Buktikan, Sambo dan kroninya dihukum berat dan bersihkan Institusi Polri. Rakyat menanti bukti.'	Negatif
@DivHumas_Polri Jualan teroris teros ... maaf pak masalah utama institusi ini gk kelar2 di tambah ini teroris ...	Negatif
@DivHumas_Polri Pak kapolri semoga sehat2 pak..lanjutkan apa yg bpk lakukan demi kebaikan institusi polri dan negara RI'	Positif
@Dennysiregar7 Kasus sambo justru menjadi bukti bahwa ada yg salah di dalam institusi Polri.	Negatif
RT @Hunterjagar3: Pak Kapolri @ListyoSigitP tolong jajarannya yang merusak citra Kepolisian di proses sesuai dengan hukum yang berlaku.	Positif
@Tita83079013 Rakyat tidak terdidik tak punya budaya keadilan hukum masih bisa dimengerti.	Negatif