

Analisis Sentimen terhadap Komisi Pemilihan Umum 2024 di Indonesia melalui Twiter menggunakan Algoritma *Support Vector Machine (SVM)*

Sentiment Analysis of the 2024 General Election Commission in Indonesia through Twiter using the Support Vector Machine (SVM) Algorithm

¹Mhd.Rizki Maulana*, ²Raissa Amanda Putri

^{1,2}Program Studi Ilmu Komputer, Universitas Islam Negeri Sumatera Utara
Jl. Lap. Golf, Kp. Tengah. Kec. Pancur Batu, Kab. Deli Serdang, Sumatera Utara 20353,
Medan, Indonesia

e-mail: *muh.rizkimaulana02@gmail.com¹, raissa.ap@uinsu.ac.id²

(received: 01 May 2024, revised: 13 May 2024, accepted: 15 May 2024)

Abstrak

Indonesia bisa menjadi negara berkeadilan berdasarkan Pancasila dan UUD 1945. Segala sesuatu yang berkaitan dengan pemerintahan telah diatur dalam Undang-Undang. Indonesia sendiri telah melewati tiga fase pemerintahan sejak memproklamasikan Kemerdekaan pada tanggal 17 Agustus 1945. Ketiga tahapan ini, khususnya, orde lama, orde baru, dan reformasi, masih berdampak hingga saat ini. Hal ini berbeda dengan dua tahap sebelumnya, dimana semua pilihan terpusat pada pemerintah. Pada masa perubahan ini, segala pilihan ada di tangan rakyat atau lebih dikenal dengan pemerintahan mayoritas. Dalam penelitian ini metode yang digunakan dalam mengklasifikasi komisi pemilihan umum 2024 adalah Support Vector Machine (SVM). Metode SVM merupakan salah satu kelebihan yang dapat diimplementasikan relatif mudah, karena proses penentuan support vector dapat di rumuskan dalam confusion matrix. Analisis sentimen dan opinion mining adalah bidang studi yang menganalisis pendapat seseorang, sentimen seseorang, evaluasi seseorang. Namun dalam pelaksanaannya banyak sekali hal-hal yang dinilai oleh para pemilih bahwa Pemilu ini berjalan tidak sesuai dengan aturan yang ada. Mulai dari aroma kecurangan banyak disuarakan netizen, ketidaknetralan para penegak hukum terhadap salah satu calon pasangan presiden dan wakil presiden, serta pelaksanaan yang tidak JURDIL atau Jujur dan Adil. Semuanya banyak disuarakan lewat media sosial salah satunya adalah twitter. Diketahui bahwa dari 2084 data yang diperoleh dari tweet mengenai KPU terdapat 29.5% merupakan sentimen yang bersifat positif dan 70.5% sentimen yang bersifat negatif. Hasil klasifikasi sentimen mengenai KPU menggunakan algoritma Support Vector Machine yang sesuai dengan data sebenarnya berjumlah 307 data dari total 418 data uji. Nilai akurasi dari klasifikasi sentimen mengenai KPU menggunakan algoritma Support Vector Machine (SVM) yang diperoleh yaitu 73%.

Kata Kunci : Sentimen ; KPU ; Twiter ; SVM ; Indonesia.

Abstract

Indonesia can be a country of justice based on Pancasila and the 1945 Constitution. Everything related to government has been regulated in the Law. Indonesia itself has gone through three phases of government since proclaiming Independence on August 17, 1945. These three phases, specifically, the old order, the new order, and the reformation, still have an impact today. This is different from the previous two phases, where all choices were centered on the government. In this study, the method used in classifying the 2024 general election commission is Support Vector Machine (SVM). The SVM method is one of the advantages that can be implemented relatively easily, because the process of determining the support vector can be formulated in the confusion matrix. Sentiment analysis and opinion mining are fields of study that analyze a person's opinion, a person's sentiment, a person's evaluation, but in its implementation there are many things that are considered by voters that this election is not in accordance with existing rules. Starting from the aroma of fraud voiced by netizens,

<http://sistemasi.ftik.unisi.ac.id>

the non-neutrality of law enforcers against one of the candidates for the presidential and vice presidential pairs, and the implementation that is not JURDIL or Honest and Fair. Everything is widely voiced through social media, one of which is twitter. It is known that of the 2084 data obtained from tweets about the KPU, 29.5% are positive sentiments and 70.5% are negative sentiments. The results of sentiment classification about KPU using the Support Vector Machine algorithm that matches the actual data amounted to 307 data out of a total of 418 test data The accuracy value of the sentiment classification about KPU using the Support Vector Machine (SVM) algorithm obtained is 73%.

Keywords: *Sentiment; KPU; Twitter; Svm; Indonesia*

1 Pendahuluan

Indonesia bisa menjadi negara berkeadilan berdasarkan Pancasila dan UUD 1945. Segala sesuatu yang berkaitan dengan pemerintahan telah diatur dalam Undang-Undang. Indonesia sendiri telah melewati tiga fase pemerintahan sejak memproklamasikan Kemerdekaan pada tanggal 17 Agustus 1945. Ketiga tahapan ini, khususnya, orde lama, orde baru, dan reformasi, masih berdampak hingga saat ini. Hal ini berbeda dengan dua tahap sebelumnya, dimana semua pilihan terpusat pada pemerintah. Pada masa perubahan ini, segala pilihan ada di tangan rakyat atau lebih dikenal dengan pemerintahan mayoritas. Dimana seluruh rakyat Indonesia berhak mengutarakan kesimpulan dan pertimbangannya, termasuk dalam menentukan pionir bangsa [1].

Bangsa Indonesia belum lama ini telah melakukan hajatan besar demokrasi yaitu PEMILU (Pemilihan Umum). Yang berbeda dengan Pemilu sebelumnya adalah pelaksanaan untuk Pemilihan Calon Legislatif maupun Presiden dilakukan secara bersamaan. Hal ini merupakan amanat Undang-Undang (UU) Nomor 42/2008 tentang Pilpres. Namun dalam pelaksanaannya banyak sekali hal-hal yang dinilai oleh para pemilih bahwa Pemilu ini berjalan tidak sesuai dengan aturan yang ada. Mulai dari aroma kecurangan banyak disuarakan netizen, ketidaknetralan para penegak hukum terhadap salah satu calon pasangan presiden dan wakil presiden, serta pelaksanaan yang tidak JURDIL atau Jujur dan Adil. Semuanya banyak disuarakan lewat media sosial salah satunya adalah twitter [2]. Ada juga dengan bentuk demonstrasi (unjuk rasa) sebagai bentuk kekecewaan dari para sebagian pemilih. KPU (Komisi Pemilihan Umum) selaku penyelenggara hajatan besar ini tak lepas dari tuduhan-tuduhan yang dilontarkan oleh sebagian pemilih yang merasa tidak puas dengan pelaksanaan Pemilu kali ini [3]. Sentimen yang diberikan kepada KPU lewat media sosial merupakan gambaran kekecewaan tersebut, lewat *tweet* yang dishare oleh pengguna aplikasi micro blogging ini bisa kita ambil dan dijadikan analisis sentimen terhadap KPU selaku penyelenggara Pemilu [4][5].

Manfaat yang didapat dengan melakukan sentiment analyst adalah menemukan informasi berharga yang dibutuhkan orang lain dari data yang tidak terstruktur. Maka diharapkan dengan penelitian ini dapat diketahui sentimen para netizen atau pengguna twitter terhadap Komisi Pemilihan Umum. Saat ini penggunaan media sosial adalah terutama *Twitter* di Indonesia telah banyak memberikan dampak besar dalam membangun opini, pandangan, sentimen, dan preferensi politik publik terutama menjelang pemilihan umum Capres 2024 [6]. Pada saat ini tahapan pemilihan umum, para pengguna media sosial atau para pendukung masing-masing kandidat juga memberikan opini mereka melalui media sosial tentang jalannya debat menggunakan hashtag tertentu pada *tweet* yang ada di *Twitter* Untuk mengetahui dan menentukan kecenderungan pengguna *Twitter* dalam memposting *Tweet*, perlu untuk melakukan analisis sentimen [7][8].

Pengguna media sosial yang ada di Indonesia saat ini mencapai 170 juta pengguna pada awal tahun 2021 lalu, dengan media sosial yang mencakup favorit adalah *twitter* dengan pembuktiannya menghasilkan survei bahwa Indonesia menempati posisi ke-6 dengan

peringkat penggunaan *twitter*, selain dari itu Indonesia memiliki pengguna dengan jenjang umur 16 – 64 tahun mencapai tingkat 63,6% bersesuaian dengan sudah layak menjadi pemilih dalam pemilihan umum. Karena banyaknya pendapat yang di sampaikan dari warganet Indonesia Khususnya di media sosial *twitter* ini akan menghasilkan berbagai macam reaksi. Untuk mengetahui dan menentukan kecenderungan pengguna *Twitter* dalam memposting *tweet*, perlu untuk melakukan analisis sentimen [9]. Dalam konteks di media sosial, analisis sentimen adalah bagaimana melakukan analisa terhadap orang yang di media sosial, analisis sentimen adalah bagaimana melakukan analisa terhadap orang yang mengekspresikan opini – opini mereka pada berbagai topik di media sosial [10].

Dalam penelitian ini metode yang digunakan dalam mengklasifikasi komisi pemilihan umum 2024 adalah Support Vector Machine (SVM) [11]. Metode SVM merupakan salah satu kelebihan yang dapat diimplementasikan relatif mudah, karena proses penentuan support vector dapat di rumuskan dalam confusion matrix [12] [13]. Analisis sentimen dan *opinion mining* adalah bidang studi yang menganalisis pendapat seseorang, sentimen seseorang, evaluasi seseorang [14][15][16].

2 Tinjauan Literatur

Hasil penelitian oleh Handayani, [17] tentang Analisis Sentimen Terhadap Bakal Calon Presiden RI 2024 di Twitter menggunakan Algoritma SVM Berdasarkan hasil dan pembahasan yang telah diuraikan di atas, proses analisis sentimen terhadap sentimen twitter mengenai bakal calon presiden republik Indonesia tahun 2024 dapat dilakukan dengan metode Support Vector Machine. Total sentimen awal yang dianalisis berjumlah 1719 buah sentimen. Dengan persebaran 597 data untuk Aniesbaswedan, 627 data untuk Ganjarpranowo dan 495 data untuk Prabowo Subianto. Proses analisis yang dilakukan memiliki nilai akurasi 75% pada sentimen calon 1, 86% pada calon 2 dan 72% pada calon 3. Sehingga rata-rata akurasi untuk ketiga calon adalah sebesar 78%. Hasil akurasi terhadap penerapan metode support vector machine dipengaruhi oleh beberapa hal, salah satunya adalah jumlah dataset dan komposisi jumlah sentimen positif dan negatif.

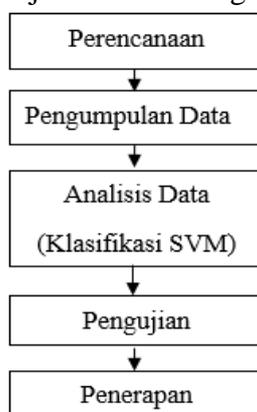
Hasil penelitian oleh Pratama *et al*, [18] tentang Analisis Sentimen *Twitter* Debat Calon Presiden Indonesia Menggunakan Metode *Fined-Grained Sentiment Analysis* yang mendapatkan sentimen yang di tunjukkan publik dunia maya pada saat debat berlangsung cenderung bersentimen positif dari 15.000 percakapan pada saat debat berlangsung. Berdasarkan hasil pada tabel dan barplot pada program R, hasil presentase, baik tweet yang menggunakan hastag #JokowiAminMenangDebat maupun hastag #PrabowoIndonesiaMenang menunjukkan cenderung mengandung sentimen positif lebih banyak dari pada negatif dan netral pada setiap tweet. Untuk sentimen negatif pada kedua hastag, jumlahnya cukup banyak pada setiap tweet yang diposting saat debat berlangsung, walaupun memang tidak sebanyak sentimen positif.

Hasil penelitian oleh Lukmana *et al*, [19] tentang Analisis Sentimen Terhadap Calon Presiden 2019 Dengan *Support Vector Machine* di *Twitter* sentiment positif yang diperoleh Jokowi lebih sedikit dari sentimen negatif. Pelabelan dilakukan pada 10.000 *tweet* yang berkaitan dengan kata kunci “Jokowi”, diperoleh data sebanyak 2.774 *tweet* yang mengandung sentimen positif dan negatif. Sebanyak 1.117 *tweet* berlabel positif dan 1.657 *tweet* berlabel negatif. Proporsi *sentiment* positif yang diperoleh Prabowo lebih sedikit dari sentimen negatif seperti yang Pelabelan dilakukan pada 10.000 *tweet* yang berkaitan dengan kata kunci “Prabowo”, diperoleh data sebanyak 1.632 *tweet* yang mengandung sentimen positif dan negatif. Sebanyak 814 *tweet* berlabel positif dan 818 *tweet* berlabel negatif.

Berdasarkan penelitian diatas memiliki data 1719 buah sentimen maka penilitian ini dibuat dengan perbedaan yang memiliki data 2084 buah sentimen, kemudian akan di analisis bagaimana tingkat akurasi menggunakan metode SVM dengan data yang lebih besar.

3 Metode Penelitian

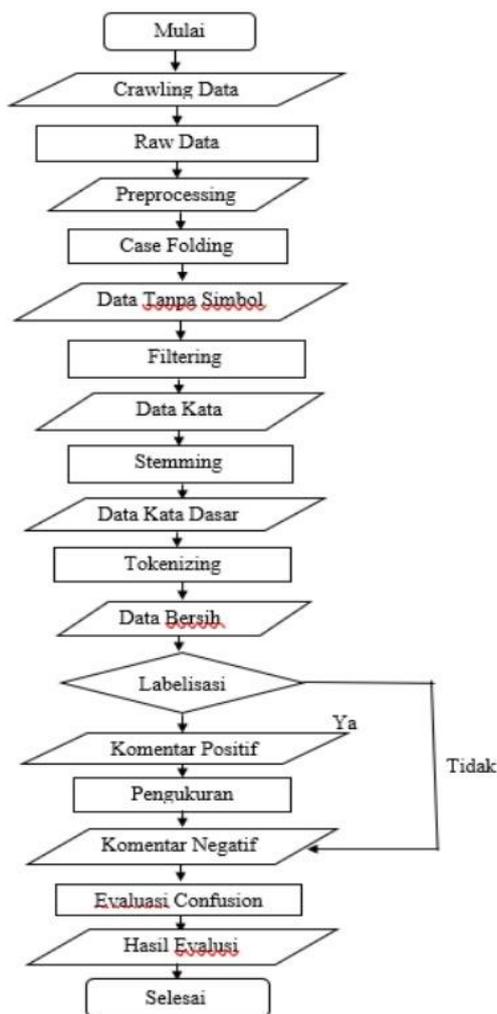
Tempat penelitian guna mendapatkan data penelitian yang akurat dan sebenarnya. Lokasi penelitian ini mengambil data dari Twitter sebagai tempat penelitian karna sesuai dengan topik yang diangkat peneliti mengenai analisis sentimen, selain itu peneliti juga ingin mengetahui bagaimana klasifikasi terhadap komisi pemilihan umum 2024. Sesuai dengan topik yang diangkat oleh peneliti Twitter berperan aktif sebagai media sosial yang digunakan masyarakat dalam menyampaikan opininya sehingga peneliti dapat mengklasifikasikan pendapat masyarakat tentang KPU 2024. Proses penelitian ini diawali dengan perencanaan yaitu menentukan topik yang akan dibahas. Topik penelitian ini yaitu analisis sentimen KPU 2024 di Indonesia melalui twitter menggunakan algoritma Support vektor machine. Penelitian ini menggunakan kuantitatif yaitu penelitian yang menguji teori dengan meneliti Pada kerangka penelitian terdapat didalamnya gambaran langkah-langkah yang dilakukan pada saat melakukan penelitian, agar penelitian yang di lakukan secara sistematis dan tujuan yang di harapkan dapat dicapai [20] dijelaskan dalam gambar 1.



Gambar 1. Kerangka Penelitian

Penelitian ini menggunakan sebanyak 2000 data tweet. dan 1200 data akan menjadi data latih dan 800 data lagi akan dijadikan sebagai data uji. Dalam pengumpulan data *tweet* menggunakan bahasa Indonesia yang bertagar komisipemilihanumum2024. Proses pengambilan data pada penelitian ini menggunakan *library Tweepy*. Penggunaan *library* ini membutuhkan akses OAuth 1 yang bisa didapatkan dari *website* twitter developer. Pengguna OAuth 1 pada *library* ini mendapatkan izin untuk melakukan akses menggunakan API yang tersedia. *Library* ini memanfaatkan API twitter berupa *search* bisa mencari *tweet* yang memiliki kecocokan dengan kata kunci yang diberikan.

Pada penelitian ini analisis data yakni penyusunan data *tweet* sebagai proses awal untuk pengolahan data, proses ini digunakan untuk menggambarkan *tweet* pengguna *twitter* tentang komisi pemilihan umum 2024. Dengan mendisplay data, akan memudahkan untuk meneliti dan menyusun data tersebut. Kemudian akan dilakukan pelabelan kelas dengan ketentuan apabila *tweet* kurang dari 0 maka kelas tersebut negatif dan lebih dari 0 maka positif. Terhadap dalam melakukan analisis sentimen ini dilakukan dengan menggunakan metode SVM dimulai dengan input data yang berupa data latih dan data uji yang kemudian diproses pada tahapan preprocessing hingga proses klasifikasi. Proses ini dilakukan untuk membentuk model yang akan digunakan untuk mengklasifikasi data baru [22] [23]. dijelaskan dalam gambar 2.



Gambar 2. Flowchart Klasifikasi SVM

4. Hasil dan Pembahasan

Pada bagian ini menjelaskan proses analisis data yang di dapatkan pada saat penelitian dengan analisis sebagai berikut:

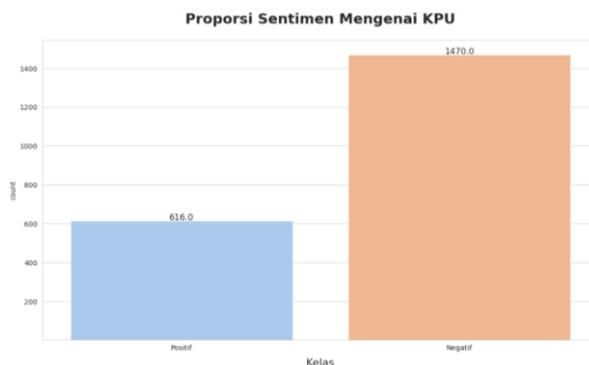
4.1 Pelabelan Sentimen

Pelabelan sentimen dilakukan secara manual, dimana proses pemberian label memanfaatkan penalaran pikiran manusia dalam menentukannya. Pelabelan sentimen manual berarti manusia, biasanya para ahli bahasa, membaca setiap data teks dan memberikan label pada sentimen yang terkandung. Misalnya, kalimat diberi label "positif" jika isinya memuji suatu produk, "negatif" jika isinya berisi kritik. Berikut adalah contoh pelabelan yang dilakukan pada 2 sampel sentimen dijelaskan dalam tabel 1.

Tabel 1. Contoh pelabelan manual

No.	Sentimen Latih	Label
1.	Jaga persatuan dan kesatuan dengan tidak mudah terprovokasi pemberitaan negatif, nantikan pengumuman resmi dari KPU	Positif
2.	sudah telak.....KPU ikut Curang	Negatif

Sentimen 1 dikategorikan positif karna di dalam kalimatnya ada pujian, sementara sentimen 2 termasuk kategori negatif karna terdapat kritikan. Berikut persebaran label sentimen dari keseluruhan *dataset* yang digunakan. dijelaskan dalam gambar 3.



Gambar 3. Persebaran data

4.2 Pembobotan TF-IDF

Tahap selanjutnya setelah melewati tahapan pelabelan dan pembersihan sentimen adalah tahap pembobotan TF-IDF (*Term Frequency-Inverse Document Frequency*), dimana pada tahapan ini menggunakan teknik perhitungan setiap pembobotan kata (*term*) yang ada didalam data dokumen dihitung dari setiap kata dan setiap kata kemudian akan dikalikan idf. Berikut merupakan sampel perhitungan nilai TF dan nilai DF dari 5 buah data latih dan 2 data uji dijelaskan dalam tabel 2, 3 dan 4.

Tabel 2. Sampel Data Latih

Sentimen Latih	Kelas
['hukum', 'kpu', 'indonesia']	Positif
['sahkan', 'provokasi', 'curang', 'ganggu', 'negatif']	Positif
['yakin', 'curang']	Positif
['dapat', 'angket', 'komisi', 'pemilu', 'kampanye', 'partai']	Negatif
['langsung', 'masyarakat', 'tuduh', 'jujur', 'kpu', 'masyarakat', 'umum']	Negatif

Tabel 3. Sampel Data Uji

Sentimen Uji
['waspada', 'tuntut', 'sahkan', 'yakin', 'curang', 'provokasi', 'yakin', 'sebar', 'takut', 'sahkan']
['curang', 'orang', 'umum', 'rugi']

Tabel 4. Nilai DF dari Data Latih

Term	TF					DF
	D1	D2	D3	D4	D5	
hukum	1	0	0	0	0	1
kpu	1	0	0	0	1	2
indonesia	1	0	0	0	0	1
...
umum	0	0	0	0	1	1

Setelah nilai *TF* (*term frequency*) diperoleh, langkah selanjutnya adalah mencari nilai dari *IDF*. Berikut ada rumus untuk menentukan nilai *IDF* dari setiap kata dengan persamaan (1).

$$IDF = \ln \left(\frac{D+1}{df+1} \right) + 1 \dots\dots\dots (1)$$

Berikut adalah sampel dalam menerapkan rumus tersebut pada data pertama :

$$IDF = \ln \left(\frac{D+1}{df+1} \right) + 1 = \ln \left(\frac{5+1}{1+1} \right) + 1 = \ln 3 + 1 = 2.099, \text{ sehingga hasilnya dijelaskan dalam tabel 5.}$$

Tabel 5. Nilai IDF dari Data Latih

Term	DF	IDF
hukum	1	2.099
kpu	2	1,693
indonesia	1	2.099
...

Term	DF	IDF
umum	1	0.699

Setelah nilai TF dan IDF diperoleh barulah dapat dihitung untuk nilai TF-IDF nya. Untuk mencari nilai TF-IDF dengan persamaan (2).

$$W = TF \times IDF \dots\dots\dots (2)$$

Berikut adalah sampel dalam menerapkan rumus tersebut pada data pertama :

$$W = TF \times IDF = 1 \times 2.099 = 2.099, \text{ sehingga hasilnya dijelaskan dalam tabel 6.}$$

Tabel 6. Nilai TF-IDF dari data latihan

Term	TF-IDF				
	D1	D2	D3	D4	D5
hukum	2.099	0	0	0	0
kpu	1.693	0	0	0	1.693
indonesia	2.099	0	0	0	0
...
umum	0	0	0	0	2.099

Selanjutnya nilai TF-IDF dinormalisasikan untuk menyamakan interval dari setiap data, adapun persamaan yang digunakan untuk menormalisasi data adalah dengan persamaan (3).

$$TF_{norm}(t, d) = \frac{TF(t,d)}{\sqrt{\sum_i (TF(t,d))^2}} \dots\dots\dots (3)$$

Sampel dalam menerapkan rumus tersebut, diantaranya :

$$TF_{norm}(t, d) = \frac{TF(t,d)}{\sqrt{\sum_i (TF(t,d))^2}} = \frac{2.099}{\sqrt{(2.099)^2 + (1.693)^2 \dots + (2.099)^2}} = \frac{2.099}{\sqrt{103.98}} = 0.209$$

Berikut adalah hasil perhitungan normalisasi data yang dilakukan dijelaskan dalam tabel 7.

Tabel 7. Normalisasi Data

No.	D1	D2	D3	D4	D5
1	0.206	0	0	0	0
2	0.166	0	0	0	0.166
3	0.206	0	0	0	0
...
20	0	0	0	0	0.206

4.3 Klasifikasi Support Vector Machine

Setelah data dibersihkan dan distrukturkan, langkah selanjutnya adalah melakukan klasifikasi dengan menggunakan algoritma *Support Vector Machine*. Tahap pertama dari proses klasifikasi adalah membagi data ke data latihan dan data uji. Pada penelitian ini menggunakan rasio 8:2 untuk perbandingan data latihan dan data uji. Data latihan digunakan untuk mempelajari karakteristik dan perbedaan kedua kelas positif dan negatif sedangkan data uji digunakan untuk melihat persentase keberhasilan ketika mengklasifikasikan dengan benar.

a. Menghitung Kernel

Pada klasifikasi yang dilakukan, jenis kernel yang digunakan adalah kernel linear karena data yang dimasukkan merupakan data linear. Berikut adalah persamaan yang digunakan untuk menghitung nilai kernel linear dengan persamaan (4).

$$K(x, y) = x \cdot y \dots\dots\dots (4)$$

Berikut contoh representase data pada 5 buah data dijelaskan dalam tabel 8.

Tabel 8. Sampel Representase

	x1	x2	x3	x4	x5
y1	K(x1,y1)	K(x2,y1)	K(x3,y1)	K(x4,y1)	K(x5,y1)
y2	K(x1,y2)	K(x2,y2)	K(x3,y2)	K(x4,y2)	K(x5,y2)
y3	K(x1,y3)	K(x2,y3)	K(x3,y3)	K(x4,y3)	K(x5,y3)
y4	K(x1,y4)	K(x2,y4)	K(x3,y4)	K(x4,y4)	K(x5,y4)
y5	K(x1,y5)	K(x2,y5)	K(x3,y5)	K(x4,y5)	K(x5,y5)

Adapun hasil dari perhitungan kernel linear dari data sampel yang dimiliki sebagaimana yang disajikan dijelaskan dalam tabel 9.

Tabel 9. Perhitungan Kernel Linear

No.	1	2	3	..	18	19	20
1	0.042	0.034	0.042	..	0	0	0
2	0.034	0.055	0.034	..	0.034	0.068	0.034
3	0.042	0.034	0.042	..	0	0	0
...
20	0	0.034	0	..	0.042	0.085	0.042

Contoh perhitungan untuk kolom 1 baris 1.

$$K(x,y) = x * y = (t1d1 * t1d1 + t1d2 * t1d2 + t1d3 * t1d3 + t1d4 * t1d4 + t1d5 * t1d5) = (0.206 * 0.206 + 0*0 + 0*0 + 0*0 + 0*0)$$

$$D_{11} = 0.042$$

Setelah nilai kernel diketahui, langkah selanjutnya adalah perhitungan matrix Hessian. Sebelum melakukan perhitungan matrix Hessian akan ditetapkan beberapa parameter yang antara lain adalah α_i , C, γ , λ dan iterasi maksimum. Berikut uraian dari parameter yang akan digunakan pada tahap perhitungan matrix Hessian dijelaskan dalam tabel 10.

Tabel 10. NilaiParameter

Parameter	Nilai
α_i	0
C	1
γ	0.1
Λ	0.5
iterasi	2

b. Perhitungan Matriks Hessian

Langkah perhitungan matriks Hessian dimulai dengan menginisiasi nilai $\alpha = 0$ kemudian melakukan perhitungan menggunakan persamaan (5).

$$D_{ij} = y_i y_j (K(x_i, x_j) + \lambda^2) \dots\dots\dots (5)$$

Berikut hasil perhitungan matriks Hessian dijelaskan dalam tabel 11.

Tabel 11. Hasil Perhitungan Matriks Hessian

No.	1	2	3	..	18	19	20
1	0.292	0.284	0.292	..	0.25	0.25	0.25
2	0.284	0.305	0.284	..	0.284	0.318	0.284
3	0.292	0.284	0.292	..	0.25	0.25	0.25
...
20	0.25	0.284	0.25	..	0.292	0.335	0.292

Contoh perhitungan nilai matriks Hessian pada kolom 1 baris 1

$$D_{11} = y_i y_j (K(x_i, x_j) + \lambda^2) = 1 * 1(0.042) + 0.5^2 = 0.292$$

c. Sequential Training SVM

Langkah selanjutnya setelah perhitungan nilai matriks hessian adalah melakukan perhitungan *sequential training* dengan menggunakan persamaan (6).

$$E_i \sum_{j=1}^n \alpha_j D_{ij} \dots\dots\dots (6)$$

Pada perhitungan awal, iterasi dimulai dari iterasi 0. Karena nilai α awal masih bernilai 0. Maka hasil perhitungan *sequential training* adalah dijelaskan dalam tabel 12.

Tabel 12. Sequential Training iterasi 0

No.	E_i
1	0
2	0
3	0
...	...
20	0

dilihat dari matriks hessian kolom ke 11 dan 10. Sehingga akan diperoleh nilai kernel dari setiap kelas sebagai berikut.

$$K(x_i, x^+) = \sum \alpha_i y_i D_i$$

$$= (0.2579 * 1 * 0.25) + (0.2555 * 1 * 0.25) + (0.2579 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25669 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25695 * 1 * 0.25) + (0.25837 * 1 * 0.25) + (0.34323 * -1 * -0.292) + (0.25923 * 1 * 0.292) + (0.25923 * 1 * 0.292) + (0.25923 * 1 * 0.292) + (0.34323 * -1 * -0.292) + (0.34323 * -1 * -0.292) + (0.25599 * 1 * 0.25) + (0.25599 * 1 * 0.25) + (0.25599 * 1 * 0.25) + (0.25274 * 1 * 0.25) + (0.25599 * 1 * 0.25)$$

$$= 1.425463$$

$$K(x_i, x^-) = \sum \alpha_i y_i D_i$$

$$= (0.2579 * 1 * -0.25) + (0.2555 * 1 * -0.25) + (0.2579 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25669 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25695 * 1 * -0.25) + (0.25837 * 1 * -0.25) + (0.34323 * -1 * 0.292) + (0.25923 * 1 * -0.292) + (0.25923 * 1 * -0.292) + (0.25923 * 1 * -0.292) + (0.34323 * -1 * 0.292) + (0.34323 * -1 * 0.292) + (0.25599 * 1 * -0.25) + (0.25599 * 1 * -0.25) + (0.25599 * 1 * -0.25) + (0.25274 * 1 * -0.25) + (0.25599 * 1 * -0.25) = -1.42546$$

sehingga hasilnya dijelaskan dalam tabel 17.

Tabel 17. Nilai x^- dan x^+

x^+	x^-
1.425463	1.42546

Nilai $K(x_i, x^+)$ dan $K(x_i, x^-)$ sudah dihitung, langkah selanjutnya adalah menghitung nilai bias dengan menggunakan persamaan (8).

$$b = \frac{1}{2} [\sum_{i=1} \alpha_i y_i K(x_i x^+) + \sum_{i=1} \alpha_i y_i K(x_i x^-)] \dots\dots\dots (8)$$

Dari persamaan diatas, maka diperoleh nilai b sebagai berikut.:

$$b = -\frac{1}{2} [w. x^+ + w. x^-] = -\frac{1}{2} [1.425463 + (-1.42546)] = -\frac{1}{2} (0.000003) = -0.0000015$$

Setelah nilai bias diperoleh, kemudian masuk ketahap pengujian. Langkah pertama pada tahap pengujian adalah melakukan perhitungan nilai TF-IDF pada seluruh di data uji dijelaskan dalam tabel 18.

Tabel 18. Nilai TF-IDF Data Uji

No.	Term	TF	DF	IDF	TF-IDF	Norm
1	waspada	1	0	2.791	2.791	0.2673
2	tuntut	1	0	2.791	2.791	0.2673
3	sahkan	2	1	2.098	4.196	0.4019
...
11	rugi	1	0	2.791	2.791	0.2673

Selanjutnya dilakukan perhitungan kernel dari setiap data uji dengan data latih sebelumnya. Berikut hasil perhitungan kernel data latih terhadap data uji dijelaskan dalam tabel 19.

Tabel 19. Hasil Perhitungan Kernel Data Latih terhadap Data Uji

No.	waspada	tuntut	sahkan	yakin	curang	provokasi	sebar	takut	orang	umum	rugi
1	0.055	0.055	0.083	0.083	0.067	0.055	0.055	0.055	0.055	0.055	0.055
2	0.089	0.089	0.133	0.133	0.108	0.089	0.089	0.089	0.089	0.089	0.089
3	0.055	0.055	0.083	0.083	0.067	0.055	0.055	0.055	0.055	0.055	0.055
...
20	0.055	0.055	0.083	0.083	0.067	0.055	0.055	0.055	0.055	0.055	0.055

Perhitungan untuk kolom 1 baris 1.

$$K(x,y) = (x1 * y1) + (x1 * y2) + .. + (x1 * y5) = (0.2673 * 0.2058) * (0.2673 * 0) * (0.2673 * 0) * (0.2673 * 0) * (0.2673 * 0) = 0.055$$

dirasakan oleh masyarakat. Deskripsi wordcloud ini dapat membantu KPU memahami sentimen publik dan menggunakan informasi tersebut untuk meningkatkan kinerja mereka. Ini adalah kesempatan bagi KPU untuk merespons secara konstruktif terhadap masukan dari masyarakat dan memperkuat kepercayaan publik dalam sistem pemilu Indonesia.

4.4 Implementasi Sistem Analisis Sentimen

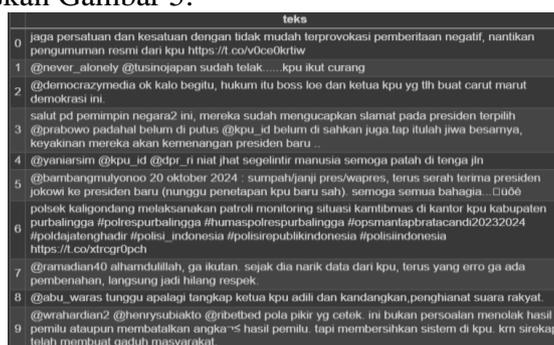
Semua tahapan pada analisis akan diimplementasikan ke dalam suatu sistem yang bertujuan untuk melakukan analisis sentimen. Proses analisis sentimen dilakukan menggunakan Google Collab dan bahasa pemrograman *python*.

1. *Preprocessing Data*

Proses ini menghasilkan dataset baru dengan dalam bentuk file berformat *.csv* yang akan dianalisis. Adapun beberapa tahapan pada *preprocessing* adalah sebagai berikut:

a. *Casefolding*

Pada tahap ini program akan merubah seluruh huruf yang ada pada data menjadi huruf kecil (*lower case*) dijelaskan Gambar 5.



The image shows a list of text entries, likely tweets or social media posts, displayed in a dark-themed interface. The text is as follows:

```
teks
0 jaga persatuan dan kesatuan dengan tidak mudah terprovokasi pemberitaan negatif, nantikan
  pengumuman resmi dari kpu https://t.co/V0ca0krtw
1 @never_alonely @tusingapan sudah telak... kpu ikut curang
2 @democrazymedia ok kalo begitu, hukum itu boss loe dan ketua kpu yg tih buat carut marut
  demokrasi ini.
3 salut pd pemimpin negara ini, mereka sudah mengucapkan selamat pada presiden terpilih
  @prabowo padahal belum di putus @kpu id belum di sahkan juga tap itulah jiwa besarnya,
  keyakinan mereka akan kemenangan presiden baru...
4 @yaniansim @kpu_id @dpr_id niat jhat segelintir manusia semoga patah di tengah jln
5 @bambangmuliono 20 oktober 2024 : sumpah/janji pres/wapres, terus serah terima presiden
  jokowi ke presiden baru (nunggu penetapan kpu baru sah) semoga semua bahagia... 🙏🏻
6 polsek kaligondang melaksanakan patroli monitoring situasi kamtibmas di kantor kpu kabupaten
  purbalingga #polrespurbalingga #humaspolrespurbalingga #opsmantapbratacandi20232024
  #poldajalahadir #polisi_indonesia #polisirepublikindonesia #polisiindonesia
  https://t.co/xtrcgrpch
7 @ramadian40 alhamdulillah, ga ikutan sejak dia narik data dari kpu, terus yang erro ga ada
  pembenahan, langsung jadi hilang respek.
8 @abu_waras tunggu apalagi langkap ketua kpu adili dan kandangkan, penghianat suara rakyat.
  @wrahardian2 @henrysubiako @ribelbed pola pikir yg cetek. ini bukan persoalan menolak hasil
  pemilu ataupun membatalkan angka 's hasil pemilu. tapi membersihkan sistem di kpu. km sirekap
  telah membuat gaduh masyarakat.
```

Gambar 5. Proses *case folding*

Adapun kode python untuk melakukan proses *casefolding* dijelaskan Gambar 6.



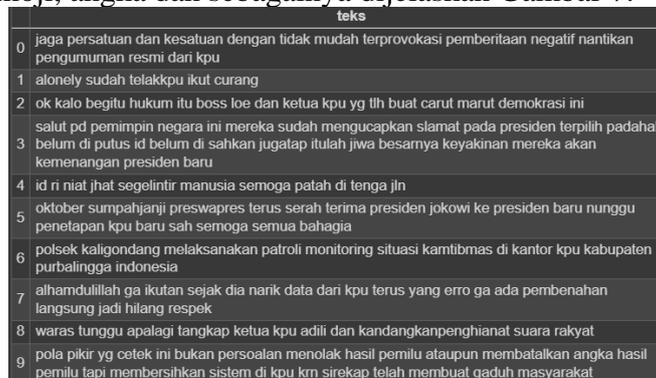
```
[11] def lowercase(teks):
    low = teks.lower()
    return low

kpu.teks = kpu.teks.apply(lambda low:lowercase(str(low)))
```

Gambar 6. Kode untuk *lowercase*

b. *Cleansing*

Pada tahap ini program akan menghilangkan seluruh karakter yang tidak diperlukan, seperti tanda baca, emoji, angka dan sebagainya dijelaskan Gambar 7.



The image shows the same list of text entries as in Gambar 5, but after the cleansing process. The text is as follows:

```
teks
0 jaga persatuan dan kesatuan dengan tidak mudah terprovokasi pemberitaan negatif nantikan
  pengumuman resmi dari kpu
1 alonely sudah telakkpu ikut curang
2 ok kalo begitu hukum itu boss loe dan ketua kpu yg tih buat carut marut demokrasi ini
3 salut pd pemimpin negara ini mereka sudah mengucapkan selamat pada presiden terpilih padahal
  belum di putus id belum di sahkan jugatap itulah jiwa besarnya keyakinan mereka akan
  kemenangan presiden baru
4 id ri niat jhat segelintir manusia semoga patah di tengah jln
5 oktober sumpah/janji pres/wapres terus serah terima presiden jokowi ke presiden baru nunggu
  penetapan kpu baru sah semoga semua bahagia
6 polsek kaligondang melaksanakan patroli monitoring situasi kamtibmas di kantor kpu kabupaten
  purbalingga indonesia
7 alhamdulillah ga ikutan sejak dia narik data dari kpu terus yang erro ga ada pembenahan
  langsung jadi hilang respek
8 waras tunggu apalagi langkap ketua kpu adili dan kandangkan penghianat suara rakyat
9 pola pikir yg cetek ini bukan persoalan menolak hasil pemilu ataupun membatalkan angka hasil
  pemilu tapi membersihkan sistem di kpu km sirekap telah membuat gaduh masyarakat
```

Gambar 7. Proses *cleansing*

Adapun kode python untuk melakukan proses *cleansing* dijelaskan Gambar 8.

```
def repetition(review_text):
    repeat = re.sub(r'\b(\w+)?:\W\b\b+', r'\1', review_text, flags=re.IGNORECASE)
    return repeat
kpu.teks = kpu.teks.apply(lambda word: repetition(word))
def remove_special_signs(text):
    text = text.replace('\t', " ").replace('\n', " ").replace('\u', " ").replace('\', "'")
    text = text.encode('ascii', 'replace').decode('ascii')
    text = ' '.join(re.sub("([@#][A-Za-z0-9+])+(\w+:\w+\S+)", " ", text).split())
    return text.replace("http://", " ").replace("https://", " ")
kpu.teks = kpu.teks.apply(remove_special_signs)

def remove_number(text):
    return re.sub(r"\d+", "", text)
kpu.teks = kpu.teks.apply(remove_number)

def remove_punctuation(text):
    return text.translate(str.maketrans("", "", string.punctuation))
kpu.teks = kpu.teks.apply(remove_punctuation)
```

Gambar 8. Kode untuk *cleansing*

c. *Tokenizing*

Pada tahap ini akan dilakukan pemecahan kalimat menjadi kata per kata. Hal ini dilakukan untuk mempermudah pembersihan teks pada substansi kata, seperti untuk menghilangkan imbuhan, mendeteksi kata non-formal dan mendeteksi kata yang tidak memiliki nilai informasi untuk dianalisis dijelaskan Gambar 9.

	teks
0	[jaga, 'persatuan', 'dan', 'kesatuan', 'dengan', 'tidak', 'mudah', 'terprovokasi', 'pemberitaan', 'negatif', 'nantikan', 'pengumuman', 'resmi', 'dan', 'kpu']
1	[aloney, 'sudah', 'telakkpu', 'ikut', 'curang']
2	[ok, 'kalo', 'begitu', 'hukum', 'itu', 'boss', 'loe', 'dan', 'ketua', 'kpu', 'yg', 'tli', 'buat', 'carut', 'marut', 'demokrasi', 'ini']
3	[salut, 'pd', 'pemimpin', 'negara', 'ini', 'mereka', 'sudah', 'mengucapkan', 'selamat', 'pada', 'presiden', 'terpilih', 'padahal', 'belum', 'di', 'putus', 'id', 'belum', 'di', 'sahkan', 'juga', 'itulah', 'jawa', 'besarnya', 'keyakinan', 'mereka', 'akan', 'kemenangan', 'presiden', 'baru']
4	[id, 'ni', 'niat', 'jhal', 'segelintir', 'manusia', 'semoga', 'patah', 'di', 'tenga', 'jln']
5	[oktober, 'sumpahjani', 'preswapres', 'terus', 'serah', 'terima', 'presiden', 'jokowi', 'ke', 'presiden', 'baru', 'menunggu', 'penetapan', 'kpu', 'baru', 'sah', 'semoga', 'semua', 'bahagia']
6	[polsek, 'kaligondang', 'melaksanakan', 'patroli', 'monitoring', 'situasi', 'kamtibmas', 'di', 'kantor', 'kpu', 'kabupaten', 'purbalingga', 'indonesia']
7	[alhamdulillah, 'ga', 'ikutan', 'sejak', 'dia', 'narik', 'data', 'dan', 'kpu', 'terus', 'yang', 'erro', 'ga', 'ada', 'pembenahan', 'langsung', 'jadi', 'hilang', 'respek']
8	[waras, 'tunggu', 'apalagi', 'langkap', 'ketua', 'kpu', 'adili', 'dan', 'kandangkepghianat', 'suara', 'rakyat']
9	[pola, 'pikir', 'yg', 'cetek', 'ini', 'bukan', 'persoalan', 'menolak', 'hasil', 'pemilu', 'ataupun', 'membatalkan', 'angka', 'hasil', 'pemilu', 'tapi', 'membersihkan', 'sistem', 'di', 'kpu', 'krm', 'sirekap', 'telah', 'membuat', 'gaduh', 'masyarakat']

Gambar 9. Proses *tokenizing*

Adapun kode python untuk melakukan proses *tokenizing* dijelaskan Gambar 10.

```
def tokenizingText(text):
    text = word_tokenize(text)
    return text
kpu.teks= kpu.teks.apply(tokenizingText)
kpu.drop(kpu[kpu.teks.str.len() < 1].index, inplace=True)
kpu.reset_index(drop=True, inplace=True)
```

Gambar 10. Kode untuk *tokenizing*

d. *Filtering(slangword)*

Pada tahap ini program akan mengubah kata-kata non-formal menjadi formal dijelaskan Gambar 11.

	teks
0	[jaga, 'persatuan', 'dan', 'kesatuan', 'dengan', 'tidak', 'mudah', 'terprovokasi', 'pemberitaan', 'negatif', 'nantikan', 'pengumuman', 'resmi', 'dan', 'kpu']
1	[aloney, 'sudah', 'telakkpu', 'ikut', 'curang']
2	[ok, 'kalo', 'begitu', 'hukum', 'itu', 'bos', 'kamu', 'dan', 'ketua', 'kpu', 'yang', 'tli', 'buat', 'carut', 'marut', 'demokrasi', 'ini']
3	[salut, 'pada', 'pemimpin', 'negara', 'ini', 'mereka', 'sudah', 'mengucapkan', 'selamat', 'pada', 'presiden', 'terpilih', 'padahal', 'belum', 'di', 'putus', 'id', 'belum', 'di', 'sahkan', 'juga', 'itulah', 'jawa', 'besarnya', 'keyakinan', 'mereka', 'akan', 'kemenangan', 'presiden', 'baru']
4	[id, 'ni', 'niat', 'jhal', 'segelintir', 'manusia', 'semoga', 'patah', 'di', 'tenga', 'jalan']
5	[oktober, 'sumpahjani', 'preswapres', 'terus', 'serah', 'terima', 'presiden', 'jokowi', 'ke', 'presiden', 'baru', 'menunggu', 'penetapan', 'kpu', 'baru', 'sah', 'semoga', 'semua', 'bahagia']
6	[polsek, 'kaligondang', 'melaksanakan', 'patroli', 'monitoring', 'situasi', 'keamanan dan ketertiban masyarakat', 'di', 'kantor', 'kpu', 'kabupaten', 'purbalingga', 'indonesia']
7	[alhamdulillah, 'enggak', 'ikutan', 'sejak', 'dia', 'menarik', 'data', 'dan', 'kpu', 'terus', 'yang', 'erro', 'enggak', 'ada', 'pembenahan', 'langsung', 'jadi', 'hilang', 'respek']
8	[waras, 'tunggu', 'apalagi', 'langkap', 'ketua', 'kpu', 'adili', 'dan', 'kandangkepghianat', 'suara', 'rakyat']
9	[pola, 'pikir', 'yang', 'cetek', 'ini', 'bukan', 'persoalan', 'menolak', 'hasil', 'pemilu', 'ataupun', 'membatalkan', 'angka', 'hasil', 'pemilu', 'tapi', 'membersihkan', 'sistem', 'di', 'kpu', 'karena', 'sirekap', 'telah', 'membuat', 'gaduh', 'masyarakat']

Gambar 11. Proses *converting slangword*

Adapun kode python untuk melakukan proses *converting slangword* dijelaskan Gambar 12.

```
def convertToSlangword(slangw):
    kamusSlang = eval(open("Slang.txt").read())
    pattern = re.compile(r'\b( ' + ' |' .join (kamusSlang.keys())+r')\b')
    content = []
    for kata in slangw:
        filterSlang = pattern.sub(lambda x: kamusSlang[x.group()],kata)
        content.append(filterSlang.lower())
    slangw = content
    return slangw
kpu.teks = kpu.teks.apply(convertToSlangword)
```

Gambar 12. Kode untuk *converting slangword*

e. *Removing stopwords*

Pada tahap ini program akan menghilangkan seluruh kata yang dianggap tidak bermanfaat atau non-informatif dijelaskan Gambar 13.

	teks
0	[jaga, 'persatuan', 'kesatuan', 'mudah', 'terprovokasi', 'pemberitaan', 'negatif', 'nantikan', 'pengumuman', 'resmi', 'kpu']
1	['alonly', 'telakkpu', 'curang']
2	[ok, 'hukum', 'bos', 'ketua', 'kpu', 'lh', 'carut', 'marut', 'demokrasi']
3	['salut', 'pemimpin', 'negara', 'selamat', 'presiden', 'terpilih', 'putus', 'id', 'sahkan', 'jugatap', 'jiwa', 'besarnya', 'keyakinan', 'kemenangan', 'presiden']
4	[id, 'ni', 'niat', 'jhat', 'segelintir', 'manusia', 'semoga', 'patah', 'tenga', 'jalan']
5	[oktober, 'sumpahjanji', 'preswapres', 'serah', 'terima', 'presiden', 'jokowi', 'presiden', 'menunggu', 'penetapan', 'kpu', 'sah', 'semoga', 'bahagia']
6	[polsek, 'kaligondang', 'melaksanakan', 'patroli', 'monitoring', 'situasi', 'keamanan dan ketertiban masyarakat', 'kantor', 'kpu', 'kabupaten', 'purbalingga', 'indonesia']
7	['alhamdulillah', 'ikutan', 'menarik', 'data', 'kpu', 'erro', 'pembenahan', 'langsung', 'hilang', 'respek']
8	[waras, 'tunggu', 'tangkap', 'ketua', 'kpu', 'adili', 'kandanganpenghianat', 'suara', 'rakyat']
9	['pola', 'pikir', 'cetek', 'menolak', 'hasil', 'pemilu', 'membatalkan', 'angka', 'hasil', 'pemilu', 'membersihkan', 'sistem', 'kpu', 'sirekap', 'gaduh', 'masyarakat']

Gambar 13. Proses *removing stopwords*

Adapun kode python untuk melakukan proses *removing stopwords* dijelaskan Gambar

14.

```
daftar_stopword = stopwords.words('indonesian')
daftar_stopword.extend(['yg', 'sib', 'allah', 'suf', 'pakehat', 'maha', 'dg', 'rt', 'labbh', 'csat', 'rg', 'kadroom', 'oon', 'sfb', 'pks', 'ajam', 'cux', 'khotimah', 'ya', 'allah', 'rg', 'only', 'menang', 'naka', 'da', 'sa', 'iya', 'kearifanullah', 'alidzia', 'di', 'porno', 'khotimah', 'nakula', 'derajat', 'berdapat', 'su', 'sh', 'cipok', 'nyamem', 'delia', 'delto', 'kiron', 'tear', 'an', 'ca', 'discotik', 'diskotik', 'bokap', 'an', 'mentri', 'entri', 'adh', 'agamy', 'als', 'kayak', 'raharjo', 'alibkan', 'nykitti', 'surai', 'ko', 'negri', 'nya', 'kontra', 'yes', 'tu', 'mbuat', 'dekade', 'dgalasan', 'bentar', 'klooon', 'agamy', 'mka', 'deh', 'skali', 'dawa', 'genggang', 'daba', 'ditaiti', 'pempaga', 'luka', 'tuel', 'jip', 'apresiasi', 'vigilansman', 'perangi', 'benge', 'deslarah', 'dini', 'skali', 'gakk', 'poing', 'banyag', 'malagi', 'selai', 'no', 'gak', 'ga', 'yah', 'aj', 'albon', 'ahy', 'ahudalah', 'ah', 'abhh', 'adobhh', 'ad', 'abu', 'abud', 'ahok', 'dun', 'abdul', 'amin', 'amin', 'amir', 'abictu', 'anjg', 'wid', 'hutu', 'amp', 'tbtb', 'sapi', 'spout', 'huliyen', 'pirot', 'ami', 'terkenalpubya', 'yuan', 'adjen', 'grhatana', 'banyasaaktemm', 'fuck', 'fck', 'tae', 'tai', 'planga', 'plongo', 'wei'])
daftar_stopword = set(daftar_stopword)
def stopwordsText(words):
    return [word for word in words if word not in daftar_stopword]
kpu.teks = kpu.teks.apply(stopwordsText)
```

Gambar 14. Kode untuk *removing stopwords*

f. *Stemming*

Pada tahap ini seluruh kata pada seluruh dokumen akan dinormalisasi, artinya akan mengubah seluruh kata menjadi bentuk dasar dengan menghilangkan imbuhan pada kata dijelaskan Gambar 15.

	teks
0	[jaga, 'satu', 'satu', 'mudah', 'provokasi', 'berita', 'negatif', 'nanti', 'umum', 'resmi', 'kpu']
1	['alonly', 'telakkpu', 'curang']
2	[ok, 'hukum', 'bos', 'ketua', 'kpu', 'lh', 'carut', 'marut', 'demokrasi']
3	['salut', 'pimpin', 'negara', 'selamat', 'presiden', 'pilih', 'putus', 'id', 'sahkan', 'jugatap', 'jiwa', 'besar', 'yakin', 'menang', 'presiden']
4	[id, 'ni', 'niat', 'jhat', 'gelintir', 'manusia', 'moga', 'patah', 'tenga', 'jalan']
5	[oktober, 'sumpahjanji', 'preswapres', 'serah', 'terima', 'presiden', 'jokowi', 'presiden', 'tunggu', 'tetap', 'kpu', 'sah', 'moga', 'bahagia']
6	[polsek, 'kaligondang', 'laksana', 'patroli', 'monitoring', 'situasi', 'aman dan tertib masyarakat', 'kantor', 'kpu', 'kabupaten', 'purbalingga', 'indonesia']
7	['alhamdulillah', 'ikut', 'tarik', 'data', 'kpu', 'erro', 'benah', 'langsung', 'hilang', 'respek']
8	[waras, 'tunggu', 'tangkap', 'ketua', 'kpu', 'adil', 'kandanganpenghianat', 'suara', 'rakyat']
9	['pola', 'pikir', 'cetek', 'tolak', 'hasil', 'milu', 'batal', 'angka', 'hasil', 'milu', 'bersih', 'sistem', 'kpu', 'sirekap', 'gaduh', 'masyarakat']

Gambar 15. Proses *stemming*

Adapun kode python untuk melakukan proses *stemming* dijelaskan Gambar 16.

```
factory = StemmerFactory()
stemmer = factory.create_stemmer()
def stemmed_wrapper(term):
    return stemmer.stem(term)
term_dict = {}
for document in kpu.teks:
    for term in document:
        if term not in term_dict:
            term_dict[term] = ''
for term in term_dict:
    term_dict[term] = stemmed_wrapper(term)
def stemmingText(document):
    return [term_dict[term] for term in document]
kpu.teks = kpu.teks.swifter.apply(stemmingText)
```

Gambar 16. Kode untuk *stemming*

2. *Splitting Data*

Tahapan ini adalah tahapan dimana dataset akan dibagi menjadi data latih (*training*) dan data uji (*testing*). Pada analisis yang dilakukan, perbandingan antara data latih dan data uji adalah 8 : 2. Dengan dataset setelah *preprocessing* yang berjumlah 2086 data, 1668 diantaranya akan dijadikan sebagai data latih sedangkan 418 sebagai data uji dijelaskan Gambar 17.

```
(1668, 4021)
(418, 4021)
```

Gambar 17. Proses *Splitting Data*

Adapun kode python untuk melakukan proses *splitting* data dijelaskan Gambar 18.

```
X_train, X_test, y_train, y_test = train_test_split(kpu01['teks'], kpu01['label'], test_size = 0.2, stratify=kpu01['label'], random_state=2)
vectorizer = TfidfVectorizer()
X_train = vectorizer.fit_transform(X_train)
X_test = vectorizer.transform(X_test)
print(X_train.shape)
print(X_test.shape)
```

Gambar 18. Kode untuk *splitting* data

3. Proses Support Vector Machine

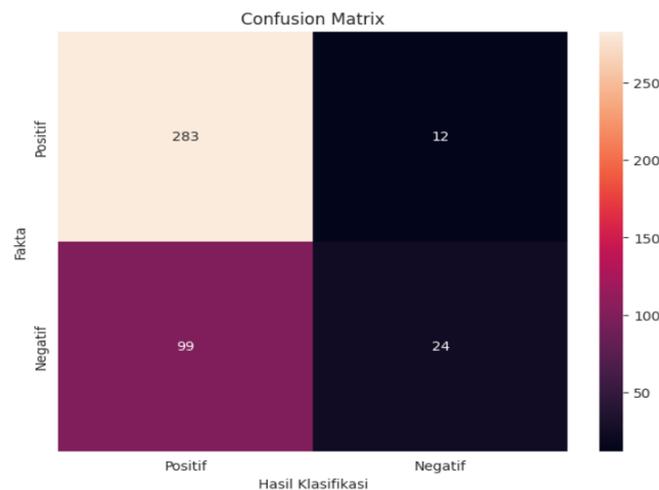
Tahap akhir dari analisis yang dilakukan adalah penerapan algoritma support vector machine. Proses klasifikasi akan melatih model yang telah dibuat pada saat proses pelatihan untuk melakukan prediksi pada data uji yang telah ditetapkan dijelaskan Gambar 19.

```
clf = svm.SVC(kernel="linear")
clf.fit(X_train, y_train)
predict=clf.predict(X_test)
print(mt.classification_report(y_test, clf.predict(X_test)))
```

Gambar 19. Proses *Support Vector Machine*

4.5 Evaluasi Hasil

Setelah selesai melakukan proses pengujian pada algoritma Support Vector Machine maka akan diperoleh hasil. Dimana hasil yang dimaksud adalah label dari data uji yang diperoleh dari model pada saat proses pelatihan. Hasil klasifikasi data uji yang berupa kelas sentimen yang didapatkan dari program akan dibandingkan dengan data kelas sebenarnya sehingga akan diketahui nilai *accuracy*, *precision*, *recall* dan *f1-score* dari model yang digunakan terhadap dataset. Setelah selesai melakukan proses pengujian pada algoritma *Support Vector Machine* maka akan diperoleh hasil. Dimana hasil yang dimaksud adalah label dari data uji yang diperoleh dari model pada saat proses pelatihan. Hasil klasifikasi data uji yang berupa kelas sentimen yang didapatkan dari program akan dibandingkan dengan data kelas sebenarnya sehingga akan diketahui nilai *accuracy*, *precision*, *recall* dan *f1-score* dari model yang digunakan terhadap dataset. Untuk melihat nilai *accuracy*, *precision*, *recall* dan *F1-score* secara keseluruhan dapat menggunakan *classification report*. Berikut *classification report* dan *confusion matrix* dari hasil analisis sentimen yang dilakukan menggunakan algoritma *Support Vector Machine* dijelaskan Gambar 20.



Tabel 20. Confussion matrix

5. Kesimpulan

Berdasarkan hasil akhir dari pemecahan masalah dari penelitian, maka penulis akan menguraikan beberapa kesimpulan. Adapun kesimpulan yang dimaksud antara lain adalah Diketahui bahwa dari 2084 data yang diperoleh dari tweet mengenai KPU terdapat 29.5% merupakan sentimen yang bersifat positif dan 70.5% sentimen yang bersifat negatif. Hasil klasifikasi sentimen mengenai KPU menggunakan algoritma *Support Vector Machine* yang sesuai dengan data sebenarnya berjumlah 307 data dari total 418 data uji. Nilai akurasi dari klasifikasi sentimen mengenai KPU menggunakan algoritma *Support Vector Machine* (SVM) yang diperoleh yaitu 73%.

Referensi

- [1] A. Nur and A. Rusnali, "Debat Capres Cawapres Dalam Perspektif Syariat Islam (Tinjauan Analisis Semiotika Debat Capres Cawapres Pemilu 2019)," *Al-Din: Jurnal Dakwah dan Sosial Keagamaan*, 5(2), 1-15. 2019. doi:<https://doi.org/10.35673/ajdsk.v5i2.582>.
- [2] A. Bin Muhammad Alkatiri, Z. Nadiah, And A. N. S. Nasution, "Opini Publik Terhadap Penerapan New Normal Di Media Sosial Twitter ", *Coverage*, Vol. 11, No. 1, Pp. 19-26, Sep. 2020.
- [3] W. H. Silitonga dan J. I. Sihotang, "Analisis Sentimen Pemilihan Presiden Indonesia Tahun 2019 Di Twitter Berdasarkan Geolocation Menggunakan Metode Naïve Bayesian Classification", *Jurnal TeIKa (Teknologi Informasi dan Komunikasi)*, vol. 9, no. 2, hlm. 115-127, Okt 2019.
- [4] I. Santoso, Windu Gata, and Atik Budi Paryanti, "Penggunaan Feature Selection di Algoritma Support Vector Machine untuk Sentimen Analisis Komisi Pemilihan Umum ", *JURNAL RESTI (Rekayasa Sistem Teknologi Informasi)* , vol. 3, no. 3, pp. 364 - 370, Dec. 2019.
- [5] R. Ferdiana, F. Jatmiko, D. Dwi Purwanti, A. Sekar Tri Ayu, and W. Fajar Dicka, "Dataset Indonesia untuk Analisis Sentimen", *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 8, no. 4, pp. 334-339, Nov. 2019.
- [6] D. Rusdianan and D. Rosiyadi, "Analisa Sentimen Terhadap Tokoh Publik Menggunakan Metode Naïve Bayes Classifier Dan Support Vector Machine," *CESS (Journal of Computer Engineering System and Science)*, Vol. 4 No. 2. 2019.
- [7] R. Hidayatillah, M. Mirwan, M. Hakam, and A. Nugroho, "Levels of Political Participation Based on Naive Bayes Classifier," *IJCCS (Indonesian Journal of*

- Computing and Cybernetics Systems*), vol. 13, no. 1, p. 73, Jan. 2019, doi: 10.22146/ijccs.42531.
- [8] W. H. Silitonga dan J. I. Sihotang, “Analisis Sentimen Pemilihan Presiden Indonesia Tahun 2019 Di Twitter Berdasarkan Geolocation Menggunakan Metode Naïve Bayesian Classification”, *TeIKA*, vol. 9, no. 2, hlm. 115-127, Okt 2019.
- [9] S. Fendyputra Pratama, R. Andrean, and A. Nugroho, “Analisis Sentimen Twitter Debat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis,” *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 4, no. 2, pp. 2541–3619, 2019, doi: 10.31328/jo.
- [10] H. Sujono. “Analisis Sentimen Pada Tweet Dengan Tagar #Bpjsrasarentenir Menggunakan Metode Support Vectore Machine (SVM).” *Skripsi: Universitas Islam Riau*. 2021.
- [11] M. S. Hasibuan and A. Serdano, “Analisis Sentimen Kebijakan Pembelajaran Tatap Muka Menggunakan Support Vector Machine dan Naive Bayes,” *JRST (Jurnal Riset Sains dan Teknologi)*, vol. 6, no. 2, p. 199, Nov. 2022, doi: 10.30595/jrst.v6i2.15145.
- [12] F. Fathonah and A. Herliana, “Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode Naïve Bayes,” *Jurnal Sains dan Informatika*, vol. 7, no. 2, pp. 155–164, Dec. 2021, doi: 10.34128/jsi.v7i2.331.
- [13] M. Furqan, S. Mayang Sari, and S. Sari, “Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia. *Techno.Com*, 21(1), 51-60. doi:<https://doi.org/10.33633/tc.v21i1.5446>
- [14] R. Binsar Sinaga, H. Ramanizar Al Fajri, H. Mubarak, A. Dwi Pangestu, D. Sandya Prasvita, “Analisis Sentimen Pengguna Twitter terhadap Konflik antara Palestina dan Israel Menggunakan Metode Naïve Bayesian Classification dan Support Vector Machine,” *Journal Of Intelligent Systems And Computation*, Vol 2, No. 2. 2021.
- [15] L. Setiyani, M. Wahidin, D. Awaludin, and S. Purwani, “Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review,” *Faktor Exacta*, vol. 13, no. 1, p. 35, Jun. 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [16] J. Homepage, A. Roihan, P. Abas Sunarya, and A. S. Rafika, “IJCIT (Indonesian Journal on Computer and Information Technology) Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper,” 2019.
- [17] A. Handayani and I. Zufria, “Analisis Sentimen Terhadap Bakal Capres RI 2024 di Twitter Menggunakan Algoritma SVM”, *josh*, vol. 5, no. 1, pp. 53-63, Oct. 2023.
- [18] S. Pratama, f. Septian, et al. "Analisis Sentimen Twitter Debat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis." *Journal of Information Technology and Computer Science*, vol. 4, no. 2, 2019, pp. 39-44, doi:[10.31328/jointecs.v4i2.1004](https://doi.org/10.31328/jointecs.v4i2.1004).
- [19] D. T. Lukmana, S. Subanti, & Y. Susanti. “Analisis Sentimen Terhadap Calon Presiden 2019 Dengan Support Vector Machine di Twitter”. *Seminar & Conference Proceedings of UMT*. 2019.
- [20] J. R. Fauzi, “Algoritma dan Flowchart Dalam Menyelesaikan Suatu Masalah,” *Jurnal Hukum Progresif*. 2023.
- [21] S. Mariko, “Aplikasi Website Berbasis Html Dan Javascript Untuk Menyelesaikan Fungsi Integral Pada Mata Kuliah Kalkulus,” *Jurnal Inovasi Teknologi Pendidikan*, vol. 6, no. 1, pp. 80–91, 2019, doi: 10.21831/jitp.v6.1.22280.
- [22] A. Amarulloh, “Analisis Perbandingan Performa Web Service Rest Menggunakan Framework Laravel, Django, Dan Node Js Pada Aplikasi Berbasis Website,” *Journal Of Informatics And Telecommunication Engineering*, Vol 7. No. 2. 2023.