

ANALISA SENTIMEN KOMENTAR VIDEO YOUTUBE DI CHANNEL TVONENEWS TENTANG CALON PRESIDEN PRABOWO SUBIANTO MENGGUNAKAN SUPPORT VECTOR MACHINE

Edi Tohidi¹, Reza Perdana Herdiansyah², Edi Wahyudin³, Kaslani⁴

¹Teknik Informatika, ²Komputerisasi Akuntansi, ³Manajemen Informatika STMIK IKMI CIREBON
Jalan Perjuangan No.10B, Karyamulya Cirebon, Indonesia
editohidi@gmail.com

ABSTRAK

Indonesia merupakan negara demokrasi di mana rakyat memilih presiden melalui Pemilihan Umum Presiden (Pilpres) yang dilakukan 5 tahun sekali. Pada Pilpres 2024, ada 3 kandidat capres yaitu Anies Baswedan, Prabowo Subianto, dan Ganjar Pranowo. Youtube menjadi platform utama masyarakat menyampaikan opini politik. Penelitian ini menganalisis sentimen komentar video Youtube TVOneNews tentang calon presiden Prabowo sebagai capres 2024 dengan SVM. Tujuan penelitian ini adalah mengukur akurasi SVM dalam mengklasifikasi sentimen komentar video Youtube TVOneNews berjudul "Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024" serta melihat sentimen masyarakat terhadap calon presiden prabowo subianto. Metode penelitian ini menggunakan Knowledge Discovery in Database (KDD) yang terdiri dari lima tahapan yaitu Selection, Preprocessing, Transformation, Data Mining dan Interpretation. Datadalam penelitian ini berjumlah 927 komentar yang didapatkan melalui crawling setelah preprocessing tersisa 877 data dengan label positif 530 dan label negatif 346. Hasil penelitian menunjukkan terdapat sebuah perbedaan jumlah label sentimen awal dengan hasil SVM, dimana sentimen positif bertambah dari 530 menjadi 542 dan sentimen negatif berkurang dari 346 menjadi 334. Dan hasil klasifikasi SVM mendapatkan nilai akurasi 85%, presisi 87% dan recall 89%.

Kata kunci: Support Vector Machine, analisa sentimen, Vader.

1. PENDAHULUAN

Indonesia merupakan negara demokrasi dimana rakyat memilih presiden melalui Pemilihan Umum Presiden (Pilpres) yang dilakukan 5 tahun sekali. Pada Pilpres 2024, ada 3 kandidat capres yaitu Anies Baswedan, Prabowo Subianto, dan Ganjar Pranowo. Youtube menjadi platform utama masyarakat menyampaikan opini politik. Penelitian ini menganalisis sentimen komentar video Youtube TVOneNews tentang deklarasi Prabowo sebagai capres 2024 dengan SVM.

Tujuan penelitian ini adalah mengukur akurasi SVM dalam mengklasifikasi sentimen komentar video Youtube TVOneNews berjudul "Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024". Hasil penelitian diharapkan memberi pemahaman mendalam tentang efektivitas SVM mengklasifikasi sentimen masyarakat terhadap Prabowo. SVM merupakan algoritma pembelajaran mesin untuk klasifikasi data yang memiliki akurasi tinggi. SVM menemukan *hyperplane* terbaik antar kelas dengan mempertimbangkan vektor dukungan terdekat *hyperplane* [1]. Kelebihan SVM adalah mengidentifikasi *hyperplane* berbeda untuk memaksimalkan jarak antar kelas [2].

Oleh karena itu, penelitian ini penting untuk mengetahui efektivitas SVM dalam kasus klasifikasi sentimen komentar Youtube tentang politik. Penelitian ini diharapkan bermanfaat bagi masyarakat penggiat politik dan penelitian selanjutnya untuk memahami sentimen publik terhadap kandidat capres berdasarkan analisis komentar Youtube dengan SVM.

Penelitian ini akan menerapkan *library Vader Sentiment* untuk menentukan kalimat positif-negatif dan mengimplementasikan SVM untuk mengklasifikasi data tersebut guna mengukur akurasi algoritma SVM dalam kasus ini.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian [3]. Menjelaskan berbagai metode klasifikasi data seperti *Decision Tree*, *Naïve Bayes*, SVM (*Support Vector Machine*), dan Jaringan Saraf Tiruan. Menurut penelitian tersebut Klasifikasi adalah pengelompokan data berdasarkan ciri atau karakteristik objek.

Selanjutnya Penelitian yang dilakukan oleh [4] menganalisis sentimen tentang akun *Apex Legend* menggunakan VADER Sentimen untuk menentukan sentimen positif, negatif, dan netral. Sentimen yang dimaksud adalah evaluasi atau pendapat yang diekspresikan dalam teks atau dokumen. VADER menilai sentimen frasa dan kalimat menggunakan kamus *lexicon* tanpa melihat konteks lainnya. nilai 0.5 berdasarkan ambang batas skor yang ditentukan. Sentimen dikategorikan positif jika nilai *compound* >0.5, negatif jika <-0.5, dan netral jika -0.5 hingga 0.5.

Selanjutnya Penelitian yang dilakukan oleh [5]. Melakukan analisis sentimen pada layanan internet di Indonesia menggunakan Support Vector Machine dengan fitur seleksi RFE dan metode SMOTE. Pengujian pada dataset MyTelkomsel menghasilkan akurasi SVM tertinggi 99,7% dengan precision 100%,

recall 100% dan f1-score 99,68% menggunakan kernel RBF, SMOTE dan RFE.

Selanjutnya penelitian yang dilakukan [6]. Penelitian tersebut bertujuan membandingkan algoritma SVM dan NBC untuk analisis sentimen pilkada 2019 di Twitter. Hasilnya menunjukkan akurasi NBC 81,7%, recall 81,7%, precision 80%, sedangkan SVM akurasi 80,7%, recall 80,7%, precision 84%.

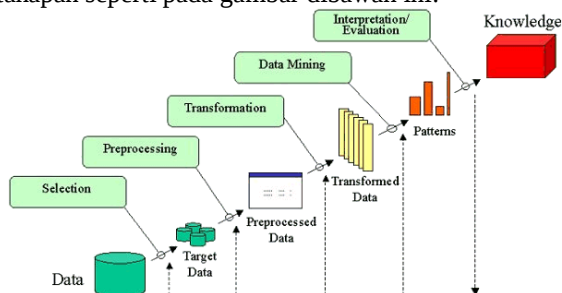
Penelitian yang dilakukan oleh [7]. Pada Penelitian tersebut bertujuan untuk menganalisis persepsi masyarakat terhadap Ustadz Abdul Somad dengan memanfaatkan ulasan di kolom komentar Youtube. Analisis dilakukan mengimplementasikan model algoritma Naive Bayes. Data yang digunakan sebanyak 1000 ulasan komentar di 10 video Youtube tentang Ustadz Abdul Somad. Hasil penelitian menunjukkan mayoritas komentar bersifat positif yaitu sebanyak 67%, kemudian netral 27%, dan negatif 6%. Berdasarkan pengujian didapatkan tingkat akurasi 87%, presisi 91% dan recall 97% untuk algoritma Naive Bayes dalam mengklasifikasi sentimen pada penelitian ini.

Selanjutnya penelitian yang dilakukan oleh [8]. menjelaskan tentang penggunaan metode Naive Bayes untuk mengelompokkan kemandirian siswa/i SMA/MA Double Track berdasarkan data yang telah diubah menjadi bentuk yang pantas untuk diproses mining. Metode ini dipilih karena kebutuhan jumlah data pelatihan yang kecil dan tingkat akurasi yang baik. Tahapan metode Naive Bayes meliputi perhitungan jumlah label, jumlah kasus per kelas, semua variabel kelas dikalikan, dan hasil kelas akan dibandingkan. Hasil akan dinyatakan kedalam bentuk tabel untuk mengetahui probabilitas tertinggi. Hasil dari penelitian tersebut adalah tingkat akurasi sebesar 70% dalam memprediksi status kemandirian siswa menggunakan algoritma Naive Bayes.

3. METODE PENELITIAN

3.1. Knowledge Discovery in Database

Metode penelitian merupakan prosedur sistematis yang disusun untuk melaksanakan penelitian guna mencapai tujuan yang diinginkan. Penelitian ini menggunakan metode Knowledge Discovery in Database (KDD) sebagai metodenya. KDD didefinisikan sebagai proses mengekstraksi informasi bermanfaat dari sekumpulan data menurut [9]. Proses dalam metode KDD terdiri dari beberapa tahapan seperti pada gambar dibawah ini:



Gambar 1. Metode Penelitian.

Pada Gambar 3. Menjelaskan tentang tahap tahap metode KDD, yang dimana urutannya *Selection*, *Preprocessing*, *Transformation*, *Data Mining* dan *interpretation*.

3.1.1. Selection

Pada tahapan selection KDD dalam penelitian ini melakukan crawling data komentar youtube. Data yang diperoleh dari hasil crawling mendapatkan 927 data dari komentar channel youtube TvOneNews yang berjudul “Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024 | Kabar Pemilu tvOne”

3.1.2. Preprocessing

Pada tahap prerocessing pada penelitian ini melakukan pembersihan data yang meliputi beberapa tahap yaitu tahap pengoreksian data, *cleaning*, hapus duplikat data, *tokenize*, *stopword*, *steeming*, *translate* dan *labeling* menggunakan *Vader*.

3.1.3. Transform data

pada tahap transform dalam penelitian ini mengubah data label yang berupa *string* kedalam *numeric*. 1 disumsikan sebagai positif dan 0 diasumsikan sebagai negatif.

3.1.4. Data Mining

Pada tahap data mining pada penelitian ini melakukan klasifikasi menggunakan model algoritma *Support Vector Machine*.

3.1.5. interpretation

pada tahap evaluasi dalam penelitian ini melakukan evaluasi akurasi, presisi dan recall terhadap model menggunakan *Confusion Mtarix*.

3.2. Sumber Data

Data bersumber dari komentar video *Youtube* pada channel TvOneNews yang berjudul “Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024 | Kabar Pemilu tvOne”. Sumber data pada penelitian ini bersifat publik yang diambil dengan teknik *Crawling* menggunakan *API Youtube* dengan mengambil id (wu_wqHg1YLk) pada url vidio yang ingin di *Crawling* komentar nya https://www.youtube.com/watch?v=wu_wqHg1YLk

3.3. Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang bersifat publik. Teknik pengambilan data sekunder yang dilakukan adalah *Crawling*, yaitu mengumpulkan dataset komentar YouTube melalui API. Sebanyak 927 data teks komentar berhasil diambil sejak video diunggah. Teknik crawling ini menyimpan hasil komentar dalam format CSV.

3.4. Teknik Analisis Data

Pada penelitian ini akan menganalisis data akurasi, presisi dan recall algoritma *Support Vector*

Machine dalam mengklasifikasi data. Dibawah ini adalah penjelasan tentang akurasi, presisi, dan recall : Akurasi. Akurasi adalah perbandingan antar data yang terklasifikasikan dengan benar pada keseluruhan data [10]. Dibawah ini adalah rumus menghitung akurasi:

$$akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Presisi

Presisi adalah rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif [11]. Dibawah ini adalah rumus menghitung presisi :

$$presisi = \frac{TP}{TP+FP} \quad (2)$$

Recall

Recall adalah prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif [12]. Dibawah ini adalah rumus untuk mencari recall:

$$recall = \frac{TP}{TP+FN} \quad (3)$$

Penjelasan :

TP (*True Positive*) : data positif yang terklasifikasi dengan benar.

TN (True Negative) : data negatif yang terklasifikasi dengan benar.

FN (False Negative) : data negatif tapi terklasifikasi salah oleh sistem.

FP (False Positive) : data positif tapi terklasifikasi salah oleh sistem.

4. HASIL DAN PEMBAHASAN

Hasil dan pembahasan dari penelitian ini berupa klasifikasi sentimen positif dan sentimen negatif tentang calon presiden prabowo subianto berdasarkan komentar video dari channel youtube TvOneNews yang berjudul “Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024 | Kabar Pemilu tvOne” menggunakan algoritma Support Vector Machine

4.1. Selection

Tahap *Selection* dalam penelitian ini melakukan *Crawling* data dari komentar Youtube menggunakan API Key dari *Google Cloud* yang dieksekusi menggunakan program dengan tools *Google Collab*. Data komentar yang diambil berdasarkan judul video pada channel Youtube TvOneNews yang berjudul “Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024 | Kabar Pemilu tvOne”. Data yang berhasil didapatkan sebanyak 927 data pada tanggal 15 desember 2023. Berikut ini adalah hasil *Crawling* data:

Tabel 1. Hasil *Crawling* Komentar

No	publishedAt	authorDisplayName	Text	likeCount
1.	08/12/2023 17.06	@hsssisopirbentor7399	PRABOWO GIBRAN SUKSES SELALU DNG RAHMAT ALLAH SWT. DI PILPRES 2024.i	0
2.	06/12/2023 02.35	@DodyDody-wl1zj	âœ€Eï, ðœ€Eï, ð	0
3.	04/12/2023 12.24	@pawirotamiyo9131	Halahh Tv One Meneh Nggone Golkar Meneh Huuu Ngapusi	0
4.	03/12/2023 01.34	@globalchannel2547	Ketika rezim orde baru bangkit lagi..maka kalian sejatinya melanjutkan kemiskinan orgtua kalian....dasar generasi tdk berpengalaman sejarah.	0
5.	02/12/2023 23.58	@rahmateffendi744	A Min presiden 2024 â • ðâ • ðâ • ð pilihan kami kaum muda milenialâ • ðâ • ðâ • ð jauh dari Dinasti politik	0

4.2. Preprocessing

Setelah melakukan tahap *selection* selanjutnya dalam penelitian ini akan melakukan tahap *preprocessing*. Tujuan dari *preprocessing* adalah membersihkan, mengorganisir, dan mempersiapkan data sehingga dapat diolah dengan baik dan

menghasilkan analisis yang lebih Tepat. tahap *preprocessing* dalam penelitian ini meliputi pengoreksian data, *cleaning*, hapus duplikat data, *tokenize*, *stopword*, *steeming*, *translate* dan *labeling* menggunakan Vader berikut tahapannya :

Tabel 2. Hasil Koreksi data.

No	Sebelum dikoreksi	Sedudah dikoreksi
1.	PRABOWO GIBRAN SUKSES SELALU DNG RAHMAT ALLAH SWT. DI PILPRES 2024.i	PRABOWO GIBRAN SUKSES SELALU DENGAN RAHMAT ALLAH SWT. DI PILPRES 2024
2.	âœ€Ēĭ, □ âœ€ĒĒĭ, □	Halahh Tv One selalu bohong soal Golkar
3.	Halahh Tv One Meneh Nggone Golkar Meneh Huuu Ngapusi	Ketika rezim orde baru bangkit lagi.. maka kalian sejatinya melanjutkan kemiskinan orang tua kalian... dasar generasi tldak berpengalaman sejarah.
4.	Ketika rezim orde baru bangkit lagi..maka kalian sejatinya melanjutkan kemiskinan orgtua kalian...dasar generasi tdk berpengalaman sejarah.	AMin presiden 2024 â • ðâ • ðâ • □ pilihan kami kaum muda milenialâ • ðâ • ðâ • □ jauh dari Dinasti politik
5.	A Min presiden 2024 â • ðâ • ðâ • □ pilihan kami kaum muda milenialâ • ðâ • ðâ • □ jauh dari Dinasti politik	relawan prabowo.. Buka mata batin kalian.. Apa kasus Mahkamah Konstitusi, dan yg terburu intervensi kasus setnov belum membuka hati dan pikiran kalian ?

Pada Tabel 2 adalah hasil dari koreksi data yang meliputi penghapusan komentar berupa simbol, komentar berupa URL, mengubah komentar bahasa

daerah menjadi bahasa baku, memperbaiki kosakata yang disingkat atau menambahkan kosakata yang kurang.

Tabel 3. Hasil cleaning.

NO	Sebelum <i>dicleaning</i>	Sesudah <i>dicleaning</i>
1.	PRABOWO GIBRAN SUKSES SELALU DENGAN RAHMAT ALLAH SWT. DI PILPRES 2024	prabowo gibran sukses selalu dengan rahmat allah swt di pilpres
2.	Halahh Tv One selalu bohong soal Golkar	halahh tv one selalu bohong soal golkar
3.	Ketika rezim orde baru bangkit lagi.. maka kalian sejatinya melanjutkan kemiskinan orang tua kalian... dasar generasi tidak berpengalaman sejarah.	ketika rezim orde baru bangkit lagi maka kalian sejatinya melanjutkan kemiskinan orang tua kalian dasar generasi tidak berpengalaman sejarah
4.	AMin presiden 2024 • • • • • pilihan kami kaum muda milenial • • • • • jauh dari Dinasti politik	amin presiden pilihan kami kaum muda milenial jauh dari dinasti politik
5.	woi relawan prabowo.. Buka mata batin kalian.. Apa kasus MK, dan yg terbaru intervensi kasus setnov blm membuka hati dan pikiran kalian?	woi relawan prabowo buka mata batin kalian apa kasus mk dan yg terbaru intervensi kasus setnov blm membuka hati dan pikiran kalian

Pada tabel 3 adalah hasil sebelum dan sesudah cleaning. Tahap cleaning meliputi pembuangan karakter HTML seperti
, *quot, menghapus data yang duplikat, menghapus simbol pada data komentar

berupa @, #, emoji, karakter URL, dan menubah huruf kapital menjadi huruf kecil.

Tabel 4. Hasil Tokenize

No	Sebelum Tokenize	Sesudah Tokenize
1.	prabowo gibran sukses selalu dengan rahmat allah swt di pilpres	['prabowo', 'gibran', 'sukses', 'selalu', 'dengan', 'rahmat', 'allah', 'swt', 'di', 'pilpres']
2.	halahh tv one selalu bohong soal golkar	['halahh', 'tv', 'one', 'selalu', 'bohong', 'soal', 'golkar']
3.	ketika rezim orde baru bangkit lagi maka kalian sejatinya melanjutkan kemiskinan orang tua kalian dasar generasi tidak berpengalaman sejarah	['ketika', 'rezim', 'orde', 'baru', 'bangkit', 'lagi', 'maka', 'kalian', 'sejatinya', 'melanjutkan', 'kemiskinan', 'orang', 'tua', 'kalian', 'dasar', 'generasi', 'tidak', 'berpengalaman', 'sejarah']
4.	amin presiden pilihan kami kaum muda milenial jauh dari dinasti politik	['amin', 'presiden', 'pilihan', 'kami', 'kaum', 'muda', 'milenial', 'jauh', 'dari', 'dinasti', 'politik']
5.	woi relawan prabowo buka mata batin kalian apa kasus mk dan yg terbaru intervensi kasus setnov blm membuka hati dan pikiran kalian	['woi', 'relawan', 'prabowo', 'buka', 'mata', 'batin', 'kalian', 'apa', 'kasus', 'mk', 'dan', 'yg', 'terbaru', 'intervensi', 'kasus', 'setnov', 'blm', 'membuka', 'hati', 'dan', 'pikiran', 'kalian']

Pada Tabel 4 adalah hasil sebelum dan sesudah Tokenize. Tokwnize adalah mengubah teks menjadi

unit yang lebih kecil. Token dapat berupa rasa atau kata dari teks.

Tabel 5. Hasil StopWord

NO	Sebelum distopword	Sesudah distopword
1.	['prabowo', 'gibran', 'sukses', 'selalu', 'dengan', 'rahmat', 'allah', 'swt', 'di', 'pilpres']	['prabowo', 'gibran', 'sukses', 'selalu', 'rahmat', 'allah', 'swt', 'pilpres']
2.	['halahh', 'tv', 'one', 'selalu', 'bohong', 'soal', 'golkar']	['tv', 'one', 'selalu', 'bohong', 'soal', 'golkar']
3.	['ketika', 'rezim', 'orde', 'baru', 'bangkit', 'lagi', 'maka', 'kalian', 'sejatinya', 'melanjutkan', 'kemiskinan', 'orang', 'tua', 'kalian', 'dasar', 'generasi', 'tidak', 'berpengalaman', 'sejarah']	['rezim', 'orde', 'baru', 'bangkit', 'kalian', 'sejatinya', 'melanjutkan', 'kemiskinan', 'orang', 'tua', 'kalian', 'dasar', 'generasi', 'berpengalaman', 'sejarah']
4.	['amin', 'presiden', 'pilihan', 'kami', 'kaum', 'muda', 'milenial', 'jauh', 'dari', 'dinasti', 'politik']	['amin', 'presiden', 'pilihan', 'kaum', 'muda', 'milenial', 'jauh', 'dinasti', 'politik']
5.	['woi', 'relawan', 'prabowo', 'buka', 'mata', 'batin', 'kalian', 'apa', 'kasus', 'mk', 'dan', 'yg', 'terbaru', 'intervensi', 'kasus', 'setnov', 'blm', 'membuka', 'hati', 'dan', 'pikiran', 'kalian']	['relawan', 'prabowo', 'buka', 'mata', 'batin', 'kalian', 'apa', 'kasus', 'mk', 'terbaru', 'intervensi', 'kasus', 'setnov', 'blm', 'membuka', 'hati', 'pikiran', 'kalian']

Pada tabel 5 adalah hasil sebelum dan sesudah Stopword.

Tabel 6. Hasil Stemming.

NO	Sebelum distemming	Sesudah distemming
1.	['prabowo', 'gibran', 'sukses', 'selalu', 'rahmat', 'allah', 'swt', 'pilpres']	['prabowo', 'gibran', 'sukses', 'selalu', 'rahmat', 'allah', 'swt', 'pilpres']
2.	['tv', 'one', 'selalu', 'bohong', 'soal', 'golkar']	['tv', 'one', 'selalu', 'bohong', 'soal', 'golkar']
3.	['rezim', 'orde', 'baru', 'bangkit', 'kalian', 'sejatinya', 'melanjutkan', 'kemiskinan', 'orang', 'tua', 'kalian', 'dasar', 'generasi', 'berpengalaman', 'sejarah']	['rezim', 'orde', 'baru', 'bangkit', 'kalian', 'sejati', 'lanjut', 'miskin', 'orang', 'tua', 'kalian', 'dasar', 'generasi', 'alam', 'sejarah']
4.	['amin', 'presiden', 'pilihan', 'kaum', 'muda', 'milenial', 'jauh', 'dinasti', 'politik']	['amin', 'presiden', 'pilih', 'kaum', 'muda', 'milenial', 'jauh', 'dinasti', 'politik']
5.	['relawan', 'prabowo', 'buka', 'mata', 'batin', 'kalian', 'apa', 'kasus', 'mk', 'terbaru', 'intervensi', 'kasus', 'setnov', 'blm', 'membuka', 'hati', 'pikiran', 'kalian']	['rawan', 'prabowo', 'buka', 'mata', 'batin', 'kalian', 'apa', 'kasus', 'mk', 'baru', 'intervensi', 'kasus', 'setnov', 'blm', 'buka', 'hati', 'pikir', 'kalian']

Pada tabel 6 adalah hasil sebelum dan sesudah proses *stemming*. *Stemming* adalah kata-kata umum

yang sering muncul dan cenderung tidak memberikan makna.

Tabel 7. Hasil Translate.

NO	SEBELUM DITRANSLATE	SESUDAH DITRANSLATE
1.	prabowo gibran sukses selalu rahmat allah swt pilpres	Prabowo Gibran's success is always the grace of Allah SWT in the presidential election
2.	tv one selalu bohong soal golkar	TV One always lies about Golkar
3.	rezim orde baru bangkit kalian sejati lanjut miskin orang tua kalian dasar generasi alam sejarah	the new order regime arose, you really continued to be poor, your parents were the basis of the natural generation of history
4.	amin presiden pilih kaum muda milenial jauh dinasti politik	Amen, the president chooses millennial youth from a political dynasty
5.	rawan prabowo buka mata batin kalian apa kasus mk baru intervensi kasus setnov blm buka hati pikir kalian	Rawan Prabowo, open your mind's eyes, what is the new MK case that intervened in the Setnov case, haven't you opened your mind yet?

Pada tabel 7 adalah hasil sebelum dan sesudah data ditranslate. Karena pada penelitian ini akan

menggunakan *Vader*. Karena *VADER* menggunakan kamus lexicon berbahasa inggris.

Tabel 8. Hasil pelabelan *Vader*

NO	Text	compound_score	label
1.	Prabowo Gibran's success is always the grace of Allah SWT in the presidential election	5,263194444	positif
2.	TV One always lies about Golkar	-0.4215	negatif
3.	the new order regime arose, you really continued to be poor, your parents were the basis of the natural generation of history	-0.1531	negatif
4.	Amen, the president chooses millennial youth from a political dynasty	00.00	negatif
5.	Rawan Prabowo, open your mind's eyes, what is the new MK case that intervened in the Setnov case, haven't you opened your mind yet?	00.00	negatif

Pada tabel 8 adalah hasil pelabelan menggunakan *Vader*. *Vader* akan menghasilkan compound score terhadap teks komentar. Pada penelitian ini jika *compound score* > 0.5 akan diberi positif termasuk juga *compound score* bernilai >= 0.5 dan <= 0.0 akan diberi label positif. jika *compound score* < -0.5 akan diberi label negatif termasuk juga nilai *compound score* >= 0.5 dan == 0.0.

4.3. Transformation

Setelah melakukan tahap preprocessing, selanjutnya peneliti melakukan tahap transformation. Pada tahap transformation dalam penelitian ini mengubah data label yang berupa *string* positif dan negatif menjadi numerik. Label positif akan diubah menjadi 1 dan label negatif diubah menjadi 0. Berikut hasil nya :

Tabel 9. Hasil *Transform Text category to Numeric*.

label	label_encoded
positif	1
negatif	0
negatif	0
negatif	0
negatif	0

Pada tabel 9 adalah hasil sebelum dan sesudah *ditransform text category to Numeric*.

4.4. Data Mining

Pada tahap *data mining*, proses klasifikasi data dilakukan dengan menerapkan algoritma Support Vector machine(SVM) terhadap data text hasil preprocessing dan label yang sudah dikonvert kedalam *numerik*. Tahap klasifikasi juga menghasilkan metrik

evaluasi seperti skor presisi, laporan klasifikasi dan *confusion matrix*.

```

Accuracy: 0.85

Confusion Matrix:
[[53 14]
 [12 97]]

Classification Report:

```

	precision	recall	f1-score	support
0	0.82	0.79	0.80	67
1	0.87	0.89	0.88	109
accuracy			0.85	176
macro avg	0.84	0.84	0.84	176
weighted avg	0.85	0.85	0.85	176

Gambar 2. Hasil program klasifikasi SVM.

Gambar 2 hasil skor akurasi, laporan klafikasi dan tabel *Confusion Matrix*.

Tabel 10. Hasil klasifikasi algoritma SVM

label	label_encoded	hasil_klasifikasi_svm
positif	1	positif
negatif	0	negatif
negatif	0	negatif
negatif	0	negatif
negatif	0	negatif

Pada tabel 10 adalah hasil dari klasifikasi algoritma Support Vector Machine berdasarkan kolom teks yang sudah *dipreprocessing* dan kolom yang sudah di *encoding*.

4.5. Interpretation

Setelah tahap *data mining* akan melakukan *interpretation* atau *evaluation*. Pada tahapan *evaluation* dalam penelitian ini menghitung hasil klasifikasi algoritma *Support Vector Machine* menggunakan metode *Confusion matrix* secara manual. *Confusion matrix* adalah sebuah tabel yang menunjukkan kinerja model dalam prediksi yang benar (*True*) dan salah (*False*). Dengan *confusion matrix* peneliti dapat mengitung nilai akurasi, presisi dan *recall*. Berikut adalah hasil perhitungan menggunakan rumus akurasi, presisi dan *recall* secara manual :

Akurasi :

$$\begin{aligned}
 \text{akurasi} &= \frac{97 + 53}{97 + 53 + 14 + 12} \\
 &= \frac{150}{176} = 0,85 \\
 &= 0,85 * 100 = 85\%
 \end{aligned}$$

Presisi :

$$\begin{aligned}
 \text{presisi} &= \frac{97}{97 + 14} \\
 &= \frac{97}{111} = 0,87 \\
 &= 0,87 * 100\% = 87\%
 \end{aligned}$$

Recall :

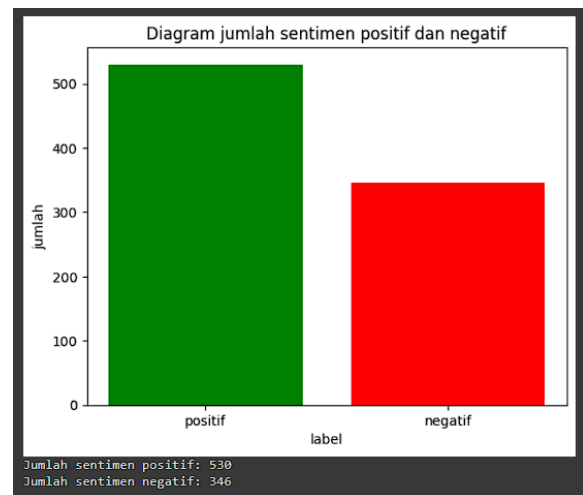
$$\begin{aligned}
 \text{recall} &= \frac{97}{97 + 12} \\
 &= \frac{97}{111} = 0,8 \\
 &= 0,89 * 100\% = 89\%
 \end{aligned}$$

Hasil evaluasi perhitungan manual akurasi, presisi dan *recall* menggunakan *confusion matrix* mendapatkan nilai akhir dari akurasi 85%, presisi 87% dan *recall* 89%.

4.6. PEMBAHASAN

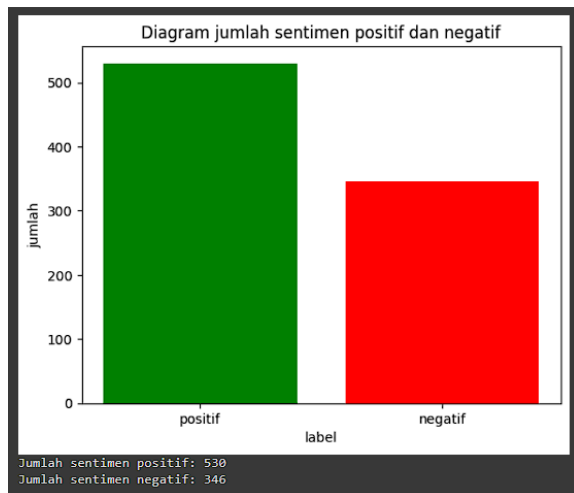
Pembahasan hasil analisis yang dilakukan dengan proses pengumpulan data menggunakan teknik *crawling* pada komentar vidio channel youtube TvOneNews yang berjudul “Relawan dari Berbagai Daerah Deklarasikan Prabowo sebagai Capres 2024 | Kabar Pemilu tvOne”. menghasilkan data berjumlah 927 sebelum *diprocessing* ketika data sudah *diprocessing* terjadi penyusutan data menjadi 877 data. Data diberi label sentimen positif dan negatif berdasarkan *compound score* yang didapatkan menggunakan *vader sentimen*.

Jika *Compound score* lebih dari 0.5 atau lebih dari sama dengan 0.5 maka diberi label positif dan jika *compound score* lebih dari -0.5 atau kurang dari -0.5 maka diberi label negatif dan *compound score* yang bernilai 0.0 maka dilabeli negatif. Terdapat sebuah perbedaan sentimen awal dengan sentimen hasil prediksi algoritma *Support Vector Machine*. Label awal memiliki jumlah sentimen positif sebesar 530 dan label sentimen negatif berjumlah sebesar 346. Setelah proses analisa menggunakan algoritma *Support Vector Machine* hasil label positif 542 dan label negatif 334. berikut hasil pelabelan sentiment awal menggunakan *vader* :



Gambar 3. Hasil sentimen Vader

Gambar 3 adalah hasil sentimen awal dengan *Vader*.



Gambar 4. hasil prediksi sentimen SVM

Pada gambar 4 adalah hasil grafik pelabelan sentimen menggunakan *Support Vector Machine*. Jumlah sentimen positif berjumlah 530 dan jumlah sentimen negatif 346 menggunakan rasio 80:20 untuk pembagian data test dan data uji.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil dari penelitian ini, dapat disimpulkan algoritma *Support Vector Machine* (SVM) telah berhasil diimplementasikan untuk analisis sentimen pada komentar video YouTube tentang Calon Presiden Prabowo Subianto di channel YouTube TVOneNews. Dari 927 komentar yang diambil dengan teknik crawling, terdapat 530 komentar positif dan 346 komentar negatif berdasarkan compound score dari VADER. Setelah dianalisis menggunakan SVM, didapatkan 542 komentar positif dan 334 komentar negatif. Hal ini menunjukkan bahwa sebagian besar komentar masyarakat terhadap pencalonan Prabowo Subianto adalah positif, meskipun ada komentar negatif. Dari sisi performa, SVM menghasilkan akurasi 85%, presisi 87%, dan recall 89%. Dengan demikian dapat disimpulkan bahwa SVM cukup efektif untuk mengklasifikasikan sentimen komentar YouTube tentang topik politik seperti pencalonan presiden. Metode ini dapat memberikan gambaran tentang pendapat publik di media sosial terhadap tokoh politik tertentu. Pada penelitian ini menyadari bahwa penelitian ini masih memiliki kelemahan dan keterbatasan tertentu. Oleh karena itu, untuk penelitian di masa depan, peneliti mempertimbangkan beberapa saran berikut ini: Pertama, dapat mencoba metode analisis sentimen selain *Support Vector Machine* (SVM) seperti *Naive Bayes*, *Random Forest*, atau *Deep Learning* untuk membandingkan performa dan akurasi metode lainnya. Kedua, dapat melakukan optimasi parameter SVM seperti kernel, C, dan gamma menggunakan teknik optimasi *Grid Search* atau *Random Search* untuk memastikan model SVM yang dihasilkan memberikan performa terbaik. Dengan mempertimbangkan saran ini, diharapkan penelitian

selanjutnya dapat memberikan hasil yang lebih baik lagi dalam hal akurasi dan performa metode analisis sentimen komentar Youtube.

DAFTAR PUSTAKA

- [1] A. Sentimen, C. Twitter, and T. Penerapan, "G-Tech : Jurnal Teknologi Terapan," vol. 6, no. 2, pp. 387–394, 2022.
- [2] Z. Hao, L. Shaohong, and S. Jinping, "Unit model of binary SVM with DS output and its application in multi-class SVM," *Proc. - 2011 4th Int. Symp. Comput. Intell. Des. Isc. 2011*, vol. 1, no. 2, pp. 101–104, 2011, doi: 10.1109/ISCID.2011.34.
- [3] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama, Muhammad Fathony Akbar, "Metode-metode Klasifikasi," *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 134, 2018.
- [4] D. Abimanyu, E. Budianita, E. P. Cynthia, F. Yanto, and Y. Yusra, "Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 3, pp. 423–431, 2022, doi: 10.32672/jnkti.v5i3.4382.
- [5] D. Saputra and M. R. Pribadi, "Analisis Sentimen Masyarakat terhadap Layanan Provider Internet di Indonesia menggunakan SVM," *MDP Student Conf.*, vol. 2, no. 1, pp. 32–41, 2023, doi: 10.35957/mdp-sc.v2i1.4554.
- [6] E. S. R. Br.Situmorang, M. K. Anam, R. Rahmadden, and A. N. Ulfah, "Perbandingan Algoritma Svm Dan Nbc Dalam Analisa Sentimen Pilkada Pada Twitter," *CSRID (Computer Sci. Res. Its Dev. Journal)*, vol. 13, no. 3, p. 169, 2021, doi: 10.22303/csr.13.3.2021.169-179.
- [7] H. A. R. Harpizon, R. Kurniawan, Iwan Iskandar, R. Salambue, E. Budianita, and F. Syafria, "Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes," *JNKTI (Jurnal Nas. Komputasi dan Teknol. Informasi)*, vol. 5, no. 1, pp. 131–140, 2022, [Online]. Available: <http://repository.uin-suska.ac.id/59746/>
- [8] M. Walid and F. Halimiyah, "Klasifikasi Kemandirian Siswa SMA/MA Double Track Menggunakan Metode Naive Bayes," *J. ICT Inf. Commun. Technol.*, vol. 22, pp. 190–197, 2022.
- [9] T. Silwattanansarn, "Data Mining and Its Applications for Knowledge Management: A Literature Review from 2007 to 2012," *Int. J. Data Min. Knowl. Manag. Process.*, vol. 2, no. 5, pp. 13–24, 2012, doi: 10.5121/ijdkp.2012.2502.
- [10] R. R. Khaerullah, N. Suarna, and O. Nurdian, "Analisa Pengelompokan Dataset Komputer Menggunakan Algoritma X-Means," *J. Inform. dan Teknol. Inf.*, vol. 1, no. 2, pp. 125–132, 2023, doi: 10.56854/jt.v1i2.135.

- [11] T. Prasetya, I. Ali, C. L. Rohmat, and O. Nurdiawan, "Klasifikasi Status Stunting Balita Di Desa Slangit Menggunakan Metode K-Nearest Neighbor," *INFORMATICS Educ. Prof. J. Informatics*, vol. 5, no. 1, p. 93, 2020, doi: 10.51211/itbi.v5i1.1431.
- [12] Syafi'i, O. Nurdiawan, and G. Dwilestari, "Penerapan Machine Learning Untuk Menentukan Kelayakan Kredit Menggunakan Metode Support Vektor Machine," *J. Sist. Inf. dan Manaj.*, vol. 10, no. 2, pp. 1–6, 2022.