

---

## Perbandingan Analisis Sentimen pada Aplikasi SIREKAP dengan aplikasi SITUNG di Media Sosial X Menggunakan Algoritma *Support Vector Machine*

Taufik Akbar<sup>1</sup>, Rahmi Imanda<sup>2</sup>

akbar.taufik062@gmail.com<sup>1</sup>, rahmi.imanda@uhamka.ac.id<sup>2</sup>

<sup>1,2</sup> Universitas Muhammadiyah Prof. Dr. HAMKA

---

Informasi Artikel	Abstrak
Diterima : 31 Mei 2024 Direvisi : 7 Jul 2024 Disetujui : 8 Agu 2024	<i>Indonesia adalah negara yang menganut sistem demokrasi, yang mana untuk pemilihan pemimpin negara harus melalui PEMILU. Sampai saat ini KPU sudah membuat 2 aplikasi untuk membantu dalam PEMILU yaitu Sistem Informasi Pemilihan Suara (SITUNG) dan Sistem Informasi Rekapitulasi Elektronik (SIREKAP). Namun ternyata pelaksanaan pemilu dengan kedua aplikasi ini menimbulkan berbagai macam sentimen dari masyarakat. Penelitian ini bertujuan untuk mengetahui sentimen masyarakat terhadap SITUNG dan SIREKAP serta apakah jumlah sentimen negatif SIREKAP lebih banyak daripada SITUNG. Penelitian ini mengambil data dari media sosial X dan menggunakan SVM untuk implementasi algoritma. Jumlah sentimen yang digunakan untuk SITUNG sebanyak 1469 yang terdiri dari 832 sentimen positif serta 637 sentimen negatif, lalu untuk SIREKAP sebanyak 1897 yang terdiri dari 1092 sentimen positif dan 805 sentimen negatif. Hasil nilai precision, recall dan f1-score pada SITUNG sekitar 70%, 69%, 69% sedangkan SIREKAP sekitar 67%, 66%, 66%.</i>
<b>Kata Kunci</b>	
Analisis Sentimen, Situng, Sirekap, Support Vector Machine	

---

---

Keywords	Abstract
Sentiment Analysis, Situng, Sirekap, Support Vector Machine	<i>Indonesia is a democratic system, the election must through PEMILU. Until now, the KPU has made 2 applications for the elections, the Vote Selection Information System (SITUNG) and the Electronic Recapitulation Information System (SIREKAP). However, it turns out that the implementation of elections with these two applications raises various kinds of sentiments from the community. This research aims to determine the public sentiment towards SITUNG and SIREKAP and whether the number of negative sentiments of SIREKAP is more than SITUNG. This research takes data from X social media and uses SVM for algorithm. The number of sentiments used for SITUNG is 1469 which consists of 832 positive sentiments and 637 negative sentiments, then for SIREKAP is 1897 which consists of 1092 positive sentiments and 805 negative sentiments. The results of the precision, recall and f1-score values on SITUNG are around 70%, 69%, 69% while SIREKAP is around 67%, 66%, 66%.</i>

---

## A. Pendahuluan

Indonesia merupakan salah satu negara yang mengikuti sistem demokrasi. Salah satu pilar demokrasi adalah suatu pemerintahan dipimpin oleh rakyat, dari rakyat, dan untuk rakyat [1]. Jika suatu negara menganut sistem demokrasi maka pemilihan pemerintahan diselenggarakan Pemilihan Umum (Pemilu). Tujuan utama dari pemilu adalah memberikan kesempatan kepada warga negara untuk menyampaikan suara mereka dan memilih para pemimpin yang akan mewakili mereka di pemerintahan [2]. Terdapat enam asas pemilu di Indonesia yaitu langsung, umum, bebas, rahasia, jujur dan adil (LUBER JURDIL). Di era perkembangan teknologi dan informasi seperti sekarang dengan sistem Pemilu, perlu dilakukan se-penting mungkin agar masyarakat dapat percaya bahwa perubahan sistem Pemilu dilakukan untuk memudahkan masyarakat untuk berpartisipasi dalam Pemilu [3].

Pada pemilu tahun 2019 Sistem Informasi Perhitungan Suara (SITUNG) menjadi awal dari proses perubahan penyelenggaraan pemilu, dari manual ke digital. Tujuan dari SITUNG dibuat untuk mempermudah masyarakat dalam mencari informasi terkait hasil pemilu 2019 dan mempermudah kinerja komisi pemilihan umum (KPU) dalam menghitung suara. Sistem kerja SITUNG yaitu menggunakan formulir C1 asli sebagai data utama, kemudian formulir C1 tersebut diserahkan ke Komisi Pemilihan Umum (KPU) Kabupaten/Kota dan diupload ke server KPU melalui SITUNG. Pada pemilu 2019, SITUNG dianggap sebagai salah satu penyebab keributan pasca-Pemilu 2019 karena ada klaim bahwa perhitungan di SITUNG terdapat banyak kesalahan data [4].

Pada pemilu tahun 2020 KPU merilis aplikasi Sistem Informasi Rekapitulasi Elektronik (SIREKAP) pada Pemilihan Kepala Daerah (Pilkada) serentak tahun 2020, SIREKAP digunakan sebagai perangkat pendukung rekapitulasi manual berjenjang. Setelah digunakan dalam Pilkada, akhirnya KPU menginisiasikan untuk Pemilu 2024 menggunakan aplikasi SIREKAP. SIREKAP terbagi menjadi 2 jenis yaitu, SIREKAP Mobile dan SIREKAP Web, yang masing-masing memiliki tugas dan fungsi yang berbeda. SIREKAP Mobile dimanfaatkan oleh KPPS untuk menghitung atau merekapitulasi hasil pemungutan suara di setiap TPS. Fungsinya sangat penting sebagai sumber data utama perolehan suara yang terdokumentasi dalam Formulir C. Hasil-KWK [5]. KPPS memanfaatkan SIREKAP Mobile untuk memasukkan data suara yang akan diolah dan dihitung oleh sistem. SIREKAP Web dipakai oleh Panitia Pemilihan Kecamatan (PPK) dan anggota KPU di tingkat Kota/Kabupaten dan Provinsi. Peran utamanya adalah dalam pengelolaan dan pengumpulan semua data yang berasal dari berbagai TPS.

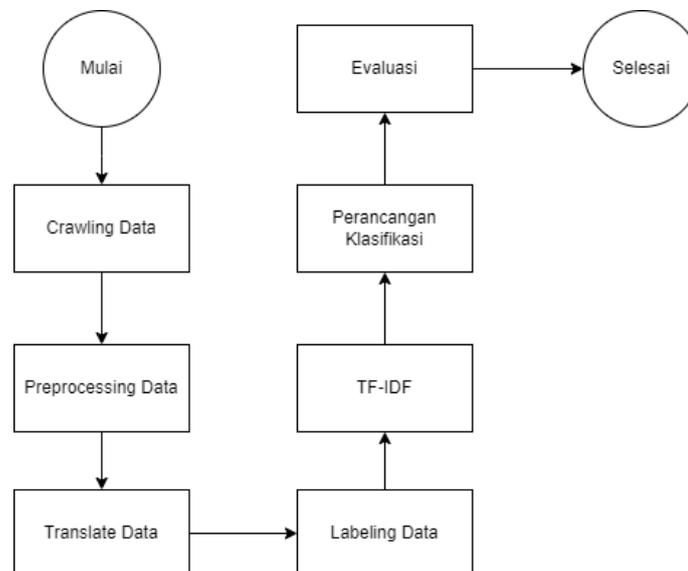
Pada penyelenggaraan pemilu 2024 menggunakan SIREKAP, hasil upload formulir C1 ke sistem menggunakan teknologi *Optical Character Recognition* (OCR). Teknologi ini merupakan sebuah pembandingan antara SIREKAP dengan SITUNG dimana angkanya diinput secara manual. Menurut Marsudi Wahyu Kisworo yang merupakan ahli IT, ada tiga masalah yang terjadi pada SIREKAP. Pertama, tulisan tangan formulir C1 berbeda-beda tiap TPS, ada tulisan tangan yang bagus adapun tidak yang membuat sulit dibaca oleh OCR. Kedua, merk hp dan kualitas foto formulir C1 tiap TPS berbeda-beda. Ada yang bagus, remang-remang, warna putih, serta kekuning-kuningan. Ketiga, masalah pada kertas formulir C1 yang kondisinya terlipat bisa menjadikan kesalahan baca oleh OCR [6].

Media sosial X adalah layanan jejaring sosial yang membantu penggunanya mengirim dan membaca pesan berbasis teks hingga 140 karakter [7]. Banyak masyarakat yang memberikan pendapat mereka melalui media sosial X terkait permasalahan yang terjadi pada Pemilu 2019 menggunakan SITUNG serta Pemilu tahun 2024 menggunakan SIREKAP. Berikut adalah contoh sentimen masyarakat terhadap SIREKAP “Permasalahan SIREKAP sudah jelas adanya ketidakcocokan antara hasil C1 plano dengan hasil yg ada pada angka di SIREKAP. Di KP hal itu tidak terjadi karena hasil inputan angka sudah sesuai dengan C1 Plano. Terima kasih tim KP terima kasih mas @ainunnajib”. Selain itu masyarakat juga memberikan sentiment mereka terhadap Pemilu 2019 menggunakan SITUNG seperti “C1 SDH dikawl bhkan smp prhitungn ditingkt kcmtan.. Tp di SITUNG KPU sekian ribu salah input. ms kt hrs menthelengin 813.000 inputan KPU tiap saat? Yg ini dibenrin yg lain ganti dt LG. kayak kucing2 an.. Blm LG yg C1 nya tdk asli. Mesti ada tindakan psti dr BPN”. Perlu kita ketahui bahwa sentimen seperti ini penting terutama bagi aplikasi yang dibuat oleh pemerintah. Sentimen masyarakat terhadap kedua aplikasi ini terlalu banyak diposting di media sosial X untuk diproses secara manual. Oleh karena itu, diperlukan suatu metode untuk memahami pendapat masyarakat terhadap aplikasi tersebut dan mengkategorikan pendapat tersebut ke positif atau negatif. Maka diperlukan analisis sentimen pada data ulasan atau opini-opini tersebut [8].

Analisis sentimen adalah proses untuk menganalisis opini seseorang terhadap suatu topik yang menentukan apakah opini tersebut positif atau negatif [9]. Tujuan dari penelitian ini untuk mengetahui sentimen masyarakat terhadap dua aplikasi ini serta apakah jumlah sentimen negatif pada SIREKAP lebih banyak daripada SITUNG.

## B. Metode Penelitian

Metode penelitian ini terdiri dari beberapa tahap. Gambar 1 menampilkan urutan atau rangkaian penelitian ini



**Gambar 1.** Metode Penelitian

### **B.1. Crawling Data**

Teknik crawling data yang digunakan pada penelitian ini dikumpulkan dari media sosial X menggunakan library tweet harvest. Tweet harvest adalah library untuk mengambil data dari media sosial X dengan menggunakan access token dari akun yang sudah terdaftar di media sosial X. Data dengan kata SIREKAP dikumpulkan dari tanggal 13 febuari 2024 sampai 14 maret 2024 dan data dengan kata SITUNG dikumpulkan dari tanggal 17 april 2019 sampai 16 mei 2019.

### **B.2. Preprocessing Data**

Pada tahap preprocessing, langkah-langkah tertentu harus dilakukan sebelum melanjutkan ke tahap berikutnya. Proses preprocessing akan berguna untuk membersihkan data, sehingga mempermudah dalam analisis selanjutnya [10].

#### **B.2.1. Cleansing**

Cleansing adalah proses evaluasi kualitas data dengan mengubah, memodifikasi, atau menghapus data yang dianggap tidak diperlukan dalam mengolah data [11].

#### **B.2.2. Case Folding**

Case folding adalah proses mengubah atau mengkonversi semua huruf menjadi huruf besar ataupun kecil. Dalam penelitian ini semua huruf akan diubah ke huruf kecil.

#### **B.2.3. Normalization**

Normalization adalah proses mengubah suatu kata dalam teks menjadi bentuk yang lebih umum. Tujuan dari proses ini yaitu mengurangi variasi dan menghasilkan kata yang konsisten dengan menghapus variasi kata yang memiliki makna sama [12].

#### **B.2.4. Stopword Removal**

Stopword removal adalah proses menghapus kata-kata yang tidak mempunyai arti signifikan untuk analisis data. Contoh kata yang sering muncul seperti “dan”, “yang”, “adalah”, “di” dan sebagainya. Proses ini menggunakan library sastrawi.

#### **B.2.5. Tokenizing**

Tokenizing bertujuan untuk memecah kalimat menjadi kata-kata yang memiliki makna, sehingga dapat meningkatkan akurasi data.

#### **B.2.6. Stemming**

Stemming adalah suatu proses mengubah kata berimbuhan menjadi kata dasar. Contoh kata yang diubah seperti “menahan => tahan”, “berbalas-balasan => balas”, dan sebagainya. Proses ini menggunakan library sastrawi.

### **B.3. Translate Data**

Pada tahap ini, dataset SIREKAP dan SITUNG yang awalnya berbahasa indonesia akan diubah ke dalam bahasa inggris. Tujuan translate data adalah untuk mempermudah proses labelling data secara otomatis menggunakan library TextBlob.

#### **B.4. Labelling Data**

Pada tahap ini, setiap data akan diberikan label untuk menentukan data tersebut termasuk ke dalam sentimen positif atau negatif. Proses ini menggunakan library TextBlob.

#### **B.5. TF-IDF**

TF-IDF adalah suatu metode algoritma yang bermanfaat dalam menghitung bobot setiap kata yang digunakan. Metode ini akan menghitung nilai Term Frequency (TF) dan Inverse Document Frequency (IDF) pada setiap kata di setiap dokumen dalam korpus [13]. Output dari metode ini adalah matriks fitur TF-IDF yang memberikan representasi numerik dari kata-kata dalam dokumen.

#### **B.6. Perancangan Klasifikasi**

Penelitian ini menggunakan algoritma SVM (*Support Vector Machine*) untuk melakukan pengklasifikasi data berdasarkan estimasi probabilitas. SVM adalah salah satu algoritma machine learning yang bertujuan untuk menemukan hyperplane optimal yang dapat memisahkan titik-titik data input [14]. Algoritma klasifikasi SVM memanfaatkan data traning untuk mengembangkan sebuah model klasifikasi. Model yang dibuat kemudian digunakan untuk memprediksi kelas data baru yang belum pernah dilihat sebelumnya, yang disebut data testing [15].

#### **B.7. Evaluasi**

Confusion matrix berguna dalam menilai performa algoritma klasifikasi yang dipakai untuk menganalisis sentimen. Confusion matrix dibagi menjadi 4 bagian, yaitu: True Positif (TP) adalah jumlah prediksi yang tepat sesuai dengan nilai aktual; True Negatif (TN) adalah jumlah prediksi yang tepat sesuai dengan nilai aktual; False Positif (FP) adalah jumlah prediksi yang salah tidak sesuai dengan nilai aktual; False Negatif (FN) adalah jumlah prediksi yang salah tidak sesuai dengan nilai aktual [16].

### **C. Hasil dan Pembahasan**

#### **C.1. Crawling Data**

Pada penelitian ini menggunakan 2 dataset yaitu data SIREKAP dan data SITUNG yang berasal dari media sosial X. Pengumpulan data menggunakan library tweet harvest. Tweet harvest menggunakan command atau perintah dalam mengumpulkan data dari media sosial X, oleh karena itu jalankan command **'SIREKAP lang:id since:2024-02-13 until:2024-03-14'** untuk mengambil data dengan kata SIREKAP. Setelah itu jalankan command **'SITUNG lang:id since:2019-**

04-17 until:2019-05-16' untuk mengambil data dengan kata SITUNG. Hasil pengumpulan data dengan kata SIREKAP mendapatkan data sebanyak 2022 data sementara itu hasil pengumpulan data dengan kata SITUNG mendapatkan data sebanyak 1503 data.

**Tabel 1.** Penjelasan command SIREKAP

No.	Kata	Keterangan
1.	sirekap	Tweet harus mengandung kata 'sirekap'
2.	lang:id	Tweet harus berbahasa indonesia
3.	since:2024-02-13	Tweet yang dikumpulkan dimulai dari tanggal 13 febuari 2024
4.	until:2024-03-14	Tweet yang dikumpulkan diakhiri tanggal 14 maret 2024

**Tabel 2.** Penjelasan command SITUNG

No.	Kata	Keterangan
1.	situng	Tweet harus mengandung kata 'situng'
2.	lang:id	Tweet harus berbahasa indonesia
3.	since:2019-04-17	Tweet yang dikumpulkan dimulai dari tanggal 17 april 2019
4.	until:2019-05-16	Tweet yang dikumpulkan diakhiri tanggal 16 mei 2019

## C.2. Preprocessing Data

Preprocessing data adalah tahapan untuk mengatasi berbagai kendala yang dapat menghambat saat pemrosesan data [17]. Beberapa tahapan yang terjadi pada saat preprocessing data sebagai berikut.

### C.2.1. Cleansing Data

Pada tahap cleansing dilakukan tujuh tahapan antara lain sebagai berikut: pertama remove duplicate data agar bersifat unik; kedua menghapus simbol mention tweet yang ada pada data; ketiga menghapus simbol hastag yang ada pada data; keempat menghapus simbol RT/retweet yang ada pada data; kelima menghapus link url yang ada pada data; keenam menghapus spasi berlebih yang ada pada data; ketujuh menghapus emoji atau emoticon yang ada pada data. Berikut adalah hasil dari proses cleansing data:

**Tabel 3.** Cleansing Data

Sebelum Cleansing	Sesudah Cleansing
ALHAMDULILLAH MSH ADA ORG YG SELALU MENCERMATI SIREKAP YG LAKNAT ITU <a href="https://t.co/V9qCosyFJQ">https://t.co/V9qCosyFJQ</a>	ALHAMDULILLAH MSH ADA ORG YG SELALU MENCERMATI SIREKAP YG LAKNAT ITU

### C.2.2. Case Folding

Pada tahap case folding dilakukan konversi keseluruhan kata atau huruf kedalam huruf kecil. Berikut adalah hasil dari proses case folding:

**Tabel 4.** Case Folding

Sebelum Case Folding	Sesudah Case Folding
ALHAMDULILLAH MSH ADA ORG YG SELALU MENCERMATI SIREKAP YG LAKNAT ITU	alhamdulillah msh ada org yg selalu mencermati SIREKAP yg laknat itu

### C.2.3. Normalization

Pada tahap normalization kata singkatan, gaul, dan slang word yang ada pada data akan dirubah menjadi kata baku. Berikut adalah contoh list kata yang dirubah menjadi kata baku:

**Tabel 5.** Contoh list kata normalization

No	Singkatan	Hasil
1.	yg	yang
2.	gw	saya
3.	plg	pulang
4.	grgr	gara-gara
5.	sbg	sebagai
6.	dll	dan lain-lain
7.	jgn	jangan

Kata-kata yang mengandung singkatan seperti “msh” akan diubah menjadi “masih”, “org” diubah menjadi “orang” dan kata “yg” akan diubah menjadi “yang”. Berikut dibawah ini hasil dari proses normalization.

**Tabel 6.** Normalization

Sebelum Normalization	Sesudah Normalization
alhamdulillah msh ada org yg selalu mencermati SIREKAP yg laknat itu	alhamdulillah masih ada orang yang selalu mencermati SIREKAP yang laknat itu

### C.2.4. Stopword Removal

Pada proses stopword removal kata-kata yang mengandung informasi tidak penting akan dihapus. Kata yang dihapus termasuk kedalam kata bantu yang sering muncul, contohnya seperti “di”, “dan”, “juga” dan lain-lain. Pada proses stopword removal ini menggunakan library sastrawi. Berikut hasil dari proses stopword removal:

**Tabel 7.** Stopword removal

Sebelum Stopword Removal	Sesudah Stopword Removal
alhamdulillah masih ada orang yang selalu mencermati SIREKAP yang laknat itu	alhamdulillah ada orang selalu mencermati SIREKAP laknat

### C.2.5. Tokenize

Pada proses ini kalimat yang ada pada data akan dipisah menjadi kata-kata yang memiliki makna, sehingga dapat meningkatkan akurasi data. Hasil dari tokenize ini menghasilkan data SIREKAP sebanyak 1897 data dan dataset SITUNG sebanyak 1469 data. Berikut hasil dari proses tokenize:

**Gambar 8.** Tokenize

Sebelum Tokenize	Sesudah Tokenize
alhamdulillah ada orang selalu mencermati SIREKAP laknat	'alhamdulillah' 'ada' 'orang' 'selalu' 'mencermati' 'SIREKAP' 'laknat'

### C.2.6. Stemming

Pada proses ini mengubah kata yang berimbuhan menjadi kata dasar. Tujuannya adalah menghapus kata yang berimbuhan seperti prefiks, suffiks dan

konfiks yang ada pada setiap kata. Proses stemming menggunakan library sastrawi. Berikut hasil dari proses stemming:

**Gambar 9. Stemming**

Sebelum Stemming	Sesudah Stemming
'alhamdulillah' 'ada' 'orang' 'selalu' 'mencermati' 'SIREKAP' 'laknat'	alhamdulillah ada orang selalu cermat SIREKAP laknat

### C.3. Translate Data

Pada tahap ini, dataset SIREKAP dan SITUNG yang awalnya berbahasa indonesia akan diubah ke dalam bahasa inggris. Tujuan translate data adalah untuk mempermudah proses labelling data secara otomatis menggunakan library TextBlob. Proses translate ini menggunakan library GoogleTranslator. Berikut adalah hasil dari proses translate data:

**Gambar 10. Translate Data**

Sebelum Translate	Sesudah Translate
alhamdulillah ada orang selalu cermat SIREKAP laknat	alhamdulillah there are people who are always careful about cursing

### C.4. Labelling Data

Pada tahap ini, dataset SIREKAP dan SITUNG akan diberikan label berdasarkan dari kata yang sering muncul yang berisifat sentimen positif atau negatif. Proses ini menggunakan library TextBlob. Proses ini menghasilkan sentimen positif berjumlah 1092 data dan sentimen negatif berjumlah 805 data pada dataset SIREKAP, sementara itu pada dataset SITUNG menghasilkan sentimen positif berjumlah 832 data dan sentimen negatif berjumlah 637 data. Berikut hasil dari proses labelling data:

**Gambar 11. Translate Data**

Text Indonesia	Text English	Label
alhamdulillah ada orang selalu cermat SIREKAP laknat	alhamdulillah there are people who are always careful about cursing	Negatif

Hasil dari proses ini selanjutnya akan divisualisasikan dari keseluruhan data yang bernilai positif dan negatif yang disebut sebagai wordcloud. Wordcloud adalah visualisasi dari kata-kata yang ada dan kata yang sering muncul akan ditampilkan dengan ukuran font yang lebih besar. Berikut adalah hasil dari wordcloud positif dan negative pada SIREKAP:



Gambar 2. Hasil wordcloud sentimen positif dan negatif pada dataset SIREKAP

Selanjutnya dibawah ini adalah hasil wordcloud sentiment positif dan negatif terhadap dataset SITUNG:



Gambar 3. Hasil wordcloud sentimen positif dan negatif pada dataset SITUNG

**C.5. TF-IDF**

Pada proses TF-IDF, dataset akan dihitung bobot kata setiap dokumen dan menghitung berapa banyak kata yang muncul setiap dokumen. Proses ini menggunakan TfidfVectorizer dari library scikit-learn yang berfungsi untuk mengubah teks menjadi matrix TF-IDF. Selanjutnya matrix ini akan diubah menjadi DataFrame dengan nama kolom yang sama seperti vektor TF-IDF. Berikut hasil dari proses TF-IDF:

	0	1	2	3	4	5	6	7	8	9	...	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.520561	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0

5 rows x 1960 columns

**Gambar 4.** Hasil TF-IDF pada dataset SIREKAP

Selanjutnya dibawah ini adalah hasil TF-IDF pada dataset SITUNG:

	0	1	2	3	4	5	6	7	8	9	...	1550	1551	1552	1553	1554	1555	1556	1557	1558
0	0.0	0.0	0.0	0.124318	0.231811	0.112157	0.0	0.0	0.0	0.373947	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.179838	0.000000	0.324493	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

5 rows x 1560 columns

**Gambar 5.** Hasil TF-IDF pada dataset SITUNG

**C.6. Perancangan Klasifikasi**

Sebelum melakukan implementasi algoritma SVM (*Support Vector Machine*) perlu memisahkan dataset menjadi dua bagian. Data training (70%) dan data test (30%). Setelah dataset berhasil dipisah menjadi dua bagian, dataset akan dilatih menggunakan algoritma SVM dari library scikit-learn. Hasil akurasi pada dataset SIREKAP berjumlah 68% dan untuk hasil akurasi pada dataset SITUNG berjumlah 70%.

	precision	recall	f1-score	support
Negatif	0.63	0.55	0.59	238
Positif	0.70	0.77	0.74	332
accuracy			0.68	570
macro avg	0.67	0.66	0.66	570
weighted avg	0.67	0.68	0.67	570

**Gambar 6.** Hasil akurasi dataset SIREKAP

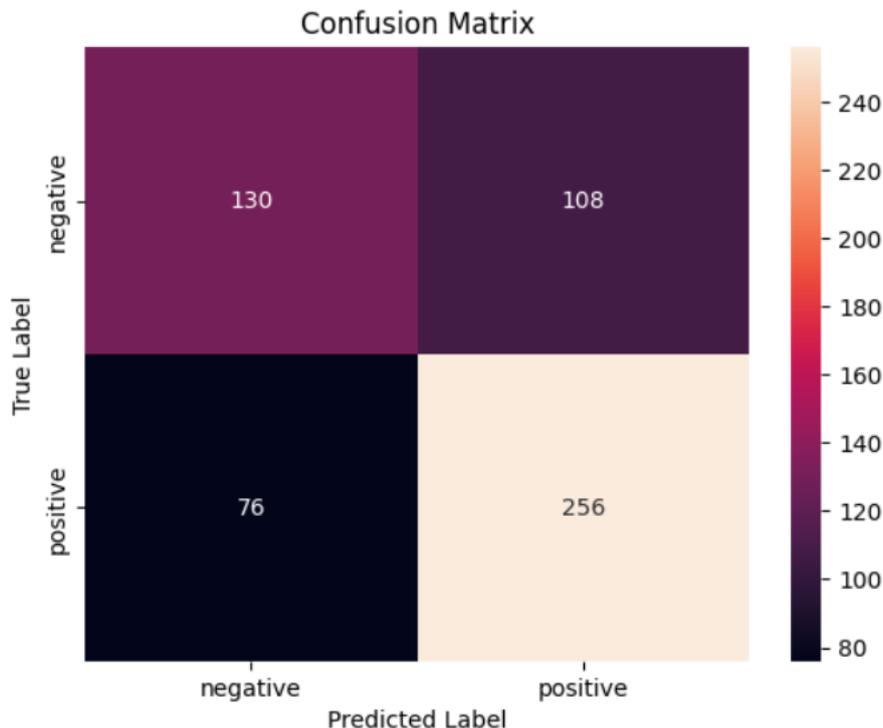
Berikut dibawah ini adalah hasil akurasi algoritma SVM pada dataset SITUNG:

	precision	recall	f1-score	support
Negatif	0.71	0.58	0.64	202
Positif	0.69	0.80	0.74	239
accuracy			0.70	441
macro avg	0.70	0.69	0.69	441
weighted avg	0.70	0.70	0.70	441

**Gambar 7.** Hasil akurasi dataset SITUNG

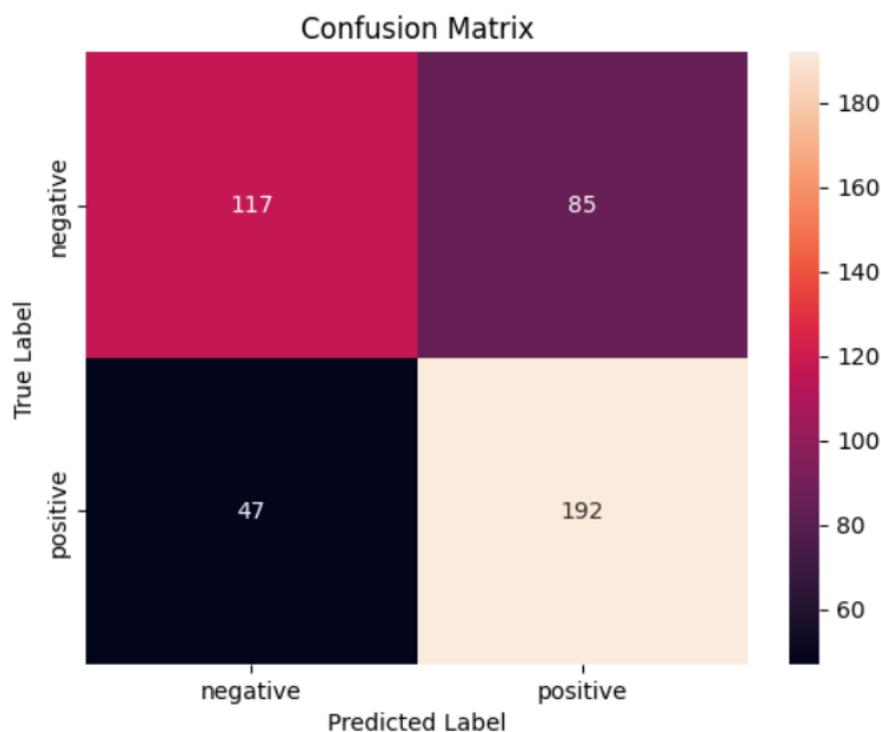
### C.7. Evaluasi

Proses ini diperoleh melalui confusion matrix yang menunjukkan bahwa algoritma SVM memiliki tingkat akurasi yang cukup baik. Berikut confusion matrix pada dataset SIREKAP:



**Gambar 8.** Hasil confusion matrix SIREKAP

Pada confusion matrix untuk dataset SIREKAP, terdapat 130 kasus True Negative yang menunjukkan bahwa klasifikasi data negatif dengan benar, 256 kasus True Positif yang menunjukkan klasifikasi data positif dengan benar, 76 kasus False Negatif yang berarti ada beberapa data positif yang salah diklasifikasi sebagai data negatif, dan 108 False Positif yang berarti ada beberapa data negatif yang salah diklasifikasi sebagai data positif. Selanjutnya adalah hasil confusion matrix pada dataset SITUNG.



**Gambar 9.** Hasil confusion matrix SITUNG

Pada confusion matrix SITUNG, terdapat 117 kasus True Negative yang menunjukkan bahwa klasifikasi data negatif dengan benar, 192 kasus True Positif yang menunjukkan klasifikasi data positif dengan benar, 47 kasus False Negatif yang berarti ada beberapa data positif yang salah diklasifikasi sebagai data negatif, dan 85 False Positif yang berarti ada beberapa data negatif yang salah diklasifikasi sebagai data positif. Tingkat akurasi pada dataset SITUNG memiliki nilai yang tinggi sekitar 70% dibandingkan dengan tingkat akurasi pada dataset SIREKAP dengan nilai 68%.

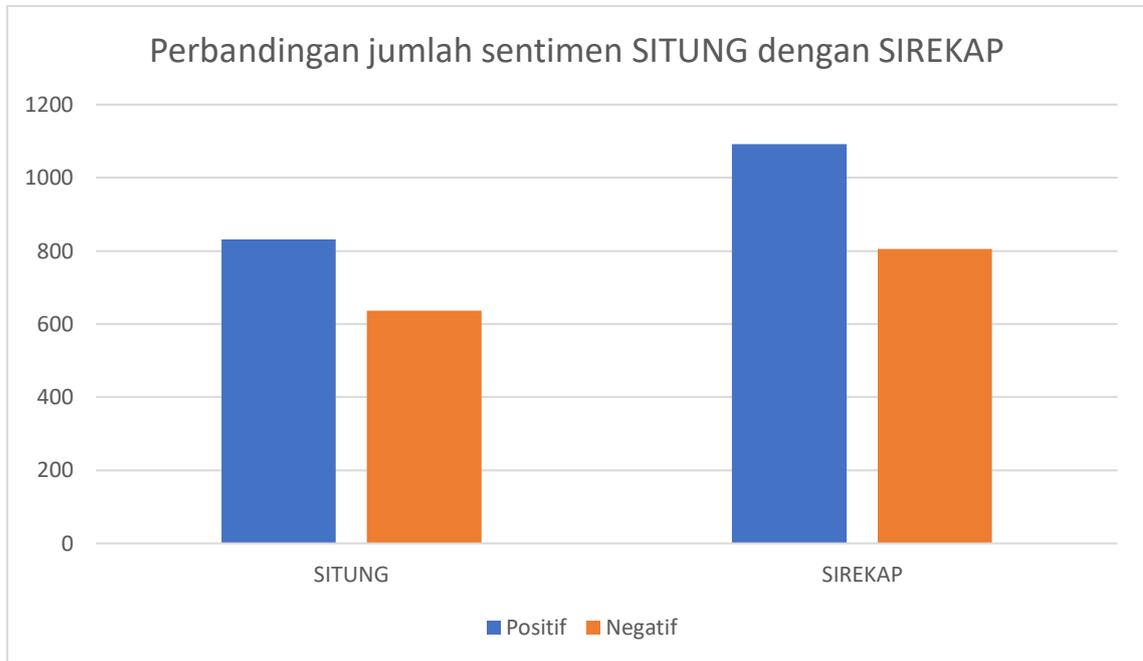
#### D. Kesimpulan

Berdasarkan hasil dari proses penelitian ini menggunakan algoritma SVM (*Support Vector Machine*) dapat ditarik kesimpulan antara lain:

1. Sebanyak 1897 data pada dataset SIREKAP dengan sentimen positif berjumlah 1092 data dan sentimen negatif berjumlah 805, sementara itu dataset SITUNG memiliki data sebanyak 1469 dengan sentimen positif berjumlah 832 data dan sentimen negatif berjumlah 637 data.
2. Pada dataset SIREKAP memiliki sentimen positif sebesar 57.5% dan sentimen negatif sebesar 42.4%, sementara itu pada dataset SITUNG memiliki sentimen positif sebesar 56.6% dan sentimen negatif sebesar 43.3%. Dataset SIREKAP memiliki sentimen positif sedikit lebih besar dari dataset SITUNG yang berarti sentimen aplikasi SIREKAP sedikit lebih baik dari aplikasi SITUNG.
3. Dari pengujian menggunakan algoritma SVM dengan memisahkan dataset menjadi dua bagian. Data training (70%) dan data test (30%) terhadap dataset SIREKAP dan SITUNG. Dataset SIREKAP mendapatkan nilai akurasi

sebesar 68% sedangkan untuk dataset SITUNG mendapatkan nilai akurasi lebih tinggi sebesar 70%.

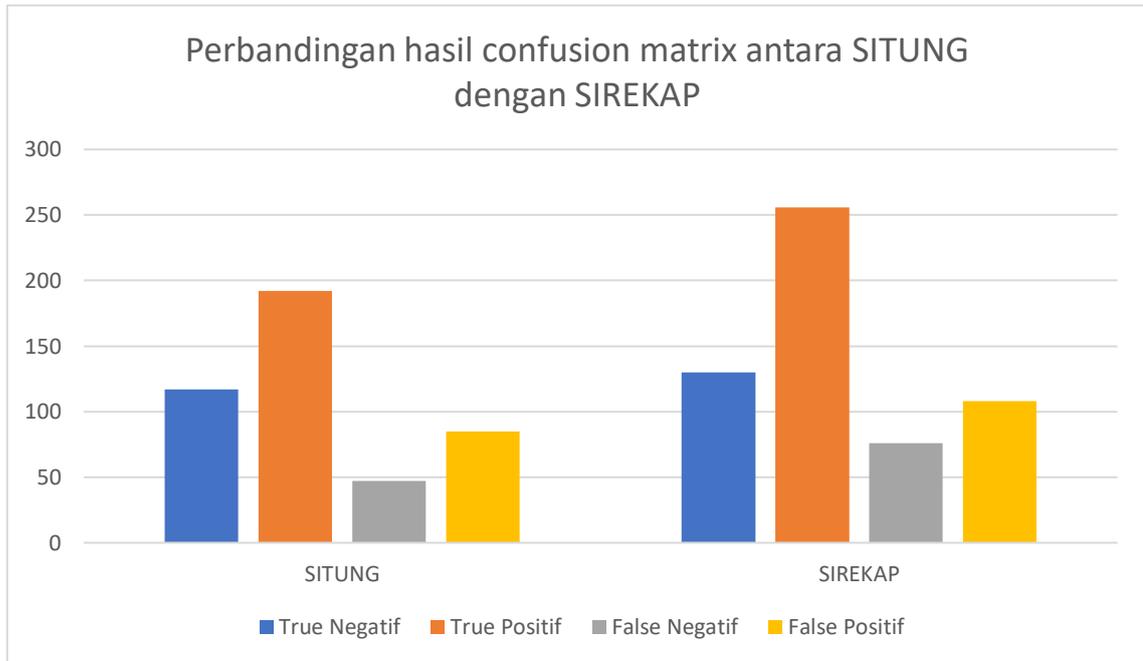
Berikut adalah perbandingan jumlah sentimen pada aplikasi SIREKAP dan SITUNG dibawah ini:



**Gambar 10.** Chart perbandingan jumlah sentiment SITUNG dengan SIREKAP

Dari perbandingan kedua aplikasi ini ternyata jumlah sentimen positif pada aplikasi SIREKAP mendapatkan lebih banyak dibandingkan aplikasi SITUNG, yaitu sebanyak 1092 pada SIREKAP dan 832 pada SITUNG, dengan selisih 260. Begitupun dengan jumlah sentiment negatif pada aplikasi SIREKAP yaitu sebanyak 805, sedangkan pada aplikasi SITUNG hanya mendapatkan 637, dengan selisih 168.

Selanjutnya adalah perbandingan hasil confusion matrix antara SITUNG dengan SIREKAP dibawah ini:



**Gambar 11.** Chart perbandingan hasil confusion matrix SITUNG dengan SIREKAP

Dari Gambar 11 dapat ditarik kesimpulan bahwa aplikasi SIREKAP mendapatkan hasil TRUE POSITIF lebih banyak dibandingkan aplikasi SITUNG, yaitu sebanyak 256 pada SIREKAP dan 192 pada SITUNG. Begitupun dengan hasil TRUE NEGATIF pada aplikasi SIREKAP yaitu sebanyak 130, sedangkan pada aplikasi SITUNG hanya mendapatkan 117.

### E. Referensi

- [1] Ari Welianto, "Sistem Demokrasi di Indonesia." Accessed: Apr. 25, 2024. [Online]. Available: <https://www.kompas.com/skola/read/2019/12/11/201742369/sistem-demokrasi-di-indonesia?page=all>
- [2] Annisa Medina Sari, "Pengertian Pemilu, Fungsi dan Prinsipnya," Fakultas Hukum Universitas Muhammadiyah Sumatera Utara.
- [3] M. Hardiyanti, Praditya Arcy Pratama, Aura Diva Saputra, Mila Mar'atus Sholehah, and M. Rizieq Aditya R, "URGENSI SISTEM E-VOTING DAN SIREKAP DALAM PENYELENGGARAAN PEMILU 2024," *JOURNAL EQUITABLE*, vol. 7, no. 2, pp. 249–271, Nov. 2022, doi: 10.37859/jeq.v7i2.4257.
- [4] Ananda Ridho Sulistya, "Inilah Perbedaan Cara Kerja SIREKAP dengan SITUNG," *Tempo*.
- [5] Devira Prastiwi, "4 Fakta Terkait SIREKAP KPU pada Pemilu 2024, Mulai dari Apa Itu hingga Cara Kerja," *Liputan 6*.
- [6] Utami Argawati, "Pakar IT Ungkap Tiga Sumber Masalah SIREKAP," Jakarta, Apr. 23, 2024.
- [7] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 10, no. 1, Jan. 2022, doi: 10.23960/jitet.v10i1.2262.

- [8] R. Wahyudi and G. Kusumawardana, "Analisis Sentimen pada Aplikasi Grab di Google Play Store Menggunakan *Support Vector Machine*," *Jurnal Informatika*, vol. 8, no. 2, pp. 200–207, Sep. 2021, doi: 10.31294/ji.v8i2.9681.
- [9] M. R. Fahlevvi, "ANALISIS SENTIMEN TERHADAP ULASAN APLIKASI PEJABAT PENGELOLA INFORMASI DAN DOKUMENTASI KEMENTERIAN DALAM NEGERI REPUBLIK INDONESIA DI GOOGLE PLAYSTORE MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*," *Jurnal Teknologi dan Komunikasi Pemerintahan*, vol. 4, no. 1, pp. 1–13, Jun. 2022, doi: 10.33701/jtkp.v4i1.2701.
- [10] N. Cahyono and Dewi Setiyawati, "Analisis Sentimen Pengguna Sosial Media Twitter Terhadap Perokok Di Indonesia," *Indonesian Journal of Computer Science*, vol. 12, no. 1, Feb. 2023, doi: 10.33022/ijcs.v12i1.3154.
- [11] D. Darwis, N. Siskawati, and Z. Abidin, "PENERAPAN ALGORITMA NAIVE BAYES UNTUK ANALISIS SENTIMEN REVIEW DATA TWITTER BMKG NASIONAL," *Jurnal Tekno Kompak*, vol. 15, no. 1, p. 131, Feb. 2021, doi: 10.33365/jtk.v15i1.744.
- [12] Fauzan Baehaqi and N. Cahyono, "Analisis Sentimen Terhadap Cyberbullying Pada Komentar Di Instagram Menggunakan Algoritma Naïve Bayes," *Indonesian Journal of Computer Science*, vol. 13, no. 1, Jan. 2024, doi: 10.33022/ijcs.v13i1.3301.
- [13] Delta Sierra, "Algoritma TF-IDF."
- [14] Dadan Dahman W., "*Support Vector Machine (SVM)*."
- [15] E. Suryati, A. Ari Aldino, N. Penulis Korespondensi, and E. Suryati Submitted, "Analisis Sentimen Transportasi Online Menggunakan Ekstraksi Fitur Model Word2vec Text Embedding Dan Algoritma *Support Vector Machine (SVM)*," vol. 4, no. 1, pp. 96–106, 2023, doi: 10.33365/jtsi.v4i1.2445.
- [16] I. H. Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 8, no. 3, pp. 302–307, Sep. 2023, doi: 10.30591/jpit.v8i3.5734.
- [17] Universitas Bina Nusantara, "Teknik pre-processing dan classification dalam data science."