

Penerbit: LPPM Universitas Bina Sarana Informatika  
Journal of Accounting Information System  
Website: <https://jurnal.bsi.ac.id/index.php/jais>

## Analisis Sentimen Terhadap Pengguna Aplikasi Instagram di Google Play Store Menggunakan Support Vector Machine

Ratih Yulia Hayuningtyas<sup>1</sup>, Retno Sari<sup>2</sup>

<sup>1</sup> Sistem Informasi, Universitas Bina Sarana Informatika

<sup>2</sup> Informatika, Universitas Nusa Mandiri

<sup>1</sup> [ratih.ryl@bsi.ac.id](mailto:ratih.ryl@bsi.ac.id), <sup>2</sup> [retno.mnr@nusamandiri.ac.id](mailto:retno.mnr@nusamandiri.ac.id)

### INFORMASI ARTIKEL

#### Riwayat Artikel:

Submit 01/05/2024

Revisi 02/06/2025

Terima 04/06/2025

Terbit Online 04/06/2025

#### Kata Kunci :

RBF, Instagram, Google Playstore, SVM, Sentimen

**Keyword :** RBF, Instagram, Google Playstore, SVM, Sentiment.

### ABSTRAK

Ulasan dari pengguna merupakan salah satu faktor pendukung dalam menentukan sebuah keputusan. Hal ini dapat dilihat dengan banyaknya ulasan positif akan berbanding lurus dengan banyaknya pengguna aplikasi. Penelitian ini melakukan eksperimen dengan melakukan analisis sentiment terhadap pengguna aplikasi Instagram di Google Play Store menggunakan Support Vector Machine. Tujuan penelitian ini dilakukan untuk mengetahui pengaruh parameter C dan gamma pada hasil akurasi analisis sentimen terhadap Instagram di *Google Play Store* dengan metode *Support Vector Machine* (SVM) menggunakan kernel Linear, Sigmoid, Polynomial dan RBF. Tahapan pada penelitian ini yaitu crawling data, preprocessing, transformation, modeling, evaluasi dan wordcloud. Banyaknya data pada penelitian ini yaitu 10.000 ulasan. Preprocessing yang dilakukan yaitu filtering, labeling, tokenizing, stopword dan stemming. Pada tahap modeling dilakukan dengan membagi dua bagian yaitu data latih dan data uji, dengan membuat 2 model setelah itu dilakukan pengujian kernel SVM yaitu linier, Polinomial, sigmoid dan RBF dengan parameter C:1 dan gamma 1. Dari hasil pengujian didapatkan nilai akurasi tertinggi berada pada Model 2 dengan kernel RBF dengan nilai 0.745.

### ABSTRACT

*User reviews are one of the supporting factors in determining a decision. This can be seen from the number of positive reviews directly proportional to the number of application users. This study experimented by performing sentiment analysis on Instagram application users on the Google Play Store using a Support Vector Machine. The purpose of this study was to determine the effect of C and gamma parameters on the accuracy of sentiment analysis on Instagram on the Google Play Store with the Support Vector Machine (SVM) method using Linear, Sigmoid, Polinomial, and RBF kernels. The stages in this study are data crawling, preprocessing, transformation, modeling, evaluation, and word cloud. The amount of data in this study is 10,000 reviews. The preprocessing carried out is filtering, labeling, tokenizing, stopword, and stemming. The modeling stage, is done by dividing two parts, namely training data and test data, by creating 2 models after which the SVM kernel test is carried out, namely linear, Polinomial, sigmoid, and RBF with parameters C: 1 and gamma 1. From the test results, the highest accuracy value is in Model 2 with the RBF kernel with a value of 0.745.*

### Penulis Koresponden:

Retno Sari,

Email: [retno.mnr@nusamandiri.ac.id](mailto:retno.mnr@nusamandiri.ac.id)

## Pendahuluan

Pada tahun 2020 pengguna Instagram menempati peringkat 4 di dunia dengan jumlah pengguna sebanyak 73 juta pengguna (Asshiddiqi & Muslim, 2020). Menurut data yang didapat dari hootsuite *we are social*, di Indonesia pengguna Instagram sebanyak 87,7% dari total masyarakat di tahun 2021 (Azizah, 2023). Dengan berkembangnya media elektronik saat ini, media tradisional jarang digunakan oleh masyarakat (Sayogo et al., 2023).

Saat ini setiap orang dapat memperlihatkan aktivitas kesehariannya serta berkomunikasi dengan sesama pengguna media sosial secara online (Wijaya & Widyaningrum, 2024). Instagram banyak digunakan untuk mengemukakan pendapat yang ditujukan kepada sebuah instansi, tokoh publik atau lembaga lainnya (Asshiddiqi & Muslim, 2020). Instagram memiliki kelebihan untuk memperluas jaringan sosial dengan berteman kepada sesama pengguna (Nugraha & Astuti, 2023)

Banyaknya pengguna Instagram berbanding lurus dengan ulasan yang diberikan pengguna terkait aplikasi ini, dapat berupa ulasan positif maupun ulasan negatif. Ulasan yang diberikan menjadi bahan pertimbangan untuk penggunaan dan untuk eksistensi aplikasi ini. Dari ulasan-ulasan yang diberikan dapat dimanfaatkan untuk menyimpulkan ulasan terbanyak yang diutarakan pada google play store (Zulfiqri et al., 2024). Ulasan pengguna merupakan hal yang penting ditindak lanjuti untuk memperbaiki kinerja aplikasi sehingga aplikasi tetap bertahan eksistensinya.

Dengan adanya permasalahan diatas maka diperlukan analisis sentiment untuk mengetahui seberapa puas pengguna Instagram. Analisis sentimen merupakan proses meringkas, mengolah data berupa teks dengan tujuan menemukan informasi sentimen pada sebuah ulasan (Arsi & Waluyo, 2021), mengelompokkan ulasan berdasarkan polaritas teks yang ada di dalamnya (Winarti et al., 2023) dan untuk melihat pendapat masyarakat akan isu yang saat ini sedang berkembang (Anggraini et al., 2021). Pemanfaatan data yang diambil secara publik merupakan sebuah alternatif untuk menggantikan survey tradisional (Rachman & Pramana, 2020).

Terdapat penelitian yang membandingkan algoritma Support Vector Machine dengan Logistic Regression pada analisis sentimen sosial media Universitas, didapatkan bahwa Support Vector Machine dengan memakai TF-IDF mendapatkan hasil lebih baik dari pada Logistic Regression yaitu rata-rata nilai F1 74.6% sedangkan Logistic Regression dengan Bag-of-Words mendapatkan nilai F1 74% (Wicaksono et al., 2022).

Terdapat penelitian sebelumnya melakukan analisis sentiment *cyberbullying* pada komentar Instagram dengan menggunakan komentar sebanyak 1000 data dan dilakukan eksperimen menggunakan algoritma *Support Vector Machine* dengan membagi data latih dan data uji, pada penelitian ini ini didapatkan tingkat akurasi mencapai 92% dengan pembagian data latih dan uji yaitu 90% data latih dan 10% data uji (Tampubolon & Devesa, 2025).

Penelitian sentiment analisis terhadap Rohingya di Instagram membandingkan antara algoritma *Support Vector Machine* dengan metode tradisional dengan membagi kategori sentimen menjadi tiga yaitu positif, negatif dan netral didapatkan hasil bahwa *Support Vector Machine* mendapatkan nilai akurasi 88,8%, precision 88,3%, recall 91% dan F1-score 89,6% sedangkan tradisional method nilai akurasinya 74,9%, precision 74%, recall 76,5% dan F1-Score 75,12% (Santoso, 2024).

Dilakukan analisis komentar terhadap layanan aplikasi SeaBank Indonesia pada Instagram dengan *Support Vector Machine* dengan jumlah data yang dipreprocessing sebanyak 1201 data yang membagi data latih sebanyak 960 dan data uji sebanyak 241 dengan melakukan pelabelan data menjadi 2 yaitu positif dan negatif dan menghasilkan nilai akurasi 89%, precision sebesar 93%, recall 83% dan f1-score sebesar 88%, (Arrafiq & Kurniawan, 2024).

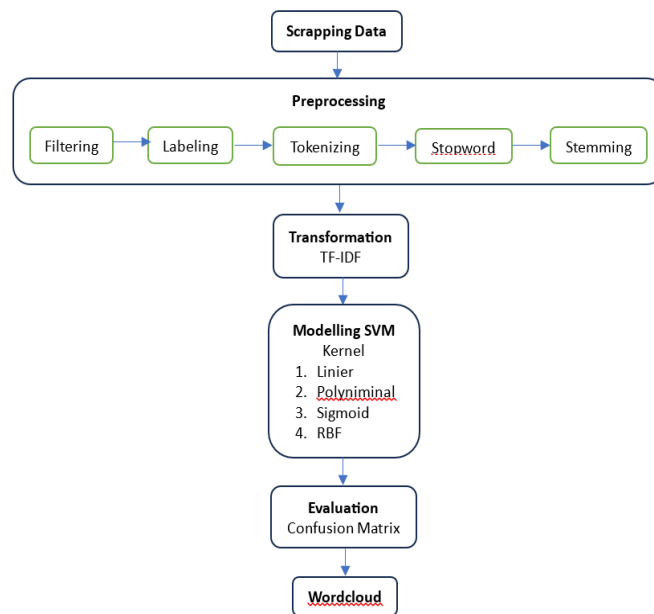
*Support Vector Machine* merupakan algoritma yang paling populer untuk analisis sentiment (Gupta & Rattan, 2023) dan telah berhasil digunakan dalam klasifikasi teks dan analisis sentiment (Borg & Boldt, 2020). SVM merupakan algoritma klasifikasi yang efisien yang banyak digunakan (Wang & Zhao, 2020) dan dinilai merupakan algoritma dengan pendekatan terbaik untuk analisis sentiment karena menghasilkan akurasi yang lebih baik (Isnain et al., 2023). Kernel pada algoritma SVM memiliki peranan penting untuk bekerja lebih baik dalam klasifikasi (Nurhaliza Agustina et al., 2024), kernel memiliki dampak terhadap akurasi (Praghakusma & Charibaldi, 2021).

Dari hasil penelitian yang sudah dilakukan, maka pada penelitian ini membahas mengenai analisis sentimen dengan menggunakan *Support Vector Machine* yang bertujuan untuk mengetahui pengaruh parameter C dan gamma terhadap hasil akurasi analisis sentimen terhadap Instagram di *Google Play Store* dengan metode *Support Vector Machine* (SVM) menggunakan kernel Linear, Sigmoid, Polynomial dan RBF. Penelitian ini

bermanfaat sebagai referensi penelitian analisis sentiment untuk dapat dikembangkan agar menghasilkan akurasi yang lebih baik.

## Metode Penelitian

Penelitian ini memiliki tahapan dari pengumpulan data sampai dengan tahapan akhir. Penelitian ini memiliki beberapa tahapan dimulai dari scrapping data, preprocessing, proses transformationa, modelling dengan metode *Support Vector Machine* dan pengujian dengan beberapa kernel *Support Vector Machine*, evaluasi dengan confusion matrix dan menampilkan *wordcloud*. Gambar dibawah ini menunjukkan tahapan penelitian .



Gambar 1. Tahapan Penelitian

### A. Pengumpulan data

Dataset merupakan data yang akan dipakai dalam eksperimen ini untuk proses *text mining* dengan metode *support vector machine*. Dalam penelitian, data dikumpulkan sebanyak 10.000 data yang diambil dengan cara *scrapping* pada aplikasi *google play store* terhadap *review* pengguna instagram. Dataset akan dilakukan *preprocessing* sebelum dilakukan proses klasifikasi.

### B. Preprocessing

Digunakan untuk mengolah dokumen mentah menjadi dokumen yang siap untuk dilakukan (Puji Astuti et al., 2022). Tahapan ini sangat berdampak untuk analisis sentimen terutama untuk data yang mempunyai banyak noise dan perlu dilakukan *filtering*, *label*, *tokenizing*, *stopword* dan *stemming* (Afdal & Elita, 2022).

Tabel 1. Tahapan *Preprocessing*

No	Tahapan	Penjelasan
1	Filtering	Membersihkan data yang tidak diperlukan serta membersihkan data yang duplikat (Widya Utami & Artana, 2022)
2	Label	Dataset yang ada akan diberikan label yaitu positif dan negatif (Alhaq et al., 2021)
2	Tokenizing	Proses menghilangkan simbol, tanda baca, dan angka, yang ada pada dataset (Novia Lestari et al., 2023)
3	Stopword	Menghilangkan kata yang dianggap kurang di dalam kalimat sehingga mengurangi jumlah teks dengan tujuan proses <i>stopword</i> menjadi lebih cepat (Hermawan et al., 2023)

4	Stemming	Menghilangkan kata imbuhan dalam kalimat sehingga menjadi kata dasar (Astuti et al., 2024)
---	----------	--

C. Transformation

Pembobotan TF-IDF merupakan proses pembobotan dari setiap kata, semakin banyak kata yang tampil dalam suatu dataset maka semakin tinggi nilai bobotnya. TF (*Term Frequency*) pembobotan terhadap kata yang tampil dalam kalimat berdasarkan jumlah yang tampil pada seluruh kalimat pada dokumen (*Inverse Document Frequency-IDF*) (Arifin et al., 2021). TF-IDF disebut juga dengan pengujian matematis, dimana mengukur seberapa penting sebuah kata dalam suatu kalimat, kepentingan kata akan mempengaruhi frekuensi (Atika et al., 2022).

D. Modeling

Pada tahapan modelling dilakukan penerapan metode yang dipakai pada proses klasifikasi dalam penelitian yaitu *Support Vector Machine*. Metode ini salah satu metode klasifikasi yang memisahkan dua kelas pada suatu data dengan membuat *hyperlane*. Metode ini sering dipakai dalam memperkirakan dari sebuah kelas dan memberikan sebuah label (Setiawan & Suryono, 2024). *Support Vector Machine* memiliki kelebihan yaitu memiliki akurasi yang baik (Darmawan et al., 2023).

Pada support vector machine diperlukan sebuah kernel tambahan yang bertujuan untuk mentransformasikan data, kernel pada *support vector machine* yaitu kernel linier, Polinomial, sigmoid dan RBF (Praghakusma & Charibaldi, 2021). Setiap kernel memiliki fungsi parameter yang spesifik untuk mencapai hasil yang optimal, berikut adalah penjelasan dari kernel:

Tabel 2. Fungsi Kernel

Kernel	Keterangan
Linier	Dengan kernel ini <i>support vector machine</i> dapat menangani data yang dipisahkan secara linier tanpa transformasi nonlinier
Polinomial	Digunakan baik untuk <i>hard</i> maupun <i>soft</i> margin dalam menangani pola nonlinier
RBF	RBF memiliki kekurangan dan kelebihan yang mempengaruhi kinerja <i>support vector machine</i>
Sigmoid	Salah satu alat dalam <i>support vector machine</i> yang dapat menghasilkan klasifikasi secara optimal

E. Evaluation

*Confusion matriks* disebut juga dengan matriks evaluasi yang digunakan untuk memahami sejauh mana metode yang digunakan pada penelitian ini berhasil mengklasifikasikan sebuah dataset. Dalam matriks evaluasi ada beberapa katagerori dalam pengklasifikasian yaitu *True Positif* (TP), *True Negative* (TN), *False Positif* (FP), dan *False Negative* (FN) (I Komang Dharmendra et al., 2023). Pada proses ini akan menghasilkan nilai akurasi, nilai presisi, dan nilai recall dari sebuah matriks evaluasi. Akurasi digunakan sebagai pengukuran bagaimana metode dapat mengklasifikasikan data dengan .betul presisi digunakan untuk memperkirakan sejauh mana memprediksi positif dengan benar sedangkan recall untuk memperkirakan sampai mana metode dapat membaca instance positif yang sesungguhnya (I Komang Dharmendra et al., 2023).

F. Wordcloud

Setelah itu kita menggunakan wordcloud untuk melihat kata yang sering tampil, *wordcloud* merupakan salah satu metode untuk menampilkan kata-kata yang populer yang ada pada Kumpulan data, metode ini

menampilkan kata dengan cara visual (Astuti et al., 2024) Semakin besar huruf yang ada pada *wordcloud* maka semakin sering frekuensi kemunculan data tersebut pada kelasnya (Muttaqin & Kharisudin, 2021).

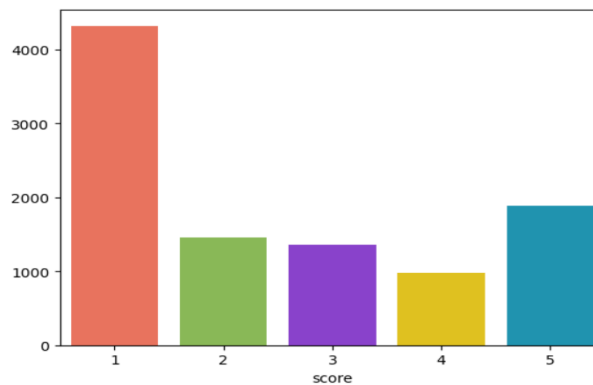
## Hasil dan Pembahasan

### A. Data

Data yang digunakan yaitu data *review* dari pengguna aplikasi Instagram pada google play store. Dataset diperoleh dengan teknik scrapping, data yang digunakan sebanyak 10.000 data dari tahun 2018-2024, data tersebut memiliki nilai 1 dan 2 termasuk *review* negative, nilai 3 termasuk netral, *review* dengan nilai 4 dan 5 termasuk *review* positif. Dari pengumpulan data ini akan dilakukan *preprocessing*, sehingga data yang ada sudah siap untuk dilakukan pengolahan. Berikut adalah data *review* pengguna aplikasi instragram.

Tabel 3. Dataset

	reviewId	userNa me	userImage	content	sc or e	thumbsU pCount	reviewCre atedVersi on	at	replyCo ntent
0	1a2eb812-68f3-44b6-8ac0-13e1d9c9696b	Penggu na Google	https://play- lh.googleusercontent.com/EGemoI2N...	Mudah refresh terus, foto yg dilihat langsung ...	1	0	215.0.0.27. 359	12/9/2 021 18:48	NaN
1	3006c0ce-5fbf-4cac-b320-67c33e55cf20	Penggu na Google	https://play- lh.googleusercontent.com/EGemoI2N...	Setelah update malah banyak bug. Di story tuli...	1	0	167.1.0.25. 120	11/26/ 2023 16:09	NaN
2	6f4fe7d4-dd6a-4897-a5b9-936d8b6a69dd	Penggu na Google	https://play- lh.googleusercontent.com/EGemoI2N...	Banyak bug nya dan sampe sekarang belum diperb...	1	8	185.0.0.38. 116	5/1/20 21 21:57	NaN



Gambar 2. Diagram Dataset

### B. Preprocessing

Data yang sudah dikumpulkan maka dilakukan tahapan selanjutnya yaitu *preprocessing*, agar data dapat langsung digunakan untuk pengolahan dengan metode *support vector machine*. Berikut adalah tahapan *preprocessing* pada penelitian ini.

#### 1. Filtering

Dari hasil pengumpulan data atribut yang diperoleh yaitu, *reviewId*, *userName*, *userImage*, *content*, *score*, *thumbsUpCount*, *reviewCreatedVersion*, *at*, *replyContent*. Pada tahap *filtering* menghilangkan atribut yang tidak digunakan dan menyisipkan atribut tahun, bulan dan tanggal.

Tabel 4. Hasil Filtering

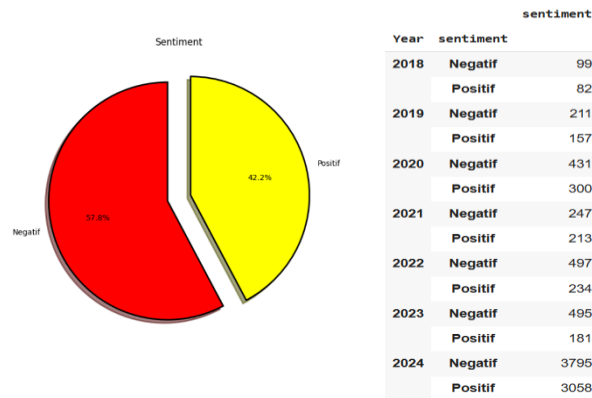
	reviewId	content	score	thumbsUpCount	Year	Month	Day
0	1a2eb812-68f3-44b6-8ac0-13e1d9c9696b	Mudah refresh terus, foto yg dilihat langsung ...	1	0	2021	12	9
1	3006c0ce-5fbf-4cac-b320-67c33e55cf20	Setelah update malah banyak bug. Di story tuli...	1	0	2023	11	26
2	6f4fe7d4-dd6a-4897-a5b9-936d8b6a69dd	Banyak bug nya dan sampe sekarang belum diperb...	1	8	2021	5	1

2. Labeling

Menentukan kelas dari setiap review, untuk menentukan review masuk kedalam kelas positif atau negatif kita melihat kolom score pada dataset. Berikut adalah hasil dari labeling.

Tabel 5. Hasil Labeling

	reviewId	content	score	thumbsUpCount	Year	Month	Day	sentiment
0	1a2eb812-68f3-44b6-8ac0-13e1d9c9696b	Mudah refresh terus, foto yg dilihat langsung ...	1	0	2021	12	9	Negatif
1	3006c0ce-5fbf-4cac-b320-67c33e55cf20	Setelah update malah banyak bug. Di story tuli...	1	0	2023	11	26	Negatif
2	6f4fe7d4-dd6a-4897-a5b9-936d8b6a69dd	Banyak bug nya dan sampe sekarang belum diperb...	1	8	2021	5	1	Negatif



Gambar 3. Diagram Labeling

3. Tokenizing

Pada tahapan ini dalam penelitian dilakukan proses penghapusan simbol, penghapusan url, penghapusan mention, penghapusan hasthag, penghapusan tanda baca, dan angka:

Tabel 6. Hasil Tokenize

	reviewId	content	score	thumbsUpCount	Year	Month	Day	sentiment	Tokenize
0	1a2eb812-68f3-44b6-8ac0-13e1d9c9696b	mudah refresh terus foto yg dilihat langsung h...	1	0	2021	12	9	Negatif	[mudah, refresh, terus, foto, yg, dilihat, lan...

	reviewId	content	score	thumbsUpCount	Year	Month	Day	sentiment	Tokenize
1	3006c0ce-5fbf-4cac-b320-67c33e55cf20	setelah update malah banyak bug di story tulis...	1	0	2023	11	26	Negatif	[setelah, update, malah, banyak, bug, di, stor...
2	6f4fe7d4-dd6a-4897-a5b9-936d8b6a69dd	banyak bug nya dan sampe sekarang belum diperb...	1	8	2021	5	1	Negatif	[banyak, bug, nya, dan, sampe, sekarang, belum...

4. Stopword

Setelah dataset dilakukan *tokenizing*, maka proses selanjutnya dilakukan *stopword*. Pada tahapan ini *stopword* dilakukan penghapusan kata yang kurang dari 4 kata dari sebuah kalimat. Hasil proses *stopword* dapat dilihat pada tabel 7.

Tabel 7. Hasil *stopword*

	reviewId	content	score	thumbsUpCount	Year	Month	Day	sentiment	Tokenize	Stopword
0	1a2eb812-68f3-44b6-8ac0-13e1d9c9696b	mudah refresh terus foto yg dilihat langsung h...	1	0	2021	12	9	Negatif	[mudah, refresh, terus, foto, yg, dilihat, lan...	[mudah, refresh, foto, langsung, hilang, kebaw...
1	3006c0ce-5fbf-4cac-b320-67c33e55cf20	setelah update malah banyak bug di story tulis...	1	0	2023	11	26	Negatif	[setelah, update, malah, banyak, bug, di, stor...	[update, story, tulisan, musik, temukan, story...
2	6f4fe7d4-dd6a-4897-a5b9-936d8b6a69dd	banyak bug nya dan sampe sekarang belum diperb...	1	8	2021	5	1	Negatif	[banyak, bug, nya, dan, sampe, sekarang, belum...	[sampe, diperbaiki, report, masi, perubahan, t...

5. Stemming

Selanjutnya tahapan proses *stemming* menghilangkan imbuhan dari setiap kata yang ada di dalam dataset sehingga menjadi kata dasar. Berikut adalah hasil dari proses *stemming*.

Tabel 8. Hasil Stemming

	reviewId	content	score	thumbsUpCount	Year	Month	Day	sentiment	Tokenize	Stopword	Stemming
0	1a2eb812-68f3-44b6-8ac0-13e1d9c9696b	mudah refresh terus foto yg dilihat langsung h...	1	0	2021	12	9	Negatif	[mudah, refresh, terus, foto, yg, dilihat, lan...	[mudah, refresh, foto, langsung, hilang, kebaw...	[mudah, refresh, foto, langsung, hilang, bawah...
1	3006c0ce-5fbf-4cac-b320-67c33e55cf20	setelah update malah banyak bug di story tulis...	1	0	2023	11	26	Negatif	[setelah, update, malah, banyak, bug, di, stor...	[update, story, tulisan, musik, temukan, story...	[update, story, tulis, musik, temu, story, ora...

reviewId	content	score	thumbsUpCount	Year	Month	Day	sentiment	Tokenize	Stopword	Stemming
2 6f4fe7d4-dd6a-4897-a5b9-936d8b6a69dd	banyak bug nya dan sampe sekarang belum diperb...	1	8	2021	5	1	Negatif	[banyak, bug, nya, dan, sampe, sekarang, belum...	[sampe, diperbaiki, report, masi, perubahan, t...	[sampe, baik, report, mas, ubah, temu, fitur, ...

C. Transformation

Setelah dilakukan preprocessing tahapan selanjutnya yaitu TF-IDF untuk menghitung kemunculan dari setiap kata pada dokumen. Berikut adalah tampilan dari TF-IDF.

Tabel 9. Hasil FD-IDF

Index	Kata	TF-IDF
0	Response	0.301201
0	Updaye	0.301201
0	Closed	0.288369
0	Then	0.279264
0	Susah	0.276222
...	...	...
9999	Susah	0.121907
9999	Coba	0.102024
9999	Kalo	0.094680
9999	Udah	0.082760
9999	Fitur	0.078869

D. Modeling

Pada proses modeling, data yang sudah diberikan nilai bobot disetiap kata dilaksanakan pengolahan dengan metode *support vector machine*. Sebelum data diolah akan dibagi menjadi dua bagian yaitu data latih dan data uji. Pada tabel 3 dapat dilihat pembagian data.

Tabel 10. Pembagian Data

Model	Latih : Uji	Data Latih	Data Uji
Model 1	90 : 10	9000	1000
Model 2	80 : 20	8000	2000

Data latih dan data uji akan dilakukan eksperimen terhadap masing-masing kernel pada *support vector machine*. Berikut masing-masing model dari kernel

Tabel 11. Pengujian Data

Model	Kernel dan Parameter
Model 1 (90 : 10)	Kernel : Linear, Parameter C: 1, gamma : 1
	Kernel : Sigmoid, Parameter C: 1, gamma : 1
	Kernel : Polinomial, Parameter C: 1, gamma : 1
	Kernel : RBF, Parameter C: 1, gamma : 1
Model 2 (80 : 10)	Kernel : Linear, Parameter C: 1, gamma : 1
	Kernel : Sigmoid, Parameter C: 1, gamma : 1
	Kernel : Polinomial, Parameter C: 1, gamma : 1
	Kernel : RBF, Parameter C: 1, gamma : 1

Masing-masing kernel di implementasikan dengan metode *support vector machine* sehingga diperoleh nilai akurasi. Berikut hasil dari masing-masing model kernel *support vector machine*.

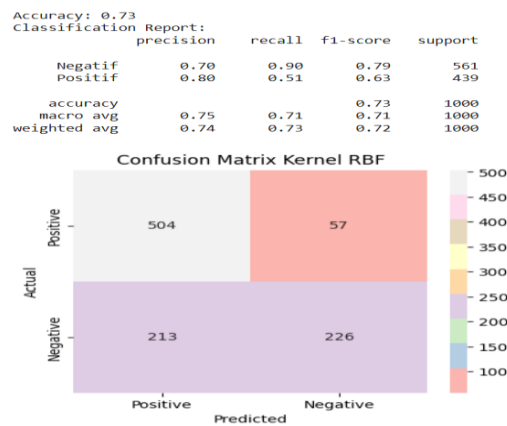
Tabel 12. Hasil Pengujian Kernel SVM

Model	Kernel dan Parameter	Nilai Akurasi
Model 1	Kernel : Linear, Parameter C: 1, gamma : 1	0.727

(90 : 10)	Kernel : Sigmoid, Parameter C: 1, gamma : 1	0.724
	Kernel : Polynomial, Parameter C: 1, gamma : 1	0.66
	Kernel : RBF, Parameter C: 1, gamma : 1	0.73
<b>Model 2</b> (80 : 10)	Kernel : Linear, Parameter C: 1, gamma : 1	0.736
	Kernel : Sigmoid, Parameter C: 1, gamma : 1	0.732
	Kernel : Polynomial, Parameter C: 1, gamma : 1	0.669
	Kernel : RBF, Parameter C: 1, gamma : 1	0.745

#### E. Evaluation

Dilihat dari hasil pengujian di atas, dari model 1 dengan menggunakan data latih dan data uji 90 : 10 nilai akurasi yang terbesar yaitu pada kernel RBF dengan parameter C: 1 dan gamma : 1 sebesar 0.73. Selanjutnya dilakukan evaluasi menggunakan *confusion matrix* yang menghasilkan keluaran sebagai berikut:



Gambar 4. Confusion Matrix Kernel RBF

Pada gambar diatas confusion matriks memberikan informasi data berupa *true positive* (TP) sebanyak 504, *false negative* sebanyak 57, *false positive* sebanyak 213 dan *true negative* 226. Dari gambar *confusion matrix* diatas diperoleh nilai akurasi 0.73, presisi 0.8, recall 0.51 dan f1-score 0.63.

Pada model 2 menggunakan data latih dan data uji yaitu 80 : 20, terlihat nilai akurasi terbesar pada RBF sebesar 0.745 dengan parameter C : 1 dan gamma : 1. Evaluasi model menggunakan *confusion matrix*, dapat dilihat pada gambar di bawah ini.



kernel antara lain linear, sigmoid, Polinomial, RBF dengan parameter C:1 dan gamma:1. Hasil akurasi dari pengujian tersebut yaitu untuk model 1 kernel linear 0.727, sigmoid 0.724, Polinomial 0.66 dan RBF 0.73, sedangkan untuk model 2 hasil akurasi untuk kernel linear 0.736, sigmoid 0.732, Polinomial 0.669 dan RBF 0.745. Dari hasil diatas dapat dilihat nilai akurasi terbesar yaitu model 2 dengan menggunakan kernel RBF, Parameter C: 1, gamma:1 dengan nilai 0.745. Kontribusi dalam penelitian ini mengusulkan parameter yang tepat pada parameter C dan Gamma pada kernel *Support vector Machine* untuk menghasilkan akurasi yang baik. Pada penelitian selanjutnya dapat dilakukan optimasi untuk menghasilkan akurasi yang lebih baik lagi.

## Referensi

- Afdal, M., & Elita, L. R. (2022). Penerapan Text Mining Pada Aplikasi Tokopedia Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, 8(1), 78. <https://doi.org/10.24014/rmsi.v8i1.16595>
- Alhaq, Z., Mustopa, A., Mulyatin, S., & Santoso, J. D. (2021). Penerapan Metode Support Vector Machine Untuk Analisis Sentimen Pengguna Twitter. *JOISM: Jurnal of Information System Manajemen*, 3(1), 16–21. <https://doi.org/10.37595/mediainfo.v20i2.59>
- Anggraini, N., Surya Negara Harahap, E., & Basuki Kurniawan, T. (2021). Text Mining-Text Analysis Related to COVID-19 Vaccination Issues. *Jurnal Ilmu Pengetahuan Dan Teknologi Komunikasi*, 23(2), 141–153. <http://dx.doi.org/10.33169/iptekom.23.2.2021.141-153>
- Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 6(2), 129. <https://doi.org/10.30998/string.v6i2.10133>
- Arrafiq, M. S., & Kurniawan, R. (2024). Analisis Sentimen Pengguna Terhadap Layanan Aplikasi Seabank Indonesia di Instagram Menggunakan Metode Support Vector Machine. *Journal of Information System ...*, 5(4), 1280–1291. <https://doi.org/10.47065/josh.v5i4.5620>
- Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(1), 147. <https://doi.org/10.25126/jtiik.0813944>
- Asshiddiqi, M. F., & Muslim, L. K. (2020). perbandingan metode Decision Tree dan Support Vector Machine untuk Analisis Sentimen pada Instagram Mengenai Kinerja PSSI. *E-Proceeding of Engineering*, 7(3), 9936–9948.
- Astuti, K. C., Firmansyah, A., & Riyadi, A. (2024). Implementasi Text Mining Untuk Analisis Sentimen Masyarakat Terhadap Ulasan Aplikasi Digital Korlantas Polri pada Google Play Store. *REMIK: Riset Dan E-Jurnal Manajemen Informatika Komputer*, 8(1), 383–394.
- Atika, D., Ari Aldino, A., Informasi, S., Pagar Alam No, J., Ratu, L., & Kedaton, K. (2022). Term Frequency-Inverse Document Frequency Support Vector Machine Untuk Analisis Sentimen Opini Masyarakat Terhadap Tekanan Mental Pada Media Sosial Twitter. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 3(4), page-page. <http://jim.teknokrat.ac.id/index.php/JTSI>
- Azizah, H. N. (2023). Visualisasi Analisis Sentimen Siberbullying Pada Post Instagram Menggunakan Orange Data Mining. *The Indonesian Journal of Computer Science Research*, 2(1), 42–48. <https://doi.org/10.59095/ijcsr.v2i1.15>
- Borg, A., & Boldt, M. (2020). Using VADER sentiment and SVM for predicting customer response sentiment. *Expert Systems with Applications*, 162, 113746. <https://doi.org/10.1016/j.eswa.2020.113746>
- Darmawan, I. P. D. W., Pradnyana, G. A., & Pascima, I. B. N. (2023). Optimasi Parameter Support Vector Machine Dengan Algoritma Genetika Untuk Analisis Sentimen Pada Media Sosial Instagram. *SINTECH (Science and Information Technology) Journal*, 6(1), 58–67. <https://doi.org/10.31598/sintechjournal.v6i1.1245>
- Gupta, V., & Rattan, P. (2023). Improving Twitter Sentiment Analysis Efficiency with SVM-PSO Classification and EFWS Heuristic. *Procedia Computer Science*, 230(2023), 698–715. <https://doi.org/10.1016/j.procs.2023.12.125>
- Hermawan, A., Jowensen, I., Junaedi, J., & Edy. (2023). Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine. *JST (Jurnal Sains Dan Teknologi)*, 12(1), 129–137. <https://doi.org/10.23887/jstundiksha.v12i1.52358>
- I Komang Dharmendra, Ricky Aurelius Nurtanto Diaz, Muhamad Samsudin, I Gusti Agung Ngurah Rai Semadi, & I Made Agus Wirahadi Putra. (2023). Text Mining Untuk Mendeteksi Emosi Pengguna Terhadap “Nusantara” Sebagai Nama Ikn. *Jurnal Teknologi Informasi Dan Komputer*, 9(5), 457–463. <https://doi.org/10.36002/jutik.v9i5.2639>
- Isnan, M., Elwirehardja, G. N., & Pardamean, B. (2023). Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model. *Procedia Computer Science*, 227, 168–175. <https://doi.org/10.1016/j.procs.2023.10.514>
- Muttaqin, M. N., & Kharisudin, I. (2021). Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor. *UNNES Journal of Mathematics*, 10(2), 22–27. <http://journal.unnes.ac.id/sju/index.php/ujm>
- Novia Lestari, Ozzy Secio Riza, & Reno Ardin. (2023). Implementation Of Text Mining And Pattern Discovery With Naive Bayes Algorithm For Classification Of Text Documents. *Jurnal Teknologi Informasi & Komunikasi Digital Zone*, 14(x), 88–102.
- Nugraha, D., & Astuti, P. (2023). Analisis Sentimen Cyberbullying pada Sosial Media Instagram Menggunakan Metode

- Support Vector Machine. *Information System for Educators and Professionals*, 8(2), 153–164.
- Nurhaliza Agustina, C. A., Novita, R., Mustakim, & Rozanda, N. E. (2024). The Implementation of TF-IDF and Word2Vec on Booster Vaccine Sentiment Analysis Using Support Vector Machine Algorithm. *Procedia Computer Science*, 234, 156–163. <https://doi.org/10.1016/j.procs.2024.02.162>
- Praghakusma, A. Z., & Charibaldi, N. (2021). Komparasi Fungsi Kernel Metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi). *JSTIE (Jurnal Sarjana Teknik Informatika) (E-Journal)*, 9(2), 88. <https://doi.org/10.12928/jstie.v9i2.20181>
- Puji Astuti, A., Alam, S., & Jaelani, I. (2022). Komparasi Algoritma Support Vector Machine dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo. *Jurnal Bangkit Indonesia*, 11(2), 1–6. <https://doi.org/10.52771/bangkitindonesia.v11i2.196>
- Rachman, F. F., & Pramana, S. (2020). Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter. *Indonesian Health Information Management Journal*, 8(2), 100–109. <https://inohim.esaunggul.ac.id/index.php/INO/article/view/223/175>
- Santoso, H. (2024). Sentiment Analysis Towards Rohingya on Instagram Using Support Vector Machine Method. *Journal of Agrocomplex and Engineering*, 1(1), 32–38.
- Sayogo, D. S., Irawan, B., & Bahtiar, A. (2023). Analisis Sentimen Ulasan Instagram Di Google Play Store Menggunakan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3314–3319. <https://doi.org/10.36040/jati.v7i6.8178>
- Setiawan, A., & Suryono, R. R. (2024). Analisis Sentimen Ibu Kota Nusantara menggunakan Algoritma Support Vector Machine dan Naïve Bayes. *Edumatic: Jurnal Pendidikan Informatika*, 8(1), 183–192. <https://doi.org/10.29408/edumatic.v8i1.25667>
- Tampubolon, T. A. S., & Devega, M. (2025). Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 7(1), 26–36. <http://j-ptiik.ub.ac.id>
- Wang, D., & Zhao, Y. (2020). Using News to Predict Investor Sentiment: Based on SVM Model. *Procedia Computer Science*, 174(2019), 191–199. <https://doi.org/10.1016/j.procs.2020.06.074>
- Wicaksono, F. A., Romadhony, A., & Hasmawati. (2022). Sentiment Analysis of University Social Media Using Support Vector Machine and Logistic Regression Methods. *Ind. Journal on Computing*, 7(2), 15–24. <https://doi.org/10.34818/indojc.2022.7.2.638>
- Widya Utami, N., & Artana, M. (2022). Text Mining Dalam Analisis Sentimen Pembelajaran Daring Di Masa Pandemi Covid 19 Menggunakan Algoritma K-Nearest Neighbor. *Jurnal Informatika Teknologi Dan Sains*, 4(2), 140–148. <https://doi.org/10.51401/jinteks.v4i2.2034>
- Wijaya, D. S., & Widyaningrum, D. (2024). Komparasi metode algoritma klasifikasi pada analisis sentimen komentar cyberbullying di instagram. *Jurnal TEKINKOM*, 7(1), 466–472. <https://doi.org/10.37600/tekinkom.v7i1.1392>
- Winarti, D., Salsabila, F. A., Cahyani, F. A., Saputra, N., Tentua, M. N., & Riyadi, A. (2023). Analisis Sentimen Instagram Vaksinasi Masa Pandemi Covid-19 Menggunakan Metode Naïve Bayes. *JuSiTik : Jurnal Sistem Dan Teknologi Informasi Komunikasi*, 6(1), 42–47. <https://doi.org/10.32524/jusitik.v6i1.787>
- Zulfiqri, R., Sari, B. N., & Padilah, T. N. (2024). Analisis Sentimen Ulasan pengguna Aplikasi Media Sosial Instagram pada Situs Google Play Stor menggunakan Naive Bayes Classifier. *JITET (Jurnal Informatika Dan Teknik Elektro Terapan)*, 12(3), 2965–2973.