

IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN KOMENTAR APLIKASI TANDA TANGAN DIGITAL

IMPLEMENTATION OF SUPPORT VECTOR MACHINE FOR SENTIMENT ANALYSIS COMMENTS DIGITAL SIGNATURE APPLICATION

Mita Oktafani¹, Putri Taqwa Prasetyaningrum²

E-mail : ¹mita.oktafani36@gmail.com , ²putri@mercubuana-yogya.ac.id

^{1,2}Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Mercu Buana Yogyakarta

Abstrak

Ulasan atau review pengguna diolah untuk memahami keluhan pengguna terhadap aplikasi agar tersampaikan kepada pihak pengembang untuk dapat memelihara aplikasi. Dari ulasan yang terdapat di Google Playstore, PT. Privy ID sebagai pengembang aplikasi tanda tangan digital ingin mengetahui tingkat kepuasan publik atau pengguna sekaligus sebagai bahan evaluasi bagi perusahaan. Dari ulasan pengguna dilakukan identifikasi dan klasifikasi, namun hal tersebut dapat menghabiskan waktu dalam proses pengklasifikasiannya karena ulasan pengguna yang berjumlah banyak bahkan bertambah setiap saat. Oleh karena itu, dibutuhkan sistem otomatis analisa sentimen ulasan aplikasi. Sistem ini menerapkan algoritma Support Vector Machine (SVM) sebagai metode untuk menentukan sebuah ulasan mempunyai sentimen positif atau negatif. Analisis dilakukan dengan membandingkan fungsi kernel Radial Basis Function (RBF) dan kernel Linear yang dikombinasikan dengan eksperimen penambahan fitur selection, parameter, dan n-gram. Kesimpulannya adalah penyetelan parameter, dan fitur-fitur n-gram mampu meningkatkan nilai akurasi. Dari hasil percobaan dihasilkan ulasan aplikasi tanda tangan digital kelas negatif sebanyak 4.137 dan kelas positif sebanyak 1.816 pada kernel RBF. Sedangkan pada kernel Linear diperoleh kelas negatif sebanyak 4.049 dan kelas positif sebanyak 1.904 data.

Kata Kunci : Analisis Sentimen, Support Vector Machine, Kernel, RBF, Linear

Abstract

User reviews are processed to understand user complaints about the application so that they are conveyed to the developers to enable them to maintain the application. From the reviews on Google Playstore, PT. Privy ID as a digital signature application developer wants to know the level of public or user satisfaction as well as an evaluation material for the company. From user reviews, identification and classification are carried out, but this can take time in the classification process because of the large number of user reviews which increases all the time. Therefore, we need an automated system for application reviews sentiment analysis. This system applies the Support Vector Machine (SVM) algorithm as a method to determine whether a review has a positive or negative sentiment. The analysis is carried out by comparing the Radial Basis Function (RBF) kernel and Linear kernel function combined with the experiment of adding selection, parameters, and n-gram features. The conclusion shows that parameter adjustment, and n-gram features can increase the accuracy value. From the experimental results, there were 4,137 negative classes and 1,816 positive classes on the RBF kernel. While the Linear kernel obtained 4,049 negative classes and 1,904 data positive classes.

Keywords : Sentiment Analysis, Support Vector Machine, Kernel, RBF, Linear

1. PENDAHULUAN

Dalam situasi masyarakat berbasis data dan informasi, data digital merupakan komponen yang penting dan membutuhkan sistem keamanan yang memadai baik saat penyimpanan ataupun saat pendistribusian data. Sejak adanya pandemi Covid-19, tingkat *fraud* (tindakan kecurangan) dalam ekosistem digital semakin meningkat hingga mencapai 35%. Transaksi *e-commerce* terindikasi mengalami kecurangan yaitu terdapat korban *online financial fraud* (kecurangan transaksi daring) sebanyak 26% dan sebanyak 57% dari masyarakat meyakini adanya *fraud* pada transaksi daring. Dalam permasalahan ini, Pemerintah telah mengeluarkan regulasi untuk melindungi masyarakat saat melakukan transaksi ataupun membuat perjanjian secara elektronik yaitu dengan menggunakan tanda tangan digital. Penggunaan tanda tangan digital dinilai sebagai bentuk transformasi dalam upaya meminimalisir pemalsuan dokumen dan transaksi elektronik[1]. Masyarakat baik berupa individu maupun perusahaan dapat memanfaatkan tanda tangan digital sebagai bentuk perwakilan identitas digital terverifikasi yang sah dan tervalidasi. Saat ini, telah banyak aplikasi yang menawarkan jasa tanda tangan digital di Indonesia antara lain Vida, Digisign, dan PrivyId. PrivyId merupakan salah satu aplikasi tanda tangan digital tersertifikasi pertama di Indonesia pada Tahun 2016. Dengan meningkatnya jumlah pengguna aplikasi tiap tahun maka jumlah ulasan yang ada pada Google Playstore juga ikut meningkat. Ulasan yang diberikan bermacam-macam, dari pengguna yang mengalami kesulitan dalam mengoperasikan, pengguna yang mengalami bug error, hingga pengguna yang sangat terbantu dengan adanya aplikasi tanda tangan digital. Ulasan atau review pengguna sering kali digunakan sebagai acuan dalam mencari informasi mengenai suatu produk atau aplikasi[2]. Ulasan yang diberikan sangat banyak dan beragam serta dapat memberikan dampak positif maupun negatif terhadap pihak pengembang. Dengan jumlah ulasan yang sangat banyak maka diperlukan adanya sistem otomatis untuk mengolah data ulasan. Teknik tersebut adalah analisis sentimen. Analisis sentimen adalah proses pengolahan data tekstual secara otomatis guna memperoleh informasi sentimen dalam suatu kalimat opini. Analisis ini dilakukan untuk melihat kecenderungan opini atau pendapat terhadap sebuah masalah atau objek yaitu kecenderungan opini bersifat negatif atau positif[3].

Pada penelitian terkait, telah dilakukan sentimen analisis terhadap aplikasi RuangGuru melalui media sosial twitter dengan membandingkan tiga algoritma klasifikasi serta penambahan optimasi PSO. Dari hasil analisis sentimen tersebut diperoleh nilai akurasi tertinggi yaitu algoritma SVM dengan optimasi PSO[4]. Rokhman dkk. [5] membandingkan kinerja metode Support Vector Machine dengan metode Decision Tree dalam klasifikasi teks ulasan aplikasi transportasi online. Setelah dilakukan proses klasifikasi dan perbandingan diperoleh hasil akurasi 90.20% untuk metode Support Vector Machine. Amuis dan Affandes [6] dalam penelitiannya melakukan klasifikasi teks yang ada pada twitter sebagai tweet iklan atau bukan iklan dengan menerapkan Metode SVM dengan menggunakan kernel RBF dan penyetelan parameter nilai C dan γ menunjukkan hasil nilai akurasi stabil pada rentang nilai $0 \leq C \leq 3$ dan $0.01 \leq \gamma \leq 10$, sedangkan pada data yang belum dilakukan pemilihan fitur dan akurasi stabil pada rentang nilai $0 \leq C \leq 300$ dan $0.01 \leq \gamma \leq 10$. Penelitian Irma et.al [7] tentang prediksi rating pada ulasan produk kecantikan dengan menggunakan metode klasifikasi Naïve Bayes dan N-gram, menunjukkan hasil bahwa model dengan menerapkan full preprocessing dan kombinasi n-gram mampu menghasilkan akurasi yang lebih baik sebesar 97% pada pengujian toleransi 1 dan sebesar 96% pada pengujian sentimen ulasan. N-gram berpengaruh terhadap hasil akurasi model. Ananda dan Pristyanto [8] melakukan klasifikasi menggunakan algoritma SVM, kernel Linear, dan kernel RBF dalam menangani analisis sentimen kepuasan pengguna layanan internet Biznet. Pengujian dilakukan dengan 3 skenario yaitu pada skenario 1 menggunakan 800, skenario 2 menggunakan 900 data dan skenario 3

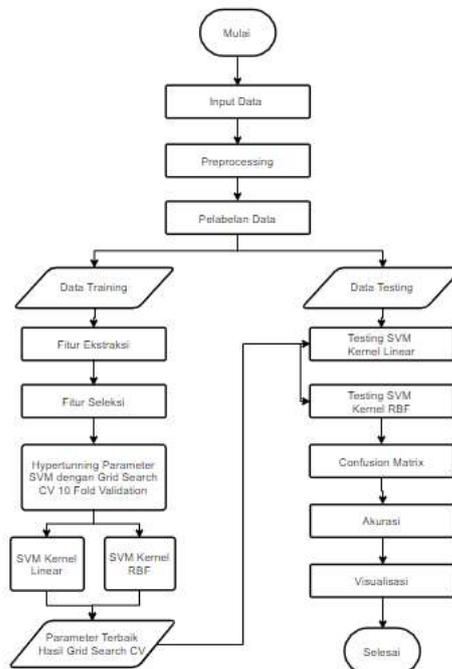
menggunakan 1000 data, dan dilakukan pembagian proporsi data 90% data training dan 10% untuk data testing. Hasil pengujian dari metode Support Vector dengan kedua kernel mempunyai hasil kinerja akurasi, presisi, dan recall yang relatif sama.

Penelitian ini difokuskan pada klasifikasi teks pada analisis data ulasan komentar aplikasi tanda tangan digital dengan membandingkan kinerja metode Support Vector Machine serta mengukur pengaruh dari berbagai eksperimen atau uji coba untuk menemukan metode yang mempunyai performa akurasi terbaik dalam mengklasifikasikan teks sebagai ulasan positif atau negatif.

Harapan peneliti yaitu hasil penelitian dapat membantu pihak pengembang aplikasi dalam penilaian kekurangan dan kelebihan dari aplikasi mereka serta dapat digunakan sebagai bahan evaluasi terhadap pihak pengembang di kemudian hari.

2. METODOLOGI

Pada penelitian ini, terdapat tahap analisis data yang ditunjukkan pada Gambar 1.



Gambar. 1. Diagram Analisis Data

Tahap diawali dengan data scrapping. Tahap kedua yaitu *preprocessing*. Tahap ketiga yaitu pelabelan data dengan menggunakan kamus lexicon. Setelah itu, dilanjutkan tahap keempat yaitu pembagian data menjadi data latih dan uji menggunakan *K-fold Validation* dengan nilai $k=10$. Tahapan kelima yaitu fitur ekstraksi dengan TF-IDF. TF-IDF berfungsi untuk menghitung bobot tiap kata dan mengetahui seberapa sering kata muncul dalam suatu dokumen.[9] Dimana semua kata dalam data diubah menjadi daftar dan ditetapkan ke setiap dokumen sebagai vektor dengan menerapkan n-gram untuk memperoleh hasil fitur terbaik, sementara penggunaan fitur n-gram memberikan hasil yang lebih baik dalam akurasi model klasifikasi teks seperti analisis sentimen.[10]. Tahap keenam adalah proses klasifikasi dengan metode Support Vector Machine. SVM adalah algoritma machine learning dan bekerja sesuai prinsip *Structural Risk Minimization (SRM)* yang bertujuan menemukan hyperplane terbaik serta memaksimalkan jarak antar kelas atau margin.[11]. Tahap ketujuh adalah melakukan pengujian dan evaluasi terhadap data uji dan data testing tersebut untuk

mengetahui kinerja dari model klasifikasi melalui tabel confusion matrix. Selanjutnya dilakukan perhitungan akurasi dari hasil confusion matrix dan terakhir visualisasi.

2.1 Input Data (Dataset)

Dalam penelitian ini menggunakan dataset yang diperoleh dari hasil scrapping situs Google Play Aplikasi Privyid. Dataset tersebut berisikan kumpulan ulasan dari pengguna aplikasi PrivyId. Data yang diambil sebanyak 6.000 data ulasan yang kemudian dilakukan preprocessing data menghilangkan missing value sehingga jumlah data menjadi 5.953.

2.2 Data Preprocessing

Data preprocessing merupakan proses untuk mengubah data tak terstruktur menjadi data terstruktur [12] sehingga proses analisis lebih mudah dan hasil klasifikasi lebih akurat. Berbagai macam metode yang diaplikasikan dalam proses *data preprocessing* adalah sebagai berikut:

a. Case Folding

Digunakan untuk menyeragamkan data agar berbentuk *lowercase*.

```
Case Folding Result :
0 tolong perbaiki, saya sudah mencoba untuk ungg...
1 masuk ke email sy tgl 8 mei dan baru sy baca,,...
2 aplikasi yang sangat bagus.. memudahkan pekerj...
3 verifikasi dokumen memakan waktu yang lama. be...
4 good application.. sangat rekomend dan bagus u...
5 mohon maaf min, ini kenapa saya udah terdaftar?...
6 pertama daftar , di bilang ktp saya udh daftar...
7 aplikasi sampah, masa untuk verifikasi aja sus...
8 wah aplikasi sangat membantu sekali, apalagi d...
9 ini kenapa, ntk ktp saya tidak ditemukan terus ...
Name: content, dtype: object
```

Gambar. 2. Hasil Case Folding

b. Tokenizing

Digunakan untuk menghapus teks spesial, spasi, mention, tag, link, hashtag, menghapus url yang tidak lengkap, menghapus angka, tanda baca, menghapus jarak atau spasi pada setiap kata, menghapus emoji, menghapus duplikat karakter atau huruf.

```
Tokenizing Result :
0 [tolong, perbaiki, saya, sudah, mencoba, untuk...
1 [masuk, ke, email, sy, tgl, mei, dan, baru, sy...
2 [aplikasi, yang, sangat, bagus, memudahkan, pe...
3 [verifikasi, dokumen, memakan, waktu, yang, la...
4 [god, application, sangat, rekomend, dan, bagus...
5 [mohon, maaf, minini, kenapa, saya, udah, torda...
6 [pertama, daftar, di, bilang, ktp, saya, udh, ...
7 [aplikasi, sampah, masa, untuk, verifikasi, aj...
8 [wah, aplikasi, sangat, membantu, sekali, apal...
9 [ini, kenapa, ntk, ktp, saya, tidak, ditemukan, ...
Name: content, dtype: object
```

Gambar. 3. Hasil Tokenizing

c. Convert Slangword (Normalisasi)

Digunakan untuk mengubah kata menjadi bentuk baku.

```
Convert Slangword Result :
0 [tolong, perbaiki, saya, sudah, mencoba, untuk...
1 [masuk, ke, email, saya, tanggal, mei, dan, ba...
2 [aplikasi, yang, sangat, bagus, memudahkan, pe...
3 [verifikasi, dokumen, memakan, waktu, yang, la...
4 [god, application, sangat, rekomend, dan, bagus...
5 [mohon, maaf, minini, kenapa, saya, sudah, terd...
6 [pertama, daftar, di, bilang, ktp, saya, sudah...
7 [aplikasi, sampah, masa, untuk, verifikasi, sa...
8 [wah, aplikasi, sangat, membantu, sekali, apal...
9 [ini, kenapa, ntk, ktp, saya, tidak, ditemukan, ...
Name: content, dtype: object
```

Gambar. 4. Hasil Slangword

d. Negation Handling (Negasi)

Digunakan untuk mengatasi permasalahan negasi terhadap suatu kata.

```

Negation Handling Result :
0 [tolong, perbaiki, saya, sudah, mencoba, untuk...
1 [masuk, ke, email, saya, tanggal, mei, dan, ba...
2 [aplikasi, yang, sangat, bagus, memudahkan, pe...
3 [verifikasi, dokumen, memakan, waktu, yang, la...
4 [god, application, sangat, rekomend, dan, bagus...
5 [mohon, maaf, minini, kenapa, saya, sudah, terd...
6 [pertama, daftar, di, bilang, ktp, saya, sudah...
7 [aplikasi, sampah, masa, untuk, verifikasi, sa...
8 [wah, aplikasi, sangat, membantu, sekali, apal...
9 [ini, kenapanik, ktp, saya, tidak, tidak_ditem...
Name: content, dtype: object
    
```

Gambar. 5. Hasil *Slangword*

e. *Stemming*

Digunakan untuk mengubah kata menjadi kata dasar.

```

Streaming output truncated to the last 5000 lines.
penjelasanyapas : penjelasanyapas
passwordtidak : passwordtidak
balas : balas
keluhan : keluhan
apalah : apa
kerjasama : kerjasama
kejelatannya : jasa
tidak_banyak : tidak banyak
nasabah : nasabah
bahasanya : bahasa
cal : cal
center : center
pantas : pantas
tidak_untuk : tidak untuk
bekerja : kerja
publik : publik
memberi : beri
sebaliknya : balik
rahasia : rahasia
terjadi : jadi
    
```

Gambar. 6. Hasil *Stemming*

f. *Stopwords Removal*

Digunakan untuk menghilangkan kata-kata yang tidak penting.

```

Stopword Removal Result :
0 [tolong, baik, coba, ungha, dokumen, kali, tid...
1 [masuk, email, tanggal, mei, baca, email, tand...
2 [aplikasi, bagus, mudah, pekerjaan, khusus, ker...
3 [verifikasi, dokumen, makan, kali, ulang, alas...
4 [god, application, rekomend, bagus, sukses, moga...
5 [mohon, maaf, min, terdaftarpadahal, blom, daft...
6 [daftar, ktp, daftar, belum daftar, coba, klik...
7 [aplikasi, sampah, verifikasi, susah, ampun, c...
8 [aplikasi, bantu, sat, pandemi, seperti, aplika...
9 [kenapanik, ktp, tidak temu, dukcapil, asli, d...
Name: content, dtype: object
    
```

Gambar. 7. Hasil *Stopword Removal*

g. *Joint Text*

Digunakan untuk mengembalikan kata yang terpecah menjadi bentuk kalimat.

```

0 tolong baik coba ungha dokumen kali tidak bisa...
1 masuk email tanggal mei baca email tanda tanga...
2 aplikasi bagus mudah pekerjaan khusus kerja kan...
3 verifikasi dokumen makan kali ulang alas foto ...
4 god application rekomend bagus sukses moga pelay...
5 mohon maaf min terdaftarpadahal blom daftarturus...
6 daftar ktp daftar belum daftar coba klik rest ...
7 aplikasi sampah verifikasi susah ampun cek ema...
8 aplikasi bantu sat pandemi seperti aplikasi pri...
9 kenapanik ktp tidak temu dukcapil asli dukcapi...
Name: content, dtype: object
    
```

Gambar. 8. Hasil *Join Text*

h. *NaN Removal*

Digunakan untuk menghilangkan data kosong atau null.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5953 entries, 0 to 5952
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   userName    5953 non-null   object
1   content     5953 non-null   object
2   at          5953 non-null   object
3   sentiment   5953 non-null   int64
dtypes: int64(1), object(3)
memory usage: 186.2+ KB
    
```

Gambar. 9. Hasil *NaN Removal*

2.3 Labelisasi Dataset

Setelah proses preprocessing selesai, maka proses selanjutnya adalah labelisasi data ulasan. Proses ini dilakukan dengan mengaplikasikan kamus lexicon [13] yang diperoleh dari <https://github.com/masdevi/ID-OpinionWords>. Kamus lexicon yang digunakan mempunyai lebih dari 5.000 kata umum dalam Bahasa Indonesia yang disertai sentimen negatif dan sentimen positif.

Penentuan label suatu ulasan didasarkan pada nilai penjumlahan secara keseluruhan kata. Jika suatu kata mempunyai nilai = 0, maka kata tersebut tidak termasuk dalam kamus lexicon, sementara jika mempunyai nilai = 1 maka termasuk kedalam kelas positif sedangkan sebaliknya apabila nilai yang diperoleh -1 maka termasuk kedalam kelas negatif.

Hasil proses labelisasi data ulasan dapat dilihat pada Tabel. 1.

Tabel 1. Hasil Labelisasi Data Ulasan

Kelas	Jumlah Data
Positif	4.036
Negatif	1.917

2.4 Splitting Data

Proses klasifikasi yang baik didukung beberapa hal agar dapat memperoleh model machine learning yang mampu memprediksi kelas data baru, yaitu diperlukannya data untuk proses training dan testing. Dataset dibagi menjadi data latih dan data uji dengan menggunakan K-Fold Validation dimana nilai k=10, dan menggunakan ratio 0,1. Hal ini menunjukkan bahwa sebesar 90% data digunakan untuk proses training dan 10% untuk proses testing[14]. K-Fold Validation ini digunakan agar proses training lebih akurat sehingga hasil prediksi akan lebih baik. Hasil proses splitting data dapat dilihat pada Tabel. 2.

Tabel.2. Hasil Pembagian Data

Data	Jumlah
Training	5.357
Testing	596

2.5 Klasifikasi Pada Support Vector Machine (SVM)

Proses ini bertujuan untuk menentukan kelas negatif dan kelas positif. Dalam proses klasifikasi teks, SVM mempunyai fungsi kernel dan parameter [15] yang dapat ditentukan nilainya dan perbedaan nilai pada parameter akan berpengaruh pada kinerja dari model. Oleh karena itu, dilakukan proses pengujian berupa parameter tuning [16] pada setiap fungsi kernel sehingga mampu menghasilkan model machine learning terbaik. Pengujian yang dilakukan yaitu pemberian nilai pada masing-masing parameter yang digunakan yaitu nilai C, Gamma, Percentile, dan N-gram.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Eksperimen Support Vector Machine (SVM)

Hasil pengujian dibuktikan dengan menggunakan tabel confusion matrix pada ke kedua kernel. Proses pemodelan algoritma SVM ini dengan membandingkan data dengan penyetulan parameter tuning, k-fold validation agar didapatkan model klasifikasi terbaik. Hasil pengujian dapat dilihat pada Tabel. 3 – Tabel. 5.

Tabel.3. Perbandingan Confusion Matrix

Model Klasifikasi	Kelas Aktual	Confusion Matrix	
		Kelas Prediksi (0)	Kelas Prediksi (1)
SVM kernel linear dengan N-gram (1,2), seleksi fitur Chi-Square (percentile 75%), parameter C = 1 dan gamma = 0.01	0	183	20
	1	23	370
SVM kernel linear tanpa N-gram, seleksi fitur, parameter C dan gamma	0	168	35
	1	23	370
SVM kernel RBF dengan N-gram (1,2), seleksi fitur Chi-Square (percentile 100%), parameter C = 100 dan gamma = 1	0	179	24
	1	18	375
SVM kernel RBF tanpa N-gram, seleksi fitur, parameter C dan gamma	0	164	39
	1	18	375

Tabel. 4. Perbandingan Akurasi

Model Klasifikasi	Kelas	Akurasi (%)		
		Presisi	Recall	F1-Score
SVM kernel Linear	0	88.83	90.15	89.49
	1	94.87	94.15	94.51
SVM kernel RBF	0	90.86	88.18	89.50
	1	93.98	95.42	94.70

Tabel.5. Hasil Eksperimen

Model Klasifikasi	Akurasi (%)	
	Training	Testing
SVM kernel linear dengan N-gram (1,2), seleksi fitur Chi-Square (percentile 75%), parameter C = 1 dan gamma = 0.01	97.40	92.80
SVM kernel linear tanpa N-gram, seleksi fitur, parameter C dan gamma	95.81	90.26
SVM kernel RBF dengan N-gram (1,2), seleksi fitur Chi-Square (percentile 100%), parameter C = 100 dan gamma = 1	98.90	93.00
SVM kernel RBF tanpa N-gram, seleksi fitur, parameter C dan gamma	98.00	90.43

Hasil di atas menunjukkan bahwa pada kernel RBF mempunyai nilai akurasi terbaik. Pemanfaatan kernel RBF dengan parameter gamma=1 dan C=100, N-Gram (1,2), seleksi fitur Chi Square (Percentile 100%) mampu menghasilkan nilai akurasi sebesar 98.90% pada data training dan 93.00% pada data *testing*.

Precision Recall Curve pada tahap pengklasifikasian data agar didapatkan nilai akurasi yang lebih baik.

5. DAFTAR RUJUKAN

- [1] T. Yuniati and M. F. Sidiq, "Literature Review: Legalisasi Dokumen Elektronik Menggunakan Tanda Tangan Digital sebagai Alternatif Pengesahan Dokumen di Masa Pandemi," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 6, 2020, doi: 10.29207/resti.v4i6.2502.
- [2] N. Herlinawati, Y. Yuliani, S. Faizah, W. Gata, and S. Samudi, "Analisis Sentimen Zoom Cloud Meetings di Play Store Menggunakan Naïve Bayes dan Support Vector Machine," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, p. 293, 2020, doi: 10.24114/cess.v5i2.18186.
- [3] I. Rozi, S. Pramono, and E. Dahlan, "Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi," *J. EECCIS*, vol. 6, no. 1, pp. 37–43, 2012.
- [4] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *J. Teknoinfo*, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.
- [5] U. A. Purwokerto and K. Kunci, "PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN DECISION TREE UNTUK ANALISIS SENTIMEN REVIEW KOMENTAR PADA APLIKASI TRANSPORTASI ONLINE Abstraksi Keywords : Pendahuluan Tinjauan Pustaka," *JOISM J. Inf. Syst. Manag. Vol. 3, No. 1.*, vol. 3, no. 1, pp. 1–7, 2021.
- [6] M. . Imelda A.Muis & Muhammad Affandes, "Penerapan Metode Support Vector Machine (SVM) Menggunakan Kernel Radial Basis Function (RBF) Pada Klasifikasi Tweet," *Sains, Teknol. dan Ind. Sultan Syarif Kasim Riau*, vol. 12, no. 2, pp. 189–197, 2015.
- [7] I. Pujadayanti, M. A. Fauzi, and Y. A. Sari, "Prediksi Rating Otomatis pada Ulasan Produk Kecantikan dengan Metode Naïve Bayes dan N-gram," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4421–4427, 2018.
- [8] F. D. Ananda and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine Sentiment Analysis of Twitter Users on Internet Service Providers Using Support Vector Machine Algorithm," vol. 20, no. 2, pp. 407–416, 2021, doi: 10.30812/matrik.v20i2.1130.
- [9] M. Nurjannah and I. Fitri Astuti, "PENERAPAN ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK TEXT MINING Mahasiswa S1 Program Studi Ilmu Komputer FMIPA Universitas Mulawarman Dosen Program Studi Ilmu Komputer FMIPA Universitas Mulawarman," *J. Inform. Mulawarman*, vol. 8, no. 3, pp. 110–113, 2013.
- [10] S. Khomsah and A. S. Aribowo, "Model Text-Preprocessing Komentar Youtube Dalam Bahasa Indonesia," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 1, no. 10, pp. 648–654, 2021.
- [11] F. Rahutomo, P. Y. Saputra, and M. A. Fidyawan, "Implementasi Twitter Sentiment Analysis Untuk Review Film Menggunakan Algoritma Support Vector Machine," *J. Inform. Polinema*, vol. 4, no. 2, p. 93, 2018, doi: 10.33795/jip.v4i2.152.
- [12] N. Saputra, "(Sentiment Analisis With Lexicon Preprocessing)," *Din. Inform.*, vol. 7, no. 1, pp. 45–57, 2019.
- [13] M. K. S. S. T and V. E. S. T, "Abstrak," vol. 2, no. 1, pp. 1237–1247, 2015.
- [14] K. Akromunnisa and R. Hidayat, "Klasifikasi Dokumen Tugas Akhir (Skripsi) Menggunakan K-Nearest Neighbor," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol.

- 4, no. 1, p. 69, 2019, doi: 10.14421/jiska.2019.41-07.
- [15] T. Bayu, S. A. Yogyakarta, and O. Arifin, “IMPLEMENTASI METODE FORWARD SELECTION PADA ALGORITMA SUPPORT VECTOR MACHINE (SVM) DAN NAIVE BAYES CLASSIFIER KERNEL DENSITY (STUDI KASUS KLASIFIKASI JALUR MINAT SMA) IMPLEMEN SUPPORT VECTOR MACHINE (SVM) DAN NAIVE BAYES CLASSIFIER MACHINE (SVM) AND NAÏVE BAYES KERNEL DENSITY ALGORITHM,” no. May, 2019, doi: 10.25126/jtiik.201961000.
- [16] R. Puspitasari, “Optimasi Algoritma Klasifikasi Biner dengan Tuning Parameter pada Penyakit Diabetes Mellitus,” pp. 50–59, 2019, doi: 10.30864/eksplora.v9i1.257.