

Analisis Sentimen Layanan Perwalian Mahasiswa UMSIDA Menggunakan Metode Support Vector Machine (SVM)

¹Angga Wibawa Saputra, ²Hamzah Setiawan, ³Rohman Dijaya
^{1,2,3}Teknik Informatika, Universitas Muhammadiyah Sidoarjo, Indonesia
[1anggawibawasaputra45@gmail.com](mailto:anggawibawasaputra45@gmail.com); [2hamzah@umsida.ac.id](mailto:hamzah@umsida.ac.id); [3rohman.dijaya@umsida.ac.id](mailto:rohman.dijaya@umsida.ac.id);

Article Info

Article history:

Received, xxx xx xxxx
Revised, xxx xx xxxx
Accepted, xxx xx xxx

Kata Kunci:

Analisis Sentimen
Ekstraksi Fitur TF-IDF
Layanan Perwalian
SVM

ABSTRAK

Aplikasi myUMSIDA mendukung kegiatan akademik di Universitas Muhammadiyah Sidoarjo, namun ulasan mahasiswa menunjukkan adanya keluhan terkait layanan dan fasilitas. Analisis sentimen diperlukan untuk mengklasifikasikan ulasan ini menjadi kategori positif atau negatif, sehingga dapat memberikan wawasan untuk meningkatkan kualitas layanan. Penelitian ini menggunakan algoritma *Support Vector Machine (SVM)* dengan kernel linier yang dikenal memiliki akurasi tinggi, dikombinasikan dengan metode ekstraksi fitur *TF-IDF* untuk meningkatkan klasifikasi teks. Sebanyak 1.300 ulasan dari tahun 2023 diproses melalui tahapan pelabelan, *preprocessing*, *transformasi*, dan *klasifikasi*. Data dibagi menjadi tiga skenario: 70:30, 60:40, dan 50:50 untuk pelatihan dan pengujian. Evaluasi kinerja dilakukan menggunakan *akurasi*, *presisi*, *recall*, dan *F1-Score*. Hasil terbaik diperoleh pada skenario 70:30 dengan *akurasi* sebesar 86,92%, *presisi* 86,60%, *recall* 84,72%, dan *F1-Score* 85,7%. Penelitian ini menunjukkan efektivitas SVM dengan kernel linier dan *TF-IDF* dalam menganalisis sentimen, memberikan dasar untuk meningkatkan layanan pada aplikasi myUMSIDA.

ABSTRACT

The myUMSIDA application supports academic activities at Universitas Muhammadiyah Sidoarjo, but student reviews reveal complaints about its services and facilities. Sentiment analysis is necessary to classify these reviews into positive or negative categories, providing insights to improve service quality. This study uses the Support Vector Machine (SVM) algorithm with a linear kernel, known for its high accuracy, combined with the TF-IDF feature extraction method to enhance text classification. A total of 1,300 reviews from 2023 were processed through labeling, preprocessing, transformation, and classification. Data were split into three scenarios: 70:30, 60:40, and 50:50 for training and testing. Performance was evaluated using accuracy, precision, recall, and F1-Score. The best results were achieved in the 70:30 scenario, with an accuracy of 86.92%, precision of 86.60%, recall of 84.72%, and F1-Score of 85.7%. This study highlights the effectiveness of SVM with a linear kernel and TF-IDF in analyzing sentiment, offering a basis for enhancing the myUMSIDA application's services.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



Penulis Korespondensi:

Angga Wibawa Saputra,
Program Studi Informatika,
Universitas Muhammadiyah Sidoarjo, Indonesia,
Email: anggawibawasaputra45@gmail.com

1. PENDAHULUAN

Berbagai aspek kehidupan tidak dapat lepas dari penggunaan teknologi informasi, termasuk pendidikan. Universitas Muhammadiyah Sidoarjo sebagai sebuah lembaga pendidikan tinggi juga memanfaatkan teknologi informasi dalam kegiatan operasionalnya[1]. Salah satu layanan Universitas Muhammadiyah Sidoarjo yang memanfaatkan adanya teknologi informasi dan sering digunakan oleh mahasiswanya berupa Sistem Informasi Akademik Mahasiswa atau biasa disebut dengan myUMSIDA.

Aplikasi myUMSIDA merupakan sebuah sistem informasi yang digunakan untuk melayani keperluan akademik mahasiswa. Salah satu layanan yang terdapat pada myUMSIDA adalah layanan perwalian dimana mahasiswa memberi ulasan negatif atau positif. Setiap pendapat yang ditulis oleh pengguna atau mahasiswa melalui sebuah keluhan atau ulasan yang memiliki sentimen positif atau negatif.

Copyright © Universitas Muhammadiyah Sidoarjo. This preprint is protected by copyright held by Universitas Muhammadiyah Sidoarjo and is distributed under the Creative Commons Attribution License (CC BY). Users may share, distribute, or reproduce the work as long as the original author(s) and copyright holder are credited, and the preprint server is cited per academic standards.

Authors retain the right to publish their work in academic journals where copyright remains with them. Any use, distribution, or reproduction that does not comply with these terms is not permitted.

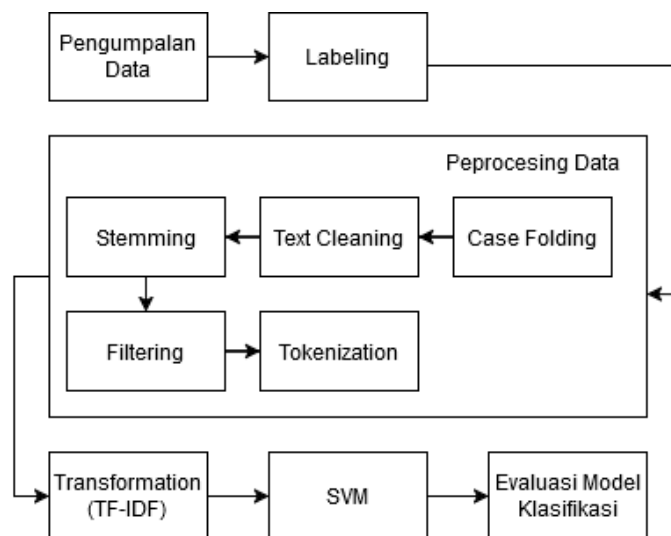
Untuk itu dibutuhkan proses analisis sentimen untuk mengklasifikasikan pendapat yang ditulis melalui keluhan atau ulasan baik dalam pelayanan prasarana dan layanan akademik[2]. Beberapa penelitian sebelumnya telah menunjukkan bahwa metode *Support Vector Machine (SVM)* sangat efektif dalam analisis sentimen. Misalnya, penelitian oleh [3] menunjukkan bahwa *SVM* memiliki akurasi yang tinggi dalam mengklasifikasikan ulasan dari aplikasi Quipper. Penelitian lain oleh [4] menemukan bahwa *SVM* dengan kernel sigmoid mencapai akurasi tertinggi sebesar 80% dalam klasifikasi ulasan e-commerce. Berdasarkan temuan-temuan ini, penggunaan *SVM* dalam penelitian ini dianggap tepat untuk mengklasifikasikan sentimen mahasiswa terhadap layanan perwalian pada aplikasi myUMSIDA.

Analisis Sentimen adalah proses yang dilakukan dengan tujuan untuk melakukan analisis terhadap pendapat atau sebuah opini dan menentukan opini yang diberikan merupakan opini positif atau negatif[5]. Analisis Sentimen digunakan untuk menemukan informasi yang berharga yang dibutuhkan dari data yang tidak terstruktur, sehingga pada penelitian ini juga diharapkan dapat mengetahui apakah opini bersifat negatif atau positif[6]. Metode yang digunakan pada penelitian ini adalah *Support Vector Machine (SVM)* dan penambahan fitur ekstraksi menggunakan *TF-IDF* yang digunakan untuk mengukur sejauh mana sebuah kata penting di sebuah dokumen[7].

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap pendapat atau opini yang diungkapkan melalui keluhan atau ulasan terkait dengan layanan myUMSIDA di Universitas Muhammadiyah Sidoarjo. Dengan menggunakan metode *Support Vector Machine (SVM)* dan penambahan fitur ekstraksi menggunakan *TF-IDF*, penelitian ini bertujuan untuk mengklasifikasikan pendapat yang terungkap dalam keluhan atau ulasan sebagai positif atau negatif.

2. METODE PENELITIAN

Diagram alur penelitian yang ditampilkan pada Gambar 1 memberikan gambaran rinci mengenai setiap tahapan dalam proses analisis sentimen.



Gambar 1 Alur Rancangan Penelitian

Pengumpulan Data

Penelitian ini memanfaatkan data opini yang bersumber dari aplikasi myUMSIDA. Proses Pengumpulan data ini dilakukan dengan cara *Collect Data* dimana pengumpulan data menggunakan metode Data Sekunder dimana data yang telah ada. Pada tahap pengumpulan data, peneliti melakukan pengumpulan data pada tahun 2023 data opini sebanyak 1300 data baik data bersifat negatif atau positif. Data yang terkumpul diambil sebanyak 6 atribut berupa saran, akademik, sarana prasarana, dosen, pelayanan, sentimen.

Labeling

Memberikan kategori atau label pada setiap data teks untuk mengidentifikasi sentimen yang terkandung dalam teks tersebut. Sentimen umumnya dapat dibagi menjadi kategori seperti positif, negatif. Proses labeling dimulai dengan membaca dan memahami konten teks secara keseluruhan untuk menentukan sentimen yang terkandung di dalamnya. Ini melibatkan analisis terhadap pemahaman bahasa, konteks, dan makna dari teks tersebut. Dengan memberikan label pada setiap data teks, analisis sentimen dapat dilakukan secara efektif untuk memahami pandangan, opini, atau perasaan yang terkandung dalam teks tersebut.

Preprocessing

Karena data belum terstruktur dengan baik, data ulasan yang sedang dalam proses labeling belum sepenuhnya siap untuk diproses. Oleh karena itu, preprocessing yang bertujuan untuk menyeragamkan bentuk kata harus dilakukan sehingga data akan lebih terstruktur. Dengan kata lain, teks dilipat menjadi kusus dan semua huruf kapital diubah menjadi huruf kecil.,selanjutnya dilakukan *text cleaning*, *stemming*, *filtering*, dan yang terakhir *tokenization*[8]-[13].

Transformation

Pada tahap transformasi, data teks diubah menjadi bentuk yang mungkin digunakan untuk pengolahan data. Proses ini dilakukan dengan mengonversi teks menjadi vektor berbobot menggunakan metode *TF-IDF* (*Term Frequency-Inverse Document Frequency*)[14], yang berfungsi untuk menentukan seberapa penting suatu kata dalam sebuah dokumen dibandingkan dengan dokumen lain dalam kumpulan data. Setelah transformasi selesai, data dibagi secara acak menjadi data pelatihan dan data pengujian. Data pelatihan digunakan untuk melatih model guna mendapatkan parameter a dan b , yang diperlukan dalam fungsi klasifikasi[15]. Kemudian, data pengujian dimanfaatkan untuk menilai performa model dalam mengelompokkan data berdasarkan sentimen positif atau negatif.

SVM

Support Vector Machine (*SVM*) digunakan untuk mengelompokkan data opini ke dalam kategori positif atau negatif. Penelitian ini memanfaatkan kernel linear karena data dianggap dapat dipisahkan dengan garis pemisah linier. Pelatihan model dilakukan menggunakan data pelatihan yang dihasilkan dari tiga skenario pembagian data[16]. Salah satu parameter yang diuji dalam model ini adalah Cost (C). Parameter C berperan dalam mengendalikan proses pencarian hyperplane terbaik, sekaligus memastikan margin antara kelas positif dan negatif dapat dimaksimalkan untuk meningkatkan keakuratan klasifikasi.

Evaluasi Model Klasifikasi

Evaluasi Model Klasifikasi Evaluasi model dilakukan dengan memanfaatkan matriks kebingungan untuk mengukur kinerja model pada data pengujian. Metode ini menghitung jumlah prediksi yang salah dan benar, serta memberikan informasi terkait kesalahan yang terjadi selama proses klasifikasi[17]. Melalui matriks kebingungan, hasil prediksi dapat dikelompokkan ke dalam dua kategori utama: positif (kelas 1) dan negatif (kelas -1). Tabel 1 berikut menyajikan struktur matriks kebingungan.

Tabel 1 Struktur Matriks Kebingungan		
Hasil Prediksi	Kelas Aktual 1 (Positif)	Kelas Aktual -1 (Negatif)
Kelas 1	True Positif (TP)	False Positif (FP)
Kelas -1	False Negatif (FN)	True Negatif (TN)

3. HASIL DAN PEMBAHASAN

Pengumpulan Data

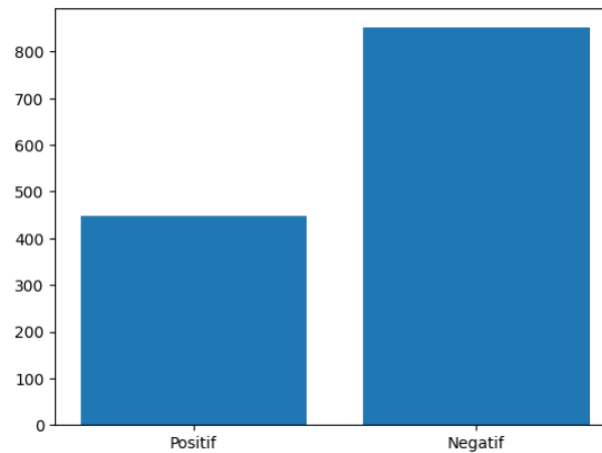
Pada Tabel 1 Menunjukkan data review layanan perwalian aplikasi myUMSIDA data yang telah dikumpulkan menggunakan cara *Collect Data* dimana data yang sudah ada pada laporan akademik aplikasi myUMSIDA dengan total data sebanyak 1300 data keluhan dan ulasan, terkumpul dari tahun 2023 didalamnya memiliki 6 atribut diantaranya saran, akademik, sarana prasarana, dosen, pelayanan, sentimen gambar berikut menunjukkan hasilnya.

	Saran	Akademik	Sarana Pr/Dosen	Pelayanan	Sentimen
1		1	1	0	0 Negatif
2	Fasilitas kurang mendukung	0	0	0	1 Negatif
3	Wifi kurang lancar	0	1	0	0 Negatif
4	Usahakan kalo kepada mahasiswa yg lupa atau ketinggalan membawa stnk tolong diringankan jangan menunggu hingga jam tengah malam kasih or	0	0	0	0 Negatif
5	Mohon untuk wifi di perbaiki lagi	0	0	0	0 Negatif
6	Parkiran kampus	0	1	0	1 Positif
7	Parkiran harus d tata dengan benar msuk rapih tapi baliknya mtorinya dempetan jdi bnyak goresann	0	1	0	0 Negatif
8	sarana dan prasarana penambahan lift	0	1	0	0 Positif
9	Pelayanan baik	0	0	0	1 Positif
10	Kualitas wifi yg tidak memadai	0	1	0	0 Negatif
11	Wifi gak bisa nyambuh	0	0	0	0 Negatif
12	Kampus fasilitas tolong dibenahi Terutama dari wifi yalg sering error gimana mahasiswa bisa maju Karena semua sekarang pakai internet dr tugas sist	0	1	0	0 Negatif
13	Wifi nya lambat	0	0	1	0 Negatif
14	Pelayanan	1	0	0	0 Positif
15	Perpustakaan	0	0	0	1 Positif
16	Wifi lambat	0	0	1	0 Negatif
17	Wifi nya sering tidak stabil	0	0	1	0 Negatif
18	acnya ga dingin	0	1	0	0 Negatif
19	Untuk perpustakaan buku bisa dioerbanyak lagi	0	0	0	1 Negatif
20	pelayanan di perbaiki	1	0	0	0 Negatif
21	Wifi nya kurang lancar untuk digunakan	0	0	1	0 Negatif
22	PUSDAKOM	0	0	1	0 Positif
23	PUSDAKOM	0	0	1	0 Positif
24	Wifi lambat	0	0	1	0 Negatif
25	Pusdakom	0	0	1	0 Positif
26	wifinya	0	0	1	0 Positif

Gambar 2. Data Keluhan dan Ulasan Hasil *Collect Data*

Labeling

Data dari hasil pengumpulan dengan *Collect Data* pada layanan perwalian aplikasi myUMSIDA yang berjumlah 1300 data, menghasilkan pelabelan data berupa hasil negatif sebanyak 851, dan 449 data dengan label positif, gambar berikut menunjukkan hasilnya.



Gambar 3. Grafik Label Data Positif Dan Negatif

Berdasarkan hasil labeling terhadap 1.300 data layanan perwalian aplikasi myUMSIDA, tercatat bahwa sebanyak 851 data (65,46%) tergolong dalam sentimen negatif, sementara 449 data (34,54%) memiliki sentimen positif. Temuan ini mengindikasikan bahwa sebagian besar ulasan atau opini mengenai layanan perwalian bernuansa negatif.

Preprocessing Data

Setelah data selesai melalui proses pelabelan, langkah selanjutnya melibatkan lima tahap pemrosesan, yaitu *case folding*, *text cleaning*, *stemming*, *filtering*, dan *tokenization*. Tabel menunjukkan hasil dari kelima tahapan pengolahan teks tersebut.

Tabel 2. Hasil dari proses *Case Folding*, *Text Cleaning*, *Stemming*, *Filtering*, *Tokenization*

Proses	DataSet	Hasil
<i>Case Folding</i>	Kualitas Wifi yang Tidak memadai Di gedung GKB 5 kampus Candi %*()	kualitas wifi yang tidak memadai di gedung gkb 5 kampus candi %*()
<i>Text Cleaning</i>	kualitas wifi yang tidak memadai di gedung gkb 5 kampus candi %*()	kualitas wifi yang tidak memadai di gedung gkb kampus candi
<i>Stemming</i>	kualitas wifi yang tidak memadai di gedung gkb kampus candi	kualitas wifi yang tidak ada di gedung gkb kampus candi
<i>Filtering</i>	kualitas wifi yang tidak ada di gedung gkb kampus candi	kualitas wifi tidak memadai gedung gkb kampus candi
<i>Tokenization</i>	kualitas wifi tidak memadai gedung gkb kampus candi	['kualitas', 'wifi', 'tidak', 'memadai', 'gedung', 'gkb', 'kampus', 'candi']

- *Case folding* mengubah semua teks menjadi huruf kecil untuk mengurangi perbedaan antara huruf besar dan huruf kecil, sehingga kata-kata seperti "Wifi" dan "wifi" dianggap sama. Misalnya, teks "Kualitas wifi yang Tidak memadai Di gedung GKB 5 kampus Candi %*()" diubah menjadi "kualitas wifi yang tidak memadai di gedung gkb 5 kampus candi %*()".
- *Text cleaning* membersihkan teks dengan menghapus elemen yang tidak relevan, seperti URL dan karakter non-alfabet. Misalnya, teks "kualitas wifi yang tidak memadai di gedung gkb 5 kampus candi %*()" setelah dibersihkan menjadi "kualitas wifi yang tidak memadai di gedung gkb kampus candi".
- *Stemming* adalah proses mengurangi kata-kata ke bentuk dasarnya. Fungsi Porter Stemmer digunakan untuk mengubah kata-kata menjadi bentuk dasar. Misalnya, "memadai" menjadi "ada". Teks "kualitas wifi yang tidak memadai di gedung gkb kampus candi" diubah menjadi "kualitas wifi yang tidak ada di gedung gkb kampus candi".
- *Filtering* menghilangkan kata-kata umum yang tidak mengandung informasi penting, seperti stopwords ("yang", "di"). Fungsi ini digunakan untuk menyaring kata-kata yang tidak relevan dari

teks. Misalnya, "kualitas wifi yang tidak ada di gedung gkb kampus candi" menjadi "kualitas wifi tidak memadai gedung gkb kampus candi".

- *Tokenization* adalah proses memecah teks menjadi unit-unit kecil, yaitu token (kata). Misalnya, "kualitas wifi tidak memadai gedung gkb kampus candi" dipecah menjadi token: ['kualitas', 'wifi', 'tidak', 'memadai', 'gedung', 'gkb', 'kampus', 'candi'].

Transformation Data

Setelah data sudah melakukan *preprocessing* data, kemudian melakukan pembobotan kata dengan *TF-IDF*. Hasil pembobotan kata pada data, gambar berikut ini menunjukkan hasilnya.

	saran	Sentimen	absensi	ac	acnya	ad	ada	adain	adanya	admin	agak	agama	agar	aik	air	airnya	irnyahara	akademik	akan
1	Fasilitas ki	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Wifi kuran	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Usahakan	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Mohon ur	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	Parkiran k	Positif	0	0	0	0	0,415986	0	0	0	0	0	0	0	0	0	0	0	0
6	Parkiran h	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	sarana da	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	Pelayanan	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	Kualitas v	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	Wifi gak b	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	Kampus fa	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	Wifi nya i	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	Pelayanan	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	Perpustakaan	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	Wifi lamb	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	Wifi nya s	Negatif	0	0	0	0	0,333902	0	0	0	0	0	0	0	0,549139	0	0	0	0
17	acnya ga c	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	Untuk per	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	pelayanan	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	Wifi nya k	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	PUSDAKO	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	PUSDAKOI	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	Wifi lemof	Negatif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	Pusdakon	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	wifinya	Positif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Gambar 4 Data Hasil Dari TF-IDF

Dataset dibagi menjadi dua bagian data pengujian dan data pelatihan pada tahap transformasi data dengan tiga rencana pemisahan yang berbeda. Fungsi *train_test_split* digunakan untuk melakukan pemisahan hasil ini. Tabel berikut menunjukkan rincian dari tiga skenario pemisahan.

Scenarion Perbandingan	Training	Testing
70:30	910	390
60:40	780	520
50:50	650	650

Dalam penelitian ini, data dibagi menjadi tiga skenario berdasarkan perbandingan antara data pelatihan dan data pengujian. Skenario pertama memiliki rasio 70:30, di mana 910 data digunakan untuk pelatihan dan 390 data untuk pengujian. Pendekatan ini memberikan lebih banyak data pelatihan untuk membangun model, meskipun data pengujiannya lebih terbatas. Skenario kedua menggunakan rasio 60:40, yang membagi data menjadi 780 untuk pelatihan dan 520 untuk pengujian. Pembagian ini menciptakan keseimbangan yang lebih baik antara jumlah data pelatihan dan pengujian, sehingga evaluasi model dapat dilakukan dengan lebih komprehensif. Skenario terakhir membagi data secara merata dengan perbandingan 50:50. Sebanyak 650 data dialokasikan untuk pelatihan dan 650 lainnya untuk pengujian

Klasifikasi Support Vector Machine

Setelah menyelesaikan tahap preprocessing dan transformation, proses pelatihan dilakukan dengan menggunakan metode SVM. Dalam Langkah ini metode *SVM* digunakan dengan fungsi kernel linier. Pada penggunaan kernel linier dalam model *SVM*, Dalam kernel linier, terdapat satu parameter yang akan diuji, yaitu nilai Cost (C). Angka parameter C ketika diuji mencakup 0,01; 0,2; 0,4; 0,6; dan 1 untuk data pelatihan. Hasil dari pengujian mengenai pengaruh nilai Cost terhadap model linier tabel berikut menunjukkan hasilnya.

Skenario	Evaluasi Model	Cost (C)				
		0.01	0.2	0.4	0.6	1
70:30	Akurasi Keseluruhan	0.65	0.81	0.85	0.86	0.86
60:40	Akurasi Keseluruhan	0.65	0.81	0.84	0.86	0.86
50:50	Akurasi Keseluruhan	0.64	0.80	0.83	0.84	0.85

Pengaruh nilai Cost pada C rendah (0.01) terhadap akurasi keseluruhan cenderung lebih rendah pada semua skenario pembagian data (70:30, 60:40, 50:50), untuk C tinggi (0.06 dan 1) akurasi keseluruhan lebih tinggi, dengan nilai C (0.6 dan 1) memberikan hasil yang serupa pada berbagai skenario, mencapai akurasi sekitar 0.86.

Evaluation

Langkah evaluasi dilakukan dengan memanfaatkan matriks kebingungan (*confusion matrix*). Perbandingan hasil evaluasi dari seluruh skenario klasifikasi metode SVM bersamaan kernel linier tabel berikut menunjukkan hasilnya.

Tabel 5. Evaluasi Skenario

Skenario	Accuracy	Precision	Recall	F1-Score
70:30	86.92%	86.60%	84.72%	85.07%
60:40	86.92%	87.23%	83.54%	84.83%
50:50	85.53%	85.67%	82.41%	83.45%

Secara keseluruhan, model SVM dengan kernel linier bekerja dengan sangat baik pada pembagian data 70:30 dan 60:40, memberikan hasil yang optimal dalam hal accuracy, precision, recall, dan F1-Score, sementara pada pembagian data 50:50, meskipun kinerjanya masih baik, terdapat penurunan di semua metrik, terutama recall, yang menunjukkan kesulitan model dalam mendeteksi kelas positif dengan proporsi data yang lebih seimbang. Hasil evaluasi terbaik di tunjukkan pada skenario ke- 1 menunjukkan performa tertinggi dalam semua metrik evaluasi menghasilkan *accuracy* sebesar 86.92%, *precision* 86.60%, *recall* 84.72%, *f1-score* 85.7%.

4. KESIMPULAN

Sentimen mahasiswa terhadap pengguna aplikasi myUMSIDA menunjukkan hasil 449 ulasan positif dan 851 negatif. Dengan menggunakan tiga skenario pembagian data dan evaluasi kinerja algoritma *SVM* bersamaan kernel linier dalam klasifikasi ulasan aplikasi myUMSIDA menunjukkan bahwa skenario ketiga memiliki hasil terbaik yang menggunakan rasio perbandingan data pelatihan 70% dan data pengujian 30% dan menghasilkan *accuracy* sebesar 86.92%, *precision* 86.60%, *recall* 84.72%, *f1-score* 85.7%.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada dosen pembimbing, Bapak Hamzah Setiawan dan Bapak Rohman Dijaya, yang telah memberikan nasihat dan petunjuk.

REFERENSI

- [1] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "ANALISIS SENTIMEN APLIKASI RUANG GURU DI TWITTER MENGGUNAKAN ALGORITMA KLASIFIKASI," *Jurnal Teknoinfo*, vol. 14, no. 2, p. 115, Jul. 2020, doi: 10.33365/jti.v14i2.679.
- [2] T. J. Pamungkas and A. Romadhony, "Analisis Sentimen Berbasis Aspek Terhadap Ulasan Restoran Berbahasa Indonesia menggunakan Support Vector Machines." Accessed: Nov. 19, 2024. [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15238>
- [3] A. A. Annur, A. Alim Murtopo, and N. Fadilah, "ANALISIS SENTIMEN APLIKASI E-LEARNING QUIPPER SELAMA PANDEMI COVID-19 DENGAN MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM)," 2022.
- [4] A. Rafid Rizqullah, A. Wedhasmara, R. Izwan Heroza, A. Putra, and P. Putra, "ANALISIS MASALAH PADA DATA REVIEW APLIKASI TERHADAP LAYANAN E-COMMERCE MENGGUNAKAN METODE TEXT CLASSIFICATION," 2022.
- [5] J. Liu, X. Luo, P. Lin, and Y. Fan, "Fine-Grained Sentiment Analysis: Recent Progress," 2022. Accessed: Nov. 19, 2024. [Online]. Available: <https://publications.waset.org/10012408/pd>.
- [6] M. Alkaff, A. Rizky Baskara, and Y. Hendro Wicaksono, "Sentiment Analysis of Indonesian Movie Trailer on YouTube Using Delta TF-IDF and SVM," in *2020 5th International Conference on Informatics and Computing, ICIC 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/ICIC50835.2020.9288579.

- [7] D. Darwis, E. Shintya Pratiwi, A. Ferico, and O. Pasaribu, "PENERAPAN ALGORITMA SVM UNTUK ANALISIS SENTIMEN PADA DATA TWITTER KOMISI PEMBERANTASAN KORUPSI REPUBLIK INDONESIA," 2020.
- [8] T. Meisya, P. Aulia, N. Arifin, and R. Mayasari, "PERBANDINGAN KERNEL SUPPORT VECTOR MACHINE (SVM) DALAM PENERAPAN ANALISIS SENTIMEN VAKSINISASI COVID-19," 2021, [Online]. Available: <https://doi.org/10.31598>
- [9] V. W. Fitri, E. Cahyanti, and Q. G. Adiwijaya, "On The Feature Extraction For Sentiment Analysis of Movie Reviews Based on SVM," 2020. doi: 10.1109/ICoICT49345.2020.9166397.
- [10] S. S. Vavilapalli, P. Reddykorepu, S. Saggam, M. Pentyala, and S. A. Devi, "Summarizing Sentiment Analysis on Movie Critics Data," in *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021. doi: 10.1109/ICICT50816.2021.9358563.
- [11] F. Firmansyah *et al.*, "Comparing Sentiment Analysis of Indonesian Presidential Election 2019 with Support Vector Machine and K-Nearest Neighbor Algorithm," in *6th International Conference on Computing, Engineering, and Design, ICCED 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICCED51276.2020.9415767.
- [12] M. Iqbal Ahmadi *et al.*, "SENTIMENT ANALYSIS ONLINE SHOP ON THE PLAY STORE USING METHOD SUPPORT VECTOR MACHINE (SVM)," *Seminar Nasional Informatika*, vol. 2020, 2020, Accessed: Nov. 19, 2024. [Online]. Available: <https://publications.waset.org/10012408/fine-grained-sentiment-analysis-recent-progress>
- [13] Y. Handayani, A. R. Hakim, and Muljono, "Sentiment analysis of Bank BNI user comments using the support vector machine method," in *Proceedings - 2020 International Seminar on Application for Technology of Information and Communication: IT Challenges for Sustainability, Scalability, and Security in the Age of Digital Disruption, iSemantic 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020, pp. 202–207. doi: 10.1109/iSemantic50169.2020.9234230.
- [14] A. Dwi Hartanto, E. Utami, S. Adi, S. Raharjo, M. Yusa, and A. Kamaludin, "Classifying User Personality Based on Media Social Posts Using Support Vector Machine Algorithm Based on DISC Approach," in *2020 2nd International Conference on Cybernetics and Intelligent System, ICORIS 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICORIS50180.2020.9320760.
- [15] *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 2020.
- [16] R. Dave, M. Khare, and S. K. Mitra, "Mutual Information Based Kernel for Image Classification," in *2020 7th International Conference on Soft Computing and Machine Intelligence, ISCMi 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 218–223. doi: 10.1109/ISCMi51676.2020.9311567.
- [17] A. A. Sa'id, Z. Rustam, V. V. P. Wibowo, Q. S. Setiawan, and A. R. Laeli, "Linear Support Vector Machine and Logistic Regression for Cerebral Infarction Classification," in *2020 International Conference on Decision Aid Sciences and Application, DASA 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 827–831. doi: 10.1109/DASA51403.2020.9317065.