

ANALISIS SENTIMEN MAHASISWA TERHADAP SIASAT UKSW BERDASARKAN KUESIONER MENGGUNAKAN METODE LOGISTIC REGRESSION

Robby Adrian Fajar Sulistya^a, Magdalena Ariance Ineke Pakere^b

^a*Fakultas Teknologi Informasi/Universitas Kristen
Satya Wacana, Jawa Tengah*

^a672022146@student.uksw.edu, ^bineke.pakereng@uksw.edu

ABSTRAK

SIASAT adalah sistem informasi akademik utama di Universitas Kristen Satya Wacana (UKSW) yang digunakan oleh mahasiswa untuk keperluan administrasi akademik. Penelitian ini bertujuan menganalisis sentimen mahasiswa terhadap SIASAT menggunakan metode Logistic Regression. Data dikumpulkan melalui kuesioner daring dari 58 responden dan diberi label positif atau negatif. Tahapan penelitian meliputi preprocessing teks (lowercasing, cleaning, tokenisasi, stopword removal, stemming), pembobotan kata menggunakan TF-IDF, penanganan ketidakseimbangan kelas dengan Random Oversampling, serta klasifikasi menggunakan Logistic Regression. Evaluasi dilakukan menggunakan Stratified K-Fold Cross-Validation (5 fold), dengan oversampling diterapkan hanya pada data latih di setiap fold untuk menghindari data leakage. Model menghasilkan rata-rata accuracy sebesar 70% dan F1-score weighted sebesar 70%. Temuan ini mengindikasikan bahwa Logistic Regression cukup menjanjikan untuk klasifikasi sentimen teks berbahasa Indonesia pada dataset sederhana, meskipun terbatas oleh ukuran sampel yang kecil.

Kata kunci : *analisis sentimen, logistic regression, machine learning, SIASAT, TF-IDF*

ABSTRACT

SIASAT is the primary academic information system at Satya Wacana Christian University (UKSW) used by students for academic administrative activities. This study aims to analyze students' sentiment toward SIASAT using a Logistic Regression classifier based on open-ended responses collected through an online questionnaire. A total of 58 textual responses were manually labeled into two classes, namely positive and negative. The research stages include text preprocessing (lowercasing, cleaning, tokenization, stopword removal, and stemming), feature weighting using TF-IDF, handling class imbalance through Random Oversampling, and classification using Logistic Regression. Evaluation was performed using Stratified 5-Fold Cross-Validation, with oversampling applied exclusively to the training data within each fold to prevent data leakage. The model achieved an average accuracy of 70% and a weighted F1-score of 70%. These results suggest that Logistic Regression provides a promising baseline performance for Indonesian text sentiment classification on a small-scale dataset, although its generalizability remains limited due to the small sample size and the manual labeling process.

Keywords: *logistic regression, machine learning, sentiment analysis, SIASAT, TF-IDF*

1. PENDAHULUAN

Perkembangan teknologi telah mendorong peningkatan pemanfaatan sistem informasi dalam bidang pendidikan, baik di tingkat sekolah maupun perguruan tinggi. Sistem informasi pendidikan dirancang untuk mengelola data serta menyajikan informasi yang berkaitan dengan kegiatan akademik [1]. SIASAT (Sistem Informasi Akademik Satya Wacana) merupakan salah satu layanan web utama di Universitas Kristen Satya Wacana (UKSW) yang digunakan oleh mahasiswa dan dosen untuk berbagai keperluan, seperti pendaftaran mata kuliah, pengecekan hasil studi, absensi, pembayaran tagihan, hingga pemantauan keaktifan akademik mahasiswa.

Analisis sentimen merupakan salah satu cabang dari pemrosesan bahasa alami yang berfokus pada penggalian opini, perasaan, dan emosi yang tertuang dalam teks [2]. Analisis ini mengelompokkan ulasan menjadi dua kategori, yaitu positif dan negatif [3]. Penerapan sistem analisis sentimen dapat membantu instansi dalam meningkatkan kualitas layanan berdasarkan umpan balik pengguna yang autentik dan spesifik. Hasil analisis tersebut dapat dimanfaatkan untuk memahami persepsi serta pengalaman pengguna terhadap suatu sistem atau layanan yang digunakan.

Logistic Regression merupakan metode statistik yang digunakan untuk menganalisis hubungan antara variabel bebas dan variabel terikat yang bersifat nominal atau ordinal [4]. Metode ini termasuk dalam regresi linier yang umum digunakan untuk mempelajari keterkaitan antara sejumlah variabel numerik dengan variabel dependen yang bersifat biner atau probabilistik. Dalam

penerapannya, variabel terikat pada Logistic Regression berbentuk biner, dengan pengkodean 1 (Ya) dan 0 (Tidak) [5]. Tujuan utama dari Logistic Regression adalah memperkirakan peluang terjadinya suatu peristiwa berdasarkan variabel-variabel prediktor yang relevan [6].

Penelitian ini menggunakan ulasan atau review dari mahasiswa untuk menggambarkan pengalaman mereka dalam menggunakan web SIASAT. Hasil analisis sentimen yang diperoleh diharapkan dapat memberikan masukan bagi pihak pengembang dalam upaya peningkatan kualitas dan kenyamanan penggunaan web SIASAT. Secara khusus, penelitian ini menerapkan metode klasifikasi Logistic Regression untuk mengidentifikasi dan mengelompokkan ulasan pengguna terhadap web SIASAT.

Beberapa penelitian terdahulu memiliki kemiripan atau kesamaan subjek dan objek yang bisa dijadikan sebagai acuan untuk penelitian ini.

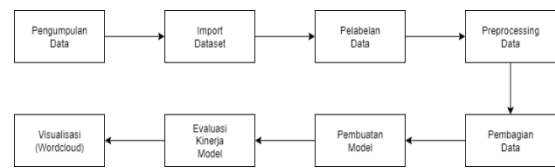
Tabel 1 Penelitian Terdahulu

No	Peneliti & Tahun	Objek Penelitian	Metode	Data	Hasil Utama	Gap yang Ditutupi Penelitian Ini
1	Ratomo & Parker (2018)	SIASAT	User Experience Questionnaire (UEQ)	54 mahasiswa	Mayoritas kategori norma—bad; perspicuity	Belum menegakan opini teks bebas

Berdasarkan Tabel 1, penelitian-penelitian sebelumnya terhadap SIASAT lebih banyak menggunakan pendekatan evaluasi berbasis kuesioner terstruktur seperti UEQ, PIECES, dan Heuristic Evaluation yang menghasilkan skor kuantitatif. Namun, belum terdapat penelitian yang secara khusus mengklasifikasikan sentimen mahasiswa berdasarkan komentar terbuka menggunakan pendekatan Natural Language Processing. Oleh karena itu, penelitian ini bertujuan untuk mengklasifikasikan sentimen mahasiswa terhadap SIASAT ke dalam dua kelas (positif dan negatif) berdasarkan 58 komentar terbuka yang dikumpulkan melalui kuesioner daring. Kontribusi penelitian ini meliputi penerapan Logistic Regression berbasis TF-IDF sebagai baseline classifier untuk teks berbahasa Indonesia, evaluasi stabilitas model menggunakan Stratified 5-Fold Cross-Validation, serta identifikasi kata-kata dominan yang merepresentasikan persepsi mahasiswa terhadap sistem. Ruang lingkup penelitian dibatasi pada komentar kuesioner mahasiswa UKSW dengan pelabelan manual dan ukuran dataset yang relatif kecil.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif, dengan tahapan yang meliputi pengumpulan data, *import dataset*, pelabelan data, *preprocessing data*, pembagian data, pembuatan model menggunakan *Logistic Regression*, evaluasi kinerja menggunakan *Confusion Matrix*, visualisasi hasil dengan *Word Cloud*, serta penyusunan kesimpulan dan saran sebagaimana ditunjukkan pada Gambar 1.



Gambar 1 Tahapan Penelitian

Tahap awal penelitian dimulai dengan pengumpulan data berupa komentar terbuka (open-ended responses) dari mahasiswa Universitas Kristen Satya Wacana (UKSW) terhadap Sistem Informasi Akademik Terpadu (SIASAT). Data dikumpulkan secara sukarela dan anonim melalui Google Form dari 58 responden mahasiswa aktif yang seluruhnya pernah menggunakan SIASAT. Pertanyaan yang diberikan meminta mahasiswa menuliskan pengalaman, pendapat, atau evaluasi mereka terhadap sistem. Data yang diperoleh digunakan hanya untuk keperluan penelitian dan tidak memuat identitas pribadi responden. Responden berasal dari berbagai angkatan (tahun masuk 2020–2024) dan program studi di lingkungan UKSW. Kriteria inklusi penelitian adalah mahasiswa aktif yang telah menggunakan SIASAT minimal satu semester. Pertanyaan utama yang diberikan bersifat terbuka, yaitu: *“Bagaimana pendapat atau pengalaman Anda dalam menggunakan web SIASAT?”* sehingga responden dapat menyampaikan opini secara bebas. Setelah seluruh data terkumpul, langkah berikutnya adalah melakukan proses *import dataset* ke dalam lingkungan kerja *Python*. Dataset yang umumnya berformat *.csv* atau *.xlsx* ini diunggah dan dianalisis untuk kemudian diproses lebih lanjut sesuai kebutuhan model analisis sentimen. Tahap ketiga adalah pelabelan data (data labeling). Dalam analisis sentimen, pelabelan merupakan proses pemberian label terhadap teks berdasarkan emosi atau

opini yang terkandung di dalamnya. Tujuannya adalah untuk mengidentifikasi apakah suatu ulasan mengandung sentimen positif atau negatif yang relevan terhadap topik tertentu [10]. Pelabelan dilakukan oleh satu peneliti berdasarkan definisi operasional sentimen positif dan negatif. Sentimen positif didefinisikan sebagai komentar yang mengandung apresiasi, kepuasan, atau penilaian baik terhadap fitur dan kinerja sistem, sedangkan sentimen negatif didefinisikan sebagai komentar yang berisi keluhan, kritik, atau ketidakpuasan terhadap aspek tertentu dari SIASAT. Hasil pelabelan menghasilkan distribusi sebanyak 35 data positif dan 23 data negatif. Untuk meminimalkan bias subjektivitas, proses pelabelan dilakukan dengan pedoman klasifikasi yang telah ditentukan sebelumnya dan dilakukan pengecekan ulang (self-review) secara menyeluruh terhadap seluruh data setelah proses labeling selesai. Keterbatasan penggunaan satu anotator diakui sebagai salah satu keterbatasan penelitian.

Tahap keempat adalah preprocessing, yaitu langkah untuk mempersiapkan data mentah agar siap digunakan dalam proses analisis maupun pembelajaran mesin. Proses preprocessing mencakup pembersihan data, normalisasi teks, penghapusan tanda baca, stopword removal, tokenization, serta konversi data menjadi format yang sesuai untuk pemrosesan lebih lanjut [11]. Preprocessing dilakukan dengan tahapan lowercasing, pembersihan karakter non-alfabet menggunakan regular expression, tokenisasi, penghapusan stopword bahasa Indonesia, serta stemming untuk mengembalikan kata ke bentuk dasarnya. Proses preprocessing dilakukan menggunakan Python dengan bantuan library re untuk pembersihan karakter, NLTK untuk tokenisasi dan stopword

removal, serta Sastrawi untuk stemming bahasa Indonesia. Setelah preprocessing, teks diubah menjadi representasi numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) dengan parameter $ngram_range=(1,2)$ dan $max_features=500$ untuk membatasi jumlah fitur agar sesuai dengan ukuran dataset.

Tahap kelima adalah pembagian dataset menjadi data latih (training data) dan data uji (testing data). Pembagian dilakukan menggunakan metode Stratified K-Fold Cross-Validation sebanyak 5 fold, yang dipilih karena mampu menjaga proporsi label agar tetap seimbang pada setiap fold. Dengan demikian, distribusi data antara kelas positif dan negatif tetap konsisten, sehingga evaluasi model menjadi lebih andal. Untuk menangani potensi ketidakseimbangan kelas, diterapkan teknik Random Oversampling yang diintegrasikan dalam satu pipeline bersama TF-IDF dan Logistic Regression menggunakan `imblearn.pipeline.Pipeline`. Oversampling dilakukan hanya pada data latih di setiap fold untuk menghindari data leakage, sedangkan data uji tetap dalam kondisi asli. Tahap keenam adalah pembuatan model klasifikasi menggunakan algoritma *Logistic Regression*. Metode ini digunakan untuk menganalisis hubungan antara variabel bebas dan variabel terikat yang bersifat biner, di mana variabel bebas dapat berupa data kontinu maupun kategorik [12]. Proses ini dibagi menjadi dua tahap, yaitu *training*, di mana model dibangun berdasarkan data latih untuk mempelajari pola hubungan antar variabel, dan *testing* yaitu tahap evaluasi kinerja model menggunakan data uji untuk menilai akurasi dan kemampuan generalisasi model. Pada penelitian ini digunakan parameter $C=1.0$, $max_iter=1000$, dengan regularisasi L2 (default), serta solver 'liblinear' yang sesuai

untuk klasifikasi biner pada dataset berukuran kecil. Logistic Regression dipilih karena sesuai untuk klasifikasi dua kelas serta memiliki interpretabilitas yang baik pada dataset berukuran kecil.

Secara matematis, *Logistic Regression* memodelkan probabilitas suatu kelas dengan fungsi sigmoid sebagai berikut.

$$P(y = 1 | x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad (1)$$

Fungsi sigmoid mengubah kombinasi linier variabel input menjadi nilai probabilitas antara 0 dan 1. Model kemudian menentukan kelas 1 jika probabilitas melebihi ambang batas (misalnya 0,5), dan kelas 0 jika berada di bawah ambang batas. Pendekatan ini sangat cocok untuk tugas klasifikasi dua kelas seperti analisis sentimen positif dan negatif.

Tahap ketujuh adalah evaluasi model menggunakan *Confusion Matrix*, yang berfungsi untuk mengukur tingkat akurasi hasil klasifikasi. Pengukuran kinerja model dilakukan dengan beberapa metrik utama, yaitu *accuracy*, *precision*, *recall*, dan *weighted F1-score* [13].

Confusion Matrix merupakan tabel berukuran 2x2 yang menggambarkan hasil prediksi model dibandingkan dengan label aktual. Tabel ini berisi empat nilai utama yaitu

- *True Positif (TP)*
- *True Negatif (TN)*
- *False Positif (FP)*
- *False Negatif (FN)*

Dari matriks tersebut, dihitung metrik evaluasi sebagai berikut:

- $Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$

- $Precision = \frac{TP}{TP + FP} \quad (3)$

- $Recall = \frac{TP}{TP + FN} \quad (4)$

- $F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$

Nilai-nilai ini menunjukkan seberapa tepat model dalam mengklasifikasikan data berdasarkan sentimen yang sebenarnya.

Tahap kedelapan adalah visualisasi hasil menggunakan *Word Cloud*. *Word Cloud* merupakan visualisasi teks yang menampilkan kata-kata dengan ukuran dan warna berbeda sesuai dengan frekuensi kemunculannya dalam dokumen [14]. Visualisasi ini membantu mengidentifikasi kata-kata yang paling sering muncul pada ulasan pengguna, sehingga dapat memberikan gambaran umum mengenai opini dominan terhadap SIASAT. Format visual yang interaktif ini juga mempermudah pembaca dalam memahami pola dan tema utama yang muncul dalam data teks [15].

3. HASIL DAN PEMBAHASAN

Pengumpulan data analisis sentimen pengguna web SIASAT menggunakan metode *Logistic Regression* mendapatkan 58 responden dengan menggunakan *google form*. Responden memberikan persetujuan secara sukarela sebelum mengisi kuesioner. Data bersifat anonim dan hanya digunakan untuk kepentingan penelitian akademik ini. *Dataset* dari responden ini kemudian dimasukkan ke format *file .xlsx*. Kemudian dataset ini dilakukan *labelling* secara manual yang hanya menggunakan 2 label yaitu positif & negatif.

Setelah dilakukan *labelling*, dataset kemudian dilakukan *preprocessing data*, pada tahap *preprocessing* ini dibagi menjadi beberapa tahap, tahap pertama adalah *lowercasing* yang berguna untuk mengubah semua karakter dalam teks menjadi huruf kecil. Tahap kedua adalah *cleaning non-*

alphabet characters yang menggunakan *regular expression*, tahap ini berguna untuk pembersihan teks dengan menghapus seluruh karakter yang bukan merupakan huruf *alphabet*. Tahap ketiga adalah tokenisasi yang berfungsi untuk memecah teks menjadi satuan kata-kata. Tahap keempat adalah *stopword removal*, tahap ini menghapus kata-kata umum yang tidak memiliki makna penting dalam analisis. Tahap kelima adalah *stemming* yang merupakan proses mengubah kata kembali ke bentuk dasarnya. Tahap keenam adalah rekonstruksi teks yang menggabungkan kembali kata-kata yang tersisa menjadi satu string bersih.

Setelah dilakukan *preprocessing data*, dilakukan pembagian data untuk data latih dan data uji. Pembagian data ini menggunakan metode *Stratified K-Fold Cross-Validation*. Tujuan menggunakan metode ini adalah memastikan setiap *subset data* yang digunakan dalam pembagian memiliki distribusi label yang seimbang. Dalam penelitian ini digunakan *5-fold cross-validation*, sehingga proses pelatihan dan pengujian dilakukan sebanyak lima kali dengan kombinasi *fold* yang berbeda, dan setiap *fold* mendapatkan giliran sebagai data pengujian.

Setelah pembagian data dilakukan pembuatan menggunakan *Logistic Regression*. Implementasi dari *Logistic Regression* dalam penelitian ini dilakukan dengan menggunakan *library scikit-learn* dengan dua parameter.

- $C=1.0$: parameter regulasi (*inverse* dari kekuatan regularisasi), di mana nilai yang lebih tinggi menunjukkan regularisasi yang lebih lemah.
- $max_iter=1000$: jumlah maksimum iterasi untuk mencapai konvergensi dalam proses pelatihan model.

Model *Logistic Regression* pada analisis ini dikemas dalam *pipeline* yang terdiri dari tiga tahap, yaitu *TF-IDF Vectorizer*, *Random Oversampling*, dan *Logistic Regression Classifier*. *TF-IDF Vectorizer* memiliki tugas mengubah data teks menjadi representasi numerik dengan memperhitungkan bobot kemunculan kata serta pentingnya kata. *Random Oversampling* memiliki tugas menangani ketidakseimbangan kelas dengan cara memperbanyak sampel dari kelas minoritas, sehingga model tidak bias terhadap kelas mayoritas. Dari total 58 data, terdapat 35 komentar positif (60,34%) dan 23 komentar negatif (39,66%). Distribusi ini menunjukkan adanya ketidakseimbangan kelas, sehingga diterapkan teknik *Random Oversampling* pada data latih untuk mengurangi potensi bias model terhadap kelas mayoritas. Sedangkan *Logistic Regression Classifier* memiliki tugas melakukan proses klasifikasi berdasarkan fitur-fitur yang telah diolah.

Evaluasi kinerja model dilakukan menggunakan teknik *Stratified K-Fold Cross-Validation* dengan *5 fold*, untuk memastikan hasil yang diperoleh bersifat stabil dan merepresentasikan performa model secara umum terhadap seluruh data. Dalam setiap *fold*, model dievaluasi menggunakan metrik akurasi dan *F1-score (weighted average)*, yang dipilih karena mampu memberikan gambaran performa yang lebih adil, khususnya ketika data memiliki distribusi kelas yang tidak seimbang. Tabel berikut menunjukkan hasil evaluasi model pada masing-masing *fold*, termasuk nilai *precision*, *recall*, dan *F1-score* untuk tiap kelas:

Tabel 2 Hasil Evaluasi tiap *Fold*

Fold	Kelas	Precision	Recall	F1-Score
1	Positif	0.67	0.80	0.73
	Negatif	0.83	0.71	0.77
2	Positif	0.67	0.40	0.50
	Negatif	0.67	0.86	0.75
3	Positif	0.80	0.80	0.80
	Negatif	0.86	0.86	0.86
4	Positif	0.40	0.50	0.44
	Negatif	0.67	0.57	0.62
5	Positif	1.00	0.40	0.57
	Negatif	0.67	1.00	0.80

Berdasarkan hasil tersebut, model menunjukkan performa yang cukup stabil, meskipun terdapat sedikit fluktuasi antar *fold*. Rata-rata akurasi keseluruhan mencapai 70%, dan *F1-score* rata-rata sebesar 70%, yang menunjukkan kemampuan model dalam menangani klasifikasi dua kelas secara seimbang.

Tabel 3 Ringkasan Performa Model

Metrik	Mean	Standar Deviasi
Accuracy	0.70	0.09
F1-Score (weighted)	0.70	0.10

Berdasarkan Tabel 3, model menghasilkan rata-rata accuracy sebesar 0.70 dengan standar deviasi 0.09. Nilai standar deviasi

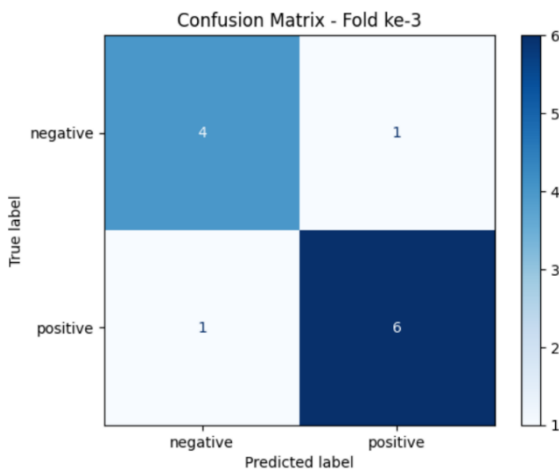
yang relatif moderat menunjukkan bahwa performa model mengalami variasi antar fold, yang kemungkinan dipengaruhi oleh ukuran dataset yang terbatas (58 data). Hal serupa terlihat pada F1-score weighted dengan rata-rata 0.70 dan standar deviasi 0.10, yang mengindikasikan bahwa model cukup konsisten dalam menangani kedua kelas.

Tabel 4 Confusion Matrix Agregat

	Prediksi Negatif	Prediksi Positif
Aktual Negatif	TN = 14	FP = 10
Aktual Positif	FN = 7	TP = 27

Berdasarkan confusion matrix agregat 5-fold cross-validation, diperoleh 27 TP, 14 TN, 10 FP, dan 7 FN. Hasil ini menunjukkan bahwa model menunjukkan kecenderungan lebih sensitif terhadap kelas positif, yang terlihat dari jumlah True Positive yang lebih tinggi dibandingkan True Negative. Namun demikian, nilai False Positive yang relatif tinggi menunjukkan bahwa model masih mengalami kesulitan dalam membedakan komentar negatif yang mengandung kata bernuansa positif. Jumlah FP (10) yang lebih tinggi daripada FN (7) mengindikasikan model cenderung salah mengklasifikasikan komentar negatif sebagai positif, kemungkinan karena adanya kata bernuansa positif dalam konteks keluhan yang tetap diberi bobot tinggi oleh TF-IDF. Meskipun Logistic Regression menunjukkan performa yang cukup baik pada dataset kecil, hasil ini mengindikasikan bahwa peningkatan masih dapat dilakukan melalui penggunaan fitur

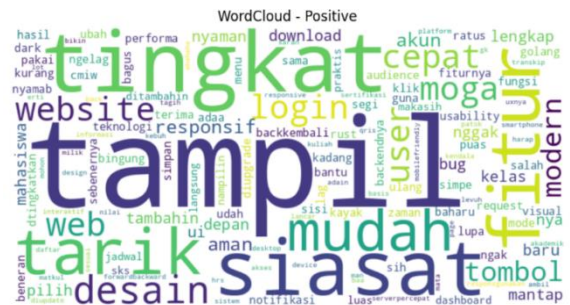
yang lebih kaya secara semantik atau pendekatan yang lebih kontekstual. Selain metrik numerik tersebut, digunakan juga *confusion matrix* untuk mengevaluasi kinerja model dalam memprediksi masing-masing kelas secara lebih rinci. *Confusion matrix* menunjukkan jumlah prediksi benar dan salah untuk tiap kelas, sehingga dapat digunakan untuk mengidentifikasi pola kesalahan model dan potensi bias terhadap kelas tertentu. Dari hasil evaluasi menggunakan *Stratified K-Fold Cross-Validation* sebanyak 5 fold, model mencapai rata-rata akurasi sebesar 70% dan *F1-score* sebesar 70%. Salah satu contoh *confusion matrix* dari proses *cross-validation* ditampilkan pada Gambar 2 sebagai ilustrasi hasil klasifikasi pada salah satu fold. Namun, evaluasi performa model dalam penelitian ini didasarkan pada agregasi seluruh fold untuk memperoleh gambaran yang lebih representatif terhadap kemampuan generalisasi model.



Gambar 2 Hasil Confusion Matrix

Hasil ini menunjukkan bahwa *Logistic Regression* cukup seimbang dalam memprediksi dua kelas, meskipun variasi antar *fold* masih terjadi akibat keterbatasan jumlah data (58 responden). Visualisasi *WordCloud* digunakan untuk menampilkan

kata-kata yang paling sering muncul pada data teks berdasarkan label sentimen yang telah ditentukan secara manual. Pada Gambar 3 dan Gambar 4, diperlihatkan dua *WordCloud* yang merepresentasikan data bersentimen positif dan negatif setelah melalui proses *preprocessing* teks, termasuk pembersihan kata umum dan proses *stemming*.



Gambar 3 Wordcloud Positif



Gambar 4 Wordcloud Negatif

Berdasarkan *WordCloud* sentimen positif, kata-kata yang dominan antara lain “mudah”, “tampil”, “krs”, “cepat”, dan “lengkap”. Dominasi kata “mudah” dan “krs” menunjukkan bahwa mahasiswa menilai proses pengisian KRS relatif mudah dilakukan melalui sistem. Kata “cepat” dan “lengkap” mengindikasikan adanya apresiasi terhadap efisiensi dan kelengkapan fitur yang tersedia. Sebaliknya, pada *WordCloud* sentimen negatif, kata-kata dominan yang muncul antara lain “nggak”, “lama”, “error”, “server”, dan “ambil”. Kemunculan kata “lama” dan “server” mengindikasikan keluhan terkait performa sistem dan stabilitas akses. Kata “error” memperkuat

indikasi bahwa permasalahan teknis menjadi salah satu sumber ketidakpuasan utama mahasiswa.

Selain keterbatasan konteks pada WordCloud, proses stemming juga berpotensi memengaruhi interpretasi sentimen. Pada beberapa kasus, kata turunan yang memiliki nuansa makna berbeda direduksi menjadi bentuk dasar yang sama. Sebagai contoh, kata “kesalahan”, “menyalahkan”, dan “salah” dapat direduksi menjadi bentuk dasar “salah”, meskipun konteks penggunaannya berbeda. Demikian pula kata “memperlambat” dan “lambat” akan dipetakan ke akar kata yang sama, padahal dalam kalimat tertentu dapat memiliki makna yang lebih spesifik. Reduksi bentuk kata ini dapat menyebabkan hilangnya informasi semantik yang penting dan berpotensi memengaruhi proses pembobotan TF-IDF serta keputusan klasifikasi model. Oleh karena itu, meskipun stemming membantu menyederhanakan variasi kata, proses ini juga dapat menjadi sumber ambiguitas dalam analisis sentimen berbasis fitur leksikal.

Secara tematik, hasil ini menunjukkan bahwa aspek kecepatan akses dan kestabilan sistem merupakan prioritas perbaikan utama, sementara aspek kemudahan penggunaan dan fitur KRS menjadi kekuatan utama SIASAT yang perlu dipertahankan. Namun demikian, perlu diperhatikan bahwa WordCloud hanya menunjukkan frekuensi kemunculan kata tanpa mempertimbangkan konteks kalimat. Misalnya, kata “tidak lambat” tetap akan memunculkan kata “lambat”, sehingga interpretasi tetap perlu dikaitkan dengan analisis kuantitatif model klasifikasi.

Berdasarkan hasil analisis sentimen dan WordCloud, dapat diidentifikasi bahwa

aspek kemudahan penggunaan dan proses KRS menjadi kekuatan utama sistem SIASAT. Hal ini terlihat dari dominasi kata “mudah” dan “krs” pada sentimen positif serta tingginya jumlah True Positive dalam confusion matrix. Sebaliknya, keluhan yang paling dominan berkaitan dengan performa sistem, seperti akses yang lambat dan error server. Hal ini menunjukkan bahwa peningkatan infrastruktur server dan optimalisasi performa sistem perlu menjadi prioritas pengelola. Dengan demikian, hasil penelitian ini tidak hanya memberikan evaluasi kuantitatif terhadap sentimen pengguna, tetapi juga menyediakan dasar rekomendasi strategis bagi pengembangan SIASAT di masa mendatang.

Model menunjukkan kecenderungan lebih sensitif terhadap kelas positif, yang terlihat dari jumlah True Positive yang lebih tinggi dibandingkan True Negative. Namun demikian, nilai False Positive yang relatif tinggi menunjukkan bahwa model masih mengalami kesulitan dalam membedakan komentar negatif yang mengandung kata bernuansa positif.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi sentimen komentar mahasiswa terhadap SIASAT menggunakan pendekatan Logistic Regression berbasis representasi TF-IDF serta penanganan ketidakseimbangan kelas melalui Random Oversampling. Berdasarkan evaluasi menggunakan Stratified K-Fold Cross-Validation (5-fold), model menghasilkan rata-rata accuracy sebesar 70% dan F1-score weighted sebesar 70%. Hasil ini menunjukkan bahwa Logistic Regression dapat digunakan sebagai baseline yang cukup stabil untuk klasifikasi sentimen teks berbahasa Indonesia pada dataset berukuran kecil.

Analisis confusion matrix agregat menunjukkan bahwa model cenderung menghasilkan prediksi positif lebih banyak, dengan 27 True Positive dan 14 True Negative, serta masih menghasilkan 10 False Positive dan 7 False Negative. Hal ini mengindikasikan bahwa model masih menghadapi tantangan dalam membedakan komentar negatif yang mengandung kata bernuansa positif. Temuan WordCloud memperlihatkan bahwa aspek kemudahan penggunaan dan proses KRS menjadi kekuatan utama SIASAT, sementara keluhan dominan berkaitan dengan performa sistem seperti akses lambat dan error server. Namun demikian, penelitian ini memiliki keterbatasan pada ukuran dataset yang relatif kecil (58 komentar) dan proses pelabelan yang dilakukan oleh satu peneliti, sehingga hasil sangat bergantung pada kualitas anotasi manual. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar (misalnya ≥ 300 komentar), melibatkan lebih dari satu anotator untuk meningkatkan reliabilitas label, serta mengeksplorasi pendekatan yang lebih kontekstual seperti model berbasis word embedding atau transformer (misalnya IndoBERT). Selain itu, analisis kesalahan (error analysis) secara kualitatif juga penting dilakukan untuk memahami pola mis-klasifikasi secara lebih mendalam.

DAFTAR PUSTAKA

- [1] E. C. Ratmoko and M. A. I. Pakereng, "Analisis User Experience Mahasiswa terhadap Sistem Informasi SIASAT Menggunakan User Experience Questionnaire," *Jurnal TEKINKOM*, vol. 5, no. 1, pp. 11–17, Jun. 2022, doi: 10.37600/tekinkom.v5i1.498.
- [2] T. M. P. Aulia, N. Arifin, and R. Mayasari, "Perbandingan Kernel Support Vector Machine (SVM) dalam Penerapan Analisis Sentimen Vaksinasi COVID-19," *SINTECH Journal*, vol. 4, no. 2, pp. 139–144, Oct. 2021, doi: <https://doi.org/10.31598/sintechjournal.v4i2.762>
- [3] F. A. Larasati, D. E. Ratnawati, and B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 9, pp. 4305–4313, Sep. 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [4] D. N. Ardedia, H. D. Arifin, S. Daniswara, dan A. P. Sari, "Klasifikasi Harga Ponsel Menggunakan Algoritma Logistic Regression," *Journal Informatics and Electronics Engineering*, vol. 4, no. 1, pp. 37–43, Jun. 2024.
- [5] F. Azimah dan K. R. N. Wardani, "Klasifikasi Deteksi Gejala Awal COVID-19 dengan Metode Logistic Regression, Random Forest Classifier dan Support Vector Machine," *Jurnal Locus: Penelitian & Pengabdian*, vol. 1, no. 6, pp. 405–418, Sep. 2022, doi: 10.36418/locus.v1i6.135.
- [6] W. N. Fitriyani, D. R. Amalia, H. H. Handayani, dan A. F. N. Masruriyah, "Aplikasi Berbasis Web Berdasarkan Model Klasifikasi Algoritma SVM dan Logistic Regression Terhadap Data Diabetes," *Riset dan E-Jurnal Manajemen Informatika Komputer (REMik)*, vol. 7, no. 4, pp. 1762–1771, Okt. 2023, doi: 10.33395/remik.v7i4.13001.
- [7] A. W. Adi and E. Maria, "Evaluasi Sistem Informasi Akademik Satya Wacana Menggunakan PIECES Framework," *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, vol. 5, no. 2, pp. 259–271, Nov. 2022. [Online]. Available: <https://ojs.cbn.ac.id/index.php/jukanti/article/view/1028/396>
- [8] S. Anwar, "Evaluasi Pengalaman Pengguna Mahasiswa terhadap Sistem Informasi SIASAT: Pendekatan dengan User Experience Questionnaire," *Jurnal*

- Ilmu Komputer Ruru*, vol. 2, no. 1, pp. 9–16, Jan. 2025, doi: <https://doi.org/10.69688/jikr.v2i1.9>.
- [9] A. L. Ma'dika and E. Mailoa, "Analisis User Experience Aplikasi Siasat Mobile dengan Metode Heuristic Evaluation," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 2, pp. 739–750, Aug. 2023, doi: <https://issn.lipi.go.id/terbit/detail/1560841321>.
- [10] O. N. Julianti, N. Suarna, and W. Prihartono, "Penerapan Natural Language Processing pada Analisis Sentimen Judi Online di Media Sosial Twitter," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 2936–2944, Jun. 2024.
- [11] S. D. Prasetyo, S. S. Hilabi, and F. Nurapriani, "Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN," *Jurnal KomtekInfo*, vol. 10, no. 1, pp. 1–7, 2023, doi: 10.35134/komtekinfo.v10i1.330.
- [12] Marketplace dan E-Commerce dengan pendekatan Binary Logistic Regression," *Jurnal Pendidikan dan Kewirausahaan*, vol. 10, no. 1, pp. 204–220, 2022, doi: 10.47668/pkwu.v9i1.281. Available: <https://journalstkipgrisitubondo.ac.id/index.php/PKWU/article/view/281>
- [13] G. Rininda, I. H. Santi, and S. Kirom, "Penerapan SVM dalam Analisis Sentimen pada Edlink Menggunakan Pengujian Confusion Matrix," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 5, pp. 3335–3340, Oct. 2023.
- [14] M. I. Putri and I. Kharisudin, "Analisis Sentimen Pengguna Aplikasi Marketplace Tokopedia pada Situs Google Play Menggunakan Metode Support Vector Machine (SVM), Naïve Bayes, dan Logistic Regression," *PRISMA, Prosiding Seminar Nasional Matematika*, vol. 5, pp. 759–766, 2022. [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [15] A. S. Afif and A. R. Pratama, "Analisis Sentimen Kebijakan Pendidikan di Masa Pandemi COVID-19 dengan CrowdTangle di Instagram," Universitas Islam Indonesia, Yogyakarta, 2021.