

Review Analysis of SatuSehat Application Using Support Vector Machine and Latent Dirichlet Allocation Modeling

Fikri Fahru Roji^{1*}, Nava Gia Ginasta², Yayan Cahyan³, Dinar Rahayu⁴, Dendi Ramdani⁵

¹Program Studi Bisnis Digital, Fakultas Ekonomi, Universitas Garut, Garut, Jawa Barat, Indonesia

²Program Studi Bisnis Digital, Fakultas Logistik, Teknologi & Bisnis, Universitas Logistik dan Bisnis Internasional, Kota Bandung, Jawa Barat, Indonesia

³Program Studi Manajemen Informatika, Fakultas Teknologi Informasi, Universitas Sebelas April, Sumedang, Jawa Barat, Indonesia

⁴Program Studi Sistem Informasi, Institut Pendidikan Indonesia, Garut, Jawa Barat, Indonesia

⁵Program Studi Rekayasa Perangkat Lunak, Fakultas Komunikasi dan Informasi, Universitas Garut, Garut, Jawa Barat, Indonesia

*Penulis koresponden, e-mail : fikri@uniga.ac.id

Abstract: *SatuSehat is a contact tracing application that replaces the PeduliLindungi application initiated by the Government of Indonesia with the aim of tracking the Covid -19 Virus. The success of the application can be known by analyzing sentiment reviews. In addition to the high number of reviews, there are also other things that need to be highlighted, namely the pattern of reviews that are not in accordance with refined spelling and diverse topics, so that identifying a topic from a collection of reviews is very difficult and takes a lot of time if done manually by humans. This research describes sentiment analysis and topic modeling on SatuSehat app user reviews. By applying Support Vector Machine (SVM) method for sentiment analysis and Latent Dirichlet Allocation (LDA) for topic modeling, this study reveals the views and trends expressed by users. The analyzed review data from Google Play Store includes 171,428 positive reviews and 131,246 negative reviews. The sentiment analysis results indicated the dominance of positive responses. LDA modeling resulted in 8 identified topics, from health concerns to app appreciation. However, negative topics included vaccination challenges, access issues, and app functionality. This research provides insight into users' perceptions of the SatuSehat app, providing a basis for further development and improvement of the app.*

Keywords: *Sentiment Analysis; Topic Modeling; OneHealth App; SVM; LDA*

Abstrak: SatuSehat merupakan aplikasi pelacakan kontak pengganti aplikasi PeduliLindungi yang digagas oleh Pemerintah Indonesia dengan tujuan untuk melakukan pelacakan terhadap Virus Covid-19. Keberhasilan aplikasi dapat diketahui dengan melakukan analisis sentiment review. Selain itu dengan jumlah ulasan yang tinggi, terdapat pula hal-hal lain yang perlu disoroti, yaitu pola ulasan yang tidak sesuai dengan ejaan yang disempurnakan dan topik yang beragam, sehingga untuk mengidentifikasi suatu topik dari kumpulan ulasan tersebut sangat sulit dan menghabiskan banyak waktu jika dilakukan secara manual oleh manusia. Penelitian ini menggambarkan analisis sentimen dan pemodelan topik pada ulasan pengguna aplikasi SatuSehat. Dengan menerapkan metode *Support Vector Machine* (SVM) untuk analisis sentimen dan *Latent Dirichlet Allocation* (LDA) untuk pemodelan topik, penelitian ini mengungkap pandangan dan tren yang diungkapkan oleh pengguna. Data ulasan dari Google Play Store yang dianalisis mencakup 171.428 ulasan positif dan 131.246 ulasan negatif. Hasil analisis sentimen mengindikasikan dominasi tanggapan positif. Pemodelan LDA menghasilkan 8 topik teridentifikasi, dari perhatian terhadap kesehatan hingga apresiasi terhadap aplikasi. Namun, topik negatif mencakup tantangan vaksinasi, masalah akses, dan fungsionalitas aplikasi. Penelitian ini memberikan wawasan tentang persepsi pengguna terhadap aplikasi SatuSehat, memberikan dasar bagi pengembangan lanjutan dan perbaikan aplikasi.

Kata kunci: Analisis Sentimen; Pemodelan Topik; Aplikasi SatuSehat; SVM; LDA

PENDAHULUAN

SatuSehat merupakan aplikasi pelacakan kontak pengganti aplikasi Peduli Lindungi yang digagas oleh Pemerintah Indonesia dengan tujuan untuk melakukan pelacakan terhadap Virus Covid-19 (Aji, Perkasa, and Puspitasari 2022). Aplikasi ini menyediakan berbagai fitur, seperti informasi kesehatan, jadwal pemeriksaan kesehatan, dan pemesanan obat-obatan. Aplikasi ini telah digunakan oleh sejumlah besar masyarakat dan mendapat respon yang beragam. Keberhasilan aplikasi dapat diketahui dengan melakukan analysis sentiment review. Selain itu dengan jumlah ulasan yang tinggi, terdapat pula hal lain yang perlu disoroti, yaitu pola ulasan yang tidak sesuai dengan ejaan yang disempurnakan dan topik yang beragam, sehingga untuk mengidentifikasi suatu topik dari kumpulan ulasan tersebut sangat sulit dan menghabiskan banyak waktu jika dilakukan secara manual oleh manusia. Padahal, kumpulan ulasan tersebut merupakan sumber data yang sangat berpotensi untuk memberikan informasi terkait keberhasilan suatu aplikasi (Aravik, Sopian, and Tohir 2023; Yusuf 2023).

Sentiment analysis atau *opinion mining* adalah bidang studi yang menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi orang terhadap entitas seperti produk, layanan, organisasi, individu, masalah, peristiwa, topik, dan atributnya (Liu 2012). Tujuannya adalah untuk mengekstrak pendapat, emosi, dan penilaian yang diungkapkan oleh individu tentang topik tertentu menggunakan teknik pemrosesan bahasa alami (Idris, Mustofa, and Salihi 2023). *Topic modelling* bertujuan untuk menemukan topik dari kumpulan dokumen. *Topic modelling* adalah teknik yang kuat dalam *text mining* yang memungkinkan untuk menemukan pola dan hubungan tersembunyi dalam data dan dokumen teks. Teknik ini telah banyak digunakan dalam berbagai bidang seperti rekayasa perangkat lunak, ilmu politik, ilmu kedokteran, dan linguistik (Jelodar et al. 2018). *Topic modelling* memberikan beberapa manfaat, termasuk mengidentifikasi topik-topik penting, otomatisasi pengkodean dan pengukuran fenomena skala besar, serta membantu dalam interpretasi dan validasi hasil (Masyhuri 2022). Teknik ini telah berhasil diterapkan dalam berbagai bidang, seperti penelitian kanker, penelitian pemasaran, dan analisis media sosial, untuk mengekstrak informasi relevan dan mengidentifikasi pola-pola (van Buchem et al. 2022; Laureate, Buntine, and Linger 2023; Valle, Osella, and Caselle 2020).

Sentiment Analysis adalah salah satu metode klasifikasi yang mana merupakan salah satu metode dalam Data Mining (Roji and Ramdani 2021). *Sentiment Analysis* dan *topic modelling* merupakan bagian dari proses *text mining*, metode klasifikasi dikarenakan proses keduanya menggunakan data tidak terstruktur atau minimal semi terstruktur. Sejumlah penelitian telah dilakukan untuk menganalisis sentimen pada berbagai aplikasi dan domain menggunakan algoritma *Support Vector Machine* (SVM). Studi oleh (Idris, Mustofa, and Salihi 2023)

memfokuskan pada analisis sentimen ulasan pengguna pada aplikasi Shopee. Studi ini berhasil mengklasifikasikan ulasan menjadi sentimen positif dan negatif dengan akurasi yang memuaskan. Studi lainnya juga mengaplikasikan SVM pada aplikasi lain seperti Grab, Ruangguru, Gopay, Pedulilindungi, iPusnas, WhatsApp, dan BRImo. Penelitian-penelitian ini menunjukkan bahwa SVM efektif dalam menganalisis sentimen pengguna dan dapat ditingkatkan dengan teknik seleksi fitur untuk meningkatkan akurasi (Astuti, Alam, and Jaelani 2022; Chairunnisa, Ernawati, and Santoni 2022; Fauzan, Junaedi, and Setyati 2022; Fitri 2020; Idris, Mustofa, and Salihi 2023; Putra 2015; Rousyati et al. 2022; Saepulrohman, Saepudin, and Gustian 2021; Septiani and Budi 2022; Wahyudi and Kusumawardana 2021).

Penelitian tentang *topic modelling* menggunakan *Latent Dirichlet Allocation* (LDA) telah dilakukan dalam berbagai domain. Misalnya, Tran (Tran et al. 2020) menganalisis literatur penelitian tentang pandemi COVID-19 menggunakan LDA untuk mengidentifikasi fokus penelitian saat ini dan mengusulkan agenda penelitian lanjutan. Mereka menemukan bahwa topik yang paling umum adalah panduan perawatan darurat, patogenesis virus, dan respons global terhadap pandemi. Santoso (Santoso et al. 2022) menggunakan LDA untuk menganalisis hashtag terkait COVID-19 di Instagram dan mengidentifikasi berbagai topik yang dibahas oleh pengguna. Noor Febriansyach (Noor Febriansyach et al. 2021) menerapkan LDA untuk menganalisis cuitan dengan hashtag #covid19 dan mengidentifikasi topik yang berbeda yang dibahas oleh akun bot. LDA juga telah digunakan dalam domain lain seperti ulasan hotel (Suparyati and Utami 2022), ulasan produk (Sugiarto et al. 2022), dan analisis tiket dukungan (Wiranto and Rosyida Uswatunnisa 2022). Studi-studi ini mengindikasikan keefektifan LDA dalam mengidentifikasi dan menganalisis topik dalam dataset yang besar.

Berdasarkan pemaparan di atas, analisis sentimen dan topik yang dibahas dari sebuah reaksi, ulasan, ataupun pendapat yang diberikan masyarakat pada ulasan aplikasi SatuSehat dapat mencerminkan seberapa efisien dan efektifnya aplikasi tersebut. Mengacu pada penelitian-penelitian terdahulu maka penulis melakukan penelitian yang bertujuan untuk menganalisis sentimen dari tanggapan pengguna terhadap aplikasi SatuSehat di Google Play Store menggunakan metode SVM. Setelah dilakukan klasifikasi penulis akan membuat analisis topik berdasarkan kategori sentimen yang dihasilkan dengan metode LDA.

KAJIAN PUSTAKA

A. Aplikasi SatuSehat

Aplikasi PeduliLindungi telah dikembangkan oleh pemerintah Indonesia sebagai respon terhadap pandemi COVID-19. Aplikasi ini berfungsi sebagai aplikasi pelacakan kontak untuk membantu mengendalikan dan membatasi penyebaran virus. Aplikasi ini telah diimplementasikan

di berbagai negara, termasuk Indonesia, dan dianggap membantu dalam mengurangi penyebaran virus (Shantika, Suryanto, and Pratama 2022). Namun, terdapat kekhawatiran terkait perlindungan data pribadi dalam aplikasi PeduliLindungi (Nurhidayati, Sugiyah, and Yuliantari 2021). Kurangnya peraturan khusus mengenai perlindungan data telah menyebabkan ketergantungan pada regulasi yang sudah ada di bidang Teknologi Informasi, Komunikasi, Kesehatan, dan Administrasi Penduduk. Selain itu, penggunaan dan pengalaman pengguna terhadap aplikasi PeduliLindungi telah menjadi subjek penelitian, dengan para peneliti mengevaluasi faktor-faktor seperti kemudahan penggunaan, privasi, kualitas sistem, dan kualitas informasi (Sudimanto, Kurniawan, and Christian 2023). Tingkat kepercayaan dan penerimaan masyarakat yang rendah terhadap aplikasi ini telah menjadi tantangan bagi pemerintah dalam mencapai tujuannya untuk mengurangi penyebaran COVID-19 (Haerani and Rahmatulloh 2021).

B. Analisis sentimen

Analisis sentimen adalah suatu teknik yang digunakan untuk menganalisis opini, sentimen, evaluasi, penilaian, sikap, dan emosi yang diungkapkan oleh individu terhadap suatu entitas seperti produk, layanan, organisasi, atau topik tertentu (Salsabila, Sihombing, and Sitorus 2022). Metode yang digunakan dalam analisis sentimen dapat bervariasi, seperti metode *Naïve Bayes*, *Support Vector Machine* (SVM), dan algoritma *Machine Learning* lainnya (Wibowo, Darmawan, and Amalia 2022). Penelitian-penelitian sebelumnya telah mengaplikasikan analisis sentimen dalam berbagai konteks, seperti analisis sentimen terhadap pembelajaran jarak jauh (As-Salafiyah and Rusydiana 2022), propaganda publik (Sihombing and Setiawan 2022), kebutuhan fitur *fast track* pada *platform streaming* (Waluyan and Hartomo 2022), opini masyarakat terkait vaksin COVID-19 (Prabowo, Pramunendar, and Megantara 2022) dan berbagai topik lainnya. Dengan menggunakan teknik analisis sentimen, penelitian ini dapat memberikan wawasan dan informasi yang berguna dalam berbagai bidang dan konteks.

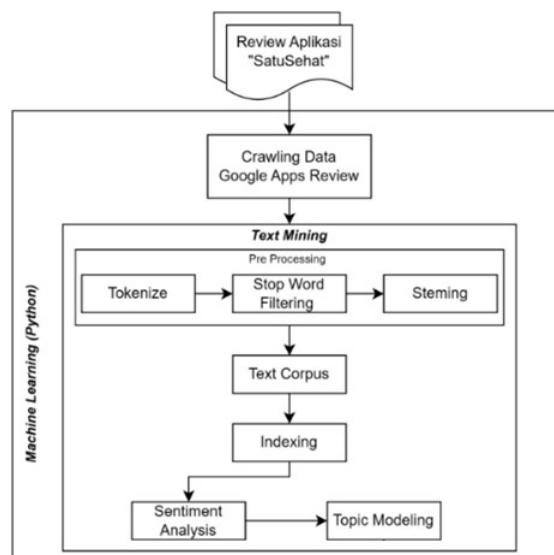
C. Pemodelan Topik

Pemodelan topik adalah teknik yang kuat dalam *text mining* yang memungkinkan penemuan pola-pola tersembunyi dan hubungan dalam data dan dokumen teks. Teknik ini telah banyak diterapkan dalam berbagai bidang seperti rekayasa perangkat lunak, ilmu politik, ilmu kedokteran, dan linguistik. Salah satu metode populer untuk pemodelan topik adalah *Latent Dirichlet Allocation* (LDA), yang telah digunakan dan diteliti secara luas dalam bidang ini (Jelodar et al. 2018). Para peneliti telah mengusulkan model-model berbeda berdasarkan LDA untuk meningkatkan pemodelan topik. Pemodelan topik telah diterapkan untuk merekomendasikan artikel ilmiah, mengelompokkan dokumen, meningkatkan pencarian bibliografis, dan menganalisis volume besar informasi yang tak terstruktur (Zuluaga 2023). Selain itu, pemodelan topik telah

digunakan dalam domain khusus seperti literatur hidrologi, ilmu bahan, dan penelitian pemasaran (Masyhuri 2022). Penerapan pemodelan topik telah menunjukkan hasil yang menjanjikan dalam hal kualitas topik, kemampuan interpretasi, dan performa prediksi (Dieng, Vidal, and Blei 2020). Secara keseluruhan, pemodelan topik telah menjadi teknik yang serbaguna dengan berbagai aplikasi dan telah berkontribusi pada kemajuan berbagai bidang penelitian.

METODE PENELITIAN

Sesuai dengan kerangka penelitian pada Gambar 1. Penelitian ini merupakan penelitian kuantitatif dan kualitatif yang menerapkan konsep dari *big data*. *Big data* merujuk pada fenomena data yang melimpah di dunia daring yang ditandai oleh ukurannya yang besar, kecepatannya, dan perbedaannya dari karakteristik basis data konvensional (Zaluchu 2020). Penggunaan metode campuran, yang melibatkan pengumpulan dan integrasi data kuantitatif dan kualitatif, dapat memberikan hasil yang lebih akurat.



Gambar 1. Metode Penelitian

A. *Crawling Data*

Proses pengambilan data ulasan dari Google Play Store menggunakan bahasa pemrograman Python dan pustaka Google Play Scraper, dengan memanfaatkan fungsi-fungsi pustaka tersebut untuk mengambil dan memanipulasi data ulasan. Pustaka Google Play Scraper menyediakan API yang tidak resmi untuk mengakses Google Play Store (Zimmeck et al. 2019). Dengan memanfaatkan fungsi-fungsi pustaka tersebut, penulis dapat mengekstrak data ulasan dari aplikasi atau kategori tertentu di Google Play Store. Proses pengumpulan data dalam penelitian ini merupakan pengambilan data primer yang diperoleh secara langsung dari sumber asli (tanpa

melalui media perantara), yang dikumpulkan secara khusus oleh peneliti untuk menjawab masalah penelitian.

B. Data Preprocessing

Pada tahap *Preprocessing* penelitian ini dilakukan menggunakan Bahasa pemrograman Python dan Microsoft Excel sebagai tempat penyimpanan data. Proses *preprocessing* data ini krusial untuk memudahkan analisis sentimen dan meningkatkan efektivitas serta efisiensi dalam pengolahan data (Nofiyani and Wulandari 2022). Tahapan proses data *preprocessing* pada sentiment analisis melibatkan beberapa langkah untuk mempersiapkan data sebelum dilakukan analisis sentimen. Beberapa tahapan yang umum dilakukan dalam preprocessing data adalah *case folding*, *tokenizing*, *normalization*, *filtering (stopword removal)*, dan *stemming* (Utami and Artana 2022). Data Transformasi dengan tahapan *Case Folding*, yaitu data-data tersebut akan diubah menjadi sama (lowercase). *Data Cleaning* dengan tahapan *Filtering*, yaitu mengeliminasi kata-kata yang tidak memiliki pengaruh atau tidak informatif. *Tokenizing*, yaitu teks pada data-data yang telah diperoleh akan dipisah untuk tiap katanya. *Normalization* adalah mengubah kata-kata gaul dan typo pada kasus ini, menggunakan Bahasa Indonesia. *Stopwords removal* yaitu dengan menghapus kata-kata yang memiliki informasi rendah dari sebuah teks. *Stemming* adalah mencari makna kata dasar. Kemudian setelah data dibersihkan maka disimpan ulang dengan nama databersih dengan ekstensi csv.

C. Klasifikasi Sentimen dengan SVM

Tahap klasifikasi dengan menggunakan metode SVM, mencari *hyperline* (batas keputusan) terbaik yang memisahkan tiap ulasan kedalam dua kelas yaitu ulasan positif dan ulasan negatif dengan menggunakan algoritma Support Vector Machine. Pada penelitian ini klasifikasi dengan menggunakan metode Support Vector Machine (SVM) merupakan model yang berasal dari teori pembelajaran statistika yang akan memberikan hasil yang lebih baik dibandingkan metode lainnya.

D. Pemodelan Topik dengan LDA

Sebelum membuat pemodelan topik dengan LDA, peneliti terlebih dahulu membuat Wordcloud masing masing dari respon positif dan negatif data agar lebih terlihat *corpus* yang ditampilkan dalam data penelitian ini. Pembuatan wordcloud menggunakan software berbasis python. Setelah corpus ditampilkan dalam bentuk Wordcloud, peneliti melakukan pemodelan topik dengan menggunakan LDA. Proses analisis pegelompokan topik dengan LDA menggunakan bantuan software berbasis python. Karena LDA menggunakan prinsip *partition based clustering*, maka peneliti menentukan terlebih dahulu jumlah kluster dari ulasan yang ada. Jumlah kluster topik pada penelitian ini adalah 5 kluster dengan tujuan agar model yang dibangun tidak

menghasilkan topik yang saling tumpang tindih antar satu dengan yang lain, atau tidak terlalu mendetail sehingga banyak topik yang harus diinterpretasikan.

E. Analisis Sentimen dan Topik Perbincangan

Peneliti melakukan analisis perbandingan antara sentimen positif dan negatif. Peneliti juga menganalisis topik-topik pembicaraan yang mencuat di wordcloud masing masing dari sentimen positif dan negatif. Peneliti juga menganalisis hasil dari LDA yaitu 3 topik pembicaraan teratas yang paling sering dibicarakan netizen melalui kata-kata yang mewakili masing-masing topik.

HASIL DAN PEMBAHASAN

A. Data Selection

Pengambilan data untuk validasi aplikasi SatuSehat dengan menggunakan teknik scraping dan library Google Play Scraper menghasilkan **386.750** data respon. Review, userName, user image, content, at, content, score, thumbsUpCount, reviewCreateVersion, response content, dan RespondAt adalah lima properti dari data yang dikumpulkan. Setelah data ulasan pengguna aplikasi SatuSehat terkumpul, langkah selanjutnya adalah memberikan label pada data dengan skala rating dari 0 hingga 2, yang dilabelkan sebagai negatif, sedangkan rating 3 dilabelkan sebagai netral, dan rating ulasan 4 dan 5 dilabelkan sebagai positif.

B. Preprocessing

Selama tahap ini, data disesuaikan dan dimodifikasi sehingga dapat diproses pada tahap selanjutnya. Tabel 1 menunjukkan contoh hasil dari enam tahap prapemrosesan: *Case Folding*, *Remove Punctuation*, *Tokenizing*, *Normalisasi*, *Filtering*, dan *Stemming*.

Tabel 1. Hasil Preprocessing

Proses	Hasil
<i>Data Collection</i>	Aplikasi gajelas, mau login susahnya mnta ampun
<i>Case Folding</i>	aplikasi gajelas, mau login susahnya mnta ampun
<i>Remove Punctuation</i>	aplikasi gajelas mau login susahnya mnta ampun
<i>Tokenizing</i>	'aplikasi' 'gajelas' 'mau' 'login' 'susahnya' 'mnta' 'ampun'
<i>Normalization</i>	'aplikasi' 'tidak' 'jelas' 'mau' 'login' 'susahnya' 'minta' 'ampun'
<i>Filtering</i>	'aplikasi' 'login' 'susahnya' 'minta' 'ampun'
<i>Stemming</i>	'aplikasi' 'login' 'susah' 'minta' 'ampun'

C. Support Vector Machine classification

Tiga kelas menentukan proses klasifikasi. Yaitu, positif, negatif dan netral dengan SVM multi-kelas satu vs satu. Dataset kemudian dibagi menjadi set pelatihan dan set pengujian dengan rasio 0,1. Artinya, 90% data akan digunakan dalam proses pelatihan dan 10% dalam proses

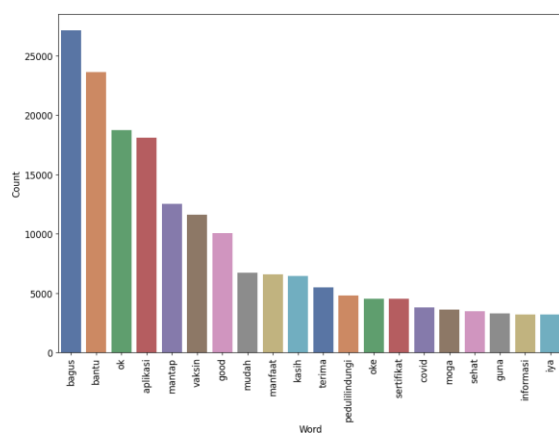
pengujian. Data pelatihan digunakan untuk membangun model menggunakan algoritma SVM. Data testing mengevaluasi performa model yang diperoleh melalui proses training. Pelatihan diawali dengan menghitung nilai awal setiap kelas, yang dilanjutkan dengan menghitung peluang term n pada sebuah dokumen, kemudian dilanjutkan dengan menghitung peluang sebuah dokumen masuk ke dalam sebuah kelas, dan tahap terakhir adalah menentukan kelas dari dokumen tersebut dengan memilih nilai peluang yang paling tinggi. Algoritma SVM memiliki fungsi dan parameter inti yang dapat digunakan selama proses klasifikasi. Berdasarkan hasil klasifikasi sentimen berbasis SVM, maka didapatkan perbandingan persentase ulasan yang cenderung positif. Perbandingan prosentase sentimen dapat dilihat pada Tabel 2. Hasil akurasi pengklasifikasian sentimen berbasis SVM sebesar 82.50%. Hasil klasifikasi ulasan aplikasi didapatkan sebanyak 171.428 ulasan positif, 131.246 ulasan negatif dan 20.819 ulasan netral.

Tabel 2. Hasil klasifikasi SVM

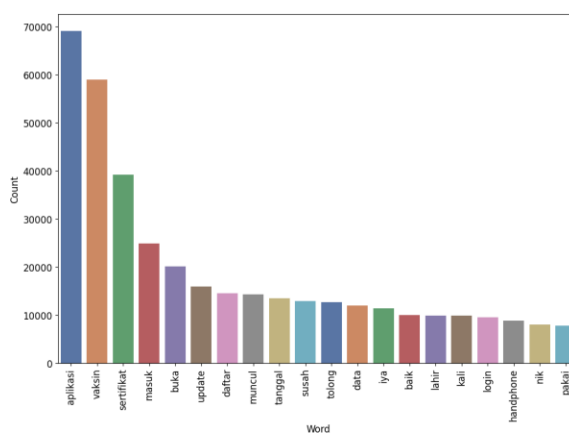
Sentimen	Jumlah Ulasan	Persentase
Positif	171.428	51.78 %
Negatif	131.246	41.87 %
Netral	20.819	6.35 %

Corpus dalam bentuk tabulasi dari data ulasan bersentimen positif dapat dilihat pada gambar

1. Corpus dalam bentuk tabulasi dari data ulasan bersentimen negatif dapat dilihat pada gambar 2.



Gambar 2. Kata paling banyak muncul dari data ulasan bersentimen positif



Gambar 3. Kata paling banyak muncul dari data ulasan bersentimen negatif

D. Topic Modeling dengan LDA

Pemodelan topik dilakukan untuk setiap kelas topik pada penelitian ini untuk mengekstrak informasi dari kumpulan opini pengguna aplikasi. Informasi tersebut akan diinterpretasikan sebagai kumpulan topik-topik sentral dalam kelas positif dan negatif. Tahap awal pemodelan topik

adalah membuat kamus dan korpus untuk kelas positif dan negatif. Data yang digunakan untuk klasifikasi SVM adalah 171.428 ulasan positif dan 131.246 ulasan negatif.

Tahapan penting dalam proses pemodelan topik adalah pembentukan kamus dan korpus untuk data kelas positif dan kelas negatif. Selanjutnya proses pembentukan model LDA dilakukan dengan menggunakan bantuan library *gensim*. Hasil pemodelan topik menggunakan LDA ditunjukkan pada tabel 3. Semakin tinggi nilai koherensi sebuah topik, semakin mudah untuk menginterpretasikan maknanya berdasarkan kumpulan kata yang menyusunnya. Dengan kata lain, semakin sering kata-kata dalam topik tersebut muncul secara bersamaan, maka semakin tinggi nilai koherensi topik tersebut. Berdasarkan tabel 3, dapat diinterpretasikan 8 topik dari pemodelan dengan data sentiment positif dan negatif. Hasil interpretasi topik digunakan untuk melihat tren ulasan pengguna aplikasi SatuSehat.

Tabel 3. Nilai dari Topik Koherensi

Sentimen	Topic	Hasil Nilai-nilai Koherensi	Perkiraan Tema
Positif	0	0.179*"ok" + 0.043*"sehat" + 0.042*"guna" + 0.041*"moga" + 0.028*"daerah" + 0.023*"cepat" + 0.021*"jaga" + 0.019*"jalan" + 0.017*"corona" + 0.014*"lancar"	Kondisi Umum dan Kesehatan
Positif	1	0.293*"bagus" + 0.111*"aplikasi" + 0.043*"oke" + 0.030*"bintang" + 0.030*"banget" + 0.028*"akurat" + 0.028*"virus" + 0.022*"waspada" + 0.018*"top" + 0.016*"lumayan"	Manfaat Penggunaan Aplikasi
Positif	2	0.057*"vaksin" + 0.027*"sertifikat" + 0.026*"aplikasi" + 0.018*"covid" + 0.018*"iya" + 0.014*"coba" + 0.014*"mohon" + 0.013*"data" + 0.013*"merah" + 0.010*"indonesia"	Aplikasi Sebagai Sarana Vaksinasi dan Sertifikat
Positif	3	0.187*"bantu" + 0.074*"mantap" + 0.061*"manfaat" + 0.059*"good" + 0.058*"kasih" + 0.054*"aplikasi" + 0.051*"terima" + 0.038*"informasi" + 0.036*"mudah" + 0.033*"zona"	Apresiasi dan penghargaan terhadap
Negatif	4	0.154*"vaksin" + 0.105*"sertifikat" + 0.029*"muncul" + 0.025*"aplikasi" + 0.024*"otp" + 0.022*"cek" + 0.021*"kode" + 0.020*"sms" + 0.016*"kirim" + 0.016*"daftar"	Sulitnya Vaksinasi dan Pendaftaran Aplikasi
Negatif	5	0.059*"aplikasi" + 0.041*"tanggal" + 0.027*"lahir" + 0.027*"data" + 0.020*"isi" + 0.018*"zona" + 0.016*"baik" + 0.015*"tolong" + 0.013*"merah" + 0.012*"kasih"	Berkaitan dengan Data dan Informasi Personal
Negatif	6	0.086*"masuk" + 0.073*"aplikasi" + 0.043*"handphone" + 0.036*"daftar" + 0.028*"nik" + 0.028*"susah" + 0.027*"pakai" + 0.022*"no" + 0.017*"login" + 0.016*"nama"	Masuk ke Aplikasi dan Identifikasi Pengguna
Negatif	7	0.102*"aplikasi" + 0.100*"buka" + 0.047*"update" + 0.024*"internet" + 0.021*"koneksi" + 0.019*"lokasi" + 0.017*"akses" + 0.016*"error" + 0.016*"lambat" + 0.015*"iya"	Fungsionalitas Aplikasi

KESIMPULAN DAN SARAN

Untuk topik ulasan pengguna SatuSehat, penelitian ini berhasil melakukan analisis sentimen menggunakan metode *Support Vector Machine* dan pemodelan topik menggunakan metode *Latent Dirichlet Allocation* (LDA). Berdasarkan analisis maka dapat disimpulkan bahwa:

1. Berdasarkan hasil analisis sentimen, diperoleh 171.428 ulasan positif, 131.246 ulasan negatif dan 20.819 ulasan netral, yang mengimplikasikan bahwa lebih banyak pengguna yang memberikan tanggapan positif daripada tanggapan negatif.
2. Pemodelan topik dengan algoritma LDA menghasilkan 4 topik ulasan positif dan 4 topik ulasan negatif.
3. Topik positif yang dihasilkan LDA berkaitan dengan perhatian terhadap kesehatan, manfaat aplikasi, peran aplikasi dalam vaksinasi, dan ungkapan apresiasi. Di sisi lain, topik negatif menyoroti tantangan dalam vaksinasi dan pendaftaran aplikasi, kekhawatiran tentang data pribadi, masalah akses ke aplikasi, dan masalah fungsionalitas.

Berdasarkan hasil penelitian ini, terdapat beberapa saran untuk penelitian selanjutnya guna lebih mendalaminya:

1. Analisis Emosi: Melakukan analisis emosi pada ulasan pengguna dapat memberikan wawasan lebih dalam mengenai perasaan dan reaksi emosional dari pengguna terhadap aplikasi. Ini akan memberikan pemahaman yang lebih kaya tentang interaksi pengguna dengan aplikasi.
2. Konteks Geografis: Meneliti sentimen dan topik berdasarkan konteks geografis dapat memberikan gambaran bagaimana tanggapan pengguna bervariasi di berbagai wilayah. Ini dapat membantu dalam pengembangan strategi pemasaran dan peningkatan aplikasi di daerah tertentu.
3. Analisis Perbandingan Metode: Menggunakan metode analisis sentimen dan topik lainnya untuk membandingkan hasil dengan metode SVM dan LDA dalam penelitian ini dapat memberikan pemahaman lebih baik tentang keefektifan dan akurasi metode yang berbeda.

DAFTAR PUSTAKA

- Aji, Arsenius Wisnu, Patria Perkasa, and Maria Puspitasari. 2022. "Public Acceptance of Peduli Lindungi Application Usage Policy." *Jurnal Kebijakan Publik* 13(2): 10–11. <https://jkp.ejournal.unri.ac.idhttps://jkp.ejournal.unri.ac.id>.
- Aravik, Havis, Ahmad Sopian, and Ahmad Tohir. 2023. "Pemanfaatan Aplikasi ResearchGate Sebagai Sumber Literasi Karya Ilmiah." *AKM: Aksi Kepada Masyarakat* 3(2): 187–206.
- As-Salafiyah, Aisyah, and Aam Slamet Rusydiana. 2022. "Analisis Sentiman Atas Pembelajaran Jarak Jauh." *Kompetensi*.
- Astuti, Anggi P, Syariful Alam, and Irsan Jaelani. 2022. "Komparasi Algoritma Support Vector

- Machine Dengan Naive Bayes Untuk Analisis Sentimen Pada Aplikasi BRImo.” *Jurnal Bangkit Indonesia*.
- van Buchem, Marieke M et al. 2022. “Analyzing Patient Experiences Using Natural Language Processing: Development and Validation of the Artificial Intelligence Patient Reported Experience Measure (AI-PREM).” *BMC Medical Informatics and Decision Making* 22(1): 183. <https://doi.org/10.1186/s12911-022-01923-5>.
- Chairunnisa, Cindy, Iin Ernawati, and Mayanda M Santoni. 2022. “Klasifikasi Sentimen Ulasan Pengguna Aplikasi PeduliLindungi Di Google Play Menggunakan Algoritma Support Vector Machine Dengan Seleksi Fitur Chi-Square.” *Informatik Jurnal Ilmu Komputer*.
- Dieng, Adji B, Joaquín R Vidal, and David M Blei. 2020. “Topic Modeling in Embedding Spaces.” *Transactions of the Association for Computational Linguistics*.
- Fauzan, Moch, Hartarto Junaedi, and Endang Setyati. 2022. “Klasifikasi AI – Qur’an Terjemahan Bahasa Indonesia Dengan Menggunakan Algoritma Support Vector Machine (Svm).” *Konvergensi*.
- Fitri, Evita. 2020. “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine.” *Jurnal Transformatika*.
- Haerani, Erna, and Alam Rahmatulloh. 2021. “Analisis User Experience Aplikasi Peduli Lindungi Untuk Menunjang Proses Bisnis Berkelanjutan.” *Satin - Sains Dan Teknologi Informasi*.
- Idris, Irma Surya Kumala, Yasin Aril Mustofa, and Irvan Abraham Salih. 2023. “Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM).” *Jambura Journal of Electrical and Electronics Engineering* 5(1): 32–35.
- Jelodar, Hamed et al. 2018. “Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey.” *Multimedia Tools and Applications*.
- Laureate, Caitlin Doogan Poet, Wray Buntine, and Henry Linger. 2023. “A Systematic Review of the Use of Topic Models for Short Text Social Media Analysis.” *Artificial Intelligence Review*. <https://doi.org/10.1007/s10462-023-10471-x>.
- Liu, Bing. 2012. *Sentiment Analysis and Opinion Mining*. Chicago: Morgan & Claypool.
- Masyhuri, Muhammad. 2022. “Topic Modelling Application for Determining Competitiveness Factors of the Small Business Firms.” *International Journal of Social Science and Business*.
- Nofiyani, Nofiyani, and Wulandari Wulandari. 2022. “Implementasi Electronic Data Processing Untuk Meningkatkan Efektifitas Dan Efisiensi Pada Text Mining.” *Jurnal Media Informatika Budidarma*.
- Noor Febriansyach, Muhammad Alif, Faza Rashif, Gautama Nirvana, and Nur Aini Rakhmawati. 2021. “Implementasi LDA Untuk Pengelompokan Topik Tweet Akun Bot Twitter Bertagar \#Covid-19.” *Cogito Smart Journal*.
- Nurhidayati, Nurhidayati, Sugiyah Sugiyah, and Kartika Yuliantari. 2021. “Pengaturan Perlindungan Data Pribadi Dalam Penggunaan Aplikasi Pedulilindungi.” *Widya Cipta - Jurnal Sekretari Dan Manajemen*.
- Prabowo, Dwi Puji, Ricardus Anggi Pramunendar, and Rama Aria Megantara. 2022. “Prediksi Sentimen Masyarakat Terhadap Penggunaan Vaksin Covid 19 Menggunakan RNN.” *Jurnal Informatika Upgris*.
- Putra, Risyad Ananda. 2015. “Rancang Bangun Website Company Profile Dan Aplikasi Pengolahan Database Customer Pada Pt Bumi Convexindo.”
- Roji, Fikri Fahu, and Dendi Ramdani. 2021. “The Study of Decision Tree Algorithm For Classification of Student Graduation (Case Study: Faculty of Economics, University of Garut).” *RISTEC: Research 2021, Undefined* 2(1): 62–72. <https://journal.institutpendidikan.ac.id/index.php/ristec/article/view/1215>.

- Rousyati, Rousyati, Windu Gata, Dany Pratmanto, and Nia K Wardhani. 2022. "Analisa Sentimen Financial Technology Peer to Peer Lending Pada Aplikasi Koinworks." *Jurnal Teknologi Informasi Dan Ilmu Komputer*.
- Saepulrohman, Asep, Sudin Saepudin, and Dudih Gustian. 2021. "Analisis Sentimen Kepuasan Pengguna Aplikasi Whatsapp Menggunakan Algoritma Naïve Bayes Dan Support Vector Machine." *Is the Best Accounting Information Systems and Information Technology Business Enterprise This Is Link for Ojs Us*.
- Salsabila, Aqila, Jeremia Jordan Sihombing, and Rizky Irvandi Sitorus. 2022. "Implementasi Algoritma Support Vector Machine Untuk Analisis Sentimen Aplikasi OLX Di Playstore." *Journal of Informatics and Data Science*.
- Santoso, Kevin Hendrawan, Asmaul Husna, Nadia Azalia Putri, and Nur Aini Rakhmawati. 2022. "Analisis Topik Tagar Covidindonesia Pada Instagram Menggunakan Latent Dirichlet Allocation." *Jiska (Jurnal Informatika Sunan Kalijaga)*.
- Septiani, Andina, and Indra Budi. 2022. "Klasifikasi Ulasan Pengguna Aplikasi: Studi Kasus Aplikasi Ipusnas Perpustakaan Nasional Republik Indonesia (PNRI)." *Jipi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*.
- Shantika, Neneng, Tri L M Suryanto, and Arista Pratama. 2022. "Analisis Faktor Pendorong Niat Menggunakan Aplikasi PeduliLindungi Menggunakan Model UTAUT Modifikasi." *Jurnal Teknik Informatika Dan Sistem Informasi*.
- Sihombing, Edy Syahputra, and Maria Evangelina Setiawan. 2022. "Kerancuan Berpikir: Propaganda Publik Media Pro-Ukraina Dalam Situasi Perang Rusia-Ukraina 2022." *Sapientia Humana Jurnal Sosial Humaniora*.
- Sudimanto, None, Callista P Kurniawan, and Marcell Christian. 2023. "Analisis Pengaruh Persepsi Kemudahan, Keamanan Privasi, Kualitas Sistem, Serta Kualitas Informasi Terhadap Kepuasan Pengguna Aplikasi Pedulilindungi." *Media Informatika*.
- Sugiarto, Dedy et al. 2022. "Analisis Ulasan Pembelian Produk Beras Di Marketplace Shopee Dengan Menggunakan Topic Modelling." *Explore Jurnal Sistem Informasi Dan Telematika*.
- Suparyati, Suparyati, and Emma Utami. 2022. "Pengamatan Tren Ulasan Hotel Menggunakan Pemodelan Topik Berbasis Latent Dirichlet Allocation." *Jiko (Jurnal Informatika Dan Komputer)*.
- Tran, Bach Xuan et al. 2020. "Studies of Novel Coronavirus Disease 19 (COVID-19) Pandemic: A Global Analysis of Literature." *International Journal of Environmental Research and Public Health*.
- Utami, Nengah W, and Made Artana. 2022. "Text Mining Dalam Analisis Sentimen Pembelajaran Daring Di Masa Pandemi Covid 19 Menggunakan Algoritma K-Nearest Neighbor." *Jurnal Informatika Teknologi Dan Sains*.
- Valle, Filippo, Matteo Osella, and Michele Caselle. 2020. "A Topic Modeling Analysis of TCGA Breast and Lung Cancer Transcriptomic Data." *Cancers* 12(12). <https://www.mdpi.com/2072-6694/12/12/3799>.
- Wahyudi, Rizki, and Gilang Kusumawardana. 2021. "Analisis Sentimen Pada Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine." *Jurnal Informatika* 8(2): 200–207.
- Waluyan, Michel Geren, and Kristoko Dwi Hartomo. 2022. "Analisis Sentimen Kebutuhan Fast Track Pada Originals Vidio Menggunakan Support Vector Machine." *Jatiji (Jurnal Teknik Informatika Dan Sistem Informasi)*.
- Wibowo, Ari, Wachid Darmawan, and Nurul Amalia. 2022. "Komparasi Metode Naïve Bayes Dan K-Nearest Neighbor Terhadap Analisis Sentimen Pengguna Aplikasi Pedulilindungi." *Ic-Tech*.

- Wiranto, Wiranto, and None Mila Rosyida Uswatunnisa. 2022. "Topic Modeling for Support Ticket Using Latent Dirichlet Allocation." *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*.
- Yusuf, Faidah. 2023. "Pelatihan Sitasi Untuk Karya Tulis Ilmiah Mahasiswa Menggunakan Mendeley Reference." *Jurnal Pengabdian Pendidikan Masyarakat (JPPM)* 4(1): 33–38.
- Zaluchu, Sonny Eli. 2020. "Strategi Penelitian Kualitatif Dan Kuantitatif Di Dalam Penelitian Agama." *Evangelikal Jurnal Teologi Injili Dan Pembinaan Warga Jemaat*.
- Zimmeck, Sebastian et al. 2019. "MAPS: Scaling Privacy Compliance Analysis to a Million Apps." *Proceedings on Privacy Enhancing Technologies*.
- Zuluaga, Martha. 2023. "Topic Modeling: Perspectives From a Literature Review." *Ieee Access*.