

Analisis Sentimen Pengguna Aplikasi Mobile JKN Menggunakan Metode Stochastic Gradient Descent (SGD)

Faris Abdi El Hakim^{1,*}, M Adamu Islam Mashuri¹

¹Fakultas Vokasi, Program Studi D4 Manajemen Informatika, Universitas Negeri Surabaya, Surabaya, Indonesia

Email: ¹farishakim@unesa.ac.id, ²mmashuri@unesa.ac.id

(*Email Corresponding Author: farishakim@unesa.ac.id)

Received: 25 Juni 2025 | Revision: 26 Juni 2025 | Accepted: 26 Juni 2025

Abstrak

Kesehatan merupakan aspek penting yang harus dimiliki setiap individu untuk menjalankan aktivitas sehari-hari dan mendukung kesejahteraan masyarakat. Dalam upaya meningkatkan layanan kesehatan, BPJS Kesehatan meluncurkan aplikasi Mobile JKN sejak 2017 untuk mempermudah akses layanan kesehatan secara online seperti pendaftaran antrean, konsultasi dokter, dan informasi fasilitas kesehatan. Namun, aplikasi ini menghadapi berbagai kendala teknis yang menyebabkan ketidakpuasan pengguna yang tercermin dari banyaknya ulasan negatif di Google Play Store. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna Mobile JKN secara otomatis menggunakan metode Stochastic Gradient Descent (SGD). Sebanyak 100.000 data ulasan diunduh dan diproses melalui tahapan pelabelan sentimen, preprocessing teks, dan ekstraksi fitur TF-IDF. Evaluasi model menggunakan K-Fold Cross Validation menghasilkan akurasi tertinggi 91.00% dengan hyperparameter optimal. Hasil ini menunjukkan bahwa metode SGD efektif dan efisien dalam mengolah data ulasan berukuran besar sehingga dapat memberikan masukan berharga bagi pengembang aplikasi Mobile JKN untuk meningkatkan kualitas layanan dan kepuasan pengguna.

Kata Kunci: Analisis Sentimen, Stochastic Gradient Descent (SGD), Mobile JKN, Wordcloud, Natural Language Processing (NLP).

Abstract

Health is a crucial aspect that every individual must have to carry out daily activities and support community well-being. To improve healthcare services, BPJS Kesehatan launched the Mobile JKN application in 2017 to facilitate online access to healthcare services such as queue registration, doctor consultations, and health facility information. However, the application faces several technical issues causing user dissatisfaction, reflected by numerous negative reviews on the Google Play Store. This study aims to automatically analyze user sentiment towards the Mobile JKN app using the Stochastic Gradient Descent (SGD) method. A total of 100,000 review data were collected and processed through sentiment labeling, text preprocessing, and TF-IDF feature extraction. Model evaluation using K-Fold Cross Validation achieved a highest accuracy of 91.00% with optimized hyperparameters. These results demonstrate that the SGD method is effective and efficient in handling large-scale review data, providing valuable insights for the Mobile JKN development team to enhance service quality and user satisfaction.

Keywords: Sentiment Analysis, Stochastic Gradient Descent (SGD), Mobile JKN, Wordcloud, Natural Language Processing (NLP).

1. PENDAHULUAN

Kesehatan adalah salah satu aspek yang penting dan harus dimiliki oleh setiap individu dalam menjalankan aktivitas tertentu serta kesejahteraan dalam masyarakat[1]. Dalam meningkatkan aspek tersebut dibutuhkan layanan kesehatan yang baik dan mudah dijangkau oleh seluruh golongan masyarakat. Salah satu solusi dalam meningkatkan layanan kesehatan dan kualitas pelayanan kesehatan dengan memanfaatkan teknologi informasi. BPJS Kesehatan sebagai badan yang dibentuk oleh pemerintah Indonesia dalam menjalankan program jaminan kesehatan telah meluncurkan aplikasi Mobile JKN sejak 2017 sebagai upaya mempermudah akses pada layanan kesehatan secara online seperti pendaftaran layanan (antrean), konsultasi dokter, informasi lokasi fasilitas kesehatan, info program JKN, dan perubahan data pasien[2], [3]. Diharapkan dengan adanya aplikasi Mobile JKN dapat mengurangi hambatan dalam mendapatkan layanan kesehatan, mempercepat proses dalam tahapan administrasi, dan meningkatkan kenyamanan pengguna saat menggunakan akses dari berbagai layanan kesehatan. Akan tetapi permasalahan muncul dalam penggunaan aplikasi Mobile JKN terdapat berbagai kesalahan sistem (bug) seperti lambatnya respon aplikasi, fitur yang tidak berjalan optimal, serta kendala dalam proses login dan navigasi yang menyebabkan penurunan kinerja aplikasi dan ketidakpuasan pengguna[2], [4]. Ketidakpuasan pengguna pada aplikasi disalurkan melalui ulasan dengan menuliskan komentar yang tersedia pada halaman Google Play Store, hal ini terlihat dari banyaknya jumlah komentar negatif dan ulasan dengan bintang rendah. Ulasan yang diperoleh dari pengguna baik berupa komentar negatif maupun positif dapat dimanfaatkan untuk mengetahui persepsi pengguna terhadap aplikasi Mobile JKN sebagai alat evaluasi terhadap performa aplikasi sehingga memberikan masukan yang berharga bagi tim pengembang di BPJS Kesehatan. Untuk mendapatkan evaluasi dalam ulasan komentar terkait aplikasi Mobile JKN secara otomatis maka dibutuhkannya analisis sentimen untuk memahami opini dan pengalaman pengguna.

Analisis Sentimen adalah teknik Natural Language Processing (NLP) untuk menganalisis suatu pendapat, opini, atau perasaan terhadap topik tertentu dengan mengelompokkan ke dalam kategori positif, negatif, atau netral[5], [6]. Penelitian sebelumnya terkait analisis sentimen publik terhadap Pilgub Sumatera Selatan 2024 dari crawling data twitter dan youtube menggunakan model Decision Tree yang menghasilkan akurasi 79,22% dan model Random Forest menghasilkan akurasi 75,32%[7]. Penelitian selanjutnya yaitu Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine menghasilkan akurasi sebesar 85 % untuk metode Naive Bayes dan 81% untuk metode Support Vector Machine[8]. Selain itu penelitian yang mengenai aplikasi Mobile JKN dengan dataset yang berjumlah 790 data menghasilkan akurasi sebesar 75,95%, precision sebesar 76,00%, recall sebesar 98,28%, dan f1-score sebesar 85,71% menggunakan metode Naive Bayes Classifier dan akurasi sebesar 79.75%, precision sebesar 79.17%, recall sebesar 98.28%, dan f1-score sebesar 87.69% untuk C4.5[3]. Penelitian lain berjudul Analisis Sentimen Pada Ulasan Mobile Jkn Berdasarkan Pada Media Sosial Twitter (X) Menggunakan Metode Support Verctor Machine(SVM) memiliki distribusi label sentimen yang terdiri dari 13,24% positif, 71,40% netral, dan 15,36% negatif menghasilkan akurasi sebesar 82%, Precision 49%, Recall 74%, dan F1-Score 62%[9]. Penelitian terkait dengan menggunakan metode Stochastic Gradient Descent (SGD) dalam klasifikasi tahu berformalin diperoleh hasil evaluasi model dengan nilai akurasi sebesar 82,6%, presisi sebesar 81,7%, recall sebesar 84,1%, dan f1-score sebesar 83,5%[10] Meskipun dari penelitian telah menerapkan berbagai metode klasifikasi seperti Support Verctor Machine(SVM), Naive Bayes Classifier, Random Forest, dan C4.5 menghasilkan akurasi yang cukup baik akan tetapi masih terbatas dalam jumlah dataset yang relatif kecil sampai sedang pada analisis sentiment membuat kemampuan model dalam mengenali pola bahasa alami secara menyeluruh kurang optimal sehingga tidak dapat mempresentasikan keseluruhan variasi kata, kalimat, dan gaya bahasa.

Sampai saat ini belum terdapat studi yang secara khusus membandingkan distribusi sentimen pengguna aplikasi Mobile JKN dengan hasil klasifikasi menggunakan algoritma Stochastic Gradient Descent (SGD) serta menghubungkannya dengan kebutuhan perbaikan fitur yang terjadi kendala berdasarkan ulasan pengguna. Penelitian ini hadir untuk mengisi kekosongan tersebut dalam menerapkan pendekatan analisis sentimen berbasis algoritma SGD yang dikenal akan kecepatan, efisiensi, dan kemampuannya dalam menangani dataset yang berukuran besar. Tujuan utama dari studi ini adalah mengklasifikasikan sentimen pengguna aplikasi Mobile JKN ke dalam tiga kategori, yakni positif, netral, dan negatif dengan menggunakan metode Stochastic Gradient Descent (SGD) dan beberapa hyperparameter. Selain itu, penelitian ini juga berupaya mengidentifikasi fitur-fitur dalam aplikasi yang paling sering memicu tanggapan baik positif maupun negatif yang nantinya dapat dijadikan sebagai acuan strategis oleh BPJS Kesehatan. Diharapkan hasil dari analisis ini dapat memberikan wawasan yang lebih mendalam bagi pengembangan dan peningkatan mutu layanan dalam aplikasi Mobile JKN. Dengan memahami opini serta pengalaman pengguna, BPJS Kesehatan dapat melakukan perbaikan dan inovasi yang lebih terarah.

2. METODOLOGI PENELITIAN



Gambar 1. Metodologi Penelitian

2.1 Pengumpulan Data

Dalam pengumpulan data pada penelitian ini menggunakan teknik scrapping atau proses mengekstraksi konten berupa data ulasan dari google play store pada aplikasi Mobile JKN dengan bantuan library google-play-scraper dan IDE google colab. Data ulasan yang diambil diantaranya berupa id ulasan, nama pengguna, gambar pengguna, ulasan pengguna, rating dari pengguna, jumlah like dari ulasan pengguna, waktu ulasan dibuat, balasan ulasan dari developer, waktu balasan dari developer, dan versi aplikasi yang digunakan oleh pengguna saat memberikan ulasan.

2.2 Pelabelan Data

Pelabelan data merupakan proses memberikan suatukategori sentimen untuk mengklasifikasikan opini dari pengguna sehingga model machine learning dapat mengklasifikasikan sesuai dengan label kategori sentimen yang telah ditentukan[11]. Pelabelan data menggunakan metode rating-based yaitu dimana skor 1-2 diklasifikasikan dalam kategori sentimen negatif, skor 3 diklasifikasikan dalam kategori sentimen netral, dan skor 4-5 diklasifikasikan dalam kategori sentimen positif[12].

2.3 Preprocessing

Preprocessing adalah tahap dimana merubah data mentah (*raw data*) menjadi data yang lebih bersih dan terstruktur agar dapat diklasifikasikan secara akurat oleh model machine learning[7]. Pada tahap pertama yaitu membersihkan data ulasan dari unsur yang tidak dibutuhkan dan mengganggu hasil akurasi dalam klasifikasi analisis sentimen seperti menghapus mention, hastag, URL, angka, simbol, emoji, karakter non-ASCII, dan menormalkan karakter berulang (misalnya “errrorr” menjadi “eror”). Selanjutnya melakukan *case folding* dimana merubah semua teks ulasan menjadi huruf kecil yang bertujuan untuk mengurangi variasi kata yang ada. Hasil dari semua proses tersebut dilakukan tokenisasi untuk mengubah teks menjadi unit-unit yang lebih kecil atau disebut token. Tahap selanjutnya hasil dari tokenisasi, menghapus kata-kata yang termasuk ke dalam list *stopword* yang tidak memiliki pengaruh dalam analisis sentimen seperti “yang”, “dengan”, “di”. Proses terakhir pada preprocessing yaitu stemming yang merubah kata ke dalam bentuk kata dasar dengan menghapus semua kata imbuhan.

2.4 Ekstraksi Fitur

Setelah tahapan preprocessing selesai maka selanjutnya dilakukan ekstraksi fitur pada teks yang diproses dengan memberikan pembobotan setiap kata untuk merubah kata menjadi bentuk numerik dengan menggunakan perhitungan TF-IDF(Term Frequency – Inverse Document Frequency)[13]. Hasil dari ekstraksi fitur tersebut digunakan dalam input dalam model klasifikasi. Langkah-langkah dalam pembobotan kata menggunakan TF-IDF :

1. Menghitung banyaknya kemunculan suatu kata atau term yang muncul dalam sebuah dokumen atau disebut dengan Term Frequency (TF) melalui persamaan (1).

$$TF(t, d) = \frac{f_{t,d}}{\sum_n f_{n,d}} \quad (1)$$

Dimana :

$f_{t,d}$ = Jumlah frekuensi term t dalam dokumen d

$\sum_n f_{n,d}$ = Jumlah total term t dalam dokumen d

2. Menghitung Inverse Document Frequency (IDF) untuk mengukur seberapa penting sebuah term atau kata dalam keseluruhan isi dokumen dan meminimalkan pengaruh dari kata-kata umum seperti “dan”, “yang”, “atau”, “lalu” agar tidak mendominasi pada hasil analisis serta memberi bobot lebih tinggi untuk kata yang jarang muncul untuk meningkatkan relevansi term tersebut dalam analisa teks.

$$IDF(t) = \log \frac{N}{nt} \quad (2)$$

Dimana:

N = Total jumlah dokumen dalam korpus

nt = Jumlah dokumen yang mengandung term t

3. Pembobotan menggunakan TF-IDF dengan cara mengalikan nilai dari TF dan IDF untuk mendapatkan bobot akhir dari suatu term.

$$TF - IDF(t, d) = TF(t, d) \times IDF(t) \quad (3)$$

2.5 Wordcloud

Wordcloud merupakan sebuah visualisasi dari data teks yang mudah dibaca dimana semakin besar ukuran kata yang terlihat maka mencerminkan semakin banyak frekuensi kemunculan dalam data teks tersebut[14]. Selain itu, wordcloud berfungsi untuk menemukan pola, tren, dan distribusi kata dalam kumpulan data teks yang besar tanpa perlu membaca semua isi teks sehingga sangat berguna dalam analisis text mining dan survei terbuka. Dalam konteks

analisis sentimen, wordcloud juga berperan penting dalam memahami penyebaran kata-kata yang mengandung emosional positif atau negatif dalam percakapan publik, meningkatkan akurasi dan efisiensi dalam mendeteksi tren sentimen di media sosial atau platform digital[15].

2.6 Klasifikasi

Setelah dilakukan pembobotan menggunakan TF-IDF, dilanjutkan dengan melakukan penerapan teknik K-Fold Cross Validation dilakukan dengan membagi dataset ke dalam k bagian yang sama dimana secara bergantian setiap bagian sebagai data pengujian (data testing) sementara sisanya sebagai data pelatihan (data training). Kemudian dilakukan tahap pembuatan model Stochastic Gradient Descent (SGD) dengan beberapa hyperparameter seperti jumlah maksimal iterasi, learning rate, loss, tol, random_state, eta0, alpha, penalty, dan early_stoppings selama pelatihan dilakukan untuk klasifikasi analisis sentimen.

Stochastic Gradient Descent (SGD) adalah algoritma optimasi iteratif yang bertujuan menemukan titik minimum dari suatu fungsi dengan melakukan penyesuaian secara bertahap menggunakan aturan gradien atau turunan fungsi yang ingin diminimalkan. SGD mampu mempercepat proses pelatihan klasifikasi karena pembaruan parameter dilakukan secara berulang dan cepat serta tidak terbatas oleh ukuran dataset latih sehingga waktu pelaksanaannya menjadi fleksibel. Proses algoritma SGD dimulai dengan menentukan nilai awal parameter θ , kemudian pada setiap iterasi nilai θ diperbarui seperti pada persamaan 4 dengan menggunakan learning rate α hingga nilai minimum dari fungsi $J(\theta)$. Pembaruan pada persamaan 5 dilakukan untuk setiap data pelatihan dimana fungsi loss L digunakan untuk mengukur kesalahan prediksi dan regularisasi R ditambahkan untuk mengontrol kompleksitas model[10].

$$\theta_j = \theta_j - \alpha \frac{\partial y}{\partial x} J(\theta) \quad (4)$$

$$J(\theta) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x)) + \alpha R(W) \quad (5)$$

2.7 Evaluasi

Dari hasil model klasifikasi yang telah dibuat dalam klasifikasi analisis sentimen dilakukan evaluasi performa model menggunakan metrik evaluasi:

a. Akurasi

Akurasi digunakan untuk mengukur jumlah prediksi benar (positif dan negatif) dari semua data yang diuji dari model yang telah dibentuk. Pada penelitian ini akurasi menunjukkan seberapa baik dalam klasifikasi analisis sentimen ke dalam kategori positif, netral, maupun negatif. Perhitungan akurasi dapat dilihat pada persamaan 6 dimana TP(*True Positive*) adalah jumlah prediksi yang benar mengenai kelas positif, TN(*True Negative*) adalah jumlah prediksi yang benar mengenai kelas negatif, FP(*False Positive*) adalah jumlah prediksi yang salah mengenai kelas positif, FN(*False Negative*) adalah jumlah prediksi yang salah mengenai kelas negatif.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

b. Presisi

Presisi adalah rasio prediksi benar dengan total hasil yang diprediksi kelas positif maka semakin tinggi nilai presisi maka semakin sedikit kesalahan dalam memprediksi kelas positif. Untuk menghitung presisi didapatkan pada persamaan 7.

$$\text{Presisi} = \frac{TP}{TP+FP} \quad (7)$$

c. Recall

Recall mengukur kemampuan model yang dihasilkan dalam memprediksi semua data pada kelas positif, dimana didefinisikan jumlah prediksi yang benar mengenai kelas positif dibagi dengan jumlah keseluruhan data kelas positif.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

d. F1-Score

F1-Score menjawab kelemahan presisi dan recall dimana perhitungan dari rata-rata dari recall dan presisi. Semakin mendekati 1 maka semakin bagus performa dari suatu model dan semakin mendekati 0 maka model yang dihasilkan semakin buruk.

$$F1 - \text{Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (9)$$

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data dan Labeling Data

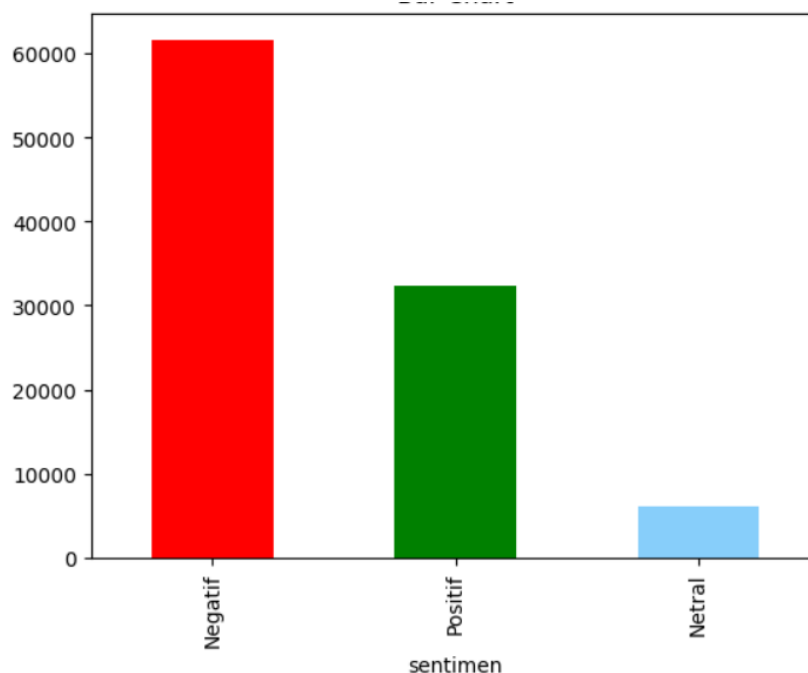
Dalam pengumpulan data menggunakan google play scrapper dengan menginstall melalui pip yang merupakan sebuah library python untuk mengambil data ulasan Mobile JKN pada google play store secara terstruktur tanpa harus menggunakan API resmi dari Google. Dengan library tersebut mengunduh 100.000 data ulasan dari aplikasi Mobile JKN berupa list of dictionary yang dikonversi menjadi dataframe melalui library pandas. Contoh data yang diambil dari ulasan Mobile JKN dapat dilihat pada gambar 2.

reviewid	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt	appVersion
73cb8d09-a8fc-4532-b1bf-8ff9cde46f2d	Pengguna Google	lh.googleusercontent.com/EGemol2N...	https://play-lama dan banyak yg mengkri...	1	1305	4.12.0	2025-06-02 13:06:24	None	NaT	4.12.0
73e3840b-adc3-418c-ac35-5538c4bba1fd	Pengguna Google	lh.googleusercontent.com/EGemol2N...	Aplikasi gak jelas, sering bug, sering error, ...	1	1181	4.12.0	2025-05-18 18:27:07	Mohon maaf atas ketidaknyamanannya. Terkait ke...	2025-05-20 14:03:19	4.12.0
71ce398f-2c14-484c-9bfe-7e26610ef400	Pengguna Google	lh.googleusercontent.com/EGemol2N...	Ini aplikasi jkn mobile nya berfungsi dengan b...	1	511	4.12.0	2025-05-21 22:36:28	Mohon maaf atas ketidaknyamanannya. Terkait ke...	2025-05-22 14:27:56	4.12.0
dd5009b4-7c0b-4f0b-af02-607c52fe983a	Pengguna Google	lh.googleusercontent.com/EGemol2N...	Ga bisa login krn pembatasan 1perangkat IMEI h...	1	67	4.12.0	2025-06-01 10:45:50	None	NaT	4.12.0
152590a9-7ca5-46d4-b32a-005a88f46c54	Pengguna Google	lh.googleusercontent.com/EGemol2N...	Ribet, ngbug terus, sms verifikasi 1 abad ha...	1	120	4.12.0	2025-05-31 06:30:44	Mohon maaf atas ketidaknyamanannya. Pastikan j...	2025-06-05 14:50:36	4.12.0

Gambar 2. Data Ulasan Mobile JKN

3.2 Pelabelan Data

Dari 100.000 data ulasan pada aplikasi Mobile JKN dilakukan pelabelan menjadi 3 kategori yaitu sentimen negatif, sentimen positif, dan sentimen netral. Dimana sentimen negatif dengan skor 1-2 berjumlah 61.575 data ulasan, sentimen netral dengan skor 3 berjumlah 6.154 data ulasan, dan sentimen positif dengan skor 4-5 berjumlah 32.271 data ulasan dapat dilihat pada Gambar 3. Distribusi data sentimen ini mempresentasikan banyak keluhan atau ketidakpuasan terhadap layanan kesehatan yang ada pada aplikasi Mobile JKN.



Gambar 3. Pelabelan Sentimen

3.3 Preprocessing

Tahap dimana proses awal pengolahan data seperti pembersihan data, transformasi data, dan pelabelan data sebelum dilakukan klasifikasi menggunakan Metode Stochastic Gradient Descent (SGD). Berikut ini merupakan tahapan preprocessing:

Data Cleaning

Data cleaning atau pembersihan data yaitu tahap penyaringan data dengan menghapus nilai-nilai yang salah dan duplikat untuk meningkatkan kualitas data agar model yang dihasilkan menghasilkan akurasi yang tinggi dengan cara menghapus teks yang mengandung mention (@), tagar (#), tautan atau link, angka, karakter emoji atau simbol-simbol, dan kata yang berulang-ulang dalam ulasan. Proses pembersihan data dapat dilihat pada Tabel 1.

Tabel 1. Output Data Cleaning

Data Ulasan	Setelah Data Cleaning
Tolong diperbaiki antara aplikasi yg di sini dengan aplikasi dimasing ² klinik, tidak ballance loh.. Di aplikasi sudah pindah faskes, Tp begitu datang ke kliniknya blm pindah katanya,, jadi ga ballance antara aplikasi dengan klinik... #sadangmedika	Tolong diperbaiki antara aplikasi yg di sini dengan aplikasi dimasing klinik tidak ballance loh Di aplikasi sudah pindah faskes Tp begitu datang ke kliniknya blm pindah katanya jadi ga ballance antara aplikasi dengan klinik

Case Folding

Setelah dilakukan pembersihan data maka data ulasan aplikasi Mobile JKN dilakukan tahap case folding yaitu merubah semua huruf kapital(upper case) yang ada pada ulasan menjadi huruf kecil(lower case) seperti pada Tabel 2.

Tabel 2. Output Case Folding

Case Folding
tolong diperbaiki antara aplikasi yg di sini dengan aplikasi dimasing klinik tidak ballance loh di aplikasi sudah pindah faskes tp begitu datang ke kliniknya blm pindah katanya jadi ga ballance antara aplikasi dengan klinik

Tokenisasi

Proses ini mengubah ulasan menjadi bentuk kecil atau disebut token. Token dapat berupa sebuah kata, frasa, atau karakter.

Tabel 3. Output Tokenisasi

Tokenisasi
['tolong', 'diperbaiki', 'antara', 'aplikasi', 'yg', 'di', 'sini', 'dengan', 'aplikasi', 'dimasing', 'klinik', 'tidak', 'ballance', 'loh', 'di', 'aplikasi', 'sudah', 'pindah', 'faskes', 'tp', 'begitu', 'datang', 'ke', 'kliniknya', 'blm', 'pindah', 'katanya', 'jadi', 'ga', 'ballance', 'antara', 'aplikasi', 'dengan', 'klinik']

Stopword Removal

Stopword removal merupakan proses menyeleksi kata yang mengganggu proses analisis sentimen seperti menghapus kata “yg”, “di”, “loh”, “tp” maka akan menghasilkan kata yang terlihat pada Tabel 4.

Tabel 4. Output Stopword Removal

Stopword Removal
['tolong', 'diperbaiki', 'aplikasi', 'aplikasi', 'dimasing', 'klinik', 'ballance', 'aplikasi', 'pindah', 'faskes', 'kliniknya', 'blm', 'pindah', 'ga', 'ballance', 'aplikasi', 'klinik']

Stemming

Stemming adalah proses mengubah kata menjadi bentuk kata dasar dan menghilangkan kata imbuhan menggunakan library Sastrawi seperti pada Tabel 5.

Tabel 5. Output Stemming

Stopword Removal dan Stemming	
['tolong', 'perbaiki', 'aplikasi', 'aplikasi', 'masing', 'klinik', 'ballance', 'aplikasi', 'pindah', 'faskes', 'klinik', 'blm', 'pindah', 'ga', 'ballance', 'aplikasi', 'klinik']	

3.4 Ekstraksi Fitur

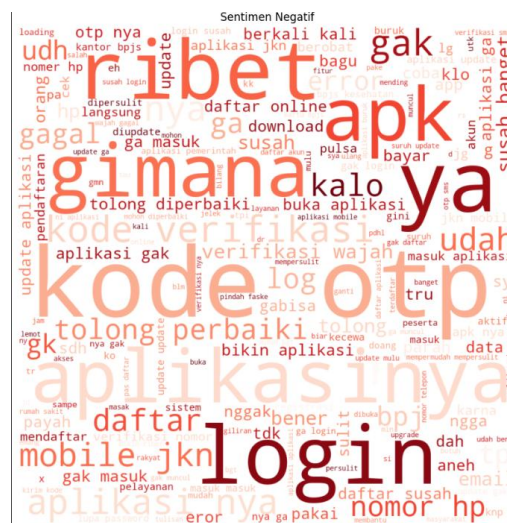
Proses ekstraksi fitur dengan menghitung bobot dari kumpulan dokumen teks menggunakan library scikit-learn. Pertama mengambil list kata kemudian menghitung kemunculan kata pada kumpulan dokumen teks atau Term Frequency (TF), menghitung banyak dokumen yang mengandung kata tersebut (DF), mengukur pentingnya suatu kata dalam keseluruhan dokumen (IDF), dan nilai bobot TF-IDF adalah hasil perkalian TF dan IDF. Untuk perhitungan TF-IDF dari list kata yang mempunyai nilai tertinggi pada ulasan Mobile JKN tercantum pada tabel 6.

Tabel 6. TF-IDF Tertinggi

No Kata	TF	DF	IDF	TF-IDF
1 Bantu	3737.956	10195	0.894703	3344.3595
2 Mudah	3341.677	11133	0.856478	2862.0723
3 Update	2739.649	9703	0.916184	2510.0222
4 Daftar	3008.099	15738	0.706140	2124.1404
5 Aplikasi	4475.677	27602	0.462149	2068.4316
6 Bagus	1690.799	4988	1.205164	2037.6896

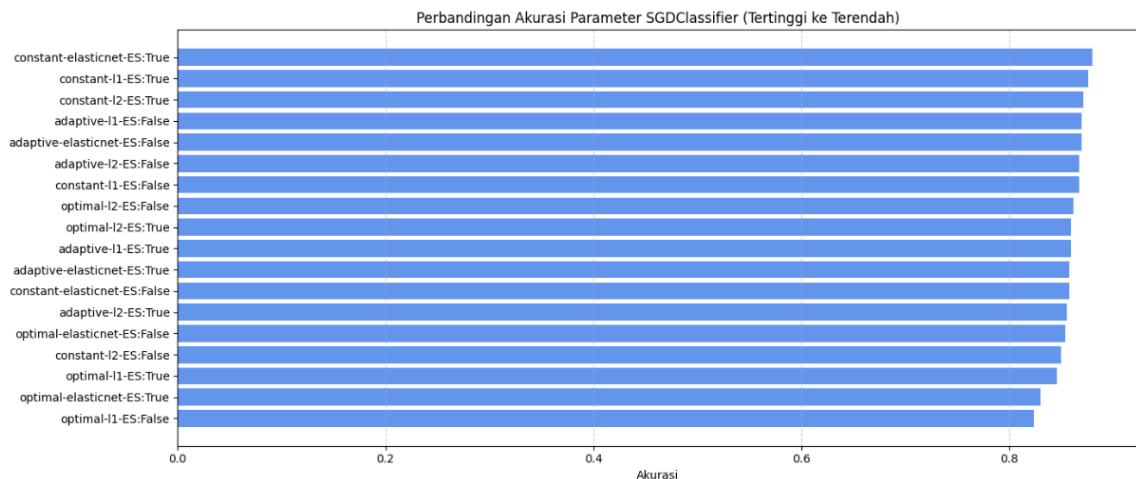
3.5 WordCloud

Visualisasi kata-kata dari ulasan aplikasi Mobile JKN menggunakan library python yaitu wordcloud dan matplotlib pyplot. Wordcloud dari 6.154 data ulasan pada google play store dengan sentimen negatif dari pengguna pada gambar 4 menunjukkan bahwa kata-kata yang paling banyak muncul seperti “kode”, “otp”, “login”, ”aplikasinya”, ”ribet” menjadi kata yang paling sering muncul, hal ini menandakan bahwa sebagian besar keluhan dari pengguna terkait masalah kode OTP yang tidak diterima oleh pengguna atau gagal diverifikasi setelah proses input kode otp, kesulitan pengguna saat proses login aplikasi Mobile JKN serta pengalaman menggunakan aplikasi yang dianggap ribet atau tidak praktis. Selain itu kata “tolong diperbaiki” menandakan adanya harapan dari pengguna agar pengembang aplikasi Mobile JKN dapat segera melakukan perbaikan bug pada fitur-fitur yang ada.



Gambar 4. Wordcloud untuk sentimen negatif

Dari kata-kata ulasan netral seperti "update," "login," dan "aplikasi," yang mengindikasikan fokus pengguna pada proses pembaruan aplikasi Mobile JKN dan fitur login. Selain itu kata seperti "susah," "gagal," dan "tolong," menandakan bahwa adanya kendala teknis dan permintaan untuk segera diperbaiki yang terlihat pada gambar 5.



Gambar 7. Wordcloud sentimen positif

3.7 Evaluasi

Evaluasi model Stochastic Gradient Descent (SGD) menggunakan teknik evaluasi K-Fold Cross Validation dengan membandingkan fold ke-1 sampai dengan fold ke-10 dengan cara setiap subset data secara bergantian sebagai data pengujian (data testing) sementara sisanya sebagai data pelatihan (data training). Hasil evaluasi menunjukkan bahwa fold ke-1 menghasilkan evaluasi model yang paling optimal dengan akurasi 91%, presisi 90.99%, recall 90.81%, dan f1-score 90.89% sedangkan evaluasi model yang kurang optimal pada fold ke-10 akurasi 85.40%, presisi 84.88%, recall 85.50%, dan f1-score 84.49%. Selain itu dari hasil metrik evaluasi dari presisi, recall, dan f1-score mempunyai selisih yang tidak terlalu signifikan dengan nilai akurasi.

Tabel 7. Perbandingan Akurasi berdasarkan K-Fold

Fold	Akurasi	Presisi	Recall	F1-Score
1	0.9100	0.9099	0.9081	0.9089
2	0.8560	0.8558	0.8562	0.8558
3	0.8840	0.8842	0.8846	0.8839
4	0.8600	0.8600	0.8603	0.8596
5	0.8640	0.8632	0.8665	0.8626
6	0.8620	0.8630	0.8640	0.8598
7	0.8601	0.8590	0.8618	0.8576
8	0.8722	0.8715	0.8741	0.8710
9	0.8642	0.8647	0.8654	0.8624
10	0.8540	0.8488	0.8550	0.8496

4. KESIMPULAN

Sebanyak 100.000 data ulasan berhasil diunduh dalam format list of dictionary dengan memanfaatkan google-play-scraper, kemudian data tersebut dikonversi menjadi dataframe menggunakan library pandas untuk memudahkan analisis. Setelah dilakukan pelabelan sentimen dari 100.000 data ulasan terdiri dari 61.575 ulasan termasuk kategori sentimen negatif (skor 1-2), 6.154 termasuk kategori sentimen netral (skor 3), dan 32.271 termasuk kategori sentimen positif (skor 4-5). Hasil analisis sentimen menggunakan metode Stochastic Gradient Descent (SGD) dari data ulasan google play store menghasilkan akurasi tertinggi menggunakan hyperparameter dari learning rate constant, penalty elasticnet, dan disertai early stopping dengan menggunakan teknik evaluasi K-Fold Cross Validation menunjukkan bahwa fold ke-1 memberikan hasil terbaik dengan akurasi 91.00%, presisi 90.99%, recall 90.81%, dan f1-score 90.89%. Sementara hasil terendah pada fold ke 10 mencapai akurasi 85.40%, presisi 84.88%, recall 85.50%, dan f1-score 84.49%. Dari hasil metrik evaluasi pada semua fold mengenai perbedaan antara metrik presisi, recall, dan f1-score dengan akurasi juga relatif kecil menunjukkan bahwa metode Stochastic Gradient Descent (SGD) adalah algoritma yang efisien dan optimal digunakan untuk memproses dataset yang berukuran besar. Selain itu, penerapan preprocessing teks dan ekstraksi

fitur menggunakan TF-IDF dengan n-grams turut meningkatkan kualitas model sehingga mampu menangani variasi bahasa dalam ulasan secara efektif dan meningkatkan performa klasifikasi sentimen secara keseluruhan.

REFERENCES

- [1] C. Kinanti and R. Lasso, “FAKTOR HAMBATAN DALAM AKSES PELAYANAN KESEHATAN PADA PUSKESMAS DI INDONESIA: SCOPING REVIEW.” [Online]. Available: <http://journal.stikeskendal.ac.id/index.php/PSKM>
- [2] M. Mauludin Rohman and S. Adinugroho, “Analisis Sentimen pada Ulasan Aplikasi Mobile JKN Menggunakan Metode Maximum Entropy dan Seleksi Fitur Gini Index Text,” 2021. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [3] C. C. Dan and L. Saufa Yardha, “ANALISIS SENTIMEN PENGGUNA APLIKASI MOBILE JKN MENGGUNAKAN ALGORITMA NAÏVE BAYES,” 2024. [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>
- [4] M. Nurmalasari, N. Anggita Temesvari, S. Ni, and matul Maula, “Analisis Sentimen terhadap Opini Masyarakat dalam Penggunaan Mobile-JKN untuk Pelayanan BPJS Kesehatan Tahun 2019,” 2020.
- [5] M. R. Fahlevvi, “Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia di Google Playstore Menggunakan Metode Support Vector Machine,” *Jurnal Teknologi dan Komunikasi Pemerintahan*, vol. 4, no. 1, pp. 1–13, 2022, [Online]. Available: <http://ejournal.ipdn.ac.id/JTKP>,
- [6] F. Fridom Mailo *et al.*, “Analisis Sentimen Data Twitter Menggunakan Metode Text Mining Tentang Masalah Obesitas di Indonesia,” 2019.
- [7] F. Panjaitan, W. Ce, H. Oktafiandi, G. Kanugrahan, Y. Ramdhani, and V. H. C. Putra, “Evaluation of Machine Learning Models for Sentiment Analysis in the South Sumatra Governor Election Using Data Balancing Techniques,” *Journal of Information Systems and Informatics*, vol. 7, no. 1, pp. 461–478, Mar. 2025, doi: 10.51519/journalisi.v7i1.1019.
- [8] Tania Puspa Rahayu Sanjaya, Ahmad Fauzi, and Anis Fitri Nur Masruriyah, “Analisis sentimen ulasan pada e-commerce shopee menggunakan algoritma naive bayes dan support vector machine,” *INFOTECH : Jurnal Informatika & Teknologi*, vol. 4, no. 1, pp. 16–26, Jun. 2023, doi: 10.37373/infotech.v4i1.422.
- [9] D. Puspitasari, N. Nazhiifah, and T. Sutabri, “Analisis Sentimen Pada Ulasan Mobile Jkn Berdasarkan Pada Media Sosial Twitter (X) Menggunakan Metode Support Verctor Machine (SVM),” *Nusantara Journal of Multidisciplinary Science*, vol. 2, no. 6, pp. 1273–1282.
- [10] F. T. Admojo and Y. I. Sulistya, “Analisis performa algoritma Stochastic Gradient Descent (SGD) dalam mengklasifikasi tahu berformalin,” *Indonesian Journal of Data and Science (IJODAS)*, vol. 3, no. 1, 2022.
- [11] A. A. Pratiwi and M. Kamayani, “Perbandingan Pelabelan Data dalam Analisis Sentimen Kurikulum Proyek di platform TikTok: Pendekatan Naïve Bayes,” *Jurnal Eksplora Informatika*, vol. 14, no. 1, pp. 96–107, Sep. 2024, doi: 10.30864/eksplora.v14i1.1093.
- [12] M. Demircan, A. Seller, F. Abut, and M. F. Akay, “Developing Turkish sentiment analysis models using machine learning and e-commerce data,” *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 202–207, Jun. 2021, doi: 10.1016/j.ijcce.2021.11.003.
- [13] H. Chyntia Morama, D. E. Ratnawati, and I. Arwani, “Analisis Sentimen berbasis Aspek terhadap Ulasan Hotel Tentrem Yogyakarta menggunakan Algoritma Random Forest Classifier,” 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [14] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, “VISUALISASI WORD CLOUD HASIL ANALISIS SENTIMEN BERBASIS FITUR LAYANAN APLIKASI GOJEK DENGAN SUPPORT VECTOR MACHINE,” *Jurnal Serina Sains, Teknik dan Kedokteran*, vol. 2, no. 1, pp. 61–70, Mar. 2024, doi: 10.24912/jsstk.v2i1.32058.

- [15] M. Galih Pradana, “PENGUNAAN FITUR WORDCLOUD DAN DOCUMENT TERM MATRIX DALAM TEXT MINING.”

- [16] X. Ying, “An Overview of Overfitting and its Solutions,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Mar. 2019. doi: 10.1088/1742-6596/1168/2/022022.