

ANALISIS SENTIMEN ULASAN APLIKASI SATU SEHAT PADA GOOGLE PLAY STORE MENGGUNAKAN NAÏVE BAYES CLASSIFIER

Muhammad Tirta Nugraha, Nina Sulistiyowati, Ultach Enri

Informatika, Universitas Singaperbangsa Karawang
Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang,
Tirtanugraha269@gmail.com

ABSTRAK

Pada tanggal 30 Desember 2022, pemerintah Indonesia telah mengumumkan pencabutan kebijakan PPKM di semua wilayah. Berselang beberapa bulan setelah pencabutan PPKM, aplikasi Peduli Lindungi resmi berganti nama menjadi satu sehat dan dilakukan proses migrasi secara bertahap. Namun dalam awal migrasinya aplikasi satu sehat, banyak pengguna aplikasi mengalami masalah, sehingga menimbulkan ulasan negatif di kolom ulasan aplikasi satu sehat pada Google Play Store. Tujuan dari penelitian ini adalah untuk mengklasifikasikan dan mengetahui sentimen pengguna aplikasi Satu Sehat menggunakan algoritma *Naïve Bayes Classifier*, dengan mengetahui performa algoritma *Naïve Bayes Classifier*. Penelitian ini, menggunakan Metode KDD (*Knowledge Discovery in Database*) yang terdiri dari tahapan, *data Scraping*, *Preprocessing*, *data Transformation*, *data mining* dan *pattern evaluation*. Hasil penelitian menunjukkan analisis sentimen ulasan aplikasi Satu Sehat dari total dataset yang berjumlah 16.245 data terdapat 12.551 data sentimen negatif dan 3.694 data sentimen positif berdasarkan *rating* bintang pengguna pada bulan Maret 2023, dapat disimpulkan bahwa hasil sentimen menunjukkan lebih banyak mengandung sentimen negatif. Hal ini terlihat bahwa banyak pengguna yang mengeluhkan tentang aplikasi yang ribet dalam penggunaannya serta masih banyak *bug* dan *error*. Hasil evaluasi performa dari algoritma *Naïve Bayes Classifier* dalam mengklasifikasi sentimen pengguna aplikasi Satu Sehat mendapatkan hasil terbaik dengan skenario pembagian data *training* sebesar 90% dan data *testing* 10% dengan nilai *accuracy* sebesar 93,00%, *precision* sebesar 93,00%, *recall* sebesar 93,00% dan *f1-score* sebesar 92,00%.

Kata kunci: Analisis Sentimen, Satu Sehat, KDD (*Knowledge Discovery in Database*), *Naive Bayes Classifier*

1. PENDAHULUAN

Adanya virus Covid-19 yang menyebar hampir di seluruh dunia pada tahun 2019 yang lalu, oleh karena itu PPKM diberlakukan di Indonesia yang membatasi aktivitas semua orang, aplikasi Peduli Lindungi menjadi salah satu aplikasi wajib bagi masyarakat untuk memeriksa sertifikat vaksin Covid-19 sebelum mengunjungi tempat umum. Pada tanggal 30 Desember 2022, pada tingkat nasional, pemerintah Indonesia telah mengumumkan pencabutan kebijakan PPKM di semua wilayah. Berselang beberapa bulan setelah pencabutan PPKM, aplikasi Peduli Lindungi resmi berganti nama menjadi satu sehat dan dilakukan proses migrasi secara bertahap mulai 1 Maret 2023 [1]

Dalam tahap migrasi aplikasi tersebut, fitur-fitur baru akan ditambahkan, termasuk fitur resume medis yang akan diarsipkan dalam satu tempat sehingga memungkinkan pemantauan riwayat kesehatan secara detail dan terstruktur dengan baik, meskipun perawatan dilakukan di lokasi yang berbeda. Pada ulasan yang diberikan oleh pengguna aplikasi satu sehat sering kali berupa saran positif dan keluhan negatif. Namun, menyortir dan memantau ulasan tersebut bisa menjadi pekerjaan yang sangat sulit karena banyaknya jumlah ulasan yang diposting di media sosial dan untuk melakukan secara manual akan memakan waktu yang cukup lama. Maka dari itu diperlukan suatu metode yang dapat memisahkan dan memantau ulasan tersebut dengan cepat dan otomatis, dengan tujuan mengkategorikan ulasan-ulasan tersebut menjadi positif atau negatif [1]

Pengguna aplikasi di Google Play Store dapat memberikan opini atau sentimen baik negatif, positif, maupun netral pada ulasan mereka. Namun, karena jumlah data yang sangat besar, maka dibutuhkan metode analisis sentimen untuk membantu proses pemahaman dan perolehan informasi dari data tersebut. Salah satu cara untuk memproses data dari opini pengunjung pada ulasan Google Play Store adalah melalui proses *Knowledge Discovery in Database* (KDD)[2].

Proses peringkasan informasi dari data opini pengguna aplikasi Satu Sehat bertujuan untuk menghasilkan pola informasi yang dapat dijadikan acuan dalam meningkatkan kualitas aplikasi. Pola informasi tersebut berbasis pengetahuan dan dapat membantu pengembang aplikasi untuk mengetahui kebutuhan pengguna serta potensi perbaikan yang dapat dilakukan KDD, sebagai salah satu metode teknis, termasuk dalam berbagai tahapan khusus yang memiliki sifat teknis yang dapat memberikan manfaat dalam penelitian ini. KDD dianggap sebagai suatu proses model yang lebih menyeluruh dan akurat [3]

Dalam rangka untuk mendapatkan informasi yang lebih menyeluruh dan akurat, diperlukan suatu algoritma dalam proses data mining, yaitu algoritma *Naïve Bayes Classifier*. Dalam penelitian ini, algoritma *Naïve Bayes Classifier* akan diterapkan untuk melakukan analisis sentimen pengguna aplikasi Satu Sehat di Google Play Store. Karena algoritma ini akan mencari nilai probabilitas tertinggi untuk mengklasifikasikan data uji ke dalam kategori yang

paling tepat. Oleh karena itu, algoritma ini banyak digunakan dan cocok untuk penelitian ini. Berdasarkan beberapa referensi penelitian, algoritma klasifikasi *Naïve Bayes* lebih cenderung dipilih karena kemudahan dan kecepatannya dalam proses pengolahan data [3].

Dalam penelitian ini, algoritma *Naïve Bayes Classifier* akan diterapkan untuk melakukan analisis sentimen pengguna aplikasi Satu Sehat di Google Play Store. Karena algoritma ini akan mencari nilai probabilitas tertinggi untuk mengklasifikasikan data uji ke dalam kategori yang paling tepat. Oleh karena itu, algoritma ini banyak digunakan dan cocok untuk penelitian ini. Berdasarkan beberapa referensi penelitian, algoritma klasifikasi *Naïve Bayes* lebih cenderung dipilih karena kemudahan dan kecepatannya dalam proses pengolahan data [3].

Berdasarkan dari pemaparan diatas, maka penelitian yang dilakukan adalah melakukan analisa terhadap ulasan pengguna aplikasi Satu Sehat dengan menggunakan algoritma *Naïve Bayes Classifier*. Penelitian ini menggunakan model *Multinomial Naive*. Penelitian dilakukan dengan mengambil data ulasan atau review dari Google Play Store. Metode yang digunakan adalah metode *Knowledge Discovery in Database (KDD)* yang merupakan metodologi standar dalam data mining dengan tools google colab dengan menggunakan bahasa pemrograman Python. Penelitian yang dilakukan berjudul "Analisis Sentimen Ulasan Aplikasi Satu Sehat Pada Google Play Store Menggunakan Algoritma *Naïve Bayes Classifier*"[4]

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining proses menggunakan metode dari bidang-bidang seperti statistik, matematika, kecerdasan buatan, dan machine learning untuk meringkaskan dan mengenali informasi yang bermanfaat dari berbagai database besar. Data mining, juga dikenal sebagai *pattern recognition*, adalah sebuah teknik yang digunakan untuk memproses data dengan tujuan mengidentifikasi pola-pola yang tersembunyi dari data yang sedang diproses. Data mining adalah sebuah metode pengolahan data yang memiliki skala besar. [5]

Teknik data mining digunakan untuk melakukan analisis dan ekstraksi pengetahuan secara otomatis dari data. Seringkali istilah "data mining" diartikan sebagai "*knowledge discovery*", yaitu ekstraksi informasi mengacu pada langkah- langkah yang digunakan untuk mengekstrak informasi yang relevan dan berguna dari kumpulan data besar. Beberapa langkah terlibat dalam proses ini untuk mengubah data yang tidak diproses menjadi pola yang bermakna, sehingga pengguna dapat memperoleh pengetahuan yang bermanfaat dari data tersebut [5]

2.1.1. Algoritma Naïve Bayes

Algoritma sederhana pengelompokan probabilistik bernama *Naïve Bayes* menghitung

berbagai probabilitas dengan cara menghitung berapa kali muncul dan dikombinasikan nilainya pada sebuah kumpulan data yang diberikan. *Naïve Bayes* dihitung dengan mengikuti langkah-langkah berikut:

1. Perhitungan menemukan rata-rata setiap kelas dan menggunakan angka itu sebagai nilai prioritas untuk kelas itu.
2. Nilai kemungkinan ditentukan untuk setiap kategori menggunakan rumus.
3. Menggunakan rumus untuk menentukan nilai posterior untuk setiap kategori. [5]

2.1.2. Text Preprocessing

Pada tahap ini data akan dibersihkan dari kelainan dan informasi yang tidak penting selama tahap pra-processing. Pembersihan data ini terjadi agar pengolahan data dapat berlangsung dan data tersebut tersedia untuk diproses lebih lanjut dengan bantuan suatu algoritma. *Text Preprocessing* secara umum terdiri dari berbagai tahapan, antara lain: [6]

1. Case folding

Case folding adalah proses mengubah kata atau huruf menjadi huruf kecil semua atau huruf besar semua. Dalam penerapan *case folding* untuk penelitian biasanya, semua teks diubah menjadi huruf kecil. Misalnya, 'Aplikasi' dan 'aplikasi' akan diperlakukan sebagai dua kata terpisah dalam kalimat apa pun; oleh karena itu, harus membuat semua kata dalam teks menjadi huruf kecil untuk menghindari duplikasi.

2. Tokenizing

Proses *tokenizing* merupakan proses membagi paragraf menjadi kalimat- kalimat komponennya dan membagi kalimat menjadi kata-kata komponennya. *Tokenizing* adalah tahapan memisahkan atau memisahkan setiap kata dalam kalimat agar dapat dibaca secara berurutan, setiap kata akan dipisah atau dilakukan sparasi diantara kata-kata.

3. Filtering

Dalam proses melakukan *text mining*, filtering dilakukan untuk mengecualikan kata-kata yang kurang bernilai atau tidak berkontribusi pada makna teks. Dalam penelitian biasanya akan menggunakan pendekatan *stopwords removal*, dimana kata-kata yang tidak penting akan dihilangkan berdasarkan data yang telah diolah dengan metode tersebut. *Library NLTK* adalah *Library* tipikal untuk menghapus stopwords. Contoh *stopword* termasuk "dan", "di", "yang", "akan" dan sebagainya.

4. Stemming

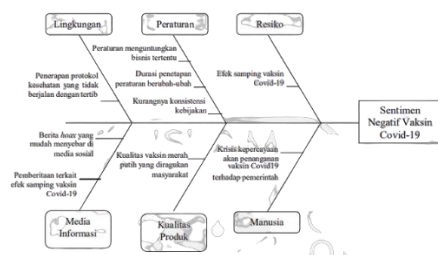
Stemming adalah proses yang digunakan untuk mengubah kata berimbuhan menjadi kata sederhana atau menghilangkan afiks. *Library NLTK* digunakan untuk mengubah kata. Dengan menggunakan *Library* sastrawi yang menjadi acuan ketentuan bahasa Indonesia, maka hasil stemming akan sesuai dengan ketentuan- ketentuan *Library* sastrawi

2.2. Analisis Sentimen

Analisis sentimen merupakan proses menganalisis dan mengolah data tekstual terkait opini produk, aplikasi, kasus, layanan, individu ataupun organisasi untuk menghasilkan informasi secara global dalam sekumpulan data. Keadaan afektif (perasaan penulis saat menulis) dan ekspresi emosional (ekspresi penulis yang ingin dia komunikasikan kepada pembaca) adalah contoh sikap [7].

2.2.1. Diagram Fishbone

Diagram *fishbone*, juga dikenal sebagai diagram *Cause and Effect* adalah diagram seperti tulang ikan yang menampilkan sebab dan akibat dari suatu situasi permasalahan. Diagram *fishbone* adalah salah satu dari tujuh alat dari *basic quality tools*. Diagram *fishbone* digunakan untuk menemukan alasan potensial dari suatu masalah, terutama ketika sebuah tim memiliki kecenderungan pemikiran rutin. Bagian tulang dari diagram *fishbone* atau diagram tulang ikan dan masalah yang sumbernya ingin ditentukan ditempatkan di kepala ikan. Setiap elemen dalam tulang memiliki penyebab mendasarnya sendiri, dan diagram *fishbone* memudahkan untuk mengidentifikasi faktor-faktor yang menjadi dasar masalah [8].



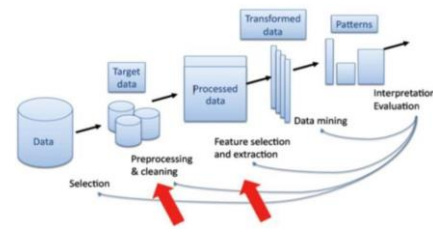
Gambar 1. Contoh diagram *fishbone*

Berikut ini adalah tahapan-tahapan yang terlibat dalam pembuatan diagram *fishbone*:

1. Di sisi kanan, tuliskan masalah utama (kepala ikan). Gambar garis dari kiri ke kanan dengan panah di ujung yang menunjuk ke masalah.
2. Tentukan semua kelas faktor utama yang menjadi penyebab pada masalah ini.
3. Gunakan panah yang lebih kecil untuk menguraikan alasan akar masalah yang mendasarinya sehingga menjadi lebih detail.

2.3. Knowledge Discovery in Database

Knowledge Discovery in Databases (KDD) adalah metode untuk menemukan dan menjelaskan pola yang bermakna dalam sejumlah besar data. Menerjemahkan output *data mining* secara akurat ke dalam bentuk yang dapat dimengerti manusia adalah inti dari prosedur KDD. [9].



Gambar 2. Proses KDD (*Knowledge Discovery in Database*)

Ada berbagai langkah yang terlibat dalam proses KDD, termasuk pemilihan data (*Selection*), pra-pemrosesan (*Preprocessing*), transformasi data (*Transformation*), data mining, dan evaluasi atau interpretasi *Interpretation/Evaluation*):

1. Pengumpulan data

Data harus dipilih dari sekumpulan data operasional sebelum langkah pengolahan informasi *Knowledge Discovery in Database (KDD)* dapat dilakukan. Data yang akan digunakan untuk prosedur proses data mining akan disimpan dalam sebuah file atau berkas yang akan disimpan terpisah dari database yang akan digunakan untuk keperluan operasional

2. Data Scraping

Pada tahap data *Scraping* dilakukan penyeleksi data set dan sampel. Ini adalah langkah sebelum tahap *Knowledge Discovery in Database (KDD)* ekstraksi informasi. File berisi informasi yang telah dipilih untuk digunakan dalam proses mining data

3. Preprocessing

Knowledge Discovery in Database (KDD) Persiapan mencakup beberapa operasi seperti pembersihan data, memeriksa inkonsistensi, dan memperbaiki kesalahan ketik. Prosedur pengayaan, yang menambahkan informasi baru ke dataset yang ada, juga merupakan bagian dari tahap *Preprocessing* dan diperlukan untuk KDD tetapi sering berasal dari sumber luar.

4. Data Transformation

Tujuan dari fase transformasi adalah untuk mengidentifikasi aspek- aspek yang relevan dari data yang diperlukan untuk tugas yang ada. Metode ini juga membantu menghilangkan faktor dan detail yang tidak perlu.

5. Data Mining

Menemukan pola atau informasi yang berguna dalam data yang telah diproses dengan cara tertentu adalah fokus pada saat ini. Data mining dapat memilih dari sejumlah metodologi dan proses yang telah terbukti. Tujuan dan prosedur *Knowledge Discovery in Database* menginformasikan pilihan teknik atau pendekatan yang dapat diterima.

6. Interpretation/Evaluation

Pola informasi yang ditemukan oleh data mining harus ditafsirkan dengan cara yang dapat dengan mudah dipahami oleh pemangku kepentingan pada tahap interpretasi. Fase ini memerlukan

penilaian validitas dan pentingnya pola atau informasi yang ditemukan, serta memverifikasi untuk melihat apakah pola atau informasi tersebut bertentangan dengan fakta atau hipotesis yang ditetapkan sebelumnya, dan kemudian menggambarkan kesimpulan analisis.

2.4. Confusion matrix

Confusion matrix adalah tabel data yang digunakan untuk melacak dan mengevaluasi kemandirian pengklasifikasi, model, dan algoritma. Ada berbagai konsep menyeluruh dalam matriks kebingungan, termasuk:

1. Istilah "*true positives*" (TP) mengacu pada tingkat di mana klasifikasi positif diidentifikasi secara akurat seperti itu.
2. Istilah "*true negatives*" (TN) mengacu pada tingkat di mana kelas negatif diidentifikasi secara akurat seperti itu.
3. Istilah "*false positive*" (FP) mengacu pada kesalahan klasifikasi kelompok negatif sebagai kelompok positif.
4. Contoh kelas positif yang disalahartikan sebagai kelas negatif adalah *false negative* (FN).[10]

Metode pengujian confusion matrix dapat menghasilkan perhitungan dengan 3 output, di antaranya yaitu;

Keakuratan sistem prediksi diukur sebagai fraksi dari semua perkiraan yang menjadi kenyataan. Adapun rumus untuk menghitung akurasi terdapat pada persamaan dibawah.

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (1)$$

Akurasi diukur dengan persentase positif sejati yang diharapkan, atau presisi. Adapun rumus untuk menghitung presisi terdapat pada persamaan dibawah.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Dalam statistik, mengingat mengukur berapa banyak positif sejati yang benar-benar diambil Rumus untuk menghitung recall terdapat pada persamaan dibawah.

Keterangan:

TP : True Positif

TN : True Negatif

FP : False Positif

FN : False Negatif

2.5. Penelitian Terdahulu

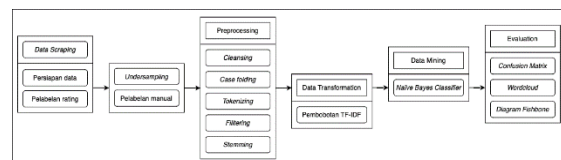
Penelitian ini merupakan dari penelitian yang sebelumnya yang dibuat oleh orang lain yang berkaitan dengan penelitian ini. Berikut ini adalah contoh penelitian terdahulu yang menjadi acuan dan bahan referensi dengan judul metode klasifikasi menggunakan algoritma naive bayes dengan metodologi penelitian KDD. Penelitian ini menjadikan twitter sebagai objek data yang berasal dari tweet pengguna. Dengan menggunakan confusion matrix, nilai accuracy tertinggi terdapat pada model 90:10.

Penelitian yang akan dilakukan saat ini analisis sentimen ulasan pengguna aplikasi satu sehat dengan

menggunakan algoritma naive bayes classifier dengan studi kasus ulasan pengguna. Tujuan dalam melakukan penelitian ini untuk mengklasifikasikan ulasan pengguna aplikasi satu sehat ke dalam kelas sentimen positif dan sentimen negatif menggunakan algoritma naive bayes classifier. Perhitungan akurasi diukur dengan confusion matrix.

3. METODE PENELITIAN

Metodologi yang diterapkan dalam penelitian ini adalah *Knowledge Discovery in Database* (KDD), yang terdiri dari beberapa tahapan seperti *Data Scraping* (proses seleksi data), *Preprocessing/Cleaning* (proses pra pengolahan data), *Transformation Data* (proses transformasi data), *Data Mining* (proses penambangan data), dan *Interpretation/Evaluation* (proses evaluasi pola). Pada tahapan data mining akan menggunakan algoritma *Naïve Bayes Classifier* dan *confusion matrix* untuk tahap evaluasi.



Gambar 3. Alur penelitian

3.1. Data Scraping

Pada tahapan ini dilakukan persiapan data dan labeling data. Data yang telah diambil menggunakan teknik web *scrapping* dengan menggunakan pemrograman python. Data sampel yang diambil merupakan ulasan pengguna aplikasi Satu Sehat pada Google Play Store, data ulasan tersebut diambil pada bulan maret 2023. Kemudian data ulasan yang sudah terkumpul maka akan dilakukan labeling data secara rating dengan 2 label yaitu positif dan negatif.

1. Undersampling Dataset

Jika terjadi data yang tidak seimbang setelah melakukan labeling secara rating maka akan dilakukan metode *Undersampling Dataset* agar data positif dan negatif menjadi seimbang, kemudian setelah data dilakukan *Undersampling* maka akan diberi label secara manual dan divalidasi oleh ahli Bahasa.

2. Text Preprocessing

Setelah tahapan text selection selesai maka akan dilakukan tahapan *Preprocessing*. Pada tahapan *Preprocessing*, dataset akan dilakukan data cleansing terlebih dahulu menggunakan metode *Preprocessing* sebelum prosedur pemodelan dilakukan. Hal ini dilakukan agar melakukan seleksi data dan mengubah data menjadi data yang terstruktur

3. Transformation

Pada tahap ini akan dilakukan proses *Transformation* melibatkan seleksi fitur dengan menggunakan metode TF-IDF (*Term Frequency Inverse Document Frequency*). Teknik seleksi fitur

dengan menggunakan TF-IDF dilakukan dengan memberikan bobot pada kemunculan kata dalam sebuah dokumen. Tujuan dari tahap ini adalah untuk mengurangi jumlah variabel dan data yang tidak diperlukan.

4. Data Mining

Pada tahap ini, Setelah dataset dibersihkan, selanjutnya diproses menggunakan berbagai metode data mining. Sebelum memulai proses pemodelan, dataset dibagi menjadi dua bagian: data training dan data testing. Percentage split data yang akan digunakan untuk penelitian kali ini adalah 70:30, 80:20 dan 90:10. Proses data mining akan menggunakan tools Google Colab dengan menggunakan bahasa pemrograman Python. Algoritma yang digunakan pada penelitian ini adalah algoritma *Naïve Bayes Classifier* model klasifikasi Multinomial *Naïve Bayes*.

5. Pattern evaluation

Pada tahapan *pattern evaluation*, evaluasi dilakukan untuk mengetahui hasil kinerja dari model yang dibangun yang menghasilkan nilai *Confusion matrix*. Metrik kinerja yang berlaku adalah *accuracy*, *precision*, *recall*, dan *f1-score* pada setiap *percentage split* data agar bisa digunakan sebagai parameter performance. Setelah mengetahui hasil sentimen positif dan negatif maka, akan dilakukan visualisasi dengan *wordcloud* terhadap kata dari ulasan sentimen yang sering muncul dan langkah selanjutnya adalah membuat diagram *fishbone* untuk menganalisis faktor-faktor masalah bagi para pengguna aplikasi dan mengidentifikasi hal-hal yang perlu dalam rangka menanggulangi permasalahan tersebut.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Penelitian ini membahas hasil analisis ulasan aplikasi Satu Sehat yang ditulis di Google Play Store. Data ulasan yang digunakan yaitu pada ulasan periode bulan Maret 2023. Hasil dari penelitian ini adalah pengembangan model klasifikasi yang dapat digunakan untuk menganalisis ulasan aplikasi. Penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD) dan algoritma *Naïve Bayes Classifier* dengan klasifikasi *Multinomial Naïve Bayes*. Implementasi model *Naïve Bayes Classifier* dilakukan menggunakan *Google Colab* dengan menggunakan bahasa pemrograman *Python*. Seluruh data yang telah diperoleh akan melalui proses KDD dengan enam tahapan, yaitu pengumpulan data, *data Scraping*, *data Preprocessing*, *data Transformation*, *data mining* dan *pattern evaluation*.

4.2. Data Scraping

Pengambilan data ulasan aplikasi Satu Sehat dilakukan dengan cara web *Scraping* dengan menggunakan *google colab* menggunakan *Python*. Teknik *Scraping* dilakukan dengan menggunakan

library *google play scraper* dengan mengambil data dari ulasan aplikasi Satu Sehat.

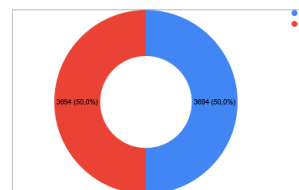
username	score	at	content	status
1. user1234	5	2023-03-01 12:34	Aplikasi ini sangat membantu...	positif
2. user5678	3	2023-03-01 13:45	Antarmuka pengguna kurang intuitif...	negatif
3. user9012	4	2023-03-01 14:56	Fitur yang ditawarkan cukup baik...	netral
4. user3456	5	2023-03-01 15:67	Rekomendasi aplikasi ini sangat tepat...	positif
5. user7890	2	2023-03-01 16:78	Ada beberapa bug yang mengganggu...	negatif
6. user1111	4	2023-03-01 17:89	Desain visualnya menarik dan modern...	positif
7. user2222	3	2023-03-01 18:90	Proses pendaftaran terlalu rumit...	negatif
8. user3333	5	2023-03-01 19:01	Fitur yang ditawarkan sangat berguna...	positif
9. user4444	4	2023-03-01 20:12	Stabilitas aplikasi cukup baik...	netral
10. user5555	3	2023-03-01 21:23	Ada beberapa fitur yang kurang...	negatif
11. user6666	5	2023-03-01 22:34	Rekomendasi aplikasi ini sangat tepat...	positif
12. user7777	4	2023-03-01 23:45	Desain visualnya menarik dan modern...	positif
13. user8888	3	2023-03-01 24:56	Proses pendaftaran terlalu rumit...	negatif
14. user9999	5	2023-03-01 25:67	Fitur yang ditawarkan sangat berguna...	positif
15. user0000	4	2023-03-01 26:78	Stabilitas aplikasi cukup baik...	netral
16. user1111	3	2023-03-01 27:89	Ada beberapa fitur yang kurang...	negatif
17. user2222	5	2023-03-01 28:90	Rekomendasi aplikasi ini sangat tepat...	positif
18. user3333	4	2023-03-01 29:01	Desain visualnya menarik dan modern...	positif
19. user4444	3	2023-03-01 30:12	Proses pendaftaran terlalu rumit...	negatif
20. user5555	5	2023-03-01 31:23	Fitur yang ditawarkan sangat berguna...	positif
21. user6666	4	2023-03-01 32:34	Stabilitas aplikasi cukup baik...	netral
22. user7777	3	2023-03-01 33:45	Ada beberapa fitur yang kurang...	negatif
23. user8888	5	2023-03-01 34:56	Rekomendasi aplikasi ini sangat tepat...	positif
24. user9999	4	2023-03-01 35:67	Desain visualnya menarik dan modern...	positif
25. user0000	3	2023-03-01 36:78	Proses pendaftaran terlalu rumit...	negatif
26. user1111	5	2023-03-01 37:89	Fitur yang ditawarkan sangat berguna...	positif
27. user2222	4	2023-03-01 38:90	Stabilitas aplikasi cukup baik...	netral
28. user3333	3	2023-03-01 39:01	Ada beberapa fitur yang kurang...	negatif
29. user4444	5	2023-03-01 40:12	Rekomendasi aplikasi ini sangat tepat...	positif
30. user5555	4	2023-03-01 41:23	Desain visualnya menarik dan modern...	positif
31. user6666	3	2023-03-01 42:34	Proses pendaftaran terlalu rumit...	negatif
32. user7777	5	2023-03-01 43:45	Fitur yang ditawarkan sangat berguna...	positif
33. user8888	4	2023-03-01 44:56	Stabilitas aplikasi cukup baik...	netral
34. user9999	3	2023-03-01 45:67	Ada beberapa fitur yang kurang...	negatif
35. user0000	5	2023-03-01 46:78	Rekomendasi aplikasi ini sangat tepat...	positif
36. user1111	4	2023-03-01 47:89	Desain visualnya menarik dan modern...	positif
37. user2222	3	2023-03-01 48:90	Proses pendaftaran terlalu rumit...	negatif
38. user3333	5	2023-03-01 49:01	Fitur yang ditawarkan sangat berguna...	positif
39. user4444	4	2023-03-01 50:12	Stabilitas aplikasi cukup baik...	netral
40. user5555	3	2023-03-01 51:23	Ada beberapa fitur yang kurang...	negatif
41. user6666	5	2023-03-01 52:34	Rekomendasi aplikasi ini sangat tepat...	positif
42. user7777	4	2023-03-01 53:45	Desain visualnya menarik dan modern...	positif
43. user8888	3	2023-03-01 54:56	Proses pendaftaran terlalu rumit...	negatif
44. user9999	5	2023-03-01 55:67	Fitur yang ditawarkan sangat berguna...	positif
45. user0000	4	2023-03-01 56:78	Stabilitas aplikasi cukup baik...	netral
46. user1111	3	2023-03-01 57:89	Ada beberapa fitur yang kurang...	negatif
47. user2222	5	2023-03-01 58:90	Rekomendasi aplikasi ini sangat tepat...	positif
48. user3333	4	2023-03-01 59:01	Desain visualnya menarik dan modern...	positif
49. user4444	3	2023-03-01 60:12	Proses pendaftaran terlalu rumit...	negatif
50. user5555	5	2023-03-01 61:23	Fitur yang ditawarkan sangat berguna...	positif
51. user6666	4	2023-03-01 62:34	Stabilitas aplikasi cukup baik...	netral
52. user7777	3	2023-03-01 63:45	Ada beberapa fitur yang kurang...	negatif
53. user8888	5	2023-03-01 64:56	Rekomendasi aplikasi ini sangat tepat...	positif
54. user9999	4	2023-03-01 65:67	Desain visualnya menarik dan modern...	positif
55. user0000	3	2023-03-01 66:78	Proses pendaftaran terlalu rumit...	negatif
56. user1111	5	2023-03-01 67:89	Fitur yang ditawarkan sangat berguna...	positif
57. user2222	4	2023-03-01 68:90	Stabilitas aplikasi cukup baik...	netral
58. user3333	3	2023-03-01 69:01	Ada beberapa fitur yang kurang...	negatif
59. user4444	5	2023-03-01 70:12	Rekomendasi aplikasi ini sangat tepat...	positif
60. user5555	4	2023-03-01 71:23	Desain visualnya menarik dan modern...	positif
61. user6666	3	2023-03-01 72:34	Proses pendaftaran terlalu rumit...	negatif
62. user7777	5	2023-03-01 73:45	Fitur yang ditawarkan sangat berguna...	positif
63. user8888	4	2023-03-01 74:56	Stabilitas aplikasi cukup baik...	netral
64. user9999	3	2023-03-01 75:67	Ada beberapa fitur yang kurang...	negatif
65. user0000	5	2023-03-01 76:78	Rekomendasi aplikasi ini sangat tepat...	positif
66. user1111	4	2023-03-01 77:89	Desain visualnya menarik dan modern...	positif
67. user2222	3	2023-03-01 78:90	Proses pendaftaran terlalu rumit...	negatif
68. user3333	5	2023-03-01 79:01	Fitur yang ditawarkan sangat berguna...	positif
69. user4444	4	2023-03-01 80:12	Stabilitas aplikasi cukup baik...	netral
70. user5555	3	2023-03-01 81:23	Ada beberapa fitur yang kurang...	negatif
71. user6666	5	2023-03-01 82:34	Rekomendasi aplikasi ini sangat tepat...	positif
72. user7777	4	2023-03-01 83:45	Desain visualnya menarik dan modern...	positif
73. user8888	3	2023-03-01 84:56	Proses pendaftaran terlalu rumit...	negatif
74. user9999	5	2023-03-01 85:67	Fitur yang ditawarkan sangat berguna...	positif
75. user0000	4	2023-03-01 86:78	Stabilitas aplikasi cukup baik...	netral
76. user1111	3	2023-03-01 87:89	Ada beberapa fitur yang kurang...	negatif
77. user2222	5	2023-03-01 88:90	Rekomendasi aplikasi ini sangat tepat...	positif
78. user3333	4	2023-03-01 89:01	Desain visualnya menarik dan modern...	positif
79. user4444	3	2023-03-01 90:12	Proses pendaftaran terlalu rumit...	negatif
80. user5555	5	2023-03-01 91:23	Fitur yang ditawarkan sangat berguna...	positif
81. user6666	4	2023-03-01 92:34	Stabilitas aplikasi cukup baik...	netral
82. user7777	3	2023-03-01 93:45	Ada beberapa fitur yang kurang...	negatif
83. user8888	5	2023-03-01 94:56	Rekomendasi aplikasi ini sangat tepat...	positif
84. user9999	4	2023-03-01 95:67	Desain visualnya menarik dan modern...	positif
85. user0000	3	2023-03-01 96:78	Proses pendaftaran terlalu rumit...	negatif
86. user1111	5	2023-03-01 97:89	Fitur yang ditawarkan sangat berguna...	positif
87. user2222	4	2023-03-01 98:90	Stabilitas aplikasi cukup baik...	netral
88. user3333	3	2023-03-01 99:01	Ada beberapa fitur yang kurang...	negatif
89. user4444	5	2023-03-01 100:12	Rekomendasi aplikasi ini sangat tepat...	positif
90. user5555	4	2023-03-01 101:23	Desain visualnya menarik dan modern...	positif
91. user6666	3	2023-03-01 102:34	Proses pendaftaran terlalu rumit...	negatif
92. user7777	5	2023-03-01 103:45	Fitur yang ditawarkan sangat berguna...	positif
93. user8888	4	2023-03-01 104:56	Stabilitas aplikasi cukup baik...	netral
94. user9999	3	2023-03-01 105:67	Ada beberapa fitur yang kurang...	negatif
95. user0000	5	2023-03-01 106:78	Rekomendasi aplikasi ini sangat tepat...	positif
96. user1111	4	2023-03-01 107:89	Desain visualnya menarik dan modern...	positif
97. user2222	3	2023-03-01 108:90	Proses pendaftaran terlalu rumit...	negatif
98. user3333	5	2023-03-01 109:01	Fitur yang ditawarkan sangat berguna...	positif
99. user4444	4	2023-03-01 110:12	Stabilitas aplikasi cukup baik...	netral
100. user5555	3	2023-03-01 111:23	Ada beberapa fitur yang kurang...	negatif

Gambar 4. Hasil *scraping* data aplikasi satu sehat

Pada Gambar 4. menunjukkan dataset yang telah disimpan dalam format csv, terdapat atribut yang telah dilakukan *Scraping* yaitu data *username*, *score*, *at*, dan *content*. Data yang berhasil diambil melalui proses *Scraping* berlangsung pada bulan Maret 2023 sejumlah 16.245 data. Data ini terdiri dari informasi yang meliputi pengguna (*username*), peringkat (*rating*), waktu, dan ulasan yang diberikan oleh pengguna.

4.3. Undersampling Dataset

Terjadi ketidak seimbangan jumlah data positif dan negatif pada data ulasan Satu Sehat yang telah di *Scraping* pada bulan maret 2023.



Gambar 5. Setelah dilakukan *undersampling*

Berdasarkan gambar 5 diatas menunjukkan hasil *undersampling*, hal ini dilakukan agar proses klasifikasi dapat menghasilkan hasil yang lebih optimal dengan jumlah data yang seimbang antara label positif dan negatif sebanyak 3.694 label positif dan 3.694 label negatif.

4.4. Preprocessing

Pada tahap *preprocessing* terdapat beberapa tahap yang dilakukan agar dataset yang berisi review yang digunakan untuk analisis sentimen memudahkan dalam pemodelan. Tahap yang dilakukan yaitu *cleansing*, *case folding*, *tokenizing*, *filtering*, *stemming*. Tahapan *preprocessing* sebagai berikut.

4.4.1. Cleansing

Hasil dataset yang telah melalui tahap *cleansing* mencakup penghapusan atribut tidak relevan dalam teks, seperti angka, simbol, URL, hashtag, emoji, tandabaca, dan spasi berlebih. Pengolahan ini bertujuan untuk mendapatkan data yang lebih bersih dan mudah dianalisis.

Gambar 6. Hasil *running* tahap *cleansing*

4.4.2. Case Folding

index	test	value
0	bug dugk apknya	negatif
1	aplikasi yang terinstall memuat update update mulai mendring him yang nyu gado kalo ya gmn install gmn aplikasi gak bnguna gak ada feednya	negatif
2	sertifikat data update pertama ketika terinstall sempat sya hilang dari aplikasi ini update ini update yang mengakibatkan sya harus nyalain vakan	negatif
3	aplikasi ini diinstall sya sya ngapain juga laa, nomer sya di diffarikan jadi langsung beranda sya	negatif
4	kemana tempat mau masuk setelah update update di permu 3 ya gaba mau masuk tempat pnyob di portali termatkash	negatif
5	gak bisa terbuksa pnyoba	negatif
6	watash di update ngn nyu gaba di buksa tolong pnyoba	negatif
7	sayu mau cek sertifikat vakan di abstrahi pnyid id dan tak munud hasil sertifikasinya mohon perbaikan sy syh vakan kek	negatif
8	tidak bisa dibuksa di hape samsung 7 2016	negatif
9	cek bksa update	negatif

Pada Gambar 7 menunjukan hasil kolom comment yang telah melalui proses case *folding* pada dataset.

Hasil *dataset* yang telah melalui tahapan *tokenizing* merupakan data yang telah dilakukan dengan cara memisahkan atau memisahkan teks kalimat menjadi kata-kata individu berdasarkan komponen penyusunnya. Hasil *tokenizing* diperoleh data ulasan daftar kata-kata yang terpisah secara individu atau kata per kata dari teks kalimat yang awalnya utuh.

0	bug.dong.apknya	
1	aplikasi yang terkesan memaksa update,update.mutu.mendring.hn.ram.yas.gde.kula.yz.ram.kecil.gamam.aplikasi.gak.berguna.gak.ada.feedahnya	
2	aplikasi tidak update,visitas.pertama.kedua.ketiga.keempat.saya.hilang.dari.aplikasi.in,update.dan.update.yang.mengkilahkan,saya.harus.melakukan.kembali.vaksin.cara	
3	saya.sedikit.dip.hanya.kidulani.ya,ya,Bagaimana.kulak.nomer.yg.ditafirkan.kidulani.kesembung.dengan.wa	
4	4 kenapa saya.mau.mutak,update.d,update.or.mutak.d,aku.saya.mutak.kembali.yang.tolong.d,perbaiki,terimakasih	
5	gak.bisa.update.punya.saya	
6	setelah.d,update.apk.yang.bisa.d,update.bisa.kidulani.d,perbaiki	
7	saya.mau.cak,sertifikasi.vaksin,d,arabudahi.pki.lidat,dan.tak.muncul.hasil.sertifikasi,mohon.pertbaikan,sy.dsh.vaksin.kak	
8	tidak.bisa.update.d,siapa.samsung.7.2016	
9	say.mau.cak	

Pada Gambar 8 menunjukan bahwa kolom *comment* yang sebelumnya kalimat utuh diubah menjadi kata per kata yang dipisahkan dengan tanda baca koma (,) pada data kalimat ulasan

Hasil *dataset* yang telah melalui tahapan *filtering* merupakan data yang telah dilakukan penghapusan kata-kata yang tidak bermanfaat dan tidak memiliki nilai dalam proses penelitian. Hasil *filtering* diperoleh dengan menggunakan metode *remove stopwords* dalam *Library NLTK*.

index	text	value
0	bug,token,apkhya	negatif
1	aplikasi,keasikan,kekakuan,update,update,mu,maand,hu,ram,ray,gede,kalo,ip,an,grmania,aplikasi,gak,beguna,gak,teakhirnya	negatif
2	serifit,ada,sekan,hejga,keempit,filang,aplikasi,update,mengabatkan,sekan,serifit	negatif
3	look,up,dikem,vis,wa,none,ip,gak,perlahan,serembang,wa	negatif
4	masuk,update,env,mu,aku,ma,ip,se,pa,folong,d,perkali,terimakasih	negatif
5	gak,terbuka	negatif
6	update,gak,ray,gak,ku,masuk,perkali	negatif
7	cek,serifit,kekakuan,d,abubakhi,pi,munul,haki,serifikasi,maon,perlahan,pi,sdh,vaikin,kn	negatif
8	redix,d3ku,hapa,samsung,7/2016	negatif

Pada Gambar 9 menunjukkan hasil kolom *comment* yang telah melalui proses *filtering* pada data kalimat ulasan.

Hasil dari *stemming* yaitu mengurangi frekuensi kemunculan kata-kata turunan dengan menghapus imbuhan yang terdapat pada awal, akhir, atau keduanya dalam sebuah kata. Dengan kata lain, *stemming* membantu menyederhanakan kata-kata sehingga bentuk dasarnya dapat diidentifikasi dengan lebih mudah.

[illegible]

Pada Gambar 10 menunjukkan hasil kolom comment berbentuk teks yang telah melalui proses stemming pada data kalimat ulasan.

Setelah melalui tahapan *Preprocessing*, didapatkan hasil *dataset* yang telahbersih yaitu *dataset* yang telah melewati langkah-langkah seperti penghapusan karakter khusus, penghilangan tanda baca, penghapusan kata-kata yang tidak relevan, normalisasi teks, dan langkah-langkah lainnya yang bertujuan untuk menghasilkan data yang lebih terstruktur dan mudah untuk dianalisis.

Dataset yang telah melalui tahapan *Preprocessing* ini memiliki peningkatan kualitas yang lebih baik dan memudahkan dalam melakukan analisis sentimen pada tahap berikutnya. *Dataset* yang telah melalui tahapan *Preprocessing* akan disimpan dalam file terpisah untuk mempermudah proses implementasi yang akan dilakukan selanjutnya. Dengan begitu, data yang telah dipersiapkan dengan baik akan siap digunakan dalam analisis lebih lanjut.

[illegible]

3598

Pada Gambar 11 menunjukkan hasil dari dataset pada kolom comment yang telah dilakukan tahap *Preprocessing* yaitu telah melewati tahap *cleansing*, *case folding*, *tokenizing*, *filtering* dan *stemming*.

Data masukan memiliki 7388 baris dan 2 kolom
Data positif sebanyak = 3045 baris
Data negatif sebanyak = 4343 baris

Gambar 12. Check Jumlah Baris Dataset

Pada Gambar 12 menunjukkan hasil *check* jumlah dataset dari tahap *Preprocessing* yang telah diproses terdiri dari 5.000 baris data. Dari total jumlah tersebut, terdapat 2.500 baris data dengan label "Positif" dan 2.500 baris data dengan label "Negatif".

4.5. Data Transformation

Hasil proses data *Transformation* yang merupakan proses pembobotan data teks dengan perhitungan TF-IDF, diperoleh dengan cara data teks diubah menjadi nilai numerik dengan menggunakan *Library Python* yaitu *Scikit-learn* dengan *CountVectorizer*. Hasil nilai TF dan IDF (*Inverse Document Frequency*) dilakukan dengan menggunakan pembobotan nilai pada setiap kata yang dihitung probabilitasnya dengan rumus yang telah ditentukan.

```
(0, 346) 0.6279111394752293
(0, 1250) 0.6201801469325723
(0, 807) 0.47021716926715074
(1, 1403) 0.3221365678565432
(1, 1656) 0.17729521526346523
(1, 1487) 0.22968381213690667
(1, 1588) 0.14570904567879966
(1, 4980) 0.13119863114445268
(1, 2078) 0.18502687667454099
(1, 1546) 0.29010362716365296
(1, 3202) 0.12210084608179006
(1, 3692) 0.5566815439265339
(1, 1778) 0.15197066204941312
(1, 2779) 0.19451187968122483
(1, 2935) 0.20823558693023278
(1, 4748) 0.21634973231227367
(1, 3353) 0.23246757111816221
(1, 2220) 0.3079617657885365
(1, 356) 0.17440389839730425
(2, 219) 0.44519716882959304
(2, 1760) 0.2480920009527512
(2, 1349) 0.4814101422009263
(2, 4531) 0.39732622193061956
(2, 4815) 0.3426917705874165
(2, 1051) 0.22791932860159314
```

Gambar 13. Hasil *running* tahap data *transformation*

Pada Gambar 13 menunjukkan hasil proses Satu Sehat TF-IDF kata-kata yang terdapat sebanyak 7.388 data ulasan aplikasi Satu Sehat yang telah diolah. Hasil *running* menunjukkan data ulasan hasil dari proses *Transformation* yang menampilkan letak kata dalam dokumen, jumlah kata dan nilai TF-IDF kata.

4.6. Data Mining

Hasil proses data *mining* dengan menerapkan algoritma *Naïve Bayes Classifier* diperoleh dengan cara pembagian data *training* dan data *testing*. Klasifikasi dalam pembuatan model penelitian ini dilakukan dengan cara tiga skenario, dengan pembagian data *training* dan data *testing* masing-masing sebesar 70:30, 80:20 dan 90:10.

Tabel 1. Pembagian Skenario

Skenario 1	Data training (70%)	Data testing (30%)
	5.171	2.2.17
Skenario 2	Data training (80%)	Data testing (20%)
	5910	1.478
Skenario 3	Data training (90%)	Data testing (10%)
	6650	738

Pada tabel 1 Informasi mengenai pembagian *dataset* berdasarkan skenario yang telah dibuat. Hasil proses klasifikasi menghasilkan nilai-nilai *accuracy*, *precision*, *recall*, dan *f1-score*.

4.7. Pattern evaluation

Pada tahap *pattern evaluation*, hasil dari ketiga skenario data mining yang telah dijalankan menggunakan algoritma *Naïve Bayes Classifier* dievaluasi. Tahap ini melibatkan penilaian terhadap klasifikasi yang telah dilakukan dan analisis terhadap hasil pemodelan yang telah dilaksanakan.

4.7.1. Confusion Matrix

Tahap *confusion matrix* merupakan tahap untuk mengukur seberapa baik kinerja dari model yang telah diimplementasikan. Dengan menggunakan *confusion matrix*, akan dihitung nilai-nilai *accuracy*, *precision*, *recall*, dan *f1-score* yang akan digunakan sebagai parameter dalam penelitian.

Hasil yang diperoleh dari pengujian model klasifikasi analisis sentimen menggunakan algoritma *Naive Bayes* dengan menggunakan pembagian data latih dan uji pada 3 skenario yang diujikan ditunjukkan pada Tabel 2.

Tabel 2. Hasil pengujian model analisis sentimen

skenario	datas et	Accura cy	Precisi on	Reca ll	F!-Scor e
Skenario 1	70:30	92%	92%	92%	91%
Skenario 2	80:20	92%	93%	92%	92%
Skenario 3	90:10	93%	93%	92%	92%

Pada Tabel 2 merupakan hasil evaluasi dengan melihat *confusion matrix* dengan pemodelan menggunakan algoritma *Naïve Bayes Classifier*. Bisa dilihat dari hasil klasifikasi dengan menggunakan skenario yang berbeda memiliki persentase yang kurang lebih sama namun untuk skenario 3 menunjukkan persentase yang paling besar. *Confusion Matrix* dengan menggunakan 3 skenario juga menunjukkan hasil yang baik untuk mengenali label dengan value positif dan value negatif yang diprediksi benar.

4.8. Pembahasan

Penelitian ini menggunakan algoritma *Naïve Bayes Classifier* untuk membuat model analisis sentimen terhadap ulasan yang diberikan oleh pengguna aplikasi Satu Sehat di Google Play Store. Metode yang digunakan dalam penelitian ini adalah *Knowledge Discovery in Database* (KDD) yang melibatkan 5 tahapan. Tahapan tersebut meliputi data *Scraping*, *Preprocessing*, *data Transformation*, *data mining* dan *pattern evaluation*.

Data ulasan yang digunakan dalam penelitian ini sebanyak 16245 data. Data berhasil diambil melalui teknik web *Scraping* menggunakan bahasa pemrograman Python dari halaman ulasan aplikasi Satu Sehat di Google Play Store. Data diperoleh dalam waktu bulan Maret 2023. Data yang berhasil terkumpul dilakukan tahap memberikan label pada data tersebut berdasarkan rating bintang. Ulasan yang diberi rating 1, 2, dan 3 akan dianggap sebagai ulasan yang negatif atau mungkin memiliki kecenderungan untuk merujuk pada aspek negatif, sedangkan ulasan yang diberi rating 4 dan 5 akan dianggap sebagai ulasan yang positif dan mungkin cenderung menggambarkan aspek-aspek yang baik

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat ditarik kesimpulan sebagai berikut: Hasil analisis sentimen ulasan positif terhadap aplikasi Satu Sehat, yaitu tidak sedikit pula yang terbantu oleh aplikasi Satu Sehat karna membutuhkan data sertifikat vaksin untuk bepergian. Selain itu, aplikasi Satu Sehat mempermudah pengguna untuk mengakses data riwayat kesehatan secara detail dan terstruktur dengan baik. Hasil analisis sentimen ulasan negative aplikasi Satu Sehat, yaitu adanya keluhan pengguna terkait beberapa aspek yaitu para pengguna mengeluhkan bahwa penggunaan aplikasi Satu Sehat terasa ribet karna adanya bug yang membuat aplikasi sering tidak bisa dibuka dan sulit pada proses login. Pengguna mengalami masalah ketika login atau daftar karna kode OTP yang tidak terkirim ke pengguna. Berdasarkan analisis sentimen pengguna aplikasi Satu Sehat memperoleh sentimen negatif, terlihat dari banyaknya pengguna yang mengeluh tentang pemakaian aplikasi Satu Sehat.

Beberapa saran yang dapat diterapkan untuk perbaikan dan pengembangan penelitian adalah sebagai berikut. Analisis sentimen ulasan aplikasi Satu Sehat, selain menggunakan algoritma *Naïve Bayes Classifier*, dapat dipertimbangkan metode klasifikasi lain seperti *Support Vector Machine*, *Random Forest*, *K-Nearest Neighbor*, dan *Decision Tree*. Sehingga hasil pengujian model dapat ditingkatkan dan menghasilkan akurasi yang lebih tinggi. Penelitian selanjutnya dapat melakukan komparasi atau perbandingan dengan menggunakan algoritma *Support Vector Machine* (SVM) yaitu metode yang cukup akurat untuk mengklasifikasikan komentar positif dan negatif dari aplikasi satu sehat

sehingga dapat menghasilkan data yang lebih valid dan akurat.

DAFTAR PUSTAKA

- [1] Rokom. (2022a, April 26). *PeduliLindungi Telah Cegah Jutaan Warga Terpapar COVID-19*. Sehatnegeriku.Kemkes.Go.Id.
- [2] Rahma, N. A., Garno, G., & Sulistiyowati, N. (2022). Analisis Sentimen Tempat Wisata Di Jakarta Pasca Covid-19 Dengan Algoritma Naïve Bayes. *Jurnal Pendidikan Dan Konseling (JPDK)*, 4(6), 5894–5908.
- [3] Praptiwi, D. Y. (2018). *Analisis Sentimen Online Review Pengguna E- Commerce Menggunakan Metode Support Vector Machine Dan Maximum Entropy (Studi Kasus: Review Bukalapak pada GooglePlay)*.
- [4] Ramadhan, B. Z., Adam, R. I., & Maulana, I. (2022). Analisis Sentimen Ulasan pada Aplikasi E-Commerce dengan Menggunakan Algoritma Naïve Bayes. *Journal of Applied Informatics and Computing*, 6(2), 220–225. <https://doi.org/10.30871/jaic.v6i2.4725>
- [5] Auliya Agustina, D., Subanti, S., Zukhronah, E., Studi Statistika, P., & SebelasMaret, U. (2020). *Implementasi Text Mining Pada Analisis Sentimen Pengguna Twitter Terhadap Marketplace di Indonesia Menggunakan Algoritma Support Vector Machine*.
- [6] Fajrin, A. A., Maulana, A., Informatika, T., Batam, U. P., & Soeprapto, J. R. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan JurnaL Ilmu Komputer (KLIK)*, 5(1), 27–36.
- [7] Rahayu, S., Yumarlin, M. Z., Bororing, J. E., & Hadiyat, R. (2022). Implementasi Metode K-Nearest Neighbor (K-NN) untuk Analisis Sentimen Kepuasan Pengguna Aplikasi Teknologi Finansial FLIP. *Edumatic J. Pendidik. Inform*, 6(1), 98–106.
- [8] Yakub, S., Boy, A. F., Mariami, I., & Widjanarko, B. (2019). Penerapan DataMining Pengaturan Pola Tata Letak Barang Pada Berkah Swalayan Untuk Strategi Penjualan Menggunakan Algoritma Apriori. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 2(1), 69–75.
- [9] Herianto, H., Syamsiyah, N., & Yahya, Y. (2021). Evaluasi Kinerja Datamining Pada Dataset Pendaftaran Mahasiswa Baru Dengan Class Yang Tidak Seimbang. *IKRA-ITH Informatika: Jurnal Komputer DanInformatika*, 5(3), 162–168.
- [10] Oktavia, R., Tata Hardinata, J., Tunas Bangsa, S., -Indonesia, P., Tunas Bangsa, A., -Indonesia Jln Sudirman Blok No, P. A., & Utara, S. (n.d.). *Penerapan Metode Algoritma K-means Dalam Pengelompokan Angka Harapan Hidup Saat*

Lahir Menurut Provinsi (Vol. 1, Issue 4).
www.bps.go.id/