

Sentiment Analysis of Sirekap Application Users Using the Support Vector Machine Algorithm

Joko Setyanto ^{1*}, Theopilus Bayu Sasongko ^{2*}

* Informatika, Universitas Amikom Yogyakarta

jsetya24@students.amikom.ac.id ¹, theopilus.27@amikom.ac.id ²

Article Info

Article history:

Received 2024-06-11

Revised 2024-06-13

Accepted 2024-06-20

Keyword:

Sentiment Analysis,
Sirekap,
Support Vector Machine,
Wordcloud

ABSTRACT

In the current era of digitalization, various activities are conducted using technology to aid their execution, including the democratic process scheduled for February 2024. The Komisi Pemilihan Umum (KPU) is utilizing a mobile-based application called Sirekap. During the previous presidential and vice-presidential elections, there were many pros and cons regarding the Sirekap application. A significant number of negative reviews were expressed by the public towards this application. This study employs the SVM algorithm to perform sentiment analysis of Sirekap application users. Before building the model, several steps were undertaken, including data labeling, data preprocessing, splitting the dataset into training and testing data, and performing transformations using Count Vectorizer. Evaluation of the SVM model results shows quite good performance with an accuracy of 81%. For the negative class, the precision and recall values are 87% and 85%, respectively, while for the positive class, the precision and recall values only reach 66% and 70%, indicating a need for improvement in the model's identification of the positive class. Five-fold cross-validation was performed with an average cross-validation score of 79.6% and a standard deviation of 2.14%, indicating the model's consistency across various training data subsets. These findings suggest that the SVM model can effectively perform text classification tasks. Based on the negative word cloud, it can be concluded that the Sirekap application still has many shortcomings affecting the democratic process in February 2024.



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

I. PENDAHULUAN

Indonesia merupakan negara demokrasi, yang artinya kedaulatan tertinggi ada di tangan rakyat. Pesta demokrasi sangat penting bagi bangsa Indonesia karena saat pesta demokrasi, rakyat dapat memberikan hak pilih kepada calon pemimpin yang akan membawa Indonesia menjadi lebih baik. Pada era digitalisasi seperti sekarang ini, beragam kegiatan dilakukan dengan memanfaatkan teknologi untuk membantu proses pelaksanaannya, termasuk pesta demokrasi yang dilaksanakan bulan Februari tahun 2024 lalu. Komisi Pemilihan Umum (KPU) menggunakan sebuah aplikasi berbasis *mobile*, yaitu Sirekap. Sirekap berfungsi untuk membantu proses rekapitulasi hasil penghitungan suara yang dilakukan KPU dari tingkat kecamatan hingga nasional. Pada tahun 2020, Sirekap pertama kali digunakan sebagai alat bantu pemilihan umum [1]. Pemilihan umum berfungsi

sebagai manifestasi nyata dari demokrasi prosedural, di mana keterlibatan aktif warga negara memainkan peran penting dalam menentukan jalannya pemerintahan [2]. Tujuan utama dari aplikasi Sirekap, selain untuk membantu proses rekapitulasi hasil penghitungan suara, adalah untuk menunjukkan transparansi proses pemilihan umum [3], karena publik bisa mengakses hasil penghitungan suara yang dilakukan di setiap Tempat Pemungutan Suara (TPS) secara langsung. Sirekap akan menampilkan C. Hasil pada setiap TPS yang telah berhasil diunggah oleh petugas Kelompok Penyelenggara Pemungutan Suara (KPPS).

Pada pemilu presiden dan wakil presiden lalu, terjadi banyak pro dan kontra mengenai aplikasi Sirekap. Banyak keluhan atau ulasan negatif yang disampaikan oleh publik terhadap aplikasi ini [4]. Beberapa masyarakat merasa bahwa aplikasi Sirekap yang dirancang untuk membantu proses pemilihan umum justru mempersulit petugas di lapangan dan

menimbulkan banyak kontroversi. Penggunaan teknologi ini dianggap tidak sepenuhnya efektif dan efisien, sering kali menyebabkan kesalahan teknis dan kebingungan di kalangan petugas KPPS. Banyak sekali perbincangan mengenai aplikasi ini, baik di sosial media maupun siaran televisi. Meskipun tujuan utamanya adalah membantu proses rekapitulasi dan meningkatkan transparansi, banyak yang merasa bahwa aplikasi ini masih belum siap digunakan dan justru menambah beban kerja petugas serta menimbulkan kecurigaan terhadap integritas proses pemilihan umum. Sehingga untuk mengatasi masalah tersebut dan memberikan gambaran yang lebih jelas mengenai persepsi masyarakat terhadap Sirekap, peneliti melakukan analisis sentimen pengguna terhadap aplikasi Sirekap menggunakan metode *Support Vector Machine* (SVM).

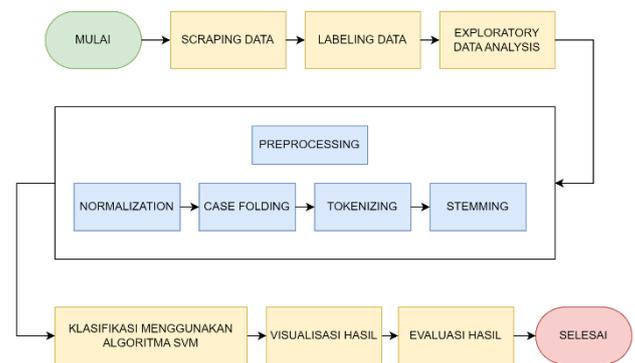
Peneliti sebelumnya [5] menggunakan algoritma *Random Forest Classifier* untuk melakukan sentimen analisis pengguna terhadap aplikasi Sirekap, dengan hasil akurasi sebesar 74%. Dalam penelitian kali ini, peneliti menggunakan algoritma *Support Vector Machine* yang dianggap efektif dalam menangani analisis sentimen, dibuktikan pada penelitian sebelumnya, beberapa peneliti telah melakukan komparasi algoritma untuk menganalisis sentimen dan menemukan bahwa algoritma SVM memberikan hasil akurasi yang paling maksimal [6]–[8]. Pada penelitian [6] melakukan komparasi antara SVM dan *Naive Bayes* (NB), menunjukkan akurasi 88,61% untuk algoritma SVM dan 88,24% untuk algoritma NB. Penelitian [7] melakukan komparasi antara SVM dan *K-Nearest Neighbour* (KNN) dengan hasil akurasi SVM sebesar 87,98% sedangkan akurasi KNN sebesar 82,14%. Pada penelitian [8] melakukan komparasi tiga algoritma yaitu SVM, KNN, dan NB, diperoleh akurasi untuk SVM, KNN, dan NB berturut-turut sebesar 76,5%, 59,1%, dan 72,3%. Walaupun pada beberapa penelitian terdapat perbedaan akurasi yang tidak terlalu signifikan, berdasarkan hasil tersebut, peneliti memutuskan untuk menggunakan algoritma *Support Vector Machine* untuk melakukan klasifikasi sentimen pengguna terhadap aplikasi Sirekap.

Pada penelitian ini, peneliti mengambil sampel data dari ulasan pada aplikasi Sirekap di *Google Play Store* [9]. Ketika peneliti mengambil sampel data, aplikasi Sirekap tercatat mendapat nilai yang cukup buruk dari penggunaannya yaitu 1.8 dari 5. Dari sampel yang telah diambil akan dikategorikan menjadi dua kelas yaitu sentimen positif dan sentimen negatif [10]. Penelitian ini bertujuan untuk menggali informasi dari ulasan pengguna aplikasi Sirekap guna mengidentifikasi kesulitan yang dialami selama proses pemilihan umum. Ulasan-ulasan tersebut akan dianalisis untuk menemukan pola keluhan dan tantangan yang dihadapi oleh pengguna. Dengan pemahaman yang lebih mendalam, diharapkan pengembang aplikasi dapat melakukan perbaikan dan pengembangan pada aplikasi Sirekap, sehingga proses pemilihan umum ke depannya dapat berjalan lebih baik dan lancar. Selain itu, penelitian ini juga bertujuan untuk menguji efektivitas algoritma SVM dalam mengklasifikasikan sentimen berdasarkan data ulasan aplikasi yang diambil dari

Google Play Store. Dengan menggunakan algoritma SVM yang telah terbukti memiliki akurasi tinggi dalam mengklasifikasikan sentimen, peneliti berharap dapat memberikan gambaran yang jelas tentang sentimen pengguna terhadap aplikasi Sirekap.

II. METODE PENELITIAN

Metode penelitian adalah panduan yang mencakup rencana serta langkah-langkah yang akan diambil dalam penelitian guna mencapai tujuan penelitian [11]. Penelitian ini menggunakan algoritma SVM untuk melakukan analisis sentimen pengguna aplikasi Sirekap. Adapun alur penelitian yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

A. Scraping Data

Data yang akan digunakan pada penelitian ini yaitu ulasan pengguna yang di ambil melalui teknik *scraping* data [12] pada *Google Play Store*. Peneliti menggunakan *library google-play-scraper* [13] dalam melakukan tahapan ini. Data diambil secara acak dan bertahap dengan jumlah 1791 data.

B. Labeling Data

Hasil dari *scraping* data selanjutnya akan diberi label positif dan negatif. Peneliti melakukan pelabelan bersama beberapa orang pelabel supaya memperoleh hasil yang maksimal [11], peneliti tidak mengacu pada *score* karena dirasa beberapa ulasan tidak relevan dengan *score* yang diberikan. Ulasan dengan label positif berisi usulan, sanjungan, dan kepuasan pengguna, sedangkan ulasan dengan label negatif berisi keluhan, protes, dan perasaan negatif seperti kekecewaan dan kemarahan pengguna aplikasi Sirekap.

C. Exploratory Data Analysis

Tujuan dari *Exploratory Data Analysis* (EDA) adalah mendapat wawasan awal ke dalam kumpulan data yang ada, sebelum menentukan proses penelitian yang lebih mendalam [14]. EDA memungkinkan ekstraksi berbagai detail awal yang berkaitan dengan data yang ditinjau, seperti volume data, kelengkapan data, jumlah data berlabel positif, dan jumlah data berlabel negatif.

D. Preprocessing Data

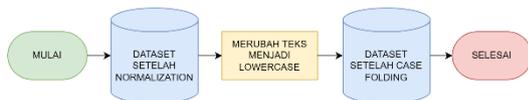
Tahap preprocessing data merupakan langkah penting yang harus dilakukan sebelum melakukan pembuatan model *text mining* [15]. Peneliti melakukan analisis singkat terhadap karakteristik data, sehingga menggunakan beberapa tahap preprocessing seperti *normalization*, *case folding*, *tokenizing*, dan *stemming*.

1) *Normalization*: Tahap *normalization* merupakan proses memperbaiki kata tidak baku menjadi kata baku pada teks ulasan pengguna [10]. Ketika melakukan pelabelan, peneliti memperhatikan terdapat beberapa kata yang mungkin kurang tepat ataupun berupa singkatan kata, sehingga diperlukan proses *normalization* untuk kata-kata tersebut. Beberapa contoh kata yang dilakukan *normalization* dapat dilihat pada Tabel I.

TABEL I
CONTOH KATA NORMALIZATION

No	Kata tidak baku/singkatan	Kata baku
1	gk, g, gak, ngak, nggak, gx, ga, tdk	tidak
2	gabisa, gbs, gbsa	tidak bisa
3	ok, oke, okk	baik
4	jos	sangat baik
5	blm, blum	belum
6	apk	aplikasi

2) *Case Folding*: Tahap *case folding* merupakan proses mengganti keseluruhan teks ulasan pengguna menjadi *lowercase* [16]. Tahap ini bertujuan agar teks yang akan diproses oleh model dapat memiliki kondisi bentuk yang sama. Proses pada tahap *case folding* dapat dilihat pada Gambar 2.



Gambar 2. Proses *case folding*

3) *Tokenizing*: Tahap *tokenizing* merupakan proses untuk memecah teks menjadi kata-kata [17]. Tahap ini bertujuan untuk merepresentasikan teks dengan cara yang mudah dipahami oleh model, dengan tahapan ini algoritma akan lebih mudah mengidentifikasi pola pada teks. Proses pada tahap *tokenizing* dapat dilihat pada Gambar 3.



Gambar 3. Proses *tokenizing*

4) *Stemming*: Tahap *stemming* merupakan proses untuk merubah kata berimbuhan menjadi kata dalam bentuk dasar [18]. Tahap ini bertujuan untuk mengurangi *unique word* yang harus diakomodasikan oleh model sehingga bisa

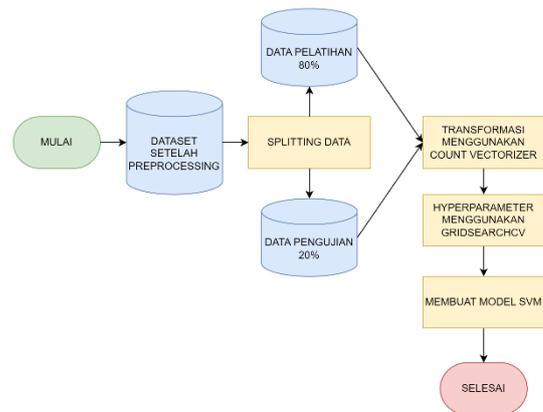
meningkatkan performa model. Proses pada tahap *stemming* dapat dilihat pada Gambar 4.



Gambar 4. Proses *stemming*

E. Klasifikasi Menggunakan Algoritma SVM

SVM merupakan model klasik untuk klasifikasi dua kelas, cara kerjanya yaitu dengan mencari *hyperplane* yang sesuai untuk memisahkan sampel data yang dikumpulkan [19]. Pada penelitian ini, peneliti menggunakan algoritma SVM untuk melakukan klasifikasi ulasan pengguna terhadap aplikasi Sirekap. Proses ini dimulai dengan membagi dataset menjadi dua yaitu 80% data pelatihan dan 20% data pengujian [20]. Setelah membagi dataset, data pelatihan diubah menjadi representasi numerik dengan *Count Vectorizer* guna mempelajari kosakata dari data pelatihan dan membuat fitur berdasarkan frekuensi kata [21]. Proses ini juga diterapkan untuk menguji data menggunakan kosakata yang telah dipelajari pada data pelatihan. Selanjutnya, dilakukan optimasi *hyperparameter* pada algoritma SVM menggunakan *GridSearchCV* guna menguji berbagai kombinasi parameter seperti nilai C dan tipe *kernel* [22]. Proses tersebut dilakukan guna mencari kombinasi parameter dengan hasil akurasi terbaik. Proses pada tahap ini dapat dilihat pada Gambar 5.



Gambar 5. Proses klasifikasi menggunakan SVM

F. Visualisasi Hasil

Dalam melakukan visualisasi hasil, peneliti menggunakan *wordcloud* yang berupa alat visual yang berfungsi untuk menampilkan kata-kata yang sering muncul [23]. Dengan menggunakan *wordcloud*, diharapkan dapat memberikan gambaran yang jelas mengenai frekuensi kemunculan kata pada ulasan pengguna.

G. Evaluasi Hasil

Penelitian ini menggunakan *Confusion Matrix* untuk menghitung akurasi, dengan membandingkan prediksi yang benar dan salah dari metode klasifikasi dengan data

sesungguhnya atau prediksi target [24]. Rumus penghitungan *confusion matrix* untuk menghitung *accuracy*, *precision*, dan *recall* dapat dilihat pada rumus berikut [25], [26].

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

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

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

Keterangan:

- TP (*True Positive*) = Jumlah data dari kelas positif yang berhasil diklasifikasikan dengan benar sebagai kelas positif.
- TN (*True Negative*) = Jumlah data dari kelas negatif yang berhasil diklasifikasikan dengan benar sebagai kelas negatif.
- FP (*False Positive*) = Jumlah data dari kelas negatif yang salah diklasifikasikan sebagai kelas positif.
- FN (*False Negative*) = Jumlah data dari kelas positif yang salah diklasifikasikan sebagai kelas negatif.

III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data ulasan pengguna aplikasi Sirekap yang tersedia di *Google Play Store*. Data yang diambil berupa teks yang berisi ulasan pengguna dan diambil secara acak dan bertahap pada 2 April 2024. Contoh hasil dari *scraping* data dapat dilihat pada Tabel II.

TABEL II
CONTOH HASIL SCRAPING DATA

No	userName	score	at	content
1	M. Mansur Sasmita	4	13/02/2024 14:22	Semoga aplikasi sirekap ini bisa digunakan dengan baik tidak error
2	Mitchie Yana	2	12/02/2024 04:52	Foto berkali" gambarnya ttep rusak. Tp datanya masuk. Ya jelas ga sesuai lah. Kalau disesuaikan mlah error. Bikin pusing.
3	Sugeng Muhamad Muklis	3	08/02/2024 11:13	Ternyata kita musti sabar ya cari waktu yg gak banyak pengguna, masih hrs dioptimalkan sistemnya biar mudah untuk mengaksesnya

Sebanyak 1791 data telah terkumpul dan selanjutnya akan dilakukan proses pelabelan data. Proses ini dilakukan secara manual dengan beberapa orang pelabel guna mendapat hasil yang maksimal. Pelabelan ini bertujuan untuk mengklasifikasikan data menjadi dua kategori, yaitu data berlabel positif dan data berlabel negatif. Contoh hasil dari pelabelan data dapat dilihat pada Tabel III.

TABEL III
CONTOH HASIL PELABELAN DATA

No	content	label
1	Semoga aplikasi sirekap ini bisa digunakan dengan baik tidak error	POSITIF
2	Foto berkali" gambarnya ttep rusak. Tp datanya masuk. Ya jelas ga sesuai lah. Kalau disesuaikan mlah error. Bikin pusing.	NEGATIF
3	Ternyata kita musti sabar ya cari waktu yg gak banyak pengguna, masih hrs dioptimalkan sistemnya biar mudah untuk mengaksesnya	NEGATIF

Setelah proses *labeling*, dilakukan *exploratory data analysis*, pada tahap ini peneliti melakukan pengecekan awal terhadap dataset. Tidak ditemukan kekosongan atau nilai *null* pada dataset, dari 1791 data yang dikategorikan, 1265 berlabel positif dan 526 berlabel negatif.

Tahap selanjutnya yaitu *preprocessing* data, proses ini merupakan tahapan penting dalam melakukan penambahan teks. Contoh hasil dari *preprocessing* data dapat dilihat pada Tabel IV.

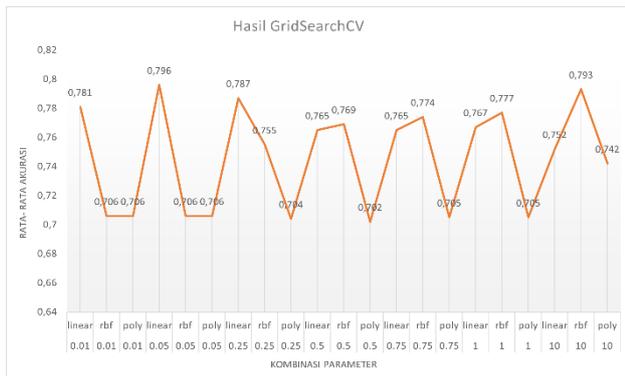
TABEL IV
CONTOH HASIL PREPROCESSING DATA

No	Tahapan	Hasil Preprocessing
1	Data Awal	Saran saja, kalo program kerja apapun dari pemerintah mau pake aplikasi itu yg sudah teruji bener? bisa lancar buat banyak orang jadi yang pake aplikasi nya nggak kesulitan dan menghambat pekerjaan, karena saya banyak mengalami hambatan saat menggunakan aplikasi nya contoh : gabisa masuk dan loading lama, jadi mohon diperbaiki kendala semacam itu tolong dipertimbangkan, terimakasih
2	Normalization	Saran saja kalo program kerja apapun dari pemerintah mau pake aplikasi itu yang sudah teruji bener bisa lancar buat banyak orang jadi yang pake aplikasi nya tidak kesulitan dan menghambat pekerjaan karena saya banyak mengalami hambatan saat menggunakan aplikasi nya contoh tidak bisa masuk dan loading lama jadi mohon diperbaiki kendala semacam itu tolong dipertimbangkan terimakasih
3	Case Folding	saran saja kalo program kerja apapun dari pemerintah mau pake aplikasi itu yang sudah teruji bener bisa lancar buat banyak orang jadi yang pake aplikasi nya tidak kesulitan dan menghambat pekerjaan karena saya banyak mengalami hambatan saat menggunakan aplikasi nya contoh tidak bisa masuk dan loading lama jadi mohon diperbaiki kendala semacam itu tolong dipertimbangkan terimakasih
4	Tokenizing	['saran', 'saja', 'kalo', 'program', 'kerja', 'apapun', 'dari', 'pemerintah', 'mau', 'pake', 'aplikasi', 'itu', 'yang', 'sudah', 'teruji', 'bener', 'bisa', 'lancar', 'buat', 'banyak', 'orang', 'jadi', 'yang', 'pake', 'aplikasi', 'nya', 'tidak', 'kesulitan', 'dan', 'menghambat', 'pekerjaan', 'karena', 'saya', 'banyak', 'mengalami', 'hambatan', 'saat', 'menggunakan', 'aplikasi', 'nya', 'contoh', 'tidak', 'bisa', 'masuk', 'dan', 'loading', 'lama', 'jadi', 'mohon', 'diperbaiki', 'kendala', 'semacam', 'itu', 'tolong', 'dipertimbangkan', 'terimakasih']

5	Stemming	['saran', 'saja', 'kalo', 'program', 'kerja', 'apa', 'dari', 'perintah', 'mau', 'pake', 'aplikasi', 'itu', 'yang', 'sudah', 'uji', 'bener', 'bisa', 'lancar', 'buat', 'banyak', 'orang', 'jadi', 'yang', 'pake', 'aplikasi', 'nya', 'tidak', 'sulit', 'dan', 'hambat', 'kerja', 'karena', 'saya', 'banyak', 'alami', 'hambat', 'saat', 'guna', 'aplikasi', 'nya', 'contoh', 'tidak', 'bisa', 'masuk', 'dan', 'loading', 'lama', 'jadi', 'mohon', 'baik', 'kendala', 'macam', 'ituu', 'tolong', 'timbang', 'terimakasih']
---	----------	--

Data yang telah diolah melalui *preprocessing* data, akan dibagi menjadi data pelatihan dan data pengujian. Dengan menggunakan perbandingan 80:20 terdapat sebanyak 1432 data pelatihan dan 359 data pengujian. Selanjutnya, dilakukan *Count Vectorizer* guna mempelajari kosakata yang ada pada data pelatihan serta melakukan pembuatan fitur berdasarkan frekuensi kemunculan kata. Proses tersebut juga dilakukan pada data pengujian menggunakan kosakata yang telah dipelajari pada data pelatihan.

Tahap berikutnya yaitu melakukan optimasi *hyperparameter* menggunakan *GridSearchCV* untuk menemukan kombinasi parameter terbaik. Parameter nilai C yang diuji yaitu 0.01, 0.05, 0.25, 0.5, 0.75, 1, dan 10. Sedangkan parameter jenis kernel yang diuji yaitu *linear*, *rbf*, dan *poly*. Hasil dari proses optimasi *hyperparameter* menggunakan *GridSearchCV* dapat dilihat pada Gambar 6.



Gambar 6. Grafik hasil *GridSearchCV*

Berdasarkan hasil tersebut, model SVM yang akan dibuat akan menggunakan parameter Nilai C 0.05 dan jenis *kernel linear*. Pemilihan parameter ini didasarkan pada performa terbaik yang diperoleh selama proses *cross validation* atau validasi silang. Dengan parameter tersebut, model SVM diharapkan dapat secara efektif dalam melakukan klasifikasi data teks ke dalam kategori positif dan negatif. Selanjutnya, model akan dilatih menggunakan data pelatihan yang telah diproses sebelumnya, lalu kemudian dievaluasi kinerjanya pada data pengujian guna memastikan akurasi yang diharapkan.

Visualisasi hasil berupa *wordcloud* negatif yang merupakan representasi visual dari data yang berlabel negatif berupa kata yang sering muncul guna menggali informasi kekurangan dari aplikasi Sirekap menurut sentimen pengguna. *Wordcloud* negatif dapat dilihat pada Gambar 7.



Gambar 7. Visualisasi *Wordcloud* Negatif

Evaluasi hasil model SVM menunjukkan kinerja yang cukup baik dengan akurasi model mencapai 81%. Pada kategori negatif, model mendapat nilai *precision* 87% dan *recall* 85% menunjukkan model sangat baik dalam mengidentifikasi ulasan negatif. Namun, pada kategori positif, *precision* hanya mencapai 66% dan *recall* 70% yang menunjukkan model masih membutuhkan perbaikan dalam mengidentifikasi ulasan positif. Selain itu, dilakukan validasi silang sebanyak 5 lipatan dengan rata-rata *cross validation* sebesar 79,6% dengan standar deviasi 2,14% menunjukkan konsistensi model pada berbagai subset data pelatihan. Temuan ini menunjukkan bahwa model SVM mampu melakukan tugas klasifikasi teks secara efektif.

IV. KESIMPULAN

Berdasarkan hasil analisis sentimen pengguna terhadap aplikasi Sirekap, dari 1791 data ulasan pengguna, 1265 ulasan merupakan opini negatif dan 526 ulasan merupakan opini positif. Hasil tersebut menunjukkan 70,63% opini yang disampaikan oleh pengguna merupakan bentuk protes, kekecewaan, maupun perasaan negatif lainnya. Berdasarkan *wordcloud* negatif, dapat ditarik kesimpulan bahwa aplikasi Sirekap masih memiliki banyak kekurangan yang berpengaruh pada proses demokrasi pada bulan Februari tahun 2024 lalu. Frekuensi kata yang sering muncul menunjukkan keluhan pengguna selama menggunakan aplikasi tersebut. Banyak keluhan pengguna yang berhubungan dengan fitur login, foto, dan kesulitan dalam mengakses aplikasi.

Hasil pengujian pada algoritma SVM diperoleh hasil bahwa dengan menggunakan parameter nilai C = 0.05 dan jenis *kernel linear* memperoleh hasil yang terbaik dengan akurasi model mencapai 81%. Terbukti bahwa algoritma *Support Vector Machine* dapat secara efektif melaksanakan tugas klasifikasi teks melalui penyempurnaan parameter yang paling sesuai.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada Program Studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang telah memberikan bimbingan dan dukungan pada penelitian ini.

DAFTAR PUSTAKA

- [1] H. KPU, "KPU-Bawaslu Rapat Persiapan Implementasi Sirekap," Komisi Pemilihan Umum. Accessed: Jun. 06, 2024. [Online]. Available: <https://www.kpu.go.id/berita/baca/8224/Guna-meningkatkan-transparansi-serta-akuntabilitas-pelaksanaan-pemilihan-maupun-pemilihan-umum--Komisi-Pemilihan-Umum--KPU--RI-bersama-Badan-Pengawas-Pemilu--Bawaslu>
- [2] D. Handita and D. L. Anggraini, "Penerapan Pemilu Online Berbasis Aplikasi Smartphone di Era Pandemi Covid-19," *Pros. Semin. Nas. Desain Sos.*, pp. 848–851, 2021.
- [3] H. KPU, "Semua TPS Gunakan Sirekap Sebagai Alat Bantu dan Publikasi," Komisi Pemilihan Umum. Accessed: Jun. 06, 2024. [Online]. Available: <https://www.kpu.go.id/berita/baca/9133/Sistem-Informasi-Rekapitulasi--Sirekap--tetap-akan-dipergunakan-Komisi-Pemilihan-Umum--KPU--pada-proses-rekapitulasi-hasil-perolehan-suara-di-Tempat-Pemungutan-Suara--TPS->
- [4] D. D. Purnamasari, "Sirekap, Alat Bantu Pemilu yang Justru Timbulkan Kegaduhan," *kompas.id*. Accessed: Jun. 06, 2024. [Online]. Available: <https://www.kompas.id/baca/polhuk/2024/02/17/sirekap-alat-bantu-pemilu-yang-justru-timbulkan-kegaduhan>
- [5] M. F. Y. Herjanto and C. Carudin, "Analisis Sentimen Ulasan Pengguna Aplikasi Sirekap Pada Play Store Menggunakan Algoritma Random Forest Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 2, pp. 1204–1210, 2024, doi: 10.23960/jitet.v12i2.4192.
- [6] K. Solecha and O. Irnawati, "Komparasi Algoritma Support Vector Machine Dan Naïve Bayes Berbasis Particle Swarm Optimization Pada Analisis Sentimen Ulasan Aplikasi Flip," *JIEET (Journal Inf. Eng. Educ. Technol.)*, vol. 07, no. 1, pp. 10–15, 2023.
- [7] M. N. Muttaqin and I. Kharisudin, "Analisis Sentimen Pada Ulasan Aplikasi Gojek Menggunakan Metode Support Vector Machine dan K Nearest Neighbor," *UNNES J. Math.*, vol. 10, no. 2, pp. 22–27, 2021, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>
- [8] A. Salma and W. Silfianti, "Sentiment Analysis of User Review on COVID-19 Information Applications Using Naïve Bayes Classifier, Support Vector Machine, and K-Nearest Neighbors," *Int. Res. J. Adv. Eng. Sci.*, vol. 6, no. 4, pp. 158–162, 2021.
- [9] K. P. Umum, "SIREKAP 2024," Google Play Store. [Online]. Available: <https://play.google.com/store/apps/details?id=id.go.kpu.sirekap2024&hl=id>
- [10] T. Safitri, Y. Umidah, and I. Maulana, "Analisis Sentimen Pengguna Twitter Terhadap Grup Musik BTS Menggunakan Algoritma Support Vector Machine," *J. Appl. Informatics Comput.*, vol. 7, no. 1, pp. 28–35, 2023, doi: 10.30871/jaic.v7i1.5039.
- [11] H. P. Doloksaribu and Yusran Timur Samuel, "Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, pp. 1–11, 2022, doi: 10.47111/jti.v16i1.3747.
- [12] R. Maulana, A. Voutama, and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC," *J. Teknol. Terpadu*, vol. 9, no. 1, pp. 42–48, 2023, doi: 10.54914/jtt.v9i1.609.
- [13] JoMingyu, "google-play-scraper." [Online]. Available: <https://pypi.org/project/google-play-scraper/>
- [14] R. Merdiansah, S. Siska, and A. Ali Ridha, "Analisis Sentimen Pengguna X Indonesia Terkait Kendaraan Listrik Menggunakan IndoBERT," *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, pp. 221–228, 2024, doi: 10.55338/jikoms.v7i1.2895.
- [15] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknol.)*, vol. 12, no. 1, pp. 129–137, 2023, doi: 10.23887/jstundiksha.v12i1.52358.
- [16] G. R. Ditami, E. F. Ripanti, and H. Sujaini, "Implementasi Support Vector Machine untuk Analisis Sentimen Terhadap Pengaruh Program Promosi Event Belanja pada Marketplace," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 3, p. 508, 2022, doi: 10.26418/jp.v8i3.56478.
- [17] M. N. B. Balit and F. S. Utomo, "Sentiment Analysis of pegipegi.com Review on Google Play Store with Naïve Bayes," *Sistemasi*, vol. 13, no. 3, pp. 1044–1053, 2024, doi: 10.32520/stmsi.v13i2.3887.
- [18] D. Safryda Putri and T. Ridwan, "Analisis Sentimen Ulasan Aplikasi Pospay Dengan Algoritma Support Vector Machine," *J. Ilm. Inform.*, vol. 11, no. 01, pp. 32–40, 2023, doi: 10.33884/jif.v11i01.6611.
- [19] B. Gaye, D. Zhang, and A. Wulamu, "Improvement of Support Vector Machine Algorithm in Big Data Background," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/5594899.
- [20] R. Tinages, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 650, 2020, doi: 10.30865/mib.v4i3.2181.
- [21] T. Turki and S. S. Roy, "Novel Hate Speech Detection Using Word Cloud Visualization and Ensemble Learning Coupled with Count Vectorizer," *Appl. Sci.*, vol. 12, no. 13, 2022, doi: 10.3390/app12136611.
- [22] M. A. Muslim *et al.*, "Support Vector Machine (SVM) Optimization Using Grid Search and Unigram to Improve E-Commerce Review Accuracy," *J. Soft Comput. Explor.*, vol. 1, no. 1, pp. 8–15, 2020, doi: 10.52465/jsosex.v1i1.3.
- [23] B. F. S. Supriyanto and S. Rosalin, "Analisis Sentimen Program Merdeka Belajar dengan Text Analysis Wordcloud & Word Frequency," *J. Minfo Polgan*, vol. 12, no. 1, pp. 25–32, 2023, doi: 10.33395/jmp.v12i1.12312.
- [24] N. Cahyono and Dewi Setiyawati, "Analisis Sentimen Pengguna Sosial Media Twitter Terhadap Perokok Di Indonesia," *Indones. J. Comput. Sci.*, vol. 12, no. 1, pp. 262–272, 2023, doi: 10.33022/ijcs.v12i1.3154.
- [25] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 697–711, 2021.
- [26] D. Krstinić, M. Braović, L. Šerić, and D. Božić-Štulić, "Multi-label Classifier Performance Evaluation with Confusion Matrix," *Comput. Sci. Inf. Technol. (CS IT)*, pp. 01–14, 2020, doi: 10.5121/csit.2020.100801.