

Penerapan Algoritma Support Vector Machine dan Particle Swarm Optimization Untuk Pembuatan Rating Aplikasi Google Playstore

Application of the Support Vector Machine and Particle Swarm Optimization Algorithms for Making Google Playstore Application Ratings

Arif Rahman^{1*}, Muhammad Imam Dinata²

^{1,2}Program Sistem Teknologi Informasi, Fakultas Teknik

Universitas Muhammadiyah Mataram

Jl. KH. Ahmad Dahlan No. 1 Pagesangan, Mataram, NTB, Indonesia

email: ^{*}arif.rahman@ummat.ac.id, ²imam.dinata@ummat.ac.id

ABSTRAK

Penggunaan *smartphone* yang besar khususnya android di indonesia sejalan dengan munculnya aplikasi untuk membantuk kegiatan masyarakat indonesia seperti gojek, grap, tokopedia, shopee dan lain sebagainya. aplikasi-aplikasi ini di unduh pada aplikasi google yang bernama google playstore, selain berfungsi sebagai penyimpanan aplikasi, google playstore bisa digunakan untuk mereview aplikasi dan memberikan rating terhadap aplikasi tersebut. seringkali *review* dan rating yang diberikan tidak searah karena hal ini rating yang diberikan pada aplikasi tidak bias, sehingga peneliti membuat sebuah penelitian untuk membuat rating berdasarkan klasifikasi *review* yang diberikan oleh user, algoritma yang digunakan untuk klasifikasi ini adalah algoritma support vectore machine dan algoritma particle swarm optimation untuk optimasinya. berdasarkan hasil penelitian, nilai akurasi dari algortma support vector machine dan particle swarm optimation rata-rata adalah 90% lalu hasil rating yang dihasilkan lebih tepat dari pada rating yang diberikan secara manual pada google playstore .

Kata Kunci: google playstore, algortima support vector machine, algoritma particle swarm optimation.

ABSTRACT

The large use of smartphones, especially Android in Indonesia, is in line with the emergence of applications to support Indonesian community activities such as Gojek, Grap, Tokopedia, Shopee, and so on. These applications are downloaded on a Google application called Google Playstore. In addition to functioning as application storage, Google Playstore can be used to review applications and provide ratings for these applications. often the reviews and ratings given are not in the same direction because this is the rating given to the application is not biased, so the researchers made a study to make a rating based on the classification of the reviews given by the user, the algorithm used for this classification is the support vectore machine algorithm and the particle swarm algorithm optimization for optimization. based on research results, the accuracy value of the support vector machine and particle swarm optimization algorithms is on average 90% and the resulting rating results are more precise than the rating given manually on google playstore.

Keywords: google playstore, algortima support vector machine, algoritma particle swarm optimation

1. PENDAHULUAN

Sebanyak 66.31% penduduk Indonesia sudah menggunakan smartphone (Kominfo, 2017). Penggunaan smartphone yang besar khususnya smartphone berjenis android ini sejalan dengan munculnya aplikasi-aplikasi untuk membantu kegiatan masyarakat Indonesia seperti aplikasi gojek dan grab yang menyediakan jasa transportasi dan juga aplikasi penyedia marketplace seperti aplikasi tokopedia dan shopee

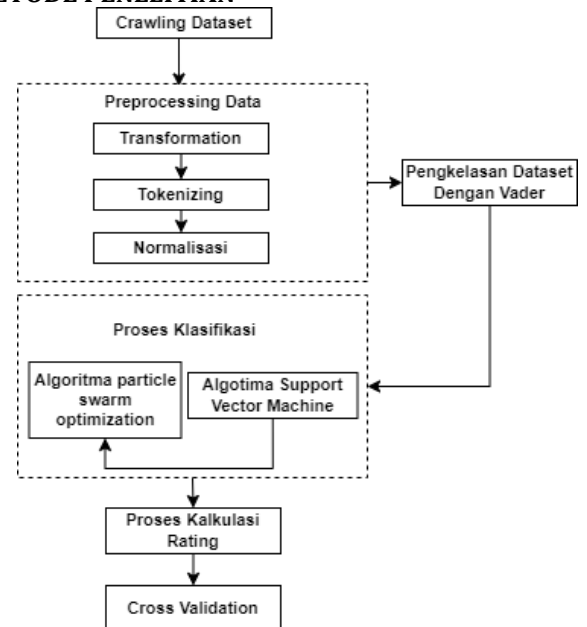
Agar masyarakat dapat mengunduh aplikasi tersebut, para penyedia aplikasi-aplikasi tersebut membagi aplikasi mereka ke google playstore bagi smartphone berbasis android dan ke apple store bagi smartphone berbasis apple. Selain berfungsi sebagai media pengunduhan aplikasi fungsi lain dari google playsote dan juga apple store adalah untuk memberikan rating, dan review bagi aplikasi yang telah diunduh.

Rating dan review yang sudah dibagikan oleh pengguna aplikasi berguna untuk mengetahui kualitas aplikasi apakah aplikasi tersebut baik atau buruk dimata penggunaan aplikasi tersebut. Tetapi terdapat beberapa kekurangan dari penggunaan rating dan review karena seringkali rating yang tinggi tidak mencerminkan aplikasi yang baik karena banyak review diaplikasi tersebut ternyata lebih banyak memberikan kritik dari pada pujian (Sadiq et al., 2021). Untuk mengatasi masalah ini digunakan Analisa setiap review menggunakan algoritma deep learning, hasil Analisa ini akan dijumlahkan sehingga menghasilkan nilai rating yang tidak bias dengan nilai akurasi sebesar 89% untuk algoritma CNN (Sadiq et al., 2021). Penggunaan algoritma lain seperti support vector machine dan Particle Swarm Optimization bisa digunakan untuk menganalisa review terhadap aplikasi di google playstore karena mendapat nilai akurasi yang tinggi sebesar 93% (Mustopa et al., 2020). Selain menggunakan algoritma support vector machine, Analisa review bisa menggunakan algoritma Logistic Regression, SGD Classifier dan K-Neighbor Classifier dengan dikombinasikan dengan metode TF/IDF dan bigrams sehingga hasil Analisa review bisa menggantikan penggunaan bintang yang tidak bisa dianalisa

karena datanya bersifat tunggal (Karim et al., 2020).

Dari beberapa sample penelitian tersebut, peneliti melakukan penelitian untuk mereview komentar pada google playstore kemudian review tersebut akan dikalkulasikan untuk mendapatkan rating aplikasi yang lebih baik. Untuk algoritma klasifikasi review, peneliti menggunakan algoritma support vector machine karena memilih kelebihan untuk menyelesaikan masalah non-linier (Santosa, 2007). Karena support vector machine memiliki kekurangan terhadap seleksi fitur (Kristiyanti, 2015), maka peneliti akan menggunakan algoritma particle swarm optimization. Dan semua prosesnya menggunakan Bahasa pemograman pyhton.

2. METODE PENELITIAN



Pertama, peneliti melakukan pengumpulan dataset mulai dari bulan januari 2022 sampai desember 2022, dataset dikumpulkan dari aplikasi twitter, Instagram, facebook dan tiktok menggunakan teknik crawling, teknik crawling berfungsi untuk mengumpulkan data terbaru dan tercepat (Ramadhan & Sastramihardja, 2018), kemudian dataset yang telah didapatkan akan masuk ke tahap preprocessing data dimana tahapanya dibagi menjadi empat yaitu Transformation berguna untuk merubah semua huruf menjadi huruf kecil (Jaka, 2015), Tokenizing berfungsi untuk menghapus karakter-karakter tertentu pada teks (Jaka, 2015), dan terakhir adalah normalisasi berfungsi untuk

menghilangkan kata tidak baku menjadi baku (Javed & Kamal, 2018). Kemudian masuk ke tahap pengkelasan dataset dengan vader, kelas yang digunakan akan searah dengan rating google playstore dan berikut daftar kelas yang akan digunakan

Tabel 1. Nilai kelas sama dengan rating

kelas	rating
sangat buruk	1
buruk	2
sedang	3
baik	4
sangat baik	5

Setelah dataset berhasil dikelaskan menggunakan vader, maka tahap selanjutnya adalah melakukan klasifikasi menggunakan algoritma support vector machine dengan cara mencari hyperlane terbaik menggunakan persamaan(Darmawan et al., 2022) :

$$f(x) = w^T x + b \tag{1}$$

$$[(w^T) + b] \geq 1 \text{ untuk } y_i = +1 \tag{1}$$

$$[(w^T) + b] \leq -1 \text{ untuk } y_i = -1 \tag{1}$$

Hasil dari algoritma support vector machine tersebut akan dibawa ke dalam algoritma particle swarm optimization, dalam algoritma ini terdapat beberapa proses untuk optimasi, pertama adalah inisialisasi dengan persamaan yang digunakan (Darmawan et al., 2022):

$$x = x_{min} + rand[0,1]x(x_{max} - x_{min}) \tag{2}$$

Kemudian update kecepatan dengan persamaan(Darmawan et al., 2022) :

$$v_{i+j}^{t+1} = w \cdot v_{i+j}^t + c1 \cdot r1 \left(Pbest_{i+j}^t \right) + c2 \cdot r2 \left(Gbest_{g+j}^t - x_{i+j}^t \right) \tag{3}$$

Lalu melakukan update posisi dan menghitung nilai fitness dengan persamaan(Darmawan et al., 2022)

$$x_{i+j}^{t+1} = x_{i+j}^t + v_{i+j}^{t+1} \tag{4}$$

Dan terakhir adalah update IBest dan gBest, nilai fitness yang lebih tinggi akan menjadi Ibest dan gBest terbaru. Tahap selanjutnya adalah proses kalkulasi rating berdasarkan hasil klasifikasi menggunakan persamaan yang peneliti buat sendiri.

$$kalkulasi_rating = \frac{nilai_review}{jumlah_review} \tag{5}$$

Kalkulasi rating ini akan menghasilkan nilai rating berdasarkan kalsifikasi yang sudah dilakukan. Lalu tahap terakhir adalah melakukan

tingkat akurasi algoritma menggunakan cross validation sebanyak 5 Kfold.

3. HASIL DAN PEMBAHASAN

Dataset aplikasi twitter, Instagram, facebook dan tiktok dikumpulkan dengan metode crawling dan hasil pengumplan ditunjukkan pada tabel 2

Tabel 2. Crawling Dataset

aplikasi	Jumlah dataset
twitter	350.121
instagram	478.094
facebook	243.901
tiktok	201.398

Setelah dataset dikumpulkan maka tahap selanjutnya adalah melakukan transform dataset, berikut sampel hasil transform dataset pada salah satu aplikasi yaitu Instagram

Tabel 3. Contoh Transform Dataset

Sebelum transform	Setelah transform
Menurut aku versi sebelumnya"nya lebih baik dari pada versi baru sekarang sekarang"	menurut aku versi sebelumnya"nya lebih baik dari pada versi baru sekarang sekarang" ini semakin di update
ini Semakin di update malah jadi banyak masalah.	malah jadi banyak masalah.
banyak masalah.	

Tahap selanjutnya adalah tokenizing dan hasilnya sebagai berikut

Tabel 4. Contoh tokenizing Dataset

Sebelum tokenizing	Setelah tokenizing
menurut aku versi sebelumnya"nya lebih baik dari pada versi baru sekarang sekarang"	menurut aku versi sebelumnya"nya lebih baik dari pada versi baru sekarang sekarang" ini semakin di update
ini semakin di update malah jadi banyak masalah.	malah jadi banyak masalah
banyak masalah.	

Terakhir adalah melakukan normalisasi dengan hasil sebagai berikut

Tabel 5. Contoh Normalisasi Dataset

Sebelum normalisasi	Setelah normalisasi
menurut aku versi sebelumnya lebih baik dari pada versi baru sekarang sekarang ini semakin di update malah jadi banyak masalah	menurut aku versi sebelum lebih baik dari pada versi baru sekarang sekarang ini semakin di update malah jadi banyak masalah

Dataset yang sudah selsai di tahap preprocessing selanjutnya masuk ketapah pengkelasan dataset dengan vader, hasil pengkelasan ini membuat dataset terbigi menjadi 5 kelas seperti yang sudah dijelaskan.

Tabel 6. Hasil Pengkelasan Dataset dengan Vader

aplikasi	Jumlah dataset	Kelas dataset
twitter	350.121	Sangat buruk = 12.983 Buruk = 34.109 Sedang = 64.129 Baik = 137.802 Sangat baik = 101.098
instagram	478.094	Sangat buruk = 45.377 Buruk = 23.375 Sedang = 78.891 Baik = 120.232 Sangat baik = 210.219
facebook	243.901	Sangat buruk = 12.891 Buruk = 45.855 Sedang = 20.443 Baik = 121.452 Sangat baik = 43.248
tiktok	201.398	Sangat buruk = 15.015

Buruk =19.405
Sedang = 30.115
Baik = 43.313
Sangat baik = 93.550

Setelah tahap pengkelasan dataset selsai,maka selanjutnya adalah melakukan klasifikasi menggunakan support vector machine dan partice swarm optimation, hasil dari klasifikasi tersebut kemudian akan dikalkulasikan dengan persamaan yang telah peneliti jelaskan sehingga hasil kalkulasi sebagai berikut :

Tabel 7. Hasil Kalkulasi

Aplikasi	Rating hasil klasifikasi
twitter	4.1
instagram	4.3
facebook	4.1
tiktok	4.4

Selanjutnya pada tahap cross validation, nilai akurasi yang didapat pada algoritma support vector machine dan partice swarm optimation cukup tinggi. Berikut nilai akurasi seperti di tunjukkan pada tabel 8

Tabel 8. Hasil Cros Validation

Aplikasi	Cross validation
twitter	Kfold 1 = 93.43%, kalkulasi rating = 4.2
	Kfold 2 = 90.12%, kalkulasi rating = 4.1
	Kfold 3 = 92.97%, kalkulasi rating = 4.2
	Kfold 4 = 95.21%, kalkulasi rating = 4.2
	Kfold 5 = 94.43%, kalkulasi rating = 4.2
instagram	Kfold 1 = 90.07%, kalkulasi rating = 4.1
	Kfold 2 = 91.30%, kalkulasi rating = 4.1
	Kfold 3 = 92.65%, kalkulasi rating = 4.2
	Kfold 4 = 96.20%, kalkulasi rating = 4.3

facebook	Kfold 5 = 95.12%, kalkulasi rating = 4.3
	Kfold 1 = 89.76%, kalkulasi rating = 4.0
	Kfold 2 = 90.42%, kalkulasi rating = 4.0
	Kfold 3 = 92.97%, kalkulasi rating = 4.0
	Kfold 4 = 97.12%, kalkulasi rating = 4.1
tiktok	Kfold 5 = 93.33%, kalkulasi rating = 4.1
	Kfold 1 = 93.43%, kalkulasi rating = 4.5
	Kfold 2 = 90.12%, kalkulasi rating = 4.5
	Kfold 3 = 92.97%, kalkulasi rating = 4.4
	Kfold 4 = 95.21%, kalkulasi rating = 4.5
	Kfold 5 = 94.43%, kalkulasi rating = 4.5

Berdasarkan hasil cross validation tersebut terlihat penggunaan support vector machine dan particle swarm optimation mendapat nilai akurasi rata-rata 90% di semua aplikasi dengan rating hasil klasifikasi yang di hasilkan rata-rata tidak jauh berbeda dari rating yang terdapat pada google playstore dan berikut perbedaan rating google playstore dan rata-rata rating klasifikasi

Tabel 9. Hasil rating

Aplikasi	Rating di Google Playstore	Rating Hasil Klasifikasi
twitter	4.0	4.2
instagram	4.3	4.2
facebook	3.9	4.0
tiktok	4.4	4.4

Berdasarkan tabel 9, terdapat perbedaan jumlah rating antara google playstore dengan rating yang dibuat berdasarkan kalsifikasi review, perbedaan ini terjadi karena user pada aplikasi tersebut memberikan rating dan review yang bertolak belakang seperti ditunjukkan pada tabel 10

Tabel 10. Contoh review dari aplikasi facebook

Review	Pemberian Rating
Aplikasi facebook makin bagus ya, tidak lemot seperti 1 atau 2 tahun lalu	2

Rating yang diberikan oleh user pada aplikasi adalah 2, sedangkan hasil pemerian rating berdasarkan review di kami adalah sebagai berikut

Tabel 11. Contoh review dari aplikasi facebook

Review	Pemberian Rating
Alikasi facebook makin bagus ya, tidak lemot seperti 1 atau 2 tahun lalu	4

Pemberian rating 4 pada kalsifikasi ini dikarenakan jenis kelas kalsifikasi pada review ini adalah bagus sehingga otomatis akan diberikan rating 4. Hal inilah yang menyebabkan rating yang dibuat menggunakan kalsifikasi review dengan goole playstore berbeda dan hasil rating yang dibuat berdasarkan review lebih baik karena hasil yang diberikan jauh lebih tepat.

4. KESIMPULAN

Berdasarkan hasil penelitian, nilai akurasi dari algoritma support vector machine cukup tinggi rata-rata berkisar 90% untuk semua aplikasi. Lalu pembuatan rating berdasarkan review jauh lebih tepat dari pada rating yang diberikan secara manual.

DAFTAR PUSTAKA

- Darmawan, R., Indra, I., & Surahmat, A. (2022). Optimalisasi Support Vector Machine (SVM) Berbasis Particle Swarm Optimization (PSO) Pada Analisis Sentimen Terhadap Official Account Ruang Guru di Twitter. *Jurnal Kajian Ilmiah*, 22(2), 143–152. <https://doi.org/10.31599/jki.v22i2.1130>
- Jaka, A. T. (2015). Preprocessing Text untuk Meminimalisir Kata yang Tidak Berarti dalam Proses Text Mining. *Informatika UPGRIS*, 1, 1–9.
- Javed, M., & Kamal, S. (2018). Normalization of unstructured and informal text in

- sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 9(10), 78–85.
<https://doi.org/10.14569/IJACSA.2018.091011>
- Karim, A., Azhari, A., Aldabbas, H., Alruily, M., Belhaouri, S. B., & Qureshi, A. A. (2020). Classification of Google Play Store Application Reviews Using Machine Learning. *Preprints, July*, 1–22.
<https://doi.org/10.20944/preprints202003.0231.v1>
- Kominfo. (2017). Survey Penggunaan TIK Serta Implikasinya Terhadap Sosial Budaya Masyarakat. *Badan Penelitian Dan Pengembangan Sumber Daya Manusia*, 1–30.
- Kristiyanti, D. A. (2015). Analisis Sentimen Review Produk Kosmetik Menggunakan Algoritma Support Vector Machine Dan Particle Swarm Optimization Sebagai. *Seminar Nasional Inovasi & Tren (SNIT) 2015 "Peluang Dan Tantangan Indonesia Dalam Menyikapi Afta 2015,"* 134–141.
http://lppm.bsi.ac.id/SNIT2015/BidangA/A22-134-141_2015-SNIT-Dinar Ajeng Kristiyanti_ALGORITMA SUPPORT VECTOR.pdf
- Mustopa, A., Hermanto, Anna, Pratama, E. B., Hendini, A., & Risdiansyah, D. (2020). Analysis of user reviews for the pedulilindungi application on google play using the support vector machine and naive bayes algorithm based on particle swarm optimization. *2020 5th International Conference on Informatics and Computing, ICIC 2020, 2*.
<https://doi.org/10.1109/ICIC50835.2020.9288655>
- Ramadhan, I., & Sastramihardja, H. (2018). Pemanfaatan Web Crawler Dalam Mengumpulkan Informasi Melalui Internet. *Konferensi Nasional Sistem Informasi 2018*, 8–9.
- Sadiq, S., Umer, M., Ullah, S., Mirjalili, S., Rupapara, V., & Nappi, M. (2021). Discrepancy detection between actual user reviews and numeric ratings of Google App store using deep learning. *Expert Systems with Applications*, 181, 115111.
<https://doi.org/10.1016/j.eswa.2021.115111>
- Santosa, B. 2007. Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. Graha Ilmu : Yogyakarta.
-