



Support Vector Machine Algorithm Optimization Using Particle Swarm Optimization for Prosperous Family Card Recipients

Optimasi Algoritma Support Vector Machine Menggunakan Particle Swarm Optimization untuk Penerima Kartu Keluarga Sejahtera

Afi Ghuran Yuda^{1*}, Angil Mulistu Karomah², Dany Tria Putra Ramadhan³,
Syarifah Intan Putri⁴, Umairah Rizky Gurning⁵

^{1,2,3,4,5}Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

⁵Puzzle Research Data Technology, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Sultan Syarif Kasim, Riau, Indonesia

E-Mail: ¹ghufrana3@gmail.com, ²angilmulistukaromah@gmail.com, ³kzero504@gmail.com,
⁴syarifahintanputri@gmail.com, ⁵11950320687@students.uin-suska.ac.id

Corresponding Author: Afi Ghuran Yuda

Abstract

Recipients of the Prosperity Card (KKS) are set quotas for the provision of assistance; however, many families submit incorrect data, making it impossible for officials to process the incomplete data. Social assistance has not achieved its goals for a very long time, despite potential growth. Support vector machine is the most commonly used classification method to categorize predictions about recipients of the Prosperous Family Card. Other techniques, such as Particle Swarm Optimization, can be used to optimize the SVM algorithm to improve accuracy. measurements based on a holdout approach to data sharing and support vector machine accuracy. Based on the findings, the accuracy of the support vector machine increased on average from 89.58% to 92.26%.

Keyword: KKS, Optimization, Prediction, PSO, SVM.

Abstrak

Penerima Kartu Sejahtera Sejahtera (KKS) ditetapkan kuota pemberian bantuan; Namun, banyak keluarga yang menyampaikan data yang tidak benar sehingga membuat aparat tidak bisa mengolah data yang tidak lengkap tersebut. Bantuan sosial belum mencapai tujuannya untuk waktu yang sangat lama, meski ada potensi pertumbuhan. Support Vector Machine adalah algoritma klasifikasi yang paling umum digunakan untuk mengklasifikasikan prediksi tentang penerima Kartu Keluarga Sejahtera. Teknik lain, seperti Particle Swarm Optimization, dapat digunakan untuk mengoptimalkan algoritma SVM guna meningkatkan akurasi. pengukuran berdasarkan pendekatan hold out untuk berbagi data, dan ketepatan algoritma SVM. Berdasarkan hasil, akurasi SVM meningkat rata-rata dari 89,58% menjadi 92,26%.

Kata Kunci: KKS, Optimasi, Prediksi, PSO, SVM.

1. PENDAHULUAN

Kemiskinan adalah rendahnya finansial seseorang yang disebabkan oleh faktor internal dan eksternal [1]. Kartu Keluarga Sejahtera (KKS) adalah bantuan yang diberikan pemerintah kepada masyarakat melalui dinas sosial [2]. KKS adalah salah satu program pemerintah untuk mempercepat penanggulangan kemiskinan yang terdapat dalam Peraturan Presiden (Perpres) No. 166 Tahun 2014 tentang penanggulangan kemiskinan [3]. Berdasarkan Basis Data Terpadu (BDT) yang diolah oleh Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2KP), penerima bantuan KKS mencapai 25% dari jumlah penduduk yang ada dengan status sosial ekonomi terendah dengan total bantuan yang diberikan sebesar Rp.200.000/keluarga/bulan [3].

Pemerintah daerah dalam melakukan pengetasan kemiskinan melalui program KKS harus didukung dengan data yang akurat sehingga dapat mengefisienkan waktu pengolahan data [4]. Permasalahan utama yang dihadapi adalah setiap Kelurahan diberikan kuota tertentu untuk mengirimkan data penerima bantuan sosial, dan banyak keluarga yang menyerahkan data yang tidak memenuhi persyaratan penerima, sehingga memperbesar kemungkinan untuk tidak tepat sasaran ataupun keterlambatan pengiriman dalam memberikan bantuan Kartu Keluarga Sejahtera (KKS), dan perlu diadakan seleksi yang lebih mendalam terhadap data-data yang diterima. Dumai adalah salah satu kota di Provinsi Riau yang memiliki penduduk lebih dari 300.000 jiwa [5]. Jumlah Data Terpadu Kesejahteraan Sosial kota Dumai per Januari 2020 SK No. 19/HUK/2020 Per Ruta/Januari 2020: 18.793, Per KK/Januari 2020: 19.321, dan Per Jiwa/Januari 2020: 78.033 [6].

Kesulitan yang ditemui pada pengenalan data penerima KKS terjadi karena belum adanya pengolahan data yang sesuai untuk mengolah data calon penerima bantuan yang cukup besar oleh tiap Kelurahan di Kota Dumai sehingga Instansi terkait masih mengalami kesulitan pada pengolahan data. Karena itu, dibutuhkan klasifikasi dari data penerima KKS dengan menggunakan metode PSO-SVM [7]. Dalam masalah tersebut maka diperlukan suatu metode yaitu data mining tepatnya klasifikasi untuk memprediksi penerima KKS [5]. Algoritma klasifikasi yang umum digunakan untuk memprediksi bantuan Kartu Keluarga Sejahtera diantaranya Naïve Bayes, Support Vector Machine (SVM) dan k-Nearest Neighbor (kNN) [8]. Dari beberapa algoritma yang umum digunakan untuk mengklasifikasi data adalah SVM [9].

SVM adalah algoritma supervised learning yang berguna untuk mengklasifikasi data. Algoritma ini adalah kasus khusus dari algoritma yang dikenal sebagai pengklasifikasi yang cocok dan mitigasi risiko yang kuat. [10]. Algoritma SVM dapat dioptimasi untuk meningkatkan akurasi dengan menggunakan algoritma lain seperti Particle Swarm Optimization (PSO) [11]. PSO adalah metode optimasi yang dapat digunakan untuk meningkatkan akurasi (Jing Wang, 2014) sehingga pada penelitian ini dilakukan percobaan dengan metode PSO-SVM untuk mendeteksi penerima bantuan KKS [12]. Sudah banyak penelitian yang menggunakan algoritma SVM. Emilia Ayu pada tahun 2021 melakukan penelitian tentang klasifikasi jenis padi menggunakan algoritma SVM berbasis PSO dan memiliki hasil akurasi sebesar 70,83% pada penelitian sebelumnya [12]. Firmansyah pada tahun 2019 juga menekankan temuan akurasi yang lebih baik dari teknik Support Vector Machine (SVM) sebesar 95% dalam perbandingan Klasifikasi SVM dan Decision Tree untuk Pemetaan Mangrove Berbasis Objek Menggunakan Citra Satelit Sentinel-2B di Gili Sulat, Lombok Timur [13]. Amalia pada tahun 2017 menyelesaikan studi penelitiannya, menggunakan algoritma SVM berbasis PSO ditemukan sebagai teknik klasifikasi data terbaik, dengan nilai akurasi 99,6% dalam mengidentifikasi kebangkrutan perusahaan [14].

Pada penelitian ini dilakukan perbandingan algoritma SVM dan algoritma SVM dengan PSO. Untuk membandingkan nilai akurasi algoritma dari data penerima KKS. Algoritma SVM digunakan karena algoritma ini dapat mengelola klasifikasi dua kelas. Namun, juga memiliki kelemahan dalam menangani data yang besar, sehingga harus dioptimalkan untuk meningkatkan akurasi hasil yaitu dengan menggunakan PSO [15]. Kombinasi kedua metode ini sangat baik digunakan dalam klasifikasi data karena dapat meningkatkan tingkat akurasi ke arah yang lebih baik [16]. Oleh karena itu, penelitian ini dapat membantu instansi untuk memprediksi dan mengolah data penerima KKS menggunakan algoritma terbaik yang paling sesuai dengan data yang tersedia.

2. METODOLOGI PENELITIAN

Penelitian ini akan menggunakan prosedur pada Gambar 1. Support Vector Machine, sebuah teknik klasifikasi yang dibangun berdasarkan prinsip-prinsip Particle Swarm Optimization, digunakan dalam penelitian ini. Pendekatan pembagian data Hold Out adalah teknik klasifikasi yang diterapkan dalam penelitian ini.

2.1 Hold-out

Hold-out adalah teknik pemecahan data langsung yang memisahkan data menjadi set pelatihan dan pengujian. Label kelas untuk dataset yang dibagi telah ditentukan. Teknik hold-out menggunakan dua bagian: satu untuk melatih classifier dan satu lagi untuk mengujinya (Said dan Mustakim, 2020).

2.2 Support Vector Machine (SVM)

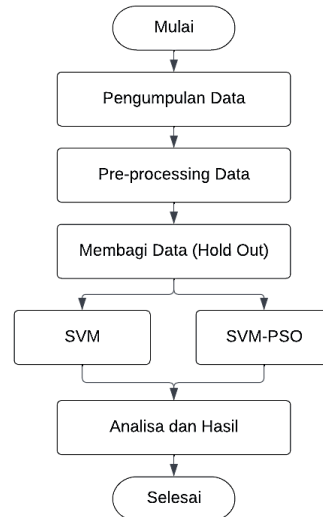
Algoritma yang umum dan banyak di aplikasikan adalah Support Vector Machine (SVM). Pengaturan parameter inti dari prosedur pelatihan SVM bersama dengan fitur yang dipilih berdampak signifikan pada akurasi klasifikasi.

Dengan menggunakan rumus Lagrangian, hyperplane margin maksimal dapat ditulis ulang sebagai batas keputusan untuk klasifikasi tes atau tupel baru seperti yang diberikan pada persamaan 1.

$$d(x^T) = \sum_{i=1}^1 y_i a_i X^T + b_0 \quad (1)$$

Keterangan:

- y_i : label kelas dari support vector machine x_i
 x^T : tupel uji
 α_i : pengganda lagrangian
 b_0 : parameter numerik
 i : jumlah vektor dukungan



Gambar 1. Metodologi Penelitian

2.3 Particle Swarm Optimization (PSO)

Metode pengoptimalan yang sangat mudah untuk mengimplementasikan dan mengubah berbagai parameter disebut Particle Swarm Optimization (PSO).

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Informasi dan Pre-processing Data

Analisis ini menggunakan Data Kesejahteraan Sosial Terpadu tahun 2020 dari Open Data Kota Dumai. Tiga tahap pra-pemrosesan data adalah pembersihan, transformasi, dan normalisasi. Normalisasi data digunakan sebagai bagian dari preprocessing data yang mengarah ke Tabel 1.

Tabel 1. Normalisasi Data Terpadu Kesejahteraan Sosial

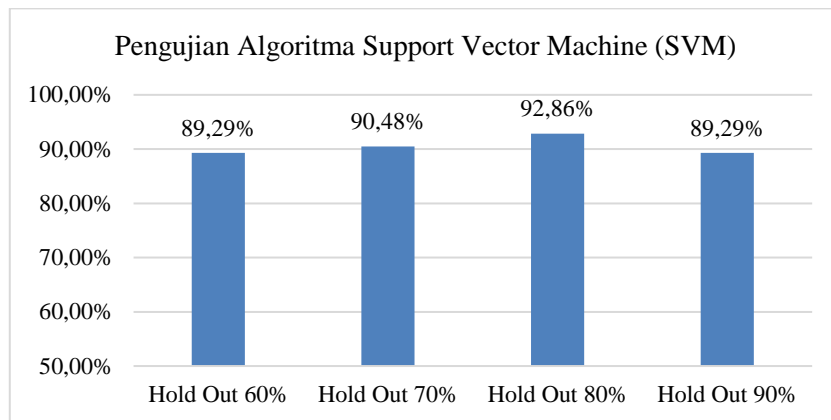
No.	Jumlah ART	STA Bangunan	..	Kesejahteraan
1	0.83	0.43	..	Ya
2	0.25	0.43	..	Ya
3	0.92	0.51	..	Ya
4	0.25	0.51	..	Ya
5	0.25	0.43	..	Ya
6	0.25	0.43	..	Tidak
7	2.00	0.43	..	Tidak
8	2.00	0.51	..	Ya
..
70	1.50	0.51	..	Ya
71	1.50	0.51	..	Ya

Atribut lengkap pada tabel 1:

- | | | | |
|---------------------|------------------------|-----------------------|--------------------------|
| 1. jumlah_art; | 11. sumber_air minum; | 21. ada_pemanas ; | 31. ada_perahu_motor; |
| 2. sta_bangunan; | 12. sumber_penerangan; | 22. ada_telepon; | 32. ada_kapal; |
| 3. sta_lahan; | 13. daya; | 23. ada_tv; | 33. aset_tak_bergerak; |
| 4. luas_lantai; | 14. bb_masak; | 24. ada_emas; | 34. rumah_lain; |
| 5. lantai; | 15. fas_bab; | 25. ada_laptop; | 35. sta_art_usaha; |
| 6. dinding; | 16. kloset; | 26. ada_sepeda; | 36. percentile; |
| 7. kondisi_dinding; | 17. buang_tinja; | 27. ada_motor; | 37. status_kesejahteraan |
| 8. atap; | 18. ada_tabung_gas; | 28. ada_mobil; | |
| 9. kondisi_atap; | 19. ada_lemari_es; | 29. ada_perahu; | |
| 10. jumlah_kamar; | 20. ada_ac; | 30. ada_motor_tempel; | |

3.2 Support Vector Machine

Pada penelitian ini akan dilakukan proses klasifikasi dengan metode Support Vector Machine (SVM) dan pembagian data dengan metode Hold Out. Dibawah ini adalah diagram pengujian algoritma pada Support Vector Machine pada Gambar 2.

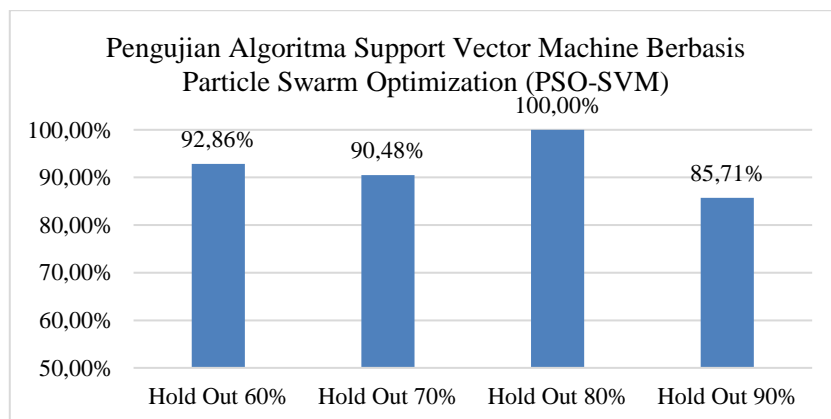


Gambar 2. Akurasi Pengujian Support Vector Machine

Setelah menambahkan rata-rata dari setiap Hold Out, akurasi rata-rata algoritma Support Vector Machine (SVM) adalah 89.58%.

3.3 Support Vector Machine Berbasis Particle Swarm Optimization (PSO-SVM)

Pada penelitian ini proses klasifikasi akan dilakukan dengan metode Support Vector Machine (SVM) dan pembagian data dengan metode Hold Out. Dibawah ini adalah diagram pengujian algoritma pada Support Vector Machine pada Gambar 3.

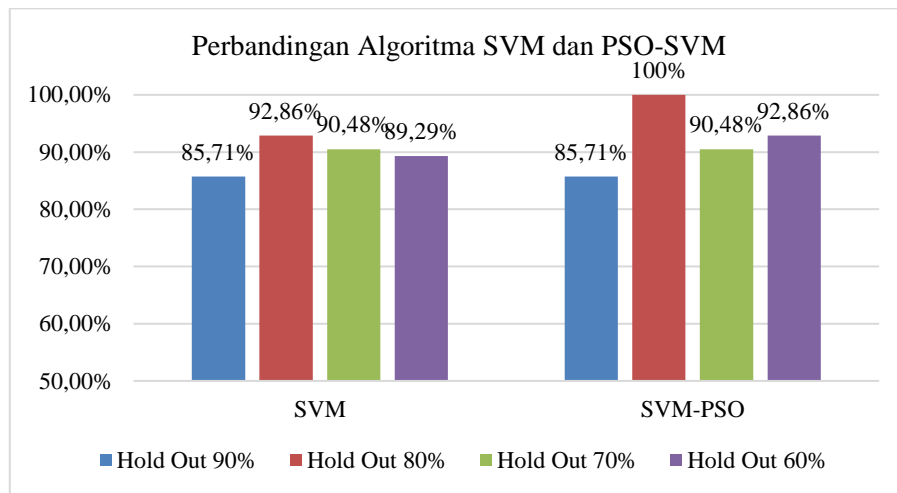


Gambar 3. Menguji algoritma Support Vector Machine berbasis Particle Swarm Optimization (PSO-SVM)

Akurasi rata-rata dari Algoritma Support Vector Machine Berbasis Particle Swarm Optimization (PSO-SVM) adalah 92.26% setelah menambahkan rata-rata dari setiap Hold Out.

3.4 Perbandingan SVM dengan SVM-PSO

Metode Support Vector Machine (SVM) dan SVM Berbasis Particle Swarm Optimization (PSO-SVM) yang digunakan pada Gambar 4 dalam algoritma Data Mining dibandingkan di bawah ini.



Gambar 4. Perbandingan Algoritma SVM dan PSO-SVM

Gambar 4 menunjukkan bagaimana pendekatan Support Vector Machine menggunakan DTKS, dan berdasarkan Particle Swarm Optimization (PSO) menunjukkan peningkatan akurasi.

4. KESIMPULAN

Sebagai hasil dari penelitian ini, pengujian dengan menggunakan metode Support Vector Machine (SVM) mampu menghasilkan pengujian dengan akurasi rata-rata sebesar 89,58%, sedangkan pengujian dengan menggunakan algoritma Support Vector Machine Based Particle Swarm Optimization (PSO-SVM) dapat menghasilkan pengujian dengan akurasi rata-rata sebesar 92,26%. Teknik Support Vector Machine Based Particle Swarm Optimization (PSO-SVM) adalah sistem klasifikasi yang paling akurat untuk Prediksi Penerima Kartu Keluarga Sejahtera dengan Menggunakan Data Terpadu Kesejahteraan Sosial, menurut temuan penelitian.

REFERENSI

- [1] T. Qurahman, I. Syukra, and U. R. Gurning, "Implementation of Technique for Order Performance by Similarity to Ideal Solution (TOPSIS) Method in Selection of Cayenne Pepper Seeds Implementasi Metode Technique for Order Performance by Similarity to Ideal Solution (TOPSIS) dalam Pemilihan Bibit C," vol. 1, no. October, pp. 85–94, 2021.
- [2] U. R. Gurning and Mustakim, "Penerapan Algoritma K-Means dan K-Medoid untuk Pengelompokkan Data Pasien Covid-19," *Build. Informatics, Technol. Sci.*, vol. 3, no. 1, p. 48–55, 2021, doi: 10.47065/bits.v3i1.1003.
- [3] H. Putri, A. I. Purnamasari, A. R. Dikananda, O. Nurdiawan, and S. Anwar, "Penerima Manfaat Bantuan Non Tunai Kartu Keluarga Sejahtera Menggunakan Metode NAÏVE BAYES dan KNN," *Build. Informatics, Technol. Sci.*, vol. 3, no. 3, pp. 331–337, 2021, doi: 10.47065/bits.v3i3.1093.
- [4] W. R. Sari Oktapia Ningse, S. Sumarno, and Z. M. Nasution, "C4.5 Algorithm Classification for Determining Smart Indonesia Program Recipients at MIS Al-Khoirot," *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 1, pp. 65–76, 2022, doi: 10.55123/jomlai.v1i1.165.
- [5] A. Prajoko, R. W. Sembiring, and S. S, "Penerapan Algoritma C4.5 Dalam Klasifikasi Penerima Kartu Keluarga Sejahtera (KKS)," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.*, vol. 6, no. 1, p. 171, 2021, doi: 10.30645/jurasik.v6i1.281.
- [6] K. Kunci, "Klasifikasi, Machine Learning, Naïve Bayes, Random Forest, PIP.," vol. 10, no. 2, 2022.
- [7] Bidara Pink, "Data Baru BPS : Jumlah Penduduk Miskin di Indonesia Berkurang," *J. Sos. Hum. Terap.*, vol. 1, no. 1, pp. 1–6, 2018.
- [8] M. Yaqoob, F. Iqbal, and S. Zahir, "Comparing predictive performance of k-nearest neighbors and support vector machine for predicting ischemic heart disease," ... *J. Adv. Sci.*, vol. 1, no. 2, pp. 49–60, 2020, [Online]. Available: <https://www.royalliteglobal.com/rjas/article/view/391%0Ahttps://www.royalliteglobal.com/rjas/article/download/391/177>
- [9] A. Nurhadi, "Klasifikasi Konten Berita Digital Bahasa Indonesia Menggunakan Support Vector Machines (SVM) Berbasis Particle Swarm Optimization (PSO)," *J. Bianglala Inform.*, vol. 3, no. 2, pp. 1–9, 2015.
- [10] Ita Rosita, "Efektivitas Program Pemberdayaan Usaha Mikro Kecil Dan Menengah (UMKM) Di Kecamatan Sungai Sembilan Kota Dumai," *J. Niara*, vol. 14, no. 3, pp. 259–265, 2021, doi:

- 10.31849/niara.v14i3.8020.
- [11] E. Pudjiarti, "Prediksi Spam Email Menggunakan Metode Support Vector Machine Dan Particle Swarm Optimization," *J. Pilar Nusa Mandiri*, vol. 12, no. 2, pp. 171–181, 2016, [Online]. Available: <http://ejournal.nusamandiri.ac.id/index.php/pilar/article/view/271>
- [12] D. A. Pisner and D. M. Schnyer, "Support vector machine," *Mach. Learn. Methods Appl. to Brain Disord.*, pp. 101–121, 2019, doi: 10.1016/B978-0-12-815739-8.00006-7.
- [13] L. Timur, "Perbandingan Klasifikasi SVM dan Decision Tree untuk Pemetaan Mangrove Berbasis Objek Menggunakan Citra Satelit Sentinel-2B di Gili Sulat, Lombok Timur," *J. Pengelolaan Sumberd. Alam dan Lingkung. (Journal Nat. Resour. Environ. Manag.*, vol. 9, no. 3, pp. 746–757, 2019, doi: 10.29244/jpsl.9.3.746-757.
- [14] H. Amalia, A. Fitria Lestari, and A. Puspita, "Penerapan Metode Svm Berbasis Pso Untuk Penentuan Kebangkrutan Perusahaan," *J. Techno Nusa Mandiri*, vol. 14, no. 2, pp. 131–136, 2017, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Qualitativ>
- [15] I. CHOFYAN, U. RUSTAN, A. HARIYANTO, and ISMAYANTI, *Laporan Akhir Penelitian Dosen Utama Upaya Mempertahankan Kabupaten Karawang Sebagai Lumbung Padi Nasional*, no. 0413097101. 2015. [Online]. Available: http://103.78.195.33/bitstream/handle/123456789/12083/fulltext_chofyan_penelitian_lppm_unisba_2015.PDF?sequence=1&isAllowed=y
- [16] Putri, E. Ucha, Sanni Irawan, and F. Rizky, "Implementasi Data Mining Untuk Prediksi Penjualan Pestisida Pada CV MITRA ARTHA SEJATI Menggunakan Algoritma Naive Bayes," *KESATRIA(J. penerapan Sist. Inf. dan Manaj.*, vol. 2, no. 1, pp. 39–46, 2021, [Online]. Available: <http://eprosiding.ars.ac.id/index.php/pti>
- [17] H. Syukron, M. F. Fayyad, and F. J. Fauzan, "Comparison K-Means K-Medoids and Fuzzy C-Means for Clustering Customer Data with LRFM Model " Perbandingan K-Means K-Medoids dan Fuzzy C-Means untuk Pengelompokan Data Pelanggan dengan Model LRFM," vol. 2, no. October, pp. 76–83, 2022.
- [18] U. R. Gurning, A. P. Pristiawati, and A. Dina, "Pengelompokan Data Loyalitas Pelanggan Model RFM pada Produk Ms Glow Dan Scarlett dengan Algoritma Fuzzy C-Means," pp. 29–34, 2022.
- [19] A. Harun and D. P. Ananda, "Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve Bayes dan Decision Tree," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 58–63, 2021.
- [20] J. Homepage, K. Pratama Simanjuntak, and U. Khaira, "MALCOM: Indonesian Journal of Machine Learning and Computer Science Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm Pengelompokkan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clusterin," *Malcom*, vol. 1, no. 1, pp. 7–16, 2021.
- [21] D. Musfiroh, U. Khaira, P. E. P. Utomo, and T. Suratno, "Analisis Sentimen terhadap Perkuliahan Daring di Indonesia dari Twitter Dataset Menggunakan InSet Lexicon: Sentiment Analysis of Online Lectures in Indonesia from Twitter Dataset Using InSet Lexicon," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 24–33, 2021.