

PENERAPAN DATA MINING UNTUK MENGLASIFIKASI PENERIMA BANTUAN PROGRAM KELUARGA HARAPAN (PKH) MENGUNAKAN METODE SUPPORT VECTOR MACHINE

Amelia Jesika¹⁾, Alexius Endy Budianto²⁾, Danang Aditya Nugraha³⁾

Universitas PGRI Kanjuruhan Malang^{1,2,3)}

email : ameliajesika167@gmail.com

Abstrak

Program Keluarga Harapan (PKH) adalah salah satu langkah strategis pemerintah Indonesia untuk mengurangi kemiskinan dan meningkatkan kesejahteraan keluarga. Namun, proses penentuan keluarga yang berhak menerima bantuan seringkali menghadapi berbagai kendala. Guna mengatasi masalah tersebut, penelitian ini mengaplikasikan teknik data mining dengan metode SUPPORT VECTOR MACHINE (SVM) untuk mengklasifikasi calon penerima PKH di wilayah Kelurahan Bangka Leleng. Studi ini memanfaatkan 1.039 data penerima bantuan periode 2019-2023 dengan lima atribut utama: usia, penghasilan, jumlah tanggungan, pekerjaan, dan status kepemilikan rumah. Pengolahan data dilakukan di lingkungan Google Colab menggunakan bahasa pemrograman Python. Tahapan riset meliputi pengumpulan data, pra-pemrosesan data, pembagian data untuk pelatihan dan pengujian, analisis, serta evaluasi hasil menggunakan Confusion Matrix. Hasil pengujian menunjukkan bahwa metode SVM sangat efektif dalam mengklasifikasi penerima PKH, dengan tingkat akurasi mencapai 96%. Akurasi tertinggi ini dicapai berkat penggunaan kernel RBF, yang performanya terbukti paling baik dibandingkan kernel lainnya. Diharapkan, penelitian ini dapat menjadi referensi untuk menciptakan metode penentuan penerima bantuan yang lebih efisien dan transparan, sehingga penyaluran bantuan dapat dilakukan secara lebih tepat sasaran.

Kata Kunci : Program Keluarga Harapan (PKH); Data Mining; Support Vector Machine (SVM); Klasifikasi; Penerima Bantuan

Abstract

The Family Hope Program (PKH) is a governmental initiative in Indonesia designed to decrease poverty and improve the welfare of families. However, the process of identifying eligible families frequently encounters difficulties. To address this, the study applies data mining techniques with the Support Vector Machine (SVM) method to classify prospective PKH recipients in Bangka Leleng Village. The research utilizes 1,039 data samples of recipients from 2019 to 2023, based on five key attributes: age, income, number of dependents, occupation, and home ownership status. Data processing was conducted using Python in the Google Colab environment. The research workflow involved data collection, preprocessing, splitting data for training and testing, analysis, and evaluation using a Confusion Matrix. The test results indicated that the SVM method is highly effective in classifying PKH recipients, achieving an accuracy rate of up to 96%. This optimal accuracy was obtained by employing the RBF kernel, which demonstrated superior performance compared to other kernels. It is anticipated that this research will provide a more efficient and transparent method for determining aid recipients, leading to a more precise distribution of assistance.

Keywords : Family Hope Program (PKH); Data Mining; Support Vector Machine (SVM); Classification; Aid Recipients

1. PENDAHULUAN

Program Keluarga Harapan (PKH) merupakan salah satu upaya pemerintah untuk mengatasi kemiskinan dengan menyalurkan dukungan finansial. Meskipun demikian, proses penentuan keluarga yang layak menerima bantuan seringkali tidak berjalan optimal. Persoalan ini menjadi perhatian khusus di Kelurahan Bangka Leleng, sehingga diperlukan metode yang lebih efektif dan terukur untuk memastikan bantuan sampai kepada pihak yang tepat.

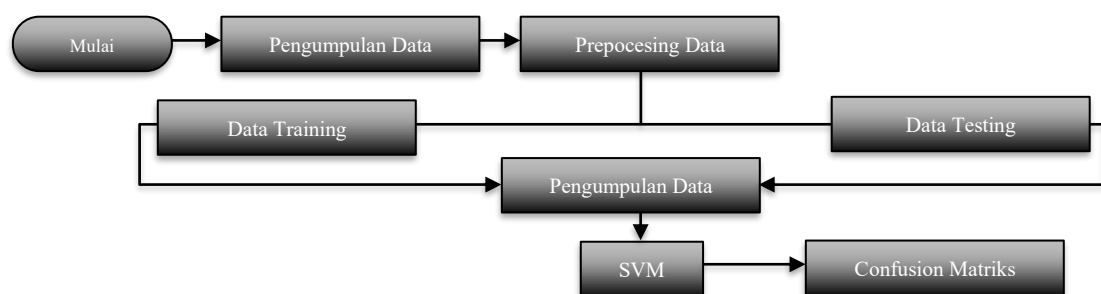
Untuk menyelesaikan masalah ini, penelitian ini mengimplementasikan teknik data mining dengan algoritma *Support Vector Machine* (SVM). Penggunaan data mining dipilih karena kemampuannya dalam mengekstrak wawasan bermanfaat dari data dalam jumlah besar. Sementara itu, SVM dianggap sebagai metode klasifikasi yang sangat efektif dan andal. SVM mampu menangani data yang kompleks dan tidak linear, memiliki akurasi tinggi, serta mahir dalam mengelola ketidakseimbangan kelas pada dataset. Implementasi SVM untuk menyeleksi penerima PKH, diharapkan tingkat ketepatan dan objektivitas dalam klasifikasi dapat meningkat. Melalui pemanfaatan data yang terstruktur seperti kondisi ekonomi, jumlah anggota keluarga, dan faktor terkait lainnya, SVM dapat meminimalkan kesalahan klasifikasi, mencegah penyalahgunaan program, dan menjamin bantuan disalurkan kepada mereka yang benar-benar membutuhkan. Oleh karena itu, pendekatan ini diharapkan dapat mewujudkan proses pengelolaan program PKH yang lebih efisien dan transparan di Kelurahan Bangka Leleng.

2. METODE / ALGORITMA

Dalam

Penelitian ini menitikberatkan pada implementasi teknik data mining untuk mengklasifikasikan penerima bantuan PKH di Kelurahan Bangka Leleng, dengan menggunakan algoritma *Support Vector Machine* (SVM). Langkah-langkah penelitian diuraikan secara kualitatif sebagai berikut:

- a) **Pengumpulan Data:** Penelitian ini mengumpulkan data dari 1.039 penerima bantuan PKH di Kelurahan Bangka Leleng pada periode 2019-2023. Data ini mencakup lima atribut pokok: usia, penghasilan, jumlah tanggungan, pekerjaan, dan status kepemilikan rumah.
- b) **Preprocessing Data:** Tahap ini bertujuan untuk membersihkan dan menyiapkan data. Prosesnya mencakup penanganan nilai yang hilang, penghapusan duplikasi data, dan normalisasi.
- c) **Pembagian Data:** Data yang sudah diproses dibagi menjadi dua set: data latih (training) untuk membangun model SVM dan data uji (*testing*) untuk mengevaluasi kinerjanya.
- d) **Penerapan Algoritma SVM:** Setelah data siap, algoritma SVM diterapkan untuk membuat model klasifikasi. Tahap ini melibatkan pemilihan fungsi kernel yang paling sesuai untuk data.
- e) **Analisis dan Evaluasi:** Hasil klasifikasi diukur menggunakan **Confusion Matrix** untuk menilai tingkat akurasi model. Hasilnya menunjukkan bahwa akurasi tertinggi mencapai **96%** dengan menggunakan kernel **RBF** yang membuktikan efektivitas metode ini.



Gambar 1 Tahapan Penelitian

3. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil dari penerapan metode *Support Vector Machine* (SVM) untuk mengklasifikasikan penerima bantuan PKH di Kelurahan Bangka Leleng. Penelitian ini menggunakan 1.039 sampel data dari tahun 2019-2023 dengan lima atribut, yakni usia, penghasilan, jumlah tanggungan, pekerjaan, dan status kepemilikan rumah.

1. Pengolahan Data

Data diolah menggunakan bahasa pemrograman Python di Google Colab. Alur pengolahan data meliputi pra-pemrosesan, pembagian data, dan implementasi algoritma SVM.

2. Hasil Klasifikasi

Setelah model dilatih, kinerjanya dievaluasi dengan Confusion Matrix. Pengujian menunjukkan bahwa kernel RBF menghasilkan kinerja superior dengan tingkat akurasi tertinggi.

3. Analisis Hasil dan Perbandingan

Tingkat akurasi yang tercatat sebesar 96% mengindikasikan bahwa metode SVM sangat efektif untuk klasifikasi penerima PKH. Temuan ini memperkuat hasil dari studi-studi sebelumnya yang juga menemukan bahwa SVM mampu memberikan klasifikasi yang akurat di berbagai bidang.

4. Kelebihan dan Kelemahan Sistem

- **Kelebihan:** Sistem yang dikembangkan menyediakan metode klasifikasi yang akurat, objektif, dan transparan. Penggunaan SVM dengan kernel RBF terbukti efisien dalam menangani data kompleks, sehingga mengurangi potensi kesalahan dalam penyaluran bantuan.
- **Kelemahan:** Kinerja model sangat bergantung pada kualitas data masukan. Apabila data tidak lengkap atau tidak akurat, hasil klasifikasi bisa saja tidak optimal. Selain itu, kompleksitas SVM bisa membutuhkan sumber daya komputasi yang besar saat memproses data dalam jumlah masif.

Tabel .1 Hasil Evaluasi

| Karnel | Akurasi | Precision | Recall | F1-score |
|------------|---------|-----------|--------|----------|
| Liniear | 90% | 0% | 0% | 0% |
| Polynomial | 95% | 92% | 57% | 71% |
| RBF | 96% | 100% | 57% | 73% |
| Sigmoid | 90% | 0% | 0% | 0% |

Hasil evaluasi, kernel RBF memiliki performa terbaik dengan akurasi 96%, precision 100%, dan F1-score 73%. Namun, recall masih bisa ditingkatkan dengan metode balancing data seperti SMOTE

4. KESIMPULAN

Berdasarkan hasil analisis, penelitian ini menyimpulkan bahwa:

Penerapan Data Mining untuk mengklasifikasi Penerima Bantuan Program Keluarga Harapan (PKH) menggunakan Metode Support Vector Machine

- Metode SVM terbukti efektif dalam mengklasifikasi data penerima PKH, dengan tingkat akurasi mencapai 96%.
- Atribut yang digunakan memiliki peran krusial dalam menentukan klasifikasi penerima bantuan.
- Penggunaan kernel RBF memberikan akurasi dan performa terbaik, menunjukkan kesesuaiannya untuk data non-linear.
- Implementasi metode ini dapat meningkatkan ketepatan dan objektivitas dalam penentuan penerima bantuan PKH, membuat penyaluran bantuan lebih transparan.

Saran : Berikut adalah beberapa rekomendasi untuk penelitian mendatang:

- Pengembangan sistem klasifikasi yang lebih maju dengan menambahkan atribut lain yang relevan (misalnya status kesehatan) dapat dilakukan untuk meningkatkan akurasi.
- Pengujian metode SVM di wilayah lain dengan karakteristik data yang berbeda dapat menguji generalisasi model.
- Disarankan untuk membandingkan kinerja SVM dengan algoritma klasifikasi lain seperti
- Random Forest atau Naive Bayes untuk evaluasi yang lebih komprehensif.

5. REFERENSI

- [1] Arifin, A. A. A., Handoko, W., & Efendi, Z. (2022). Implementasi Metode Naive Bayes Untuk Klasifikasi Penerima Program Keluarga Harapan. *J-Com (Journal of Computer)*, 2(1), 21- 26.
- [2] Arifin, N., Enri, U., & Sulistiyowati, N. (2021). Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification. *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, 6(2), 129-136.
- [3] Astuti, A. (2019). Klasifikasi Penentuan Penerima Program Keluarga Harapan (PKH) Menggunakan Algoritma Support Vector Machine (Svm) Di Kantor Dinas Sosial Lombok Timur. (Doctoral dissertation, Universitas Hamzanwadi).
- [4] Astuti, R., Husen, R. A., Triono, A., & Anam, M. K. (2023, December). Peningkatan Metode Support Vector Machines (SVM) pada Data Child-free Menggunakan Oversampling. In *SEMASTER: Seminar Nasional Teknologi Informasi & Ilmu Komputer (Vol. 2, No. 1, pp. 191- 27)*.
- [5] Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, 8(6), 219-225.
- [6] Dharma, A. S., Silaban, E. R., & Siahaan, H. M. (2023, December). Predictions using Support Vector Machine with Particle Swarm Optimization in Candidates Recipient of Program Keluarga Harapan. In *IAIC International Conference Series (Vol. 4, No. 1, pp. 115-121)*.
- [7] Fibriyanti, B. (2022). Penerapan Algoritma K-Means Untuk Menentukan Prioritas Penerimaan Bantuan PKH. *Jurnal Teknologi Terkini*, 2(12).
- [8] Isyam, M. N., Indrayana, D., & Apriandari, W. (2024). Klasifikasi Penerima Bantuan Program Keluarga Harapan menggunakan Support Vector Machine. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(5), 10416-10421.