

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Klasifikasi adalah proses menilai suatu objek data dan mengelompokkannya ke dalam kelas tertentu dari sejumlah kategori yang tersedia. Teknik ini memungkinkan pembentukan model berdasarkan data yang ada, yang kemudian dapat digunakan untuk mengolah data baru. Klasifikasi dapat dianggap sebagai metode untuk mempelajari pola dalam data guna menghasilkan informasi yang lebih mendetail dan terstruktur.[9].

Klasifikasi menjadi pilihan saat melakukan implementasi algoritma C4.5 karena algoritma ini memang dirancang khusus untuk menyelesaikan masalah klasifikasi. Tujuan utama dari C4.5 adalah membangun pohon keputusan yang dapat memetakan data input ke dalam label kelas tertentu. Dalam proses kerjanya, C4.5 memilih atribut terbaik untuk memisahkan data berdasarkan nilai *information gain* atau *gain ratio*, lalu membagi data ke dalam subset yang lebih spesifik. Proses ini dilakukan secara berulang hingga seluruh data dapat diklasifikasikan ke dalam daun pohon yang merepresentasikan kelas tertentu.

Selain itu, C4.5 mampu menangani data kategorikal maupun numerik, sehingga fleksibel untuk digunakan dalam berbagai domain seperti deteksi spam, diagnosa penyakit, atau klasifikasi berita hoaks. Hasil akhir dari algoritma ini adalah sebuah model berupa pohon keputusan yang dapat digunakan untuk mengklasifikasikan data baru dengan cepat dan mudah dipahami. Oleh karena itu, klasifikasi menjadi pendekatan yang tepat saat mengimplementasikan algoritma C4.5 karena struktur dan mekanisme algoritmanya memang difokuskan untuk menghasilkan model klasifikasi yang akurat.

Bantuan Sosial (BanSos) merupakan suatu program yang dilaksanakan oleh pemerintah untuk memberikan dukungan kepada masyarakat, khususnya bagi keluarga yang memiliki kondisi ekonomi menengah ke bawah. Program ini bertujuan untuk membantu meningkatkan taraf hidup mereka yang membutuhkan. Bantuan sosial yang diberikan sifatnya selektif dan tidak bersifat permanen, artinya bantuan tersebut diberikan dalam jangka waktu tertentu sesuai dengan kebutuhan serta kriteria yang telah ditetapkan oleh pemerintah. Ada berbagai jenis bantuan sosial yang disalurkan oleh pemerintah, seperti bantuan pangan, bantuan tunai, bantuan untuk pendidikan, serta bantuan dalam bentuk fasilitas kesehatan, yang semuanya dirancang untuk membantu meringankan beban

ekonomi masyarakat. Program BanSos ini juga diharapkan dapat mempercepat pemulihan ekonomi masyarakat dan memberikan kesempatan untuk meningkatkan kualitas hidup mereka, terutama di masa-masa sulit seperti pandemi atau krisis ekonomi [1].

Kemiskinan merupakan permasalahan yang bersifat kompleks dan mencakup berbagai aspek kehidupan, seperti rendahnya tingkat kesejahteraan, pendidikan, kesehatan, status gizi anak, serta keterbatasan akses terhadap air bersih. Dalam upaya menanggulangi permasalahan tersebut, pemerintah Indonesia telah melaksanakan berbagai program, salah satunya adalah Program Keluarga Harapan (PKH). PKH merupakan program bantuan sosial yang ditujukan bagi keluarga dengan kondisi ekonomi sangat miskin, yang bertujuan untuk meningkatkan kesadaran akan pentingnya pendidikan dan kesehatan guna menciptakan masa depan yang lebih baik bagi generasi selanjutnya.[2].

Kriteria penerima bantuan sosial mencakup berbagai kondisi yang membutuhkan perhatian khusus dari pemerintah dan lembaga terkait. Penerima bantuan sosial termasuk individu atau kelompok yang berada dalam kondisi kemiskinan, mengalami keterlantaran, atau memiliki kedisabilitas yang membatasi kemampuan mereka untuk memenuhi kebutuhan dasar. Selain itu, mereka yang tinggal di daerah terpencil dengan akses terbatas terhadap layanan dasar juga termasuk dalam kategori penerima. Bantuan sosial juga diberikan kepada mereka yang mengalami ketunaan sosial atau memiliki perilaku yang menyimpang sehingga membutuhkan intervensi sosial. Selain itu, korban bencana alam yang terdampak secara ekonomi dan sosial, serta korban tindak kekerasan, eksploitasi, dan diskriminasi yang mengalami kesulitan dalam menjalani kehidupan sehari-hari, juga berhak mendapatkan bantuan sosial [3].

Desa Pamolokan sendiri merupakan desa yang terletak di pulau Madura, tepatnya di Kabupaten Sumenep. Desa Pamolokan sendiri memiliki populasi sebesar kurang lebih 18.714 jiwa pada tahun 2024. Desa ini sendiri merupakan desa yang cukup potensial karena berada di tengah kota dengan berbagai fasilitas umum, seperti kantor kecamatan dan juga puskesmas. Permasalahan utama yang sedang di alami yaitu apakah pemberian bansos ini sudah tepat sasaran atau belum, oleh karna itu perlu dilakukan klasifikasi ini.

Undang-Undang Dasar Republik Indonesia Tahun 1945 menyatakan bahwa negara memiliki tanggung jawab untuk memelihara fakir miskin dan anak-anak yang terlantar. Hal ini tercantum jelas dalam salah satu pasal yang mengatur tentang kewajiban negara untuk

memastikan kesejahteraan bagi mereka yang membutuhkan. Selain itu, dalam Pasal 27 Ayat 2, disebutkan bahwa setiap warga negara berhak untuk mendapatkan pekerjaan dan kehidupan yang layak, sesuai dengan martabat kemanusiaan [4].

Pemerintah Indonesia terus berupaya memberikan bantuan kepada masyarakat, khususnya kepada mereka yang kurang mampu. Bantuan ini bertujuan untuk meningkatkan kesejahteraan hidup masyarakat dan membantu mereka memenuhi kebutuhan dasar sehari-hari. Upaya ini juga menunjukkan komitmen pemerintah dalam menjalankan amanat Undang-Undang Dasar untuk mewujudkan masyarakat yang adil dan sejahtera, sehingga setiap warga negara dapat hidup dengan layak dan bermartabat. Bantuan sosial yang diberikan, baik dalam bentuk uang tunai maupun barang, menjadi salah satu cara pemerintah untuk mencapai tujuan tersebut sekaligus sebagai bukti nyata pelaksanaan prinsip-prinsip dalam konstitusi negara [4].

Pentingnya topik ini membuat banyak penelitian dilakukan untuk mengkaji lebih lanjut, salah satunya penelitian [5] yang membahas klasifikasi penerima Bantuan Sosial Tunai (BST) di Desa Keramas, Gianyar, Bali. Penelitian ini bertujuan untuk memprediksi kelayakan penerima bantuan sosial, terutama karena meningkatnya angka kemiskinan akibat pandemi Covid-19. Data yang digunakan berjumlah 2.074 kepala keluarga setelah melalui proses pembersihan dari total 2.077 data awal. Proses analisis dilakukan dengan tahapan KDD (*Knowledge Discovery in Database*), yang meliputi pemilihan data, pembersihan data, transformasi data, penambangan data, dan penafsiran hasil. Dalam penelitian tersebut digunakan beberapa metode klasifikasi, yaitu *K-Nearest Neighbor*, *Naïve Bayes*, dan C4.5. Hasilnya menunjukkan bahwa algoritma C4.5 memberikan hasil terbaik dengan tingkat akurasi sebesar 97,83%. Hasil penelitian ini dapat menjadi bahan pertimbangan dalam pengambilan keputusan penyaluran bantuan sosial ke depan. Permasalahan utama yang dibahas adalah ketidaktepatan penerima BST di Desa Keramas. Proses klasifikasi didasarkan pada beberapa faktor, seperti riwayat penerimaan bantuan, status pekerjaan, tingkat pendidikan, kepemilikan kartu miskin, dan jumlah tanggungan. Faktor-faktor tersebut digunakan untuk membantu memastikan penyaluran bantuan dapat dilakukan secara lebih adil dan tepat sasaran.

Selanjutnya oleh [4] mengenai klasifikasi penerima bantuan sosial (BANSOS) di Kecamatan Malangbong, Kabupaten Garut, bertujuan untuk memberikan klasifikasi dan

rekomendasi program bantuan sosial yang dikeluarkan oleh pemerintah Indonesia. Data yang digunakan terdiri dari 100.445 entri yang telah dibersihkan, dengan fokus pada kriteria usia dan pekerjaan. Metodologi penelitian mengikuti tahapan pengumpulan data, pembersihan, transformasi, pemodelan, dan evaluasi. Algoritma yang diterapkan adalah *Naive Bayes*, yang menghasilkan akurasi sebesar 62,51%. Penelitian ini menekankan pentingnya klasifikasi yang akurat untuk meminimalkan kesalahan dalam distribusi bantuan sosial, terutama dalam konteks tantangan ekonomi yang diperburuk oleh pandemi COVID-19.

Penelitian lain yang masih berkaitan ditunjukkan oleh studi [1] yang membahas klasifikasi penerima bantuan sosial (BanSos) di Indonesia menggunakan algoritma *K-Nearest Neighbor* (K-NN). Tujuan penelitian ini adalah membantu pemerintah agar penyaluran bantuan bisa lebih adil dan tepat sasaran. Data yang digunakan mencakup beberapa faktor penting, seperti pendapatan keluarga, kondisi rumah, jumlah tanggungan, dan status pekerjaan. Proses analisis dimulai dari pengumpulan data, kemudian pengolahan data, hingga tahap klasifikasi menggunakan metode K-NN dengan nilai $K = 5$. Hasil pengujian menunjukkan bahwa metode ini mampu menentukan penerima bantuan yang layak dengan tingkat akurasi sebesar 85%. Hasil tersebut menunjukkan bahwa teknik data mining dapat membantu meningkatkan ketepatan sasaran dalam program bantuan sosial sekaligus mengurangi kesalahan dalam proses penyalurannya.

Penelitian berikutnya yang dilakukan oleh [6] mengenai klasifikasi penerima bantuan sosial (BanSos) APBD di Indonesia menggunakan metode *Support Vector Machine* (SVM) untuk meningkatkan efektivitas dan efisiensi penentuan penerima bantuan. Dataset yang digunakan mencakup 100 warga dengan delapan atribut untuk memprediksi kelayakan penerima bantuan (Layak atau Tidak Layak). Metodologi penelitian meliputi transformasi dataset menjadi nilai numerik, pembagian data menjadi set pelatihan (85%) dan set pengujian (15%), serta pelatihan model SVM menggunakan kernel linear. Hasil penelitian menunjukkan bahwa sistem ini berhasil mencapai akurasi sebesar 95%, presisi 100%, dan recall 93%. Penelitian ini bertujuan untuk mengurangi bias dalam proses seleksi penerima bantuan dan meningkatkan transparansi serta efisiensi alokasi sumber daya untuk kesejahteraan masyarakat.

Kajian terakhir adalah penelitian dilakukan oleh [7] mengenai klasifikasi penerima bantuan iuran jaminan kesehatan (BPJS-PBI) di Nusa Tenggara Barat (NTB) menggunakan metode *regresi logistik biner* dan *Naïve Bayes* untuk menentukan kelayakan penerima bantuan. Dataset yang digunakan terdiri dari 6.969 observasi dengan variabel dependen yang menunjukkan apakah individu menerima bantuan kesehatan (0 = Tidak menerima, 1 = Menerima) dan variabel independen terkait kondisi perumahan dan pendidikan. Metodologi penelitian meliputi estimasi parameter menggunakan *Maximum Likelihood Estimation* (MLE) untuk regresi logistik dan perhitungan probabilitas untuk *Naïve Bayes*. Hasil penelitian menunjukkan bahwa metode *Naïve Bayes* mencapai akurasi klasifikasi sebesar 63,9%, sedangkan regresi logistik biner mencapai akurasi 62,26%. Penelitian ini menekankan pentingnya klasifikasi yang tepat untuk memastikan bantuan diberikan kepada yang memenuhi kriteria yang ditetapkan.

Dalam penelitian ini C4.5 digunakan menjadi metode utama dalam pengklasifikasian penerima bantuan social yang dilakukan di desa Pamolokan, kabupaten Sumenep.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, maka rumusan masalah pada penelitian ini sebagai berikut :

1.2.1 Permasalahan

Penelitian ini berfokus pada dua hal utama. Pertama, bagaimana nilai akurasi yang diperoleh dari penerapan algoritma C4.5 dalam mengklasifikasikan penerima bantuan sosial sehingga dapat mendukung proses penentuan penerima bansos secara lebih cepat, objektif, dan efisien. Kedua, bagaimana pengaruh variasi rasio split data terhadap kinerja algoritma C4.5 dalam melakukan proses klasifikasi penerima bantuan sosial guna memperoleh konfigurasi data yang paling optimal. Kedua rumusan masalah ini menjadi dasar dalam menganalisis efektivitas model serta perannya dalam membantu dan mengotomatisasi proses penerimaan bansos agar penyaluran bantuan dapat dilakukan secara tepat waktu dan konsisten.

1.2.2 Solusi Permasalahan

Penggunaan metode klasifikasi berbasis data mining, seperti algoritma C4.5, diharapkan dapat meningkatkan akurasi dan efisiensi dalam pengambilan keputusan terkait penyaluran bantuan sosial.

1.2.3 Pertanyaan Penelitian

1. Berapa akurasi algoritma C4.5 dalam mengklasifikasikan penerima bantuan sosial agar distribusi bantuan lebih tepat sasaran?
2. Bagaimana pengaruh proporsi split data terhadap kinerja algoritma C4.5 dalam klasifikasi penerima bantuan sosial?

1.3 Batasan Masalah

Berikut adalah batasan masalah penelitian yang telah disusun:

1. Data yang digunakan dalam penelitian ini berasal dari dataset penerima bantuan sosial di desa Pamolokan yang merupakan bansos PKH tahun 2025 yang berjumlah 700 data.
2. Parameter yang digunakan pada dataset ini berjumlah 12, meliputi AK (anggota keluarga), Komponen, AUD (anak usia dini), SD, SMP, SMA, Disabil, Lansia, Bumil, Nominal, Status, dan pekerjaan.
3. Metode klasifikasi yang digunakan dalam penelitian ini adalah algoritma C4.5.
4. Evaluasi kinerja model dilakukan berdasarkan akurasi, presisi, dan recall.
5. Studi kasus difokuskan pada penerima bantuan sosial di Desa Pamolokan, Kabupaten Sumenep.
6. Sistem dikembangkan menggunakan bahasa pemrograman *Python* dengan bantuan *library* seperti *Scikit-learn* untuk implementasi algoritma klasifikasi.
7. Dataset yang digunakan hanya mencakup penerima bantuan sosial dari program pemerintah, baik bantuan tunai maupun non-tunai.
8. Tidak membahas aspek kebijakan atau mekanisme pengambilan keputusan dalam program bantuan sosial di tingkat pemerintah daerah atau pusat.
9. Tidak mencakup analisis prediktif terhadap perubahan penerima bantuan sosial di masa depan.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan Penelitian

1. Mengetahui nilai akurasi dari algoritma C4.5 dalam mengklasifikasikan penerima bantuan sosial agar distribusi bantuan lebih tepat sasaran

2. Mengetahui pengaruh variasi split data terhadap kinerja algoritma C4.5 dalam klasifikasi penerima bansos

1.4.2 Manfaat penelitian

Penelitian ini bermanfaat untuk mengetahui seberapa baik algoritma C4.5 dalam mengklasifikasikan penerima bantuan sosial melalui nilai akurasi yang dihasilkan, sehingga dapat digunakan sebagai dasar dalam membangun sistem yang membantu dan mengotomatisasi proses penentuan penerima bansos secara lebih cepat, objektif, dan efisien. Dengan adanya informasi ini, proses penentuan penerima bantuan dapat dilakukan dengan lebih tepat sehingga bantuan benar-benar diterima oleh masyarakat yang berhak dan dapat disalurkan tepat waktu. Selain itu, penelitian ini juga memberikan manfaat dalam mengetahui pengaruh perbedaan pembagian data latih dan data uji (split data) terhadap kinerja algoritma C4.5, sehingga dapat diketahui pembagian data yang paling efektif untuk menghasilkan model klasifikasi yang lebih akurat dan stabil. Penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi pihak terkait dalam meningkatkan sistem penyaluran bantuan sosial berbasis data serta menjadi referensi bagi penelitian selanjutnya yang berkaitan dengan klasifikasi data penerima bantuan sosial dan pengembangan sistem pendukung keputusan.

1.5 Sistematika Laporan

Sistematika penulisan proposal skripsi ini terdiri dari beberapa bab yaitu sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang, perumusan masalah yang meliputi permasalahan, metode yang diusulkan sebagai solusi permasalahan dan pertanyaan penelitian, batasan-batasan penelitian, kemudian selanjutnya terdapat tujuan dan manfaat penelitian, serta sistematika penulisan.

BAB II KAJIAN PUSTAKA

Bab ini berisi penjelasan mengenai teori-teori pendukung dari berbagai sumber yang berhubungan dengan penelitian ini dan penelitian terdahulu yang membahas topik serupa.

BAB III METODOLOGI PENELITIAN

Bab III Metode Penelitian menjelaskan tahapan dan langkah-langkah yang digunakan dalam pelaksanaan penelitian. Pada bab ini diuraikan jenis dan pendekatan penelitian, sumber dan teknik pengumpulan data, proses pengolahan data, serta metode atau algoritma yang digunakan untuk analisis. Selain itu, dijelaskan pula skenario pengujian dan teknik evaluasi yang diterapkan untuk memastikan hasil penelitian dapat diukur secara objektif dan sistematis.

BAB IV HASIL DAN PEMBAHASAN

Bab IV Hasil dan Pembahasan membahas hasil pengolahan dan pengujian data menggunakan metode yang telah ditentukan. Pada bab ini ditampilkan hasil eksperimen atau pengujian model, seperti nilai akurasi dan evaluasi kinerja, kemudian dijelaskan dan dianalisis untuk mengetahui apakah hasil tersebut telah sesuai dengan tujuan penelitian. Pembahasan dilakukan untuk menginterpretasikan hasil yang diperoleh serta menjelaskan faktor-faktor yang memengaruhi kinerja metode yang digunakan.

BAB V KESIMPULAN DAN SARAN

Bab V berisi kesimpulan dan saran sebagai penutup penelitian. Pada bagian kesimpulan, dijelaskan secara singkat hasil utama penelitian berdasarkan tujuan yang telah ditetapkan, seperti kinerja metode yang digunakan dan temuan penting dari proses analisis. Sementara itu, bagian saran memuat rekomendasi untuk pengembangan penelitian selanjutnya, baik dari sisi metode, data, maupun penerapan hasil penelitian agar dapat memberikan manfaat yang lebih optimal.