

PENERAPAN ALGORITME C4.5 TERHADAP KELAYAKAN BANTUAN PANGAN NON TUNAI (BPNT) PADA KANTOR KELURAHAN KEDAUNG

Inda Putri Fatricia^{1*}, Hendri Irawan²

^{1,2}Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, DKI Jakarta, Indonesia

Email: ^{1*}1912500806@student.budiluhur.ac.id, ²hendri.irawan@budiluhur.ac.id

Abstrak- Kemiskinan adalah keadaan saat ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan. Strategi penting untuk mendukung penanggulangan kemiskinan adalah tersedianya data kemiskinan yang akurat dan tepat sasaran. Biasanya pemerintah mengadakan program-program bantuan sembako yang ditunjukkan kepada masyarakat yang membutuhkan. Maka dari itu pemerintah membuat bantuan sosial dalam membantu masyarakat dengan program Bantuan Pangan Non Tunai (BPNT). Permasalahan pada penelitian ini pada Bantuan Sembako masih kurangnya faktor-faktor yang paling berpengaruh terhadap BPNT di desa kedaung dan belum adanya klasifikasi untuk mendapatkan BPNT di Kelurahan Kedaung. Maka dari itu penelitian ini bertujuan untuk memberikan kontribusi yang nyata tentang kelayakan penerima bantuan yaitu BPNT agar penyaluran bantuan benar-benar tepat sasaran sehingga bisa mengangkat perekonomian masyarakat di desa kedaung. Menyelesaikan permasalahan yang ada penelitian ini menggunakan metode algoritma *decision tree* C4.5 dengan mengimplementasikan data P3KE tahun 2022. Penelitian ini menghasilkan atribut atau yang berpengaruh dalam penerima BPNT di Desa Kedaung adalah Status Kawin, Hubungan Dengan Kepala Keluarga, Pekerjaan, Penerima Bantuan Pangan Non Tunai (BPNT), Penerima PKH, Penerima BPUM, Penerima BST, dan Penerima Sembako dengan tingkat Akurasi 97.66%, Presisi 98.12%, dan *Recall* 99.41%.

Kata Kunci: Klasifikasi, Bantuan pangan non tunai (BPNT), *Decision tree* C4.5

APPLICATION OF C4.5 ALGORITHM TO THE FEASIBILITY OF NON-CASH FOOD ASSISTANCE (BANTUAN PANGAN NON TUNAI (BPNT)) AT KEDAUNG VILLAGE OFFICE

Abstrack- Poverty is a state of inability to meet basic needs such as food, clothing, shelter, education and health. An important strategy to support poverty alleviation is the availability of accurate and targeted poverty data. Usually the government holds basic food assistance programs aimed at people in need. Therefore the government made social assistance to help the community with the Non-Cash Food Assistance (BPNT) program. a The problem in this research is that basic food assistance is still lacking in the factors that most influence BPNT in Kedaung village and there is no classification to get BPNT in Kedaung Village. Therefore this study aims to provide a real contribution regarding the eligibility of the recipient of assistance, namely BPNT so that the distribution of assistance is truly on target so that it can lift the economy of the people in Kedaung Village. Solving the problems in this study using the *decision tree* C4.5 algorithm method by implementing P3KE data for 2022. This research produces attributes or those that influence BPNT recipients in Kedaung Village are Marital Status, Relationship with Head of Family, Occupation, Recipients of Non-Cash Food Assistance (BPNT), PKH Recipients, BPUM Recipients, BST Recipients, and Basic Food Recipients with an accuracy rate of 97.66%, 98.12% Precision and 99.41% Recall.

Keywords: *Classification*, Non-cash food assistance (BPNT), *Decision tree* C4.5

1. PENDAHULUAN

Kemiskinan adalah keadaan saat ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan [1]. Strategi penting untuk mendukung penanggulangan kemiskinan adalah tersedianya data kemiskinan yang akurat dan tepat sasaran. Biasanya pemerintah mengadakan program-program bantuan sembako yang ditunjukkan kepada masyarakat yang membutuhkan. Maka dari itu pemerintah membuat bantuan sosial dalam membantu masyarakat dengan program Bantuan Pangan Non Tunai (BPNT). BPNT adalah bantuan sosial pangan yang disalurkan dalam bentuk non tunai dari pemerintah kepada KPM (keluarga penerima manfaat) setiap bulannya melalui mekanisme uang elektronik yang digunakan hanya

untuk membeli bahan pangan di pedagang bahan pangan atau disebut E-warong yang bekerjasama dengan Bank Penyalur [2].

Program bantuan sembako ini diberikan pada keluarga penerima manfaat setiap bulannya dan diberikan kepada seluruh masyarakat di Indonesia dengan kriteria yang telah ditentukan oleh pemerintah melalui Dinas Sosial. Pengelompokan warga miskin sangat diperlukan karena sangat membantu Kepala Kelurahan dalam memilih warga yang miskin. Berdasarkan data tahun 2022 terdapat 2.50 masyarakat miskin di Desa Kedaung. BPNT dapat menimbulkan satu permasalahan yang terjadi yaitu masih banyak warga yang menerima bantuan sembako tidak memenuhi persyaratan, maka terjadilah penerimaan bantuan sosial tidak sistematis dan terkadang tidak tepat. Berdasarkan penjelasan diatas, salah satu metode penelitian yang bisa menangani permasalahan pada penelitian ini adalah data mining.

Data mining merupakan suatu alat yang memungkinkan para pengguna untuk mengakses secara cepat data dengan jumlah yang besar [3]. Dengan salah satu teknik dari data mining yaitu klasifikasi dengan algoritma C4.5 dapat menghasilkan model dalam bentuk pohon atau yang mudah diinterpretasikan. Algoritma C4.5 adalah salah satu algoritma induksi pohon keputusan yaitu ID3 (*Iterative Dichotomiser 3*). Penelitian-penelitian terdahulu terdapat metode yang sudah melakukan penentuan strategi bantuan sosial dilakukan dengan menggunakan klasifikasi algoritma C4.5 [4]. Penelitian lainnya adalah mengelola penerima Bantuan Sosial di Desa Lurah mendapatkan nilai Pengujian algoritma C4.5 dengan akurasi perolehan 58.18. pohon keputusan dibuat dengan algoritma C4.5 [5].

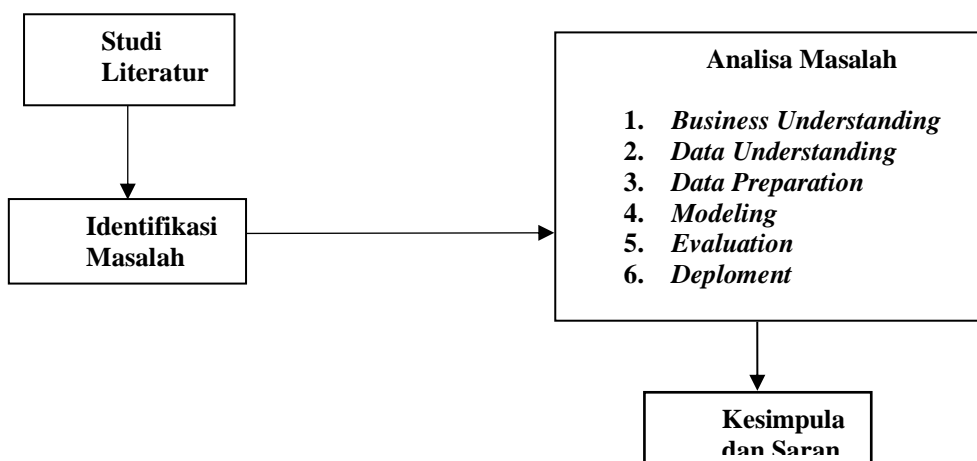
terdapat penelitian lain yang menggunakan algoritma berbeda dari yang sebelumnya yaitu penerapan data mining menggunakan metode algoritma C4.5 penelitian ini mengelola data penerima bantuan sosial Pemko di Kelurahan Martoba, dimana hasil yang diperoleh dari pengujian metode algoritma C4.5 kedalam Rapidminer memiliki validasi yang sama, dengan nilai akurasi sebesar 100% dan nilai gain sebesar 0.8474359 [6]. Penelitian lainnya yang menggunakan algoritma yang sama dengan sebelumnya adalah algoritma algoritma C4.5. Dimana hasil nilai akurasi sebesar 91.54% dan nilai AUC yaitu sebesar 0.986 [7].

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian dilakukan pada penelitian ini dapat dilihat pada

Gambar 1. Pada penelitian ini menggunakan metodologi *Cross Industry Standart Process Model for Data mining* (CRISP-DM). CRISP-DM memiliki tahapan dan kerangka kerja yang terstruktur dan sebagai pemecah masalah yang umum untuk bisnis penelitian [8].



Gambar 1. Tahapan Penelitian

2.1.1 Studi Literatur

Penelitian ini melibatkan pengumpulan data dan informasi yang relevan dengan tujuan penelitian, yaitu penerapan data mining terhadap penerimaan kelayakan penerima sembako dengan menggunakan algoritma C4.5 di Kantor Kelurahan Kedaung. Sebelum memulai penelitian ini, dilakukan studi literatur untuk memperoleh pemahaman yang komprehensif tentang klasifikasi penerimaan bantuan sosial berupa sembako. Dengan

melakukan studi literatur ini, diharapkan peneliti dapat memperoleh pemahaman yang mendalam tentang bantuan sosial, data mining, algoritma C4.5, serta strategi promosi untuk klasifikasi penerimaan bantuan sosial. Pemahaman ini akan menjadi dasar yang kuat untuk melanjutkan penelitian dan mengembangkan metodologi yang sesuai guna mencapai tujuan penelitian.

2.1.2 Identifikasi Masalah

Identifikasi masalah dilakukan dengan topik yang digunakan berdasarkan informasi yang di dapat pada studi literatur. Dalam hal ini pengumpulan data ini, penulis terjun langsung pada objek penelitian untuk mendapatkan data yang valid, maka penelitian menggunakan wawancara dan dokumentasi[9].

2.1.3 Analisa Masalah

Tahapan ini bertujuan untuk mengetahui masalah yang akan diteliti. Tahapan ini menerapkan metodologi CRISP-DM yang memiliki beberapa tahapan didalamnya. Tahapan CRISP-DM dapat dijelaskan sebagai berikut:

a. *Business Understanding*

Tahapan pertama dalam penelitian ini adalah memahami bisnis atau konteks yang akan diolah dalam dataset. Program ini bertujuan untuk memastikan bahwa bantuan pangan tepat sasaran dan efektif dalam memenuhi kebutuhan pangan berupa sembako kepada rumah tangga yang membutuhkan. Analisis data akan dilakukan untuk memahami faktor-faktor yang mempengaruhi penerimaan bantuan sosial, mengidentifikasi pola atau hubungan antara atribut-atribut tertentu dengan penerimaan Sembako.

b. *Data Understanding*

Tahapan kedua dilakukan dengan pengumpulan data penilaian bantuan sosial di kantor kelurahan kedaung pada tahun 2022. Dilanjutkan dengan pemahaman yang mendalam tentang data tersebut, dengan itu untuk mendeteksi bagian yang menarik sekiranya dapat digunakan sebagai hipotesa informasi tersembunyi. Setelah itu didapatkan beberapa atribut yang diperlukan untuk penelitian, data penerima bantuan sembako yang didapatkan sebanyak 11.295 data. Setelah mendapatkan data yang berkaitan dengan penerima bantuan penerima sembako, berikutnya dilanjutkan dengan proses mendapatkan yang mendalam tentang data, selain mendeteksi bagian menarik dari data yang dapat digunakan untuk membuat hipotesis informasi yang tersembunyi.

c. *Data Preparation*

Tahapan Proses data preparation meliputi pembersihan data, transformasi data, integrasi data, formatting data, dan split data. Tahap Data Preparation penting untuk memastikan bahwa data yang digunakan dalam proses data mining adalah valid, bersih, dan siap digunakan. Dengan melakukan pembersihan, transformasi, integrasi, dan formatting data dengan baik, dataset yang telah dipersiapkan akan memungkinkan peneliti untuk melanjutkan ke tahap selanjutnya yaitu proses data mining menggunakan algoritma C4.5 untuk menganalisis penerimaan kelayakan Sembako di Kantor Kelurahan Kedaung.

d. *Modeling*

Tahapan selanjutnya dalam penelitian ini adalah Modeling dilakukan dengan metode *Cross Validation*. Pada tahap ini, pengujian yang dilakukan dalam penelitian ini menggunakan *k-flods cross validation* dimana $k=10$ (*10-fold cross validation*).

e. *Evaluation*

Tahapan selanjutnya adalah Evaluation. Pada tahap ini, dilakukan pengukuran kinerja model algoritma untuk melakukan klasifikasi terhadap penerimaan bantuan sosial yang terjadi di Kantor Kelurahan Kedaung. Pengukuran ini akan menggunakan *confusion matrix* yang mencakup metrik-metrik penting seperti *Presisi*, *Recall*, dan Akurasi.

f. *Deployment*

Tahapan ini dilakukan setelah menyelesaikan tahapan diatas untuk menghasilkan laporan klasifikasi untuk menghasilkan penerimaan bantuan sosial yang terjadi pada Kantor Kelurahan Kedaung.

2.1.4 Kesimpulan dan saran

Pada kesimpulan dan saran berdasarkan dari hasil Presisi, *Recall*, dan Akurasi serta saran dari apa yang telah didapatkan dalam Presisi, *Recall*, dan Akurasi selanjutnya yang akan dilakukan pada masa yang akan datang.

2.2 Data Preprocessing

Tahapan yang digunakan harus melalui data *preprocessing* karena sumber data yang diperoleh masih kotor, Kegiatan *preprocessing* data yang dilakukan yaitu reduksi data atau penghapusan atribut. kemudian dilakukan pemfilteran untuk menyaring data-data yang tidak sesuai. Data yang tidak sesuai merujuk pada data yang tidak memenuhi persyaratan atau tidak konsisten dengan kriteria yang ditetapkan dalam konteks penelitian. Dalam kasus ini disebutkan, ada beberapa atribut yang mengalami penyaringan untuk menyaring data yang tidak sesuai. Jumlah data yang diperoleh tahun 2022 sebanyak 11.295 data.

2.3 Penentuan Data Latih (*Data Training*) dan Data Uji (*Data Testing*)

Pada tahap ini, dilakukan pemisahan data menjadi data latih (*data training*) dan data uji (*data testing*). Penentuan data ini akan dilakukan 3 kali percobaan dengan pembagian data sebesar 60:40, 70:30, 80:20 untuk pengujian. Dengan dilakukannya perbandingan akan menghasilkan nilai akurasi terbaik.

2.4 Pengembangan Model

Dalam tahap pengembangan model, penelitian ini melakukan eksplorasi dengan menggunakan algoritma Decision Tree C4.5 Algoritma C4.5 merupakan algoritma yang sangat populer yang digunakan oleh banyak peneliti di dunia, Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang di ciptakan oleh J. Rose Quinlan[10]. Proses ini bertujuan untuk melakukan analisis terhadap masyarakat penerimaan bantuan sosial di Kelurahan Kedaung. Proses ini bertujuan untuk mengembangkan model yang dapat digunakan untuk melakukan klasifikasi terhadap masyarakat penerimaan bantuan sosial di kelurahan tersebut.

2.5 Teknik Pengujian

Teknik pengujian yang dilakukan pada penelitian ini menggunakan metode Confision Matrix mempresentasikan evaluasi model. Matriks konfusi adalah matriks dua dimensi yang barisnya menunjukkan label sebenarnya dan kolomnya menunjukkan label yang diprediksi oleh pengklasifikasi. Dalam mencari evaluasi Confusion Matrix menghasilkan nilai Akurasi, Presisi, dan Recall. *Confusion Matrix* adalah sebuah tabel yang digunakan untuk menganalisis kinerja model klasifikasi.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Pada penelitian ini pengumpulan data dilakukan dengan cara wawancara dengan pihak desa Kedaung untuk mendapatkan data mengenai penerima bantuan sembako. Dengan melakukan wawancara dengan pihak desa dan mendapatkan jumlah record sebanyak 11.295 record. Data atau record yang dikumpulkan adalah data P3KE pada tahun 2022 pada Kantor Kelurahan Kedaung. Data yang diambil dalam penelitian ini masih menggunakan data mentah sebelum dilakukan *data preprocessing* yang dimana terdiri dari 32 atribut yaitu No, No, Id Keluarga P3KE, Provinsi, Kabupaten, Kecamatan, Desa/Kelurahan, Kode Kemdagri, Desil Kesejahteraan, Alamat, Id Individu, Nama, Nik, Pandan Dukcap, Jenis Kelamin, Hubungan dengan Kepala Keluarga, Tanggal Lahir, Status Kawin, Pekerjaan, Pendidikan, Usia dibawah 7 Tahun, Usia 7-12, Usia 13-15, Usia 116-18, Usia 19-21, Usia 22-59, Usia diatas 60 Tahun, Penerima Bantuan Pangan Non Tunai (BPNT), Penerima BPUM, Penerima, PKH, Penerima BST, dan Penerima Sembako.

3.2 Data Preprocessing

Data *preprocessing* dapatkan dilakukan dengan data yang berkualitas, pada penelitian ini dilakukan teknik data *reduction* dan data transformation pada tahap preprocessing. Melakukan data *reduction* untuk melengkapi atau menghilangkan data yang tidak lengkap (Missing Value), pemfilteran data, dan menyaring data yang tidak konsisten.

3.1.1 Data Reduction

Berdasarkan raw data yang tersedia sebanyak 32 atribut untuk klasifikasi algoritma C4.5 dalam penentuan sembako. Kegiatan *preprocessing* data yang dilakukan yaitu *reduksi* data atau penghapusan atribut. Penghapusan atribut (kolom) terdapat atribut yang tidak digunakan yaitu sebanyak 23 atribut yaitu no, id keluarga P3KE, provinsi, kabupaten, kecamatan, desa, alamat, nama, nik, kode kemdagri, desil kesejahteraan, id individu.

tanggal lahir, jenis kelamin, Pendidikan, usia dibawah 7 tahun, usia 7-12 tahun, usia 13-15 tahun, usia 16-18, usia 19-21 tahun, usia 22-59 tahun, usia 60 tahun keatas juga dihapus dikarenakan tidak berpengaruh.. kemudian dilakukan pemfilteran untuk menyaring data-data yang tidak sesuai.

Atribut "hubungan dengan kepala keluarga", kategori yang ada yaitu Kepala Keluarga, Istri, Anak dan lainnya. Dilakukan penyaringan pada kategori "lainnya". Data dengan hubungan "lainnya" dihapus karena tidak jelas apakah statusnya adalah mertua, cucu, keponakan, atau kategori lainnya yang tidak dapat diidentifikasi dengan pasti. Pada data kategori yang dilakukan penyaringan dapat dilihat pada Gambar 2 Pemfilteran Atribut

Ya	Laki-laki	Anak	05/06/2006 0:00	Belum Kawin
Ya	Perempuan	Lainnya	17/07/1967 0:00	Cerai Mati
Ya	Laki-laki	Anak	26/04/1998 0:00	Belum Kawin
Ya	Perempuan	Anak	15/12/1996 0:00	Belum Kawin
Ya	Laki-laki	Kepala Keluarga	10/08/1988 0:00	Kawin
Tidak	Perempuan	Istri	24/12/1994 0:00	Kawin
Tidak	Laki-laki	Anak	05/06/2006 0:00	Belum Kawin
Tidak	Perempuan	Lainnya	17/07/1967 0:00	Cerai Mati
Ya	Laki-laki	Kepala Keluarga	10/11/1977 0:00	Kawin

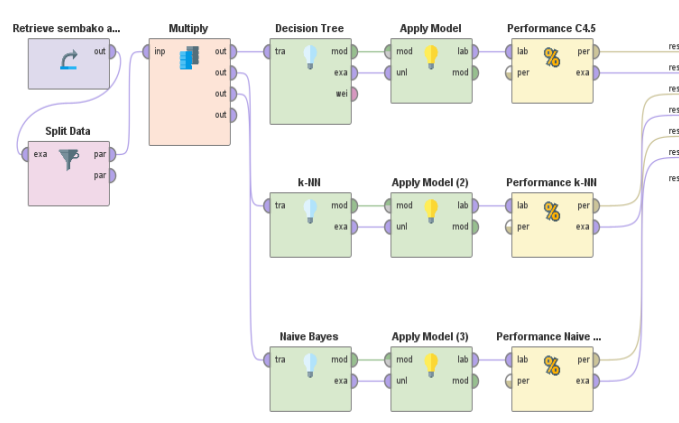
Gambar 2 Pemfilteran Atribut

3.1.2 Data Transformation

Pada tahap transformasi ini untuk mengubah seluruh data yang terkumpul agar memudahkan ketika diolah dengan menggunakan decision tree C4.5. terdapat teknik untuk melakukan data transformation adalah pemilihan atribut. Adapun 8 atribut yang digunakan pada penelitian ini yaitu Status kawin, Hubungan dengan kepala keluarga, pekerjaan, penerima Bantuan Pangan Non Tunai (BPNT), penerima PKH, penerima BPUM, penerima BST, dan penerima sembako.

3.3 Hasil Data Latih (Data Training) dan Data Uji (Data Testing)

Setelah dilakukan preprocessing data, selanjutnya akan dilakukan penentuan data *Training* (data latih) dan data *Testing* (data uji). Data *Training* adalah data yang digunakan untuk melatih mesin agar dapat mengenali pola, sedangkan data *Testing* adalah data yang digunakan untuk menguji hasil dari pelatihan yang telah dilakukan pada mesin. Dalam penelitian ini data *Training* dan data *Testing* dibagi menjadi tiga perbandingan yaitu 80:20, 70:30, dan 60:40. Sebaran data *Training* dan data *Testing* dengan menggunakan *tools* Rapidminer pisahkan data menjadi *split data* dengan secara otomatis.



Gambar 3. Permodelan dengan Teknik Split Data

Hasil akurasi dari perbandingan 60% merupakan data *training* dan 40% merupakan data *testing*, 70% merupakan data *training* dan 30% merupakan data *testing*, 80% merupakan data *training* dan 20% merupakan data *testing*. Sebaran data *Training* dan data *Testing* pada Tabel 1.

Tabel 1. Hasil Accuracy

Pembagian Data	Hasil Akurasi
60:40	97.45%
70:30	97.30%
80:20	97.60%

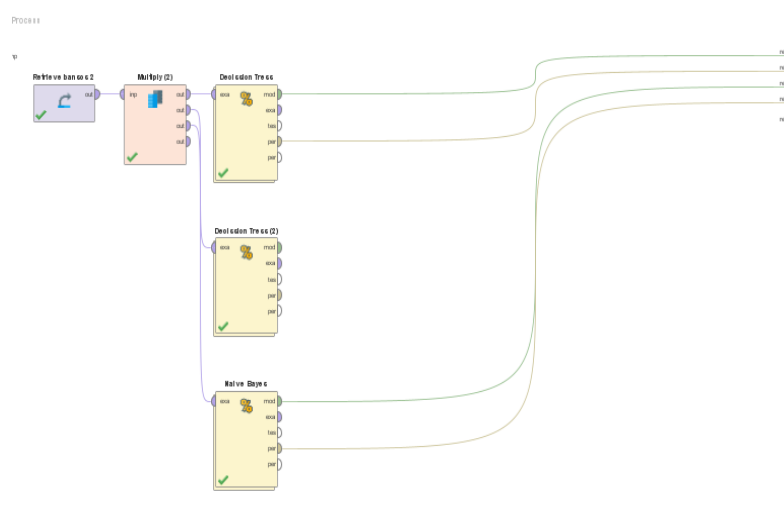
Setelah melakukan eksplorasi pada Tabel 1. Hasil dengan percobaan perbandingan data 60:40, 70:30, dan 80:20 bahwa komposisi data *training* dan *testing* yang menghasilkan akurasi tertinggi dengan pembagian data 80:20 yaitu sebesar 97.60%.

3.4 Pemodelan

Pada tahap ini melakukan tahap modeling untuk menentukan perbandingan tingkat akurasi algoritma digunakan sebagai perbandingan yaitu algoritma C4.5 dengan algoritma *Naive Bayes* dan algoritma *K-NN*. Pengujian ini dilakukan menggunakan metode *10-fold Cross Validation* dengan tools Rapidminer.

3.4.1 Hasil Komperasi Model

Pada tahap awal permodelan, dilakukan eksplorasi beberapa algoritma untuk mengetahui performance model terbaik. Parameter uji yang digunakan adalah nilai *Accuracy*, *Presisi*, *Recall*, dan AUC. Merupakan hasil akurasi algoritma C4.5 menggunakan metode *10-fold Cross Validation*. Seperti Gambar 4.



Gambar 4. Model Menggunakan *10-fold Cross Validation*

Berdasarkan pengujian beberapa algoritma *data mining* dengan menggunakan proses Rapidminer tersebut dilihat dari hasil tingkat *akurasi*. Berikut adalah ringkasan hasil dari perbandingan *akurasi* pada masing-masing algoritma dilihat pada Tabel 2.

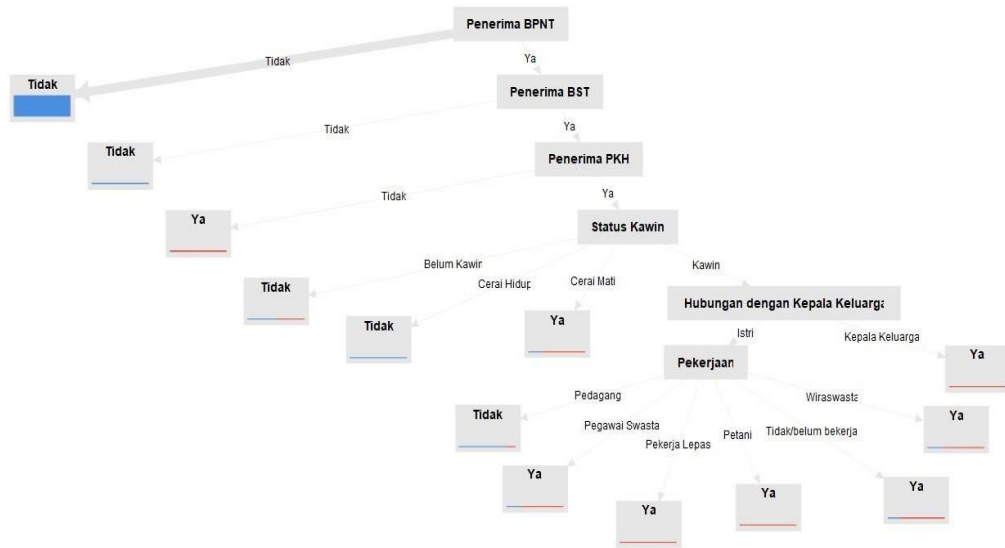
Tabel 2. Komperasi Model *10-fold Cross Validation*

Algoritma	Akurasi (%)
C4.5	97.66%
K-NN	97.37%
Naive Bayes	97.63%

Bedasarkan hasil *akurasi* dari algoritma C.45, algoritma *Naive Bayes*, dan algoritma *K-NN*, memiliki hasil komparasi model pada algoritma C.45 dengan hasil *akurasi* 97.66%. Lebih tinggi dari hasil *akurasi* algoritma *Naive Bayes* dan algoritma *K-NN*.

3.4.1 Penyajian Model Terbaik

Setelah dilakukan hasil komparasi model, selanjutnya pada bagian ini disajikan hasil pemrosesan dari model atau algoritme Decision Tree C4.5. Implementasi algoritma Decision Tree C4.5 terdiri dari, pembentukan *decision rules* dan visualisasi pohon keputusan seperti Gambar 5.



Gambar 5. Pohon Keputusan

Dihasilkan pohon keputusan seperti Gambar 5. Merupakan hasil dari pohon keputusan dan atribut Penerima Sembako adalah akar dari pohon keputusan. Penerima Sembako sebagai atribut yang paling berdampak dalam kalsifikasi penerimaan bantuan sembako dilanjutkan dengan Hubungan Dengan Kepala Keluarga. Hasil dari pola *decision rules* dalam bentuk *If-Then* pada

Tree

```
Penerima BPNT = Tidak: Tidak {Tidak=8882, Ya=59}
Penerima BPNT = Ya
| Penerima BST = Tidak: Tidak {Tidak=155, Ya=6}
| Penerima BST = Ya
| | Penerima PKH = Tidak: Ya {Tidak=5, Ya=226}
| | Penerima PKH = Ya
| | | Status Kawin = Belum Kawin: Tidak {Tidak=1, Ya=1}
| | | Status Kawin = Cerai Hidup: Tidak {Tidak=2, Ya=0}
| | | Status Kawin = Cerai Mati: Ya {Tidak=3, Ya=10}
| | | Status Kawin = Kawin
| | | | Hubungan dengan Kepala Keluarga = Istri
| | | | | Pekerjaan = Pedagang: Tidak {Tidak=4, Ya=1}
| | | | | Pekerjaan = Pegawai Swasta: Ya {Tidak=1, Ya=3}
| | | | | Pekerjaan = Pekerja Lepas: Ya {Tidak=0, Ya=3}
| | | | | Pekerjaan = Petani: Ya {Tidak=0, Ya=2}
| | | | | Pekerjaan = Tidak/belum bekerja: Ya {Tidak=30, Ya=124}
| | | | | Pekerjaan = Wiraswasta: Ya {Tidak=2, Ya=6}
| | | | Hubungan dengan Kepala Keluarga = Kepala Keluarga: Ya {Tidak=0, Ya=7}
```

Gambar 6.

Tree

```

Penerima BPNT = Tidak: Tidak {Tidak=8882, Ya=59}
Penerima BPNT = Ya
| Penerima BST = Tidak: Tidak {Tidak=155, Ya=6}
| Penerima BST = Ya
| | Penerima PKH = Tidak: Ya {Tidak=5, Ya=226}
| | Penerima PKH = Ya
| | | Status Kawin = Belum Kawin: Tidak {Tidak=1, Ya=1}
| | | Status Kawin = Cerai Hidup: Tidak {Tidak=2, Ya=0}
| | | Status Kawin = Cerai Mati: Ya {Tidak=3, Ya=10}
| | | Status Kawin = Kawin
| | | | Hubungan dengan Kepala Keluarga = Istri
| | | | | Pekerjaan = Pedagang: Tidak {Tidak=4, Ya=1}
| | | | | Pekerjaan = Pegawai Swasta: Ya {Tidak=1, Ya=3}
| | | | | Pekerjaan = Pekerja Lepas: Ya {Tidak=0, Ya=3}
| | | | | Pekerjaan = Petani: Ya {Tidak=0, Ya=2}
| | | | | Pekerjaan = Tidak/belum bekerja: Ya {Tidak=30, Ya=124}
| | | | | Pekerjaan = Wiraswasta: Ya {Tidak=2, Ya=6}
| | | | Hubungan dengan Kepala Keluarga = Kepala Keluarga: Ya {Tidak=0, Ya=7}

```

Gambar 6. Hasil Pola data berupa Peraturan *If-Then*

RULES :

1. Jika Tidak Penerima BPNT, maka tidak menerima BPNT.
2. Jika Ya Penerima BPNT = “Tidak” menerima BST, maka tidak menerimaBPNT.
3. Jika penerima BPNT = “Ya” menerima BST = “Tidak” menerima PKH, makamenerima BPNT.
4. Jika penerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, dan statuskawin = “belum kawin”, maka tidak menerima BPNT.
5. Jika penerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, dan statuskawin “cerai hidup”, maka tidak menerima BPNT.
6. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, dan statuskawin “cerai mati”, maka menerima BPNT.
7. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, dan hubungan dengan kepala keluarga sebagai “kepala keluarga”, maka menerima BPNT.
8. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “pedagang”, maka tidak menerima BPNT.
9. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “pegawai swasta”, maka tidak menerima BPNT.
10. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “pedagang”, maka tidak menerima BPNT.
11. Jika menerima BPNT = “Ya” menerima BST, menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “wiraswasta”, maka menerima BPNT.
12. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “tidak/belum bekerja”, maka menerima BPNT.
13. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “petani”, maka menerima BPNT.
14. Jika menerima BPNT = “Ya” menerima BST = “Ya” menerima PKH, status kawin “kawin”, hubungan dengan kepala keluarga sebagai “istri”, pekerjaan sebagai “pekerja lepas”, maka menerima BPNT.

3.5 Pengujian

Pada tahap ini, pengujian menggunakan metode *Confusion Matrix* dan parameter uji yang terdiri dari akurasi, *Recall*, dan presisi. Hasil pengujian disajikan dalam bentuk tabulasi, menggunakan perbandingan *10-fold Confusion Matrix*. Pengujian menggunakan *tools* rapidminer dapat dilihat pada Tabel 3.

Tabel 3. Hasil *Confusion Matrix* dari pemodelan terpilih

	<i>True</i>	<i>True</i>	<i>Class Precision</i>
--	-------------	-------------	------------------------

	Tidak	Ya	
Pred. Tidak	8888	170	98,12%
Pred. Ya	53	422	88,84%
Class Recall	99,41%	71,28%	

- *Prediction Ya - True Ya* : Jumlah data yang diprediksi Ya dan kenyataannya Ya (TP).
- *Prediction Tidak – True tidak* : Jumlah data yang diprediksi Tidak dan kenyataannya Ya (TN).
- *Prediction Ya – True tidak* : Jumlah data yang diprediksi Ya dan kenyataannya Tidak (FP).
- *Prediction Tidak – True Ya* : Jumlah data yang diprediksi Tidak dan kenyataannya Ya (FN).

Peneliti melakukan perhitungan mandiri pada Tabel 3 sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{8888 + 422}{8888 + 422 + 170 + 53} = 97,66\%$$

$$Precision = \frac{TP}{TP + FP} = \frac{8888}{8888 + 170} = 98,12\%$$

$$Recall = \frac{TP}{TP + FN} = \frac{8888}{8888 + 53} = 99,41\%$$

Precision menggambarkan Pred.Tidak kasus dari hasil yang tidak menerima bantuan sembako. *Recall* menggambarkan pred kasus dari hasil yang diidentifikasi menerima bantuan sembako. *Akurasi* menggambarkan perbandingan kasus yang diidentifikasi secara benar (Ya dan Tidak menerima bantuan sembako) dengan semua kasus.

Berdasarkan hasil perbandingan antara perhitungan mandiri dan perhitungan menggunakan *tools* Rapidminer didapatkan kesamaan, yaitu akurasi sebesar 97,66%, *Recall* 98.12%, dan presisi sebesar 99.41%. Dari hal tersebut, disimpulkan bahwa pengujian berhasil.

4. KESIMPULAN

Berdasarkan penerapan algoritma C4.5 dapat digunakan untuk mengklasifikasi faktor penerima bantuan sosial yang paling berpengaruh adalah Status Kawin, Hubungan dengan Kepala Keluarga, Pekerjaan, Penerima Bantuan pangan non tunai (BPNT), Penerima Program Keluarga Harapan (PKH), penerima Bantuan Produktif Usaha Mikro (BPUM), Penerima Baantuan Sosial Tunai (BST), Penerima manfaat BPNT. Penerapan algoritma C4.5 dapat digunakan untuk mengklasifikasi penerima Bantuan pangan non tunai (BPNT) dengan lebih akurat sebesar 97.66% dibandingkan dengan algoritma K-NN sebesar 97.37% dan Naïve Bayes sebesar 97.63%. Hal ini terbukti bahwa penerapan Algoritma C4.5 dalam menentukan penerimaan BPNT sangat tepat digunakan untuk membantu pihak-pihak yang berkepentingan dalam pengambilan keputusan karena hasil dari akurasi 97.66% menentukan warga yang layak dalam menerima BPNT dan jumlah sebesar 2.34% menentukan ketidaklayakan warga dalam menerima BPNT.

DAFTAR PUSTAKA

- [1] D. Wahyuli, H. Handrizal, I. Parlina, A. P. Windarto, D. Suhendro, dan A. Wanto, "Mengelompokkan Garis Kemiskinan Menurut Provinsi Menggunakan Algoritma K-Medoids," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, hal. 452, 2019, doi: 10.30645/senaris.v1i0.51.
- [2] I. Fadlurrohman, S. A. Nulhaqim, dan S. Sulastri, "Implementasi Program Bantuan Pangan Non Tunai (Studi Kasus Di Kota Cimahi)," *Share Soc. Work J.*, vol. 9, no. 2, hal. 122, 2020, doi: 10.24198/share.v9i2.20326.
- [3] C. Zai, "Implementasi Data Mining Sebagai Pengolahan Data," *Portal Data*, vol. 2, no. 3, hal. 1–12, 2022.
- [4] S. Badriah, M. F. E. Nugroho, N. Sanjaya, I. Rismawati, B. N. Sari, dan C. Rozikin, "Klasifikasi Algoritma C4 . 5 dalam Menentukan Penerima Bantuan Covid-19 (Studi Kasus : Desa di Karawang)," *JIP (Jurnal Inform. Polinema)*, vol. 7, no. 3, hal. 23–28, 2021.

- [5] O. N. Washilaturrizqi, “Implementasi Algoritma C4 . 5 Untuk Menentukan Penerima,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, hal. 373–377, 2023.
- [6] W. Lidysari, H. S. Tambunan, dan H. Qurniawan, “Penerapan Data Mining Dalam Menentukan Kelayakan Penerima Bantuan Sosial Pemko Dengan Algoritma C4.5 (Kasus Kantor Kelurahan Martoba),” *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 3, no. 1, hal. 53–61, 2022, doi: 10.30645/kesatria.v3i1.97.
- [7] R. A. Saputra, S. Wasiyanti, dan D. Pribadi, “Information Gain Pada Algoritma C4.5 Untuk Klasifikasi Penerimaan Bantuan Pangan Non Tunai (Bpnt),” *Indones. J. Bus. Intell.*, vol. 4, no. 1, hal. 25, 2021, doi: 10.21927/ijubi.v4i1.1757.
- [8] M. A. Hasanah, S. Soim, dan A. S. Handayani, “Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir,” vol. 5, no. 2, 2021.
- [9] A. H. S. Negoro *et al.*, *Prosiding Diseminasi Hasil Penelitian Dosen 2022*. Jember: UPT Penerbit Universitas Jember, 2022.
[Daring]. Tersedia pada:
https://www.google.co.id/books/edition/Prosiding_Diseminasi_Hasil_Penelitian_Do/FXt8EAAAQBAJ?hl=id&gbpv=1
- [10] D. Marlina dan M. Bakri, “Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 1, hal. 23–28, 2021.