

PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI PENENTUAN PENERIMA PROGRAM KARTU INDONESIA SEHAT

Cece Pratiwi, Tati Ernawati

Teknik Informatika, Politeknik TEDC Bandung
Jalan Politeknik-Pesantren KM2 Cibabat Cimahi Utara – Cimahi Indonesia
Cecepratiwi07@gmail.com

ABSTRAK

Kemiskinan merupakan satu masalah yang kompleks dan menjadi bagian yang tidak dapat dipisahkan dari permasalahan masyarakat khususnya di pedesaan. Salah satu upaya pemerintah untuk mengatasi kemiskinan adalah melalui program Kartu Indonesia Sehat (KIS) pada setiap daerah di Indonesia. Permasalahan yang terjadi di Kecamatan Pasi Kolaga di Provinsi Sulawesi Tenggara terkait bantuan KIS adalah proses penyaluran bantuan kepada masyarakat tidak tepat sasaran dikarenakan pengelolaan data yang masih dilakukan secara manual. Tujuan penelitian ini adalah mengetahui hasil klasifikasi data menggunakan algoritma C4.5 untuk penerima KIS, yang akan dibandingkan dengan sebuah simulasi aplikasi sederhana berbasis *web*. Metode penelitian mencakup sembilan tahap (menentukan masalah, menentukan tujuan dan ruang lingkup, mencari literatur, mengumpulkan data, menentukan populasi dan sampel, metode yang digunakan, proses dan hasil penelitian, analisis dan penarikan kesimpulan) proses algoritma C4.5 dilakukan terhadap 2117 data. Hasil uji *rapidminer* dan aplikasi berbasis *web* menunjukkan hasil yang sama yaitu 1916 data diterima dan 201 data ditolak dengan tingkat akurasi 97,92%. Hal ini menunjukkan aplikasi dapat digunakan sebagai alat bantu bagi perangkat desa untuk mempermudah pengelolaan data bantuan KIS.

Kata kunci : *Kartu Indonesia Sehat (KIS), Klasifikasi, Algoritma C4.5, Web*

1. PENDAHULUAN

Kemiskinan merupakan satu masalah yang kompleks dan menjadi bagian yang tidak dapat dipisahkan dari permasalahan masyarakat khususnya di pedesaan. Kemiskinan diindikasikan dengan kurangnya pendapatan dalam memenuhi standar minimum kebutuhan hidup sandang, pangan, papan, pendidikan dan kesehatan. Terkait permasalahan kemiskinan, pemerintah Indonesia telah membuat program pengetasana kemiskinan berupa bantuan sosial seperti Kartu Indonesia Sehat (KIS) [1]. Program KIS memberikan bantuan iuran yang ditujukan kepada warga yang tidak mampu dan fakir miskin sebagai satu dari beberapa usaha pemerintah dalam meningkatkan status sosial ekonomi bagi keluarga tidak mampu [2].

Sistem penerapan Algoritma C4.5 yang diterapkan pada aplikasi atau sistem informasi sebelumnya telah dibuat oleh [3] mengkaji penerapan algoritma C4.5 penentuan program bantuan pemerintah daerah yang diimplementasikan pada aplikasi berbasis desktop, hasil kajian menunjukkan aplikasi dapat memprediksi dengan tepat dan akurat tingkat penerimaan program bantuan. [4] membuat sistem informasi dengan menerapkan algoritma C4.5 untuk kelayakan penerimaan pegawai, hasilnya mengindikasikan bahwa sistem informasi yang dibangun menggunakan Adobe Dreamweaver CS3 dapat memudahkan pengolahan data untuk memprediksi kelayakan penerimaan pegawai. Studi lain dilakukan oleh [5] menerapkan algoritma C4.5 pada sistem untuk klasifikasi penerima bantuan sosial

bagi mahasiswa dengan hasil kajiannya yaitu sistem dapat memberikan klasifikasi yang tepat. Perbedaan penelitian yang dikaji dengan penelitian sebelumnya adalah membandingkan klasifikasi algoritma C.45 dengan *rapidminer* dengan aplikasi menggunakan *Hypertext Preprocessor* (PHP) dan JavaScript (JS) yang berbasis *web* pada program bantuan KIS.

Permasalahan dilapangan berdasarkan hasil wawancara dengan staf kecamatan Pasi Kolaga dengan jumlah penduduk 2117 orang dengan mayoritas mata pencaharian sebagai nelayan yang dilakukan pada tanggal 27 Januari 2024 ditemukan permasalahan yaitu kendala penyaluran bantuan KIS dimana bantuan masih belum tepat sasaran, seperti penerima yang seharusnya menerima bantuan tidak memperoleh bantuan dikarenakan kesalahan data. Terdapat sejumlah masyarakat yang seharusnya menerima bantuan tetapi tidak menerima. Penyebab permasalahan yang terjadi adalah penelolan data penerima KIS masih diolah secara manual sehingga sering terjadi kesalahan pada saat penentuan calon penerima bantuan. Merujuk pada permasalahan tersebut diajukan sebuah solusi berupa pengelolaan data KIS memanfaatkan data mining yang akan dibuat dalam sebuah aplikasi sederhana untuk memudahkan dalam pengolahan data. Hasil klasifikasi data menggunakan perangkat lunak *rapidminer* akan dibandingkan dengan hasil dari aplikasi yang dibuat.

Penelitian ini bertujuan untuk mengetahui hasil klasifikasi data menggunakan algoritma C4.5 untuk penerima KIS, yang akan dibandingkan dengan sebuah simulasi aplikasi sederhana berbasis *web*.

2. TINJAUAN PUSTAKA

2.1. Kartu Indonesia Sehat (KIS)

KIS adalah kartu yang memiliki fungsi memberikan jaminan kesehatan kepada masyarakat untuk mendapatkan pelayanan kesehatan secara gratis [6]. KIS dikelola oleh badan penyelenggara jaminan sosial kesehatan, dibuat sebagai penyempurnaan program BPJS Kesehatan untuk peserta Penerima Bantuan Iuran (PBI) [7].

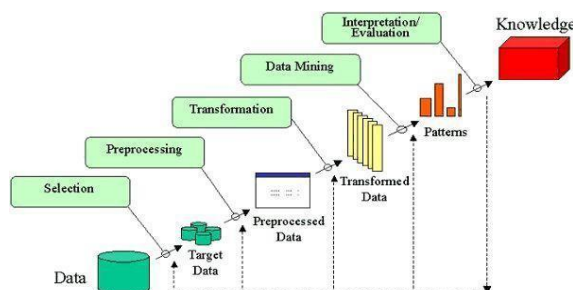
Kriteria penerima KIS berdasarkan Undang-Undang No. 40/2004 tentang Sistem Jaminan Sosial Nasional dan diimplementasikan melalui Jaminan Kesehatan Nasional (JKN) oleh BPJS Kesehatan adalah penerima Bukan Pekerja Penerima Upah (PBPU) [8].

2.2. Data Mining

Data mining didefinisikan sebagai proses mengekstrak informasi yang berguna dari kumpulan data skala besar baru untuk memecahkan masalah, *data mining* dikenal sebagai *knowledge discovery* [9]. Definisi lain *data mining*, proses ekstraksi data menjadi informasi yang belum pernah terjadi sebelumnya. Menggunakan teknik yang tepat, proses ini akan menghasilkan hasil terbaik [10].

2.2.1. Proses Data Mining

Data Mining Knowledge Discovery in Database (KDD) dapat mengungkapkan informasi yang tersembunyi dalam tumpukan data [11].



Gambar 1. Proses KDD

Keterangan[12].

a. Data Selection

Sebelum menggunakan informasi KDD, data yang akan digunakan untuk *data mining* harus dikumpulkan dan dianalisis dalam format yang berbeda dari *database* operasional.

b. Pre-processing cleaning

Data mining adalah proses menggabungkan data duplikat, mempertahankan data yang tidak konsisten, dan meningkatkan kualitas, seperti tipografi, dengan informasi yang relevan, seperti data eksternal, untuk menciptakan fokus KDD.

c. Transformation

Kode dalam KDD adalah proses kreatif dan generatif yang mengubah data menjadi seperangkat informasi, yang dapat digunakan untuk *data mining*.

d. Interpretation/Evaluasi

Pola informasi dihasilkan dari proses *data mining* harus disajikan dalam bentuk yang mudah dipahami oleh pengguna disebut interpretasi. Tahap ini mencakup menentukan apakah pola/informasi yang ditemukan bertentangan dengan fakta/hipotesa sebelumnya.

2.3. Metode Klasifikasi

Metode klasifikasi merupakan bagian dari cabang dari *discovery data mining* yang termasuk dalam metode *supervised learning*. Metode ini merupakan metode pendekatan yang memiliki data latin dan variabel target yang akan dicari hubungan di antara keduanya. Klasifikasi biasanya digunakan untuk mengelompokkan data dengan cara menentukan objek/konsep kedalam satu kelompok kategori berdasarkan objek/konsep yang bersangkutan [13].

2.4. Data Set

Data set merupakan kumpulan data yang teroganisir dalam satu kesatuan dan terdiri dari satu atau lebih variable atau atribut. Data numerik, teks, gambar, audio, dan video adalah beberapa jenis data yang dapat termasuk dalam set data. Data set biasanya digunakan dalam analisi data, pembelajaran mesin, dan pengembangan model prediktif. Data set dapat diambil dari berbagai sumber, seperti survey, *database*, *log file*, dan media sosial. Data set yang besar dan kompleks sering memerlukan pengolahan dan penyaringan sebelum dapat digunakan dalam analisi atau pengembangan model [14].

2.5. Algoritma Data Mining

2.5.1. Algoritma C4.5

Algoritma C4.5, metode statistik yang terkenal dan banyak digunakan untuk membuat keputusan dengan membandingkan berbagai fakta, memberikan bahasa yang jelas dan mudah dimengerti. Perhitungan algoritma C4.5 dilakukan dalam 2 tahap, yaitu [15].

1. Entropy

Nilai *entropy* digunakan untuk menghitung jumlah informasi pada atribut dengan menggunakan persamaan berikut:

$$Entropy(S) = -\sum_{i=1}^n p_i \cdot \log_2 p_i \dots \dots (1)$$

Keterangan:

S: himpunan kasus

A: atribut

n: jumlah partisi atribut S

Pi: proporsi dari Si terhadap S

2. Gain

Nilai *Gain* digunakan untuk mengetahui seberapa efektif atribut dalam proses klasifikasi. Ini dapat dilakukan dengan persamaan berikut:

$$Gain(S,A) =$$

$$Entropy(S) - \sum_{i=1}^n \left(\frac{S_i}{S}\right) Entropy(S_i) \dots \dots (2)$$

Keterangan:

S: himpunan kasus

A: atribut

n: jumlah partisi atribut A
 |S_i|: jumlah kasus pada partisi ke-i
 |S|: jumlah kasus dalam S

2.5.2. Confusion Matrix

Didasarkan pada banyak data pengujian yang dapat diprediksi secara benar atau salah, model penelitian metode klasifikasi. *Confusion matrix* memiliki data yang digunakan untuk membandingkan hasil klasifikasi. Persamaan sebagai berikut [16].

$$\text{Akurasi} = \frac{(TP+TN)}{TP+TN+FP+FN} \dots\dots\dots(3)$$

Keterangan:

TP: Jumlah data positif yang diklasifikasi dengan benar oleh komputer
 TN: Jumlah data negatif yang diklasifikasi dengan benar oleh komputer
 FP: Jumlah data negatif yang diklasifikasikan dengan salah oleh komputer
 FN: Jumlah data positif yang diklasifikasikan dengan salah oleh komputer

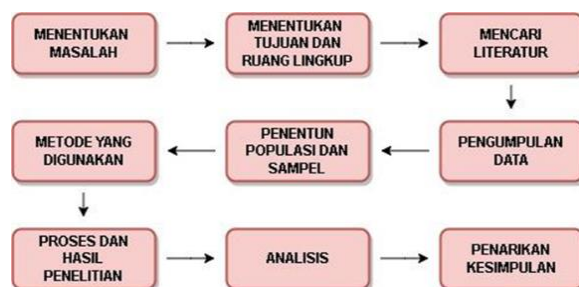
Persamaan berikut dapat digunakan untuk menghitung keakuratan model algoritma, seperti yang ditunjukkan dalam Tabel 1.

Tabel 1. Persamaan *Confusion Matrix*

Kelas Prediksi	Kelas Prediksi	
	Benar	Salah
Benar	TP	FN
Salah	FP	TN

3. METODE PENELITIAN

Metodologi yang menjadi acuan terdiri 9 (Sembilan) tahapan yaitu:



Gambar 2. Metodologi yang Digunakan

3.1. Menentukan Masalah

Masalah ditentukan berdasarkan hasil observasi di lapangan pada tanggal 27 Januari 2024 dengan pegawai di kecamatan Pasi Kolaga Kabupaten Muna.

3.2. Menentukan Tujuan dan Ruang Lingkup

Tujuan dari penelitian ini yaitu untuk menentukan hasil klasifikasi data menggunakan algoritma C4.5 untuk penerima KIS. Hasil ini akan dibandingkan dengan simulasi aplikasi sederhana menggunakan bahasa pemrograman HTML, CSS, JavaScript dan PHP versi 1.91.

Ruang lingkup merupakan batasan dari pembahasan yang dikaji yaitu:

- Data set yang di gunakan berupa KIS tahun 2022, data yang diolah tidak menggunakan data yang bersifat rahasia seperti Nomor Induk Kependudukan (NIK).
- Menggunakan perangkat lunak *RapidMiner* untuk proses pengelolaan data awal, penentuan *class* status penerimaan dan pengujian algoritma C4.5.
- Aplikasi untuk simulasi terdiri dari fitur *search* (pencarian data), pemilihan parameter dalam pohon keputusan, penampilan data *class* status penerimaan dan pengunduhan data hasil pengujian algoritma C.45.

3.3. Mencari Literatur

Mempelajari berbagai sumber referensi, termasuk artikel dan buku teks, yang berkaitan dengan topik yang dibahas mencakup penjelasan mendalam tentang konsep atau teori yang relevan.

3.4. Penentuan Populasi

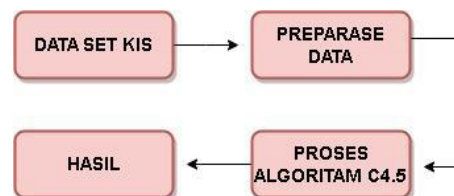
Populasi yang menjadi objek penelitian adalah kecamatan Pasi Kolaga yang terdiri dari 4 desa yaitu Mata Indaha, Lambelu, Tampunabale dan Kolese dengan sampel juga sebanyak 4 desa (Mata Indaha, Lambelu, Tampunabale dan Kolese). Jumlah KK sebanyak 2.117 Kepala Keluarga (KK). Jumlah sampel merupakan jumlah populasi, alasan jumlah sampel sama dengan jumlah populasi dikarenakan permintaan yang merupakan kebutuhan dari tempat studi kasus.

3.5. Pengumpulan Data

Metode pengumpulan data merupakan tahapan yang paling strategis dalam proses penelitian dan tujuan utamanya adalah memperoleh data. Teknik pengumpulan data dengan cara melakukan wawancara (*interview*) dan dokumen/studi Pustaka (*document*).

3.6. Metode yang Digunakan

Metode yang digunakan dalam pengelolaan *data mining* terdiri dari 4 tahapan yaitu:



Gambar 3. Metode yang Digunakan

3.7. Proses dan Hasil Penelitian

Algoritma C4.5 digunakan dalam proses pengolah data yang dilakukan oleh perangkat lunak *RapidMiner*. Nilai klasifikasi terbaik akan digunakan sebagai referensi untuk menentukan kelayakan penerima KIS pada titik tertentu. Proses algoritma C4.5 akan disimulasikan dalam aplikasi sederhana.

3.8. Analisis Hasil Penelitian

Analisis dilakukan terhadap data hasil proses algoritma C4.5 yang akan dijadikan sebagai masukan/saran alternatif solusi kepada para pemangku kepentingan di tempat studi kasus. Masukan yang diajukan dapat dijadikan sebagai bahan pertimbangan bagi pimpinan dalam membuat keputusan/kebijakan dalam upaya menyelesaikan permasalahan penerima KIS yang terjadi di kantor Kecamatan Pasi Kolaga.

3.9. Penarikan Kesimpulan

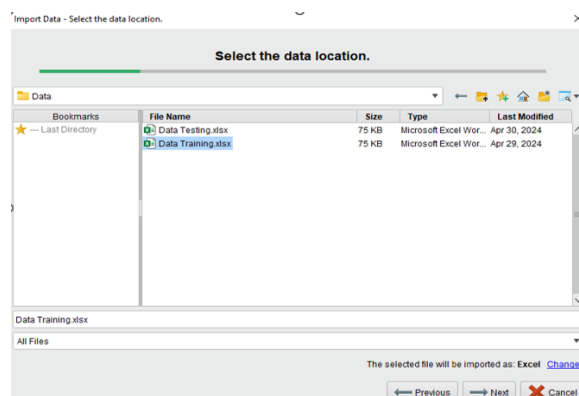
Kesimpulan dibuat dengan membandingkan proses uji coba pada *rapidminer* dengan hasil aplikasi yang dibuat.

4. HASIL DAN PEMBAHASAN

4.1. Hasil

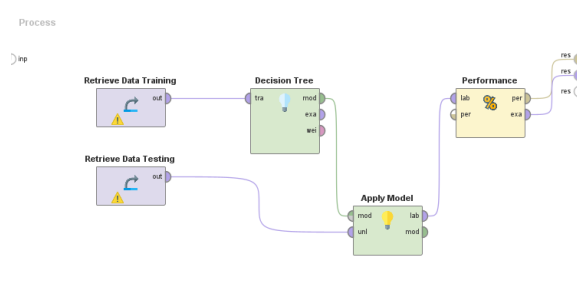
4.1.1. Hasil Algoritma C4.5

Gambar 4 adalah hasil *import* data yang telah dipreparasi menggunakan *rapidminer*.



Gambar 4. Import Data yang Telah Dipreparasi

Gambar 5 dan 6 merupakan hasil dari proses pengujian menggunakan algoritma C4.5 dari 2.117 yang diuji menghasilkan data diterima sebanyak 1916 dan data ditolak sebanyak 201 data.



Gambar 5. Proses Algoritma C4.5

Name	Type	Missing	Statistics	Filter (11 / 11 attributes)
Label	Binomial	0	Statistics: Diterima, Ditolak	Values: Diterima (1916), Ditolak (201)
prediction(Status Penerima)	Binomial	0	Statistics: Diterima, Ditolak	Values: Diterima (1916), Ditolak (141)
confidence(Diterima)	Real	0	Min: 0, Max: 1, Average: 0.925	
confidence(Ditolak)	Real	0	Min: 0, Max: 1, Average: 0.074	
Nama	Polynomial	0	Lower: P959 (1), Upper: P1 (1)	Values: P1 (1), P10 (1), ... (2115 more)
Alamat	Polynomial	0	Lower: 1 (120), Upper: 2 (951)	Values: 2 (951), 4 (947), ... (7 more)
Jenis Kelamin	Binomial	0	Statistics: L, P	Values: L (2055), P (62)
Pendidikan	Polynomial	0	Lower: 4 (180), Upper: 3 (1042)	Values: 3 (1042), 2 (150), ... (2 more)

Gambar 6. Hasil Data Pengujian Algoritma C4.5

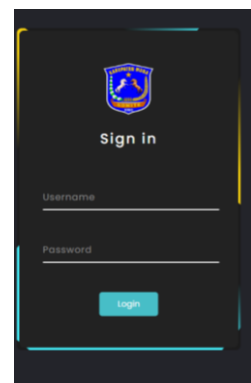
Gambar 7 hasil tingkat akurasi sebesar 97,92%, dimana *class precision* diterima sebesar 100%, *class precision* ditolak sebesar 78,11%, *class recall* diterima sebesar 97,76% dan *class recall* ditolak sebesar 100%.

accuracy: 99.15%			
	true Diterima	true Ditolak	class precision
pred. Diterima	1959	17	99.14%
pred. Ditolak	1	140	99.29%
class recall	99.95%	89.17%	

Gambar 7. Tingkat Akurasi Pengujian Algoritma C4.5

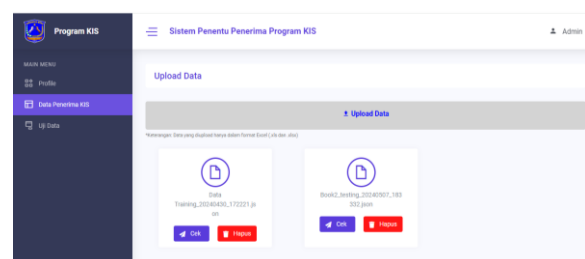
4.1.2. Hasil Pada Aplikasi

Halaman *login*, halaman untuk autentikasi berdasarkan *username* dan *password* untuk masuk ke dalam aplikasi (Gambar 8).



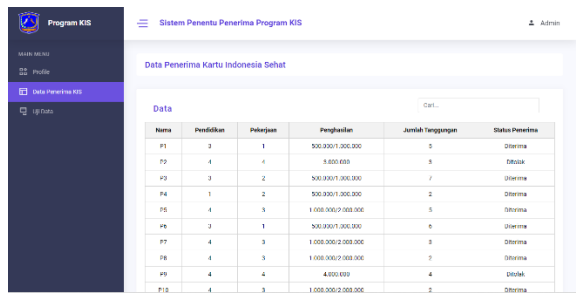
Gambar 8. Halaman Login

Gambar 9 merupakan halaman *upload* atau *import* data kedalam aplikasi berupa data excel (xlsx atau csv).



Gambar 9. Halaman Upload Data

Gambar 10 halaman uji data terhadap data yang telah di *upload*.



Nama	Pendidikan	Pekerjaan	Penghasilan	Jumlah Tanggungan	Status Penerima
P1	2	1	500.000/1.000.000	5	Diterima
P2	4	4	1.000.000	5	Ditolak
P3	2	2	500.000/1.000.000	7	Diterima
P4	1	2	500.000/1.000.000	2	Diterima
P5	4	3	1.000.000/2.000.000	5	Diterima
P6	2	1	500.000/1.000.000	6	Diterima
P7	4	3	1.000.000/2.000.000	3	Diterima
P8	4	3	1.000.000/2.000.000	7	Diterima
P9	4	4	4.000.000	4	Ditolak
P10	4	3	1.000.000/2.000.000	2	Diterima

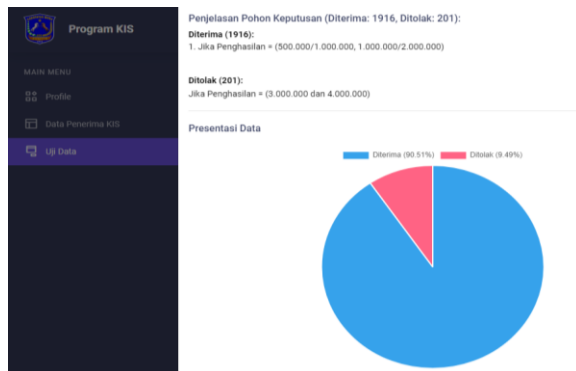
Gambar 10. Halaman Uji Data

Gambar 11 halaman yang menampilkan data pohon keputusan dari data yang diuji sesuai dengan *field* yang telah dipilih.



Gambar 11. Halaman Uji Data Pohon Keputusan

Gambar 12 adalah penjelasan dari pohon keputusan, menjelaskan *value* dari tabel yang telah diproses yang masuk kedalam klasifikasi diterima dan ditolak dan juga menampilkan jumlah datanya.



Gambar 12. Penjelasan Pohon Keputusan

4.2. Pembahasan

Hasil dari penentuan status *class* penerimaan menggunakan algoritma C4.5 *class* diterima diperoleh hasil 1916 data dan *class* ditolak 201 data, sedangkan hasil dari pengujian pada aplikasi *class* Diterima diperoleh 1916 data dan *class* ditolak 201. Dari hasil perbandingan menunjukkan tingkat akurasi sama baik menggunakan *rapidminer* maupun aplikasi. Tabel 2 menampilkan 10 data dari 2117 data.

Tabel 2. Evaluasi Perbandingan Hasil

Hasil Status <i>Class</i> Penerimaan			
Perangkat lunak (<i>RapidMiner</i>)		Aplikasi	
Nama	Status Penerima	Nama	Status Penerima
P1	Diterima	Penerima 1	Diterima
P2	Ditolak	Penerima 2	Ditolak
P3	Diterima	Penerima 3	Diterima
P4	Diterima	Penerima 4	Diterima
P5	Diterima	Penerima 5	Diterima
P6	Diterima	Penerima 6	Diterima
P7	Diterima	Penerima 7	Diterima
P8	Diterima	Penerima 8	Diterima
P9	Ditolak	Penerima 9	Ditolak
P10	Diterima	Penerima 10	Diterima
Hasil Akhir			
Perangkat lunak (<i>RapidMiner</i>)		Aplikasi	
Diterima	Ditolak	Diterima	Ditolak
1916	201	1916	201

Klasifikasi data baik menggunakan perangkat lunak *rapidminer* maupun aplikasi yang telah dibangun menampilkan hasil yang sama, hal ini mengindikasikan bahwa aplikasi dapat digunakan oleh petugas di Kecamatan Pasi Kolaga dalam mempermudah pengelolaan data penerima bantuan KIS sehingga penerima bantuan akan tepat sasaran.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini adalah penerapan algoritma C4.5 menggunakan perangkat lunak *rapidminer* dan implementasi pada aplikasi menghasilkan data yang sama, sehingga aplikasi dapat digunakan untuk mengurangi kesalahan pengolahan data. Aplikasi dapat digunakan sebagai alat bantu bagi perangkat desa untuk mempermudah pengelolaan data bantuan KIS sehingga lebih efektif dan efisien.

Penambahan fitur edit dan hapus data serta format *import* data selain xls merupakan masukan pada kajian selanjutnya.

DAFTAR PUSTAKA

- [1] J. Saintikom *et al.*, "Klasifikasi Masyarakat Miskin Layak Menerima Bantuan Sosial Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)," vol. 20, pp. 167–173, 2021.
- [2] R. Girsang, E. F. Ginting, and M. Hutasuht, "Penerapan Algoritma C4.5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 449, 2022, doi: 10.53513/jursi.v1i4.5727.
- [3] F. Erwanto, A. H. Mirza, and E. Supratman, "T Penerapan Algoritma C45 Untuk Penentuan Kelayakan," *Bina Dharma Conf. Comput. Sci. 2019*, pp. 284–291, 2019.
- [4] L. Siliyani, Iqbal Agis Junizar, Uyu Nuraeni, Edi Tohidi, and Irfan Ali, "Penerapan Algoritma

- Naive Bayes Untuk Mengetahui Kepuasan Mahasiswa Terhadap Layanan Administrasi Keuangan,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 4, no. 3, pp. 72–79, 2020, doi: 10.32485/kopertip.v4i3.122.
- [5] M. Maliangga, E. N. Walewangko, and A. T. Londa, “Pengaruh Kebijakan Pemerintah Kartu Indonesia Pintar (Kip) Dan Kartu Indonesia Sehat (Kis) Terhadap Konsumsi Rumah Tangga Miskin Di Kecamatan Dumoga Timur Kabupaten Bolaang Mongondow,” *J. Berk. Ilm. Efisiensi*, vol. 19, no. 01, pp. 32–43, 2019.
- [6] M. I. A. Aziz, “Implementasi Kebijakan Kartu Indonesia Sehat di Kecamatan Seberang Ulu I Kota Palembang,” *Univ. Sriwij.*, p. 43, 2019.
- [7] Siregar et al, *DATA MINING: Pengolahan Data Menjadi Informasi dengan RapidMiner*. 2016. [Online]. Available: <https://books.google.co.id>
- [8] Sekretaris Negara RI, “UU RI No.40 Tahun 2004 Tentang Sistem Jaminan Sosial Nasional,” *Jdih BPK RI*, pp. 1–45, 2004.
- [9] P. D. M. I. dengan RapidMiner, *No Title*. 2020. [Online]. Available: <https://books.google.co.id>
- [10] E. R. Batubara, P. P. P. A. N. . F. I. R.H.Zer, F. N. Tambunan, and R. Rosnelly, “Pengenalan Pola Kemampuan Pelanggan Dalam Membayar Air Pdam Menggunakan Algoritma Naive Bayes,” *Jurnaltimes*, vol. X, no. 2, pp. 29–38, 2022.
- [11] L. Elvitaria, “Memprediksi Tingkat Peminat Ekstrakurikuler Pada Siswa Smk Analisis Kesehatan Abdurrah Menggunakn Algoritma C4.5 (Studi Kasus: Smk Analis Kesehatan Abdurrah),” *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 2, no. 2, pp. 220–233, 2019, doi: 10.36341/rabit.v2i2.212.
- [12] Iis Dewi Ratih, S M Retnaningsih, and V M Dewi, “Klasifikasi Kualitas Tanah Menggunakan Metode Naive Bayes Classifier,” *J. Apl. Mat. dan Stat.*, vol. 1, no. 1, pp. 11–20, 2022, doi: 10.53625/jams.v1i1.4427.
- [13] S. Haryati, A. Sudarsono, and E. Suryana, “Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu),” *J. Media Infotama*, vol. 11, no. 2, pp. 130–138, 2015.
- [14] M. Yunus, H. Ramadhan, D. R. Aji, and A. Yulianto, “Penerapan Metode Data Mining C4.5 Untuk Pemilihan Penerima Kartu Indonesia Pintar (KIP),” *Paradig. - J. Komput. dan Inform.*, vol. 23, no. 2, 2021, doi: 10.31294/p.v23i2.11395.
- [15] E. Elisa and A. A. Fajrin, “Data Mining dalam Menganalisis Faktor Alasan Pemilihan Perumahan,” *Pros. Semin. Nas. Ilmu ...*, no. September, pp. 43–48, 2020, [Online]. Available: <http://ejournal.upbatam.ac.id/index.php/prosidin g/article/view/3611>
- [16] R. A. Saputra, S. Wasiyanti, and D. Pribadi, “Information Gain Pada Algoritma C4.5 Untuk Klasifikasi Penerimaan Bantuan Pangan Non Tunai (Bpnt),” *Indones. J. Bus. Intell.*, vol. 4, no. 1, p. 25, 2021, doi: 10.21927/ijubi.v4i1.1757.