

Klasifikasi Penerima Bantuan Beras Miskin Menggunakan Algoritma K-NN, NBC dan C4.5

Andani Putri Pristiawati*, Inggih Permana, Zarnelly, Fitriani Muttakin

Fakultas Sains dan Teknologi, Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹11950324487@students.uin-suska.ac.id, ²inggihpermana@uin-suska.ac.id, ³zarnelly@uin-suska.ac.id, ⁴fitriani.muttakin@uin-suska.ac.id

Email Penulis Korespondensi: 11950324487@students.uin-suska.ac.id

Submitted: 12/06/2023; Accepted: 29/06/2023; Published: 29/06/2023

Abstrak—Salah satu tugas Dinas sosial Kota Dumai adalah memberikan bantuan beras miskin kepada masyarakat yang membutuhkan. Permasalahan yang sering terjadi dalam pemberian beras miskin tersebut adalah sering terjadi salah sasaran penerima beras miskin. Dalam mengatasi permasalahan yang ada penelitian ini telah melakukan model- model klasifikasi dengan menggunakan algoritma *K-Nearest Neighbor* (K-NN), *Naïve Bayes Classifier* (NBC) dan Algoritma C4.5. Berdasarkan hasil percobaan didapatkan bahwa model klasifikasi terbaik dihasilkan oleh Algoritma K-NN dengan nilai K sama dengan 21 dengan nilai akurasi 79,00%, percobaan yang didapatkan pada model klasifikasi algoritma NBC dengan nilai akurasi 84,10% dan model klasifikasi algoritma C4.5 didapatkan akurasi 88,36%. Disamping itu algoritma C4.5 berhasil membuat sebuah pohon keputusan untuk model klasifikasi dengan kompleksitas terendah karena berhasil mereduksi atribut jumlah dari 33 menjadi 5 atribut.

Kata Kunci: Beras Miskin; C4.5; Klasifikasi; K-Nearest Neighbor; Naïve Bayes Classifier

Abstract—One of the tasks of the Dumai City Social Service is to provide poor rice assistance to people in need. The problem that often occurs in the distribution of rice to the poor is that the target recipients of poor rice often occur. In overcoming the existing problems, this research has carried out classification models using the K-Nearest Neighbor (K-NN) algorithm, Naïve Bayes Classifier (NBC), and C4.5 Algorithm. Based on the experimental results, it was found that the best classification model was produced by the K-NN Algorithm with a value of K equal to 21. Besides that, the C4.5 algorithm succeeded in making a decision tree for the classification model with the lowest complexity because it succeeded in reducing the number of attributes from 33 to 5 attributes. The decision tree can be used as material for consideration to the Social Service in making decisions on Raskin beneficiaries.

Keywords: Poor Rice; C4.5; Classification; K-Nearest Neighbor; Naïve Bayes Classifier

1. PENDAHULUAN

Dinas Sosial Kota Dumai Provinsi Riau adalah salah satu unsur penyelenggara pemerintah Kota Dumai yang tercatat pada Peraturan Daerah Kota Dumai Nomor 12 Tahun 2016 yang berisi tentang Pembentukan dan Susunan Perangkat Daerah Kota Dumai. Untuk menjalankan tugasnya Dinas Sosial Kota Dumai melakukan pemberian bantuan kepada rumah tangga sasaran berupa bantuan secara tunai maupun non tunai. Bantuan sosial biasanya diberikan oleh pemerintah atau organisasi nirlaba kepada kelompok masyarakat yang memiliki keterbatasan ekonomi atau sosial. Tujuan utama dari bantuan sosial adalah untuk membantu masyarakat yang membutuhkan agar dapat mencukupi kebutuhan dasar diantaranya makanan, pakaian, tempat tinggal, dan pendidikan [1], [2].

Bantuan beras miskin (raskin) merupakan program bantuan Dinas Sosial yang ditujukan kepada masyarakat kurang mampu guna untuk mencukupi kebutuhan pokok, terutama dalam hal pangan. Bantuan beras miskin biasanya diperuntukan kepada keluarga penerima manfaat (KPM) yang sudah ditetapkan oleh pemerintah melalui sistem klasifikasi berdasarkan kriteria sosial ekonomi [3]. Sehingga dengan adanya program ini Dinas Sosial menyiapkan anggaran beras untuk masyarakat miskin 15 kg/keluarga/bulan. Penerima bantuan raskin harus membayar dengan harga Rp. 1.600/kg pada titik distribusi [4], [5], sisanya ditanggung oleh pemerintah. Sehingga selisih antara harga normal beras yang seharusnya dibayar dan harga yang sesungguhnya beras dibayar oleh masyarakat miskin menjadi jumlah yang ditanggung oleh pemerintah per kg [6],[7].

Pelaksanaan penyaluran raskin terdapat permasalahan yang terjadi dari pihak pemerintah maupun dari pihak KPM. Permasalahan dari pihak pemerintah adalah keterlambatan penyaluran raskin, kurang tepatnya KPM, kelangkaan stok beras, kualitas beras yang buruk dan penyalahgunaan pemberian raskin [8],[9]. Sedangkan, permasalahan dari pihak KPM adalah kesadaran yang dimiliki masyarakat masih rendah akan apa yang menjadi haknya atau bukan, masyarakat yang menerima raskin banyak yang tidak berhak tetapi malah menjadi menerima raskin. Selain itu banyak juga masyarakat yang mendapatkan raskin menjual kembali beras tersebut kepada orang lain, untuk digantikan ke bahan pokok lainnya, atau menjualnya kembali karena rendahnya kualitas raskin. Salah sasaran pemberian raskin menjadi permasalahan utama bagi pemerintah dalam melaksanakan penyaluran raskin [10], [11].

Untuk mengatasi permasalahan yang telah dijabarkan sebelumnya, maka dari itu penelitian ini membuat model klasifikasi untuk menentukan apakah sebuah keluarga layak atau tidak menerima bantuan raskin. Algoritma yang digunakan untuk membuat model klasifikasi dalam penelitian ini adalah *K-Nearest Neighbor* (K-NN), *Naïve Bayes Classifier* (NBC) dan Algoritma C4.5 [12], [13]. Algoritma – algoritma tersebut dipilih karena algoritma tersebut telah banyak digunakan untuk kasus klasifikasi penerima bantuan sosial. Perfoma dari model – model klasifikasi tersebut dibandingkan dan diambil model klasifikasi dengan perfoma terbaik. Model klasifikasi yang didapatkan diharapkan meningkatkan efektivitas dalam pengambilan keputusan berdasarkan data yang telah terkumpul [14][15].

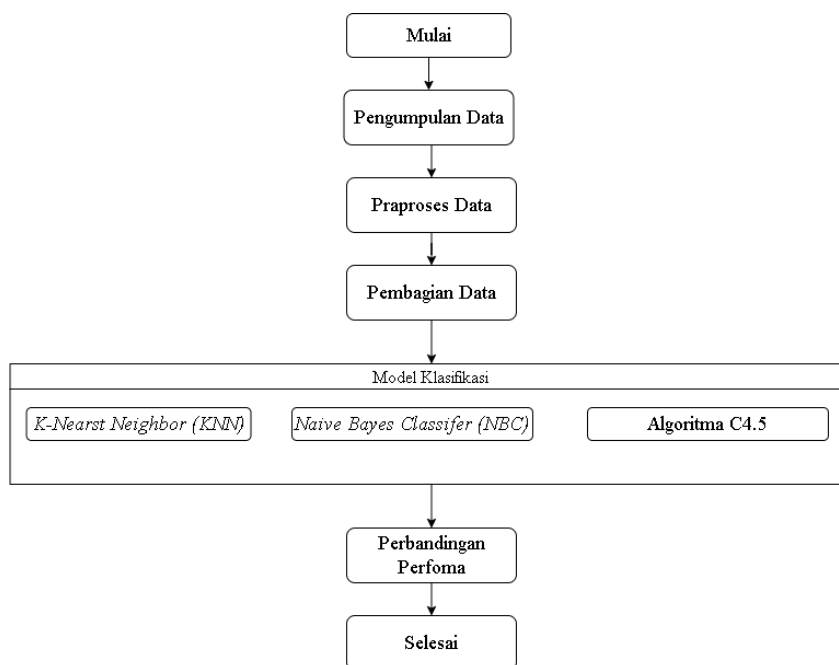
Beberapa penelitian sebelumnya telah membuat model klasifikasi untuk calon penerima raskin, seperti Hidayatullah (2021) [16], dalam penelitian tersebut dilakukan uji kelayakan penerima raskin menggunakan NBC, penelitian tersebut menggunakan 205 baris data dengan 6 atribut. Akurasi yang didapat dari penelitian tersebut adalah 96,59%.. Selanjutnya terdapat penelitian yang dilakukan oleh Nurahman dan Aminah (2022) [17], dalam penelitian tersebut menggunakan algoritma C4.5 menggunakan data sebanyak 100 baris data dan menunjukkan akurasi sebesar 93,00%.

Adapun terdapat perbedaan pada penelitian ini dengan penelitian sebelumnya, pada penelitian yang sudah dilakukan sebelumnya hanya menggunakan satu buah algoritma saja, sedangkan penelitian ini akan membandingkan tiga buah algoritma yaitu K-NN, NBC dan C4.5. Penelitian sebelumnya [13] hanya menggunakan 6 atribut sedangkan penelitian ini menggunakan 33 atribut. Perbedaan selanjutnya penelitian ini menggunakan jumlah baris data yang lebih banyak dari penelitian sebelumnya [14],[16].

Paper ini disusun dalam empat buah bab. Bab kedua berisi metodologi penelitian yang digunakan untuk menyelesaikan penelitian ini. Bab tiga berisi hasil dan pembahasan dari percobaan yang telah dilakukan dan bab terakhir berisi kesimpulan dari penelitian ini. Sehingga dengan adanya penelitian ini pihak pemerintah dapat memilih dengan tepat masyarakat yang memenuhi kriteria untuk menerima bantuan raskin.

2. METODOLOGI PENELITIAN

Penelitian ini terdiri dari 5 tahapan diantaranya adalah tahap pertama adalah pengumpulan data, tahap kedua adalah tahap praproses data, tahapan ketiga adalah tahap pembagian data, tahapan keempat adalah tahapan pembuatan model klasifikasi dan tahapan terakhir adalah tahap hasil perbandingan perfoma. Alur tahapan pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Metodologi Penelitian

2.1 Pengumpulan Data

Penelitian ini menggunakan data bantuan sosial yaitu Beras Miskin (raskin), sumber data pada penelitian ini adalah Data Terpadu Kesejahteraan Sosial (DTKS) tahun 2022 pada Dinas Sosial Kota Dumai Provinsi Riau di Kelurahan Pangkalan Sesai Kota Dumai.

2.2 Praproses Data

2.2.1 Pembersihan Data

Pada proses pembersihan ini hal yang dilakukan adalah menghilangkan baris data yang mengandung nilai NULL. pembersihan data dilakukan dengan menggunakan operator *Filter Example* pada aplikasi RapidMiner Studio 10.1.

2.2.2 Normalisasi Data

Penelitian ini melakukan normalisasi data menggunakan *Min-Max Normalization*. Nilai minimum yang digunakan adalah 0 dan maksimal yang digunakan adalah 1. Normalisasi data dilakukan dengan operator *Normalize* pada aplikasi RapidMiner Studio 10.1.

2.3 Pembagian Data

Data uji dan data latih melakukan pembagian data dengan menggunakan *K-Fold Cross Validation*. Penelitian ini dilakukan dengan percobaan $K=10$. Pengujian *K-Fold* pada penelitian ini menggunakan operator model *Cross Validation* pada aplikasi RapidMiner Studio 10.1 [18].

2.4 Pembuatan Model Klasifikasi

2.4.1 *K-Nearst Neighbor (K-NN)*

Parameter k-NN pada penelitian ini dapat dilihat pada Tabel 1. Tools yang digunakan adalah RapidMiner Studio 10.1. Operator RapidMiner yang digunakan adalah operator k-NN. Pada model k-NN mengaplikasikan parameter *measure types* yaitu *mixed measure* dan *mixed measure* yang digunakan adalah *mixed euclidean distance*.

Tabel 1. Parameter k-NN

No	Parameter	Nilai
1	Perhitungan jarak terdekat	Euclidean distance
2	Nilai k	K=3, K=5, K=7, K=9, K=11, K=13, K=15, K=17, K=19, K=21

2.4.2 *Naïve Bayes Classifier*

Penelitian ini melakukan model klasifikasi pada aplikasi RapidMiner Studio 10.1 dengan menggunakan operator model *Naïve bayes* [19], pada proses ini menggunakan parameter *laplace correction* untuk menangani nilai pada probabilitas 0.

2.4.3 *Algoritma C4.5*

Pengujian pada model klasifikasi C4.5 dilakukan menggunakan operator model *Decision Tree* pada aplikasi RapidMiner Studio 10.1 dengan menggunakan parameter seperti pada Tabel 2. Selanjutnya dilakukan proses *apply model* dan *performance* untuk mendapatkan hasil perfoma dan pohon keputusan.

Tabel 2. Parameter C4.5

No	Parameter	Nilai
1	Criterion	Information Gain
2	Maximal Depth	5
3	Apply Pruning	Confidence= 0,5

2.5 Perbandingan Perfoma

Penelitian ini dalam melakukan pengukuran perfoma menggunakan *Confusion matrix*. *Confusion Matrix* dapat menghasilkan nilai *Accuracy*, *Precision* dan *Recall*. Langkah yang dilakukan untuk mendapatkan nilai tersebut dengan menggunakan operator model *Performance* pada RapidMiner Studio 10.1. Setelah mendapatkan hasil pengukuran perfoma maka dilakukan perbandingan dari tiga model akurasi algoritma k-NN, NBC dan C4.5. Adapun akurasi pada persamaan 1, recall pada persamaan 2 dan presisi pada persamaan 3 [20], [21].

$$Accuracy = \left(\frac{TP+TN}{TP+TN+FP+FN} \right) x \quad (1)$$

$$Recall = \left(\frac{TP}{TP+FN} \right) \quad (2)$$

$$Precision = \left(\frac{TP}{TP+FP} \right) \quad (3)$$

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan Data

Hasil perolehan data yang didapatkan dari pengumpulan data di Dinas Sosial Kota Dumai adalah dalam format *Microsoft Excel*. Pada data tersebut terdapat 585 baris data calon penerima bantuan beras miskin pada Kelurahan Pangkalan Sesai. Data tersebut terdiri dari 33 atribut sebagai kriteria penentu kelayakan penerima bantuan raskin. Pada data tersebut terdiri dari 388 keluarga yang berhak menerima bantuan raskin dan 196 keluarga yang tidak berhak menerima bantuan raskin. Untuk lebih detailnya atribut-atribut data dapat dilihat Tabel 3 dan sampel data raskin dapat dilihat pada Tabel 4.

Tabel 3. Atribut- Atribut Data Raskin

Kode	Atribut	Keterangan	Nilai
------	---------	------------	-------



A1	Jumlah ART	Jumlah anggota rumah tangga	Menyesuaikan masing-masing anggota keluarga pada setiap rumah.
A2	Sta Bangunan	Status penguasaan bangunan tempat tinggal yang ditempati	1. Milik sendiri; 2. Kontrak/sewa; 3. Bebas sewa ; 4. Dinas ; 5. Lainnya
A3	Sta Lahan	Status lahan tempat tinggal yang ditempati	1. Milik sendiri; 2. Milik orang lain; 3. Tanah negara; 4. Lainnya
A4	Lantai	Jenis lantai terluas	1.Marmmergranit; 2. Keramik; 3. Parket/vinil/permadani; 4. Ubin/tegel/teraso; 5. Kayu/papan kualitas tinggi; 6.Semestara/bata merah; 7. Bambu; 8.Kayu/papan berkualitas rendah; 9. Tanah; 10. Lainnya
A5	Dinding	Jenis dinding terluas	1. Tembok; 2. Plasteran anyaman bambu/kawat; 3. Kayu; 4. Anyaman bambu; 5. Batang Kayu; 6. Bambu; 7. Lainnya;
A6	Kondisi Dinding	Kondisi dinding terluas	1. Bagus/kualitas tinggi; 2. Jelek/kualitas rendah
A7	Atap	Jenis atap terluas	1. Beton/genteng beton; 2. Genteng keramik; 3. Genteng metal; 4. Genteng tanah liat; 5. Asbes; 6. Seng; 7. Sirap; 8, Bambu; 9. Jerami/ijuk/daundaunan/rumbia; 10. Lainnya
A8	Kondisi Atap	Kondisi atap terluas	1. Bagus/kualitas tinggi; 2. Jelek/kualitas rendah
A9	Jumlah Kamar	Jumlah kamar tidur	Menyesuaikan jumlah kamar pada masing-masing rumah.
A10	Sumber Air Minum	Sumber air minum	1. Air kemasan bermerk; 2. Air isi ulang; 3. Leding meteran; 4. Leding eceran; 5. Sumur bor/pompa; 6. Sumur terlindung; 7. Sumur tak terlindung; 8. Mata air terlindung; 9. Mata air tak terlindung; 10.Airsungai/danau/waduk; 11. Air hujan; 12. Lainnya
A11	Sumber Penerangan	Sumber penerangan utama	1. Listrik PLN; 2. Listrik non PLN ;3. Bukan listrik
A12	Daya	Daya terpasang	1. 450 watt; 2. 900 watt; 3. 1.300 watt; 4. 2.200 watt; 5. > 2.200 watt; 6. Tanpa meteran
A13	BB Masak	Bahan bakar/energi utama untuk memasak	1. Listrik; 2. Gas > 3 kg; 3. Gas 3 kg; 4. Gas kota/biogas; 5. Minyak tanah; 6. Briket; 7. Arang; 8. Kayu bakar ; 9. Tidak memasak dirumah
A14	Fasbab	Penggunaan fasilitas tempat BAB	1. Sendiri; 2. Bersama; 3. Tidak ada
A15	Kloset	Jenis kloset	1. Leher angsa; 2. Plengsengan; 3. Cemplung/cubluk; 4. Tidak pakai
A16	Buang Tinja	Tempat pembungan akhir tinja	1. Tangki; 2. Plengsengan; 3. Lubang tanah; 4.Kolam/sawah/sungai/danau/laut; 5.Pantai/tanah lapang/kebun; 6. Lainnya
A17	Tabung Gas	Memiliki tabung gas 5,5 kg atau lebih	1. Ya 2. Tidak
A18	Lemari Es	Memiliki lemari es/kulkas	1. Ya 2. Tidak
A19	AC	Memiliki AC	1. Ya 2. Tidak
A20	Pemanas air	Memiliki pemanas air (water heater)	1. Ya 2. Tidak
A21	Telepon	Memiliki telepon rumah (PSTN)	1. Ya 2. Tidak
A22	TV	Memiliki televisi	1. Ya 2. Tidak

A23	Emas	Memiliki emas/perhiasan & tabungan (senilai 10 gram 10 gram emas)	1. Ya	2. Tidak
A24	Komputer/Laptop	Memiliki komputer/laptop	1. Ya	2. Tidak
A25	Sepeda	Memiliki sepeda	1. Ya	2. Tidak
A26	Motor	Memiliki sepeda motor	1. Ya	2. Tidak
A27	Mobil	Memiliki mobil	1. Ya	2. Tidak
A28	Perahu	Memiliki perahu	1. Ya	2. Tidak
A29	Motor Tempel	Memiliki motor tempel	1. Ya	2. Tidak
A30	Kapal	Memiliki kapal	1. Ya	2. Tidak
A31	Aset Tidak Bergerak	Rumah tangga memiliki aset tidak bergerak	a.Lahan 1. Ya 3. Ya	2. Tidak 4. Tidak
A32	Rumah Lain	Rumah tangga memiliki aset rumah di tempat lain	1. Ya	2. Tidak
A33	Sta Art Usaha	Apakah ada ART yang memiliki usaha sendiri/milik bersama	1. Ya	2. Tidak
Label/Class	Status	Atribut label yang bernilai kelas untuk calon penerima Raskin. Apabila bernilai Yes maka kreteria penerima dan apabila bernilai No maka kreteria tidak menerima	Yes	No

Tabel 4. Sampel Data Raskin

Jumlah ART	sta bangunan	sta lahan	...	rumah lain	sta art usaha	Status
5	2	2	...	4	1	NO
4	NULL	NULL	...	NULL	NULL	NO
1	NULL	NULL	...	NULL	NULL	YES
1	NULL	NULL	...	NULL	NULL	YES
1	NULL	NULL	...	NULL	NULL	YES
1	NULL	NULL	...	NULL	NULL	YES
4	2	2	...	4	2	YES
...
2	1	1	...	4	2	YES
4	1	1	...	4	2	NO
4	2	2	...	4	2	YES

3.2 Hasil Praproses Data

Pada data awal terdapat 180 baris data yang mengandung nilai NULL, baris data tersebut kemudian dihapus sehingga tersisa 405 data. Dari sisa data tersebut terdapat 283 baris data dengan label Yes (berhak menerima raskin) dan 122 baris data dengan labek No (tidak berhak menerima raskin). Sampel data pembersihan dapat dilihat pada Tabel 5.

Tabel 5. Sampel Hasil Pembersihan Data

No	A1	A2	A3	A4	A5	A6	...	A28	A29	A30	A31	A32	A33	Status
1	5	2	2	6	3	2	...	4	2	4	2	4	1	NO
2	4	2	2	6	3	2	...	4	2	4	2	4	2	YES
3	2	1	1	2	1	2	...	4	2	4	1	4	2	YES
4	4	1	1	8	3	2	...	4	2	4	1	4	2	NO
5	4	2	2	6	1	2	...	4	2	4	2	4	2	YES
...	NO
401	7	2	2	6	1	2	...	4	2	4	2	4	1	NO
402	7	1	1	2	1	1	...	4	1	4	1	4	2	NO
403	5	1	1	6	1	1	...	4	2	4	1	4	1	YES
404	1	5	4	2	1	2	...	4	2	4	2	4	2	YES
405	3	1	1	6	1	2	...	4	2	4	1	4	2	YES

Setelah melakukan pembersihan data dilakukan normalisasi *min-max*. Sampel data normalisasi dapat dilihat pada Tabel 6.



Tabel 6. Sampel Hasil Normalisasi Data Raskin

No	A1	A2	A3	A4	A5	A6	...	A28	A29	A30	A31	A32	A33	Status
1	0,4	0,3	0,3	0,7	1,0	1,0	...	1,0	1,0	1,0	1,0	1,0	0,5	NO
2	0,3	0,3	0,3	0,7	1,0	1,0	...	1,0	1,0	1,0	1,0	1,0	1,0	YES
3	0,1	0,0	0,0	0,1	0,0	1,0	...	1,0	1,0	1,0	0,0	1,0	1,0	YES
4	0,3	0,0	0,0	1,0	1,0	1,0	...	1,0	1,0	1,0	0,0	1,0	1,0	NO
5	0,3	0,3	0,3	0,7	0,0	1,0	...	1,0	1,0	1,0	1,0	1,0	1,0	YES
...
401	0,7	0,3	0,3	0,7	0,0	1,0	...	1,0	1,0	1,0	1,0	1,0	0,5	NO
402	0,7	0,0	0,0	0,1	0,0	0,0	...	1,0	0,0	1,0	0,0	1,0	1,0	NO
403	0,4	0,0	0,0	0,7	0,0	0,0	...	1,0	1,0	1,0	0,0	1,0	0,5	YES
404	0,0	1,0	1,0	0,1	0,0	1,0	...	1,0	1,0	1,0	1,0	1,0	1,0	YES
405	0,2	0,0	0,0	0,7	0,0	1,0	...	1,0	1,0	1,0	0,0	1,0	1,0	YES

3.3 Hasil Pembuatan Model Klasifikasi

3.3.1 K-Nearest Neighbor

Hasil dari percobaan menggunakan 10 buah K dapat dilihat pada Tabel 7.

Tabel 7. Hasil Perfoma K-NN

Perfoma k-NN				
No	K	Accuracy	Precision	Recall
1	3	75,33%	79,37%	87,69%
2	5	78,27%	80,53%	91,23%
3	7	77,06%	79,75%	90,53%
4	9	78,28%	78,93%	94,41%
5	11	78,28%	78,42%	95,45%
6	13	77,53%	77,43%	95,17%
7	15	77,55%	77,85%	95,15%
8	17	77,52%	77,61%	95,82%
9	19	79,00%	78,48%	96,86%
10	21	76,29%	75,92%	97,20%

Berdasarkan Tabel 7 dapat dilihat bahwa akurasi tertinggi terjadi ketika nilai K pada K-NN sama dengan 19, yaitu dengan akurasi 79,00%. Presisi tertinggi terjadi Ketika K pada K-NN sama dengan 5, yaitu dengan nilai presisi 80,53%, sedangkan recall tertinggi ketika nilai K pada K-NN sama dengan 21, yaitu dengan nilai recall 97,20%. Karena akurasi, presisi dan recall tertinggi berada pada nilai K yang berbeda, maka penelitian ini cenderung memilih K-NN dengan nilai K pada recall terbaik yaitu 21. Hal ini karena pada model klasifikasi K=21 bisa meminimalkan kesalahan calon yang berhak menerima menjadi tidak berhak menerima. Hal tersebut dapat terlihat dari nilai recall yang tinggi. Hasil confusion matrix dari K=21 dapat dilihat pada Tabel 8.

Tabel 8. Confusion Matrix k-NN

Confusion Matrix		
	true No	true Yes
pred. No	34	8
Pres. Yes	88	275

3.3.2 Naïve Bayes Classifier

Hasil perfoma dari pengujian model NBC dapat dilihat pada tabel 9 dan *confusion matrix* dapat dilihat pada tabel 10.

Tabel 9. Hasil Perfoma NBC

Perfoma NBC			
	Accuracy	Precision	Recall
NBC	84,10%	88,12%	89,42%

Tabel 10. Confusion Matrix NBC

Confusion Matrix		
	true No	true Yes
pred. No	80	28
Pres. Yes	32	238

Berdasarkan pada Tabel 9 didapatkan hasil dari performa algoritma *naive bayes* mendapatkan hasil akurasi sama dengan 84,10%, presisi sama dengan 88,12% dan recall sama dengan 89,42%.

3.3.3 Algoritma C4.5

Hasil performa dari model C4.5 dapat dilihat pada Tabel 11 dan *confusion matrix* dapat dilihat pada Tabel 12 berikut. Pohon keputusan algoritma C4.5 dapat dilihat pada gambar 2.

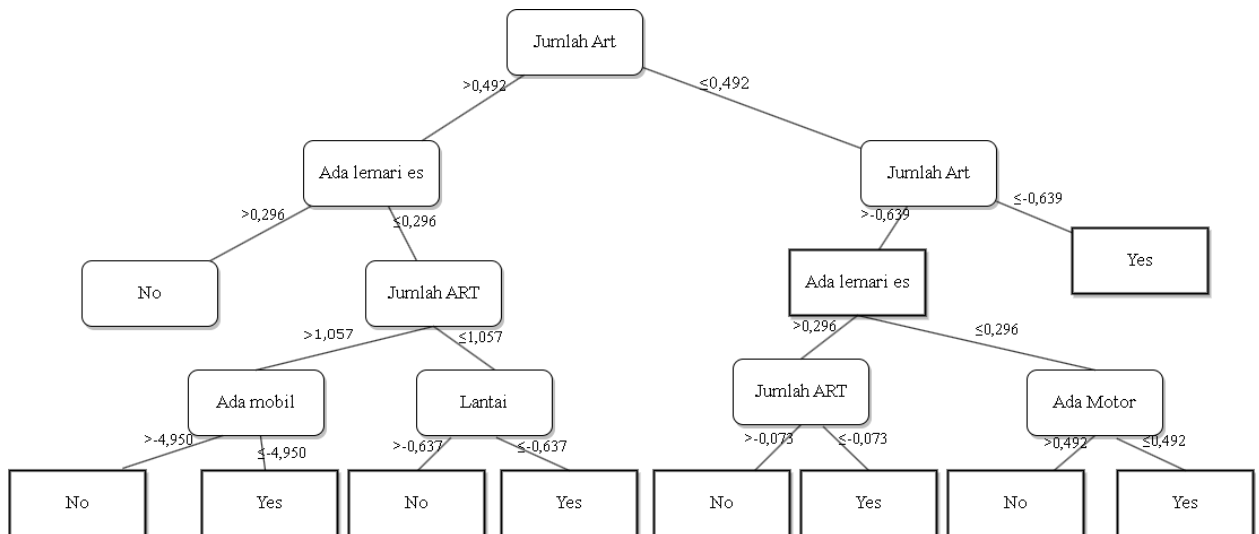
Tabel 11. Hasil Performa C4.5

Performa C4.5			
	Accuracy	Precision	Recall
C4.5	88,36%	93,10%	90,08%

Tabel 12. Confusion Matrix C4.5

Confusion Matrix		
	true No	true Yes
pred. No	103	28
Pres. Yes	19	255

Berdasarkan Tabel 11 didapatkan hasil dari performa algoritma C4.5 nilai akurasi yang didapatkan adalah 88,36%, nilai presisi yang didapatkan adalah 93,10% dan nilai recall yang didapatkan adalah 90,08%.

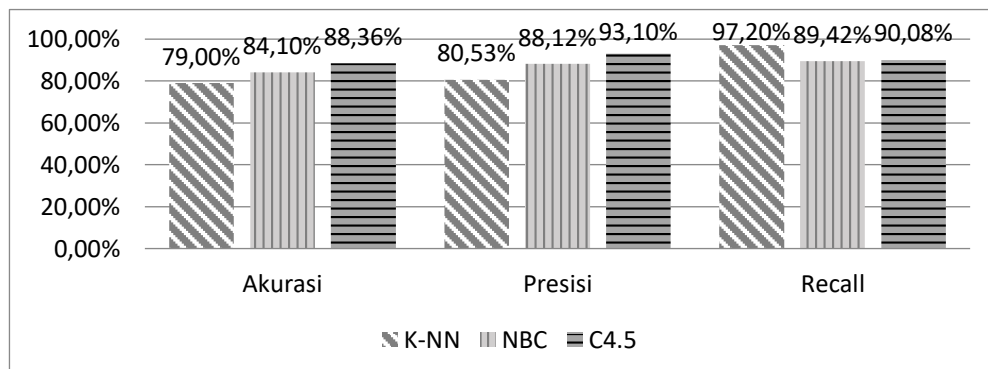


Gambar 2. Pohon Keputusan

Berdasarkan Gambar 2 dapat dilihat bahwa meskipun pada awal terdapat 33 atribut tetapi setelah dilakukan algoritma C4.5 terdapat 5 atribut yang signifikan, yaitu “Jumlah Art”, “ada lemari es”, “ada mobil”, “lantai” dan “ada motor”. Jumlah atribut yang lebih sedikit ini tentu saja membuat kompleksitas algoritma C4.5 menjadi lebih sederhana.

3.4 Hasil Perbandingan Performa Model Klasifikasi

Hasil perbandingan algoritma klasifikasi dapat dilihat pada Gambar 3.



Gambar 3. Performa Hasil Pengujian Algoritma K-NN, NBC dan C4.5

Pada gambar tersebut diketahui akurasi dan presisi tertinggi pada model klasifikasi yang dihasilkan oleh algoritma C4.5. Akan tetapi pada kasus dipenelitian ini lebih cenderung pada model klasifikasi dengan recall tertinggi yaitu model klasifikasi yang dihasilkan oleh K-NN. Hal ini dikarenakan recall yang tinggi dapat meminimalkan orang yang berhak menerima raskin menjadi tidak berhak menerima raskin. Meskipun disisi lain model klasifikasi K-NN memiliki presisi yang terendah yang berarti dibandingkan dengan model klasifikasi yang dihasilkan oleh algoritma yang lain, ada lebih banyak yang seharusnya tidak berhak menerima menjadi berhak menerima. Meskipun begitu algoritma C4.5 memiliki keunggulan dari segi kesederhanaan kompleksitas perhitungan. Hal ini dikarenakan model klasifikasi yang dihasilkan oleh algoritma C4.5 berhasil mereduksi jumlah atribut dari 33 menjadi 5 atribut.

4. KESIMPULAN

Hasil pengujian ini didapatkan bahwa pada model klasifikasi tiga algoritma nilai recall tertinggi diperoleh algoritma k-NN, yaitu dengan nilai 97,20% pada percobaan $k=21$. Akurasi dan presisi tertinggi diperoleh pada algoritma C4.5, yaitu dengan nilai akurasi 88,36% dan presisi dengan nilai 93,10%. Penelitian ini cenderung memilih nilai recall karena dapat meminimalkan orang yang berhak menerima raskin menjadi tidak berhak menerima raskin. Selain itu, algoritma C4.5 menghasilkan sebuah pohon keputusan yang berhasil mereduksi 33 atribut menjadi 5 atribut. Algoritma C4.5 bekerja secara sederhana dalam hal kompleksitas perhitungan, pohon keputusan tersebut dapat dijadikan sebagai pertimbangan untuk mengambil keputusan dari prediksi penerima bantuan raskin. Sehingga dengan adanya penelitian ini diharapkan model klasifikasi tersebut bisa menjadi pertimbangan dalam memprediksi penerima raskin dalam mengambil keputusan sasaran penerima raskin khususnya pada Dinas Sosial Kota Dumai.

REFERENCES

- [1] N. Nurahman and S.- Aminah, "KLASIFIKASI PENERIMA BANTUAN SOSIAL DI DESA BATUAH MENGGUNAKAN METODE ALGORITMA C4.5," *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, vol. 5, no. 2, p. 271, Dec. 2022, doi: 10.37600/tekinkom.v5i2.516.
- [2] R. K. Amru Alba, "Kebijakan-pemberian-bantuan-sosial-bagi-keluarga-miskin," UNIMAL PRESS, 2019.
- [3] C. Fadlan, S. Ningsih, and A. Perdana Windarto, "PENERAPAN METODE NAÏVE BAYES DALAM KLASIFIKASI KELAYAKAN KELUARGA PENERIMA BERAS RASTRA," 2018. doi: doi.org/10.59637/jsti.v17i2.152.
- [4] R. Meivera, M. Hutabarat, J. Waruwu, R. Meivera Siburian, and D. J. Waruwu, "SISTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN PENERIMAAN BERAS MISKIN (RASKIN) DENGAN METODE WEIGHTED PRODUCT (STUDI KASUS DESA HILIWETO KECAMATAN ONOHAZUMBA)," 2022. doi: doi.org/10.59637/jsti.v17i2.152.
- [5] TNP2K, "PROGRAM BANTUAN PEMERINTAH UNTUK INDIVIDU, KELUARGA, DAN KELOMPOK TIDAK MAMPU Menuju Bantuan Sosial Terintegrasi," Cetakan Pertama, 2018.
- [6] Pratama, Dimas Epy, Selnia M. Lamata, and Dewi Halimatuz Zahro. "PROGRAM RASKIN SEBAGAI DAKWAH PEMBERDAYAAN MASYARAKAT PROLETAR DI INDONESIA." *Dakwatun: Jurnal Manajemen Dakwah* 1, no. 2 (2022): 106-117. doi.org/10.58194/jdmd.v1i2.121.
- [7] F. Yumono, D. Efytra Yuliana, and R. N. Sarbini, "HISTOGRAM CITRA JENIS BERAS DENGAN MENYERTAKAN KERTAS PUTIH UNTUK IDENTIFIKASI AWAL JENIS BERAS DENGAN MENGGUNAKAN JARINGAN SYARAF TIRUAN," *Jurnal Informatika dan Rekayasa Perangkat Lunak (JATIKA)*, vol. 3, no. 2, pp. 129–137, 2022, doi.org/10.33365/jatika.v3i1.189.
- [8] Mustofa, Kharomatun Nuruhul, Eka Suswaini, and Nurfalinda Nurfalinda. "KLASIFIKASI PENERIMA BANTUAN RASTRA MENGGUNAKAN METODE NAIVE BAYES." *Student Online Journal (SOJ) UMRAH-Teknik 2*, no. 1 (2021): 175-185.
- [9] P. Pendidikan, P. dan Jumlah Tanggungan Terhadap Kemiskinan di Desa Aek Bolon Julu, D. Siahaan, and R. Linda Sari, "Employment and Number of Dependents on Poverty in Aek Bolon Julu Village," 2023, doi: 10.32734/lwsa.v6i1.1683.
- [10] Sandi, S. A., & Novianto, Y. (2023). Klasifikasi Kelayakan Keluarga Penerima Bantuan Langsung Tunai Menggunakan Gain Ratio Dan Naïve Bayes. *Jurnal Informatika Dan Rekayasa Komputer (Jakakom)*, 3(1), 433-442, doi.org/10.33998/jakakom.2023.3.1.794.
- [11] Q. Iman and A. W. Wijayanto, "Klasifikasi Rumah Tangga Penerima Beras Miskin (Raskin)/Beras Sejahtera (Rastra) Di Provinsi Jawa Barat Tahun 2017 Dengan Metode Random Forest Dan Support Vector Machine," *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 9, no. 2, pp. 178–184, 2021.
- [12] Supriyadi, Andy Supriyadi Andy. "Perbandingan Algoritma Naive Bayes dan Decision Tree (C4. 5) dalam Klasifikasi Dosen Berprestasi." *Generation Journal* 7, no. 1 (2023), doi.org/10.29407/gj.v7i1.19797.
- [13] A. Tangkelayuk and E. Mailoa, "Klasifikasi Kualitas Air Menggunakan Metode KNN, Naïve Bayes Dan Decision Tree," vol. 9, no. 2, pp. 1109–1119, 2022, doi.org/10.35957/jatisi.v9i2.2048.
- [14] M. Sadikin, R. Rosnelly, R. Roslina, and ..., "Penerapan Data Mining Pada Penerimaan Dosen Tetap Menggunakan Metode Naive Bayes Classifier dan C4. 5," *Jurnal Media ...*, vol. 4, pp. 1100–1109, 2020, doi: 10.30865/mib.v4i4.2434.
- [15] Ridho, Rashid, and Hendra Hendra. "Klasifikasi Diagnosis Penyakit Covid-19 Menggunakan Metode Decision Tree." *JUST IT: Jurnal Sistem Informasi, Teknologi Informasi dan Komputer* 11, no. 3 (2022), doi.org/10.24853/justit.11.3.%25p.
- [16] Hidayatulloh, Taufik, Ardi Winardi, Lestari Yusuf, Satia Suhada, and Saeful Bahri. "Feasibility Test Of Poor Rice Recipients In Bencoy Sukabumi Village Using Naive Bayes." *Jurnal Pilar Nusa Mandiri* 17, no. 1 (2021): 93-98, doi.org/10.33480/pilar.v17i1.2227.
- [17] N. Nurahman and S.- Aminah, "KLASIFIKASI PENERIMA BANTUAN SOSIAL DI DESA BATUAH MENGGUNAKAN METODE ALGORITMA C4.5," *Jurnal Teknik Informasi dan Komputer (Tekinkom)*, vol. 5, no. 2, p. 271, Dec. 2022, doi: 10.37600/tekinkom.v5i2.516.



- [18] R. Tuntun, K. Kusrini, and K. Kusnawi, “Analisis Perbandingan Kinerja Algoritma Klasifikasi dengan Menggunakan Metode K-Fold Cross Validation,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 4, p. 2111, Oct. 2022, doi: 10.30865/mib.v6i4.4681.
- [19] Q. A’yuniyah et al., “Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 1, p. 72, Sep. 2022, doi: 10.30865/json.v4i1.4781.
- [20] D. Normawati and S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *Jurnal Sains Komputer & Informatika (J-Sakti)*, vol. 5, no. 2, pp. 697–711, 2021, doi.org/10.30645/j-sakti.v5i2.369.
- [21] K. P. Siwilopo and H. Marcos, “MEMBANDINGKAN KLASIFIKASI PADA BUAH JERUK MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK DAN K-NEAREST NEIGHBOR,” *KOMPUTA : Jurnal Ilmiah Komputer dan Informatika*, vol. 12, no. 1, 2023, doi. 10.34010/KOMPUTA.V12I1.9068.