

KLASIFIKASI PENENTUAN PENERIMA BANTUAN SOSIAL TUNAI (BST) MENGGUNAKAN ALGORITMA C4.5 DI DESA KERAMAS, GIANYAR BALI

Ni Wayan Oktha Pratiwi^{1*}, Nengah Widya Utami², I Gede Juliana Eka Putra³

^{1,2,3}Sistem Informasi Akuntansi, STMIK Primakara
okthapратиwi7@gmail.com^{1*}

Abstrak: Pandemi Covid-19 saat ini menyebabkan meningkatnya angka kemiskinan di Indonesia. Untuk menanggulangi penurunan pendapatan maka kementerian sosial memberikan Bantuan Sosial Tunai (BST) kepada 9 juta KPM di Indonesia. Salah satu desa yang mendapatkan bantuan sosial tunai ini yaitu Desa Keramas. Yang menjadi permasalahannya yaitu kurang tepatnya penerima BST. Tujuan dari penelitian ini untuk mengklasifikasi kelayakan penerima BST memakai metode algoritma C4.5 dan diharapkan algoritma ini mampu memberikan rekomendasi untuk pengambilan keputusan dalam penerimaan bantuan lainnya. Hasil dari penelitian ini untuk memprediksi apakah masyarakat layak atau tidak layak dalam menerima BST. Peneliti melakukan pengujian menggunakan 5 atribut antara lain pernah mendapatkan bantuan lain, pekerjaan, pendidikan, memiliki kartu miskin, dan jumlah tanggungan. Pengujian yang menggunakan data dengan jumlah 2.074 yang terbagi menjadi 2 jenis diantaranya data *training* dengan jumlah 80% dan *testing* sebanyak 20%. Pengujian ini menghasilkan akurasi sebesar 97,83% dan membandingkan dengan algoritma *K-Nearest Neighbor* yang menghasilkan akurasi sebesar 92,29% dan *Naïve Bayes* sebesar 91,81%.

Kata Kunci : *Data Mining*, Klasifikasi, Algoritma C4.5, Rapidminer *Data Mining*, Bantuan Sosial Tunai.

Abstract: The current Covid-19 pandemic has led to an increase in the poverty rate in Indonesia. To overcome the decline in income, the Ministry of Social Affairs provides Cash Social Assistance (BST) to 9 million KPM in Indonesia. One of the villages that received cash social assistance is Keramas Village. The problem is that the BST recipient is not precise. The purpose of this study is to classify the eligibility of BST recipients using the C4.5 algorithm and the C4.5 algorithm is expected to be able to provide recommendations for decision making in receiving other assistance. The results of this study are to predict whether the community is eligible or not eligible to receive BST. Researchers tested using 5 attributes, including having received other assistance, employment, education, having a poor card, and number of dependents. The test uses 2,074 data which is divided into 2 types of data, namely 80% training data and 20% testing data. This test produces an accuracy of 97.83% and compares with the *K-Nearest Neighbor* algorithm which produces an accuracy of 92.29% and *Naïve Bayes* of 91.81%.

Keywords : *Data Mining*, Classification, C4.5 Algorithm, Rapidminer *Data Mining*, Cash Social Assistance.

PENDAHULUAN

Ketika pandemi Covid-19 terjadi, menyebabkan meningkatnya angka kemiskinan di Indonesia. Berdasarkan data pada Badan Pusat Statistik menyebutkan bahwa bulan Agustus 2020 membuat jumlah pengangguran mengalami peningkatan yang sangat drastis sebesar 2,7 juta orang [1].

Bantuan ini diberikan oleh pemerintah Indonesia dalam menanggulangi masalah penurunan pendapatan masyarakat. Bantuan Sosial Tunai ini diberikan oleh dinas sosial sebanyak lebih dari 103 ribu keluarga penerima manfaat (KPM), dan besar bantuannya Rp.300.000 perbulan yang diberikan dari tahun 2020 sampai dengan 2021 [2]. Desa Keramas juga termasuk ke dalam penerima BST, namun menurut staf desa mengalami kesulitan yang dimana adanya ketidaktepatan penerima bantuan ini.

Fayyad et al. 1996 mengatakan bahwa data mining merupakan sebuah metode analisis untuk menemukan pengetahuan dalam sebuah basis data atau juga sering disebut dengan KDD[3][4]. Menurut Garter Group data mining adalah langkah untuk menemukan relasi seperti pola yang memiliki arti serta cenderung memeriksa data yang tersimpan dengan teknik statistika dan matematika[5]. Sehingga dapat dikatakan bahwa data mining adalah sebuah proses menggali suatu informasi yang memiliki nilai tambah yang tidak ditemukan dengan manual pada sebuah basis data[6]. pohon keputusan adalah jenis prediksib yang menunjukkan faktor yang mempengaruhi keputusan dengan estimasi akhir dalam pengambilan keputusan[7]. Metode algoritma C4.5 adalah metode algoritma yang membantu untuk mengklasifikasikan data penduduk dan membandingkan dengan data penerima bantuan sosial tunai untuk membantu Desa Keramas dan dinas

sosial dalam menentukan warga tersebut layak atau tidaknya mendapatkan BST ini[8].

Adapun studi yang berhubungan dengan penelitian dengan metode klasifikasi decision tree serta algoritma C4.5. Salah satu penelitian dilakukan oleh Siti Badriah, dkk yang berjudul “Klasifikasi Algoritma C4.5 dalam Menentukan Penerima Bantuan Covid-19 (Studi Kasus: Desa di Karawang)” dengan menggunakan 430 data dengan 5 atribut yaitu: nama, jumlah anggota keluarga, pekerjaan, pendapatan sebelum Covid-19, pendapatan setelah Covid-19[9]. Selanjutnya penelitian yang dilakukan oleh Romatua Estiniaty Silalahi, Heru Satria Tambunan dan Dedi Suhendro yang berjudul “Analisa Kelayakan Penerima Program Keluarga Harapan (PKH) Menggunakan Algoritma C4.5” dengan menggunakan 163 data PKH, dengan melakukan observasi dengan 100 sampel data. Dengan 4 atribut yaitu: jumlah penghasilan, jumlah tanggungan keluarga, kepemilikan kendaraan, status tempat tinggal[10].

Penelitian ini bertujuan untuk memberikan rekomendasi kepada Desa Keramas dalam menentukan penerima bantuan selanjutnya yang sesuai dengan perhitungan yang digunakan. Penelitian ini juga diharapkan digunakan untuk masukan dalam memberikan bantuan selanjutnya kepada warga Desa Keramas.

TINJAUAN PUSTAKA

Data Mining

Penambangan data adalah sejumlah prosedur yang bertujuan menggali nilai tambah berupa informasi yang belum diketahui dengan manual dalam sebuah basis data *knowledge discovery in databases* (KDD)[4]. Fayyad et al. 1996 mengatakan bahwa penambangan data merupakan sebuah metode analisis untuk menemukan pengetahuan dalam sebuah basis data atau juga sering disebut dengan KDD[3]. Menurut Garter Group penambangan data adalah langkah untuk menemukan relasi seperti pola yang memiliki arti serta cenderung memeriksa data yang tersimpan dengan teknik statistika dan matematika [5]. Penambangan data dapat dibagi menjadi beberapa kelompok berdasarkan tugas yang dilakukan:

- 1) Uraian
Pada tahap ini dilakukan analisis secara sederhana guna menemukan pola serta kecenderungan terhadap data yang tersedia
- 2) Pertimbangan
Bentuk yang diwujudkan dengan record lengkap yang menyediakan nilai dari variabel nilai prediksi
- 3) Prediksi

Prediksi sama seperti estimasi namun perbedaannya hanya hasilnya aka nada dimasa mendatang.

- 4) Klasifikasi
Digunakan untuk mengelompokkan data berdasarkan ketertarikan data terhadap sampel.
- 5) Pengklusteran
Dipakai untuk mengelompokkan sebuah data yang memiliki kesamaan kedalam sebuah wilayah yang sama
- 6) Asosiasi
Digunakan untuk menemukan aturan asosiatif anatar suatu kombinasi item.

Klasifikasi

Han dan klasifikasi berpendapat bahwa klasifikasi adalah model data mining yang melibatkan prediksi jenis label[7]. Klasifikasi dapat dilihat sebagai suatu proses untuk mendeklarasikan suatu objek data sebagai salah satu jenisnya. Di antara beberapa algoritma klasifikasi, metode pohon keputusan digunakan. Pohon keputusan adalah algoritma yang dapat mengubah data yang sangat besar menjadi pohon keputusan yang dapat digunakan sebagai keputusan dalam representasi aturan[11].

Algoritma C4.5

Data pohon keputusan direpresentasikan sebagai tabel dengan atribut dan catatan. Atribut adalah parameter yang digunakan sebagai kriteria dalam pohon keputusan. Atribut target adalah atribut yang menunjukkan data solusi berdasarkan elemen. Versi adalah atribut yang memiliki nilai, misalnya atribut cuaca memiliki kasus seperti cerah, mendung, dan hujan. Secara umum dalam membangun pohon keputusan adalah sebagai berikut[7]:

- 1) Pilih atribut sebagai root.
- 2) Buat cabang untuk setiap nilai.
- 3) Mendistribusikan kasus di cabang-cabang.
- 4) Kembali lakukan proses sehingga kasus memiliki klasifikasi yang lama

Untuk menghitung nilai *gain* digunakan rumus, yaitu:

$$Gain(S,A) = Entropy(s) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan:

S : Himpunan kasus

A : Atribut

n : Jumlah partisipasi atribut A

|S| :Jumlah kasus pada partisi ke-i

|S_i| :Jumlah kasus dalam S

Untuk menghitung nilai entropy digunakan rumus, yaitu:

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

S : Himpunan kasus

A : Filtur

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

Confusion Matrix

Cross Validation adalah salah satu teknik dalam menilai tingkat akurasi dalam data mining. Metode cross validation juga sering disebut dengan k-fold cross validation menurut (Santosa dan Umam 2018) dalam bukunya “Data Mining dan Big Data Analytics”. Data yang dipakai dalam proses disebut training, sedangkan data yang digunakan disebut dengan data testing[12].

Tabel 1. Confusion Matrix

Kelas	Prediksi Yes	Prediksi No	Total
Aktual Yes	True Positive (TP)	False Negative (FN)	Positive (P)
Aktual No	False Positive (FP)	True Negative (TN)	Negative (N)
Total	P'	N'	P+N

Rumus:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{TN + FP}$$

Penelitian Yang Relevan

Suatu penelitian terdahulu yang berkaitan dengan klasifikasi decision tree dengan metode algoritma C4.5. Salah satu penelitian dilakukan oleh Siti Badriah, dkk yang berjudul “Klasifikasi Algoritma C4.5 dalam Menentukan Penerima Bantuan Covid-19 (Studi Kasus: Desa di Karawang)” dengan menggunakan 430 data dengan 5 atribut yaitu: nama, jumlah anggota keluarga, pekerjaan, pendapatan sebelum Covid-19, pendapatan setelah Covid-19. Yang mana nilai akurasi terbaik terdapat pada perbandingan 90% data training dan 10% data testing dan nilai akurasi sebesar 79,54% dengan hasil pengaruh yang paling tinggi terdapat pada atribut penghasilan dan yang kedua pada kondisi rumah[9].

Penelitian selanjutnya yang dilakukan oleh Romatua Estiniaty Silalahi, Heru Satria Tambunan

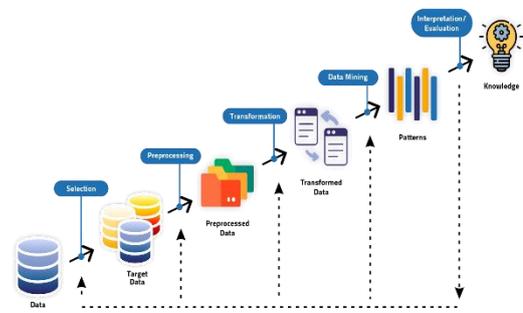
dan Dedi Suhendro yang berjudul “Analisa Kelayakan Penerima Program Keluarga Harapan (PKH) Menggunakan Algoritma C4.5” dengan menggunakan 163 data PKH, dengan melakukan observasi dengan 100 sampel data. Dengan 4 atribut yaitu: jumlah penghasilan, jumlah tanggungan keluarga, kepemilikan kendaraan, status tempat tinggal. Hasil dari penelitian ini yaitu menghasilkan enam rules dan tingkat akurasi sebesar 96,67% dan dengan adanya penerapan ini mampu memberikan solusi [10]. Namun yang menjadi perbedaan dengan penelitian terdahulu yaitu pada penelitian ini menggunakan 5 atribut sebagai penilaian dan juga membandingkan tingkat akurasi dengan menggunakan algoritma K-NN dan *Naïve Bayes*.

METODE

Jenis Data dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data kuantitatif dan kualitatif. Sumber data yang digunakan adalah data primer dan data sekunder. Data dalam bentuk angka yang dapat dihitung, hasil pengamatan langsung, dan sumber data sekunder yang digunakan sebagai pendukung sumber utama, yaitu buku, jurnal sebelumnya, dan data yang dapat dipercaya di Internet..

KDD (Knowledge Discovery in Database)



Gambar 1. Tahapan KDD

1) Pemilihan data

Langkah ini memilih data dari kumpulan data yang ada. Sebelum information mining di KDD dimulai dan data seleksi ini digunakan dalam data mining, kemudian disimpan dalam file data operasi yang terpisah [11].

2) Pembersihan

Proses ini dilakukan untuk menghilangkan duplikasi data yang ada, memverifikasi inkonsistensi data, dan

memperbaiki kesalahan data yang ada seperti salah cetak, dll.

3) Transformasi

Data diubah menjadi bentuk yang sesuai untuk penemuan, karena data ini sangat tergantung pada jenis atau pola informasi yang dicari dari database.

4) Penambangan data

Penambangan data adalah proses menemukan pola atau informasi yang menarik menggunakan data yang dipilih menggunakan metode yang telah ditentukan..

5) Interpretasi/Ulasan

Pada tahap ini hasil dari data mining yang perlu ditampilkan dalam bentuk yang mudah dipahami dan dipahami oleh banyak orang. [5].

HASIL DAN PEMBAHASAN

Pada enelitian ini mulai dengan melakukan pengumpulan data yaitu data seluruh penduduk desa, namun yang hanya digunakan yaitu data per-kepala keluarga yaitu sebanyak 2.077 data, dengan membandingkan data penerima BST pada tahun 2020 sebanyak 156 penerima dan tahun 2021 sebanyak 235 penerima. Selain itu juga menggunakan data penerima bantuan lainnya, seperti bantuan BLT pada tahun 2020 sebanyak 140 penerima, pada tahun 2021 sebanyak 200 penerima, pada tahun 2022 sebanyak 102 penerima dan untuk bantuan UMKM sebanyak 9 penerima. Data yang didapatkan adalah data yang diperoleh dari Desa Keramas berupa data mentah dalam bentuk excel.

Dari data yang sudah ada, dilakukan pembersihan data yang tidak diperlukan secara manual, secara lengkapnya dilakukan pada proses KDD (Knowledge Discovery in Database)

Selection

Langkah selanjutnya adalah pembersihan data yang tidak diperlukan dari sumber data sebelum ke tahapan cleaning atau preprocessing.

Preprocessing

Preprocessing atau *cleaning* yang dilakukan yaitu dengan menghapus dan menghilangkan data yang tidak di perlukan. Dari 2.077 data kepala keluarga yang ada hanya di digunakan 2.074 data, karena terdapat data yang kosong.

Pada tahapan ini dilakukan penghapusan beberapa atribut seperti nama, nik, usia, jenis kelamin, hubungan keluarga, nama ayah dan nama ibu. Pada tahapan ini juga dilakukan penambahan data untuk atribut yang digunakan seperti, pernah mendapatkan bantuan lain seperti BLT, UMKM dan penambahan apakah memiliki kartu miskin atau tidak.

Transformation

Transformation merupakan tahapan yang dimana data yang sudah melalui tahapan *selection* dan *cleaning*, pada tahap ini data yang ada harus sesuai dengan algoritma proses *data mining*. Pada tahapan ini yang dilakukan mengubah bentuk data jumlah tanggungan, karena sesuai dengan algoritma yang digunakan dalam bentuk numerik dan nominal dan juga atribut dari pendidikan , pekerjaan, jumlah tanggungan di sesuaikan dengan kriteria yang telah di buat sebelumnya.

Data Mining

Data mining pada penelitian ini yaitu menggunakan klasifikasi algoritma C4.5 dan tools yang digunakan adalah Rapidminer. Metode ini digunakan karena mudah untuk dimengerti, bisa divisualisasikan ke menjadi bentuk pohon keputusan, algoritma C4.5 juga menggunakan konsep information entropy dan gain sehingga pembagiannya yang optimal.

Pada tahapan ini peneliti menggunakan tools RapidMiner versi 9.10 dalam melakukan pengujian terhadap klasifikasi penentuan layak atau tidak pemberian bantuan terhadap warga Desa Keramas. Pengolahan data dibedakan menjadi 2 yaitu data testing dan data training. Analisa data dilakukan

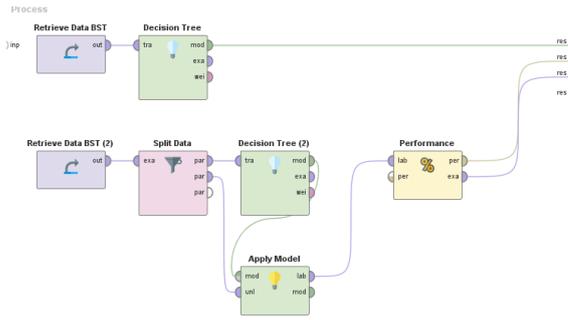
dengan perbandingan 80% : 20%. Perbandingan ini dilakukan untuk mengetahui nilai akurasi yang paling baik digunakan.

Tabel 2. Jumlah Data *Preprocessing*

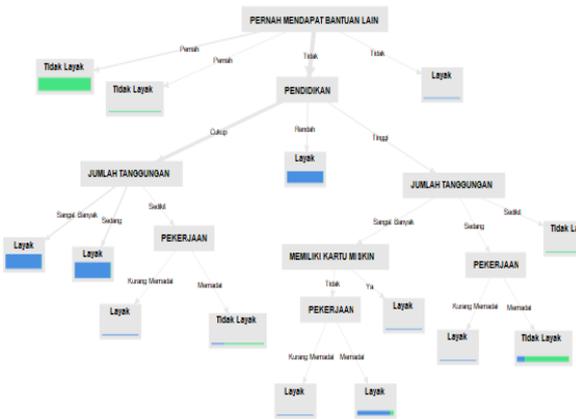
Nama Banjar	Sebelum <i>Preprocessing</i>	Sesudah <i>Preprocessing</i>
Banjar Palak	264	264
Banjar Biya	418	417
Banjar Maspait	456	456
Banjar Lodpeken	266	265
Banjar Lebah	284	284
Banjar Gelgel	389	388
Total	2.077	2.074

Interpretation & Evaluation

Pada tahap ini merupakan tahapan terakhir setelah dilakukannya semua tahapan sebelumnya. Tahap ini merupakan penggunaan aplikasi RapidMiner yang mana perubahan pola-pola yang dihasilkan menjadi sebuah informasi yang mudah untuk dipahami dan mengerti. Pada tahap ini juga pemeriksaan apakah informasi yang dihasilkan bertentangan dengan fakta atau hipotensis yang sebelumnya.



Gambar 2. Model Algoritma C4.5



Gambar 3. Decision Tree C4.5

Berdasarkan gambar 4.17 merupakan hasil dari *decision tree* algoritma C4.5

Row No.	KEPUTUSAN	prediction(KEP..)	confidence(Layak)	confidence(Tidak Layak)	confidence(Tidak Layak)	confidence(Tidak Layak)
1	Layak	Tidak Layak	0.146	0.837	0.016	0
2	Tidak Layak	Layak	0.992	0.008	0	0
3	Layak	Tidak Layak	0.146	0.837	0.016	0
4	Layak	Tidak Layak	0.146	0.837	0.016	0
5	Layak	Tidak Layak	0.188	0.812	0	0
6	Layak	Tidak Layak	0.146	0.837	0.016	0
7	Layak	Tidak Layak	0.146	0.837	0.016	0
8	Tidak Layak	Layak	0.904	0.096	0	0
9	Tidak Layak	Layak	0.904	0.096	0	0

Gambar 4. Hasil Prediksi Algoritma C4.5

Gambar 4.19 merupakan hasil dari prediksi yang telah di input dengan tingkat akurasi 80% data *training* dan 20% data *testing*. Yang dimana pada prediksi ini mengalami kesalahan sebesar 2,2% atau dari 415 data *testing* yang salah prediksi hanya 9 data. Maka dapat dikatakan bahwa prediksi menggunakan metode C4.5 dikatakan cukup berhasil.

accuracy: 97,83%

	true Layak	true Tidak Layak	true tidak Layak	true Tidak Layak	class precision
pred. Layak	291	3	0	0	98.98%
pred. Tidak Layak	6	115	0	0	95.04%
pred. tidak Layak	0	0	0	0	0.00%
pred. Tidak Layak	0	0	0	0	0.00%
class recall	97.98%	97.46%	0.00%	0.00%	

Gambar 5. Nilai Akurasi Algoritma C4.5

Adapun perhitungan manual dari nilai pada gambar 5 adalah sebagai berikut :

$$Accuracy = \frac{291 + 115}{291 + 6 + 115 + 3} = 0,9783 = 97,83\%$$

$$Precision = \frac{291}{291 + 6} = 0,9797 = 97,97\%$$

$$Recall = \frac{291}{291 + 3} = 0,9897 = 98,97\%$$

$$Specificity = \frac{115}{115 + 6} = 0,9504 = 95,04\%$$

Tabel 3. Perbandingan Akurasi Algoritma C4.5

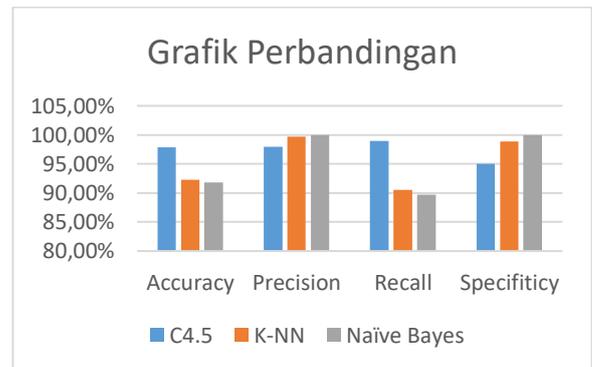
TRAINING	TESTING	HASIL AKURASI
50%	50%	97,30%
60%	40%	92,04%
70%	30%	92,77%
80%	20%	97,83%
90%	10%	96,62%

Berdasarkan hasil penerapan Algoritma C4.5 menggunakan tolls *Rapidminer* dengan operator *Split Data* dengan perbandingan data *training* dan data *testing* seperti tabel 4.8 diatas didapatkan akurasi yang paling besar yaitu perbandingan 80%;20% dengan nilai akurasi sebesar 97,83%, yang dimana artinya bahwa rule yang dihasilkan dari algoritma ini mendekati 100%.

Tabel 4. Perbandingan Algoritma

Algoritma	Accuracy	Precision	Recall	Specificity
C4.5	97,83%	97,97%	98,97%	95,04%
K-Nearest Neighbor	92,29%	99,66%	90,51%	98,86%
Naïve Bayes	91,81%	100%	89,72%	100%

Hasil dari perbandingan diperoleh nilai tertinggi dari tingkat akurasi yaitu dari algoritma C4.5 sebesar 97,83%, dari tingkat *Precision* yaitu algoritma *Naïve Bayes* sebesar 100% dari tingkat *Recall* yaitu algoritma C4.5 sebesar 98,97% dan dari tingkat *Specificity* yaitu algoritma *Naïve Bayes* sebesar 100%. Jika dirata-ratakan maka algoritma paling tinggi yaitu C4.5 sebesar 97,45%.



Gambar 6. Grafik Perbandingan Algoritma

Berdasarkan dari perbandingan antara algoritma C4.5, K- Nearest Neighbor, dan Naïve Bayes yang

telah dilakukan. Hasil dari tingkat akurasi, *precision*, *recall* dan *specificity* jika dirubah kedalam bentuk grafik adalah seperti grafik diatas.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan dari penelitian dan pengujian data menggunakan *tolls* Rapidminer dalam memprediksi kelayakan penerima bantuan sosial tunai kepada masyarakat Desa Keramas, Gianyar Bali. Dalam pengujian ini menggunakan sebanyak 2.077 data dan setelah melalui tahapan KDD didapatkan yaitu pendidikan, pekerjaan, jumlah tanggungan, memiliki kartu miskin, dan pernah mendapatkan bantuan lain. Dalam pengujian ini data dibagi menjadi 2 jenis yaitu data *training* dan data *testing*, dalam pengujian ini menggunakan perbandingan 80% *training* 20% data *testing*. Selain menggunakan algoritma C4.5 untuk mengetahui algoritma yang mana lebih tinggi tingkat akurasinya dalam penelitian ini juga menggunakan algoritma K-NN dan Naïve Bayes dan dapat ditarik kesimpulannya yaitu sebagai berikut:

- 1) Berdasarkan hasil dari pohon keputusan (*decision tree*) didapatkan yang menjadi *root node* yaitu “pernah mendapatkan bantuan lain”, dan yang menjadi *leaf node* yaitu “pekerjaan”. Sehingga dapat ditarik kesimpulannya bahwa masyarakat yang berhak mendapatkan bantuan ini masyarakat yang tidak pernah mendapatkan bantuan lainnya.
- 2) Berdasarkan hasil prediksi dengan algoritma C4.5 yang telah dilakukan bahwa algoritma ini cocok untuk memprediksi kelayakan penerima bantuan sosial tunai karena menghasilkan akurasi sebesar 97,83%.
- 3) Hasil perbandingan algoritma C4.5 dengan *K-Nearest Neighbor* dan Naïve Bayes menunjukkan bahwa algoritma C4.5 menghasilkan nilai terbaik. Algoritma C4.5 menghasilkan akurasi sebesar 97,83%, *precision* sebesar 97,97%, *recall* sebesar 98,97%, *specificity* sebesar 95,04%. Algoritma *K-Nearest Neighbor* menghasilkan akurasi sebesar 92,29%, *precision* sebesar 99,66%, *recall* sebesar 90,51%, *specificity* sebesar 98,86%. Algoritma *Naïve Bayes* menghasilkan akurasi sebesar 91,81%, *precision* sebesar 100%, *recall* sebesar 89,72%, *specificity* sebesar 100%.

Saran

Dari hasil penelitian penerapan data mining yang telah dijelaskan, adapun saran yang sekiranya dapat menjadi bahan pengembangan selanjutnya yaitu sebagai berikut:

- 1) Penelitian ini bisa digunakan dalam penentuan bantuan lainnya, jika atribut yang digunakan itu sama. Karena penelitian ini dilaksanakan pada masa pandemi covid-19.

- 2) Mencoba menggunakan data yang lebih besar.
- 3) Penambahkan variabel lain agar hasil dari klasifikasi yang dihasilkan lebih akurat seperti penghasilan, jumlah aset atau yang lainnya.
- 4) Pengembangan lebih lanjut untuk membuat sistem pendukung keputusan berbasis web maupun *mobile* agar mempermudah Desa Keramas, Gianyar Bali dalam menentukan kelayakan penerima bantuan sosial tunai atau bantuan lainnya.

DAFTAR PUSTAKA

- [1] “Situasi Kemiskinan Selama Pandemi | The SMERU Research Institute.” <https://smeru.or.id/id/content/situasi-kemiskinan-selama-pandemi> (accessed Oct. 07, 2021).
- [2] “BANTUAN SOSIAL MASYARAKAT UNTUK BALI - Pemerintah Provinsi Bali.” <https://www.baliprov.go.id/web/bantuan-sosial-masyarakat-untuk-bali/> (accessed Oct. 13, 2021).
- [3] Dr. Suyanto, *Data Mining Untuk Klasifikasi dan Klasterisasi Data*, 1st ed. Bandung: Informatika Bandung, 2019.
- [4] N. W. Utami and A. Paramitha, “Penerapan Data Mining Untuk Mengetahui Pola Pemilihan Program Studi Di Stmik Primakara Menggunakan Algoritma K-Means ...,” *J. Teknol. Inf. dan ...*, vol. 3, pp. 456–463, 2021, [Online]. Available: <http://jurnal.undhirabali.ac.id/index.php/jutik/article/view/1540>.
- [5] K. A. E. T. Luthfi, *Algoritma Data Mining Yogyakarta*, no. February. Yogyakarta: C.V Andi Offset, 2009.
- [6] N. W. Utami, I. N. Sukajaya, I. Made Candiasa, and E. G. A. Dewi, “The implementation of data mining to show UKT (students’ tuition) using fuzzy C-means algorithm: (Case study: Universitas Pendidikan Ganesha),” *2019 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2019*, pp. 101–106, 2019, doi: 10.1109/ICACSIS47736.2019.8979933.
- [7] R. Vlandari, *Data Mining (Teori dan Aplikasi Rapidminer)*, 1st ed. Surakarta: Gava Media, 2017.
- [8] N. K. M. Astuti, N. W. Utami, and I. G. P. K. Juliharta, “Classification of Blood Donor Data Using C4.5 and K-Nearest Neighbor

- Methods (Case Study: Utd Pmi Bali Province),” *J. Pilar Nusa Mandiri*, vol. 18, no. 1, pp. 9–16, 2022, doi: 10.33480/pilar.v18i1.2790.
- [9] S. Badriah, M. Fajar, E. Nugroho, N. Sanjaya, and I. Rismawati, “Klasifikasi Algoritma C4 . 5 dalam Menentukan Penerima Bantuan Covid-19 (Studi Kasus : Desa di Karawang),” *JIP (Jurnal Inform. Polinema)*, vol. 19, pp. 23–28, 2019.
- [10] M. A. Tanjung, P. P, and H. Qurniawan, “Analisa Kelayakan Penerima Program Keluarga Harapan (PKH) Menggunakan Algoritma C4.5,” *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 6, no. 1, p. 217, 2021, doi: 10.30645/jurasik.v6i1.286.
- [11] Yuli Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika,” *J. Edik Inform.*, vol. 2, 2019.
- [12] S. Hendrian, “Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan,” *Fakt. Exacta*, vol. 11, no. 3, pp. 266–274, 2018, doi: 10.30998/faktorexacta.v11i3.2777.