

PENERAPAN DATA MINING DENGAN ALGORITMA C4.5 DALAM PENENTUAN KELAYAKAN NASABAH KREDIT (STUDI KASUS: LPD DESA ADAT SUMERTA)

AA Ayu Wulan Agustini^{a1}, I Putu Mahendra Adi Wardana^{a2}, Kadek Oky Sanjaya^{a3}
Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Sains, Universitas Hindu
Indonesia

¹wulanagustini2000@gmail.com, ²mahendrawardana@unhi.ac.id, ³kadekoki@unhi.ac.id

Abstrak

LPD Desa Adat Sumerta merupakan sebuah organisasi yang terletak pada desa adat yang melayani nasabah pada jasa kredit. Di era perkembangan teknologi saat ini LPD Desa Adat Sumerta masih memberikan kredit menggunakan sistem yang manual. Terdapat sebuah permasalahan mengenai pemberian kredit di LPD Desa Adat Sumerta karena masih berdasarkan sistem manual dan pemberian kreditnya masih terdapat kesalahan. Dari itu dilakukan penelitian mengenai data mining menggunakan algoritma C4.5 dalam penentuan kelayakan kredit di LPD Desa Adat Sumerta untuk mengurangi terjadinya kesalahan dalam pemberian kredit. Penelitian ini menggunakan observasi, wawancara, dokumentasi dan studi kepustakaan. Data didapatkan secara langsung dari ketua LPD dan kepala bagian kredit pada LPD Desa Adat Sumerta. Algoritma C4.5 membantu dalam pengklasifikasian nasabah yang akan menghasilkan sebuah pohon keputusan atau decision tree. Hasil dari perhitungan algoritma C4.5 adalah menggunakan Microsoft Excel mendapatkan nilai pohon keputusan node terakhir yaitu kriteria pendapatan dengan nilai >12.500.000 (5%) adalah nasabah yang diprioritaskan mendapatkan kredit dan nilai <=12.500.000 (95%). Setelah itu dengan dataset yang telah dimiliki juga dilakukan pengujian confusion matrix dengan hasil nilai accuracy sebesar 81.00%, nilai precision 100.00%, recall 100.00%.

Kata Kunci: Data Mining, C4.5, Kredit, Kelayakan, LPD

Abstract

LPD Desa Adat Sumerta is an organization located in a traditional village that serves customers in credit services. In the current era of technological development, LPD Desa Adat Sumerta still provides credit using a manual system. There is a problem regarding the provision of credit at LPD Desa Adat Sumerta because it is still based on a manual system and the provision of credit still has errors. Therefore, research was conducted on data mining using the C4.5 algorithm in determining creditworthiness at LPD Desa Adat Sumerta to reduce errors in granting credit. This research uses observation, interviews, documentation, and literature study. Data was obtained directly from the head of the LPD and the head of the credit section at the LPD Desa Adat Sumerta. The C4.5 algorithm helps in classifying customers which will produce a decision tree. The results of the C4.5 algorithm calculation are using Microsoft Excel to get the value of the last node decision tree, namely the income criteria with a value of > 12,500,000 is a customer who is prioritized to get credit and a value of < = 12,500,000. After that, with the dataset that has been owned, confusion matrix testing is also carried out with the results of an accuracy value of 81.00%, a precision value of 100.00%, a recall of 100.00%.

Keywords: Data Mining, C4.5, Credit, Eligibility, LPD

1. Pendahuluan

LPD Desa Adat Sumerta memiliki konsep desa adat dan bersinergi kuat dan memiliki kondisi-kondisi tertentu yang tidak dapat ditemukan pada penyediaan layanan jasa keuangan di daerah lain, dikarenakan LPD mengayomi masyarakat desa adat dengan cara pendekatan yang kekurangan dalam pemberian kredit, dan akan diusahakan dibantu jika nasabah tersebut memiliki bobot penilaian yang sesuai dengan yang ada pada LPD Desa Adat Sumerta [1], [2]. Di dalam LPD Desa Adat Sumerta tentunya terdapat beberapa jenis nasabah yang meminjam kredit seperti, nasabah dengan kredit lancar, kurang lancar, dan kredit macet [3]–[5]. Masalah kredit macet adalah salah satu masalah yang sudah lumrah

didalam perkreditan, dikarenakan kurangnya analisa terhadap nasabah pemohon kredit). Kualitas nasabah yang meminjam kredit pada LPD Desa Adat Sumerta juga mengalami penurunan yang dipengaruhi oleh pendapatan perkapita nasabah, seperti pendapatan yang kurang dan gaya sosial yang tinggi.

LPD Desa Adat Sumerta memiliki beberapa kriteria tertentu untuk nasabah yang ingin melakukan peminjaman kredit, tentunya dalam proses peminjaman kredit sering terjadi kendala seperti kredit macet, dikarenakan kurangnya perhatian terhadap nasabah dalam pemberian kredit. Mengurangi terjadinya kesalahan dalam pemberian kredit diperlukan perhitungan kelayakan nasabah menggunakan penerapan data mining, apakah nasabah tersebut layak untuk diberikan kredit atau tidak. Data Mining merupakan suatu solusi pemecahan masalah dengan teknik klasifikasi yang memiliki beberapa 2 karakteristik seperti penentuan nilai *gain* dan *entropy* yang kemungkinan setiap kriterianya menjadi acuan keputusan dilanjutkan dengan pemeringkatan dan hasil keputusan [6], [7].

Penelitian ini akan menggunakan prosedur dari tahapan model Cross-Industry Standard Process for Data Mining atau yang biasa disebut dengan CRISP-DM karena proses pada model ini sangat terstruktur dan mudah digunakan [8]. Pemilihan metode yang digunakan dalam data mining adalah algoritma C4.5 yaitu untuk menentukan nasabah yang layak mendapatkan kredit. Diharapkan penelitian ini dapat memberikan hasil yang baik dalam perhitungan penentuan pemberian kelayakan kredit kepada LPD Desa Adat Sumerta. Algoritma C4.5 belum banyak digunakan dalam penelitian penentuan kelayakan kredit. Algoritma C4.5 merupakan evolusi dari algoritma Interative Dichomizer (ID3) [9]. Algoritma ini memiliki input contoh dalam membangun pohon keputusan yang akan dilakukan pengujian dan bidang data yang digunakan sebagai parameter untuk menentukan peringkat data [10]. Algoritma C4.5 terdiri dari pembuatan pohon dengan percabangan awal sebagai atribut terpenting, yang bercabang sampai aturan terpenuhi. Setelah mendapatkan hasil selanjutnya akan dilakukan pengujian akurasi dengan menggunakan *confusion matrix* untuk mengetahui nilai *accuracy*, *recall* dan *precision* pada perhitungan menggunakan algoritma C4.5.

Penelitian ini bertujuan untuk mendapatkan klasifikasi nasabah yang akan mendapatkan kredit di prioritaskan dan mendapatkan kredit lebih akhir dengan menggunakan penerapan data mining menggunakan algoritma C4.5. Dalam sebuah pemodelan data mining, analisis dilakukan untuk membantu LPD Desa Adat Sumerta memberikan keputusan yang baik dan efektif serta tepat sasaran terkait dengan kelayakan penerima kredit LPD Desa Adat Sumerta.

2. Metodologi Penelitian

Penelitian ini akan menggunakan prosedur dari tahapan model CRISP-DM yang mana merupakan kerangka kerja dalam penelitian data mining karena alurnya mudah dipahami dan digunakan. Adapun tahapan CRISP-DM dalam penelitian penerapan data mining dalam menentukan kelayakan nasabah kredit ini, yaitu pemahaman bisnis (*business understanding*), pemahaman data (*data understanding*), persiapan data (*data preparation*), pemodelan (*modelling*), evaluasi (*evaluation*), penyebaran (*deployment*) [11].



Gambar 1. Tahapan Model CRISP-DM

2.1. Pemahaman Bisnis (*Business Understanding*)

Tahapan pemahaman bisnis hal yang dilakukan yaitu memahami kebutuhan serta penentuan tujuan dari data mining. Penelitian ini dimulai dari pemahaman dalam studi kasus yaitu penentuan kelayakan nasabah kredit, dilanjutkan dengan pemilihan metode data mining dan akan dilakukan pengujian dengan RapidMiner. Pemilihan metode yang digunakan dalam data mining adalah algoritma C4.5 yaitu untuk menentukan nasabah yang layak mendapatkan kredit. Diharapkan penelitian ini dapat memberikan hasil yang baik dalam perhitungan penentuan pemberian kelayakan kredit kepada LPD Desa Adat Sumerta. Algoritma C4.5 belum banyak digunakan dalam penelitian penentuan kelayakan kredit.

2.2. Pemahaman Data (*Data Understanding*)

Tahapan pemahaman data Langkah yang dilakukan yaitu mengumpulkan data lalu selanjutnya data diproses untuk dilakukan Analisa serta melakukan evaluasi terhadap efisiensi data yang digunakan. Proses pengumpulan data yang akan digunakan sebagai sampel penelitian dilakukan dengan metode yang telah ditentukan dan sesuai dengan tujuannya. Metode yang digunakan yaitu wawancara dan observasi. Setelah data terkumpul maka selanjutnya dilakukan verifikasi kualitas data untuk melakukan pengecekan kelengkapan data, kesalahan data, hingga apakah terdapat data yang hilang.

2.3. Persiapan Data (*Data Preparation*)

Pada tahap ini meliputi pemilihan data yang akan digunakan dan data yang akan dikeluarkan untuk dimasukkan dalam perhitungan data mining. Pada tahapan ini juga dilakukan data cleaning untuk memperbaiki, menghapus atau mengabaikan noise pada data. proses ini menghasilkan data nilai alternatif. Data yang telah dimiliki pada tahap ini adalah data kriteria, data pembobotan kriteria, serta data nilai alternatif, yang digunakan pada tahap selanjutnya.

2.4. Pemodelan (*Modelling*)

Pemodelan merupakan tahap dimana dilakukannya pemodelan metode yang digunakan, peneliti menggunakan metode Algoritma C4.5.

2.5. Evaluasi (*Evaluation*)

Pada tahap evaluasi dilakukan pengujian menggunakan confusion matrix dari hasil data mining. Analisis akurasi digunakan untuk menentukan data apakah dengan metode yang digunakan akan menghasilkan perbedaan hasil dengan data riil.

2.6. Penyebaran (*Deployment*)

Pada tahapan ini akan dilakukan penyebaran hasil berdasarkan pengujian evaluasi yang telah dilakukan sebelumnya. Jika hasil dari pengujian menunjukkan nilai akurasi yang baik maka dapat disebarkan ke pihak LPD Desa Adat Sumerta.

3. Kajian Pustaka

Penelitian yang dilakukan menggunakan beberapa teori atau kajian pustaka untuk mendukung penelitian agar informasi yang diberikan lebih relevan. Berikut beberapa kajian pustaka yang digunakan dalam penelitian yang dilakukan adalah sebagai berikut.

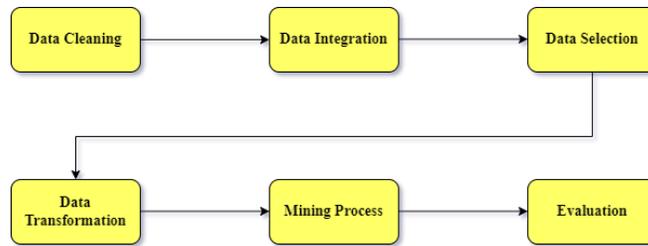
3.1 Data Mining

Data Mining atau pengumpulan data dengan cara menambang data merupakan suatu model atau metode yang digunakan untuk mendapatkan banyak informasi dari banyaknya data yang akan dilakukan dengan menggunakan Teknik statistik, matematika maupun pengenalan sebuah pola [12], [13]. Penerapan data mining memiliki tahapan yang digunakan dalam penyelesaian untuk memprediksi kelayakan pemberian kredit. Melalui tahap-tahap yang terstruktur dengan baik akan memudahkan dalam menentukan hasil dari perhitungan data mining.

3.2 Algoritma C4.5

Algoritma C4.5 merupakan metode yang digunakan dalam menyelesaikan perhitungan data mining untuk menghasilkan klasifikasi yang memiliki sifat prediktif. Algoritma C4.5 akan menghasilkan pola pohon keputusan (*decision tree*) [14]–[16]. Pohon keputusan yang dihasilkan berguna untuk mengeksplorasi data tersembunyi. Kelebihan dari Algoritma C4.5 adalah sebagai berikut.

- a. Algoritma C4.5 dapat melakukan perhitungan meskipun terdapat atribut yang kosong (Missing Value).
- b. Algoritma C4.5 dapat menangani atribut kontinu dan diskrit.
- c. Algoritma C4.5 dapat memangkas pohon keputusan untuk mengatasi overfitting.



Gambar 2. Tahapan Proses Data Mining

Gambar 2. Merupakan alur dari proses data mining dari tahapan awal yaitu *data cleaning* hingga tahapan terakhir *evaluation*. Berikut penjelasan dari tahapan proses data mining, yakni:

1. Pembersihan Data (Data Cleaning), merupakan tahapan proses menghilangkan noise ataupun missing value pada sebuah data. Apabila pada data terdapat data yang tidak digunakan atau kosong maka baris tersebut akan dihilangkan atau tidak digunakan kembali.
2. Integrasi Data (Data Integration), merupakan tahapan menggabungkan data dari banyak sumber yang akan digabungkan ke dalam penyimpanan yang baru.
3. Seleksi Data (Data Selection), Tahapan ini dilakukan pemilihan data operational, karena dari banyaknya data tidak semua akan digunakan. Maka dari itu perlu dilakukan seleksi data yang akan dipilih dan sesuai untuk dilakukan analisis.
4. Data Transformasi (Transformation Data), merupakan proses yang dilakukan untuk mengubah data sesuai dengan format yang akan digunakan dalam proses data mining.
5. Proses Mining (Mining Process), Proses data mining merupakan tahapan yang paling utama, dimana algoritma yang digunakan akan diimplementasikan untuk mendapatkan hasil yang bermanfaat dan tersembunyi dari data yang dilakukan analisis. Algoritma C4.5 merupakan salah satu Teknik decision tree yang digunakan untuk membuat pohon keputusan agar akurasi dapat dibuat lebih baik. Algoritma C4.5 menggunakan *gain ratio* sebagai parameter untuk mendapatkan value variable yang digunakan untuk membuat sub bagian dari pohon keputusan. Rumus untuk mendapatkan *Gain Ratio* adalah sebagai berikut:

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)}$$

Langkah selanjutnya adalah menghitung SplitInfo, menentukan nilai entropy dengan rumus:

$$SplitInfo(S, A) = - \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S}$$

Langkah selanjutnya adalah menghitung nilai Gain dengan menggunakan rumus:

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

Langkah selanjutnya adalah menghitung nilai Entropy dengan rumus:

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

6. Evaluasi (Evaluation), Tahapan evaluasi merupakan tahapan terakhir pada proses data mining. Proses ini akan melakukan evaluasi dari hasil data mining menggunakan metode yang ditentukan untuk mencari apakah pola yang didapatkan sesuai dengan hipotesa yang diharapkan atau diteliti.

3.3 Confusion Matrix

Pengujian akurasi pada perhitungan data mining biasanya menggunakan confusion matrix yang aman merupakan alat yang digunakan dalam melakukan pengujian evaluasi hasil model klasifikasi untuk menilai apakah hasil akurat atau tidak [17], [18]. Tabel kelas prediksi dapat dilihat pada Tabel 1.

Tabel 1. Tabel Confusion Matrix

		Kelas Prediksi	
		1	0
Kelas	1	TP	FN
Sebenarnya	0	FP	TN

Perhitungan *confusion matrix* menggunakan rumus sebagai berikut.

- a. Akurasi diperoleh dari nilai TP + nilai TN lalu dibagi dengan total nilai keseluruhan [19].

$$Accuracy = \frac{TP+TN}{Total}$$

- b. Presisi digunakan untuk mengukur data yang telah sesuai, presisi diperoleh dari nilai TP dibagi dengan nilai TP+FP [20].

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

- c. Recall digunakan untuk mengetahui nilai keterberhasilan dan diperoleh dari nilai TP dibagi dengan nilai TP+FN [21].

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

4. Hasil dan Pembahasan

4.1. Hasil

Perhitungan Algoritma C4.5 berdasarkan dataset yang digunakan terdapat beberapa kriteria antara lain, Pekerjaan, Pinjaman, Jaminan, Pendapatan, Pengeluaran, Karakter, Kemampuan dan Terealisasi. Kriteria yang digunakan akan dihitung untuk mendapatkan hasil Gain Ratio terbesar yang mana akan digunakan sebagai batas dalam pohon keputusan. Pertama akan menghitung nilai total *entropy* dan hasilnya dapat dilihat pada Tabel 2.

Tabel 2. Hasil Perhitungan Total *Entropy*

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Ya)	Entropy (Tidak)	Entropy (Total)
100	73	27	0,32	0,51	0,84

Selanjutnya menghitung nilai *Entropy*, *Gain*, *SplitInfo* dan *Gain Ratio* pada setiap jaminan. Hasil dari perhitungan nilai tersebut pada masing-masing kriteria dapat dilihat pada Tabel 3 sampai dengan Tabel 9. Dari hasil perhitungan akan dilanjutkan terus-menerus sehingga akan mendapatkan node yang akan menghasilkan pohon keputusan akhir yang akan digunakan untuk LPD Desa Adat Sumerta sebagai pertimbangan dalam pengambilan keputusan pemberian kredit agar tidak terjadinya kredit macet kedepannya.

Tabel 3. Nilai *Entropy*, *Gain*, *SplitInfo*, dan *Gain Ratio* Pada Kriteria Pekerjaan

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	0.030	2.03	0.014

Tabel 4. Nilai *Entropy*, *Gain*, *SplitInfo*, dan *Gain Ratio* Pada Kriteria Pinjaman

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	1.62	0.12	13.17

Tabel 5. Nilai Entropy, Gain, SplitInfo, dan Gain Ratio Pada Kriteria Jaminan

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	1.41	1.83	0.77

Tabel 6. Nilai Entropy, Gain, SplitInfo, dan Gain Ratio Pada Kriteria Pendapatan

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	1.66	0.080	20.59

Tabel 7. Nilai Entropy, Gain, SplitInfo, dan Gain Ratio Pada Kriteria Pengeluaran

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	1.65	0.14	11.73

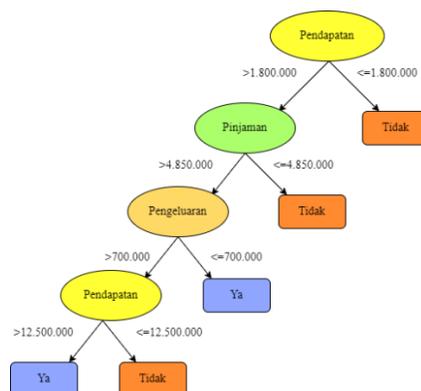
Tabel 8. Nilai Entropy, Gain, SplitInfo, dan Gain Ratio Pada Kriteria Karakter

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	0	0	0

Tabel 9. Nilai Entropy, Gain, SplitInfo, dan Gain Ratio Pada Kriteria Kemampuan

Total Kasus	Jumlah (Ya)	Jumlah (Tidak)	Entropy (Total)	Gain	SplitInfo	Gain Ratio
100	73	27	0.84	1.59	0.36	4.36

Hasil dari perhitungan yang telah didapatkan dari seluruh nilai dari setiap kriteria selanjutnya akan dijadikan node-node yang akan menjadi sebuah pohon keputusan. Node yang dipilih untuk menjadi node ataupun sub node yaitu dari nilai Gain Ratio tertinggi dari perhitungan setiap node. Untuk node pertama pertama yaitu terdapat pada kriteria Pendapatan dengan nilai Gain Ratio tertinggi yaitu 20,59 dengan nilai kriteria pendapatan $\leq 1.800.000$, dan perhitungan selanjutnya data yang digunakan sebagai ambang batas perhitungan yaitu nilai pendapatan $\leq 1.800.000$. Perhitungan dilakukan sampai dengan mendapatkan node terakhir dan seluruh data terhitung. Berikut hasil dari pohon keputusan yang telah didapatkan dalam perhitungan data mining menggunakan algoritma C4.5 dalam penentuan kelayakan kredit dapat dilihat pada Gambar 3.



Gambar 3. Hasil Pohon Keputusan Penentuan Kelayakan Kredit

Dari hasil diatas didapatkan bahwa kriteria pendapatan dengan nilai >12.500.000 (5%) adalah nasabah yang diprioritaskan mendapatkan kredit dan nilai <=12.500.000 (95%) adalah nasabah yang lebih akhir mendapatkan kredit. Minimal pendapatan nasabah yang bisa mendapatkan kredit yaitu 1.800.000.

Hasil yang didapatkan akan dilakukan pengujian yaitu menggunakan *confusion matrix* untuk mendapatkan nilai akurasi, presisi, dan recall. Pengujian *confusion matrix* akan membentuk sebuah matriks yang akan berisi nilai *true positive*, *true negative*, *false positive* dan *false negative*. Hasil dari pengujian yang dilakukan menggunakan *dataset* yang dilakukan perhitungannya sebelumnya mendapatkan bahwa Total nilai TP (73), FN (0), FP (19), dan TN (8). Nilai yang dihasilkan dari matriks diatas yaitu nilai akurasi yaitu 81%, nilai presisi yaitu 100% dan nilai recall yaitu 100%. Hasil dari pengujian *confusion matrix* dapat dilihat pada Gambar 4.

accuracy: 81.00%

	true Ya	true Tidak	class precision
pred. Ya	73	19	79.35%
pred. Tidak	0	8	100.00%
class recall	100.00%	29.63%	

Gambar 4. Hasil Pengujian Confusion Matrix

4.2. Pembahasan

Penerapan data mining dalam melakukan penyelesaian sebuah masalah yang terdapat pada LPD Desa Adat Sumerta mengenai penentuan kelayakan kredit yang diselesaikan menggunakan Algoritma C4.5 menghasilkan hasil yang baik dan sesuai dengan harapan peneliti. Penelitian yang dilakukan sudah sesuai dengan alur penelitian dan teori-teori yang digunakan sebelumnya sehingga dapat berjalan dengan baik. Hasil dari penelitian yang baik dapat memudahkan pihak LPD menggunakan hasil dari perhitungan yang dilakukan guna untuk mengurangi kredit macet kedepannya. Proses perhitungan harus dilalui dengan beberapa tahap dari awal pengumpulan data hingga tahapan akhir yaitu hasil dari pengujian apakah data yang digunakan memiliki akurasi yang baik atau tidak. Penelitian masih hanya menggunakan perhitungan dengan aplikasi *Microsoft excel*. Selanjutnya penelitian akan dikembangkan dengan menggunakan aplikasi data sains lainnya serta metode yang lebih banyak agar menghasilkan perbandingan manakah metode terbaik dan aplikasi terbaik yang digunakan melakukan penerapan data mining.

5. Simpulan

Penelitian yang telah dilakukan mengenai penerapan data mining dengan algoritma C4.5 dalam penentuan kelayakan nasabah kredit (studi kasus: LPD Desa Adat Sumerta) dapat disimpulkan Perhitungan algoritma C4.5 menggunakan aplikasi Microsoft excel (spreadsheet) adalah dengan menyiapkan dataset dan menentukan nilai Entropy, Gain, SplitInfo, dan Gain Ratio dari masing-masing atribut. Nilai Gain Ratio tertinggi akan digunakan sebagai nilai node atau hasil pohon keputusan node. Perhitungan menggunakan beberapa kriteria yaitu Pekerjaan, Pinjaman, Jaminan, Pendapatan, Pengeluaran, Karakter, Kemampuan dan Terealisasi. Selain itu hasil dari perhitungan algoritma C4.5 menggunakan Microsoft Excel mendapatkan nilai pohon keputusan node terakhir yaitu kriteria pendapatan dengan nilai >12.500.000 (5%) adalah nasabah yang diprioritaskan mendapatkan kredit dan nilai <=12.500.000 (95%) adalah nasabah yang lebih akhir mendapatkan kredit. Minimal pendapatan nasabah yang bisa mendapatkan kredit yaitu 1.800.000. Setelah itu dengan dataset yang telah dimiliki juga dilakukan pengujian confusion matrix dengan hasil nilai *accuracy* sebesar 81.00%, nilai *precision* 100.00%, *recall* 100.00%.

Daftar Pustaka

- [1] I. A. U. Dewi, I. K. N. A. Jaya, and I. D. K. L. Digita, "Sistem Informasi Geografis (SIG) Sebaran LPD Di Kota Denpasar Berbasis Web Menggunakan Framework Laravel," *KARMAPATI - Kumpul. Artik. Mhs. Pendidik. Tek. Inform.*, vol. 2 No 3, pp. 224–232, 2022, doi: <https://doi.org/10.23887/karmapati.v11i3.54546>.

- [2] I. D. K. L. Digita and K. O. Sanjaya, "Perancangan Sistem Informasi Analisa Kredit Berbasis Web Menggunakan Metode AHP (Studi Kasus: LPD Desa Adat Sumerta)," *RESI J. Ris. Sist. Inf.*, vol. 1, pp. 10–20, Jul. 2022, doi: <https://doi.org/10.32795/resi.v1i1.2945>.
- [3] LPD Desa Adat Sumerta, "Rencana Kerja dan Anggaran Pendapatan - Belanja," Dec. 2021.
- [4] K. A. K. Saputra and P. B. Anggiriawan, "Kebangkrutan Lembaga Perkreditan Desa (Lpd) Berdasarkan Analisis Berbagai Faktor," *J. Ilm. Akunt.*, vol. 4, no. 1, p. 23, 2019.
- [5] D. M. W. P. Swari, K. K. Yogantara, and I. P. A. A. Negara, "Pengaruh Komponen Pengendalian Intern Terhadap Efektivitas Pemberian Kredit Pada Lembaga Perkreditan Desa (LPD) Di Kabupaten Tabanan," *J. Res. Account.*, vol. 2, pp. 132–145, Jun. 2021.
- [6] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques Third Edition*, Third Edition. [www.elsevier.com, 2012.:](http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf) <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>
- [7] I. Jaya, Fauzi, A. Suryana, A. Widianoro, and I. Digita, "Data Visualization Of House Of Worship Distribution In The IKN Nusantara Region Using Python," *J. Ilm. Merpati*, vol. 11, pp. 1–12, 2023, doi: <https://doi.org/10.24843/JIM.2023.v11.i01.p01>.
- [8] M. A. Rivai, "Analysis of Corona Virus spread uses the CRISP-DM as a Framework: Predictive Modelling," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, p. 2987-2994, Jun. 2020, doi: [10.30534/ijtcse/2020/76932020](https://doi.org/10.30534/ijtcse/2020/76932020).
- [9] D. Marlina and M. Bakri, "Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5," *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 1.
- [10] A. Ramadhan, "Sistem Pendukung Keputusan Evaluasi Problematika Pendampingan Pembelajaran Daring dengan Algoritma C4. 5," *J. Sistim Inf. Dan Teknol.*, 2022, : <http://jsisfotek.org/index.php/JSisfotek/article/view/124>
- [11] R. Wirth and J. Hipp, "CRISP-DM: Towards a Standard Process Model for Data Mining," *Comput. Sci.*, p. 11, 2000.
- [12] H. Susanto and S. Sudiyatno, "Data Mining untuk Memprediksi Prestasi Siswa Berdasarkan Sosial Ekonomi, Motivasi, Kedisiplinan dan Prestasi Masa Lalu," *J. Pendidik. Vokasi*, vol. 4, no. 2, pp. 222–231, Jun. 2014, doi: [10.21831/jpv.v4i2.2547](https://doi.org/10.21831/jpv.v4i2.2547).
- [13] W. Irmayani, "Visualisasi Data Pada Data Mining Menggunakan Metode Klasifikasi Naive Bayes," *J. Khatulistiwa Inform.*, vol. 9, p. 5, 2021, doi: <https://doi.org/10.31294/jki.v9i1.9593>.
- [14] S. Febriani and H. Sulistiani, "Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5," *J. Teknol. Dan Sist. Inf.*, vol. 2, no. 4.
- [15] E. Fitriani, "Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan," *SISTEMASI*, vol. 9, no. 1, p. 103, Jan. 2020, doi: [10.32520/stmsi.v9i1.596](https://doi.org/10.32520/stmsi.v9i1.596).
- [16] I. Junaedi, N. Nuswantari, and V. Yasin, "Perancangan Dan Implementasi Algoritma C4.5 Untuk Data Mining Analisis Tingkat Risiko Kematian Neonatum Pada Bayi," *JISICOM*, vol. 3, pp. 29–44, 2019.
- [17] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," vol. 5, p. 15, 2021.
- [18] W. A. Bagwan and R. S. Gavali, "Delineating changes in soil erosion risk zones using RUSLE model based on confusion matrix for the Urmodi river watershed, Maharashtra, India," *Model. Earth Syst. Environ.*, vol. 7, no. 3, pp. 2113–2126, Sep. 2021, doi: [10.1007/s40808-020-00965-w](https://doi.org/10.1007/s40808-020-00965-w).
- [19] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, "Pengukuran Kinerja Sistem Kualitas Udara Dengan Teknologi Wsn Menggunakan Confusion Matrix," *J. Inform. Upgris*, vol. 6, no. 2, Jan. 2021, doi: [10.26877/jiu.v6i2.6552](https://doi.org/10.26877/jiu.v6i2.6552).

- [20] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode Naïve Bayes dan Support Vector Machine pada Analisis Sentimen Twitter," *SMATIKA J.*, vol. 10, no. 02, pp. 71–76, Dec. 2020, doi: 10.32664/smatika.v10i02.455.
- [21] N. Hadianto, H. B. Novitasari, and A. Rahmawati, "KLASIFIKASI PEMINJAMAN NASABAH BANK MENGGUNAKAN METODE NEURAL NETWORK," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 163–170, Sep. 2019, doi: 10.33480/pilar.v15i2.658.