

IMPLEMENTASI ALGORITMA C5.0 UNTUK MEMPREDIKSI KETERLAMBATAN PEMBAYARAN SUMBANGAN PEMBANGUNAN PENDIDIKAN PADA SMP SWASTA AN-NAAS BINJAI

Mutia Nurhasanah¹, Ilka Zufria²

^{1,2} Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email : nurhasanahmutia131199@gmail.com

Abstrak

Pengelolaan keuangan yang efisien merupakan aspek penting dalam keberhasilan operasional lembaga pendidikan, terutama terkait pembayaran Sumbangan Pembangunan Pendidikan (SPP). Keterlambatan pembayaran SPP dapat mengganggu arus kas sekolah dan menimbulkan masalah administratif. Penelitian ini bertujuan mengembangkan dan mengimplementasikan Algoritma C5.0 untuk memprediksi keterlambatan pembayaran SPP di SMP Swasta An-Naas Binjai. Algoritma C5.0 dipilih karena kemampuannya mengolah data kompleks dan menghasilkan model prediksi yang mudah dipahami. Hasil evaluasi menunjukkan bahwa Algoritma C5.0 mampu memberikan tingkat akurasi sebesar 85%, dengan precision 88.89% dan recall 94.12%. Ini menandakan efektivitas algoritma dalam memprediksi keterlambatan pembayaran SPP serta keandalannya sebagai alat bantu pengelolaan keuangan sekolah. Penelitian ini juga menghasilkan sistem yang diharapkan dapat meningkatkan efisiensi pengelolaan keuangan dan memberikan wawasan lebih baik mengenai faktor-faktor yang mempengaruhi keterlambatan pembayaran SPP

Kata Kunci: Algoritma C5.0, Prediksi, SPP.

Abstract

Efficient financial management is a crucial aspect of the successful operation of educational institutions, particularly concerning the payment of Educational Development Contributions (SPP). Delays in SPP payments can disrupt the school's cash flow and create administrative issues. This research aims to develop and implement the C5.0 Algorithm to predict delays in SPP payments at SMP Swasta An-Naas Binjai. The C5.0 Algorithm was chosen for its ability to process complex data and generate easily interpretable predictive models. Evaluation results show that the C5.0 Algorithm can achieve an accuracy rate of 85%, with a precision of 88.89% and a recall of 94.12%. This indicates the algorithm's effectiveness in predicting SPP payment delays and its reliability as a tool for school financial management. The research also produced a system expected to enhance financial management efficiency and provide better insights into the factors influencing SPP payment delays.

Keywords: C5.0 Algorithm, Prediction, Tuition Fees (SPP).



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

Pendahuluan

Pendidikan adalah kebutuhan mendasar bagi setiap individu. Pendidikan melibatkan upaya untuk mengembangkan, memperbaiki, dan mengubah pengetahuan serta perilaku individu atau kelompok, dengan tujuan untuk meningkatkan kecerdasan manusia melalui aktivitas pembelajaran [1].

Komponen penting dalam mendukung sistem pendidikan adalah pembayaran Sumbangan Pengembangan Pendidikan (SPP) di sekolah. SPP berfungsi sebagai sumber pendapatan utama untuk berbagai kebutuhan, termasuk pengembangan infrastruktur sekolah, gaji guru, dan biaya operasional

lainnya. Sekolah swasta umumnya membebankan SPP kepada siswa sebagai sumber pendapatan, berbeda dengan sekolah negeri yang menerima dana dari pemerintah. Di SMP Swasta An-Naas Binjai, keterlambatan pembayaran SPP oleh siswa atau orang tua/wali mereka menjadi masalah signifikan yang dapat membahayakan stabilitas keuangan sekolah dan kualitas pendidikan yang ditawarkan [2].

Mengatasi masalah ini memerlukan pendekatan yang proaktif dan berbasis data. Salah satu solusi yang mungkin adalah mengembangkan alat atau model analitis yang dapat memprediksi keterlambatan pembayaran SPP di SMP Swasta An-Naas Binjai. Dengan memprediksi potensi keterlambatan pembayaran secara akurat, sekolah dapat mengidentifikasi siswa atau keluarga yang berisiko dan mengambil tindakan pencegahan, seperti memberikan peringatan dini atau menawarkan bantuan keuangan [3].

Salah satu strategi yang dapat diterapkan adalah penggunaan Data Mining. Data mining adalah proses untuk mengekstraksi informasi yang berguna dari kumpulan data besar guna membantu dalam pengambilan keputusan. Ini melibatkan teknik pembelajaran mesin yang mengeksplorasi pola dalam data untuk mendapatkan wawasan yang berharga. Data mining menggunakan satu atau lebih teknik pembelajaran untuk mengidentifikasi pola dalam data dan memperoleh informasi yang berguna [4]. Selain itu, digunakan untuk menganalisis kumpulan data besar dengan mengungkapkan hubungan yang jelas dan menyimpulkan wawasan yang sebelumnya tidak diketahui [5].

Salah satu algoritma dari data mining adalah algoritma C5.0. Algoritma C5.0 adalah sebuah algoritma pembelajaran mesin yang menghasilkan pohon keputusan untuk analisis prediktif. Algoritma C5.0 adalah alat penambangan data yang secara khusus digunakan untuk membangun pohon keputusan dan merupakan pengembangan dari algoritma sebelumnya yang dikembangkan oleh Ross Quinlan pada tahun 1987, yaitu ID3 dan C4.5 [6]. Selama proses pembentukan pohon keputusan, Algoritma C5.0 mengidentifikasi atribut yang paling penting untuk mengklasifikasikan objek ke dalam kategori yang berbeda, dengan memprioritaskan atribut yang memberikan informasi tertinggi [7].

Dalam penelitian sebelumnya yang berjudul "Aplikasi Algoritma C5.0 untuk Memprediksi Kelulusan Mahasiswa pada Mata Kuliah Arsitektur Sistem Komputer," proses prediksi menggunakan algoritma klasifikasi C5.0 berdasarkan atribut seperti Nilai Kehadiran, Nilai Tugas, Nilai Ujian Tengah Semester, dan Nilai Ujian Akhir. Hasilnya menunjukkan bahwa kinerja algoritma C5.0 mencapai tingkat akurasi yang tinggi sebesar 93,33% [8]. Penelitian lain, "Prediksi Kenaikan Jabatan Karyawan Menggunakan Algoritma C5.0 (Studi Kasus di PT. Mataram Cakra Buana Agung)," menganalisis 403 data karyawan dan menemukan bahwa algoritma C5.0 menghasilkan hasil dengan tingkat akurasi sebesar 90,00%. Hal ini menunjukkan bahwa algoritma C5.0 efektif dalam memprediksi kenaikan jabatan karyawan menggunakan metode pohon keputusan [9].

Dalam penelitian ini, penulis memilih Algoritma C5.0 sebagai model analisis untuk mengatasi masalah keterlambatan pembayaran SPP di SMP Swasta An-Naas Binjai. Algoritma ini dipilih karena reputasinya sebagai alat pembelajaran mesin yang andal dan efektif untuk membuat model prediktif, terutama dalam pemodelan pohon keputusan yang dapat memproyeksikan kemungkinan keterlambatan pembayaran SPP dengan akurasi tinggi.

Berdasarkan studi-studi yang disebutkan di atas, diharapkan penelitian ini tentang "Implementasi Algoritma C5.0 untuk Memprediksi Keterlambatan Sumbangan Pengembangan Pendidikan di SMP Swasta An-Naas Binjai" akan mampu memprediksi keterlambatan pembayaran dengan mempertimbangkan faktor-faktor seperti kondisi ekonomi keluarga dan kebiasaan pembayaran sebelumnya. Hal ini akan memungkinkan institusi pendidikan untuk mengambil tindakan pencegahan tepat waktu, memperkuat efisiensi administrasi keuangan, serta meningkatkan stabilitas keuangan dan kualitas pendidikan.

Metodologi Penelitian

1. Pengumpulan Data

Metode pengumpulan data yang diterapkan dalam penelitian ini dilakukan sebagai berikut:

a. Survey (Kuesioner)

Dalam hal ini penulis akan melakukan survei kepada orang tua atau wali murid yang bertujuan untuk mengidentifikasi penyebab utama keterlambatan pembayaran SPP. Kuesioner yang digunakan mencakup beberapa pertanyaan kunci yang dirancang untuk menggali informasi mengenai frekuensi keterlambatan pembayaran, alasan Utama keterlambatan, serta prioritas

pengeluaran keluarga. Responden diminta untuk menjawab pertanyaan dengan jujur dan terbuka, sehingga data yang diperoleh dapat merepresentasikan kondisi sebenarnya.. Survei dapat dilakukan secara daring atau menggunakan kuesioner. Pertanyaan pada kuesioner sendiri meliputi :

- a. Berapa penghasilan keluarga per bulan?
 - b. Berapa jumlah tanggungan keluarga?
 - c. Apa pekerjaan Ayah?
- b. Wawancara dan Observasi

Dalam hal ini penulis melakukan wawancara dengan pihak terkait di SMP Swasta An-Naas Binjai, seperti manajemen sekolah, staf administrasi, guru, atau orang tua siswa, untuk mendapatkan pemahaman yang lebih mendalam tentang faktor-faktor yang memengaruhi keterlambatan pembayaran. Observasi langsung juga dapat dilakukan untuk mengamati praktik-praktik pembayaran di sekolah.

c. Pengumpulan Data Sekolah

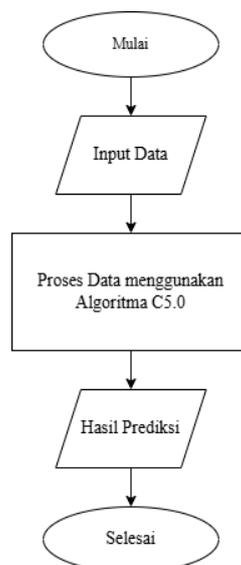
Dalam hal ini penulis mengumpulkan data yang diperoleh dari SMP Swasta An-Naas Binjai. Data ini bertujuan untuk mengidentifikasi penyebab utama keterlambatan pembayaran SPP dan mengumpulkan informasi yang diperlukan untuk analisis menggunakan Algoritma C5.0. Data yang dikumpulkan mencakup beberapa informasi kunci yang dirancang untuk menggali informasi mengenai kondisi ekonomi keluarga siswa, pekerjaan orang tua, dan riwayat pembayaran SPP. Atribut-atribut yang dikumpulkan dari data sekolah meliputi:

- a. Nama Siswa
 - b. Kelas
 - c. Status Pembayaran
- d. Studi Pustaka

Studi Pustaka ialah cara untuk mempelajari tentang apa yang akan penulis teliti, yaitu dengan cara mempelajari kajian terdahulu seperti jurnal, skripsi, tesis dan sebagainya.

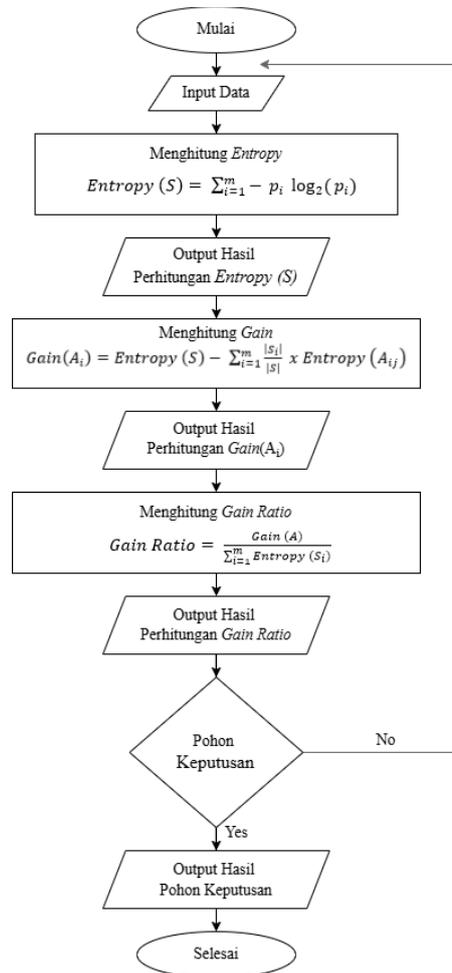
2. Algoritma C5.0

Algoritma C5.0, sebuah metode berbasis pohon keputusan, digunakan untuk menghasilkan model prediktif dari data yang dikumpulkan. Algoritma ini menggunakan information gain untuk memilih atribut yang akan diproses, sehingga meminimalkan kesalahan dalam pengambilan keputusan [10]. Algoritma ini bekerja dengan menghitung rasio gain berdasarkan entropi, memilih atribut dengan rasio gain tertinggi sebagai akar untuk node berikutnya [11]. Pendekatan terstruktur ini diterapkan untuk menganalisis persentase keterlambatan pembayaran, membantu mengidentifikasi faktor kunci yang mempengaruhi keterlambatan ini berdasarkan atribut yang dipilih.



Gambar 1. Diagram Alir Sistem

Algoritma C5.0 menghasilkan output berupa prediksi, yang dapat berupa klasifikasi atau keputusan berdasarkan model yang dikembangkan. Proses ini melibatkan perhitungan entropi, gain, dan rasio gain untuk membangun pohon keputusan yang akurat, sehingga hasil prediksi dapat digunakan untuk mengidentifikasi pola dan tren yang relevan terkait masalah yang sedang dipelajari.



Gambar 2. Diagram Alir Algoritma C5.0

Data dikumpulkan dari administrasi dan orang tua dari 100 siswa di SMP Swasta An-Naas Binjai, termasuk atribut seperti pendapatan, jumlah tanggungan, pekerjaan ayah, dan status pembayaran. Indikator untuk atribut-atribut ini meliputi pendapatan bulanan keluarga, jumlah tanggungan keluarga, jenis pekerjaan ayah, dan status pembayaran sebelumnya.

Tabel 1. Atribut pada *Dataset*

No	Atribut Data	Indikator
1	Pendapatan	Pendapatan Bulanan Keluarga
2	Tanggungan Keluarga	Jumlah Tanggungan Keluarga
3	Pekerjaan Ayah	Tipe Pekerjaan Ayah
4	Status Pembayaran	Riwayat Pembayaran Sebelumnya

Hasil dan Pembahasan

1. Implementasi Algoritma C5.0

Pada proses pembentukan pohon klasifikasi dengan algoritma C5.0, tahap pertama adalah menentukan node akar, kemudian dilanjutkan dengan penentuan cabang untuk masing-masing node. Selanjutnya, dilakukan pembagian kelas pada cabang yang telah diperoleh, dan proses tersebut diulang hingga setiap cabang memiliki kelas. Sebagai contoh, data yang digunakan untuk proses pembentukan

pohon klasifikasi adalah 80% dari keseluruhan data, yaitu sebanyak 80 data sebagai data pelatihan (training data), sedangkan sisanya, yaitu 20% dari keseluruhan data atau sebanyak 20 data, digunakan sebagai data pengujian (testing data) untuk pohon klasifikasi yang telah terbentuk.

1. Langkah pertama dalam proses pembentukan pohon klasifikasi adalah menghitung nilai *entropy*. Berikut perhitungan *entropy* total dan *entropy* pada setiap variabel

a. Menghitung *entropy* total

$$Entropy(total) = \left(\left(-\frac{60}{80} \right) \times \log_2 \left(\frac{60}{80} \right) + \left(-\frac{20}{80} \right) \times \log_2 \left(\frac{20}{80} \right) \right) = 0,8113$$

b. Menghitung *entropy* tiap kategori dari variabel Pekerjaan

$$Entropy(Wiraswasta) = \left(\left(-\frac{13}{16} \right) \times \log_2 \left(\frac{13}{16} \right) + \left(-\frac{3}{16} \right) \times \log_2 \left(\frac{3}{16} \right) \right) = 0,6962$$

$$Entropy(PNS) = \left(\left(-\frac{14}{16} \right) \times \log_2 \left(\frac{14}{16} \right) + \left(-\frac{2}{16} \right) \times \log_2 \left(\frac{2}{16} \right) \right) = 0,5436$$

$$Entropy(TNI / Polri) = \left(\left(-\frac{7}{12} \right) \times \log_2 \left(\frac{7}{12} \right) + \left(-\frac{5}{12} \right) \times \log_2 \left(\frac{5}{12} \right) \right) = 0,9799$$

$$Entropy(Buruh) = \left(\left(-\frac{5}{7} \right) \times \log_2 \left(\frac{5}{7} \right) + \left(-\frac{2}{7} \right) \times \log_2 \left(\frac{2}{7} \right) \right) = 0,8631$$

$$Entropy(Lainnya) = \left(\left(-\frac{21}{29} \right) \times \log_2 \left(\frac{21}{29} \right) + \left(-\frac{8}{29} \right) \times \log_2 \left(\frac{8}{29} \right) \right) = 0,8498$$

2. Langkah selanjutnya adalah menentukan Gain

$$\begin{aligned} Gain(total,pekerjaan) &= Entropy(total) - \left(\left(\frac{16}{80} \right) \times Entropy(Wiraswasta) \right) + \left(\frac{16}{80} \right) \times \\ &\quad Entropy(PNS) + \left(\frac{12}{80} \right) \times Entropy(TNI / Polri) + \left(\frac{7}{80} \right) \times \\ &\quad Entropy(Buruh) + \left(\frac{29}{80} \right) \times Entropy(Lainnya) \right) \\ &= 0,8113 - \left(\left(\frac{16}{80} \right) \times 0,6962 \right) + \left(\frac{16}{80} \right) \times 0,5436 + \left(\frac{12}{80} \right) \times \\ &\quad 0,9799 + \left(\frac{7}{80} \right) \times 0,8631 + \left(\frac{29}{80} \right) \times 0,8498 \right) \\ &= 0,0328 \end{aligned}$$

3. Langkah terakhir adalah menentukan Gain Ratio

$$\begin{aligned} Gain\ Ratio &= \frac{Gain(total, pekerjaan)}{(E(Wiraswasta) + E(PNS) + E(TNI/Polri) + E(Buruh) + E(Lainnya))} \\ &= \frac{0,0328}{((0,6962) + (0,5436) + (0,9799) + (0,8631) + (0,8498))} = 0,0083 \end{aligned}$$

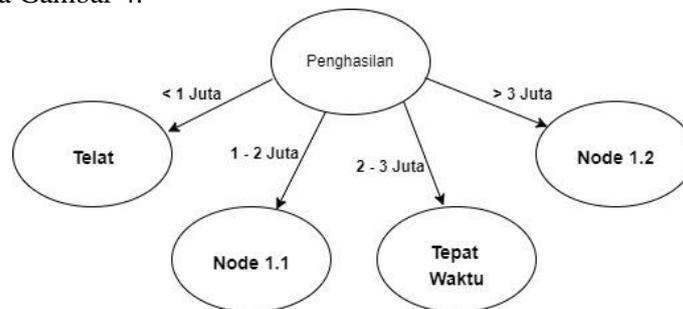
Berikut adalah tabel perhitungan dari Algoritma C5.0 dengan seluruh variabel yang tersedia

Tabel 2. Hasil Perhitungan Node 1

Atribut	Nilai	Jumlah	Tepat Waktu	Terlam-bat	Entropy	Gain	Gain Ratio
Pekerjaan	Wiraswasta	16	13	3	0.6962	0.0328	0.0083
	PNS	16	14	2	0.5436		
	TNI/Polri	12	7	5	0.9799		
	Buruh	7	5	2	0.5917		
	Lainnya	29	21	8	0.5917		

Total		80	60	20	0.8113		
Penghasilan	< 1 Juta	6	1	5	0.65	0.1939	0.0731
	1 – 2 Juta	15	7	8	0.9968		
	2 – 3 Juta	25	23	2	0.4022		
	> 3 Juta	34	29	5	0.6024		
Total		80	60	20	0.8113		
Tanggung- ngan	Banyak	12	9	3	0.8113	0.027	0.011
	Cukup	44	36	8	0.684		
	Sedikit	24	15	9	0.9544		
Total		80	60	20	0.8113		

Dari hasil perhitungan nilai Entropy, Gain dan Gain Ratio seperti terlihat pada tabel di atas dapat diketahui bahwa atribut yang memiliki nilai Gain Ratio tertinggi adalah atribut Penghasilan yaitu sebesar 0,0731 dengan demikian, atribut Penghasilan menjadi node akar. Ada empat nilai yang terdapat pada atribut yaitu < 1 Juta, 1 – 2 Juta, 2 – 3 Juta dan > 3 Juta. Dari keempat nilai tersebut nilai < 1 Juta dinyatakan Telat dan nilai 2 – 3 Juta dinyatakan Tepat Waktu. Sedangkan untuk mencari node 1.1 pada 1 – 2 Juta dan node 1.2 > 3 Juta dilakukan pada iterasi kedua. Dari hasil tersebut dapat digambarkan pohon keputusan pada Gambar 4.



Gambar 3. Pohon Keputusan Node 1 (Penghasilan)

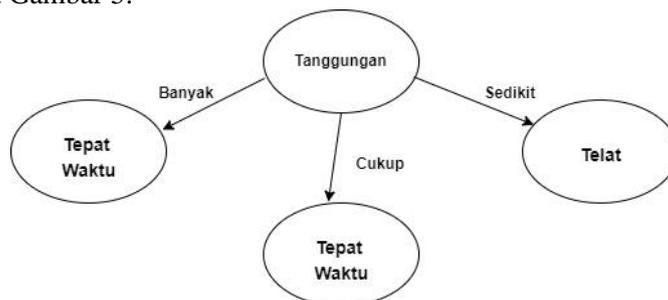
Selanjutnya adalah Berdasarkan penghasilan 1 – 2 Juta, yang dijabarkan pada tabel di bawah ini

Tabel 3. Hasil Perhitungan Node 1.1

Atribut	Nilai	Jumlah	Tepat Waktu	Terlam-bat	Entropy	Gain	Gain Ratio
Pekerjaan	Wiraswasta	4	2	2	1	0.1222	0.0463
	PNS	0	0	0	0		
	TNI/Polri	0	0	0	0		
	Buruh	6	4	2	0.9183		
	Lainnya	5	1	4	0.7219		
Total		15	7	8	0.9968		
Tanggung- ngan	Banyak	2	2	0	0	0.2228	0.1294
	Cukup	8	4	4	1		
	Sedikit	5	1	4	0.7219		
Total		15	7	8	0.9968		

Dari hasil perhitungan nilai Entropy, Gain dan Gain Ratio seperti terlihat pada tabel di atas dapat diketahui bahwa atribut yang memiliki nilai Gain Ratio tertinggi adalah atribut Tanggungan yaitu sebesar 0,1294 dengan demikian, atribut Tanggungan menjadi node akar. Ada tiga nilai yang terdapat pada atribut yaitu Banyak, Sedang, dan Kecil. Dari ketiga nilai tersebut nilai Banyak dan Cukup dinyatakan Tepat Waktu dan nilai Sedikit dinyatakan Telat. Dari hasil tersebut dapat digambarkan

pohon keputusan pada Gambar 5.



Gambar 4. Pohon Keputusan Node 1.1 (Tanggung)

Selanjutnya adalah Berdasarkan penghasilan > 3 Juta, yang dijabarkan pada tabel di bawah ini

Tabel 4. Hasil Perhitungan Node 1.2

Atribut	Nilai	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain	Gain Ratio
Pekerjaan	Wiraswasta	3	3	0	0	0.1157	0.0766
	PNS	14	12	2	0.5917		
	TNI/Polri	9	6	3	0.9183		
	Buruh	0	0	0	0		
	Lainnya	8	8	0	0		
Total		34	29	5	0.6024		
Tanggung	Banyak	8	5	3	0.9544	0.0812	0.0461
	Cukup	16	15	1	0.3373		
	Sedikit	10	9	1	0.369		
	Total	34	29	5	0.6024		

Dari hasil perhitungan nilai Entropy, Gain dan Gain Ratio seperti terlihat pada tabel di atas dapat diketahui bahwa atribut yang memiliki nilai Gain Ratio tertinggi adalah atribut Pekerjaan yaitu sebesar 0,0766 dengan demikian, atribut Pekerjaan menjadi node akar. Ada lima nilai yang terdapat pada atribut yaitu Wiraswasta, PNS, TNI/Polri, Buruh dan Lainnya. Dari kelima nilai tersebut nilai Wiraswasta, PNS dan Lainnya dinyatakan Tepat Waktu. . Sedangkan untuk mencari node 1.2.1 pada TNI/Polri dilakukan pada iterasi Selanjutnya. Dari hasil tersebut dapat digambarkan pohon keputusan pada gambar 6.



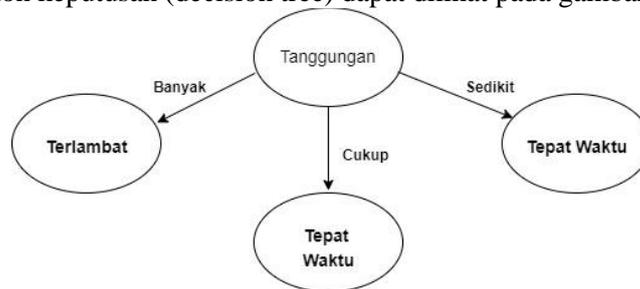
Gambar 5. Pohon Keputusan Node 1.2 (Pekerjaan)

Selanjutnya adalah Berdasarkan pekerjaan TNI/Polri, yang dijabarkan pada tabel di bawah ini

Tabel 5. Hasil Perhitungan Node 1.2.1

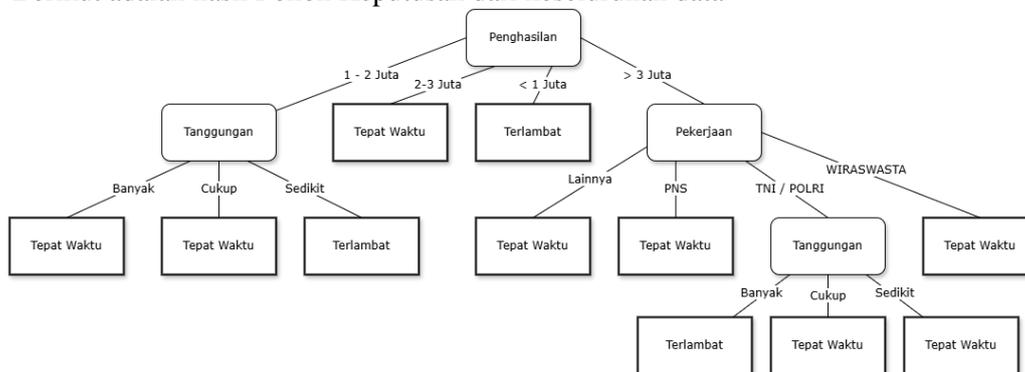
Atribut	Nilai	Jumlah	Tepat Waktu	Terlambat	Entropy	Gain	Gain Ratio
Tanggung	Banyak	3	1	2	0.9183	0.3061	0.1667
	Cukup	3	2	1	0.9183		
	Sedikit	3	3	0	0		
Total		9	6	3	0.9183		

Nilai gain ratio tertinggi pada penghitungan untuk mencari node 1.2.1 diperoleh atribut Tanggungan dengan gain ratio 0,1667. Atribut Tanggungan memiliki 3 nilai dimana nilai Sedikit dan Cukup dinyatakan Tepat Waktu dan Nilai Banyak dinyatakan dengan Terlambat. Dari hasil tersebut dapat digambarkan pohon keputusan (decision tree) dapat dilihat pada gambar 4.4.



Gambar 6. Pohon Keputusan Node 1.2.1 (Tanggungun)

Berikut adalah hasil Pohon Keputusan dari keseluruhan data



Gambar 7. Pohon Keputusan Keseluruhan

Dari pohon keputusan yang terbentuk pada gambar, diperoleh aturan-aturan sebagai berikut:

- a. Jika Penghasilan berada pada kategori 1 - 2 Juta, dan Tanggungan banyak atau cukup, maka pembayaran akan Tepat Waktu. Namun, jika Tanggungan sedikit, pembayaran akan Terlambat.
- b. Jika Penghasilan berada pada kategori 2 - 3 Juta, maka pembayaran akan Tepat Waktu.
- c. Jika Penghasilan berada pada kategori < 1 Juta, maka pembayaran akan Terlambat.
- d. Jika Penghasilan berada pada kategori > 3 Juta, maka kita perlu melihat Pekerjaan:
 - 1) Jika Pekerjaan adalah Lainnya atau PNS, maka pembayaran akan Tepat Waktu.
 - 2) Jika Pekerjaan adalah TNI/POLRI, maka kita perlu melihat Tanggungan:
 - a) Jika Tanggungan banyak atau sedikit, maka pembayaran akan Tepat Waktu.
 - b) Jika Tanggungan cukup, maka pembayaran akan Terlambat.
 - 3) Jika Pekerjaan adalah Wiraswasta, maka pembayaran akan Tepat Waktu.

2. Confusion Matrix

Confusion matrix berguna untuk mengukur hasil evaluasi dari algoritma yang digunakan berupa tabel pada program. Tabel ini menampilkan jumlah data yang diklasifikasikan benar atau salah dengan nilai yang dihasilkan diantaranya True Positive, True Negative, False Positive, dan False Negative.

Tabel 6. Tabel Confusion Matrix

Correct Classification	Classification	
	Positif	Negatif
Positif	TP (True Positive)	TN (True Negative)
Negatif	FP (False Positive)	FN (False Negative)

Dimana :

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$\text{Presisi} = \frac{TP}{TP + FP} \times 100\%$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\%$$

Keterangan:

1. TP (True Positive) = Menampilkan jumlah data positif yang diprediksi benar sesuai dengan nilai aktual
2. FP (False Positive) = Menampilkan jumlah data aktual Negative tetapi nilai prediksi menampilkan hasil positif
3. FN (False Negative) = Menampilkan jumlah data actual positive tetapi nilai prediksi menampilkan hasil negatif
4. TN (True Negative) = Menampilkan jumlah data yang diprediksi negatif sesuai dengan nilai aktual negatif

Berikut tabel *confusion matrix* untuk mengukur tingkat ketepatan, tingkat presisi dan tingkat sensitifitas dari hasil klasifikasi dengan menggunakan algoritma C5.0 dan menggunakan data *testing*

Tabel 7. *Confusion Matrix* dari model C5.0

Correct Classification	Classification	
	Tepat Waktu	Terlambat
Tepat Waktu	16	1
Terlambat	2	1

Dari tabel di atas, diperoleh hasil akurasi, presisi dan *recall* sebagai berikut:

1. Akurasi

$$\text{Akurasi} = \frac{16 + 1}{16 + 2 + 1 + 1} \times 100\%$$

$$\text{Akurasi} = 0.85 * 100\%$$

$$\text{Akurasi} = 85 \%$$

2. Presisi

$$\text{Presisi} = \frac{16}{16 + 2} \times 100\%$$

$$\text{Presisi} = 0.89 * 100\%$$

$$\text{Presisi} = 89\%$$

3. Recall

$$\text{Recall} = \frac{16}{16 + 1} \times 100\%$$

$$\text{Recall} = 0.94 * 100\%$$

$$\text{Recall} = 94 \%$$

Kesimpulan

Berdasarkan hasil akhir dari percobaan yang dilakukan dalam menyelesaikan masalah dari penelitian maka peneliti akan menguraikan beberapa kesimpulan dari penelitian terkait implementasi algoritma C5.0 untuk memprediksi keterlambatan pembayaran sumbangan pembangunan pendidikan di SMP Swasta An-naas Binjai

1. Algoritma C5.0 berhasil diimplementasikan untuk memprediksi keterlambatan pembayaran SPP. Hasil pengujian menunjukkan bahwa model ini memiliki tingkat akurasi sebesar 85%, precision sebesar 88.89%, dan recall sebesar 94.12%. Hasil ini menandakan bahwa Algoritma C5.0 memiliki kemampuan yang baik dalam mengidentifikasi siswa yang berpotensi mengalami keterlambatan pembayaran SPP, sekaligus meminimalisir kesalahan prediksi.
2. Dengan tingkat akurasi 85%, precision 88.89%, dan recall 94.12%, Algoritma C5.0 terbukti efektif dalam memprediksi keterlambatan pembayaran SPP. Tingkat recall yang tinggi

menunjukkan bahwa algoritma ini sangat baik dalam mendeteksi kasus keterlambatan pembayaran, sedangkan precision yang tinggi menunjukkan bahwa prediksi yang dihasilkan cukup tepat sasaran.

3. Sistem yang dirancang mampu menganalisis keterlambatan pembayaran SPP dengan menggunakan Algoritma C5.0 secara efisien. Sistem ini tidak hanya memudahkan dalam mendeteksi potensi keterlambatan tetapi juga membantu lembaga pendidikan dalam mengambil langkah-langkah preventif untuk mengatasi masalah tersebut..

Ucapan Terima Kasih

Pertama-tama, saya mengucapkan terima kasih kepada Allah SWT, Tuhan seluruh alam, yang telah mengizinkan saya menyelesaikan penelitian ini. Selanjutnya, terima kasih kepada para dosen yang selalu mendukung dan memberikan bimbingan selama proses penelitian ini. Tak lupa, terima kasih kepada orang tua yang senantiasa memberikan semangat dalam menyelesaikan penelitian ini.

Daftar Pustaka

- [1] E. Ruli and E. Indarini, “Meta Analisis Pengaruh Model Pembelajaran Problem Based Learning Terhadap Kemampuan Berpikir Kritis Dalam Pembelajaran Matematika Di Sekolah Dasar,” *Jurnal Pendidikan dan Konseling*, vol. 4, no. 4, pp. 221–228, 2022.
- [2] V. S. Ginting, Kusri, and E. Taufiq, “IMPLEMENTASI ALGORITMA C4.5 UNTUK MEMREDIKSI KETERLAMBATAN PEMBAYARAN SUMBANGAN PEMBANGUNAN PENDIDIKAN SEKOLAH MENGGUNAKAN PYTHON,” *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 1, pp. 36–44, 2020.
- [3] D. M. Alfiansyah and H. Soetanto, “Prediksi Keterlambatan Pembayaran SPP Siswa dengan Pendekatan Metode Naive Bayes dan K-Nearest Neighbors,” *Technology and Science (BITS)*, vol. 5, no. 4, pp. 706–719, 2024, doi: 10.47065/bits.v5i4.4643.
- [4] R. N. Amalda, N. Millah, and I. Fitriana, “Implementasi Algoritma C5.0 Dalam Menganalisa Kelayakan Penerima Keringanan UKT Mahasiswa ITK,” *Teorema (Teori dan Riset Matematika)*, vol. 7, no. 1, 2022.
- [5] A. I. Putri, M. Furqan, and Suhardi, “Application of Data Mining to Predict Birth Rates in Medan City Using the K-Nearest Neighbor Method,” *Journal of Computer Science, Information Technology and Telecommunication Engineering (JCoSITTE)*, vol. 5, no. 1, pp. 504–509, 2024, doi: 10.30596/jcositte.v5i1.17991.
- [6] D. P. Utomo and S. Aripin, “Penerapan Algoritma C5.0 Untuk Mengetahui Pola Kepuasan Mahasiswa di Masa Pembelajaran Daring,” *Prosiding Seminar Nasional Riset Dan Information Science (SENARIS)*, vol. 3, pp. 7–12, 2021.
- [7] N. Zamasi, “Implementasi Algoritma C 5.0 Pada Analisa Data Potensi Pertanian dan Perternakan,” *TIN: Terapan Informatika Nusantara*, vol. 2, no. 4, pp. 184–190, 2021, [Online]. Available: <https://ejurnal.seminar-id.com/index.php/tin>
- [8] M. S. Sungkar and M. T. Qurohman, “Penerapan Algoritma C5.0 Untuk Prediksi Kelulusan Pembelajaran Mahasiswa Pada Matakuliah Arsitektur Sistem Komputer,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 3, pp. 1166–1172, Jul. 2021, doi: 10.30865/mib.v5i3.3116.
- [9] I. Nawangsih, I. Melani, and S. Fauziah, “PREDIKSI PENGANGKATAN KARYAWAN DENGAN METODE ALGORITMA C5.0 (STUDI KASUS PT. MATARAM CAKRA BUANA AGUNG),” *Jurnal Pelita Teknologi*, vol. 16, no. 2, pp. 24–33, 2021.
- [10] M. D. Kahfi, F. R. Umbara, and H. Ashaury, “Prediksi Pengangguran Menggunakan Decision Tree Dengan Algoritma C5.0 Pada Data Penduduk Kecamatan Caringin Kabupaten Bogor,” *INFORMATICS AND DIGITAL EXPERT (INDEX)*, vol. 4, no. 2, pp. 75–80, 2022, [Online]. Available: <https://e-journal.unper.ac.id/index.php/informatics>
- [11] K. S. Ningsih and I. Zufria, “Penerapan Algoritma C5.0 Untuk Memprediksi Tingkat Kepuasan Siswa Terhadap Kinerja Guru MAN Simalungun,” *JURNAL FASILKOM*, vol. 13, no. 3, pp. 406–415, 2023.