

Analisa Algoritma C.45 Terhadap Penentuan Rekomendasi Penerima Beasiswa SMP Swasta Methodist-8 Medan

Sardo Pardingotan Sipayung¹, Tulus Pramita Sihaloho², Ade Purba³, Jenheri Rejeki Tarigan⁴

^{1,2,3,4}Universitas Mandiri Bina Prestasi

Jl. Letjend. Djamin Ginting No.285-287, Padang Bulan, Medan Baru, Kota Medan, Sumatera Utara – Indonesia 20155

¹pinsarsiphom@gmail.com, ²sihalohotulus@gmail.com, ³adelinhar31@gmail.com, ⁴jenherirejekitarigan@gmail.com

DOI: 10.58918/lofian.v3i1.217

Abstrak

Beasiswa merupakan bantuan keuangan dengan maksud untuk dipergunakan sebagai sarana melanjutkan pendidikan yang ditempuh dan biasanya diberikan oleh yayasan, perusahaan maupun lembaga pemerintah. Beasiswa berupa dana yang digunakan untuk membantu siswa kurang mampu dalam melanjutkan tugasnya menyelesaikan pendidikan. Oleh karena itu pemberian beasiswa harus tepat sasaran kepada penerima yang benar-benar layak dan pantas mendapatkannya. Banyaknya calon penerima menjadikan proses seleksi memakan waktu yang lama. Dalam hal ini penggunaan metode data mining dapat dijadikan salah satu solusi untuk mempermudah proses seleksi. Algoritma C4.5 merupakan algoritma yang akan dipakai dalam penelitian ini. Data yang digunakan adalah nama siswa, kelas, penghasilan orang tua, tanggungan orang tua, dan nilai rata-rata rapor. Pemrosesan data mining pada suatu data training akan menghasilkan pohon keputusan. Metode evaluasi yang dilakukan dalam pengujian ini didapatkan data nilai akurasi 93.33 %, hal ini bisa menjadi bukti bahwa algoritma C4.5 cukup akurat untuk memberikan rekomendasi beasiswa.

Kata Kunci: Algoritma C4.5, Beasiswa, Data Mining, Klasifikasi, Pohon Keputusan.

1. Pendahuluan

Dalam kehidupan pendidikan memiliki peranan yang sangat penting, Pendidikan itu penting bagi seseorang karena dengan adanya pendidikan, maka seseorang dapat menjadi produktif. Pendidikan sangat penting karena pendidikan dapat memberi seseorang sebuah keterampilan dan hal yang dibutuhkan untuk bisa berhasil dalam hidup. Inilah sebabnya mengapa pendidikan memainkan peran besar tidak hanya bagi siswa tetapi juga orang dewasa. Dengan adanya pendidikan, maka seseorang bisa menafkahi keluarga mereka. Pendidikan membantu individu membuat keputusan yang baik dan meningkatkan peluang mereka untuk berhasil dalam hidup. Ini juga penting bagi masyarakat. Ini dapat membantu orang tumbuh sebagai anggota masyarakat yang produktif, berdampak positif pada ekonomi dan mengurangi tingkat kejahatan secara signifikan. Namun pendidikan seringkali tidak berjalan dengan baik sebab beberapa faktor problematis menjadi penyebab putus sekolah, seringkali alasan finansial menjadi alasan utama putus sekolah.

Dalam menyikapi permasalahan perekonomian yang membuat banyak anak putus sekolah, Pemerintah meluncurkan program beasiswa.

Pada Tahun 2003 tentang Sistem Pendidikan Nasional, menyebutkan bahwa setiap peserta didik pada setiap satuan pendidikan berhak mendapatkan beasiswa bagi yang berprestasi yang orang tuanya tidak mampu membayai pendidikannya. Maka bagi Para siswa yang kurang mampu memiliki hak dan kesempatan untuk memperoleh biaya pendidikan dan bagi mereka yang memiliki prestasi akademik maupun non akademik juga memiliki hak untuk mendapatkan beasiswa (Heni Sulistiani, 2018)

Banyaknya siswa dan siswi yang kurang mampu di SMP Methodist-8 Medan dan tidak sanggup untuk membayar biaya pendidikan sehingga proses penentuan penerimaan bantuan beasiswa dilihat dari keikutsertaan dalam program Indonesia Pintar (PIP), nilai rata-rata rapor, penghasilan dan jumlah tanggungan orang tua.

Pemerintah Kota Medan memiliki program yang bernama Beasiswa berprestasi dengan tujuan untuk mengurangi halangan siswa kurang mampu untuk bersekolah dengan memberi bantuan kepada siswa tersebut untuk memperoleh pendidikan layak. Adapun Calon Penerima Beasiswa berprestasi merupakan siswa yang belum menerima bantuan beasiswa

Program Indonesia Pintar dari pemerintah pusat. Calon penerima Beasiswa Berprestasi diutamakan adalah siswa yang duduk di kelas VII.

Selama proses pengajuan calon peserta penerima Beasiswa Berprestasi SMP Methodist-8 Medan, pihak sekolah mengalami kendala saat menentukan calon penerima Beasiswa Berprestasi, hal ini disebabkan kuota yang ditetapkan oleh pemerintah kota untuk sekolah dibatasi sedangkan total jumlah siswa tahun pelajaran 2022/2023 di kelas VII ada 56 siswa. Prosedur pengolahan data meliputi penghimpunan data, verifikasi, seleksi data dan penyusunan laporan dilakukan secara manual oleh wali kelas. Sehingga proses ini memakan waktu yang relatif lama hingga beberapa pekan.

Melihat hasil uraian dari latar belakang masalah, maka dapat diketahui rumusan masalah dalam penelitian ini, yaitu bagaimana menerapkan metode Algoritma C4.5 untuk memperkirakan calon penerima beasiswa berprestasi di SMP Methodist-8 Medan sehingga bisa menjadi pendukung keputusan oleh pihak sekolah dalam proses pemberian beasiswa yang akan datang. Beasiswa mendukung biaya pendidikan bagi pelajar sehingga mereka dapat melanjutkan pendidikan yang lebih tinggi dengan mempertimbangkan prestasi dan potensi akademik (Hapsari et al., 2018) walaupun di satu sisi beasiswa diberikan kepada pelajar yang kurang mampu.

Data mining memiliki makna suatu aktifitas penguraian dari sekumpulan data menjadi suatu informasi yang berpotensi secara tersirat (implisit) yang sebelumnya belum diketahui (Lailil Muflikhah, Dian Eka Ratnawati, 2018). Data mining merupakan sebuah analisis yang memiliki sifat langsung dari data besar atau kompleks dengan maksud untuk mendapatkan kecenderungan atau pola penting dan umumnya keberadaannya kurang disadari (Aprilia C et al., 2013).

Klasifikasi dapat berarti suatu proses menemukan fungsi ataupun model yang menjadi pembeda dan menggambarkan kelas ataupun konsep data (Anjar Wanto, Muhammad Noor Hasan Siregar, Agus Perdana Windarto, Dedy Hartama, Ni Luh Wiwik Sri Rahayu Ginantra, Darmawan Napitupulu, Edi Surya Negara, Muhammad Ridwan Lubis, Sarini Vita Dewi, 2020).

Rapidminer adalah sebuah perangkat lunak opensource dan diciptakan oleh Raif Klinkenberg dan Dr. Markus Hofmann dari Institute of Technology Blanchardstown bersama rapid-i.com dengan Graphical User Interface (GUI) sehingga mempermudah para penggunanya dalam memanfaatkan perangkat lunak tersebut. Rapidminer dibuat dengan javascript dan mampu dioperasikan disemua sistem operasi komputer. Aplikasi ini tidak

membutuhkan keterampilan pengkodean khusus, karena semua fasilitas atau fitur sudah tersedia dan biasanya digunakan memproses data mining (Hendrian, 2018).

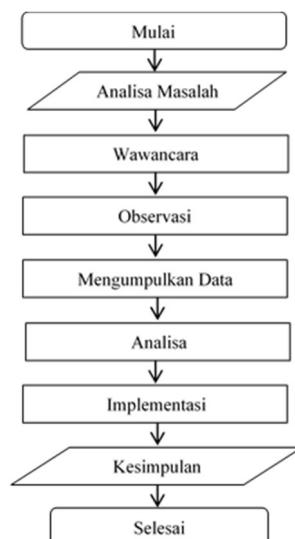
C4.5 adalah sebuah algoritma pembentuk pohon keputusan dan sebuah tatacara prediksi serta klasifikasi terkenal dan terkuat. Metode pohon keputusan atau yang biasa disebut decision tree mampu mentransformasikan fakta sangat luas menjadi suatu pohon keputusan dengan aturan-aturan yang mudah dipahami (Yunus et al., 2021).

Decision tree adalah metode yang dipakai untuk proses penalaran supaya diperoleh solusi dari suatu masalah. Pohon yang terbentuk tidak mesti berbentuk biner. Jika fasilitas dalam dataset memakai dua jenis nilai kategori maka pohon yang diperoleh akan berwujud pohon biner namun apabila memakai jenis numerik atau berisi lebih dari dua jenis nilai kategori maka wujud dari pohon yang didapat umumnya bukan berwujud pohon biner (Prasetyowati, 2017).

2. Metode Penelitian

2.1. Rancangan Penelitian

Rancangan penelitian merupakan beberapa rancangan atau alur dari Analisa Algoritma C.45 Terhadap penentuan Rekomendasi Penerima Beasiswa SMP Swasta Methodist-8 Medan Jl. K.L. Yosudarso yang dimulai dari pengambilan data hingga pengolahan data yang akan dibuat sehingga outputnya lebih jelas, adapun rancangan ini dapat dilihat dalam rancangan flowchart sebagai berikut pada Gbr. 1. berikut.



Gbr. 1. Rancangan Penelitian

2.2. Algoritma C.45

Algoritma C.4.5 adalah algoritma yang digunakan untuk membentuk pohon keputusan, yang merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturanaturan dapat dengan mudah dimengerti dalam bahasa alami. (Sugiyarti, n.d.).

Dengan penjelasan di atas algoritma C.4.5 digunakan untuk mempermudah peneliti memprediksi sebuah nilai menjadi informasi, dan menurut peneliti algoritma C.4.5 merupakan perngembangan dari ID3 dengan membentuk pohon keputusan untuk memprediksi atau memperkirakan sebuah kejadian mendapatkan informasi.

Hasil dari algoritma C.4.5 adalah pohon keputusan, sehingga mudah untuk dipahami dan proses penggunaan algoritma ini untuk menyelesaikan kasus adalah sebagai berikut (Azahari, 2021):

1) Mencari nilai Gain

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

Dimana keterangan dari rumus tersebut adalah:

S = Himpunan Kasus
A = atribut n = jumlah partisi
Pi = porsi dari si kepada s
|si| = jumlah kasus pada nilai atribut/kriteria
|s| = jumlah kasus atau total kasus

2) Mencari nilai Entropy

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

Keterangan:

S : Himpunan Kasus
N : Jumlah Partisi S
Pi : Proporsi dari Si terhadap S

3) Decision Tree

Decision tree dalam (Prasetyowati, 2017) adalah suatu metode pengelompokan atau classification data yang sangat mudah dimengerti. Dalam perkembangan decision tree dijadikan sebagai acuan penalaran untuk mempermudah memahami dari permasalahan yang dihadapi. Metode ini sangat membantu untuk menggambarkan aturan yang dapat dengan mudah dipahami dengan bahasa alami dan juga dapat diekspresikan dalam bentuk bahasa basis data seperti

Structured Query Language untuk mencari record pada kategori tertentu.

Decision tree menurut peneliti merupakan metode pengelompokan atau klasifikasi untuk mempermudah model klasifikasi yang tidak relevan dan mudah dimengerti. Salah satu metode decision tree yang peneliti ambil yaitu C.4.5 yang termasuk dari bagian metode decision tree terpopuler.

Dalam referensi (Prasetyowati, 2017) Pohon Keputusan adalah pohon yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari masalah yang dimasukkan. Pohon yang dibentuk tidak selalu berupa pohon biner. Jika semua fitur dalam data set menggunakan 2 macam nilai kategorikal maka bentuk pohon yang didapatkan berupa pohon biner. Jika dalam fitur berisi lebih dari 2 macam nilai kategorikal atau menggunakan tipe numerik maka bentuk pohon yang didapatkan biasanya tidak berupa pohon biner.

2.3. Teknik pengumpulan data

Teknik pengumpulan data di penelitian ini untuk mendapatkan sebuah informasi yang dibutuhkan dalam mencapai sebuah penelitian. Teknik pengumpulan data dalam penelitian ini menggunakan observasi, studi Pustaka, dan wawancara.

1) Observasi

Dalam penelitian ini, peneliti terlibat dengan orang yang sedang diamati atau digunakan sebagai sumber data penelitian (Prof.Dr.Sugiono, 2017). Metode pengumpulan data yang digunakan untuk mengamati langsung, melihat dan mengambil suatu data di sekolah SMP Swasta Methodist-8 Medan.

2) Studi Pustaka

Data dalam penelitian ini kebanyakan diperoleh dari sumber manusia atau human resources, melalui observasi. Metode ini di gunakan untuk mengumpulkan data berupa sejarah singkat sekolah, data-data siswa, sarana dan prasarana sekolah. Teknik ini untuk mendapatkan data yang berkenan dengan judul penelitian ini.

3) Wawancara

Untuk memahami aturan yang perlu ditetapkan dalam mengidentifikasi calon penerima Beasiswa, penelitian ini mewawancarai wakil kepala sekolah bidang kesiswaan di SMP Swasta Methodist-8 Medan.

2.4. Tahapan Penelitian

Dalam penelitian ini analisis data di lakukan dengan tahapan proses Discovery in Databases (KDD), tahapan tersebut dijelaskan sebagai berikut:

1) Data Selection

Tahapan ini data yang didapat dari SMP Swasta Methodist-8 Medan adalah data calon penerima beasiswa, dengan jumlah 50 data dan atribut variabel yang ada sebanyak 3 atribut data yang digunakan sebagai berikut:

Tabel 1
 Keterangan Atribut

Atribut	Penggunaan
Nama Siswa	No
Tanggungan Orang Tua	Variabel Bebas
Penghasilan Orang Tua	Variabel Bebas
Menerima	Variabel Target

2) Processing (Cleaning)

Dengan 3 atribut predictor dan 1 atribut target dan ada perubahan nama atribut yang bertujuan mempersingkat penyebutannya data terbut di urakan pada tabel dibawah ini:

Tabel 2
 Tanggungan Orang Tua

1	Sedikit
2-3	Cukup
>4	Banyak

Tabel 3
 Penghasilan Orang Tua

< 1.000.000	Sedikit
1.000.000 – 2.000.000	Cukup
> 2.000.000	Banyak

3) Transformation

Dalam tahap ini data tidak ada yang di bersihkan, setelah dilakukan proses tersebut dilakukan pembagian data set dengan menggunakan process split data untuk menghasilkan data training dan data testing dengan rasio 70:30. Data training berjumlah 30 dan data testing berjumlah 20 dari total keseluruhan yaitu 50 data.

4) Data Mining

Pada tahap ini, data diproses, diklasifikasikan menurut model, dan kemudian dibuat beberapa aturan.

Penelitian ini menggunakan langkah-langkah dalam algoritma C4.

3. Hasil dan Pembahasan

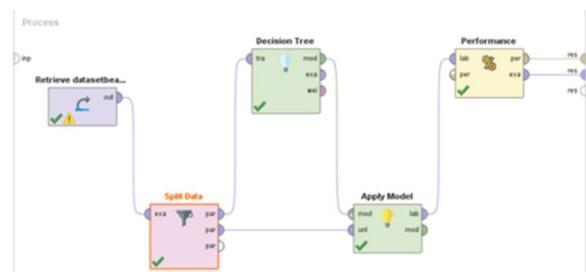
Hasil penelitian ini diharapkan untuk menentukan nilai akurasi Algoritma C4.5 dalam penyeleksian calon penerima beasiswa yang memenuhi kriteria, dengan mengolah data dan memilih atribut yang diperlukan, kemudian menguji data tersebut dengan perhitungan manual dan perangkat lunak Rapidminer. Hasil dari penelitian ini berupa proses perhitungan berdasarkan algoritma C4.5.

3.1. Pengujian algoritma C4.5

Pembuatan algoritma C4.5 dilakukan pada kumpulan dataset, dengan data yang sudah di jelaskan sebelumnya pengujian data dapat dihitung dengan beberapa langkah-langkah perhitungan Entropy dan Gain pada setiap atribut yang dimiliki diterima yang sebagai variabel target. Dengan menghitung nilai gain disetiap atribut, lalu dengan membandingkan nilai gain tersebut yang nilainya tertinggi akan digunakan sebagai simpul akar atau atribut akar.

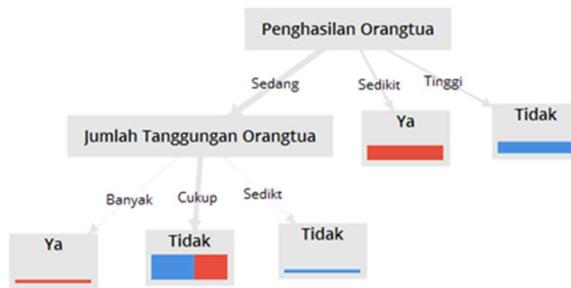
3.2. Pengujian Data Dengan RapidMiner

Pada proses pengujian data yang sudah di import, peneliti menggunakan data training dan data testing untuk menentukan sebuah model decision tree dengan menggunakan perbandingan 70% untuk data training dan 30% testing.



Gbr. 2. Proses Pengujian Data

3.3. Hasil Pemodelan Pohon Keputusan



Gbr. 3. Hasil Pohon Keputusan

Pohon keputusan C4.5 yang sudah di lihat pada gambar 2 dibaca dari atas ke bawah atau dari akar (simpul pertama paling atas) sampai ke daun (simpul terluar yang tidak lagi memiliki cabang). Berikut cara membacanya dengan mengacu pada tiap simpulnya. Jika Penghasilan orangtua Sedikit maka diterima dan jika penghasilan orangtua tinggi maka tidak diterima tanpa perlu lagi melihat atribut lainnya. Jika Penghasilan Orangtua sedang maka terdapat tiga kemungkinan, pertama jika Tanggungan orangtua sedikit maka tidak diterima, kedua jika tanggungan orangtua cukup maka tidak diterima, ketiga jika tanggungan orangtua banyak maka diterima.

3.4. Hasil Accuracy

accuracy: 93.33%			
	true Tidak	true Ya	class precision
pred. Tidak	5	0	100.00%
pred. Ya	1	9	90.00%
class recall	83.33%	100.00%	

Gbr. 4. Hasil Accuracy

Hasil pengukuran hasil accuracy data yang diperoleh dari data training dapat dilihat pada gambar 4 dengan nilai accuracy mencapai 93,33%. Dari tabel dapat diketahui prediksi data siswa yang diterima dengan true diterima mencapai 9 siswa dan true tidak sebanyak 1 data, dengan hasil nilai presisi sebesar 90%. Sedangkan untuk prediksi data siswa yang tidak diterima untuk true diterima mencapai 0 siswa dan true tidak sebanyak 5 siswa, dengan hasil nilai presisi 100%. Untuk class recall data siswa yang sangat rendah terdapat pada true diterima mencapai 100%, sedangkan untuk class recall true tidak mencapai 83.33%.

4. Kesimpulan

Penerapan algoritma C4.5 dalam menentukan peserta beasiswa bagi siswa berprestasi terbukti dapat

menghasilkan nilai akurasi yang tinggi dalam proses penentuan pemberian beasiswa Kartu Medan Pintar (KBP) dengan data awal sebanyak 50 data yang menghasilkan nilai akurasi sebesar 93,33 % dari 30 data training dan 20 data testing. Pengolahan data ini mempercepat dan mempermudah pihak sekolah dalam menentukan penerimaan beasiswa yang lebih akurat dan tepat sasaran.

Saran untuk penelitian berikutnya sebaiknya menggunakan jumlah data dan atribut yang lebih banyak agar hasil nilai akurasi dapat lebih baik dari penelitian sebelumnya. Penelitian ini bisa dikembangkan dengan menggunakan algoritma lain kemudian membandingkan hasilnya dengan hasil penelitian menggunakan algoritma C4.5 yang digunakan dalam penelitian ini.

Referensi

- [1] Herlinawati, E. Heriyati, Sudiyono, and A. B. Susanto, Kajian Program Indonesia Pintar (PIP): Strategi Penjangkauan Anak Tidak Sekolah (ATS) Untuk Mengikuti Pendidikan Melalui Program Indonesia Pintar (PIP). 2018.
- [2] Hapsari, D. T., Harini, & Nugroho, J. A. Pengaruh Beasiswa PPA dan Kebiasaan Belajar Terhadap Prestasi Belajar Mahasiswa FKIP UNS Penerima Beasiswa PPA Periode Januari – Juni 2017. BISE: Jurnal Pendidikan Bisnis Dan Ekonomi, Vol. 4, No.1 , 2018.
- [3] Yunus, M., Ramadhan, H., Aji, D. R., & Yulianto. Penerapan Metode Data Mining C4.5 Untuk Pemilihan Penerima Kartu Indonesia Pintar (KIP). Paradigma - Jurnal Komputer Dan Informatika, Vol. 23, No. 2, September 2021.
- [4] Wahyu Susantol dan Astriana Mulyani. Analisa Algoritma C4.5 Terhadap Penentuan Rekomendasi Penerima Beasiswa. OKTAL : Jurnal Ilmu Komputer dan Science, Vol 1, No. 10, Oktober 2022
- [5] Angga Pebdika, Ruli Herdiana dan Dodi Solihudin. Klasifikasi Menggunakan Metode Naive Bayes Untuk Menentukan Calon Penerima Pip. JATI (Jurnal Mahasiswa Teknik Informatika) Vol. 7 No. 1, Februari 2023
- [6] Sitanggang, E. D. (2023). Analisa Sistem Pakar Penyakit Menular Pada Anak-Anak Dengan Metode Forward Chaining. LOFIAN: Jurnal Teknologi Informasi Dan Komunikasi, 2(2), 20–25. <https://doi.org/10.58918/lofian.v2i2.207>
- [7] Weni Ratna Sari Oktapia Ningse , S Sumarno , Zulaini Masruro Nasution. Klasifikasi Algoritma C4.5 untuk Penentuan Penerima Program Indonesia Pintar pada MIS Al-Khoirot. JOMLAI: Journal of Machine Learning and Artificial Intelligence, Vol. 1, No. 1, Maret 2022.
- [8] Azahari, N. Rekomendasi Penerimaan Beasiswa Yayasan Untuk Siswa Baru SMK TI Airlangga dengan Algoritma C4 .Vol 5. No. 5, April 2021.
- [9] Vulandari, R. T. Data Mining Teori Dan Aplikasi Rapidminer. Yogyakarta, 2017.
- [10] Muhammad Arhani dan Muhammad Nasir. Algoritma dan Implementasi Data Mining. Penerbit Andi 2020.