

IMPLEMENTASI ALGORITMA C4.5 UNTUK PENENTUAN PENERIMAN BEASISWA (Studi Kasus: SMA N 2 Cepu)

Dian Febriani ^{a*}, Retno Wahyusari^b

^{a,b} Sekolah Tinggi Teknologi Ronggolawe Cepu

*Febrianid335@gmail.com, retnowahyusari@gmail.com

Diterima : 10-11-2021

Direview : 29-21-2021

Diterbitkan : 31-21-2021

Intisari

Salah satu Hak Asasi Manusia adalah Hak untuk mendapatkan pendidikan. Dalam menyelenggarakan pendidikan yang bermutu perlu biaya yang cukup besar. Beasiswa merupakan cara untuk mengatasi masalah biaya bagi mereka yang kurang mampu. Beasiswa adalah bantuan dari pemerintah maupun swasta berupa sejumlah uang yang diberikan kepada siswa yang sedang ataupun yang akan mengikuti pendidikan sekolah. Pemberian beasiswa ini sangat penting, guna memberikan ketenangan siswa dalam mengikuti belajar-mengajar. Karena banyaknya siswa yang mengajukan beasiswa dan indikator yang terlalu banyak mengakibatkan penerimaan beasiswa yang kurang tepat sasaran. Dalam mengatasi hal tersebut, dibuatlah suatu sistem pendukung keputusan penerimaan beasiswa menggunakan algoritma C4.5 yang nantinya akan mempermudah tim seleksi SMA 2 Cepu untuk menyeleksi siswa yang berhak mendapatkan beasiswa karena di SMA 2 Cepu proses seleksi masih menggunakan cara yang manual dan membutuhkan waktu yang lama. Algoritma C4.5 merupakan perkembangan dari Algoritma ID3 yang berbasis *Supervised Learning*. Dengan menggunakan metode *Decision Tree* Algoritma C4.5 pemberian beasiswa dapat tepat sasaran. Penelitian ini menggunakan data kelas 11 dan kelas 12 sebanyak 501 yang menjadi data *training* untuk menentukan pohon keputusan di RapidMiner, sedangkan kelas 10 sebanyak 280 menjadi data testing untuk di proses Matlab. Terdapat 99 siswa kelas 11 dan 12 yang Layak mendapatkan beasiswa, sedangkan 402 siswa kelas 11 dan 12 yang Tidak Layak mendapatkan beasiswa. Hasil penelitian aplikasi sistem pendukung keputusan dengan metode decision tree untuk rekomendasi siswa layak atau tidak layak menerima beasiswa menggunakan Algoritma C4.5 dapat menghasilkan tingkat akurasi yang tinggi yaitu sebesar 96,80%.

Kata kunci: Beasiswa, *Decision Tree*, Algoritma C4.5, Sistem Pendukung keputusan

Abstract

One of human rights is the right to education. In organizing quality education requires substantial costs. Scholarships are a way to overcome cost problems for those who are less well off. Scholarships are assistance from the government or the private sector in the form of a sum of money given to students who are currently or who will attend school education. The granting of this scholarships is very important, in order to give students peace of mind in attending teaching and learning. Because many students apply for scholarships and too many indicators result in receiving scholarships that are not well targeted. To overcome this problem, a scholarships acceptance Decision Support System was created using the C4.5 Algorithm which will make it easier for the 2 Cepu High School selection team to select students who are entitled to scholarships because in High School 2 Cepu the selection proses still uses manual methods and requires a long time. C4.5 Algorithm is a development of the ID3 Algorithm based on Supervised Learning. By using the Decision Tree method C4.5 Algorithm it is hoped that the awarding of scholarships can be right on target. In this study the researchers used class 11 and class 12 as training data to determine the decision tree in RapidMiner, while class 10 data became the testing data for the matlab proses. There are 99 students in grades 11 and 12 who are eligible for scholarships, while 402 grade 11 and 12 students who do not deserve a scholarship. The results of the decision tree support system research with the decision tree method for the recommendation of students to be eligibel or not to receive sholarships using the C4.5 arlgorithm can prosuce a high level of accuracy that is equal to 96,80%.

Keywords: Scholarships, *Decision Tree*, Algorithm C4.5, Decision Support

1. Pendahuluan

Hak Asasi Manusia yang ditulis dan disepakati salah satunya adalah hak atas mendapatkan pendidikan. Hal ini juga diperkuat oleh UUD 1945 pasal 31 ayat 1-2 “Bahwa setiap warga negara berhak mendapatkan pendidikan”. Dalam menyelenggarakan pendidikan yang bermutu diperlukan biaya yang cukup besar, namun tidak dapat dipungkiri bahwa masih terdapat warga Indonesia yang penghasilannya rendah untuk membiayai pendidikan. Maka dari itu pemerintah memberikan bantuan dana berupa beasiswa.

Penelitian yang dilakukan Deny Triwibowo dkk (Triwibowo, Kurniadi and Hartinah, 2019) beasiswa merupakan pemberian berupa bantuan keuangan yang diberikan kepada perorangan, mahasiswa atau pelajar yang digunakan demi keberlangsungan pendidikan yang ditempuh. Diharapkan adanya beasiswa ini siswa dapat

menyelesaikan pendidikan tanpa ada gangguan terutama biaya.

Sekolah Menengah Atas 2 Cepu merupakan sekolah yang menyelenggarakan program beasiswa kurang mampu bagi siswanya. Beasiswa tersebut berupa bantuan keringanan pembayaran uang SPP (Sumbangan Pembinaan Pendidikan). Karena banyak siswa Sekolah Menengah Atas 2 Cepu yang mengajukan beasiswa dan indikator yang terlalu banyak mengakibatkan penerimaan beasiswa kurang tepat sasaran dan juga proses seleksi yang dilakukan masih secara manual sehingga waktu yang dibutuhkan cukup lama dan kurang efisien. Jumlah siswa di SMA Negeri 2 Cepu sebanyak 709 siswa dari kelas 10, 11, 12 pada kelas 10 yang berhak mendapatkan beasiswa hanya 69 siswa, sedangkan untuk kelas 11 dan 12 yang berhak untuk mendapatkan beasiswa sebanyak 99 siswa.

Hal ini menuntut penelitian yang mampu mengatasi masalah tersebut. Banyak penelitian yang telah dilakukan untuk mengatasi masalah tersebut salah satunya adalah data mining klasifikasi yang digunakan dalam mengolah data yang sangat besar untuk memudahkan proses recording.

Penelitian yang dilakukan oleh Heni Sulistiani dan Yohana Tri Utami (Sulistiani and Utami, 2018) menyatakan bahwa klasifikasi merupakan salah satu teknik data mining untuk menganalisis data penting dan dapat secara otomatis memprediksi kelas dari data lain yang belum diklasifikasikan. Beberapa contoh klasifikasi seperti Decision Tree, Bayesian Network, Adaptive Bayesian Network, dan Naive Bayes. Dalam penelitian ini penulis menerapkan salah satu teknik klasifikasi Decision tree (pohon keputusan).

Decision tree (Pohon Keputusan) adalah pohon klasifikasi yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari permasalahan pengambilan keputusan secara kompleks dapat diubah menjadi sederhana. Beberapa algoritma yang digunakan untuk membangun pohon keputusan yang berbasis algoritma induksi pohon keputusan seperti ID3, C4.5 dan CART. Muhammad Faisal, dan dkk (Faisal, Nasution and Ta, 2017).

Penelitian yang dilakukan oleh Panji Bimo Nugroho Setio, dkk (Setio, Saputro and Bowo Winarno, 2020) Algoritma C4.5 adalah salah satu metode untuk membuat decision tree berdasarkan training data yang telah disediakan. Algoritma C4.5 merupakan pengembangan dari ID3 yang berbasis supervised learning. Beberapa pengembangan yang dilakukan C4.5 antara lain bisa mengatasi missing value, bisa mengatasi continue data, dan pruning. Rizky Haqmanullah Pambudi, dkk (Rizky Haqmanullah Pambudi, Setiawan and Indriati, 2018) C4.5 karena dapat melakukan prediksi dengan memberikan tingkat nilai akurasi yang ideal untuk memprediksikan kemampuan siswa. Priati (Assiroj, 2018) dengan penelitian membandingkan 3 algoritma yaitu algoritma C4.5, Naive Bayes dan CART dengan pengukuran menggunakan metode pengujian confusion matrix dan kurva ROC menyatakan bahwa algoritma C4.5 dan algoritma CART memiliki akurasi yang sama tinggi serta paling rendah adalah akurasi Naive Bayes.

Algoritma C4.5 adalah salah satu metode untuk membuat *decision tree* berdasarkan training data yang telah disediakan. Algoritma C4.5 merupakan pengembangan dari ID3 yang berbasis *supervised learning*. Beberapa pengembangan yang dilakukan C4.5 antara lain bisa mengatasi *missing value*, bisa mengatasi *continue data*, dan *pruning*

Atribut yang digunakan dalam menentukan penerimaan beasiswa Nama, Jumlah Pemasukan Orang Tua, Jumlah Tanggungan Orang Tua, Jumlah Pengeluaran Orang Tua, Nilai, Prestasi, Transportasi, Jarak, dan Rekomendasi. Pada penelitian ini menggunakan data kuantitatif dari kelas 11 dan kelas 12 sebanyak 501 untuk data training pada RapidMiner, sedangkan untuk kelas 10 sebanyak 280 untuk data testing sebagai proses dari Matlab.

2. Kerangka Teori

3.1. Beasiswa

Menurut Desi Ayu Ningsih, dkk (Ningsih, Hartama and Dewi, 2020) beasiswa adalah bentuk penghargaan yang diberikan kepada individu agar dapat melanjutkan pendidikan ke jenjang yang lebih tinggi. Adapun tujuan pemberian beasiswa yaitu diantaranya:

- Untuk membantu para pelajar supaya dapat mencari ilmu yang sesuai dengan bidang yang hendak dikuasai, yang paling utama bagi yang memiliki masalah dalam pembiayaan.
- Membuat generasi baru yang lebih cerdas dan pintar. Karena dengan adanya bantuan beasiswa ini maka seseorang terutama kaum muda dapat memiliki kesempatan untuk memperoleh pendidikan pada jenjang yang lebih tinggi.
- Meningkatkan kesejahteraan.

3.2. Sistem Pendukung Keputusan

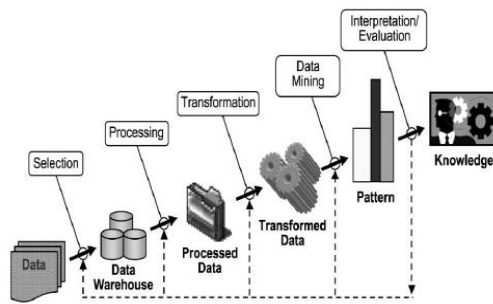
Konsep Sistem Pendukung Keputusan (SPK) atau *Decision Support System* (DSS) pertama kali diungkapkan pada awal tahun 1970-an oleh Michael S. Scott Morton dengan istilah *Management Decision System*. Sistem tersebut adalah suatu sistem yang berbasis komputer yang ditujukan untuk pengambilan keputusan dengan memanfaatkan data dan model tertentu untuk memecahkan berbagai persoalan yang tidak terstruktur.

3.3. Data Mining

Menurut Larose (Larose, 2005) data mining adalah proses menemukan korelasi, pola, dan tren baru yang bermakna dengan menyaring sejumlah besar data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola secara teknik statistik dan matematika. Terdapat beberapa istilah lain yang memiliki makna sama dengan data mining yaitu Knowledge Discovery in Database (KDD). Pengelompokan data mining dibagi menjadi beberapa kelompok yaitu:

1. Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya.
2. Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas.
3. Prediksi memiliki kemiripan dengan klasifikasi, akan tetapi data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang.

Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap. Tahap-tahap tersebut bersifat interaktif dimana pemakai terlibat langsung atau dengan perantaraan knowledge base.



Gambar 1 Tahapan Data Mining

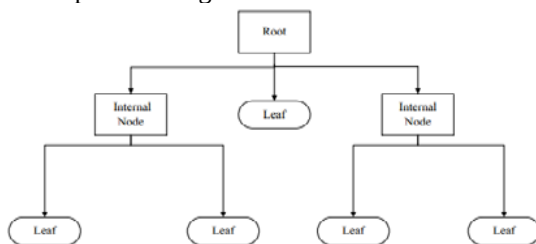
3.4. Klasifikasi

Menurut Dewi Sartika dan Dana Indra Sensuse (Sartika and Sensuse, 2017) Klasifikasi adalah bentuk analisis data untuk mengekstrak model yang akan digunakan untuk memprediksi label kelas. Kelas dalam klasifikasi merupakan atribut dalam satu set data yang paling unik yang merupakan variabel bebas dalam statistik. Klasifikasi data terdiri dari dua proses yaitu yang pertama tahap pembelajaran merupakan tahapan dalam pembentukan model klasifikasi, sedangkan yang kedua tahap pengklasifikasian merupakan tahapan pengguna model klasifikasi untuk memprediksi label kelas dari suatu data.

Tujuan dari klasifikasi adalah untuk menemukan model dari *training set* yang membedakan atribut kedalam kategori atau kelas yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasikan atribut yang kelasnya belum diketahui.

3.5. Decision Tree

Menurut Panji Bimo Nugroho Setio, dkk (Setio, Saputro and Bowo Winarno, 2020) *decision tree* merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Data yang berupa data fakta diubah menjadi sebuah pohon keputusan yang lebih berisi aturan dan tentunya dapat lebih mudah dipahami dengan bahasa alami



Gambar 2 Decision Tree

3.6. Algoritma C4.5

Algoritma C4.5 adalah salah satu metode untuk membuat Decision Tree berdasarkan Training data yang telah disediakan. Algoritma C4.5 yang dibuat oleh Ross Quilan yang merupakan pengembangan dari Algoritma ID3 yang juga dibuat oleh Quilan. Output dari algoritma C4.5 adalah sebuah pohon keputusan atau yang dikenal dengan *decision tree*. Ada beberapa tahapan dalam membuat sebuah *decision tree* dengan Algoritma C4.5 yaitu:

- 1) Mempersiapkan *data training*. Data training biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.

- 2) Menghitung akar dari pohon. Akar akan diambil dari atribut yang akan dipilih, dengan cara menghitung nilai *gain* dari masing-masing atribut, nilai *gain* yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dahulu nilai *entropy*. Untuk menghitung nilai *entropy* digunakan rumus:

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

Keterangan

S = himpunan kasus

n = jumlah partisi S

P_i = proporsi S_i terhadap

- 3) Menghitung nilai *Gain* menggunakan rumus:

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad (2)$$

Keterangan

S = himpunan kasus

A = atribut

n = jumlah partisi A

|S_i| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

- 4) Ulangi langkah mencari nilai *gain* hingga semua *record* terpartisi.
- 5) Proses partisi *decision tree* akan berhenti saat:
 - a) Semua *record* dalam simpul N mendapatkan kelas yang sama.
 - b) Tidak ada atribut didalam *record* yang dipartisi lagi.
 - c) Tidak ada *record* di dalam cabang yang kosong.

3.7. Confusion Matrix

Pengukuran terhadap kinerja suatu sistem klasifikasi merupakan hal yang penting. Kinerja sistem menggambarkan seberapa baik sistem dalam mengklasifikasi data. *Confusion Matrix* merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi.

Pada pengukuran kinerja menggunakan *Confusion Matrix*, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)*. Nilai *True Positive (TP)* merupakan data positif yang terdeteksi benar, Nilai *True Negative (TN)* merupakan jumlah data negatif dengan benar. Sedangkan *False Positive (FP)* merupakan data negatif namun terdeteksi sebagai data positif, dan nilai *False Negative (FN)* merupakan kebalikan dari *True Positive*, sehingga data positif namun terdeteksi sebagai data negatif.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (3)$$

$$presisi = \frac{TP}{FP+TP} \times 100\% \quad (4)$$

$$Recall = \frac{TP}{FN + TP} \times 100\% \quad (5)$$

3.8. RapidMiner

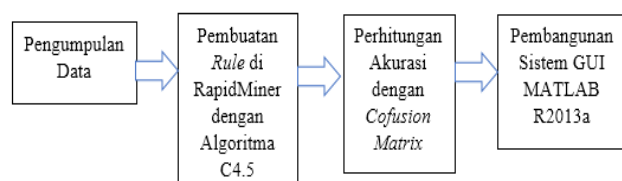
Menurut Asep Muhidin dan Liswanti (Muhidin et al., 2019) RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi.

3.9. Matlab

Menurut Anita Sindar Sinaga, dkk (Sinaga, Sitio and Sijabat, 2020) Matlab sebuah bahasa dengan (high-performance) kinerja tinggi untuk komputasi masalah teknik. Matlab mengintegrasikan komputasi, visualisasi, dan pemrograman dalam suatu model yang sangat mudah untuk pakai dimana masalah-masalah dan penyelesaiannya diekspresikan dalam notasi matematika yang familiar. Nama MATLAB merupakan singkatan dari *Matrix Laboratory*. Dalam lingkungan perguruan tinggi teknik matlab merupakan perangkat standart untuk memperkenalkan dan mengembangkan penyajian matematika, rekayasa dan keilmuan. Fitur-fitur MATLAB sudah banyak dikembangkan, dan lebih kita kenal dengan nama toolbox. Sangat penting bagi seorang pengguna MATLAB, toolbox mana yang mendukung untuk learn dan apply technology yang sedang dipelajarinya.

3. Metodologi

Tahapan dalam penelitian ini diawali dengan pengambilan data dari lapangan yaitu Sekolah Menengah Atas 2 Cepu, setelah data yang dibutuhkan terkumpul kemudian penulis menggunakan perangkat lunak RapidMiner dan MATLAB R2013a untuk melakukan optimasi dengan menggunakan Algoritma C4.5 untuk mendukung sistem penentuan keputusan siswa penerima beasiswa yang ada di Sekolah Menengah Atas 2 Cepu.



Gambar 3 Prosedur Penelitian

Hasil pengumpulan data kemudian dilakukan *Integrasi Data*. Data yang digunakan dalam penelitian ini adalah data kualitatif yang diperoleh dari Sekolah Menengah Atas 2 Cepu. Dari data yang dikumpulkan yaitu data siswa ajaran 2018/2019 kelas 11 dan 12 sebagai data *Training* dan kelas 10 sebagai data *Testing*. Atribut data yang didapat sebelum dilakukan transformasi data terlihat pada Tabel 1.

Tabel 1 Keterangan Data Siswa

Atribut	Keterangan Atribut
NIS	Atribut NIS yang dimiliki SMA Negeri 2 Cepu
Nama	Atribut yang difungsikan untuk ID
Kelas	Atribut kelas yang di tempati
Agama	Atribut ini yang menginformasikan agama dari siswa SMA Negeri 2 Cepu
Jenis Kelamin	Atribut ini yang menginformasikan Jenis Kelamin dari siswa SMA Negeri 2 Cepu
Jumlah Penghasilan Orang Tua (JPO)	Atribut ini yang menginformasikan Jumlah

Jumlah Tanggungan Orang Tua (JTO)	Penghasilan Orang Tua dari siswa SMA Negeri 2 Cepu Atribut ini yang menginformasikan Jumlah Tanggungan Orang Tua dari siswa SMA Negeri 2 Cepu
Jumlah Pengeluaran Orang Tua (PO)	Atribut ini yang menginformasikan Jumlah Pengeluaran Orang Tua dari siswa SMA Negeri 2 Cepu
Nilai	Atribut ini yang menginformasikan Nilai dari siswa SMA Negeri 2 Cepu
Prestasi	Atribut ini yang menginformasikan Prestasi dari siswa SMA Negeri 2 Cepu
Transportasi	Atribut ini yang menginformasikan Transportasi dari siswa SMA Negeri 2 Cepu
Jarak	Atribut ini yang menginformasikan Jarak dari siswa SMA Negeri 2 Cepu
Rekomendasi	Atribut ini yang menginformasikan sebagai Label

Hasil pengumpulan data dilanjutkan dengan seleksi data. Atribut awal terdapat 12 atribut, akan tetapi hanya 9 atribut yang digunakan yaitu Nama, JPO, JTO, PO, Nilai, Prestasi, Transportasi, Jarak dan Rekomendasi. Tabel 3.2 merupakan pemilihan atribut, 7 atribut yang akan menjadi nilai model untuk RapiMiner dan 1 atribut untuk ID dan Rekomendasi sebagai label.

Tabel 2 Pemilihan Atribut

Atribut	Detail Penggunaan	
NIS	X	Tidak
Nama	√	ID
Kelas	X	Tidak
Agama	X	Tidak
Jenis Kelamin	X	Tidak
JPO	√	Nilai Model
JTO	√	Nilai Model
PO	√	Nilai Model
Nilai	√	Nilai Model
Prestasi	√	Nilai Model
Transportasi	√	Nilai Model
Jarak	√	Nilai Model
Rekomendasi	√	Label Target

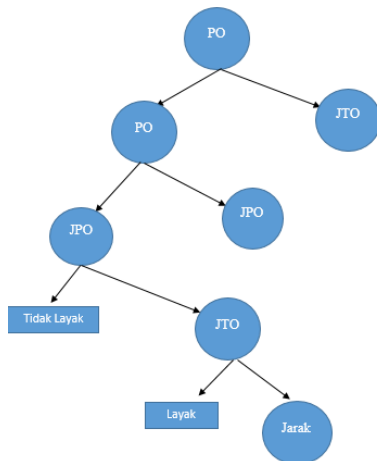
Setelah data dilakukan seleksi data kemudian dilakukan perhitungan manual. Hasil perhitungan manual kemudian dibandingkan dengan hasil pada Tool RapidMiner. Jika hasil sama maka *rule* yang didapat digunakan sebagai acuan dalam membuat program.

4. Hasil dan Pembahasan

4.1. Hasil

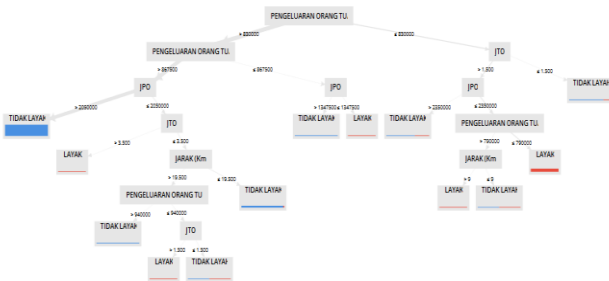
Penelitian ini bertujuan untuk menentukan Layak atau Tidak layaknya siswa SMA N 2 Cepu mendapatkan beasiswa menggunakan data training berjumlah 501 data dan data

testing berjumlah 280. Hasil perhitungan secara manual menghasilkan pohon keputusan seperti pada Gambar 3.



Gambar 3. Pohon Keputusan Hasil Perhitungan Manual.

Hasil pohon keputusan hasil perhitungan manual, dibandingkan dengan hasil pohon keputusan yang terbentuk dari pemanfaatan tool RapidMiner. Gambar 4 merupakan bentuk pohon keputusan hasil pengolahan data menggunakan RapidMiner.



Gambar 4. Pohon Keputusan Hasil RapidMiner.

RapidMiner selain dapat menampilkan struktur pohon keputusan, juga menghasilkan rule. Rule/aturan yang terbentuk dijadikan acuan dalam membuat aplikasi. Gambar 5 merupakan rule/aturan yang terbentuk dari pohon keputusan.

Tree

```

PENGELUARAN ORANG TUA > 830000
|
|  PENGELUARAN ORANG TUA > 867500
|  |
|  |  JPO > 2050000: TIDAK LAYAK (TIDAK LAYAK=343, LAYAK=0)
|  |  |
|  |  |  JPO ≤ 2050000
|  |  |  |
|  |  |  |  JTO > 3.500: LAYAK (TIDAK LAYAK=0, LAYAK=2)
|  |  |  |  |
|  |  |  |  |  JTO ≤ 3.500
|  |  |  |  |  |
|  |  |  |  |  |  JARAK (km) > 19.500
|  |  |  |  |  |  |
|  |  |  |  |  |  |  PENGELUARAN ORANG TUA > 940000: TIDAK LAYAK (TIDAK LAYAK=6, LAYAK=0)
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  PENGELUARAN ORANG TUA ≤ 940000
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  JTO > 1.500: LAYAK (TIDAK LAYAK=0, LAYAK=3)
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  JTO ≤ 1.500: TIDAK LAYAK (TIDAK LAYAK=1, LAYAK=1)
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  JARAK (km) ≤ 19.500: TIDAK LAYAK (TIDAK LAYAK=40, LAYAK=1)
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  PENGELUARAN ORANG TUA ≤ 867500
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  JPO > 1347500: TIDAK LAYAK (TIDAK LAYAK=5, LAYAK=0)
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  JPO ≤ 1347500: LAYAK (TIDAK LAYAK=0, LAYAK=3)
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  PENGELUARAN ORANG TUA ≤ 830000
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  JTO > 1.500
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  JPO > 2350000: TIDAK LAYAK (TIDAK LAYAK=2, LAYAK=1)
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  JPO ≤ 2350000
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  PENGELUARAN ORANG TUA > 790000
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  JARAK (km) > 9: LAYAK (TIDAK LAYAK=0, LAYAK=3)
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  JARAK (km) ≤ 9: TIDAK LAYAK (TIDAK LAYAK=1, LAYAK=1)
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  PENGELUARAN ORANG TUA ≤ 790000: LAYAK (TIDAK LAYAK=0, LAYAK=83)
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  JTO ≤ 1.500: TIDAK LAYAK (TIDAK LAYAK=4, LAYAK=1)

```

Gambar 5 Rule/Aturan Hasil Pohon Keputusan

Setelah mendapatkan rule dari proses pada aplikasi RapidMiner dilanjutkan dengan pembuatan GUI yang

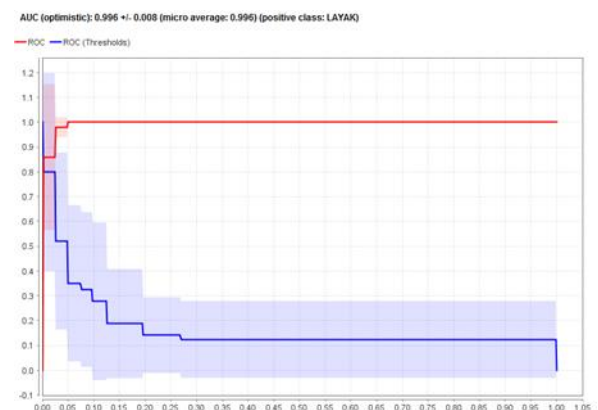
bertujuan untuk mempermudah proses input data dari siswa penerima beasiswa. Rule yang didapatkan dari RapidMiner diubah dalam bentuk sintak pada bahasa pemrograman. Berdasarkan dari sintak rule RapidMiner dapat membangun interface dari matlab yang akan mempermudah untuk menyeleksi siswa penerima beasiswa.

Gambar 6 Interface siswa yang Layak menerima Beasiswa

4.2. Pembahasan

Hasil pengukuran tingkat keakuratan klasifikasi dengan Confusion Matrix didapatkan hasil akurasi yang tinggi yaitu sebesar 96,80%. Dimana TP adalah (True Positive) jumlah data yang benar tetapi terklasifikasikan dengan benar oleh sistem. TN adalah (True Negative) jumlah data yang salah tetapi terklasifikasikan dengan benar oleh sistem. FP (False Positive) jumlah data yang benar tetapi terklasifikasikan dengan salah oleh sistem. FN (False Negative) jumlah data yang salah tetapi terklasifikasikan dengan salah oleh sistem.

Hasil yang di dapat dari pengolahan ROC untuk Algoritma C4.5 dengan menggunakan data Training 501 sebesar 0.996 dapat dilihat pada Gambar 7 dengan tingkat diagnosa excellent classification.



Gambar 7 Kurva ROC

5. Simpulan

Hasil penelitian menunjukkan tingkat akurasi yang tinggi yaitu sebesar 96,80% dengan nilai *class Precision* sebesar 97,54% dan 93,68% dan nilai *class Recall* sebesar 98,51% dan 89,90%. Sistem pendukung keputusan yang dibangun dapat membantu efektivitas dan efisiensi proses pengambilan keputusan seleksi beasiswa oleh pihak SMA Negeri 2 Cepu.

Saran untuk mengembangkan penelitian selanjutnya dapat menambahkan beberapa atribut dalam menentukan pemberian beasiswa serta menerapkan metode Algoritma CART, guna membandingkan tingkat akurasi.

Daftar Pustaka

- Assiroj, P. (2018) 'Kajian Perbandingan Teknik Klasifikasi Algoritma C4.5, Naïve Bayes Dan Cart Untuk Prediksi Kelulusan Mahasiswa (Studi Kasus : Stmik Rosma Karawang) (Studi Kasus : Stmik Rosma Karawang)', 15(March), pp. 1–17. doi: 10.5281/zenodo.1184054.
- Faisal, M., Nasution, Y. N. and Ta, F. D. (2017) 'Perbandingan Kinerja Metode Klasifikasi Chi-square Automatic Interaction Detection (CHAID) dengan Metode Klasifikasi Algoritma C4.5 pada Studi Kasus : Penderita Diabetes Melitus Tipe 2 Di Samarinda Tahun 2015', *jurnal EKSPONENSIAL*, 8(2015), pp. 119–124.
- Larose, D. T. (2005) *Discovering Knowledge In Data : an introduction to data mining. In Automotive Industries AI*. Canada: A John Wiley & Sons.Inc.
- Muhidin, A. *et al.* (2019) 'Klasifikasi Penduduk Tidak Mampu Desa Mandiraha Wetan Menggunakan Algoritma C4.5', *Jurnal Pelitabangsa*, 9(3), pp. 13–18.
- Ningsih, D. A., Hartama, D. and Dewi, R. (2020) 'Penerapan Metode VIKOR Pada Pengambilan Keputusan Seleksi Calon Penerima Beasiswa di SMK TPI Al-Hassanah Pematang Bandar', *Brahmana : Jurnal Penerapan Kecerdasan Buatan*, 2(1), pp. 25–32. doi: 10.30645/brahmana.v2i1.45.
- Rizky Haqmanullah Pambudi, Setiawan, B. D. and Indriati (2018) 'Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal', *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(7), pp. 2637–2643. Available at: <http://j-ptiik.ub.ac.id>.
- Sartika, D. and Sensuse, D. I. (2017) 'Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian', *jastisi*, 1(2), pp. 151–161.
- Setio, P. B. N., Saputro, D. R. S. and Bowo Winarno (2020) 'Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5', *PRISMA, Prosiding Seminar Nasional Matematika*, 3, pp. 64–71.
- Sinaga, A. S., Sitio, A. S. and Sijabat, P. (2020) 'Pengenalan Dasar Pengkodean Secara Daring pada SMK Pemda Lubuk Pakam', *Abdimas Universal*, 2(2), pp. 95–99. doi: 10.36277/abdimasuniversal.v2i2.74.
- Sulistiani, H. and Utami, Y. T. (2018) 'Penerapan algoritma klasifikasi sebagai pendukung keputusan pemberian beasiswa mahasiswa', *SNTI*, 6, pp. 300–305.
- Triwibowo, D. N., Kurniadi, A. and Hartinah, S. (2019) 'Sistem Pendukung Keputusan Pemilihan Mahasiswa Penerima Beasiswa dengan K-NN dan ELECTRE', *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, 9(2), p. 89. doi: 10.35585/inspir.v9i2.2531.