

PENERAPAN DATA MINING UNTUK REKOMENDASIBEASISWA PADA SMA N 1 MLONGGO MENGUNAKAN ALGORITMA C4.5

Mahindra suryaning praja S.kom (mahendrapraja7@gmail.com)

Erna Zuni Astuti, M.kom (ernazunias@yahoo.com)

Jurusan Teknik Informatika

UDINUS

Abstrak : Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh. Pemberian beasiswa kepada mahasiswa di Perguruan Tinggi merupakan wujud dari partisipasi masyarakat, instansi, pemerintah, perusahaan-perusahaan swasta dalam ikut serta membangun bangsa khususnya dalam bidang pendidikan. Oleh karena itu beasiswa harus diberikan kepada penerima yang layak dan pantas mendapatkannya. Dalam tugas akhir kali ini penulis membangun sebuah sistem untuk melakukan perekomendasi penerima beasiswa sehingga beasiswa dapat diberikan kepada penerima yang layak dan pantas mendapatkannya, dimana metode yang akan digunakan adalah Klasifikasi dilakukan menggunakan data mining algoritma C4.5 Data yang digunakan yaitu data jurusan, kelas, jumlah nilai, penghasilan orangtua, dan jumlah saudara kandung. Proses data mining pada data training akan menghasilkan pohon keputusan atau rule. Metode evaluasi yang dilakukan dalam penelitian ini yaitu menggunakan confusion matrix dan nilai akurasi, untuk sekali pengujian tingkat akurasi yang dihasilkan yaitu 70%. hal ini membuktikan bahwa algoritma C4.5 cukup akurat dalam menentukan rekomendasi beasiswa pada SMA N 1 Mlonggo.

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

Abstract : Scholarship is a gift in the form of financial assistance granted to individuals who intended to be used for the continuation of education pursued. Awarding scholarships to students in universities is a form of participation of society, institutions, governments, companies private sector in participating in the development of the nation, especially in the field of educators. Therefore scholarship should be awarded to eligible recipients and deserve her. In the final task this time the author built a system to do perekomendasi awardees so that scholarships can be awarded to eligible recipients and deserve mendapatkannya, wherein the method to be used is performed using data mining classification algorithm C4.5 The data used are the data subject, class, the number of values, parental income, and number of siblings. The process of data mining on training data will generate decision trees or rule. Evaluation method in this research is using confusion matrix and values ?? of accuracy, to a level of testing the resulting accuracy of 70%. this proves that the algorithm C4.5 quite accurate in determining scholarship recommendations in SMA N 1 Mlonggo.

Keywords : Scholarships, Data Mining, Classification, C4.5.decision tree

1. PENDAHULUAN

Pendidikan merupakan suatu kebutuhan primer yang sejak dini hingga dewasa hendaknya dirasakan oleh seluruh masyarakat. Hal ini sesuai dengan amanat UUD Negara Kita, anjuran agama, dan menjadi penentu kemajuan suatu bangsa [1]. Pendidikan juga merupakan variabel vital untuk pembangunan suatu bangsa. Suatu bangsa bisa maju dengan cepat dibandingkan dengan negara lain karena penyebaran pengetahuan (*knowledge*) yang merata keseluruh lapisan masyarakatnya. Institusi yang paling bertanggung jawab untuk penyebaran pengetahuan adalah institusi pendidikan [2].

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh [3]. Demikian juga di SMAN 1 Mlonggo mendapat program pemberian beasiswa. Tetapi cara yang digunakan masih menggunakan cara manual sehingga ada beberapa kelemahan pada sistem yang digunakan saat ini, salah satunya adalah kurang tepatnya dalam penyaluran beasiswa. Hal ini terjadi karena cara yang digunakan oleh pihak yang diberi kepercayaan dalam pengambilan keputusan melihat kriteria-kriteria yang ditentukan secara terpisah dan juga dipengaruhi oleh jumlah data calon penerima beasiswa yang masuk.

Berdasarkan dari uraian latar belakang tersebut, maka dalam pembuatan tugas akhir ini penulis membuat judul “PENERAPAN DATA MINING UNTUK

REKOMENDASI BEASISWA PADA SMA N 1 MLONGGO MENGGUNAKAN ALGORITMA C4.5 “

2. LANDASAN TEORI

2.1 Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran computer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari. *Knowledge Discovery in Databases (KDD)* adalah penerapan metode *saintifik* pada data mining. Dalam konteks ini data mining merupakan satu langkah dari proses KDD [2].

Beberapa teknik dan sifat data mining adalah sebagai berikut:

- *Classification [Predictive]*
- *Clustering [Descriptive]*
- *Association Rule Discovery [Descriptive]*
- *Regression [Predictive]*
- *Deviation Detection [Predictive]*

2.2.1 Pengelompokan Data Mining

Ada beberapa pengelompokan data mining, data mining akan dikelompokkan berdasarkan tugas-tugas yang dikerjakan, antarlain [13] :

1. Deskripsi

mencari cara untuk menentukan sebuah pola dan kecenderungan yang ada dalam data. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan

penjelasan untuk suatu pola atau kecenderungan

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik daripada kearah kategori. Model dibangun menggunakan record lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi. Kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

4. Klasifikasi

Dalam klasifikasi variabel, tujuan bersifat kategorik. Misalnya, kita akan mengklasifikasikan suhu dalam tiga kelas, yaitu suhu panas, suhu sejuk, suhu dingin.

5. Pengklusteran

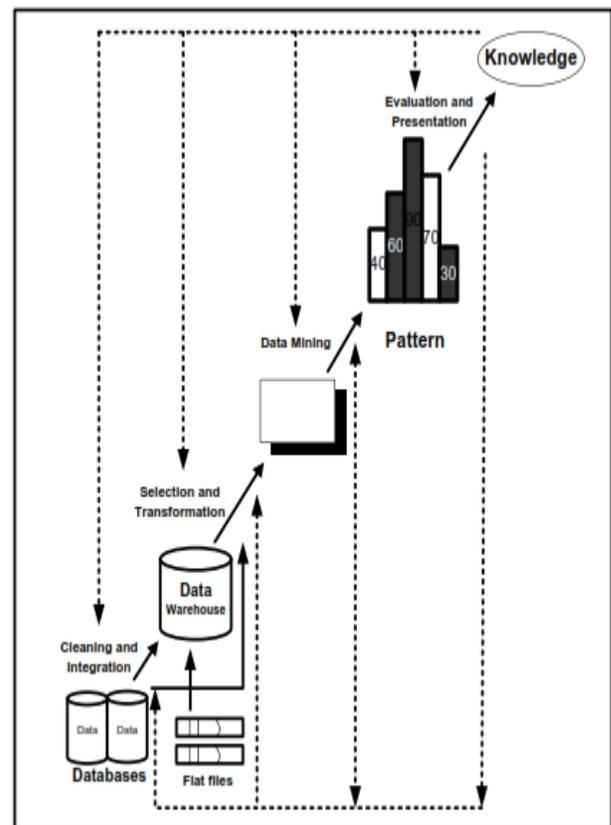
Pengklusteran adalah pengelompokkan sebuah record, pengamatan dan membentuk kelas kedalam sebuah objek yang mempunyai kemiripan. algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan record dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan record dalam kelompok lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu.

2.2.2 Tahap-Tahap Data Mining

Karena data mining adalah sebuah untaian proses, maka pecah menjadi beberapa tahap. Tahapan tersebut akan bersifat interaktif, pengguna akan terlibat langsung atau dengan perantara *KDD*[17].



Gambar 2.1 Tahapan *Data Mining*.

Tahapan *data mining* dibagi menjadi enam bagian yaitu :

1. Pembersihan data (*data cleaning*)

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus *KDD*. Proses cleaning mencakup antara lain membuang

duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

2. Integrasi data (*data integration*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk *data mining* tidak hanya berasal dari satu *database* tetapi juga berasal dari beberapa *database* atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyebabkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi Data (*Data Selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus *market basket analysis*, tidak

perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data (*Data Transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan *clustering* hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses *mining*.

adalah sebuah proses yang paling utama pada saat metode diterapkan untuk mencari pengetahuan tersembunyi dan berharga dari data.

6. Evaluasi pola (*pattern evaluation*),

Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai.

7. Presentasi pengetahuan (*knowledge presentation*),

Merupakan penyajian dan visualisasi pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses *data mining* adalah bagaimana memformulasikan keputusan atau

aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami *data mining*. Karenanya presentasi hasil *data mining* dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses *data mining*. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil *data mining* [19].

2.3 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu [13].

Cara algoritma C4.5 untuk membangun pohon keputusan yaitu:

- a. Pilih atribut yang akan digunakan sebagai akar
- b. Buatlah sebuah cabang untuk setiap nilai
- c. Bagilah kasus dalam sebuah cabang
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Ada beberapa tahap dalam membuat sebuah pohon keputusan dengan algoritma C4.5 [14].

1. Menyiapkan data training. Data ini diambil dari data yang sudah pernah ada sebelumnya dan sudah dikelompokkan kedalam kelas tertentu.
2. Setelah itu tentukan akar dari pohon. Pilih akar dari atribut, cara adalah dengan menghitung nilai gain dari semua atribut, yang menjadi akar pertama adalah nilai gain yang paling. Sebelum menentukan nilai gain, terlebih dahulu hitung nilai entropy. Untuk menentukan nilai entropy gunakan rumus

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

Keterangan :

S = himpunan kasus

n = jumlah partisi S

p_i = proporsi S_i terhadap S

3. Setelah itu tentukan nilai *gain* menggunakan rumus :

$$gain(S, A) = Entropy(s) - \sum_i^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Keterangan :

S = Himpunan kasus

A = fitur

N = jumlah partisi atribut A

$|S_i|$ = proporsi S_i terhadap S

$|S|$ = jumlah kasus dalam S

4. Setelah itu ulangilah langkah ke-2 sampai semua record terpartisi secara sempurna.

5. Proses partisi pohon keputusan akan berhenti saat :
 - a) Semua *record* dalam simpul N mendapat kelas yang sama.
 - b) Tidak ada atribut di dalam record yang dipartisi lagi.
 - c) Tidak ada *record* di dalam cabang yang kosong.

2.4 Decision Tree

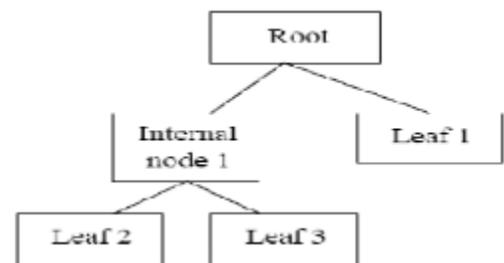
Pohon (*tree*) adalah sebuah struktur data yang terdiri dari simpul (*node*) dan rusuk (*edge*). Simpul pada sebuah pohon dibedakan menjadi tiga, yaitu simpul akar (*root node*), simpul percabangan/ internal (*branch/ internal node*) dan simpul daun (*leaf node*), [15]. Pohon keputusan merupakan representasi sederhana dari teknik klasifikasi untuk sejumlah kelas berhingga, dimana simpul internal maupun simpul akar ditandai dengan nama atribut, rusuk-rusuknya diberi label nilai atribut yang mungkin dan simpul daun ditandai dengan kelas-kelas yang berbeda [15].



Gambar 2.2 Konsep Pohon Keputusan.

Pohon keputusan adalah sebuah struktur flowchart yang setiap node nya merepresentasikan test dalam atribut (contoh, koin bila kita bolak balikan akan menghasilkan kepala, atau ekor), Setiap cabang (*branch*) mewakili hasil test dan setiap daun node (*leaf*) mewakili kelas label (hasil keputusan setelah menghitung semua atribut).

Bagian dari akar (*root*) hingga ke daun merepresentasikan dari rules (aturan) yang terbentuk. Pohon keputusan menurut saya bukanlah sebuah algoritma melainkan metode yang nantinya menghasilkan beberapa algoritma yang dapat digunakan dalam pengembangannya. Metode ini akan membantu kita untuk mengeksplorasi data, dan menemukan relasi tersembunyi antar sejumlah variabel input dan target. Pohon keputusan adalah himpunan aturan *IF...THEN*. Setiap *path* dalam *tree* dihubungkan dengan sebuah aturan, di mana premis terdiri atas sekumpulan *node-node* yang ditemui, dan kesimpulan dari aturam terdiri atas kelas yang terhubung dengan *leaf* dari *path* [16].



Gambar 2.3 Konsep Dasar Pohon Keputusan

Paling atas dari pohon keputusan disebut dengan titik akar (*root*), sedangkan setiap cabang dari pohon keputusan adalah pembagian berdasarkan hasil uji, dan titik akhir (*leaf*) merupakan pembagian kelas yang dihasilkan.

2.5 Confusion Matrix

Confusion matrix adalah suatu metode yang biasanya digunakan sebagai perhitungan akurasi pada konsep data mining. Informasi dalam confusion matrix diperlukan untuk menentukan kinerja model klasifikasi. Ringkasan informasi ini ke dalam sebuah nilai digunakan untuk

membandingkan kinerja dari model-model yang berbeda. Hal ini dapat dilakukan dengan menggunakan performace metric [17]

Tabel 2.1 Confusion Matrik

Classsification	Predicted class	
	Class = Yes	Class = No
Class = Yes	a (<i>true positive</i> – TP)	b (<i>false negative</i> – TN)
Class = No	c (<i>false positive</i> – FP)	d (<i>true negative</i> – TN)

3. METODOLOGI PENELITIAN

3.1 Metode Pengumpulan Data

Sesuai dengan tujuan penyusunan laporan Tugas Akhir dan untuk mendapatkan data yang valid dan akurat, maka dalam pengumpulan data penulis menggunakan suatu teknik yaitu:

3.2 Parameter yang diukur

Parameter yang diukur untuk pada data siswa sebagaiatribut untuk pengklasifikasian yaitu :

1. Data Jurusan

Data siswa yang merupakan data pengelompokansiswa yaitu jurusan IPA dan IPS

2. Data Kelas

Data siswa yang merupakan data pengelompokan siswa yaitu kelas A1, A2, S1,S2.

3. Data Jumlah Nilai Siswa

Data jumlah nilai siswa merupakan jumlah dari semua nilai mata pelajaran yang akan dijadikan acuan untuk menentukan apakah jumlah nilai

siswa tersebut termasuk di atas rata-rata atau dibawah rata-rata. Penilaian : jika per kelas sudah ditentukan jumlah rata-ratanya maka jumlah nilai siswa ini dijadikan acuan jika nilai siswa yang dibawah rata-rata maka siswa tidak lolos untuk rekomendasi beasiswa dan jika diatas rata-rata maka lolos rekomendasi beasiswa

4. Data Penghasilan Orang tua

Data penghasilan orang tua merupakan parameter yang sangat diperlukan untuk menentukan rekomendasi beasiswa karena menunjukkan berapa penghasilan yang didapat. Penilaian : Pada data penghasilan orangtua terdapat 3 kriteria yang sudah ditentukan oleh pihak Sma Negeri 1 Mlonggo yaitu golongan 1 dengan penghasilan = 1.000.000, golongan 2 dengan penghasilan 1.000.000-2.000.000, dan golongan 3 yaitu dengan penghasilan diatas 2.000.000.

5. Jumlah Tanggungan orang tua

Data jumlah tanggungan orang tua sangat menentukan karena digunakan untuk perbandingan dengan penghasilan orangtua. Penilaian : jika penghasilan orangtua dengan penghasilan yang rendah dan mempunyai jumlah tanggungan orang tua > 4 maka akan dipertimbangkan untuk mendapatkan rekomendai beasiswa.

3.3 Pengolahan Data (Data Preparation)

Pada penelitian ini atribut yang digunakan yaitu :

Tabel 2.2 Detail Atribut Data Penelitian

Atribut	Detail Penggunaan	
Nomor Induk Siswa	X	No
Nama	X	No
Jurusan	√	Nilai Model
Kelas	√	Nilai Model
Jumlah Nilai	√	Nilai Model
Jumlah kategori	√	Nilai Model
Pekerjaan Orangtua	X	No
Penghasilan Orangtua	√	Nilai Model
Jumlah Tanggungan orang tua	√	Nilai Model
Rekomendasi Beasiswa	√	Label Target

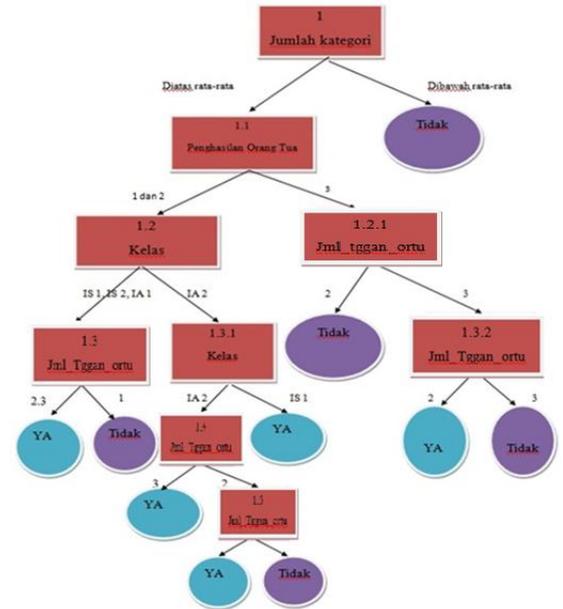
Keterangan :

(√) : Atribut yang dipakai dalam pengklasifikasian siswa

(×) : Atribut yang tidak digunakan dalam pengklasifikasian

4. HASIL DAN PEMBAHASAN

hasil pohon keputusan akan tampak pada gambar 4.4



Dari pohon keputusan tersebut didapat rules sebagai berikut :

Keputusan untuk Hasil = YA

- IF ipa jumlah_nilai ≥ 960 ,
penghasilan_ortu=1 and
jml_tggan_ortu =2 THEN Hasil = YA
- IF ips jumlah_nilai ≥ 944 ,
penghasilan_ortu=1 and
jml_tggan_ortu =2 THEN Hasil = YA
- IF jumlah_nilai = diatas_rata2,
penghasilan_ortu=1 and
jml_tggan_ortu =3 THEN Hasil = YA
- IF jumlah_nilai = diatas_rata2,
penghasilan_ortu=2 and
jml_tggan_ortu =3 THEN Hasil = YA

Keputusan untuk Hasil = TIDAK

- IF jumlah_nilai = dibawah_rata2
THEN Hasil = TIDAK
- IF penghasilan_ortu = 3 THEN Hasil
= TIDAK

3. IF jumlah_nilai = diatas_rata2, penghasilan_ortu=2 and jml_tggan_ortu = 1 THEN Hasil = TIDAK
4. IF jumlah_nilai = diatas_rata2, penghasilan_ortu=2 and jml_tggan_ortu =2 THEN Hasil = TIDAK
5. IF jumlah_nilai = diatas_rata2, penghasilan_ortu=1 and jml_tggan_ortu = 1 THEN Hasil = TIDAK

4.2 Implementasi terhadap Data Siswa

Pada penelitian ini, peneliti menggunakan bahasa pemrograman PHP untuk menguji bagaimana proses klasifikasi data siswa. Dibawah ini merupakan uraian bagaimana menentukan keputusan apakah siswa tersebut layak mendapatkan rekomendasi beasiswa.

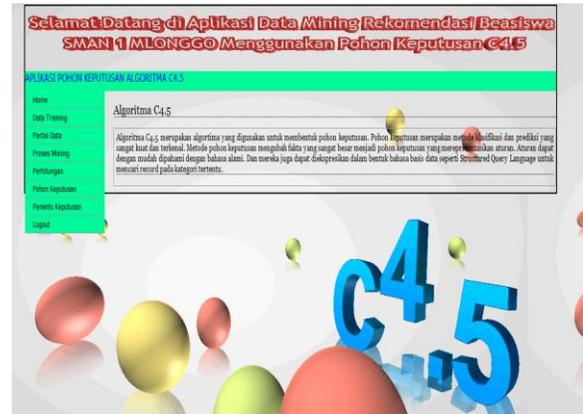
a. Tampilan Form Iog in

Untuk masuk ke dalam aplikasi harus menginputkan username dan pasword.



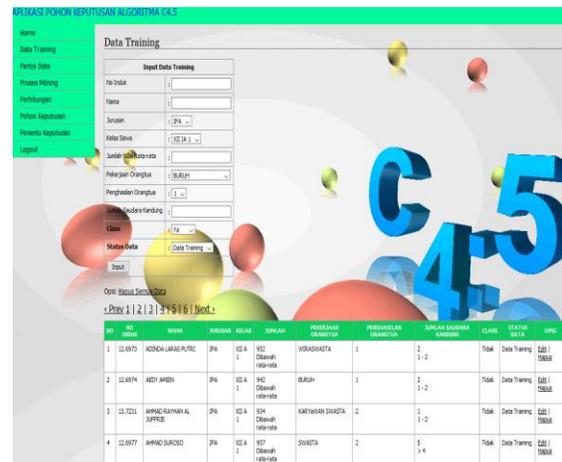
Gambar 4. 1 Tampilan log in aplikasi Penentu Keputusan

Kemudian jika username dan password sudah di inputkan, maka akan muncul halaman Home. Hasilnya akan tampak pada Gambar 4.6



Gambar 4. 2 Tampilan Home

Untuk menentukan keputusan, input data siswa yaitu Nomor Induk Siswa, Nama Siswa, Jurusan, Kelas Siswa, Jumlah Nilai Rata-Rata, Penghasilan Orangtua, dan Jumlah Saudara Kandung. Hasilnya akan tampak pada Gambar 4.7



Gambar 4. 3 Tampilan Input data

Kemudian jika data sudah di inputkan, maka akan keluar hasilnya.

5. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil pengujian metode pohon keputusan terhadap data siswa dapat disimpulkan bahwa, Penerapan metode pohon keputusan terhadap data siswa SMA N 1 Mlonggo memiliki tingkat akurasi yang cukup baik dalam menyelesaikan klasifikasi rekomendasi beasiswa. Metode pohon keputusan merupakan metode yang cukup sesuai untuk penyelesaian studi kasus dalam pemilihan siswa yang mendapatkan rekomendasi beasiswa. Tingkat akurasi yang dihasilkan oleh metode tersebut adalah 79%

5.2 Saran

Adapun saran-saran yang disampaikan berdasarkan hasil pengamatan dan analisa selama melakukan penelitian data SMA N 1 Mlonggo adalah:

1. Penelitian selanjutnya sebaiknya menggunakan data yang lebih banyak agar menghasilkan rules yang lebih akurat.
2. Penelitian selanjutnya sebaiknya menggunakan atribut yang lebih banyak agar menghasilkan data yang lebih akurat.
3. Pengujian metode ini belum sampai pada implementasi yang menghitung iterasi c4.5 kemudian menghasilkan rule dan dapat menentukan keputusan, maka perlu dibuat

sistem aplikasi yang diperuntukkan untuk pihak sekolah.

4. Untuk penelitian selanjutnya sebaiknya uji data sampai 3 kali atau lebih dengan presentase yang berbeda untuk mengetahui perbandingannya

DAFTAR PUSTAKA

- [1] L. A. Abdillah and Emigawaty, "Analisis laporan tugas akhir mahasiswa Diploma I dari sudut pandang kaidah karya ilmiah dan penggunaan teknologi informasi," *Jurnal Ilmiah MATRIK*, vol. 11, pp. 19-36, April 2009.
- [2] L. A. Abdillah, "Students learning center strategy based on e-learning and blogs," in *Seminar Nasional Sains dan Teknologi (SNST) ke-4 Tahun 2013*, Fakultas Teknik Universitas Wahid Hasyim Semarang 2013, pp. F.3.15-20.
- [3] "wikipedia," [online]. Available: <https://id.wikipedia.org/wiki/Beasiswa> [Accessed 17 September 2015].
- [4] Hermawati, F.A, *Data Mining*. Yogyakarta : ANDI, 2013.
- [5] Nofriansyah, Dicky. *Konsep Data Mining Vs Sistem Pendukung Keputusan*. Yogyakarta: Deepublish, 2014.
- [6] Yosoa Putra Raharja, "Rancang Bangun Sistem Rekomendasi Beasiswa Menggunakan Algoritma Klasifikasi C4.5 Pada Universitas Dian Nuswantoro," Universitas Dian Nuswantoro, Semarang, skripsi 2014.
- [7] Dina Maurina, " Penerapan Data Mining Untuk Rekomendasi Beasiswa

- Tepat Sasaran Menggunakan Algoritma C4.5,* Universitas Dian Nuswantoro, Semarang, skripsi 2014.
- [8] K. Hastuti, " *Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non-Aktif Semantik*, pp. pp. 241-249, 2012.
- [9] D. A. Nursela, Penerapan algoritma C4.5 untuk klasifikasi tingkat keganasan kanker payudara, semarang : universitas dianuswantoro, 2014
- [10] Joko Anik Andriani, " *Penerapan Algoritma C4.5 Pada Program Klasifikasi Mahasiswa Dropout*," AMIK BSI Jakarta, Seminar Nasional Matematika, 2012.
- [11] Swastina, L. Penerapan Algoritma C4.5 Untuk Penentuan Jurusan Mahasiswa. *GEMA AKTUALITA, Vol.2 No.1.* 93-98, Juni 2013.
- [12] Sunjana, "Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5," *Seminar Nasional Aplikasi Teknologi Informasi*, pp. D31-D34, Juni 2010.
- [13] Kusri and Emha Taufiq Luthfi, *Algoritma Data Mining*. Yogyakarta: ANDI, 2009.
- [14] Rahmadya T. H dan Herlawati Prabowo P. W, *Penerapan Data Mining dengan Matlab*. Bandung: Rekayasa Sains, 2013.
- [15] Hermawati. F. Astuti. (2013). *Data Mining*. Yogyakarta: Andi Offset.
- [16] Wibowo, Ari. (2011). *Prediksi Nasabah Potensial Menggunakan Metode Klasifikasi Pohon Biner*: Universitas Politeknik Negri Batam.
- [17] Dwi Untari, " *Data Mining Untuk Menganalisa Prediksi Mahasiswa Berpotensi Non-Aktif Menggunakan Metode decision Tree C4.5,*" Universitas Dian Nuswantoro, Semarang, pdf skripsi 2014.
- [18] Jiawei Han and Micheline Kamber, *Data Mining Concepts and Techniques*, 2nd ed. San Francisco: Morgan Kaufmann Publishers, 2006.
- [19] V Wiratna Sujarweni, *Metodologi Penelitian*. Yogyakarta: PUSTAKABARUPERSS, 2014.
- [20] Han, J. and Kamber, M, 2006, " *Data Mining Concepts and Techniques Second Edition* ". Morgan Kauffman, San Francisco.
- [21] Eko Prasetyo, *Mengolah Data Menjadi Informasi Menggunakan Matlab*, Aldo Sahala, Ed. Yogyakarta: ANDI, 2014