

PENERAPAN ALGORITMA C4.5 DALAM MEMPREDIKSI ASAL CALON MAHASISWA BERBASIS WEBSITE (STUDI KASUS : FAKULTAS HUKUM UNIVERSITAS MATARAM)

Fauzi Rahman, Hani Zulfia Zahro', F.X. Ariwibisono

Program Studi Teknik Informatika S1, Fakultas Teknologi Industri
Institut Teknologi Nasional Malang, Jalan Raya Karanglo km 2 Malang, Indonesia
1618069@scholar.itn.ac.id

ABSTRAK

Universitas Mataram (UNRAM) merupakan salah satu perguruan tinggi yang berdiri di Nusa Tenggara Barat (NTB) sejak tanggal 26 Juni 1962. Dalam perkembangannya, sejak didirikannya hingga tahun 2012 Universitas Mataram telah memiliki 9 fakultas dan salah satunya merupakan Fakultas Hukum. Banyaknya jumlah pendaftar mahasiswa baru tiap tahun yang berasal dari beberapa daerah Kota dan Kabupaten Nusa Tenggara Barat menarik rasa ingin tahu akan jumlahnya serta dominan pendaftar dari sekolah mana saja.

Algoritma C4.5 merupakan algoritma yang dapat membentuk suatu pohon keputusan dengan mengubah fakta yang besar dan dapat mempresentasikan aturan. Algoritma C4.5 merupakan salah satu model pengklasifikasian dalam data mining, dengan memanfaatkan data mahasiswa fakultas hukum universitas mataram yang memiliki beberapa atribut diharapkan dapat menghasilkan suatu informasi rules atau aturan.

Hasil dari penelitian pada bagian pengujian fungsional sistem dan web browser dapat berjalan dengan baik sesuai yang diharapkan. Sedangkan hasil pengolahan data pada sistem didapatkan hasil jumlah data dengan keterangan naik 20 data, keterangan turun 2 data dan keterangan tetap 5 data atau sekitar 74,07% merupakan keterangan naik yang dapat dilihat pada grafik hasil.

Kata Kunci : Sistem Prediksi, Data Mining, Algoritma C4.5

1. PENDAHULUAN

1.1. Latar Belakang

Universitas Mataram (UNRAM) merupakan salah satu perguruan tinggi yang berdiri di Nusa Tenggara Barat (NTB) sejak tanggal 26 Juni 1962. Lokasi tepatnya berada pada Jl. Majapahit No. 62, Mataram, Nusa Tenggara Barat. Dalam perkembangannya, sejak didirikannya hingga tahun 2012 Universitas Mataram telah memiliki 9 fakultas dan salah satunya merupakan Fakultas Hukum.

Banyaknya jumlah pendaftar mahasiswa baru tiap tahun yang berasal dari beberapa daerah Kota dan Kabupaten Nusa Tenggara Barat menarik rasa ingin tahu akan jumlahnya serta dominan pendaftar dari sekolah mana saja. Dalam mengetahui calon mahasiswa Fakultas Hukum Universitas Mataram berasal dari daerah sekolah mana saja yang banyak peminatnya, maka diperlukan prediksi asal calon mahasiswa Fakultas Hukum Universitas Mataram.

Algoritma C4.5 merupakan salah satu model pengklasifikasian dalam data mining yang dapat digunakan untuk memprediksi, dengan memanfaatkan data mahasiswa Fakultas Hukum Universitas Mataram yang memiliki beberapa atribut diharapkan dapat menghasilkan suatu informasi rules atau aturan yang akan digunakan sebagai sebuah acuan dalam memprediksi asal calon mahasiswa yang baru.

Data yang digunakan dalam pengolahan merupakan data mahasiswa yang mendaftar selama 3 tahun terakhir yaitu data mahasiswa angkatan 2017, 2018 dan 2019 yang didapatkan dari Fakultas Hukum

Universitas Mataram. Sistem prediksi yang akan dikembangkan merupakan sistem berbasis *website* dengan bahasa pemrograman *PHP*.

Berdasarkan latar belakang diatas, maka dilakukan Penerapan Algoritma C4.5 dalam memprediksi asal calon mahasiswa Fakultas Hukum Universitas Mataram dengan menggunakan data mahasiswa yang mendaftar selama 3 tahun terakhir yang kemudian diharapkan dapat berguna bagi suatu instansi tersebut.

1.2. Rumusan Masalah

Berdasarkan pada latar belakang yang telah disampaikan mengenai Fakultas Hukum Universitas Mataram dan metode data mining Algoritma C4.5, maka dapat dirumuskan permasalahannya bagaimana suatu metode Algoritma C4.5 dapat digunakan dalam memprediksi asal calon mahasiswa?

1.3. Tujuan

Adapun tujuan yang ingin dicapai dalam penelitian ini merupakan memprediksi asal calon mahasiswa menggunakan data mahasiswa angkatan 2017, 2018 dan 2019 dengan metode Algoritma C4.5.

1.4. Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut:

1. Data yang digunakan merupakan data mahasiswa angkatan 2017, 2018 dan 2019 yang didapatkan dari Fakultas Hukum Universitas Mataram.

2. Variabel dari data digunakan merupakan variabel sekolah, asal dan perbandingan dari data.
3. Metode yang digunakan merupakan Algoritma C4.5 untuk membuat pohon keputusan.
4. Platform yang digunakan merupakan berbasis website.
5. Bahasa pemrograman yang digunakan merupakan PHP.

2. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Pegawai yang baik dan berkualitas memiliki dampak yang positif terhadap perusahaan, karena itu perusahaan harus menyeleksi orang-orang yang ingin masuk dan bekerja ke dalam perusahaan dengan baik. Metode Algoritma C4.5 digunakan untuk melakukan prediksi dan klasifikasi terhadap calon pegawai yang berpotensi untuk masuk ke dalam perusahaan dengan cara membuat pohon keputusan berdasarkan data-data pegawai yang mencakup berbagai macam informasi mulai dari nama, umur, gaji, alamat, dan atribut lain. Hasil tingkat keberhasilan prediksi calon pegawai baru di PT WISE secara keseluruhan yang telah diukur menggunakan metode ten-fold cross validation adalah sebesar 71%. [1]

Kurangnya pengetahuan tentang kontrasepsi menyebabkan tidak sedikit akseptor lebih memilih menggunakan kontrasepsi berdasarkan coba-coba atau bahkan mengikuti saran dari orang lain. Dalam penelitian ini akan dilakukan analisa data penggunaan jenis kontrasepsi menggunakan klasifikasi data mining yakni Algoritma C4.5 dengan menggunakan beberapa parameter diantaranya usia, jumlah anak, tekanan darah dan riwayat penyakit yang diaplikasikan dalam bentuk web guna mempermudah para akseptor dalam mendapatkan nilai informasi yang lebih cepat dan fleksibel. Evaluasi hasil prediksi penggunaan jenis kontrasepsi dengan menggunakan Algoritma C4.5 ini, dievaluasi dengan pengujian Cross Validation menghasilkan tingkat akurasi sebesar 85,38%. [2]

Penyakit hepatitis merupakan penyakit peradangan hati karena infeksi virus yang menyerang dan menyebabkan kerusakan pada sel-sel dan fungsi organ hati. Penyakit hepatitis merupakan penyakit cikal bakal dari kanker hati. Penelitian ini menggunakan metode Algoritma C.45 dengan menghitung dan mendapatkan rule-rule kemungkinan kelangsungan hidup penderita apakah hidup atau mati. Hasil klasifikasi data mining Algoritma C4.5 menghasilkan akurasi 77,29% dan nilai AUC 0,846 yang termasuk dalam Good Classification. [3]

Pada umumnya, metode yang digunakan untuk menganalisis data nasabah adalah dengan cara mengklasifikasikan semua nasabah yang telah melunasi angsuran kreditnya ke dalam target pemasaran, sehingga metode ini menyebabkan tingginya biaya operasional marketing. Penelitian ini dilakukan untuk membantu menyelesaikan permasalahan tersebut dengan merancang sebuah

aplikasi data mining yang berfungsi untuk memprediksi kriteria nasabah kredit yang berpotensi melakukan peminjaman (kredit) terhadap bank. Algoritma yang dipakai sebagai algoritma pembentuk pohon keputusannya adalah Algoritma C4.5. Adapun data yang di proses dalam penelitian ini adalah data angsuran nasabah kredit Bank XY pada bulan Juni 2009. [4]

Tryout merupakan tahap evaluasi yang dilakukan untuk menghadapi ujian nasional dikarenakan soal-soal yang terdapat di dalam tryout mengacu pada materi dalam ujian nasional. Namun, siswa tidak semuanya dapat menyelesaikan soal-soal tryout dengan mudah dan perlu mendapat perhatian dari pihak sekolah. Penelitian dilakukan menggunakan metode Algoritma C4.5 dengan atribut data tryout selama 3 tahun kebelakang dan label yang digunakan ada 2 macam yaitu siap dan belum siap. Penggunaan algoritma ini didukung dengan simulasi yang dilakukan menggunakan aplikasi RapidMiner dan mendapatkan nilai akurasi sebesar 99,48%. [5]

2.2 Data Mining

Data mining merupakan suatu proses penambangan, penggalian maupun pengolahan sebuah kumpulan data guna mencari atau mendapatkan informasi baru yang dapat bermanfaat dan berguna dalam pemecahan suatu masalah.

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam basis data. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar. [6]

Menurut Gartner Group, data mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. [6]

2.3 Klasifikasi

Klasifikasi merupakan suatu proses pengelompokan kumpulan obyek dari sebuah basis data ke dalam suatu kelas-kelas berdasarkan kesamaan maupun klasifikasi yang telah ditentukan.

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori yaitu pendapatan tinggi, pendapatan sedang dan pendapatan rendah. [6]

2.4 Decision Tree

Pohon keputusan merupakan suatu metode klasifikasi yang mempresentasikan aturan, dimana aturan tersebut dapat dengan mudah diinterpretasi oleh manusia. Suatu model pohon keputusan terdiri

dari kumpulan aturan yang membagi jumlah kelompok yang besar menjadi lebih kecil dengan memperhatikan variabel yang ditujunya.

Pohon keputusan adalah metode yang berfungsi untuk mengubah fakta menjadi pohon keputusan yang merepresentasikan aturan yang dapat mudah dimengerti dengan bahasa alami. Proses dari pohon keputusan ini dimulai dari node akar hingga node daun yang dilakukan secara rekursif dimana setiap percabangan menyatakan kondisi dan setiap ujung pohon akan menyatakan keputusan.[7]

2.5 Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma yang dapat digunakan dalam membentuk suatu pohon keputusan, dimana metode pohon keputusan mengubah fakta yang besar menjadi sebuah pohon keputusan yang dapat mempresentasikan aturan.

Algoritma C4.5 yaitu sebuah algoritma yang digunakan untuk membangun decision tree (pengambilan keputusan). Algoritma C.45 adalah salah satu algoritma induksi pohon keputusan yaitu ID3 (Iterative Dichotomiser 3). ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel training, label training dan atribut. Algoritma C4.5 merupakan pengembangan dari ID3. Beberapa pengembangan yang dilakukan pada C4.5 adalah sebagai antara lain bisa mengatasi missing value, bisa mengatasi kontinu data, dan pruning.[6]

Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk tiap-tiap nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti dalam persamaan berikut:

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

Keterangan :

- S : himpunan kasus
- A : atribut
- n : jumlah partisi atribut A
- |S_i| : jumlah kasus pada partisi ke-i
- |S| : jumlah kasus dalam S

Setelah mendapatkan nilai gain, adapun hal lainnya yang perlu dilakukan perhitungan adalah mencari nilai entropy suatu atribut. Entropy digunakan untuk menyatakan suatu kelas dan mengukur suatu ketidakkasian himpunan kasus.

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

Keterangan :

- S : himpunan kasus
- n : jumlah partisi S
- P_j : proporsi dari S_i terhadap S

3. METODE PENELITIAN

3.1 Analisis Sistem

Sistem pada Penerapan Algoritma C4.5 Dalam Memprediksi Asal Calon Mahasiswa Fakultas Hukum Universitas Mataram merupakan suatu sistem guna memberikan informasi daerah asal mana yang memiliki banyak peminatnya dalam mendaftar sebagai calon mahasiswa di Fakultas Hukum Universitas Mataram. Sehingga dalam penerapannya dibutuhkan data-data mahasiswa pendaftar Fakultas Hukum Universitas Mataram selama 3 tahun sebelumnya.

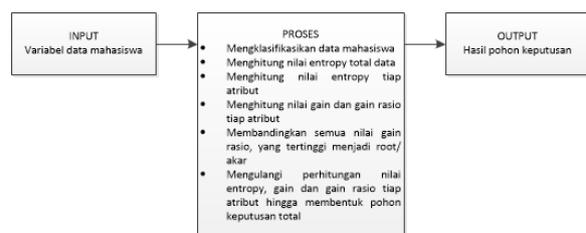
3.2 Sistem yang Akan Dibangun

Sistem yang akan dibangun merupakan suatu implementasi dari penerapan Algoritma C4.5 guna menampilkan atau memberitahukan pengguna tentang pohon keputusan dan rule-rule atau aturan yang terbentuk dari hasil prediksi asal calon mahasiswa. Oleh karenanya, aplikasi harus dapat memenuhi beberapa kebutuhan antara lain :

1. Penerapan Algoritma C4.5 dalam memprediksi asal calon mahasiswa dapat menampilkan pohon keputusan hasil perhitungan.
2. Penerapan Algoritma C4.5 dalam memprediksi asal calon mahasiswa dapat menampilkan rule-rule atau aturan yang terbentuk dari pohon keputusan.

3.3 Desain Arsitektur Sistem

Dalam penelitian ini, adapun suatu desain arsitektur sistem guna mendeskripsikan dan mendefinisikan dari aplikasi yang dibentuk, dapat dilihat pada Gambar 1 berikut ini :



Gambar 1 Desain Arsitektur Sistem

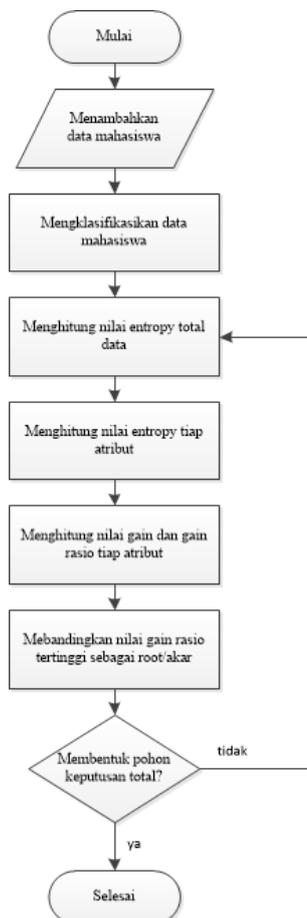
Pada desain arsitektur sistem Gambar 3.1 diatas terdapat input, proses, dan output, berikut penjelasan mengenai desain tersebut :

1. Input
 - a. Variabel data mahasiswa merupakan informasi data diri dari tiap mahasiswa.
2. Proses
 - a. Mengklasifikasikan data mahasiswa berdasarkan tiap variabel.
 - b. Menghitung nilai entropy total dari variabel.
 - c. Menghitung nilai entropy tiap atribut dari variabel.

- d. Menghitung nilai gain tiap atribut dari variabel.
 - e. Membandingkan semua nilai gain, yang tertinggi akan menjadi root/akar pohon keputusan.
 - f. Mengulangi perhitungan nilai entropy, gain dan *gain rasio* tiap atribut hingga membentuk pohon keputusan.
3. Output
- a. Menampilkan hasil pohon keputusan.

3.4 Flowchart

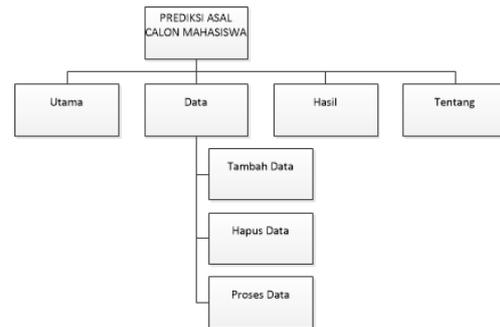
Berikut merupakan flowchart Algoritma C4.5 dari penerapan Algoritma C4.5 dalam memprediksi asal calon mahasiswa Fakultas Hukum Universitas Mataram yang telah dibuat pada Gambar 2 berikut ini:



Gambar 2 Flowchart Algoritma C4.5

3.5 Struktur Menu

Struktur menu menampilkan suatu tatanan menu yang terdapat pada aplikasi yang dibangun. adapun struktur menu dapat dilihat pada Gambar 3 berikut ini:



Gambar 3 Struktur Menu

Pada struktur menu diatas, terdapat menu utama yang dapat digunakan untuk menampilkan halaman beranda website, menu data yang dapat digunakan untuk menampilkan halaman pemrosesan data dengan menambah, mengubah dan menghapus data yang akan diproses, menu hasil yang dapat digunakan untuk menampilkan halaman hasil dari pemrosesan data dan juga menu tentang yang dapat digunakan untuk menampilkan informasi tentang pengembang.

3.6 Perancangan Database

1. Tabel Data Latih

Tabel data latih merupakan tabel yang digunakan untuk menyimpan data mahasiswa dari tahun sebelumnya atau data yang akan dibandingkan yang disimpan ke dalam database MySQL. Tabel data latih ditunjukkan pada Tabel 1.

Tabel 1 Struktur Tabel Data Latih

No	Name	Type
1	Id	Int(11)
2	Nama	Varchar(50)
3	Kelamin	Varchar (50)
4	Sekolah	Varchar (50)
5	Angkatan	Varchar (50)
6	Asal	Varchar (50)

2. Tabel Data Uji

Tabel data uji merupakan tabel yang digunakan untuk menyimpan data mahasiswa dari tahun yang baru atau data pembanding yang disimpan ke dalam database MySQL. Tabel data uji ditunjukkan pada Tabel 2.

Tabel 2 Struktur Tabel Data Uji

No	Name	Type
1	Id	Int(11)
2	Nama	Varchar(50)
3	Kelamin	Varchar (50)
4	Sekolah	Varchar (50)
5	Angkatan	Varchar (50)
6	Asal	Varchar (50)

3. Tabel Klasifikasi

Tabel klasifikasi merupakan tabel yang digunakan untuk menyimpan data klasifikasi naik,

turun dan tetap jumlahnya mahasiswa dari perbandingan data latih dan data uji. Tabel klasifikasi ditunjukkan pada Tabel 3.

Tabel 3 Struktur Tabel Klasifikasi

No	Name	Type
1	Id	Int(11)
2	Sekolah	Varchar(50)
3	Asal	Varchar (50)
4	Jumlah1	Int(11)
5	Jumlah2	Int(11)
6	Perbandingan	Int(11)
7	Keterangan	Varchar(50)

4. Tabel Atribut

Tabel atribut merupakan tabel yang digunakan untuk menyimpan data atribut dan nilai atribut apa saja yang akan diolah dalam proses perhitungan C45. Tabel atribut ditunjukkan pada tabel 4.

Tabel 4 Struktur Tabel Atribut

No	Name	Type
1	Id	Int(11)
2	Atribut	Varchar(50)
3	Nilai_atribut	Varchar (50)

5. Tabel Mining

Tabel mining merupakan tabel yang digunakan untuk menyimpan data jumlah kasus, nilai entropy, gain, splitinfo dan gain ratio dari hasil proses perhitungan C45. Tabel mining diunjukkan pada tabel 5.

Tabel 5 Struktur Tabel Mining

No	Name	Type
1	Id	Int(11)
2	Atribut	Varchar (50)
3	Nilai_atribut	Varchar (50)
4	Jml_kasus_total	Double
5	Jml_naik	Double
6	Jml_turun	Double
7	Jml_tetap	Double
8	Entropy	Double
9	Inf_gain_temp	Double
10	Inf_gain	Double
11	Split_info_temp	Double
12	Split_info	Double
13	Gain_ratio	Double

6. Tabel Iterasi

Tabel iterasi merupakan tabel yang digunakan untuk menyimpan data nilai iterasi atau data nilai perulangan pada proses mining. Tabel gain ditunjukkan pada Tabel 6.

Tabel 6 Struktur Tabel Iterasi

No	Name	Type
1	Id	Int(11)
2	Iterasi	Int(11)
3	Atribut_gain_ratio_max	Varchar (50)
4	Atribut	Varchar (50)

No	Name	Type
5	Nilai_atribut	Varchar (50)
6	Jml_kasus_total	Double
7	Jml_naik	Double
8	Jml_turun	Double
9	Jml_tetap	Double
10	Entropy	Double
11	Inf_gain	Double
12	Split_info	Double
13	Gain_ratio	Double

7. Tabel Pohon Keputusan

Tabel pohon keputusan merupakan tabel yang digunakan untuk menyimpan data rule atau aturan yang terbentuk dari hasil pohon keputusan. Tabel pohon keputusan ditunjukkan pada Tabel 7.

Tabel 7 Struktur Tabel Pohon Keputusan

No	Name	Type
1	Id	Int(11)
2	Parent	Text
3	Akar	Text
4	Keputusan	Varchar (50)

3.7 Rancangan Antarmuka Sistem

Rancangan antarmuka sistem merupakan desain tampilan dari sistem website yang akan dibangun.

1. Halaman utama merupakan halaman utama ketika membuka atau memuat sebuah aplikasi website, rancangan halaman utama dapat dilihat pada Gambar 4 berikut ini.



Gambar 4 Halaman Utama

2. Halaman data merupakan halaman untuk menambah, mengubah dan menghapus data yang digunakan dalam pemrosesan data pada aplikasi website, rancangan halaman data dapat dilihat pada Gambar 5 berikut ini.



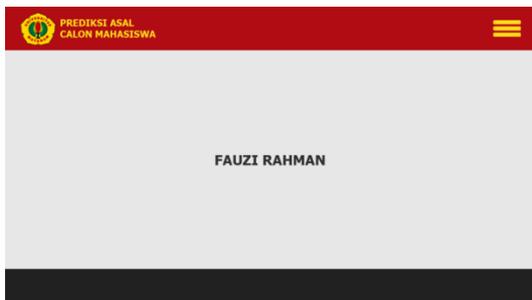
Gambar 5 Halaman Data

3. Halaman hasil merupakan halaman untuk menampilkan hasil dari pemrosesan data berupa tabel hasil dan juga chart, rancangan halaman hasil dapat dilihat pada Gambar 6 berikut ini.



Gambar 6 Halaman Hasil

4. Halaman tentang merupakan halaman untuk menampilkan informasi tentang pengembang, rancangan halaman tentang dapat dilihat pada Gambar 7 berikut ini.



Gambar 7 Halaman Tentang

4. HASIL DAN PEMBAHASAN

4.1 Pengujian Metode

1. Pengklasifikasian Data

Dalam proses pengklasifikasian data dilakukan menggunakan 2 buah data mahasiswa dengan tahun yang berbeda dengan mengambil variabel sekolah dan asal serta jumlahnya yang akan dibandingkan sehingga tercipta variabel perbandingan untuk mengetahui naik, turun ataupun tetap jumlah mahasiswa pertahunnya.

Tabel 8 Klasifikasi

Sekolah	Asal	Perbandingan	Keterangan
SMA	Mataram	60	Naik
SMA	Lombok Barat	9	Naik
SMA	Lombok Tengah	23	Naik
SMA	Lombok Timur	49	Naik
SMA	Lombok Utara	2	Naik
SMA	Sumbawa	21	Naik
SMA	Sumbawa Barat	8	Naik
SMA	Bima	31	Naik
SMA	Dompu	15	Naik
SMK	Mataram	4	Naik
SMK	Lombok Barat	0	Tetap
SMK	Lombok Tengah	4	Naik
SMK	Lombok Timur	2	Naik
SMK	Lombok Utara	0	Tetap

Sekolah	Asal	Perbandingan	Keterangan
SMK	Sumbawa	2	Naik
SMK	Sumbawa Barat	0	Tetap
SMK	Bima	1	Turun
SMK	Dompu	1	Naik
MA	Mataram	19	Naik
MA	Lombok Barat	6	Naik
MA	Lombok Tengah	5	Naik
MA	Lombok Timur	12	Naik
MA	Lombok Utara	0	Tetap
MA	Sumbawa	0	Tetap
MA	Sumbawa Barat	2	Naik
MA	Bima	5	Naik
MA	Dompu	1	Naik

2. Menghitung Nilai Entropy, Gain dan Gain Rasio

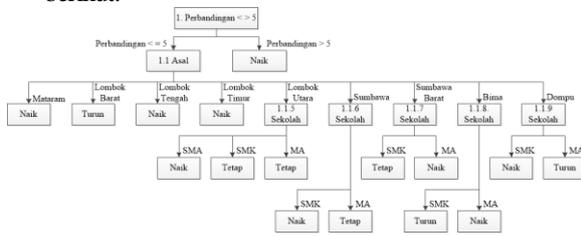
Dari data yang telah diklasifikasi dilakukan perhitungan sebuah entropy, gain dan gain rasio guna menentukan variabel mana yang akan menjadi root atau akar dari pohon keputusan dan membentuk sebuah percabangan.

Tabel 9 Nilai Entropy, Gain dan Gain Rasio

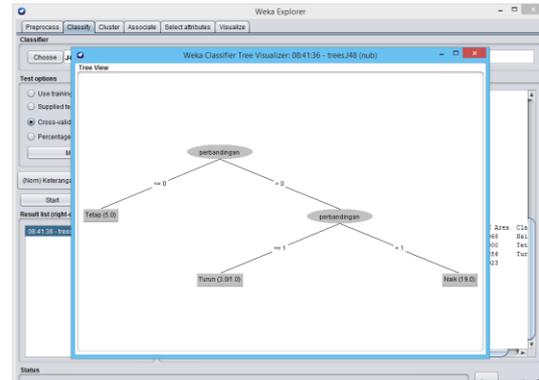
Atribut	Nilai Atribut	Jumlah Data	Naik	Turun	Tetap	Entropy	Gain	Split Info	Gain Rasio
Sekolah	SMA	9	9	0	0	0,000			
	SMK	9	5	1	3	1,352			
	MA	9	6	1	2	1,224	0,191	1,585	0,120
Asal	Mataram	3	3	0	0	0,000			
	Lombok Barat	3	2	0	1	0,918			
	Lombok Tengah	3	3	0	0	0,000			
	Lombok Timur	3	3	0	0	0,000			
	Lombok Utara	3	1	0	2	0,918			
	Sumbawa	3	2	0	1	0,918			
	Sumbawa Barat	3	2	0	1	0,918			
	Bima	3	2	1	0	0,918			
	Dompu	3	2	1	0	0,918	0,437	3,170	0,138
	Perbandingan <= 5	Perbandingan <=5	16	9	2	5	1,366		
Perbandingan >5		11	11	0	0	0,000	0,240	0,975	0,246
Perbandingan <= 10	Perbandingan <=10	19	12	2	5	1,267			
	Perbandingan >10	8	8	0	0	0,000	0,157	0,877	0,180

3. Pohon Keputusan

Dari proses pengolahan data terbentuk suatu pohon keputusan total yang dapat dilihat pada Gambar berikut.



Gambar 8 Pohon Keputusan



Gambar 9 Pohon Keputusan WEKA

4. Rule atau Aturan

Dari proses pengolahan data hingga membentuk suatu pohon keputusan total, maka diperoleh rule atau aturan sebagai berikut.

Tabel 10 Rule atau Aturan

NO	Aturan
1	if Perbandingan <= 5 and Asal = Mataram then Keterangan = Naik
2	if Perbandingan <= 5 and Asal = Lombok Barat then Keterangan = Tetap
3	if Perbandingan <= 5 and Asal = Lombok Timur then Keterangan = Naik
4	if Perbandingan <= 5 and Asal = Lombok Tengah then Keterangan = Naik
5	if Perbandingan <= 5 and Asal = Sumbawa and Sekolah = SMK then Keterangan = Naik
6	if Perbandingan <= 5 and Asal = Sumbawa and Sekolah = MA then Keterangan = Tetap
7	if Perbandingan <= 5 and Asal = Sumbawa Barat and Sekolah = SMK then Keterangan = Naik
8	if Perbandingan <= 5 and Asal = Sumbawa Barat and Sekolah = MA then Keterangan = Naik
9	if Perbandingan <= 5 and Asal = Bima and Sekolah = SMK then Keterangan = Turun
10	if Perbandingan <= 5 and Asal = Bima and Sekolah = MA then Keterangan = Naik
11	if Perbandingan <= 5 and Asal = Lombok Utara and Sekolah = SMA then Keterangan = Naik
12	if Perbandingan <= 5 and Asal = Lombok Utara and Sekolah = SMK then Keterangan = Tetap
13	if Perbandingan <= 5 and Asal = Lombok Utara and Sekolah = MA then Keterangan = Tetap
14	if Perbandingan <= 5 and Asal = Dompu and Sekolah = SMK then Keterangan = Naik
15	if Perbandingan <= 5 and Asal = Dompu and Sekolah = MA then Keterangan = Turun
16	if Perbandingan > 5 then Keterangan = Naik

5. Perhitungan WEKA

Metode Algoritma C4.5 pada aplikasi WEKA menghasilkan sebuah pohon keputusan yang dapat dilihat seperti pada tampilan Gambar 9 berikut ini.

6. Proses Perhitungan Sistem

Proses perhitungan metode Algoritma C4.5 pada sistem dapat dilihat pada halaman perhitungan.php seperti pada tampilan Gambar 10 berikut ini.

Gambar 10 Proses Perhitungan Sistem

7. Hasil Pada Sistem

Hasil daripada pengolahan sistem prediksi calon mahasiswa ini ditampilkan berupa suatu grafik data yang dapat dilihat pada tampilan Gambar 11 berikut ini.

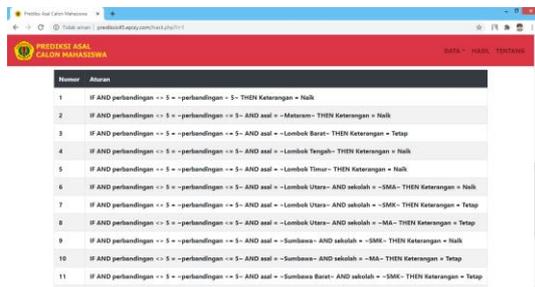


Gambar 11 Grafik Hasil

Hasil pohon keputusan dari pengolahan data pada sistem, dapat dilihat pada Gambar 12 berikut ini.

Gambar 12 Pohon Keputusan Sistem

Adapun juga hasil rule atau aturan dari pohon keputusannya dari pengolahan data yang dapat dilihat pada Gambar 13 berikut ini.



Gambar 13 Rule atau Aturan Sistem

4.2 Pengujian Fungsional Sistem

Pengujian fungsional sistem merupakan pengujian guna mengetahui fitur-fitur yang terdapat pada sistem prediksi asal calon mahasiswa dapat bekerja sesuai yang diinginkan atau tidak. Beberapa fitur yang diuji dan hasil pengujian dapat dilihat pada tabel berikut.

Tabel 11 Pengujian Fungsional Sistem

No	Fitur yang diuji	Hasil	
		Tidak Bekerja	Bekerja
1	Berita utama		✓
2	Pengumuman		✓
3	Tambah data		✓
4	Hapus data		✓
5	Proses data		✓
6	Grafik data		✓

4.3 Pengujian Web Browser

Pengujian web browser merupakan pengujian guna mengetahui sistem prediksi asal calon mahasiswa dapat bekerja dan berjalan dengan semestinya pada tiap browser yang digunakan. Beberapa browser yang diuji dan hasil pengujian dapat dilihat pada Tabel berikut.

Tabel 12 Pengujian Web Browser

No	Item yang diuji	Hasil		
		GC V:83.0.41 03.61 (32-bit)	MF V:77.0 (32-bit)	IE V:11.0.96 00.16438 (32-bit)
1	Halaman utama	✓	✓	✓
2	Halaman data latih	✓	✓	✓
3	Halaman data uji	✓	✓	✓
4	Halaman hasil	✓	✓	✓
5	Halaman tentang	✓	✓	✓

Keterangan :

GC : Google Chrome

MF : Mozilla Firefox

IE : Internet Explorer

✓ : Berjalan

× : Tidak Berjalan

4.4 Perbandingan Hasil

Perbandingan hasil merupakan perbandingan hasil dari perhitungan manual, aplikasi Weka dan sistem yang dilakukan guna mengetahui apakah terdapat kesamaan ataupun perbedaan pada hasil perhitungannya. Perbandingan hasil dapat dilihat pada Tabel 13 berikut.

Tabel 13 Perbandingan Hasil

No	Hasil Perhitungan	Hasil	
		Pohon Keputusan	Aturan
1	Manual	✓	✓
2	Aplikasi Weka	×	×
3	Sistem	✓	✓

Keterangan :

✓ : Sama

× : Tidak Sama

Analisa :

Pada perbandingan hasil pohon keputusan dan aturan yang dilakukan pada perhitungan manual dan sistem terdapat kesamaan, sedangkan pada aplikasi Weka tidak sama. Hal ini dikarenakan pada perhitungan manual dan sistem data numerik atribut perbandingan ditentukan nilai ambang batasnya yaitu perbandingan < > 5 dan perbandingan < > 10. Sedangkan pada aplikasi Weka menghitung seluruh nilai ambang batas yang dapat terbentuk pada data numerik atribut perbandingan.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan daripada hasil perancangan dan implementasi pada sistem, dapat dijabarkan beberapa kesimpulan sebagai berikut:

- Berdasarkan hasil pengolahan data pada sistem didapatkan hasil jumlah data dengan keterangan naik 20 data, keterangan turun 2 data dan keterangan tetap 5 data atau sekitar 74,07% merupakan keterangan naik yang dapat dilihat pada grafik hasil.
- Berdasarkan perbandingan hasil perhitungan manual dan sistem dibandingkan dengan perhitungan pada Weka, pohon keputusan yang dihasilkan berbeda dikarenakan perhitungan manual dan sistem data numerik atribut perbandingan ditentukan nilai ambang batasnya yaitu perbandingan < > 5 dan perbandingan < > 10. Sedangkan pada aplikasi Weka menghitung seluruh nilai ambang batas yang dapat terbentuk pada data numerik atribut perbandingan.
- Berdasarkan hasil pengujian sistem, dimana pengujian fungsional sistem fitur yang diuji dapat bekerja sedangkan pengujian web browser pada 3 peramban dapat bekerja dengan baik sesuai yang diharapkan.

5.2 Saran

Selain kesimpulan, adapun saran yang ingin penulis sampaikan agar sistem kedepannya dapat lebih baik yaitu sebagai berikut:

1. Mengembangkan sistem aplikasi prediksi asal calon mahasiswa sehingga dapat menampilkan informasi data maupun hasil yang lebih informatif.
2. Mengembangkan sistem aplikasi prediksi asal calon mahasiswa sehingga dapat diterapkan pada sistem prediksi yang lainnya.

DAFTAR PUSTAKA

- [1] Harryanto, F. F., & Hansun, S. (2017). Penerapan Algoritma C4. 5 untuk Memprediksi Penerimaan Calon Pegawai Baru di PT WISE. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 3(2), 95-103.
- [2] Wajhillah, R., & Yulianti, I. (2017). Penerapan algoritma c4. 5 untuk prediksi penggunaan jenis kontrasepsi berbasis web. *KLIK-KUMPULAN JURNAL ILMU KOMPUTER*, 4(2), 160-173.
- [3] Septiani, W. D. (2014). Penerapan Algoritma C4. 5 untuk prediksi penyakit Hepatitis. *Jurnal Techno Nusa Mandiri*, 11(1), 69-78.
- [4] Maburur, A. G. (2012). Penerapan data mining untuk memprediksi kriteria nasabah kredit. *KOMPUTA: Jurnal Komputer dan Informatika*, 1(1).
- [5] Rahma, N. Z., & Setyono, A. (2018). Penerapan Algoritma C4. 5 Dalam Memprediksi Kesiapan Siswa SMP IT PAPB Semarang Menghadapi Ujian Nasional. *SISFOTENIKA*, 8(1), 35-46.
- [6] Kusriani, E. T. L. (2009). Algoritma data mining. *Yogyakarta: Andi Offset*.
- [7] Slamet, A. (2007). Manajemen sumber daya manusia. Universitas Negeri Semarang Press. Semarang.