

KLASIFIKASI GAYA BELAJAR PADA MAHASISWA MENGGUNAKAN METODE ALGORITMA C4.5

¹Arif Rahmatullah, ²Suherman, ³Haris Triono Sigit

¹ arifrahmatullah2669@gmail.com

² suherman.unsera@gmail.com

³ haristsigit@gmail.com

^{1,2,3} Jurusan Teknik Informatika, Fakultas Teknologi Informasi, Universitas Serang Raya
Jl. Raya Serang – Cilegon Km. 05 (Taman Drangong), Serang – Banten

Intisari— Gaya belajar adalah cara individu untuk menyerap informasi. Dalam perguruan tinggi banyak kegiatan yang dibutuhkan kemampuan dalam menyerap informasi seperti ketika dosen memberikan materi, sehingga gaya belajar perlu diperhatikan. Dalam penyampaian informasi yang diberikan oleh dosen ditemukannya permasalahan dan kebutuhan terkait hal tersebut, yaitu belum tersedianya penentuan klasifikasi gaya belajar pada mahasiswa. Yang mana sekarang ini dosen mengalami kesulitan dalam memberikan penilaian pencapaian kemampuan terhadap mahasiswa berdasarkan kriteria dari gaya belajar dan mengklasifikasikannya, dan juga untuk meningkatkan kualitas metode pembelajaran. Adapun saran yang diberikan untuk tercapainya tujuan dan mengatasi permasalahan tersebut yaitu dirancangannya klasifikasi gaya belajar pada mahasiswa dan pengolahan data yang dilakukan menggunakan komputer. Penelitian ini menggunakan metode Algoritma C4.5. Algoritma C4.5 termasuk dalam klasifikasi berbasis *Decision Tree*, yang mana algoritma C4.5 merupakan metode yang dapat menangani fitur dengan tipe numerik. Metode yang digunakan dalam perancangan klasifikasi yaitu *Unified Modeling Language* (UML). Perangkat lunak pendukung yang digunakan adalah bahasa pemrograman PHP Versi 7.4, dengan bantuan *script editor Visual Studio Code*, dan *database MySQL* sebagai pembangun program. Hasil dari penelitian ini yaitu klasifikasi dengan mengimplementasikan Algoritma C4.5 yang dapat mengetahui kemampuan gaya belajar pada mahasiswa, berdasarkan kriteria yang sudah ditetapkan, sehingga klasifikasi ini dapat semakin mudah, efektif, dan efisien

Kata Kunci: Gaya Belajar, Mahasiswa, Data Mining, Algoritma C4.5

I. PENDAHULUAN

1.1. Latar Belakang

Gaya belajar adalah seperangkat karakteristik pribadi yang menjadikan efektif pengajaran bagi beberapa mahasiswa dan tidak efektif bagi yang lain (Wawan Wahyudin, 2016). Menurut De Potter & Hernacki (1999), menjelaskan secara umum gaya belajar manusia dibedakan ke dalam tiga kelompok besar, yaitu gaya belajar visual, gaya belajar auditorial dan gaya belajar kinestetik. Gaya belajar visual adalah gaya belajar dengan cara melihat, mengamati, memandang, dan sejenisnya. Kekuatan gaya belajar ini terletak pada indera penglihatan.

Pada dasarnya setiap mahasiswa memiliki gaya belajar tersebut namun tidak semuanya yang berkembang secara seimbang melainkan ada yang mendominasi dengan gaya

belajar yang dimilikinya. Hal tersebut menyebabkan mahasiswa akan menyukai pembelajaran yang bervariasi yang sesuai dengan gaya belajar yang dimilikinya. Keberagaman gaya belajar mahasiswa memerlukan suatu pemilihan strategi mengajar yang cocok agar kekuatan gaya belajar mahasiswa berkembang dengan baik (Yusri Wahyuni, 2017). Musrofi (dalam Pratiwi, 2014) mengatakan hanya 30% mahasiswa yang berhasil mengikuti pembelajaran di kelas karena mereka mempunyai gaya belajar yang sesuai dengan gaya mengajar yang diterapkan dosen di dalam kelas. Sisanya, sebanyak 70% mahasiswa mengalami kesulitan dalam mengikuti pembelajaran di kelas karena mereka memiliki gaya belajar lain, yang tidak sesuai dengan gaya mengajar yang diterapkan di dalam kelas. Artinya, 70% gaya mahasiswa tidak terakomodasi oleh gaya mengajar dosen dalam pembelajaran (Jeanete Ophilia Papilaya dan Neleke Huliselan, 2016). Menurut

(Ratna Puspita Sari Putri dan Indra Waspada, 2018) dalam Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa, algoritma C4.5 dapat digunakan untuk menghasilkan prediksi kelulusan dengan nilai rata-rata precision 63.93%, recall 60.73%, dan akurasi 60.52%. Setelah pohon keputusan dipotong dengan menggunakan metode error-based pruning, didapatkan hasil yang lebih baik.

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Algoritma ini merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Ada beberapa cara untuk mengklasifikasi gaya belajar mahasiswa, salah satunya dengan menggunakan Algoritma C4.5. Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk melakukan klasifikasi atau segmentasi atau pengelompokan dan bersifat prediktif.

Cara kerja aplikasi ini yaitu *user* memilih salah satu dari kriteria yang telah disediakan. Hasil penentuan gaya belajar mahasiswa didasarkan atas hasil perhitungan dengan menggunakan metode klasifikasi *Decision Tree* dan menggunakan algoritma C4.5. Jika hasil perhitungan tersebut cenderung pada gaya belajar visual, maka mahasiswa tersebut disarankan untuk belajar dengan menggunakan gaya visual. Begitupun jika hasil perhitungannya cenderung pada auditorial atau kinestetik maka mahasiswa tersebut di sarankan untuk belajar dengan menggunakan gaya belajar auditorial atau kinestetik.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka penelitian merumuskan masalah sebagai berikut :

- 1) Bagaimana menganalisis klasifikasi gaya belajar pada mahasiswa?
- 2) Bagaimana merancang dan menerapkan aplikasi berbasis web dalam klasifikasi gaya belajar pada mahasiswa menggunakan metode algoritma C4.5?
- 3) Bagaimana mengimplementasi algoritma C4.5 ke dalam klasifikasi gaya belajar pada mahasiswa?

II. LANDASAN TEORI

2.1 Data Mining

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi

informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban, dkk. 2005).

2.2 Klasifikasi

Klasifikasi adalah salah satu teknik pada data mining yang memetakan data ke dalam kelompok atau kelas yang telah ditentukan. Klasifikasi merupakan metode supervised learning yang membutuhkan data training berlabel untuk menghasilkan sebuah aturan yang mengklasifikasikan data uji ke dalam kelompok atau kelas yang telah ditentukan (Dunham, 2003).

2.3 Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Sedangkan pohon keputusan dapat diartikan suatu cara untuk memprediksi atau mengklarifikasi yang sangat kuat. Pohon keputusan dapat membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Rumus Algoritma C4.5 terbagi menjadi 2 rumus, yang pertama rumus untuk mencari nilai gain dan yang kedua adalah rumus untuk mencari nilai entropy (Panji Bimo Nugroho Setio, 2020).

2.4 Metode Unsupervised Discretization

Di antara metode *unsupervised Discretization*, terdapat beberapa metode yang sederhana. Seperti (*equal-width Interval Discretization* dan *equal-frequency Interval Discretization*) dan yang lebih canggih, berdasarkan analisis *clustering*, seperti *k-means discretization*. Rentang kontinu dibagi menjadi sub rentang oleh lebar yang ditentukan pengguna atau Frekuensi (Joita, 2010).

Pada penelitian ini menggunakan teknik *Equal-width interval discretization*, yaitu metode diskretisasi paling sederhana yang membagi rentang nilai yang teramati pada setiap fitur/atribut dimana k adalah parameter yang disediakan oleh pengguna. Prosesnya melibatkan pemilihan nilai-nilai yang teramati dari fitur/atribut kontinyu dan menemukan nilai minimum (V_{min}) dan maksimum (V_{max}). *Interval* dapat dihitung dengan membagi kisaran nilai yang teramati untuk variabel ke dalam k yang berukuran sama dengan menggunakan rumus berikut, dimana k adalah parameter yang diberikan oleh pengguna (Alfa Saleh dan Fina Nasari, 2018).

$$Interval = \frac{V_{max} - V_{min}}{k}$$

Kemudian batas-batasnya dapat dibangun untuk $i = 1$ k-1 dengan menggunakan persamaan berikut:

$$Boundaries = V_{min} + (i * interval)$$

2.5 Gaya Belajar

Gaya belajar merupakan salah satu yang dimiliki oleh setiap individu dalam menyerap, mengatur, dan mengolah informasi yang diterima. Gaya belajar yang sesuai adalah kunci keberhasilan siswa dalam belajar. Penggunaan gaya belajar yang dibatasi hanya dalam satu gaya, terutama yang bersifat verbal atau auditorial, tentunya dapat menyebabkan banyak perbedaan dalam menyerap informasi. Oleh karena itu dalam kegiatan belajar, siswa harus dibantu dan diarahkan untuk mengenali gaya belajar yang sesuai dengan dirinya sendiri agar hasil belajar bisa maksimal (Bire, dkk. 2014).

2.6 Klasifikasi Gaya belajar

Setiap siswa memiliki karakteristik gaya belajar masing-masing menurut De Potter dalam bukunya Tutik Rachmawati dan Daryanto yang berjudul *Teori Belajar dan Proses Pembelajaran yang Mendidik* terdapat 3 modalitas (tipe) dalam gaya belajar yaitu Visual, Auditori dan Kinestetik. Pelajar visual belajar melalui apa yang mereka lihat. Auditori belajar dengan cara mendengar dan kinestetik belajar lewat gerak dan menyentuh.

2.2.9 Confusion matrix

Menurut Eko Prasetyo (2014) *Confusion matrix* merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya *confusion matrix* mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya.

Tabel 2.1 Confusion Matrix

	<i>Clasified Negative</i>	<i>Clasified Positive</i>
Actual Negative	a	b
Actual Positive	c	d

Isian "a" berisi jumlah data negatif hasil awal yang sama dengan hasil negatif klasifikasinya. Isian "b" berisi jumlah data negatif hasil awal yang berubah menjadi hasil positif di hasil klasifikasinya. Isian "c" berisi jumlah data positif hasil awal yang berubah menjadi hasil negatif klasifikasinya. Isian "d" berisi jumlah data positif hasil awal yang sama menjadi hasil positif di hasil klasifikasinya. Total jumlah isian "b" dan "c" merupakan selisih jumlah perbedaan hasil yang dihasilkan setelah proses klasifikasi. Dari hasil selisih tersebut dihitung nilai akurasi.

a. *Precision*

Precision adalah bagian data yang di ambil sesuai dengan informasi yang dibutuhkan. Rumus *Precision* adalah :

$$P = \left(\frac{d}{b + d} \right) \times 100\%$$

b. *Recall*

Recall adalah pengambilan data yang berhasil dilakukan terhadap bagian data yang relevan dengan query. Rumus *Recall* adalah :

$$R = \left(\frac{d}{c + d} \right) \times 100\%$$

c. *Accuracy*

Accuracy adalah persentase dari total data uji coba yang benar diidentifikasi. Rumus *Accuracy* adalah :

$$A = \left(\frac{a + d}{\text{total sampel}} \right) \times 100\%$$

III. DATA PENELITIAN

3.1 Data Penelitian

Penelitian ini tentang klasifikasi gaya belajar pada mahasiswa menggunakan pohon keputusan dengan metode Algoritma C4.5. Kriteria yang digunakan dalam klasifikasi gaya belajar pada mahasiswa ditunjukkan pada tabel 3.3 berikut.

Tabel 3. 1 Kriteria Klasifikasi

Indikator	Kriteria (No soal)
- Memahami sesuatu dengan asosiasi visual	1,2,3,4,5
- Belajar dengan cara mendengar - Baik dalam aktivitas lisan	6,7,8,9,10
- Belajar melalui aktifitas fisik - Selalu berorientasi pada fisik dan banyak gerak - Menyukai kegiatan coba coba	11,12,13,14,15

Untuk menentukan klasifikasi menurut pendapat Agnes Ika Kuniawati 2013 dalam penelitian klasifikasi gaya belajar pada mahasiswa, setiap kriteria memiliki jumlah soal 5 butir yang dimana jumlah seluruh soal adalah 15 butir, maka skor tertinggi yang didapatkan kemudian diklompokkan kedalam aspek visual, audio dan kinestetik, yang mana ditunjukkan oleh table 3.2 berikut.

Tabel 3. 2 Klasifikasi

No	Klasifikasi
1	Visual
2	Audio
3	Kinestetik

3.2 Perhitungan Metode Algoritma C.45

A. Developing and understanding the application domain
Tujuan yang ingin dicapai dari proses ini yaitu klasifikasi gaya belajar mahasiswa.

B. Creating a target data set

Variabel (atribut) yang digunakan yaitu nama, semester, jurusan, mata kuliah, kelas, skor_visual, skor_audio, dan skor_kinestetik. *Data point* (sample) yang akan digunakan yaitu gaya_belajar.

C. Data cleaning and preprocessing

Pada tahap ini dilakukan proses penghapusan data yang tidak diharapkan keberadaannya (*outliers*), atribut yang dihilangkan yaitu: semester, jurusan, mata kuliah, kelas.

D. Data reduction and projection

Dari tahapan data *cleaning and preprocessing* maka atribut yang digunakan yaitu skor_visual, skor_audio, dan skor_kinestetik. Ketiga atribut tersebut merupakan kriteria numerik/kontinyu. Untuk meningkatkan tingkat akurasi metode klasifikasi, maka dilakukan proses diskritisasi menggunakan teknik *unsupervised discretization* pada kriteria yang bernilai numerik/kontinyu, tujuannya adalah untuk mentransformasikan kriteria yang bernilai numerik/kontinyu menjadi kriteria kategorikal seperti berikut :

$$Interval = \frac{Vmax - Vmin}{k}$$

$$Interval = \frac{15 - 0}{3} = 5$$

Kemudian menentukan batas-batas sebagai berikut :

$$Boundaries = Vmin + (i * interval)$$

$$Boundaries (1) = 0 + (1 * 5) = 5$$

$$Boundaries (2) = 0 + (2 * 5) = 10$$

$$Boundaries (3) = 0 + (3 * 5) = 15$$

Maka diperoleh nilai kategorikal atribut sebagai berikut :

Tabel 3.3 Nilai Kategorikal

skor_visual	skor_audio	skor_kinestetik
≤5	≤5	≤5
>5	>5	>5
≤5	≤5	≤5
>5	>5	>5
≤15	≤15	≤15
>15	>15	>15

E. Choosing the data mining task

Metode data mining yang digunakan yaitu pohon klasifikasi, karena sesuai dengan tujuan yang ingin dicapai.

F. Choosing the data mining algorithm

Algoritma yang digunakan yaitu algoritma C4.5.

G. Data Mining

Tabel 3.4 Data Training Pertama

Nama	Visual	Audio	Kinestetik	Klasifikasi Awal
Rio rizky ramadhan	8	5	2	V
Muhamad Deni Hermawan	7	3	5	V
Nico Kandilo Gumay	5	6	4	A
Muhamad Ilham Ardi	7	2	6	V
Kevin Josua S	5	5	5	V
Muhammad hilal akbar	4	7	4	A
Bella Aprillia	6	5	4	V
Miftahuroyan dien marwan	2	8	5	A
Oldi Novelan	5	6	4	A
Umair Maula	8	2	5	V
Roffi Sunarya	4	6	5	A
Hoerotu Ruwaidah	9	4	2	V
M jidan mustaqim	6	3	6	V
Hilmy Bahy Ridwan	5	3	7	K
Kevin wijaya kusuma	5	2	8	K
Nanda nur adli	7	7	1	V

Penyelesaian

1. Tentukan jumlah himpunan Visual, Audio dan Kinestetik. Pada kasus ini jumlah himpunan visual yakni sebanyak 95, audio sebanyak 20 dan kinestetik sebanyak 25.

2. Hitung nilai *entropy sample training* total

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

$$Entropy (total) = -\frac{95}{140} \log_2 \left(\frac{95}{140}\right) - \frac{20}{140} \log_2 \left(\frac{20}{140}\right) - \frac{25}{140} \log_2 \left(\frac{25}{140}\right) = 1,224$$

3. Hitung nilai *entropy* setiap atribut terhadap *entropy* total. Perhitungan *Entropy* dilakukan terhadap kelas atribut *Decision*.

a. Menghitung nilai *entropy* skor_visual ≤5 dan >5

Atribut skor_visual ≤5 dan >5 memiliki 2 buah jenis *value* yaitu ≤5 sebanyak 41 buah dan >5 sebanyak 99 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_visual \leq 5$)

$$= -\frac{6}{41} \log_2 \left(\frac{6}{41} \right) - \frac{16}{41} \log_2 \left(\frac{16}{41} \right) - \frac{19}{41} \log_2 \left(\frac{19}{41} \right) = 1,450$$

Entropy ($S, skor_visual > 5$)

$$= -\frac{89}{99} \log_2 \left(\frac{89}{99} \right) - \frac{4}{99} \log_2 \left(\frac{4}{99} \right) - \frac{6}{99} \log_2 \left(\frac{9}{99} \right) = 0,570$$

- b. Menghitung nilai *entropy* skor_visual ≤ 10 dan > 10

Atribut skor_visual ≤ 10 dan > 10 memiliki 2 buah jenis *value* yaitu ≤ 10 sebanyak 138 buah dan > 10 sebanyak 2. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_visual \leq 10$)

$$= -\frac{95}{138} \log_2 \left(\frac{95}{138} \right) - \frac{20}{138} \log_2 \left(\frac{20}{138} \right) - \frac{25}{138} \log_2 \left(\frac{25}{138} \right) = 1,234$$

$$Entropy (S, skor_visual > 10) = -\frac{2}{2} \log_2 \left(\frac{2}{2} \right) = 0$$

- c. Menghitung nilai *entropy* skor_visual ≤ 15 dan > 15

Atribut skor_visual ≤ 15 dan > 15 memiliki 2 buah jenis *value* yaitu ≤ 15 sebanyak 140 buah dan > 15 sebanyak 0 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_visual \leq 15$)

$$= -\frac{95}{140} \log_2 \left(\frac{95}{140} \right) - \frac{20}{140} \log_2 \left(\frac{20}{140} \right) - \frac{25}{140} \log_2 \left(\frac{25}{140} \right) = 1,224$$

$$Entropy (S, skor_visual > 15) = -\frac{0}{0} \log_2 \left(\frac{0}{0} \right) = 0$$

- d. Menghitung nilai *entropy* skor_audio ≤ 5 dan > 5

Atribut skor_audio ≤ 5 dan > 5 memiliki 2 buah jenis *value* yaitu ≤ 5 sebanyak 115 buah dan > 5 sebanyak 25 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_audio \leq 5$)

$$= -\frac{90}{115} \log_2 \left(\frac{90}{115} \right) - \frac{25}{115} \log_2 \left(\frac{25}{115} \right) = 0,755$$

Entropy ($S, skor_audio > 5$)

$$= -\frac{5}{25} \log_2 \left(\frac{5}{25} \right) - \frac{20}{25} \log_2 \left(\frac{20}{25} \right) = 0,722$$

- e. Menghitung nilai *entropy* skor_audio ≤ 10 dan > 10

Atribut skor_audio ≤ 10 dan > 10 memiliki 2 buah jenis *value* yaitu ≤ 10 sebanyak 139 buah dan > 10 sebanyak 1 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_audio \leq 10$)

$$= -\frac{95}{139} \log_2 \left(\frac{95}{139} \right) - \frac{19}{139} \log_2 \left(\frac{19}{139} \right) - \frac{25}{139} \log_2 \left(\frac{25}{139} \right) = 1,213$$

$$Entropy (S, skor_audio > 10) = -\frac{1}{1} \log_2 \left(\frac{1}{1} \right) = 0$$

- f. Menghitung nilai *entropy* skor_audio ≤ 15 dan > 15

Atribut skor_audio ≤ 15 dan > 15 memiliki 2 buah jenis *value* yaitu ≤ 15 sebanyak 140 buah dan > 15 sebanyak 0 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_audio \leq 15$)

$$= -\frac{95}{140} \log_2 \left(\frac{95}{140} \right) - \frac{19}{140} \log_2 \left(\frac{19}{140} \right) - \frac{25}{140} \log_2 \left(\frac{25}{140} \right) = 1,224$$

$$Entropy (S, jawaban_b > 15) = -\frac{0}{0} \log_2 \left(\frac{0}{0} \right) = 0$$

- g. Menghitung nilai *entropy* skor_kinestetik ≤ 5 dan > 5

Atribut skor_kinestetik ≤ 5 dan > 5 memiliki 2 buah jenis *value* yaitu ≤ 5 sebanyak 93 buah dan > 5 sebanyak 47 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_kinestetik \leq 5$)

$$= -\frac{72}{93} \log_2 \left(\frac{72}{93} \right) - \frac{20}{93} \log_2 \left(\frac{20}{93} \right) - \frac{1}{93} \log_2 \left(\frac{1}{93} \right) = 0,763$$

Entropy ($S, skor_kinestetik > 5$)

$$= -\frac{23}{47} \log_2 \left(\frac{23}{47} \right) - \frac{24}{47} \log_2 \left(\frac{24}{47} \right) = 1,000$$

Menghitung nilai *entropy* skor_kinestetik ≤ 10 dan > 10

Atribut skor_kinestetik ≤ 10 dan > 10 memiliki 2 buah jenis *value* yaitu ≤ 10 sebanyak 102 buah dan > 10 sebanyak 0 buah. Dengan demikian perhitungan nilai *entropy* untuk setiap *value* atribut diperoleh sebagai berikut :

Entropy ($S, skor_kinestetik \leq 10$)

$$= -\frac{95}{140} \log_2 \left(\frac{95}{140} \right) - \frac{20}{140} \log_2 \left(\frac{20}{140} \right) - \frac{25}{140} \log_2 \left(\frac{25}{140} \right) = 1,224$$

$$Entropy (S, jawaban_c > 10) = -\frac{0}{0} \log_2 \left(\frac{0}{0} \right) = 0$$

- h. Menghitung nilai *entropy* skor_kinestetik ≤ 15 dan > 15

Atribut skor_kinestetik ≤ 15 dan > 15 memiliki 2 buah jenis *value* yaitu ≤ 15 sebanyak 140 buah dan > 15 sebanyak 0 buah. Dengan demikian perhitungan nilai

entropy untuk setiap *value* atribut diperoleh sebagai berikut :

$$Entropy(S, jawaban_c \leq 15) = -\frac{95}{140} \log_2 \left(\frac{95}{140} \right) - \frac{20}{140} \log_2 \left(\frac{20}{140} \right) - \frac{25}{140} \log_2 \left(\frac{25}{140} \right) = 1,224$$

$$Entropy(S, jawaban_c > 15) = -\frac{0}{0} \log_2 \left(\frac{0}{0} \right) = 0$$

4. Menghitung nilai *information gain* tiap atribut

$$Information\ Gain(S, F_j) = Entropy(S) - \sum_{v_i \in V_{F_j}} \frac{|S_{v_i}|}{|S|} \cdot Entropy(S_{v_i})$$

$$Information\ Gain(skor_visual \leq 5\ dan > 5) = 1,224 - \frac{41}{140} \cdot 1,450 - \frac{99}{140} \cdot 0,570 = 0,397$$

$$Information\ Gain(skor_visual \leq 10\ dan > 10) = 1,224 - \frac{138}{140} \cdot 1,234 - \frac{2}{140} \cdot 0 = 0,008$$

$$Information\ Gain(skor_visual \leq 15\ dan > 15) = 1,224 - \frac{140}{140} \cdot 1,224 = 0$$

$$Information\ Gain(skor_audio \leq 5\ dan > 5) = 1,224 - \frac{115}{140} \cdot 0,755 - \frac{25}{140} \cdot 0,722 = 0,475$$

$$Information\ Gain(skor_audio \leq 10\ dan > 10) = 1,224 - \frac{139}{140} \cdot 1,213 - \frac{1}{140} \cdot 0 = 0,020$$

$$Information\ Gain(skor_audio \leq 15\ dan > 15) = 1,224 - \frac{140}{140} \cdot 1,224 = 0$$

$$Information\ Gain(skor_kinestetik \leq 5\ dan > 5) = 1,224 - \frac{93}{140} \cdot 0,763 - \frac{47}{140} \cdot 1 = 0,383$$

$$Information\ Gain(skor_kinestetik \leq 10\ dan > 10) = 1,224 - \frac{140}{140} \cdot 1,224 = 0$$

$$Information\ Gain(skor_kinestetik \leq 15\ dan > 15) = 1,224 - \frac{140}{140} \cdot 1,224 = 0$$

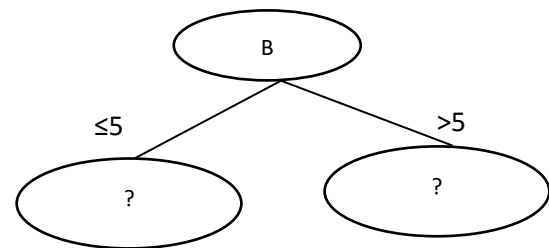
5. Menentukan *root tree*

Informasi *gain* masing-masing atribut hasil perhitungan pertama seperti di bawah ini :

Tabel 3.5 Information Gain

Atribut	Information Gain
Skor_visual ≤5 dan >5	0,397
Skor_visual ≤10 dan >10	0,008
Skor_visual ≤15 dan >15	0
Skor_audio ≤5 dan >5	0,475
Skor_audio ≤10 dan >10	0,020
Skor_audio ≤15 dan >15	0
Skor_kinestetik ≤5 dan >5	0,382
Skor_kinestetik ≤10 dan >10	0
Skor_kinestetik ≤15 dan >15	0

Berdasarkan perhitungan informasi *gain*, dapat dilihat bahwa nilai *gain* tertinggi dimiliki oleh atribut skor_audio ≤5 dan >5, sehingga dengan demikian skor_audio dijadikan sebagai root dari pohon keputusan yang akan dibentuk. Sedangkan simpul keputusan dilihat berdasarkan *value* atribut dari skor_audio, yakni jika *value* atribut memiliki class target berbeda, maka dilakukan perhitungan nilai *gain* informasi kembali terhadap atribut tertentu yang merujuk pada *value* atribut yang menjadi root pohon keputusan. Berikut merupakan ilustrasi terbentuknya root pohon keputusan dengan simpul keputusan awal menurut *value* atribut yang dimilikinya.



Gambar 3.1 Root Pohon Keputusan

- Ulangi langkah ketiga sampai kelima hingga setiap subset hanya memiliki masing-masing satu kelas.
- Setelah mendapatkan aturan *rule* dari perhitungan data *training* maka selanjutnya adalah melakukan uji *rule* menggunakan data *testing* sebanyak 20 seperti yang ditunjukkan oleh tabel:

Tabel 3.6 Data Testing

No	Nama	Visual	Audio	Kinestetik	Klasifikasi Awal	Rule	Hasil Klasifikasi
1	Muhamad Azizi	7	2	6	V	3	V
2	Fellia Marsella	11	4	0	V	3	V
3	Risa Nurlaela Safar	7	4	4	V	3	V

4	Miftah Alfawazudin	3	8	4	A	4	A
5	Dwi Rizki Gustiandini	5	7	3	A	4	A
6	Yudhistira	4	7	4	A	4	A
7	Sabrina Fauzia	6	2	7	K	3	V
8	Sandi Mohamad Tian	8	3	4	V	4	V
9	Julya Elena Hosiana	8	3	4	V	4	V
10	Titon Meisya Kresna	5	1	9	K	2	K
11	Erlangga Alfarizi F	5	9	1	A	4	A
12	Bintang Adhela Putra P	4	3	8	K	2	K
13	Mohamad Abdul B	9	3	3	V	3	V
14	Rahmi Oktavia Putri	8	2	5	V	3	V
15	Muhammad Rifky J	9	2	4	V	3	V
16	Alfiru Nur Alfian	9	1	5	V	3	V
17	Abdur Rouf	9	2	4	V	3	V
18	Taufik Hidayat M	7	4	4	V	3	V
19	Putri Helvyana	4	3	8	K	2	K
20	Wahyudin	4	3	8	K	2	K

Berdasarkan uji aturan pada table 4.33, selanjutnya adalah menghitung kinerja dari klasifikasi pada penelitian ini menggunakan metode *contusion matrix* dari hasil tersebut, sebagai berikut:

a. Akurasi :

$$\text{Akurasi} = \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{Jumlah prediksi yang dilakukan}} = \frac{19}{20} \times 100 = 95\%$$

a. Laju error :

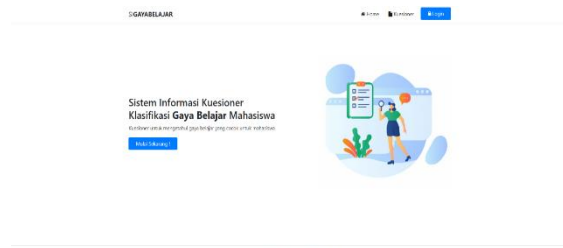
$$\text{Laju error} = \frac{\text{Jumlah data yang diprediksi secara salah}}{\text{Jumlah prediksi yang dilakukan}} = \frac{1}{20} \times 100 = 5\%$$

Hasil dari perhitungan *contusion matrix* dengan metode C4.5 mempunyai akurasi 95% dari jumlah data yang diprediksi secara benar yaitu 19 data dengan jumlah presdiksi yang dilakukan sebanyak 20 data, dan mempunyai laju error %5 dengan jumlah data yang diprediksi secara salah yaitu 1 data dengan jumlah presdiksi yang dilakukan sebanyak 20 data.

4 HASIL DAN PEMBAHASAN

4.2 Menu Utama

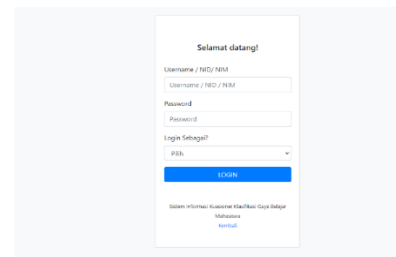
Menu utama merupakan tampilan utama pada aplikasi. Gambar menu utama ditunjukkan oleh gambar 4.1 berikut:



Gambar 4. 1 Tampilan Menu Utama

4.3 Login

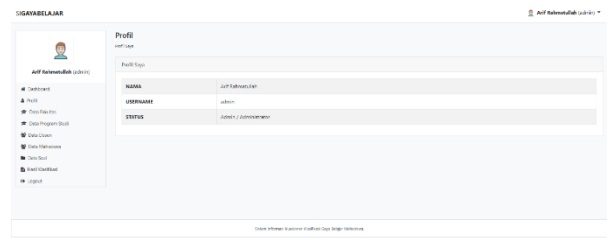
Login adalah tampilan untuk memasukan akun yang sudah didaftarkan sebelum masuk ke dalam aplikasi. Dibawah ini adalah gambar login yang ditunjukkan oleh gambar 4.2:



Gambar 4. 2 Login

4.4 Profil

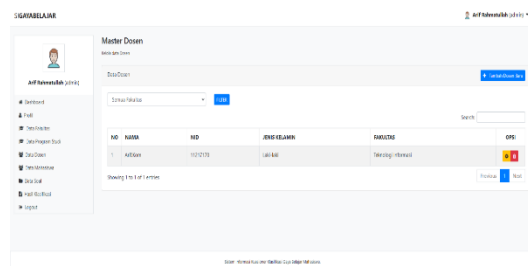
Profil adalah halaman yang menampilkan nama penguuna sebagai menu utama setelah berhasil login yang digambarkan oleh gambar 4.3:



Gambar 4. 3 Profil

4.4 Data Dosen

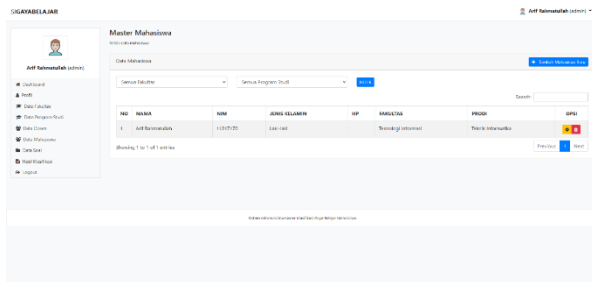
Siswa adalah halaman form untuk mengelola data dosen yang digambarkan oleh gambar4.4:



Gambar 4 4 Data Dosen

4.5 Data Mahasiswa

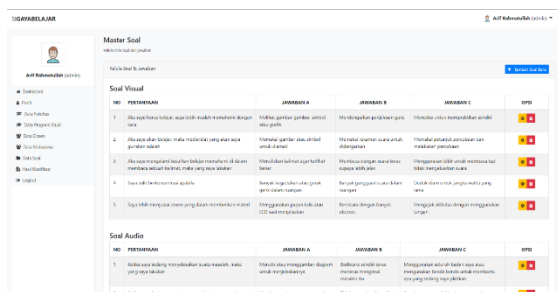
Siswa adalah halaman form untuk mengelola data mahasiswa yang digambarkan oleh gambar 4.5:



Gambar 4. 5 Data Mahasiswa

4.6 Data Soal

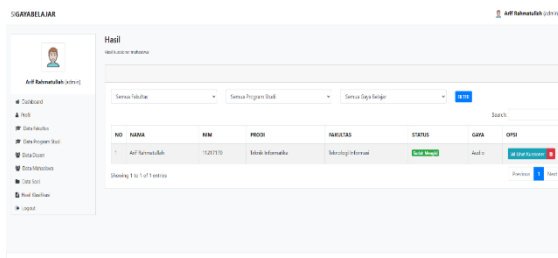
Data Soal adalah form untuk mengelola data soal yang di inputkan sebelum melakukan kuesioner yang digambarkan oleh gambar 4.6:



Gambar 4. 6 Data Soal

4.7 Hasil Klasifikasi

Hasil klasifikasi adalah halaman yang menunjukkan nama-nama siswa yang sudah diinput data dan klasifikasinya yang digambarkan oleh gambar 4.7:



Gambar 4.7 Hasil Klasifikasi

V. PUSTAKA

[1] Agnes Ika Kurniawati (2013). *Gaya belajar siswa kelas X dan XI IPA serta gaya mengajar guru di kelas tersebut dalam pembelajaran fisika di SMA*

Bhakti Karya Kaloran Kabupaten Temanggung Jawa Tengah. Skripsi thesis, Sanata Dharma University.

[2] Cholifah, W. N., Yulianingsih, & Sagita, S. M. (2018). Pengujian Black Box Testing Pada Aplikasi Action & Strategy Berbasis Android Dengan Teknologi Phonegap. *JURNAL STRING*, 3(2), 206-210.

[3] Madcoms. (2016). *Pemrograman PHP dan MySQL untuk Pemula*. Yogyakarta: Penerbit Andi.

[4] Mulyati (2015). Identifikasi Gaya Belajar Siswa Kelas V Sd Se Gugus 3 Kecamatan Pengasih Kabupaten Kulon Progo Tahun Ajaran 2014/2015. *Jurnal Pendidikan Guru Sekolah Dasar Edisi Tahun ke IV*.

[5] Munandar, T. A. (2019). *Bahan Ajar Data Mining Dengan R Edisi Revisi 3*. Serang.

[6] Panji Bimo Nugroho Setio (2020). Klasifikasi dengan Pohon Keputusan Berbasis Algoritme C4.5. *PRISMA*, Prosiding Seminar Nasional Matematika, Vol 3, 64-71.

[7] Raharjo, B. (2016). *Modul Pemrograman WEB (HTML, PHP, & MySQL)*. Bandung: Modula.

[8] Ratna Puspita Sari Putri, Indra Waspada (2018). Penerapan Algoritma C4.5 pada Aplikasi Prediksi Kelulusan Mahasiswa Prodi Informatika. *Khazanah informatika*, 4(1), 1 -7.

[9] Sari Rahayu Rahman, Firman (2019). Identifikasi Gaya Belajar Mahasiswa Pendidikan Biologi Universitas Sulawesi Barat. *Jambura Edu Biosfer journal*, 1(1), 1 – 5.

[10]. Wawan Wahyudin (2016). *GAYA BELAJAR MAHASISWA*. *Al Qalam*, 33(1), 105-120.

[11] Wenefrida Tulit Ina (2013). *Klasifikasi Data Rekam Medis Berdasarkan Kode Penyakit Internasional Menggunakan Algoritma C4.5*. *Jurnal Media Elektro*, 1(3), 105-110

[12] Widiyanto, W. W. (2018). *Analisa Metodologi Perbandingan Sistem dengan Perbandingan Model Perangkat Lunak Sistem Informasi Kepegawaian Menggunakan Waterfall Development Model, Model Prototype, dan Model Rapid Application Development (RAD)*. *Jurnal INFORMATIKA Politeknik Indonusa Surakarta*, 4(1), 34-40.

[13] Yusri Wahyuni (2017). *Identifikasi Gaya Belajar (Visual, Auditorial, Kinestetik) Mahasiswa Pendidikan Matematika Universitas Bung Hatta*. *JPPM (Jurnal Penelitian dan Pembelajaran Matematika)*, 10(2), 128-132