p-ISSN : 2406-7733
e-ISSN : 2597-9922

IMPLEMENTASI ALGORITMA C4.5 UNTUK MEMPREDIKSI MINAT *PROGRAMMING* MAHASISWA BERBASIS WEB

Asswendo Suryahadi[1], Rachmat Aulia[2], Fera Damayanti[3]
Teknik Informatika, Teknik dan Komputer, Universitas Harapan Medan
Jl. H.M. Joni Nomor 70 C Medan, Indonesia
E-mail: *asswendo@gmail.com[1], jackm4t@gmail.com[2], feradamayantii@gmail.com[3]

Abstrak - Program studi teknik informatika adalah sebuah studi yang berfokus pada pengembangan dan penerapan teknologi informasi dalam berbagai bidang, diantaranya bidang *programming*, pengembangan perangkat lunak, jaringan komputer, robotika, pengolahan data, kecerdasan buatan dan manajemen proyek IT. Dikarenakan banyaknya bidang pada program studi teknik informatika yang direpresentasikan menjadi sebuah mata kuliah wajib, mahasiswa menjadi sulit untuk menentukan minat mereka pada bidang yang ingin mahasiswa fokuskan, khususnya pada bidang *programming*. Hal itu menyebabkan banyak mahasiswa tidak dapat menguasai bidang *programming* mau itu mahasiwa yang masih aktif ataupun mahasiwa yang sudah menyelesaikan pendidikannya. Pada penelitian ini dirancang dan dibangun sebuah aplikasi prediksi berbasis *website* yang diimplementasikan sebuah metode *data mining*, yaitu algoritma C4.5. Algoritma C4.5 merupakan salah satu metode *data mining* yang bersifat klasifikasi dan prediksi, dimana algoritma tersebut menghasilkan *rules* yang bersifat prediktif yang juga dapat digunakan untuk memprediksi minat seseorang pada bidang *programming*. Data-data yang digunakan adalah data yang didapatkan dari hasil kuisioner sebanyak 62 data, dimana 50 data akan dipakai sebagai *data training* dan 12 data sebagai *data testing*. Hasil pada penelitian ini menunjukan bahwa, dari 50 *data training* yang digunakan, aplikasi berhasil mendapatkan 9 *rule* yang akan digunakan dan dari 12 *data testing* yang digunakan pada pengujian, mendapatkan hasil akurasi sebesar 83% dan laju *error* sebesar 17%. Dengan demikian, dapat disimpulkan bahwa aplikasi prediksi dapat membantu mahasiswa dalam menentukan minat mahasiswa pada bidang *programming*.

Kata Kunci: *C4.5, Data mining, Mahasiswa, Programming, Website,*

I. PENDAHULUAN

Program studi Teknik Informatika merupakan program studi yang sangat populer dan memiliki prospek karir yang sangat bagus dikarenakan tingginya kebutuhan teknologi informasi saat ini (Primayasa et al., 2020). Di dalam program studi ini, terdapat berbagai bidang studi, salah satunya adalah *programming*. Namun, pada kenyataannya, masih banyak mahasiswa maupun alumni yang belum menguasai bidang *programming*. Salah satu faktor penyebab utamanya adalah mahasiswa masih kesulitan untuk menentukan minat pada bidang yang ingin difokuskan saat kuliah, karena banyaknya bidang yang dijadikan mata kuliah wajib.

Oleh karena itu, diperlukan sebuah cara untuk membantu mahasiswa memprediksi minat bidang yang akan difokuskan, salah satunya dengan memanfaatkan teknik klasifikasi dalam *data mining*. Penelitian ini menggunakan algoritma C4.5, yang telah terbukti keandalannya dalam mengklasifikasikan data. Pemilihan algoritma ini merujuk pada penelitian sebelumnya yang ditulis oleh (Putri et al., 2021), yang menyatakan bahwa algoritma C4.5 memiliki kelebihan dalam membuat sebuah *Decision Tree* (pohon keputusan) dari *data training* untuk menghindari permasalahan klasifikasi dengan kriteria *node* yang efisien. Selain itu, penelitian lain yang dilakukan oleh Nas (2021) tentang prediksi minat calon mahasiswa menggunakan algoritma C4.5 membuktikan keunggulan metode ini dengan pencapaian tingkat akurasi hingga 100% dari 120 *data testing*.

Berdasarkan permasalahan dan solusi dari penelitian terdahulu tersebut, penelitian ini bertujuan untuk merancang dan membangun sistem prediksi berbasis *web* yang mengimplementasikan Algoritma C4.5. Sistem ini dibangun untuk memprediksi minat mahasiswa pada bidang *programming* dengan akurat. Diharapkan aplikasi ini dapat memberikan kemudahan bagi mahasiswa untuk mengetahui minat mereka, serta membantu pihak universitas dalam memetakan seberapa besar tingkat peminatan mahasiswanya pada bidang *programming*.

II. LANDASAN TEORI

Perguruan Tinggi

Perguruan tinggi adalah lembaga pendidikan tinggi yang menyelenggarakan berbagai program akademik lanjutan setelah pendidikan menengah (Karim, 2020). Peran utamanya adalah memajukan individu berkualitas tinggi guna memenuhi tuntutan perkembangan dan pembangunan masyarakat (Lian, 2019).

Mahasiswa

Mahasiswa adalah individu yang sedang menempuh pendidikan di tingkat perguruan tinggi (Purba & Nafeesa, 2022). Tugas utama mahasiswa

p-ISSN : 2406-7733
e-ISSN : 2597-9922

berfokus pada proses belajar, seperti menyelesaikan tugas, membaca materi, berdiskusi, dan menghadiri seminar (Pasaribu, 2022).

Minat

Minat dapat dijelaskan sebagai kecenderungan dan semangat yang kuat terhadap suatu hal (Nas, 2021). Minat juga merupakan perasaan preferensi dan ketertarikan terhadap sesuatu tanpa adanya paksaan (Matondang, 2018). Melalui minat tersebut, individu akan berupaya serta termotivasi untuk mencapai tujuan yang diinginkannya (Achru, 2019).

Programming

Programming adalah sebuah keterampilan dalam merancang dan membuat program komputer melalui penggunaan algoritma dan bahasa pemrograman tertentu (Naufal, 2018). Bahasa pemrograman ini kemudian diklasifikasikan dari tingkat rendah (bahasa mesin) hingga bahasa tingkat tinggi (Herwanto, 2019).

Prediksi

Prediksi merupakan proses memperkirakan atau memproyeksikan kondisi di masa depan dengan menguji situasi dan data dari masa lalu (Amalia, 2018). Prediksi juga dapat diartikan sebagai kemampuan untuk mengidentifikasi pola khusus dalam *dataset* yang diterapkan untuk memproyeksikan nilai yang belum diketahui (Damuri et al., 2021).

Data Mining

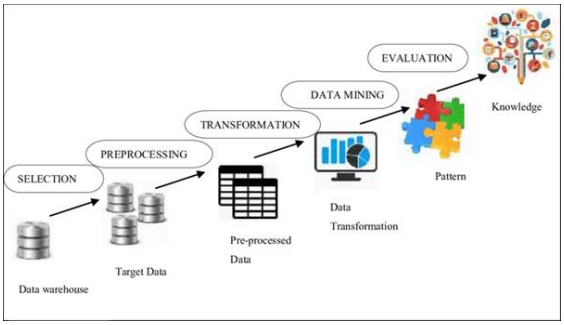
Data mining adalah proses ekstraksi informasi yang berharga dari dalam kumpulan *dataset* yang besar (Nas, 2021). Teknik ini memanfaatkan metode statistik, matematika, dan analisis pola untuk mendapatkan wawasan serta pengetahuan baru (Damuri et al., 2021).

Teknik *Data Mining*

Beberapa teknik umum yang terdapat pada *data mining* meliputi *Clustering*, *Regression*, *Association Rule Mining*, dan *Forecasting* (Zai, 2022). Selain itu, terdapat teknik *Classification* (Klasifikasi) yang secara khusus digunakan untuk meramalkan kelas dari data berdasarkan pola pelatihannya (Pramudito, 2022).

***Knowledge Discovery In Database* (KDD)**

KDD adalah tahapan sistematis dalam menggali informasi pada *data mining* (Nas, 2021). Tahapan dari KDD dapat dilihat pada Gambar 2.1 Tahapan dalam KDD yang meliputi proses: *Data selection*, *Pre-Processing & Cleaning*, *Transformation*, *Data Mining*, dan *Evaluation/Interpretation.*



Gambar 2.1 Tahapan Dalam KDD

Teknik *Classfication* (Klasifikasi)

Klasifikasi adalah suatu pendekatan penemuan model atau fungsi untuk menggambarkan dan membedakan kelas data, yang bertujuan untuk meramalkan kelas dari objek yang belum memiliki label (Nas, 2021).

Teknik *Forecasting* (Peramalan)

Teknik *forecasting* berfungsi memprediksi nilai-nilai di masa depan berdasarkan pola-pola atau tren di dalam sejumlah besar data historis, seperti prediksi permintaan pasar (Zai, 2022).

Decision Tree

Decision tree (pohon keputusan) adalah struktur grafik alir menyerupai pohon yang memecah informasi berskala besar menjadi aturan-aturan keputusan atau *rules* (Nas, 2021). Metode pohon keputusan ini memiliki kekuatan yang sangat baik untuk menyajikan aturan-aturan dalam klasifikasi dan prediksi (Putri et al., 2021).

Algoritma C4.5

Algoritma C4.5 adalah algoritma *data mining* yang andal dan mudah diartikan untuk membangun *decision tree* (Husna, 2021). Proses pembentukan pohon keputusan melibatkan serangkaian langkah-langkah, yaitu dimulai dengan menyiapkan data *training* yang telah dikelompokkan ke dalam kategori tertentu, seperti contoh yang ditunjukkan pada Tabel 1.

Tabel 1. Data Cuaca Dan Keputusan Bermain atau Tidak

Cuaca $x1$	Temperatur $x2$	Kelembaban $x3$	Angin $x4$	Main atau Tidak s
Cerah	Panas	Tinggi	Kecil	Tidak
Cerah	Panas	Tinggi	Besar	Tidak
Hujan	Sedang	Tinggi	Kecil	Tidak
Mendung	Dingin	Normal	Besar	Ya
Cerah	Sedang	Tinggi	Kecil	Tidak
Cerah	Dingin	Normal	Kecil	Ya
Hujan	Sedang	Normal	Kecil	Ya
Cerah	Sedang	Nirmal	Besar	Ya
Mendung	Sedang	Tinggi	Besar	Ya
Mendung	Panas	Normal	Kecil	Ya
Hujan	Sedang	Tinggi	Besar	Tidak

p-ISSN : 2406-7733
e-ISSN : 2597-9922

Cuaca $x1$	Temperatur $x2$	Kelembaban $x3$	Angin $x4$	Main atau Tidak s
Mendung	Panas	Tinggi	Kecil	Ya
Hujan	Sedang	Tinggi	Kecil	Ya
Hujan	Dingin	Normal	Kecil	Ya

Langkah berikutnya adalah menentukan simpul akar dari pohon dengan menghitung entropi total dan entropi dari setiap nilai atribut (Husna, 2021). Entropi mewakili ukuran informasi yang mengindikasikan seberapa tidak pasti (ketidakmurnian) atribut dari sekumpulan data objek. Persamaan untuk menentukan nilai entropi dapat dilihat pada Persamaan 1.

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

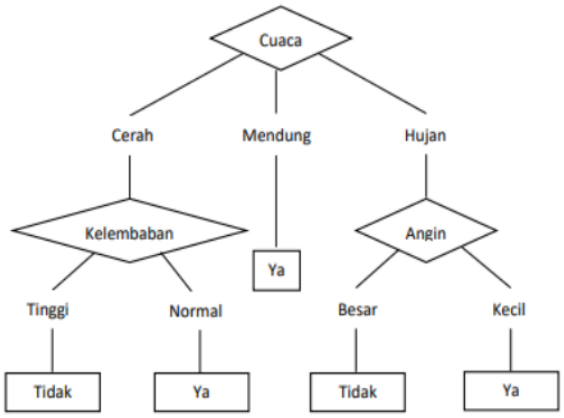
Keterangan:
s = himpunan kasus
n = jumlah partisi
p_i = proporsi kelas i dalam data yang diproses pada node s. Nilai p_i didapat dengan membagikan jumlah baris data sebagai label kelas dengan jumlah baris semua data.

Setelah menghitung entropi, langkah selanjutnya adalah menghitung gain. Gain merupakan ukuran sejauh mana suatu atribut mempengaruhi keputusan atau efektivitas suatu variabel dalam mengklasifikasikan data. Perhitungan nilai gain dilakukan menggunakan Persamaan 2.

$$Gain(s, A) = Entropy(s) - \sum^{n} \frac{|s_i|}{|s|} \ . \ Entropy(s_i) \qquad (2)$$

Keterangan:
S = himpunan kasus
A = nilai yang mungkin dari atribut A
n = jumlah partisi atribut A
$|S_i|$ = subset dari S di mana A mempunyai nilai i
$|S|$ = himpunan kasus

Dari nilai *entropy* dan *gain* yang telah didapatkan pada perhitungan, maka akan terbentuk sebuah pohon keputusan yang dicontohkan pada Gambar 2.



Gambar 2. *Decision Tree* untuk data cuaca

Website

MySQL merupakan Sistem Manajemen Basis Data Relasional (RDBMS) yang cepat dan *user-friendly* (Sulistiati et al., 2020). Aplikasi ini sangat populer di kalangan *programmer web* karena berbasis *open source* sehingga dapat digunakan secara luas (Sonny, 2021).

Database

Basis data adalah koleksi data yang terhubung dan terintegrasi secara logis guna memfasilitasi proses pencarian data yang diperlukan oleh suatu entitas (Sulistiati et al., 2020).

MySQL

MySQL merupakan Sistem Manajemen Basis Data Relasional (RDBMS) yang cepat dan *user-friendly* (Sulistiati et al., 2020). Aplikasi ini sangat populer di kalangan *programmer web* karena berbasis *open source* sehingga dapat digunakan secara luas (Sonny, 2021).

PHP

PHP adalah bahasa pemrograman skrip yang dapat diintegrasikan dengan HTML untuk mengembangkan situs web dinamis (HENDI, 2021). Bahasa ini sangat diminati oleh pengembang *web* dan dapat dimanfaatkan secara gratis (Candra, 2021).

Laravel

Laravel adalah *framework open source* berbasis PHP dengan arsitektur MVC (*Model-View-Controller*) yang berkinerja lebih cepat, stabil, memiliki tingkat keamanan tinggi, serta dilengkapi fitur *library* komprehensif (Widhi et al., 2019).

Flowchart

Flowchart (diagram alir) adalah representasi visual dari alur dan urutan langkah-langkah dalam sebuah program menggunakan grafik (RS, 2022). Diagram ini menggunakan simbol-simbol yang mewakili berbagai proses dan dihubungkan oleh

p-ISSN : 2406-7733
e-ISSN : 2597-9922

garis penghubung (Rosaly & Prasetyo, 2019). Simbol-simbol yang digunakan pada *flowchart* yang meliputi simbol seperti *Terminal*, *Input/Output*, *Processing*, *Decision*, dan lainnya.

III. ANALISIS DAN PERANCANGAN SISTEM

Analisis Masalah

Banyak mahasiswa program studi Teknik Informatika yang kesulitan menentukan minat bidang yang ingin difokuskan, khususnya pada bidang *programming*. Hal ini mengakibatkan banyak mahasiswa tidak mampu melakukan *programming* meskipun telah menjadi mata kuliah wajib maupun peminatan. Oleh karena itu, diusulkan sebuah sistem berbasis *web* yang mengimplementasikan metode klasifikasi *data mining* menggunakan Algoritma C4.5 untuk memprediksi minat *programming* mahasiswa.

Analisis Sistem

1. Analisis Kebutuhan Fungsional

Kebutuhan fungsional yang terdapat pada sistem meliputi:

1) Sistem dapat melakukan *import* data *training* dan data *testing* lalu menyimpannya ke dalam *database*.
2) Sistem dapat melakukan proses *mining* menggunakan algoritma C4.5 untuk membangun pohon keputusan.
3) Sistem dapat menghasilkan dan menampilkan *rule-rule* pembentukan *decision tree*.
4) Sistem dapat memprediksi minat *programming* mahasiswa berdasarkan kriteria yang diinputkan.

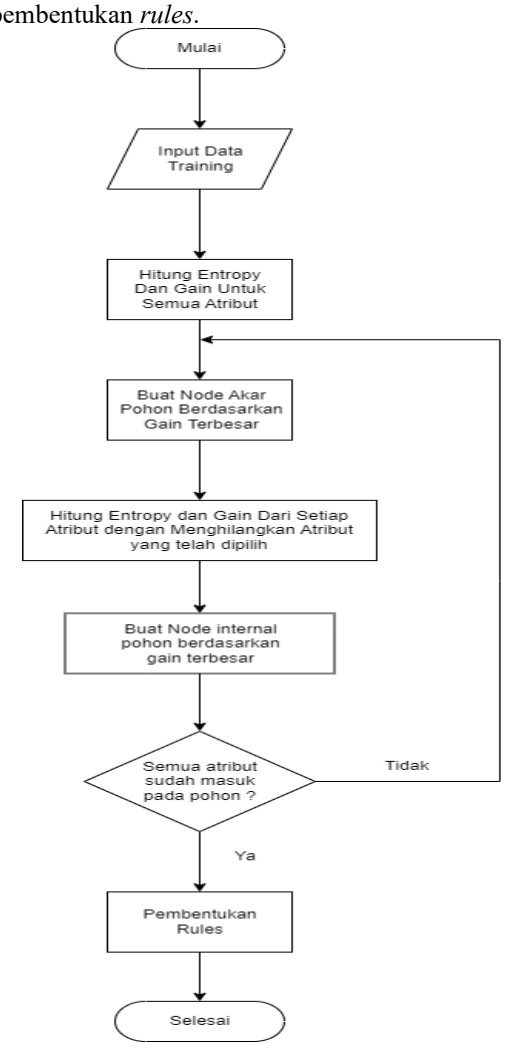
2. Analisis Kebutuhan Non-Fungsional

Kebutuhan perangkat keras (*hardware*) yang digunakan adalah Laptop (*Processor* Apple M1, RAM 8 GB, SSD 256 GB). Kebutuhan perangkat lunak (*software*) meliputi MacOS Sonoma, Visual Studio Code, PHP, *Framework* Laravel, Google Chrome, MySQL, XAMPP, dan Composer.

3. Analisis Algoritma C4.5

Data didapatkan melalui observasi, wawancara, dan penyebaran kuesioner kepada mahasiswa Teknik Informatika. Atribut yang ditentukan meliputi: *Work* (berkuliah sambil bekerja atau tidak), *Difficulty* (tingkat kesulitan dipelajari), *Prospek* (prospek karir), *Knowing programming* (pengetahuan awal), dan *Can programming* (kemampuan *coding*).

Langkah-langkah implementasi algoritma C4.5 direpresentasikan pada Gambar 3.1 yang mencakup proses input data, menghitung *Entropy* dan *Gain*, membuat *node* akar, mempartisi atribut, hingga pembentukan *rules*.



Gambar 3. *Flowchart* Alur Algoritma C4.5

Dataset kuesioner digunakan sebagai data *training* yang sampelnya disajikan pada Tabel 2.

Tabel 2. Data Training

No	*Name*	*Work*	*Difficulty*	*Prospect*	*Knowing Programming*	*Can Programming*	Kelas
1	Valdi Valdhano Oka	*work*	Tinggi	Tinggi	*Yes*	*Yes*	Minat
2	Muhammad Raihan	*not work*	Tinggi	Sedang	*Yes*	*Yes*	Minat
3	Mhd marquis	*work*	Tinggi	Tinggi	*No*	*No*	Minat
4	Mico Nicola	*work*	Sedang	Tinggi	*Yes*	*Yes*	Minat
5	Muhammad Al Ikhsan	*work*	Tinggi	Tinggi	*Yes*	*Yes*	Tidak Minat

p-ISSN : 2406-7733
e-ISSN : 2597-9922

No	Name	Work	Difficulty	Prospect	Knowing Programming	Can Programming	Kelas
6	Musliadi	not work	Sedang	Sedang	Yes	No	Minat
7	Syahendra surya	work	Rendah	Tinggi	Yes	No	Minat
8	fahri ardhana	not work	Sedang	Tinggi	Yes	Yes	Minat
9	Budi Sahputra	not work	Sedang	Tinggi	Yes	Yes	Minat
10	Daniel	not work	Sedang	Sedang	Yes	No	Minat
11	Halimansyah Putra, S. Kom	not work	Sedang	Tinggi	Yes	Yes	Minat
12	RAYHAN ALFARID SAYUTI	work	Sedang	Tinggi	No	No	Tidak Minat
13	Mhd Fadhil	not work	Sedang	Tinggi	Yes	No	Minat
14	M.Ilham Ramadhan	work	Sedang	Tinggi	Yes	Yes	Minat
15	Indra Abadi Dalimunthe	not work	Sedang	Tinggi	Yes	No	Minat
16	Nizmaldy Astoro Witjati	not work	Sedang	Tinggi	Yes	Yes	Minat
17	Muhammad Nawawi Nasution	not work	Tinggi	Sedang	Yes	Yes	Minat
18	Arya Penangsang	not work	Sedang	Tinggi	Yes	Yes	Minat
19	Aldi Ardhana	not work	Sedang	Tinggi	Yes	Yes	Minat
20	Joko Kesuma	not work	Sedang	Sedang	Yes	Yes	Minat
21	Nurhayaty Br Berutu	not work	Tinggi	Tinggi	Yes	No	Minat
22	Wafiq Fadila	work	Tinggi	Tinggi	Yes	No	Minat
23	Anugrah Wira Perdana Siregar	not work	Sedang	Sedang	Yes	No	Minat
24	Tesalonika Br Barus	not work	Tinggi	Tinggi	Yes	No	Tidak Minat
25	Amira Fahira	not work	Tinggi	Tinggi	Yes	Yes	Tidak Minat
26	Febi alfiani putri harahap	not work	Sedang	Sedang	Yes	Yes	Minat
27	Lily Putri Utami	not work	Sedang	Tinggi	Yes	Yes	Minat
28	Arya Dwi Wahyuda	not work	Sedang	Tinggi	Yes	Yes	Minat
29	Muhammad Yasir Lubis	not work	Tinggi	Tinggi	Yes	Yes	Minat
30	Fahri Akbar	work	Tinggi	Tinggi	Yes	Yes	Tidak Minat

No	Name	Work	Difficulty	Prospect	Knowing Programming	Can Programming	Kelas
31	Algazali Raihan	not work	Sedang	Rendah	Yes	No	Tidak Minat
32	Mariana Halawa	work	Sedang	Sedang	Yes	Yes	Minat
33	Rizky	work	Sedang	Tinggi	Yes	Yes	Minat
34	Akbar S	work	Tinggi	Tinggi	Yes	No	Tidak Minat
35	Ade feruzi	not work	Rendah	Rendah	No	No	Tidak Minat
36	Muhammad arief	not work	Rendah	Sedang	No	No	Minat
37	Laila Tusyfa barus	work	Sedang	Tinggi	Yes	Yes	Minat
38	Eza Azhari Chaniago	not work	Rendah	Tinggi	Yes	Yes	Minat
39	Taufik Firmansyah	work	Tinggi	Tinggi	Yes	Yes	Minat
40	Alif Fiqri Harahap	not work	Tinggi	Tinggi	Yes	No	Minat
41	Dian Alfa Ridho	work	Sedang	Tinggi	Yes	Yes	Minat
42	Febriansya	not work	Tinggi	Tinggi	Yes	No	Minat
43	Muhammad Iqbal	not work	Sedang	Tinggi	Yes	Yes	Minat
44	Muhammad Abdillah	not work	Sedang	Tinggi	Yes	No	Minat
45	Feza Rahman	not work	Tinggi	Tinggi	Yes	No	Minat
46	Azhar badri	work	Tinggi	Tinggi	Yes	Yes	Minat
47	Muhammad adnan	work	Tinggi	Tinggi	Yes	Yes	Minat
48	Muhammad Agri Febriansyah	work	Sedang	Sedang	Yes	No	Minat
49	Lastri Pratiwi	not work	Sedang	Sedang	No	No	Minat
50	Khairul sani	not work	Tinggi	Tinggi	Yes	Yes	Minat
51	ilham rizki fadhillah	work	Sedang	Sedang	Yes	Yes	Minat
52	Mhd.Hudan Alif Muhajir	work	Tinggi	Tinggi	Yes	No	Minat
53	Iswanto	work	Tinggi	Rendah	Yes	No	Minat
54	M Ichsan	work	Tinggi	Tinggi	Yes	Yes	Minat
55	Haikal	work	Sedang	Tinggi	Yes	Yes	Tidak Minat
56	Liandi Haikal	work	Sedang	Tinggi	Yes	Yes	Minat
57	Muthia Sabrina Asry	not work	Sedang	Tinggi	Yes	No	Minat
58	Lady Syahri	not work	Tinggi	Tinggi	Yes	No	Minat
59	Nurul Fida Fazria	work	Tinggi	Tinggi	Yes	Yes	Minat
60	Chita Asry	work	Sedang	Sedang	Yes	No	Tidak Minat

p-ISSN : 2406-7733
e-ISSN : 2597-9922

No	*Name*	*Work*	*Difficulty*	*Prospect*	*Knowing Programming*	*Can Programming*	Kelas
61	Sai Prashanth	*work*	Sedang	Tinggi	*Yes*	*Yes*	Minat
62	Wahyu prayuda	*not work*	Sedang	Sedang	*Yes*	*Yes*	Minat
63	Rizky dwi putra	*work*	Tinggi	Tinggi	*Yes*	*Yes*	Minat
64	Rizki purnama siburian	*not work*	Sedang	Sedang	*Yes*	*Yes*	Minat
65	Muhamad rizky amra	*not work*	Sedang	Rendah	*No*	*No*	Tidak Minat
64	Raihan abdilah	*not work*	Tinggi	Tinggi	*Yes*	*Yes*	Minat
67	RiKy Budi Ramdhani	*not work*	Sedang	Tinggi	*Yes*	*Yes*	Minat
68	Andrian Rajab Harahap	*not work*	Tinggi	Tinggi	*Yes*	*No*	Minat
69	Harun Al Rasyid	*not work*	Tinggi	Tinggi	*Yes*	*No*	Minat
70	Nanda Hardiansyah	*not work*	Tinggi	Tinggi	*Yes*	*Yes*	Minat
71	Ayu Anjelli Br Berutu	*not work*	Sedang	Tinggi	*Yes*	*Yes*	Minat
72	Yurika Yulistina Br Karo	*not work*	Tinggi	Tinggi	*Yes*	*No*	Minat

Tahapan klasifikasi iterasi Algoritma C4.5 dilakukan dengan persamaan nilai *Entropy* dan *Gain* sebagai berikut:

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

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

Iterasi 1

Menghitung *entropy* total keseluruhan kasus. Entropy (Total) $= \left(-\frac{61}{72} * \log_2\left(\frac{61}{72}\right)\right) + \left(-\frac{11}{72} * \log_2\left(\frac{11}{72}\right)\right) = 0{,}616748$ Menghitung *Gain* setiap atribut. Contoh *Gain* atribut *Knowing Programming*: Gain (Total, Knowing Programming) $= 0{,}680077 - \left(\frac{3}{50} * 0\right) + \left(\frac{47}{50} * 0{,}550984\right) = 0{,}162152$

Tabel 3. Hasil perhitungan Iterasi 1

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Total Data		50	41	9	0,680077	
Work	*Work*	19	15	4	0,742488	0,002752
	Not Work	31	26	5	0,637387	
Difficulty	Rendah	4	3	1	0,811278	0,027455
	Sedang	27	24	3	0,503258	
	Tinggi	19	14	5	0,831474	
Prospek	Rendah	2	0	2	0	0,110189
	Sedang	11	10	1	0,439497	
	Tinggi	37	31	6	0,639457	
Knowing Programming	*No*	3	0	3	0	0,162152
	Yes	47	41	6	0,550984	
Can Programming	*No*	20	14	6	0,881291	0,046163
	Yes	30	27	3	0,468996	

Atribut *Knowing Programming* terpilih sebagai *root* (Gain terbesar 0,162152). Sub-atribut "No" mendefinisikan "Tidak Minat", sehingga sub-atribut "Yes" yang diekspansi.

Iterasi 2

Tabel 4. Hasil perhitungan Iterasi 2

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Knowing programming-Yes		47	41	6	0,550984	
Work	*Work*	18	15	3	0,650022	0,005973
	Not Work	29	26	3	0,479832	
Difficulty	Rendah	3	3	0	0	0,085978
	Sedang	25	24	1	0,242292	
	Tinggi	19	14	5	0,831474	
Prospek	Rendah	1	0	1	0	0,105717
	Sedang	10	10	0	0	
	Tinggi	36	31	5	0,581321	
	No	17	14	3	0,672295	0,008455

p-ISSN : 2406-7733
e-ISSN : 2597-9922

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Can Programming	*Yes*	30	27	3	0,468996	

Atribut *Prospek* terpilih (Gain 0,105717). Sub-atribut "Tinggi" diekspansi karena sub-atribut lainnya sudah menghasilkan keputusan.

Iterasi 3

Tabel 5. Hasil perhitungan Iterasi 3

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Knowing programming - *Yes* Prospek - Tinggi		36	31	5	0,581321	
Work	*Work*	15	12	3	0,721928	0,01585
	Not Work	21	19	2	0,453716	
Difficulty	Rendah	2	2	0	0	0,168608
	Sedang	17	17	0	0	
	Tinggi	17	12	5	0,873981	
Can Programming	*No*	17	14	3	0,672295	0,008455
	Yes	24	21	3	0,543564	

Atribut *Difficulty* terpilih (*Gain* 0,168608), di mana sub-atribut "Tinggi" diekspansi untuk iterasi berikutnya.

Iterasi 4

Tabel 6. Hasil perhitungan Iterasi 4

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Knowing programming - *Yes* Prospek – Tinggi *Difficulty* - Tinggi		17	12	5	0,873981	
Work	*Work*	9	6	3	0,918296	0,006046
	Not Work	8	6	2	0,811278	
Can Programming	*No*	8	6	2	0,811278	0,006046
	Yes	9	6	3	0,918296	

Karena *Work* dan *Can Programming* memiliki *Gain* yang sama, atribut *Work* (0,006046) dipilih sebagai *root* perhitungan berikutnya. Sub-atribut "*Work*" dan "*Not Work*" keduanya diekspansi.

Iterasi 5

Tabel 7. Hasil perhitungan Iterasi 5 sub-atribut *work*

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Knowing programming - *Yes* Prospek – Tinggi *Difficulty* – Tinggi *Work* - *work*		9	6	3	0,918296	
Can Programming	*No*	3	2	1	0,918296	0
	Yes	6	4	2	0,918296	

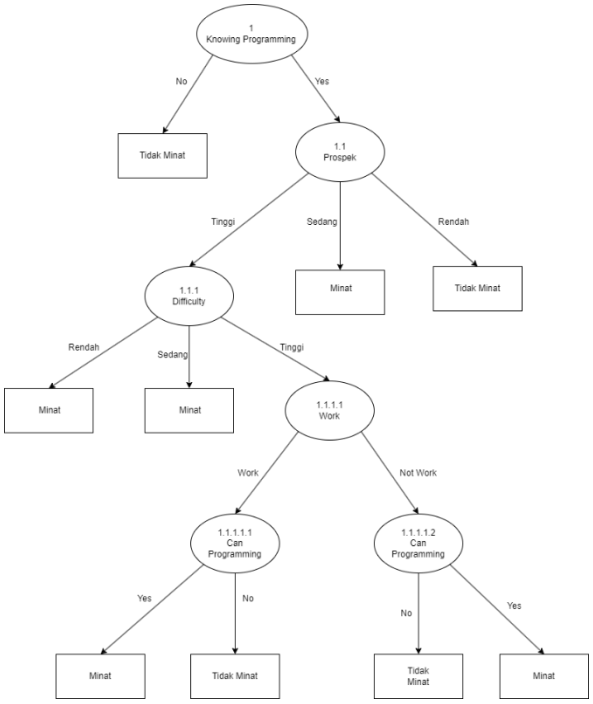
Tabel 8. Hasil perhitungan Iterasi 5 sub-atribut *not work*

Atribut		Jumlah	Minat	Tidak Minat	*Entropy*	*Gain*
Knowing programming – *Yes* Prospek – Tinggi *Difficulty* – Tinggi *Work* – *not work*		8	6	2	0,811278	
Can Programming	*No*	5	4	1	0,721928	0,015712
	Yes	3	2	1	0,918296	

Pada tahap ini, sub-atribut "*No*" (*Can Programming*) didefinisikan sebagai "Tidak Minat", dan "*Yes*" didefinisikan sebagai "Minat".

Hasil dari partisi yang telah dilakukan divisualisasikan pada Gambar 4 Hasil *Decision Tree* di bawah ini. Dari pohon keputusan, didapatkan 9 *rules* yaitu:

p-ISSN : 2406-7733
e-ISSN : 2597-9922



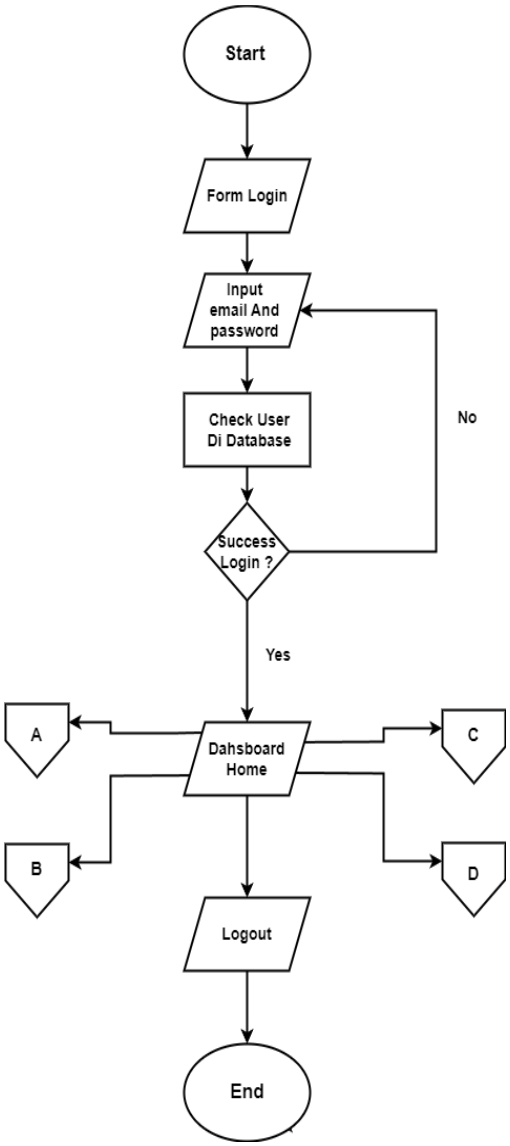
Gambar 4. Hasil *Decision Tree*

1. IF (knowing programming = No) THEN Label = Tidak Minat
2. IF (knowing_programming = Yes) AND (prospek = Rendah) THEN Label = Tidak Minat
3. IF (knowing_programming = Yes) AND (prospek = Sedang) THEN Label = Minat
4. IF (knowing_programming = Yes) AND (prospek = Tinggi) AND (difficulty = Rendah) THEN Label = Minat
5. IF (knowing_programming = Yes) AND (prospek = Tinggi) AND (difficulty = Sedang) THEN Label = Minat
6. IF (knowing_programming = Yes) AND (prospek = Tinggi) AND (difficulty = Tinggi) AND (work = work) AND (can_programming = No) THEN Label = Tidak Minat
7. IF (knowing_programming = Yes) AND (prospek = Tinggi) AND (difficulty = Tinggi) AND (work = work) AND (can_programming = Yes) THEN Label = Minat
8. IF (knowing_programming = Yes) AND (prospek = Tinggi) AND (difficulty = Tinggi) AND (work = not work) AND (can_programming = No) THEN Label = Tidak Minat
9. IF (knowing_programming = Yes) AND (prospek = Tinggi) AND (difficulty = Tinggi) AND (work = not work) AND (can_programming = Yes) THEN Label = Minat

Perancangan Sistem

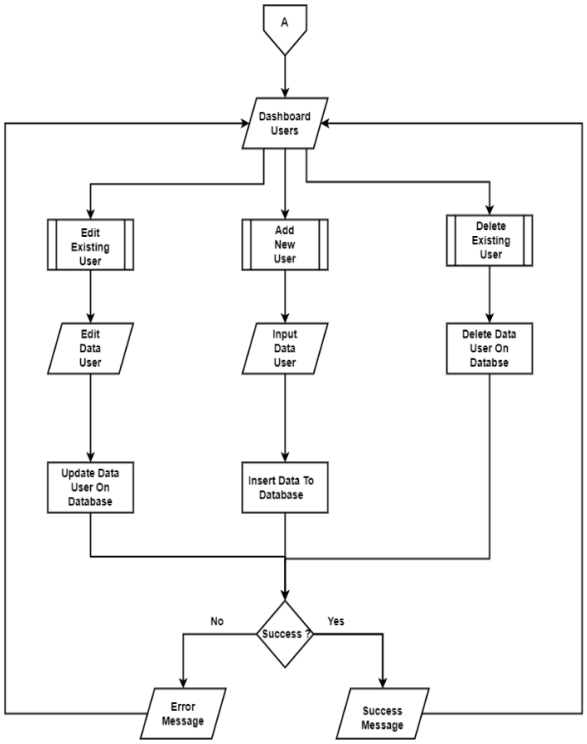
Pemodelan sistem dibangun melalui *Flowchart* dan relasi *Database*. *Flowchart* dirancang dari dua sisi, yakni admin dan *user*:

1. *Flowchart* Admin Gambar 5 - 9 mencakup menu *Login*, *Users*, *Mining*, *Data Testing*, dan *Decision Tree*.

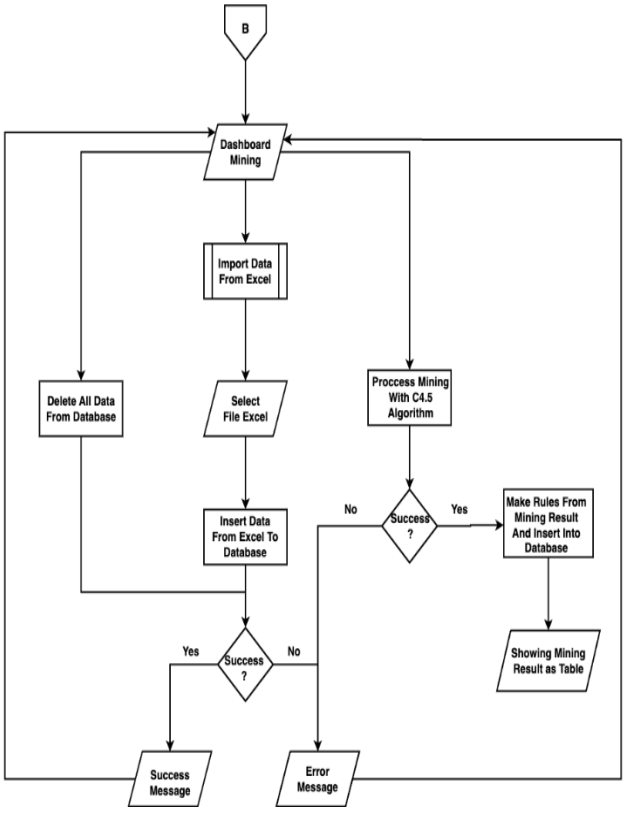


Gambar 5. *FlowChart Menu Login* Sisi *Admin*

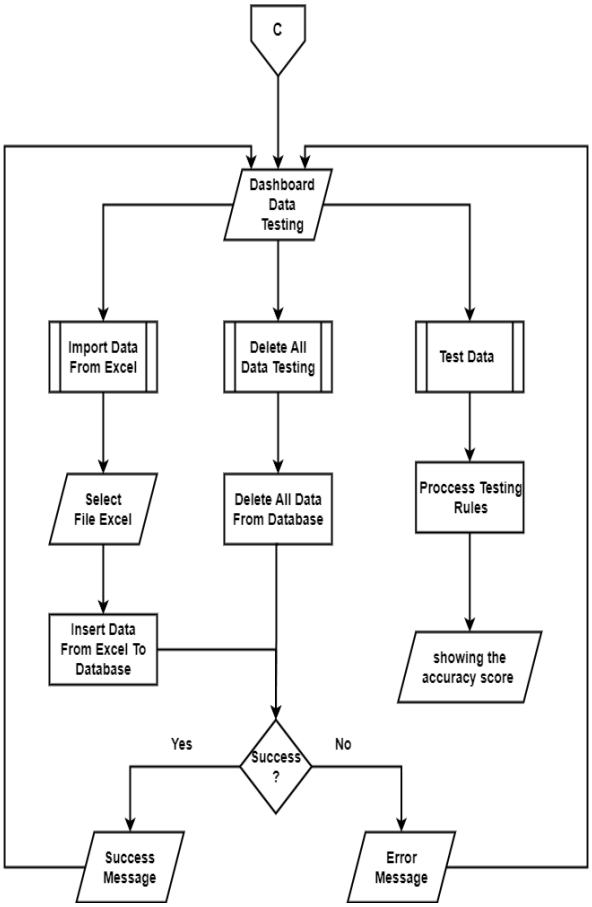
p-ISSN : 2406-7733
e-ISSN : 2597-9922



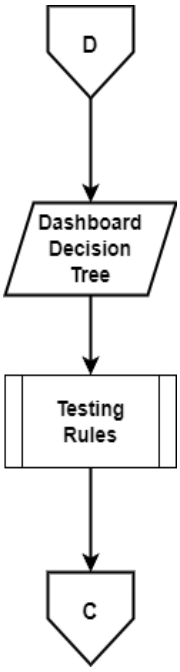
Gambar 6. *FlowChart Menu Users*



Gambar 7. *FlowChart Menu Mining*



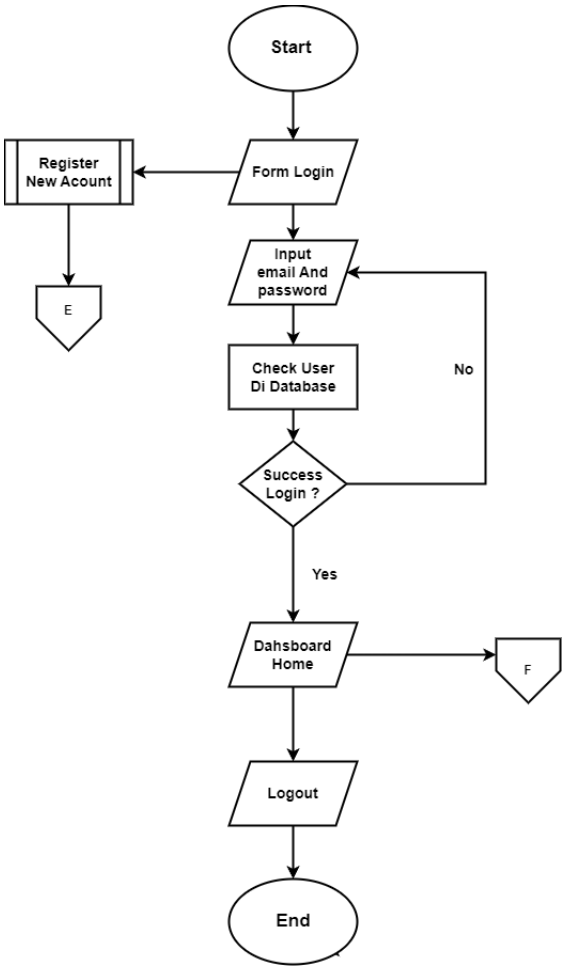
Gambar 8. *FlowChart Menu Data Testing*



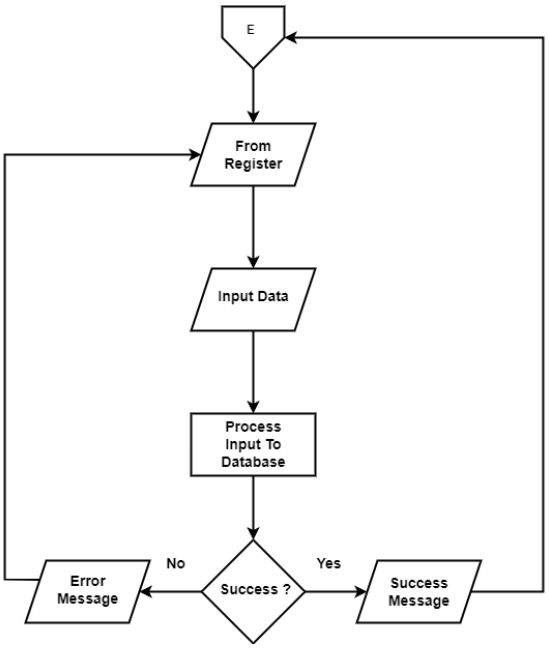
Gambar 9. *FlowChart Menu Decision Tree*

p-ISSN : 2406-7733
e-ISSN : 2597-9922

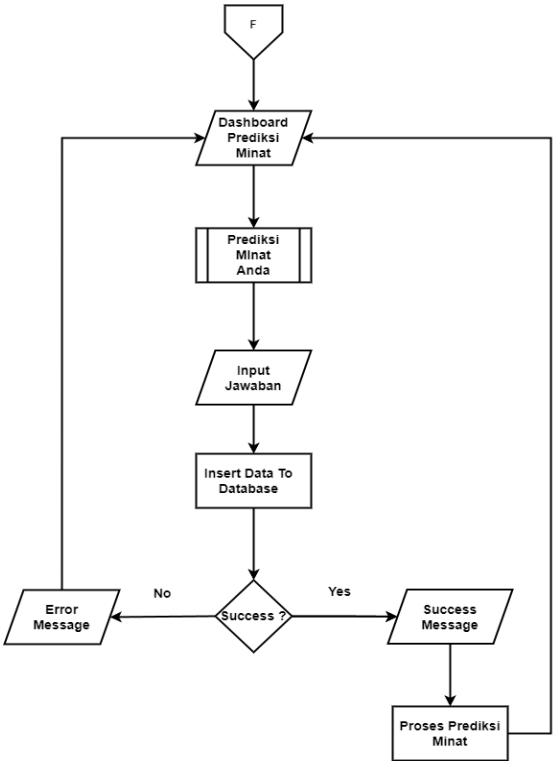
2. *Flowchart User* Gambar 3.8 - 3.10 mencakup menu *Login*, *Register User Baru*, dan *Prediksi Minat*.



Gambar 10. *FlowChart Menu Login Sisi User*

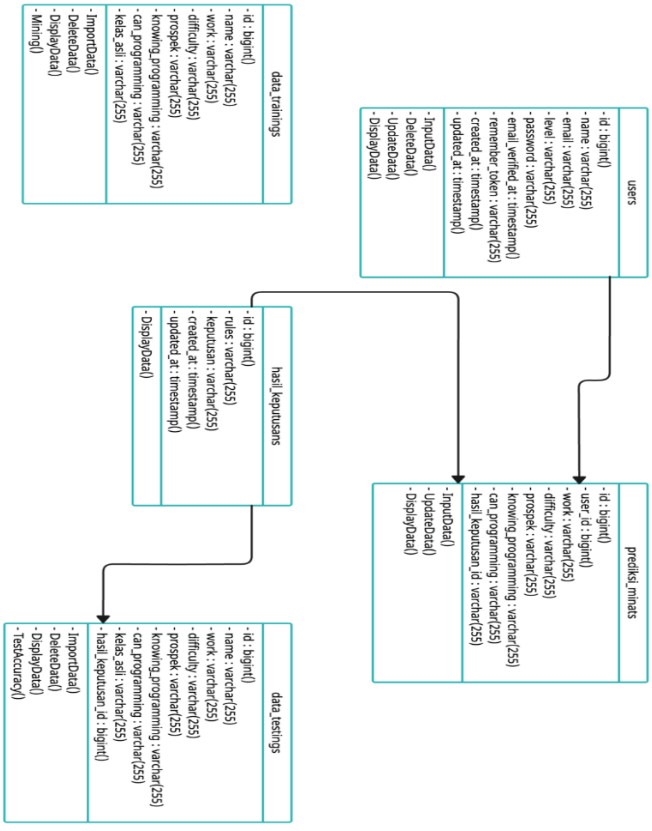


Gambar 11. *FlowChart Menu Register User Baru*



Gambar 12. *FlowChart Menu Prediksi Minat*

Struktur *Relational Database* dipresentasikan pada Gambar 13 *Relational Database*, yang memuat tabel: *users*, *data_trainings*, *data_testings*, *hasil_keputusans*, dan *prediksi_minats*.



Gambar 13. *Relational Database*

p-ISSN : 2406-7733
e-ISSN : 2597-9922

IV. HASIL DAN PEMBAHASAN

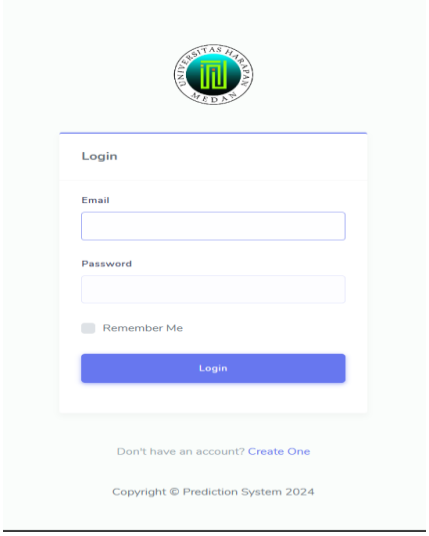
Implementasi Sistem

Implementasi algoritma C4.5 untuk sistem prediksi minat *programming* mahasiswa direalisasikan dalam bentuk aplikasi berbasis *web*. Pengembangan aplikasi ini didukung oleh perangkat keras (*hardware*) berupa laptop Macbook Pro M1 dengan spesifikasi *Processor* Apple M1, RAM 8 GB, dan SSD 256 GB. Sementara untuk perangkat lunak (*software*), sistem ini menggunakan sistem operasi MacOS Sonoma, *browser* Google Chrome, Visual Studio Code sebagai *text editor*, PHP versi 8.2, dan sistem manajemen basis data MySQL

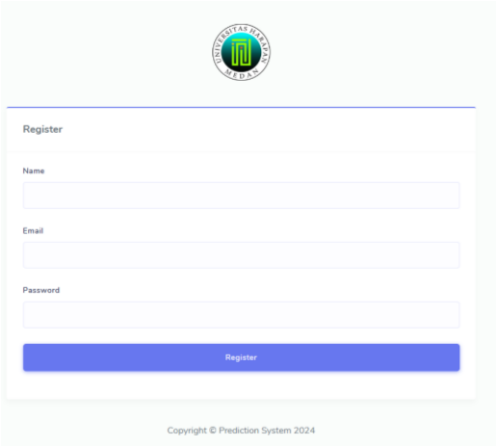
Implementasi Antarmuka Aplikasi

Tahap ini merupakan penerapan nyata dari rancangan desain antarmuka yang telah dibuat pada tahap perancangan sistem. Antarmuka aplikasi yang diimplementasikan meliputi:

1. Halaman Login dan Register: Halaman awal yang digunakan pengguna untuk masuk ke dalam aplikasi atau mendaftarkan akun baru dengan *role user* Gambar 15 dan Gambar 16.



Gambar 15. Halaman *Login*

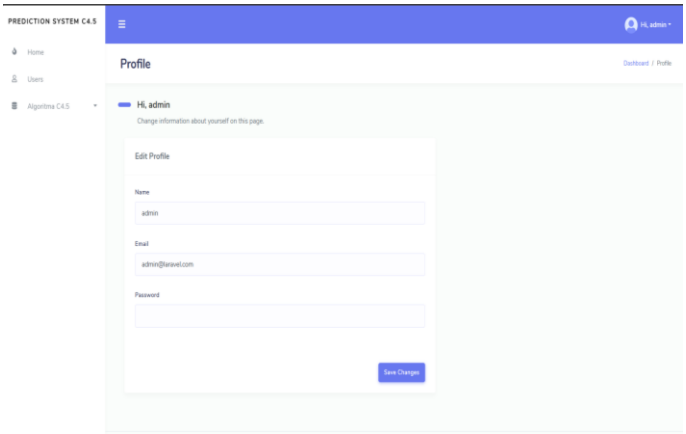


Gambar 16. Halaman *Register*

2. Halaman Home dan Profile: Dasbor utama yang menyambut pengguna saat berhasil *login*, serta halaman untuk memperbarui informasi profil seperti nama, email, dan *password* Gambar 17 dan Gambar 18.

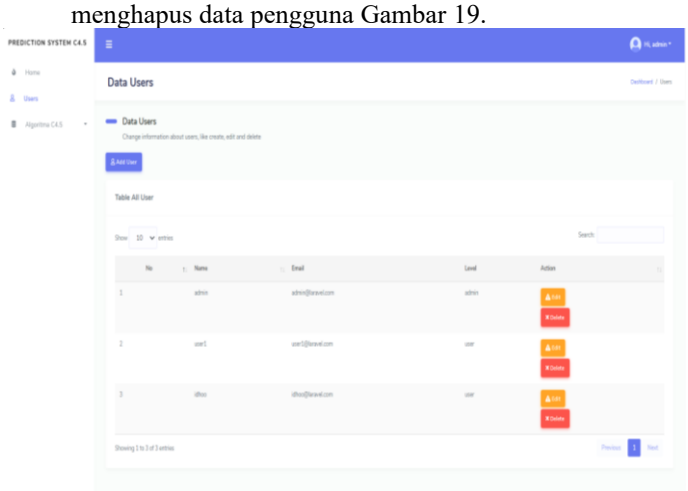


Gambar 17. Halaman *Home* Sisi Admin



Gambar 18. Halaman *Profile*

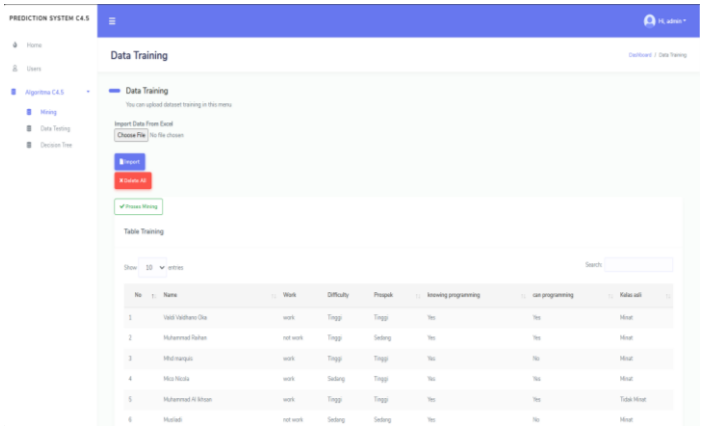
3. Halaman Users: Halaman khusus *admin* yang berfungsi untuk menambah, mengedit, dan menghapus data pengguna Gambar 19.



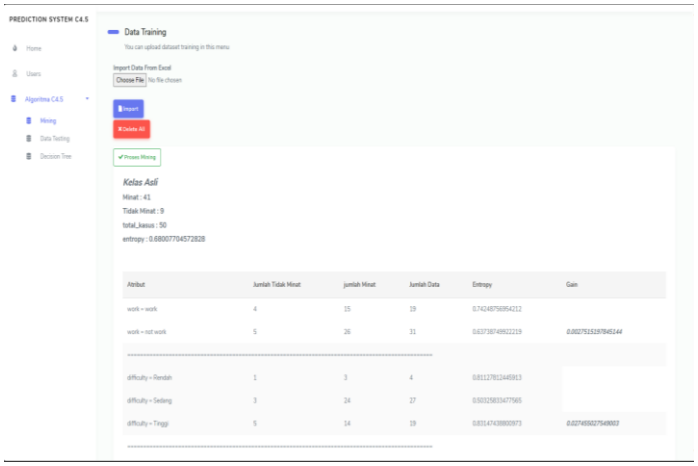
Gambar 19. Halaman *Users*

4. Halaman Mining: Halaman *admin* yang digunakan untuk melakukan proses klasifikasi data *training* (yang diimpor melalui *file* Excel) guna mendapatkan *rules* prediksi Gambar 20 dan Gambar 21 Hasil *Mining*.

p-ISSN : 2406-7733
e-ISSN : 2597-9922

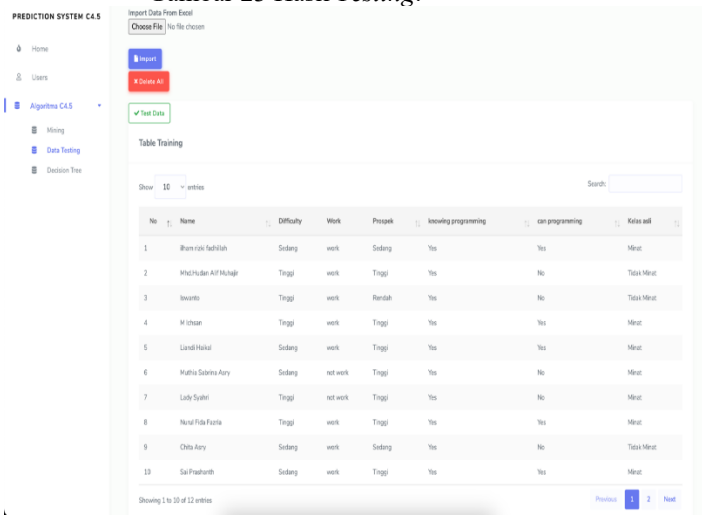


Gambar 20. Halaman *Mining*

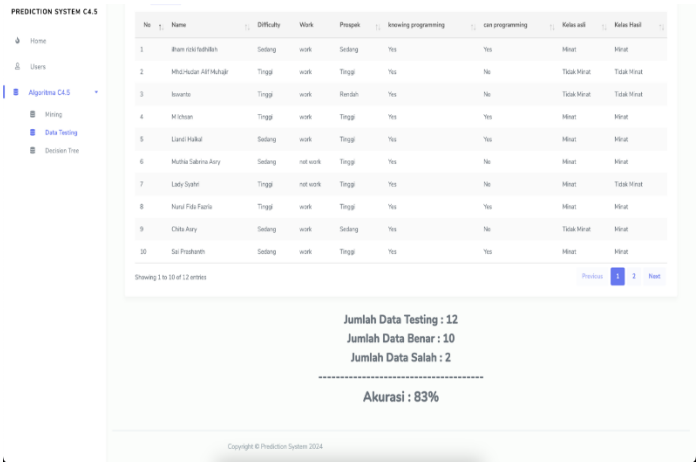


Gambar 21. Hasil *Mining*

5. Halaman *Data Testing*: Halaman fungsional untuk menguji tingkat akurasi dari *rule-rule* yang telah dihasilkan pada proses *mining* menggunakan data *testing* Gambar 22 dan Gambar 23 Hasil *Testing*.

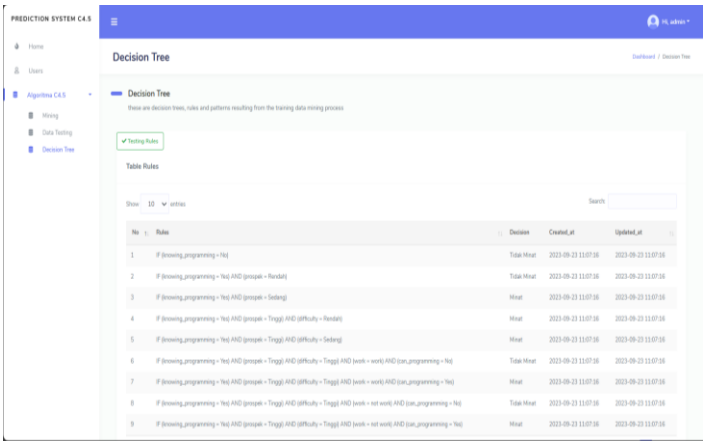


Gambar 22. Halaman Data *Testing*



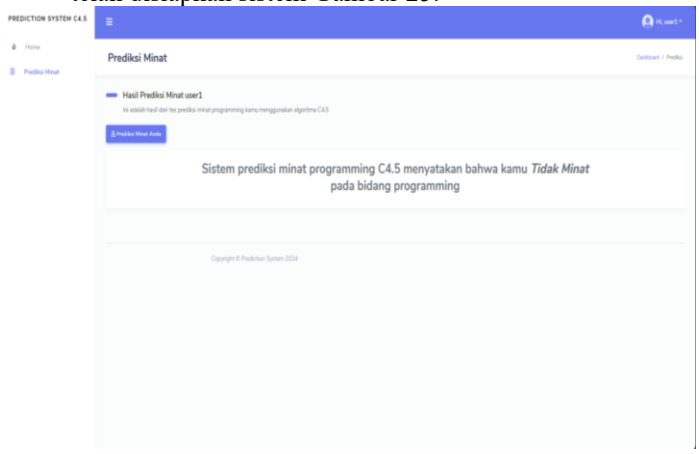
Gambar 23. Hasil *Testing*

6. Halaman Decision Tree: Halaman yang menampilkan daftar *rule* prediksi minat yang diperoleh dari hasil *mining* algoritma C4.5 Gambar 24.



Gambar 24. Halaman *Decision Tree*

7. Halaman Prediksi Minat: Halaman khusus bagi pengguna (*user*) untuk memprediksi minat mereka pada bidang *programming* dengan cara menjawab beberapa pertanyaan prasyarat yang telah disiapkan sistem Gambar 25.



Gambar 25. Halaman Prediksi Minat

p-ISSN : 2406-7733
e-ISSN : 2597-9922

Pengujian Sistem

Pengujian dilakukan untuk memvalidasi bahwa sistem yang dibangun dapat berjalan dengan baik sesuai fungsi yang diharapkan seperti terlihat pada. Tabel 9.

Tabel 9. Skenario dan Hasil Pengujian

No	Skenario Pengujian	Hasil yang Diharapkan	Hasil Pengujian	Validasi
1.	Melakukan *Import* data melalui *File* Excel pada *data training* dan *data testing*.	Data masuk ke *database* dan tampil pada menu terkait.	Aplikasi Berhasil melakukan *import* data melalui *file* excel dan masuk ke dalam *database* serta ditampilkan dalam bentuk tabel.	Valid
2.	Menjalankan Proses *Mining* Untuk Melakukan Klasifikasi Pada *Data Training* Menggunakan Algoritma C4.5	*Data Training* sukses di klasifikasi menggunakan algoritma C4.5 dan aplikasi memunculkan hasil perhitungannya dalam bentuk tabel.	Aplikasi Berhasil melakukan proses klasifikasi dan data perhitungan berhasil ditampilkan dalam bentuk tabel.	Valid
3.	*Rules* yang diperlukan tercipta setelah selesai melakukan proses *mining*	*Rules* sukses dihasilkan dan dapat ditampilkan dengan kalimat yang mudah di mengerti.	Aplikasi berhasil menampilkan *rules* yang dihasilkan dan ditampilkan dalam bentuk tabel pada menu *decision tree*.	Valid
4.	Proses *testing* akurasi pada *rules* yang ada menggunakan *data testing* yang telah di *import*.	Proses *testing* berhasil di dilakukan dan aplikasi berhasil menunjukan akurasi dan laju *error* pada *rules* yang ada.	Aplikasi berhasil melakukan *testing* terhadap *rules* yang ada menggunakan *data testing* yang telah di tetapkan dan menampilkan seberapa besar akurasi dan laju *error* yang di dapatkan dalam bentuk persentase.	Valid
5.	Proses prediksi minat dengan mengisi beberapa pertanyaan yang telah disiapkan	Proses prediksi berhasil dilakukan dan aplikasi menampilkan hasil dari prediksi.	Aplikasi berhasil memprediksi minat programming seseorang yang telah mengisi terlebih dahulu pertanyaan-pertanyaan yang ada pada form dan hasil dari prediksi ditampilkan dalam bentuk kalimat.	Valid

Dilakukan lima skenario pengujian dengan hasil sebagai berikut:

1. Skenario 1 (*Import* Data): Aplikasi berhasil melakukan *import* data *training* dan *testing* dari *file* Excel untuk disimpan ke *database* serta menampilkannya ke dalam tabel (Status: Valid).
2. Skenario 2 (Proses *Mining*): Aplikasi berhasil mengeksekusi proses klasifikasi *data training* menggunakan algoritma C4.5 dan memunculkan hasil perhitungannya dalam bentuk tabel (Status: Valid).
3. Skenario 3 (Pembentukan *Rules*): Aplikasi sukses menghasilkan *rules* yang diperlukan setelah proses *mining* dan menampilkannya dengan kalimat yang mudah dipahami pada halaman *decision tree* (Status: Valid).
4. Skenario 4 (Proses *Testing* Akurasi): Aplikasi berhasil menguji *rules* menggunakan 12 data *testing*. Dari pengujian tersebut, didapatkan jumlah data yang diprediksi benar sebanyak 10 data dan salah 2 data, sehingga menghasilkan skor akurasi sebesar 83% (Status: Valid).
5. Skenario 5 (Proses Prediksi Minat): Aplikasi berhasil memprediksi minat *programming* pengguna berdasarkan form pertanyaan yang diisi, dan menampilkannya dalam bentuk kalimat keputusan (Status: Valid).

V. KESIMPULAN DAN SARAN

Kesimpulan

Setelah melakukan perancangan, analisis, implementasi dan pengujian terhadap aplikasi prediksi minat programming mahasiswa, maka dapat diambil kesimpulan pada penelitian ini adalah sebagai berikut:

1. Perancangan dan pembuatan aplikasi prediksi minat programming mahasiswa berbasis *web* menggunakan bahasa pemograman PHP dan *framework* Laravel yang diimplementasi algoritma C4.5 telah berhasil tercipta dan telah sesuai dengan tujuan yang diharapkan.
2. Hasil klasifikasi data *training* yang dilakukan secara manual dan *rules* yang dihasilkan telah

p-ISSN : 2406-7733
e-ISSN : 2597-9922

sesuai dengan perhitungan yang dihasilkan oleh aplikasi yang dibangun.
3. Proses *testing* yang dilakukan pada *rules* yang tercipta menggunakan sebanyak 12 data, mendapatkan skor akurasi sebesar 83% dan laju *error* sebesar 17%.
4. Aplikasi algoritma C4.5 untuk memprediksi minat programming mahasiswa berhasil memprediksi minat seseorang menggunakan *rules* yang telah dihasilkan pada proses *mining*.

Saran

Adapun saran yang dapat penulis berikan untuk pengembangan pada penelitian ini ke depannya adalah sebagai berikut:

1) Pada penelitian selanjutnya diharapkan aplikasi sistem prediksi minat programming mahasiswa dapat dikembangan dalam bentuk aplikasi *mobile*.
2) Untuk pengembangan selanjutnya, penulis berharap aplikasi prediksi minat programming ini, bisa ditujukan pada semua orang tidak hanya pada lingkup mahasiswa atau perguruan tinggi saja, tentunya dengan perubahan yang telah disesuaikan

DAFTAR PUSTAKA

Achru, A. (2019). Pengembangan minat belajar dalam pembelajaran. *Jurnal Idaarah*, *3*(2), 205–215.

Amalia, Y. R. (2018). *Penerapan Data Mining Untuk Prediksi Penjualan Produk Elektronik Terlaris Menggunakan Metode K-Nearest Neighbor (Studi Kasus: PT. Bintang Multi Sarana Palembang)*. UIN Raden Fatah Palembang.

Candra, M. A. A. (2021). Sistem Informasi Berprestasi Berbasis Web Pada SMP Negeri 7 Kota Metro. *Jurnal Mahasiswa Ilmu Komputer*, *1*(2), 175–189.

Damuri, A., Riyanto, U., Rusdianto, H., & Aminudin, M. (2021). Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako. *JURIKOM (Jurnal Riset Komputer)*, *8*(6), 219–225.

HENDI, I. N. (2021). Rancang Bangun Sistem Informasi Desa Binaan universitas Pembangunan Panca Budi Berbasis Web. *Kumpulan Karya Ilmiah Mahasiswa Fakultas Sains Dan Tekhnologi*, *2*(2), 54.

Herwanto, H. (2019). Diagnosa Statistik Pemetaan Pemahaman Bahasa Pemograman Sebagai Acuan Untuk Mempersiapkan Penelitian Mahasiswa. *NUANSA INFORMATIKA*, *13*(2), 33.

Husna, F. (2021). *Implementasi data mining menggunakan algoritma c4. 5 pada klasifikasi penjualan hijab*. Universitas Islam Negeri Maulana Malik Ibrahim.

Karim, B. A. (2020). Pendidikan Perguruan Tinggi Era 4.0 Dalam Pandemi Covid-19 (Refleksi Sosiologis). *Education and Learning Journal*, *1*(2), 102–112.

Lian, B. (2019). Tanggung jawab Tridharma perguruan tinggi menjawab kebutuhan masyarakat. *Prosiding Seminar Nasional Program Pascasarjana Universitas PGRI Palembang*.

Matondang, A. (2018). Pengaruh Antara Minat Dan Motivasi Dengan Prestasi Belajar. *Bahastra: Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, *2*(2), 24–32.

Nas, C. (2021). Data Mining Prediksi Minat Calon Mahasiswa Memilih Perguruan Tinggi Menggunakan Algoritma C4. 5. *Jurnal Manajemen Informatika (JAMIKA)*, *11*(2), 131–145.

Naufal, M. F. (2018). Analisis Teknik Pembelajaran dan Pengajaran Pemrograman pada Universitas dan Industri. *Jurnal Informatika Dan Multimedia*, *10*(2), 76–83.

Pasaribu, E. P. B. (2022). *Hubungan Efikasi diri dengan Penyesuaian Diri pada Mahasiswa Baru Fakultas Psikologi Universitas Medan Area*. Universitas Medan Area.

Pramudito, D. K. (2022). Data Mining Implementation On Java North Coast Weather Forecast Dataset Using C4. 5 Algorithm. *Jurnal SIGMA*, *13*(3), 139–148.

Primayasa, W., Arifin, I., & Baharsyah, M. Y. (2020). Pengaruh salah pilih jurusan terhadap rasa putus asa mahasiswa Teknik Informatika. *Nathiqiyyah*, *3*(1), 22–26.

Purba, A. D., & Nafeesa, N. (2022). Hubungan Antara Konsep Diri Dengan Kepercayaan Diri Pada Mahasiswa/I Stambuk 2019 Universitas Medan Area. *Jouska: Jurnal Ilmiah Psikologi*, *1*(1), 16–24.

Putri, S. U., Irawan, E., & Rizky, F. (2021). Implementasi Data Mining Untuk Prediksi Penyakit Diabetes Dengan Algoritma C4. 5. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, *2*(1), 39–46.

Rosaly, R., & Prasetyo, A. (2019). Pengertian Flowchart Beserta Fungsi dan Simbol-simbol Flowchart yang Paling Umum Digunakan. *Program Studi Teknik Informatika Politeknik Purbaya*.

RS, M. A. (2022). Implementasi machine learning dengan metode algoritma decision tree c4. 5 untuk pemilihan bandwidth internet wifi rumahan. In *2022*. Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta.

Sonny, S. (2021). *Pengembangan Sistem Presensi Karyawan Dengan Teknologi GPS Berbasis Web Pada PT BPR Dana Makmur Batam*. Prodi

p-ISSN : 2406-7733
e-ISSN : 2597-9922

Teknik Informatika.

Sulistiati, T., Yuliansyah, F., Romzi, M., & Aryani, R. (2020). Membangun Website Toko Online Pempek NTHREE Menggunakan PHP dan MYSQL. *JTIM: Jurnal Teknik Informatika Mahakarya*, *3*(1), 35–44.

Widhi, A. N., Sutanta, E., & Nurnawati, E. K. (2019). Pemanfaatan Framework Laravel Untuk Pengembangan Sistem Informasi Toko Online Di Toko New Trend Baturetno. *Jurnal Script*, 232–238.

Zai, C. (2022). Implementasi Data Mining Sebagai Pengolahan Data. *Jurnal Portal Data*, *2*(3).