

BAB II KERANGKA TEORITIS

A. Landasan Teori

1. Data Mining

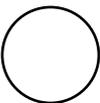
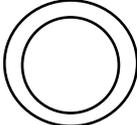
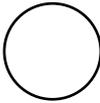
Menurut (Deny Jollyta, William Ramdhan, Muhammad Zarlis 2020, 22); Data mining adalah proses untuk mendapatkan pengetahuan dan informasi baru dengan cara menggali pengetahuan dan informasi tersebut dari data yang berjumlah banyak pada data warehouse. Pada proses ini mencari pola dari sebuah data dimana pola tersebut akan dianalisis, data mana yang memiliki pengetahuan dan informasi yang berguna bagi penelitian ini. Pola data yang memiliki informasi dan pengetahuan kemudian di presentasikan dengan teknik penafsiran dalam bentuk gambar, tulisan, grafik dan lain lain.

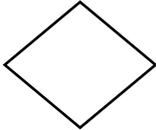
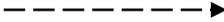
2. Business Process Model and Notation (BPMN)

Menurut (Syafrial Fachri Pane 2020, 196); BPMN adalah grafis yang merepresentasikan sebuah proses bisnis, tujuan dari BPMN adalah menyediakan manajemen proses bisnis untuk pengguna teknis dan bisnis dengan memberikan mark up yang intuitif kepada pelanggan bisnis. BPMN bertujuan untuk menyediakan pengguna teknis dan bisnis manajemen, proses bisnis mengekspresikan kualitas sari proses yang kompleks. Notasi digunakan dalam BPMN untuk menjelaskan proses proses yang terjadi.

Berikut Notasi pemodelan proses bisnis menurut (Syafrial Fachri Pane 2020, 196);:

Tabel 2. 1 Notasi BPMN

No	Simbol	Keterangan
1		Start Event
2		Intermediete Event
3		End Event

No	Simbol	Keterangan
4		Task/Activity
5		Gateway
6		Sequence Flow
7		Message Flow
8		Group

3. Unified Modeling Language (UML)

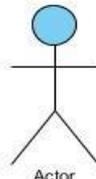
Menurut (Rosa A.S, M. Shalahuddin 2016, 137); UML dapat memenuhi kebutuhan pemodelan berbentuk visual untuk menspesifikasikan, menggambarkan, membangun, dokumentasi dari perangkat lunak. UML berfungsi sebagai pemodelan dan komunikasi sebuah sistem dengan menggunakan diagram dan teks pendukung sebagai bahasa visual. UML dapat digunakan untuk mempermudah pengembangan aplikasi dan prototype.

a. Usecase Diagram

Menurut (Sholiq 2006, 7); diagram usecase menggambarkan interaksi antara usecase dan aktor dimana aktor dapat berupa orang, sistem, atau peralatan yang berinteraksi dengan sistem yang dibangun. use case menggambarkan fungsi sistem dan persyaratan persyaratan yang harus dipenuhi dari sudut pandang pemakai. Diagram use case digunakan untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem dan siapa saja yang berhak menggunakan fungsi-fungsi tersebut.

Berikut adalah notasi Use Case menurut (Sholih 2006, 7);

Tabel 2. 2 Notasi Use Case

Simbol	Deskripsi
Use case Nama Use	Fungsionalitas yang disediakan sistem sebagai unit-unit yang saling bertukar pesan antar unit atau aktor; biasanya dinyatakan dengan menggunakan kata kerja diawal frase nama Use Case
Aktor / actor 	Orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat diluar sistem informaaasi yang akan dibuat itu sendiri, jadi walaupun simbol dari aktor adalah gambar orang, tapi aktor belum tentu merupakan orang: biasanya dinyatakan menggunakan kata benda diawalfrase nama aktor
Asosiasi / assosiation 	Komunikasi antara aktor dan use case yang berpartisipasi pada use case atau use case memiliki interaksi dengan aktor
Ekstensi / extend <<extend>>	Relasi use case tambahan kesebuah use case dinamakan use case yang ditambahkan dapat berdiri sendiri walau tanpa use case tambahan itu; mirip dengan prinsip <i>intherince</i> padapemrograman berorientasi objek; biasanya use case tambahan memiliki nama depan yang sama dengan use case yang ditambahkan.
Generalisasi / generalization 	Hubungan generalisasi dan spesialisasi (umum-khusus) antara dua buah use case dimana fungsi yang satu adalah fungsi yang lebih umum dari lainnya.

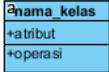
b. Class Diagram

Menurut (Sholih 2006, 13); diagram class menggambarkan interaksi antara kelas dan sistem, kelas memiliki informasi dan tingkah laku (Behavior) yang

berhubungan dengan informasi tersebut. Sebuah kelas pada diagram kelas dibuat untuk tipe objek pada diagram sequence. Dapat memberikan gambaran mengenai sistem atau perangkat lunak serta relasi-relasi yang terkandung di dalamnya.

Berikut adalah notasi class diagram menurut (Sholiq 2006, 13);

Tabel 2. 3 notasi Class Diagram

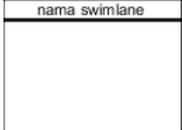
Simbol	Deskripsi
Kelas 	Kelas pada struktur sistem
Antarmuka / <i>Interface</i>  Nama_interface	Sama dengan konsep <i>interface</i> dalam pemrograman berorientasi objek
Asosiasi / <i>Association</i> 	Relasi antar kelas dengan makna umum, asosiasi biasanya juga disertai dengan <i>multiplicity</i>
Asosiasi berarah / <i>Directed Association</i> 	Relasi antar kelas dengan makna kelas satu digunakan oleh kelas yang lain, asosiasi biasanya juga disertai dengan <i>multiplicity</i>
Generalisasi 	Relasi antar kelas dengan makna generalisasi-spesialisasi (umum khusus)
Kebergantungan / <i>Dependency</i> 	Relasi antar kelas dengan makna kebergantungan antarkelas
Agregasi / <i>Aggregation</i> 	Relas antarkelas dengan makna semua-bagian (whole-part)

c. Activity Diagram

Menurut (Sholih 2006, 8); diagram activity menggambarkan sebuah aliran fungsionalitas sistem, diagram activity juga dapat dipakai untuk menggambarkan aliran kerja bisnis (Business Workflow) dan dapat juga dipakai untuk menggambarkan aliran kejadian (Flow of Event) di dalam usecase. Activity Diagram dibuat berdasarkan sebuah atau beberapa use case. Menggambarkan proses dan urutan aktivitas dalam sebuah proses.

Berikut adalah Notasi Activity Diagram Menurut (Sholih 2006, 8);

Tabel 2. 4 Notasi Activity Diagram

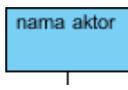
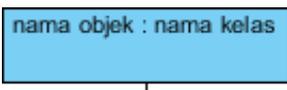
Simbol	Deskripsi
Status awal 	Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal
Aktivitas 	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja
Percabangan / <i>decision</i> 	Asosiasi percabangan dimana jika ada pilihan aktivitas lebih dari satu
Penggabungan / <i>join</i> 	Asosiasi penggabungan dimana lebih dari satu aktivitas digabungkan menjadi satu
Status akhir 	Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status akhir
Swimlane  Atau 	Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi

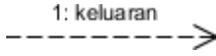
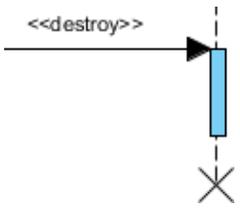
d. Sequence Diagram

Diagram sequence menurut (Sholih 2006, 9-10); Berfungsi untuk menggambarkan aliran fungsionalitas dalam usecase diagram ini juga menunjukkan langkah langkah proses yang terjadi dalam usecase. Diagram ini biasa digunakan untuk menggambarkan skenario atau rangkaian langkah – langkah yang dilakukan sebagai respon dari sebuah event untuk menghasilkan output tertentu. perubahan apa saja yang terjadi secara internal dan output apa yang dihasilkan.

Berikut adalah notasi Sequence Diagram menurut (Sholih 2006, 9-10);

Tabel 2. 5 Notasi Sequence Diagram

Simbol	Deskripsi
<p>Aktor</p>  <p>Atau</p> 	Orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat diluar sistem informasi itu sendiri, jadi walaupun simbol dari aktor adalah orang, tapi aktor belum tentu merupakan orang; biasanya dinyatakan menggunakan kata benda diawal <i>frase</i> nama aktor
<p>Garis hidup / Lifeline</p> 	Menyatakan kehidupan suatu objek
<p>Objek</p> 	Menyatakan objek yang berinteraksi pesan
<p>Waktu aktif</p> 	Menyatakan objek dalam keadaan aktif dan berinteraksi, semua yang terhubung dengan waktu aktif ini adalah sebuah tahapan yang dilakukan didalamnya.
<p>Pesan tipe create</p>  <p><<create>></p>	Menyatakan suatu objek membuat objek yang lain, arah panah mengarah pada objek yang dibuat

Simbol	Deskripsi
	<p>Menyatakan suatu objek memanggil operasi / metode yang ada pada objek lain atau dirinya sendiri.</p> <p>Arah panah mengarah pada objek yang memiliki operasi / metode, karena ini memanggil operasi / metode maka operasi / metode yang dipanggil harus ada pada diagram kelas sesuai dengan kelas objek yang berinteraksi</p>
	<p>Menyatakan bahwa suatu objek mengirimkan data / masukan / informasi ke objek lainnya, arah panah mengarah pada objek yang dikirim</p>
	<p>Menyatakan suatu objek yang telah menjalankan suatu operasi atau metode menghasilkan suatu kembalian ke objek tertentu, arah panah mengarah pada objek yang menerima kembalian</p>
	<p>Menyatakan suatu objek mengakhiri hidup objek yang lain, arah panah mengarah pada objek yang diakhiri, sebaiknya jika ada create maka ada destroy</p>

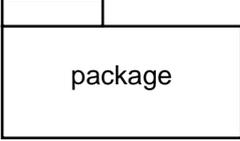
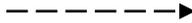
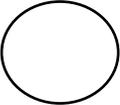
e. Simbol Component Diagram

Diagram komponen menurut (Sholiq 2006, 15-16); berfungsi untuk menggambarkan model secara fisik komponen berupa perangkat lunak pada sistem dan keterhubungan antara perangkat lunak dan sistem, terdapat dua komponen pada diagram komponen yaitu executable dan kode pustaka (libraries code). Kegunaan Component Diagram Sebuah component diagram dibuat berdasarkan kegunaan sebagai berikut. Sebagai source code perangkat lunak (software).

Komponen executable yang dilepas dan digunakan oleh user. Bagian dari basis data secara fisik.

Berikut adalah notasi Component Diagram menurut (Sholiq 2006, 15-16);

Tabel 2. 6 Notasi Component Diagram

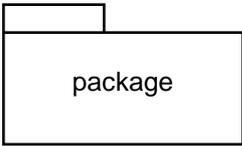
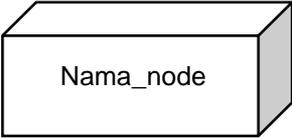
Simbol	Deskripsi
package 	Package adalah sebuah wadah dari satu atau lebih komponen
komponen 	Berisi koponen komponen dalam sistem
Kebergantungan/dependency 	Kebergantungan antar komponen, arah panah mengarah pada komponen
Antarmuka/Interface 	Antarmuka komponen
Link 	Relasi antar package atau komponen

f. Simbol Deployment Diagram

Diagram deployment menurut (Sholiq 2006, 17); menggambarkan rancangan fisik jaringan dimana akan terdapat banyak komponen dalam diagram ini. terdapat node pada diagram ini dimana node berisi banyak sub sistem yang dijalankan pada peralatan fiisik yang terpisah. Deployment diagram digunakan untuk memvisualisasikan hubungan antara software dan hardware.

Berikut adalah Notasi Deployment Diagram menurut (Sholiq 2006, 17);

Tabel 2. 7 Notasi Deployment Diagram

Simbol	Deskripsi
package 	Package adalah sebuah wadah dari satu atau lebih komponen
Node 	Berisi perangkat lunak atau perangkat keras
Kebergantungan/dependency 	Kebergantungan antar node, arah panah mengarah pada node
Link 	Relasi antar node

4. Database

a. Database

Menurut (Rosa A.S 2016, 43); basis data adalah sistem yang sudah terkomputerisasi yang berfungsi untuk memelihara data atau informasi yang sudah di olah sebelumnya dan membuat informasi atau data tersebut tersedia saat di butuhkan di masa yang akan datang. Fungsi Database Mengelompokan data dan informasi, Memudahkan dalam identifikasi data. Memudahkan proses akses, menyimpan, pembaharuan, dan penghapusan data Menjadi alternatif terkait masalah penyimpanan ruang dalam suatu aplikasi.

b. MySQL

Menurut (Dr. Gede Indrawan 2021, 3); Mysql adalah sebuah server database yang bersifat open source atau gratis digunakan. perangkat lunak database ini berbentuk database relational yang menggunakan bahasa permintaan yang bernama SQL (Structured query language). Kegunaan atau fungsi MySQL sendiri adalah untuk data warehousing (gudang data), yaitu pengumpulan data dari berbagai sumber, untuk e-commerce, maupun aplikasi logging

5. Web Server

a. Webserver

Menurut (Yudho yudhanto 2018, 10); Webserver memiliki dua pengertian jika merujuk pada hardware webserver berfungsi untuk menyimpan data seperti HTML dokumen, gambar sedangkan jika merujuk pada software fungsi webserver adalah pusat yang mengontrol proses permintaan yang diterima oleh browser. Fungsi web server adalah untuk menerima permintaan HTTP atau HTTPS dari komputer klien. Yang dimaksud sebagai klien di sini adalah browser, seperti Google Chrome dan Firefox. Selanjutnya, server akan mengirimkan respon atas permintaan tersebut kepada klien dalam wujud halaman website.

b. Apache

Menurut (Tim EMS 2015, 17); Apache adalah sebuah perangkat lunak server yang free dan open source, Apache dapat dijalankan sebagai layanan/ service yang memungkinkan penggunaanya dapat menghidupkan dan mematikannya. Apache adalah sebuah nama web server yang bertanggung jawab pada request-response HTTP dan logging informasi secara detail (kegunaan dasarnya). Selain itu, Apache juga diartikan sebagai suatu web server yang kompak, modular, mengikuti standar protokol HTTP, dan tentu saja sangat digemari.

6. Bahasa Pemrograman

a. PHP

Menurut (Supono 2016, 3); PHP adalah suatu bahasa pemrograman yang menghasilkan kode mesin dari baris kode program yang di terjemahkan yang mana kode mesin tersebut dapat dimengerti oleh komputer yang bersifat server side yang dapat di tambahkan ke HTML. fungsi HTML adalah untuk mengelola serangkaian data dan informasi sehingga suatu dokumen dapat diakses dan ditampilkan di Internet melalui layanan web. Fungsi HTML yang lebih spesifik yaitu Membuat halaman web. Menampilkan berbagai informasi di dalam sebuah browser Internet.

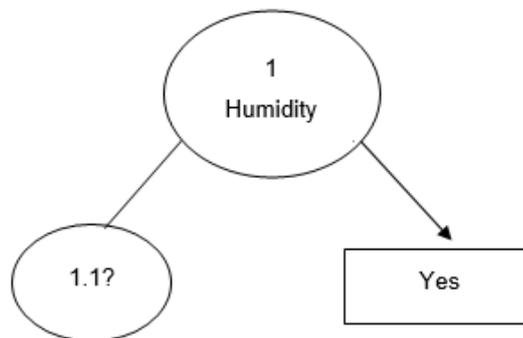
b. Hypertext Markup Language (HTML)

Menurut (Wardana 2016, 3); HTML adalah bahasa pemrograman yang berfungsi untuk pengembangan web statis atau website yang tidak memiliki interaksi dengan user dimana web statis jarang dikembangkan oleh programmer web karena permintaan user yang lebih tertarik pada web yang mempunyai interaksi tinggi. fungsi HTML adalah untuk mengelola serangkaian data dan informasi sehingga suatu dokumen dapat diakses dan ditampilkan di Internet melalui layanan web. Fungsi HTML yang lebih spesifik yaitu Membuat halaman web. Menampilkan berbagai informasi di dalam sebuah browser Internet

7. Algoritma C4.5

Algoritma C 4.5 adalah salah satu metode algoritma yang berfungsi untuk membuat model pohon keputusan. Pohon keputusan berguna untuk memprediksi selain itu pohon keputusan juga berfungsi untuk mengeksplorasi data, menemukan hubungan antara variabel input dengan variabel target; Struktur dari pohon keputusan ini digunakan untuk membagi data menjadi record yang lebih kecil dan beberapa aturan. Data dalam pohon keputusan akan di presentasikan dalam bentuk tabel dimana tabel berisikan dengan atribut dan record, atribut disini adalah sebuah nilai ukur yang dibuat untuk menentukan kriteria dalam pembentukan pohon keputusan; Setelah data sudah dibuat menjadi model pohon yang kemudian akan menghasilkan beberapa rule atau aturan, dan kemudian rule rule tersebut akan di sederhanakan. (Luthfi 2009,13-14);

Menurut (Dicky Nofriansyah 2015, 58-59); Didalam pohon keputusan terdapat beberapa elemen : Rule, node, Relationship



Dalam algoritma C4.5 Terdapat 2 Elemen yang yaitu Gain dan Entropy. Untuk memilih atribut sebagai akar berdasar pada nilai gain tertinggi dari atribut atribut yang ada. Berikut adalah rumus untuk menghitung nilai gain :

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

keterangan :

- S : Hipunan Kasus
- A : Atribut
- n : Jumlah partisi atribut A
- |S_i| : Jumlah kasus pada partisi ke i
- |S| : Jumlah kasus dalam S

untuk perhitungan entropi dapat menggunakan rumus :

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \dots (2)$$

keterangan :

S : Himpunan kasus

A : Fitur

n : Jumlah partisi S

pi : Proporsi dari Si terhadap S

Berikut adalah contoh kasus yang bersumber dari (Dicky Nofriansyah 2015, 59-76); untuk menentukan kelayakan mahasiswa yang akan mengikuti semester pendek dibutuhkan sebuah metode khusus, berikut adalah table dari data mahasiswa yang akan di proses

No	NIRM	Nama	P1	P2	P3	P4	Hasil
1	2011020215	Brahina Syahputra	MS	TS	MS	MS	Tidak Layak
2	2012010415	Evikaswina Br Ginting	MS	MS	MS	MS	Layak
3	2011030195	Fahri Rezeki Pane	MS	MS	MS	MS	Layak
4	2011010394	Fahrul Panjaitan	TS	MS	MS	MS	Tidak Layak
5	2010020016	Fery Indriawan	MS	MS	MS	MS	Layak
6	2010030067	Gusti Pratama	MS	MS	MS	MS	Layak
7	2010010529	Hadi Santosa	MS	MS	MS	MS	Layak
8	2010030227	Juli Kutnia Ningsih	MS	TS	MS	MS	Tidak Layak
9	2012010246	Lidyatik	MS	MS	MS	MS	Layak
10	2010030209	M. Rizky Satryo bagus	MS	TS	MS	MS	Tidak Layak
11	2012010245	Mega Sihombing	MS	MS	MS	MS	Layak
12	2012010253	Muhammad Rapi pane	MS	MS	MS	MS	Layak
13	2012010545	Novita Anggraini	MS	MS	MS	MS	Layak
14	2012010845	Puspita	MS	MS	MS	MS	Layak
15	2012010355	Rahmat	MS	MS	MS	MS	Layak
16	2012010125	Rini Siti Winarsih	MS	MS	MS	MS	Layak
17	2012010123	Rudi Salam Siregar	MS	TS	MS	MS	Tidak Layak
18	2012010112	Sheila Aditia Putri	MS	MS	MS	MS	Layak
19	2012010290	Sri Rezeki Nurmah	MS	TS	MS	MS	Tidak Layak
20	2012010324	Sri Wahyuni	MS	MS	MS	MS	Layak

keterangan :

P1 = absensi

P2 = nilai

P3 = kelengkapan administrasi

P4 = jumlah matakuliah

MS = memenuhi syarat

TS = tidak memenuhi syarat

(tabel penyelesaian) dari table diatas :

No	Node	X	Jumlah Kasus	Layak	Tidak Layak	Entropy	Gain
1	Total		20	14	6	0,881291	
	1. Absensi	MS	19	14	5	0,831474	0,091390
		TS	1	0	1	0,000000	
	2. Nilai	MS	15	14	1	0,353359	0,616271
		TS	5	0	5	0,000000	
	3. Administrasi	MS	20	14	6	0,881291	0,000000
		TS	0	0	0	0,000000	
	4. Jumlah Kasus	MS	20	14	6	0,881291	0,000000
		TS	0	0	0	0,000000	

untuk total keseluruhan terdapat 20 data mahasiswa yang diri dari 14 yang dinyatakan layak dan 6 yang dinyatakan tidak layak. Berikut hitungan entropy total :

$$Entropy(Total) = \left(-\frac{14}{20} * \log_2 \left(\frac{14}{20} \right) \right) + \left(-\frac{6}{20} * \log_2 \left(\frac{6}{20} \right) \right)$$

maka hasilnya 0,881291

dari 4 kriteria yang ada dijabarkan nilai gain terbesar yaitu pada kriteria nilai yang terdiri dari 2 parameter :

a. Parameter 1 memenuhi syarat

dari Total 15 terdapat 14 yang layak dan 1 yang tidak layak maka hitungan entropy nya :

$$Entropy(Nilai Ms) = \left(-\frac{14}{15} * \log_2 \left(\frac{14}{15} \right) \right) + \left(-\frac{1}{15} * \log_2 \left(\frac{1}{15} \right) \right)$$

maka hasilnya 0,353359

b. Parameter 2 tidak memenuhi syarat

dari total 5 terdapat 0 yang layak dan 5 yang tidak layak maka hitungan entropy nya :

$$Entropy(\text{Nilai } Ts) = \left(-\frac{0}{5} * \text{Log}_2\left(\frac{0}{5}\right)\right) + \left(-\frac{5}{5} * \text{Log}_2\left(\frac{5}{5}\right)\right)$$

Maka Hasilnya 0,000000

dari kedua entropy tersubt dapat di simpulkan nilai gainnya adalah :

$$Gain(\text{Nilai}) = Entropy(\text{total}) - \left(\frac{15}{20} * 0,353359\right) + \left(-\frac{5}{20} * 0,000000\right)$$

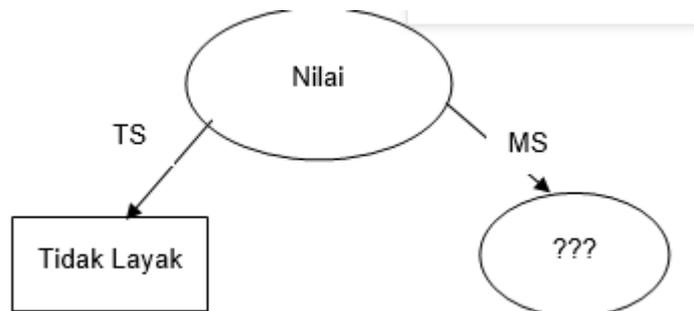
maka diperoleh nilai gain = 0,616271 (nilai gain terbesar)

No	Kriteria	Nilai Gain	Peringkat
1	Nilai	0,616271398	1
2	Absensi	0,091390231	2
3	Administrasi	0,000000	3
4	Jumlah matakuliah	0,000000	4

dari 4 kriteria terbentuklah 4 peringkat berdasarkan nilai gain terbesar, Dengan kriteria nilai total gainnya 0,616271398 dan seterusnya.

2. Nilai	MS	15	14	1	0,353359	0,616271
	TS	5	0	5	0,000000	

pada tabel diatas perlu di hitung kembali nilai entropy dan gain pada node nilai dengan parameter MS. Parameter TS tidak perlu dihitung kembali karena sudah mutlak "Tidak Layak". Berikut adalah gambar pohon keputusannya :



berdasarkan pohon keputusan diatas berikut adalah tabel Decision System :

No	NIRM	Nama	P1	P2	P3	P4	Hasil
1	2012010415	Evikaswina Br Ginting	MS	MS	MS	MS	Layak

2	2011030195	Fahri Rezeki Pane	MS	MS	MS	MS	Layak
3	2011010394	Fahrul Panjaitan	TS	MS	MS	MS	Tidak Layak
4	2010020016	Fery Indriawan	MS	MS	MS	MS	Layak
5	2010030067	Gusti Pratama	MS	MS	MS	MS	Layak
6	2010010529	Hadi Santosa	MS	MS	MS	MS	Layak
7	2012010246	Lidyatik	MS	MS	MS	MS	Layak
8	2012010245	Mega Sihombing	MS	MS	MS	MS	Layak
9	2012010253	Muhammad Rapi pane	MS	MS	MS	MS	Layak
10	2012010545	Novita Anggraini	MS	MS	MS	MS	Layak
11	2012010845	Puspita	MS	MS	MS	MS	Layak
12	2012010355	Rahmat	MS	MS	MS	MS	Layak
13	2012010125	Rini Siti Winarsih	MS	MS	MS	MS	Layak
14	2012010112	Sheila Aditia Putri	MS	MS	MS	MS	Layak
15	2012010324	Sri Wahyuni	MS	MS	MS	MS	Layak

setelah itu ditentukan kebalikan tabel penyelesaiannya

No	Node		Jumlah Kasus	Layak	Tidak Layak	Entropy	Gain
1	Total	X	15	14	1	0,353359	
	1. Absensi	MS	14	14	0	0,000000	0,353359
		TS	1	0	1	0,000000	
	2. Nilai	MS	15	14	1	0,353359	0,000000
		TS	0	0	0	0,000000	
	3. Administrasi	MS	15	14	1	0,353359	0,000000
		TS	0	0	0	0,000000	
	4. Jumlah Kasus	MS	15	14	1	0,353359	0,000000
		TS	0	0	0	0,000000	

untuk total keseluruhan terdapat 15 data mahasiswa yang terdiri dari 14 yang dinyatakan layak dan 1 yang dinyatakan tidak layak. Berikut hitungan entropy total :

$$Entropy(Total) = \left(-\frac{14}{15} * \log_2 \left(\frac{14}{15} \right) \right) + \left(-\frac{1}{15} * \log_2 \left(\frac{1}{15} \right) \right)$$

maka hasilnya 0,353359

dari 4 kriteria yang ada dijabarkan nilai gain terbesar yaitu pada kriteria nilai yang terdiri dari 1 parameter:

a. Parameter 1 memenuhi syarat

dari Total 15 terdapat 14 yang layak dan 1 yang tidak layak maka hitungan entropy nya:

$$Gain(Absensi) = Entropy (total) \left(\frac{14}{15} * 0,000000 \right) + \left(- \frac{1}{15} * 0,000000 \right)$$

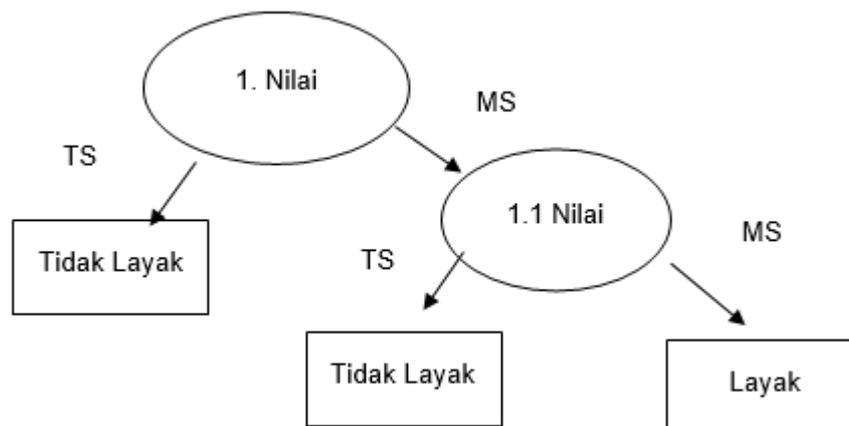
maka Hasilnya 0,353359

No	Kriteria	Nilai Gain	Peringkat
1	Absensi	0,353359	1
2	Nilai	0,000000	2
3	Administrasi	0,000000	3
4	Jumlah matakuliah	0,000000	4

dari 4 kriteria terbentuklah 4 peringkat berdasarkan nilai gain terbesar, Dengan kriteria nilai total gainnya 0,353359 dan seterusnya.

1. Absensi	MS	14	14	0	0,000000	0,353359
	TS	1	0	1	0,000000	

tabel diatas mendeskripsikan bahwa tidak ada yang perlu di hitung kembali nilai entropy dan gainnya. Jadi, Parameter absensi = MS nilai keputusannya mutlak "Layak", sedangkan parameter absensi = TS nilai keputusannya mutlak "Tidak Layak". Berikut adalah pohon Keputusannya:



berdasarkan pohon keputusan diatas maka dapat disimpulkan Mahasiswa layak atau tidak mengikuti semester pendek yang sesuai dengan rule:

a. nilai = TS Then hasil = Tidak

No	NIRM	Nama
1	2011010394	Fahrul Panjaitan

b. nilai = MS Then Hasil = Layak or Tidak Layak

No	NIRM	Nama
1	2012010415	Evikaswina Br Ginting
2	2011030195	Fahri Rezeki Pane
3	2011010394	Fahrul Panjaitan
4	2010020016	Fery Indriawan
5	2010030067	Gusti Pratama
6	2010010529	Hadi Santosa
7	2012010246	Lidyatik
8	2012010245	Mega Sihombing
9	2012010253	Muhammad Rapi pane
10	2012010545	Novita Anggraini
11	2012010845	Puspita
12	2012010355	Rahmat
13	2012010125	Rini Siti Winarsih
14	2012010112	Sheila Aditia Putri
15	2012010324	Sri Wahyuni

c. nilai = MS and absensi = MS then hasil = layak

No	NIRM	Nama
1	2012010415	Evikaswina Br Ginting
2	2011030195	Fahri Rezeki Pane
3	2010020016	Fery Indriawan
4	2010030067	Gusti Pratama
5	2010010529	Hadi Santosa
6	2012010246	Lidyatik
7	2012010245	Mega Sihombing
8	2012010253	Muhammad Rapi pane
9	2012010545	Novita Anggraini
10	2012010845	Puspita
11	2012010355	Rahmat
12	2012010125	Rini Siti Winarsih

13	2012010112	Sheila Aditia Putri
14	2012010324	Sri Wahyuni

d. nilai = MS and absensi = TS then hasil = tidak layak.

No	NIRM	Nama
1	2011010394	Fahrul Panjaitan

B. Tinjauan Studi

- (a.) Penelitian yang dilakukan oleh (Fitriani, Defit, and Nurcahyo 2021) dengan judul “Prediksi Hasil Belajar Siswa Secara Daring pada Masa Pandemi COVID-19 Menggunakan Metode C4.5”. Penelitian ini untuk mengklasifikasi data nilai siswa kelas XII Jurusan Multimedia di SMKN 2 Padang Panjang dan hasil klasifikasi dapat memprediksi hasil belajar siswa pada masa pandemi. Data nilai siswa yang diolah diambil dari 1 (satu) mata pelajaran sebagai sampel data penelitian. Nilai hasil belajar siswa dianalisa menggunakan Metode C4.5 untuk mendapatkan pengetahuan baru dari data nilai hasil belajar siswa yang dilaksanakan pada masa pandemi COVID-19. Data yang dianalisa terdiri dari atribut absensi, tugas, Ulangan Harian (UH) dan nilai ujian yang mempengaruhi kriteria keputusan hasil belajar siswa pada pelajaran yang dilaksanakan secara daring. Kriteria keputusan hasil belajar terdiri dari “Memuaskan” dan “Kurang Memuaskan”.
- (b.) Penelitian di lakukan oleh (Siahaan et al. 2020) dengan Judul “Penerapan Algoritma C4.5 dalam Menentukan Faktor yang Dapat Meningkatkan Kemampuan Bahasa Inggris pada Mahasiswa.” Penelitian ini menghasilkan akurasi untuk menentukan faktor yang dapat meningkatkan kemampuan bahasa inggris pada mahasiswa. Penelitian dilakukan oleh penulis untuk meningkatkan bahasa inggris pada mahasiswa, karena pentingnya bahasa inggris untuk dipelajari dan dikuasai.
- (c.) Penelitian di lakukan oleh (Irnanda and Windarto 2020) “Penerapan Klasifikasi C4.5 Dalam Meningkatkan Kecakapan Berbahasa Inggris dalam Masyarakat”. Penelitian ini bertujuan untuk meningkatkan kecakapan bahasa inggris di kalangan masyarakat dengan menggunakan teknik datamining C4.5. Sumber data penelitian diperoleh dari beberapa kalangan masyarakat seperti mahasiswa/i, siswa/i dan para pekerja melalui wawancara dan pembagian angket secara

random sebanyak 150 orang. Atribut yang digunakan sebagai parameter penilaian untuk meningkatkan kecakapan berbahasa Inggris antara lain: Membaca(C1), Praktek(C2), Pelafalan(C3), Lingkungan(C4) dan Pemanfaatan Teknologi(C5). Hasil pengolahan C4.5 dengan menggunakan bantuan software RapidMiner adalah atribut C2 (Praktek) menjadi atribut yang paling berpengaruh untuk meningkatkan kecakapan berbahasa Inggris dalam masyarakat dan data performance yang ditunjukkan terhadap kesesuaian metode C4.5 akurasi adalah 95%.

- (d.) Penelitian dilakukan oleh (Abdillah and Setyanto 2020) dengan judul "Implementasi Decision Tree Algoritma C4.5 Untuk Memprediksi Kesuksesan Pendidikan Karakter". Tujuan dari penelitian ini adalah mengidentifikasi calon peserta didik, serta memprediksi kesuksesannya dalam sistem pendidikan karakter dengan metode yang dipakai C 4.5 karena dapat digunakan untuk melakukan prediksi dan klasifikasi terhadap calon siswa dengan cara membuat pohon keputusan berdasarkan data-data yang sudah ada dan melakukan prediksi terhadap calon siswa baru dengan data awal yang diambil dari data mahasiswa Unires Yogyakarta sebagai objek penelitian. Dengan hasil diketahui bahwa kombinasi atribut Bidang Bahasa dan Sosial, Latar belakang pendidikan agama dan orang tua yang menjadi seorang pendidik/guru, serta kemampuan untuk membaca Al-Qur'an berkorelasi positif terhadap kesuksesan pendidikan karakter. Nilai accuracy sebesar 60,91%.
- (e.) Penelitian dilakukan oleh (OSCARIO, JASMIR, and NOVIANTO 2019) dengan judul "Penerapan Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : SD Sariputra Jambi)". Peneliti menemukan ada masalah pada proses pembelajaran di SD Sariputra Jambi, masih banyak murid yang tidak fokus atau tidak memiliki keinginan untuk belajar. Dimana peneliti berasumsi kemungkinan terjadi akibat ketidakcocokan gaya belajar dan metode belajar guru tersebut. Berangkat dari masalah tersebut peneliti melakukan analisis data mining terhadap data murid dan kepribadiannya agar data tersebut dapat digunakan menjadi informasi pengetahuan bagi SD Sariputra. Informasi ini diharapkan membantu meningkatkan proses belajar mengajar di SD Sariputra. Sampel yang peneliti ambil dari siswa-siswi kelas 3 & 4 sebanyak 155 data siswa-siswi. Metode yang digunakan adalah Decision Tree dengan Algoritma C4.5 dengan 11 atribut. Analisis tersebut menghasilkan 23 Rule. Rule tersebut dihasilkan dari hasil akhir bentuk Decision Tree.

- (f.) Penelitian dilakukan oleh (Rizky Haqmanullah Pambudi, Setiawan, and Indriati 2018) dengan judul “Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal”. Penelitian beranggapan faktor eksternal berpengaruh pada kegagalan siswa dalam menyelesaikan bidang studi khususnya bidang studi matematika. Metode yang dipakai oleh peneliti adalah algoritma C4.5 yang merupakan salah satu metode data mining. Tujuan peneliti adalah memprediksi kemampuan siswa dalam menyelesaikan bidang studi dilihat dari faktor eksternal siswa. Metode algoritma C4.5 digunakan peneliti untuk mengetahui tingkat akurasi prediksi kemampuan siswa sekolah menengah. Hasil yang dikeluarkan dari pengujian dan analisis menunjukkan bahwa Algoritma Decision Tree C4.5 akurat diterapkan untuk prediksi nilai akhir siswa sekolah menengah dengan memiliki tingkat akurasi 60%.
- (g.) Penelitian dilakukan oleh (Malik and Faid 2016) dengan judul “PREDIKSI PRESTASI SISWA SMP NURUL JADID MENGGUNAKAN ALGORITMA C4.5”. Penelitian bertujuan untuk membangun sistem prediksi Prestasi siswa yang lebih akurat dalam memudahkan dalam pengambilan keputusan untuk pembinaan siswa lebih dalam lagi tentang prestasi yang ada pada Lembaga Smp Nurul Jadid. Peneliti mengumpulkan data primer dengan cara wawancara langsung dengan pihak Kepala Sekolah Lembaga Nurul Jadid. Peneliti menganalisis dengan melakukan training pada data beberapa tahun 2015/2016 guna membangun model C4.5. Karenanya, peneliti menyimpulkan bahwa data mining dengan teknik C4.5 dapat menghasilkan prediksi prestasi yang lebih tepat sehingga dapat membantu Lembaga Smp Nurul Jadid dalam pengambilan keputusan.
- (h.) Penelitian dilakukan oleh (Surianti. S 2020) dengan judul “Analisa Prediksi Jumlah Kelulusan Siswa Paket B Dengan Menggunakan Algoritma C4.5 Pada Pusat Kegiatan Belajar Masyarakat Hati Nurani Baru (PKBM Hanuba) Medan”. Peneliti mengungkapka perlunya dilakukan prediksi pada kelulusan siswa paket B dengan tujuan untuk memprediksi kelulusan untuk tahun berikutnya. Tema yang diambil dalam penelitian ini adalah data mining, dimana data mining adalah serangkaian proses untuk menggali nilai tambah dengan mencari pola atau aturan tertentu dari data yang besar untuk mendapatkan sebuah pengetahuan atau informasi baru yang selama ini tidak diketahui secara manual. Peneliti mengharapkan dapat dijadikan suatu pola keputusan yang nantinya akan digunakan dalam menganalisa prediksi jumlah kelulusan siswa paket B. Metode yang digunakan

peneliti adalah algoritma C4.5 karena digunakan dalam menganalisa jumlah kelulusan siswa paket B berdasarkan kriteria yang akan ditetapkan sebelumnya dan selanjutnya menghasilkan suatu pohon keputusan, dari pohon keputusan inilah akan diambil pengetahuan-pengetahuan baru berupa rules atau aturan-aturan yang akan dijadikan pola keputusan.

- (i.) Peneliti dilakukan oleh (Novika et al. 2021) dengan judul “Penerapan Data Mining Klasifikasi Tingkat Pemahaman Siswa Pada Pelajaran Matematika”. Tujuan dari penelitian ini adalah untuk mengklasifikasi konsep pemahaman siswa pada pelajaran Matematika. Metode yang dipakai dalam penelitian ini adalah Algoritma C4.5 dan peneliti menggunakan software RapidMiner sebagai alat bantu. Sumber data diambil dengan cara kuesioner tentang pemahaman matematika siswa kelas delapan di SMP Swasta Tamansiswa Tapian Dolok dengan menggunakan atribut yaitu minat siswa, cara belajar siswa, motivasi siswa, cara mengajar guru, media pembelajaran, dan sarana prasarana. Setelah dilakukan beberapa proses menghasilkan entropy dan gain atribut diperoleh 15 rule keputusan pemahaman matematika dengan 9 rules berstatus paham dan 6 rules berstatus tidak paham. Peneliti juga mendapatkan akurasi dari Algoritma C4.5 pada RapidMiner diperoleh akurasi sebesar 96.00%.

- (j.) Peneliti dilakukan oleh (Irnanda, Hartama, and Windarto 2021) dengan judul “Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi”. Tujuan dari penelitian adalah untuk mengklasifikasi faktor penyebab menurunnya prestasi belajar mahasiswa pada masa pandemi dengan menggunakan metode datamining C4.5. Sumber data didapatkan dengan cara melakukan wawancara dan pembagian angket terhadap mahasiswa semester 7 tahun ajaran 2020-2021 program studi sistem informasi. Peneliti menggunakan beberapa atribut yaitu: Cara Belajar (C1), Waktu Belajar (C2), Pemahaman Materi (C3), Pemberian Tugas (C4) dan Lingkungan (C5). Setelah melakukan perhitungan atribut Pemahaman Materi adalah atribut yang paling berpengaruh terhadap menurunnya prestasi belajar mahasiswa. Peneliti menggunakan software Rapidminer sebagai alat bantu dan memperoleh akurasi 97,5%.

Tabel 2. 8 Tinjauan Studi

No	Nama peneliti	Judul	permasalahan	Jurnal	kontribusi
1	Yetti Fitriani, Sarjon Defit, Gunadi Widi Nurcahyo (2021)	Prediksi Hasil Belajar Siswa Secara Daring pada Masa Pandemi COVID-19 Menggunakan Metode C4.5	Perubahan pola belajar berpengaruh terhadap hasil nilai yang di peroleh siswa dari situ peneliti melakukan klasifikasi data nilai siswa kelas XII Jurusan Multimedia di SMKN 2 Padang Panjang dan hasil klasifikasi dapat memprediksi hasil belajar siswa pada masa pandemi.	Jurnal Sistim Informasi dan Teknologi 2021 Vol. 3 No. 3 Hal: 118-125 e-ISSN: 2686-3154 http://jsisfotek.upi.yptk.ac.id	Metode yang efektif dalam melakukan prediksi dan dapat menghitung prediksi hasil belajar siswa
2	Septri Wanti Siahaan, Kristin D R Sianipar, P.P.P.A.N.W Fikrul Ilmi R.H Zer, Dedy Hartama (2020)	Penerapan Algoritma C4.5 dalam Menentukan Faktor yang Dapat Meningkatkan Kemampuan Bahasa Inggris pada Mahasiswa	Indonesia merupakan salah satu negara yang menerapkan Bahasa Inggris sebagai mata pelajaran atau mata kuliah wajib yang dipelajari oleh siswa/i dan mahasiswa/i. Mahasiswa/i wajib mengetahui bahwa mempelajari Bahasa Inggris sangat dibutuhkan untuk meningkatkan kompetensi di masa depan.	Jurnal Eksplora Informatika Vol. 10, No. 1, September 2020 p-ISSN: 2089-1814; e-ISSN: 2460-3694; DOI: 10.30864/eksploa v10i1.396 https://eksplora.stikom-bali.ac.id/index.php/eksplora/article/view/396	Mengetahui permasalahan yang ada dan metode yang efektif untuk menentukan faktor yang dapat meningkatkan kemampuan bahasa inggris pada mahasiswa
3	Khairunnissa Fanny Imanda, Agus Perdana Windarto (2020)	Penerapan Klasifikasi C4.5 Dalam Meningkatkan Kecakapan Berbahasa Inggris dalam Masyarakat	Menguasai bahasa internasional sangat diperlukan dan dibutuhkan saat ini. maka sudah sepatutnya generasi bangsa semakin maju dalam hal ilmu pengetahuan dan teknologi yang didukung dengan penguasaan.	Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) ISBN: 978-602-52720-7-3 http://seminar-id.com/prosiding/index.php/sainteks/article/view/452	Mengetahui permasalahan yang ada dan metode yang efektif dalam menghitung kecakapan berbahasa inggris dalam masyarakat

No	Nama peneliti	Judul	permasalahan	Jurnal	kontribusi
4	M.A.Abdillah, Arief Setyanto, Sudarmawan (2020)	Implementasi Decision Tree Algoritma C4.5 Untuk Memprediksi Kesuksesan Pendidikan Karakter	Kesuksesan pendidikan karakter tentu di pengaruhi dengan banyak faktor dengan permasalahan ini peneliti ingin memprediksi kesuksean pendidikan karakter untuk mengetahui faktor yang berpengaruh besar terhadap pendidikan karakter siswa	Vol. XV Nomor 2 Juli 2020 – Jurnal ilmiah Teknologi Informasi ISSN: 1907-2430 https://jti.respati.ac.id/index.php/jurnaljti/article/view/349	Mengetahui Metode untuk memprediksi kesuksesan pendidikan karakter dan proses pengumpulan data sampai data tersebut dapat diolah
5	Oscario, Jasmir, Yudi Novianto (2019)	Penerapan Algoritma C4.5 Untuk Memprediksi Kecocokan Gaya Belajar Bagi Siswa Siswi Sekolah Dasar (Studi Kasus : SD Sariputra Jambi)	Masih banyak murid yang tidak fokus atau tidak memiliki keinginan untuk belajar. Dimana peneliti berasumsi kemungkinan terjadi akibat ketidakcocokan gaya belajar dan metode belajar guru tersebut.	Jurnal Ilmiah Sistem Informasi, Teknologi Informasi dan Sistem Komputer p-ISSN : 1907-6738 e-ISSN 2528-0082 http://ejournal.stikom-db.ac.id/index.php/processor	Mengetahui metode yang efektif untuk memprediksi kecocokan gaya belajar siswa sekolah dasar dan proses yang harus dilakukan dalam metode C 4.5
6	Rizky Haqmanullah Pambudi, Budi Darma Setiawan, Indriati (2018)	Penerapan Algoritma C4.5 Untuk Memprediksi Nilai Kelulusan Siswa Sekolah Menengah Berdasarkan Faktor Eksternal	Peneliti beranggapan faktor eksternal berpengaruh pada kegagalan siswa dalam menyelesaikan bidang studi khususnya bidang studi matematika.	Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN: 2548-964X Vol. 2, No. 7, Juli 2018, hlm. 2637-2643 http://iptiik.ub.ac.id	Mengetahui metode C4.5 efektif dalam memprediksi nilai kelulusan siswa berdasarkan faktor eksternal

No	Nama peneliti	Judul	permasalahan	Jurnal	kontribusi
7	Kamil Malik, Mochammad Faid (2016)	PREDIKSI PRESTASI SISWA SMP NURUL JADID MENGGUNAKAN ALGORITMA C4.5	Adanya siswa yang rajin tetapi tidak berprestasi dapat terjadi hal seperti ini bisa terjadi karena banyak faktor, untuk itu dibutuhkan kajian lebih mendalam lagi tentang hal ini. selain itu ada pula siswa yang memiliki nilai raport yang diatas rata-rata tidak lulus ujian nasional.	Nusantara Journal of Computers and its Applications Volume 2 – No.4, Desember 2016 http://journal.csnu.or.id/index.php/njca/article/view/11	Mengetahui metode yang dapat digunakan dalam memprediksi siswa berprestasi dan bagaimana proses algoritma C4.5 dari data awal hingga menghasilkan akurasi dan pohon keputusan
8	Surianti. S (2020)	Analisa Prediksi Jumlah Kelulusan Siswa Paket B Dengan Menggunakan Algoritma C4.5 Pada Pusat Kegiatan Belajar Masyarakat Hati Nurani Baru (PKBM Hanuba) Medan	Setiap tahunnya persen kelulusan setiap sekolah dapat berubah-ubah dan mungkin saja sama dengan tahun sebelumnya. Peneliti mengukapka perlunya dilakukan prediksi pada kelulusan siswa paket B dengan tujuan untuk memprediksi kelulusan untuk tahun berikutnya..	Jurnal Informasi dan Teknologi Ilmiah (INTI) ISSN 2301-9425 (Media Cetak) Volume 7, No 3, Juni 2020 Hal: 233-241 http://ejurnal.stmi.kbudidarma.ac.id/index.php/inti/article/view/2395	Mengetahui proses algoritma C4.5 dan pengoprasian software rapid miner dalam membuat pohon keputusan
9	Tri Novika, Poningsih, Harly Okprana, Agus Perdana Windarto, Hasudungan Siahaan (2021)	Penerapan Data Mining Klasifikasi Tingkat Pemahaman Siswa Pada Pelajaran Matematika	Matematika seringkali dianggap sebagai materi pelajaran yang sulit oleh siswa. karena itu di perlukan pengklasifikasian dalam tingkat pemahaman siswa pada mata pelajaran matematika	JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 5, Nomor 1, Januari 2021, Page 9-17 ISSN 2614-5278 , ISSN 2548-8368 https://ejurnal.stmik-budidarma.ac.id/index.php/mib	Mengetahui Algoritma yang tepat untuk melakukan klasifikasi tingkat pemahaman siswa dan proses dari data mining klasifikasi C4.5

No	Nama peneliti	Judul	permasalahan	Jurnal	kontribusi
10	Khairunnissa Fanny Irnanda, Dedy Hartama , Agus Perdana Windarto (2021)	Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi	Adanya penurunan prestasi belajar mahasiswa pada saat pandemi. Prestasi mahasiswa dapat turun karena banyak faktor dari sini peneliti melakukan analisa klasifikasi terhadap faktor faktor yang mempengaruhi penurunan prestasi belajar mahasiswa.	JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 5, Nomor 1, Januari 2021, Page 327-331 ISSN 2614-5278 (media cetak), ISSN 2548-8368 http://ejurnal.stmi-k-budidarma.ac.id/index.php/inti/article/view/2395	Mengetahui proses dari data mining dan algoritma C4.5 juga adanya penurunan prestasi belajar yang dialami oleh mahasiswa

Penelitian rujukan pada penelitian ini diambil berdasarkan kesamaan metode yang digunakan yaitu algoritma C 4.5 dan ruang lingkup penelitian yang diambil adalah seputar permasalahan yang ada pada sekolah. Perbedaan dengan jurnal diatas dengan penelitian yang ditulis berada pada atribut yang di pakai berupa nilai siswa dalam berbicara, menulis, mendengar, dan membaca dari matapelajaran bahasa inggris. Dengan sistem aplikasi yang berbasis website sebagai pengembangannya, lalu diuji kembali menggunakan black box.

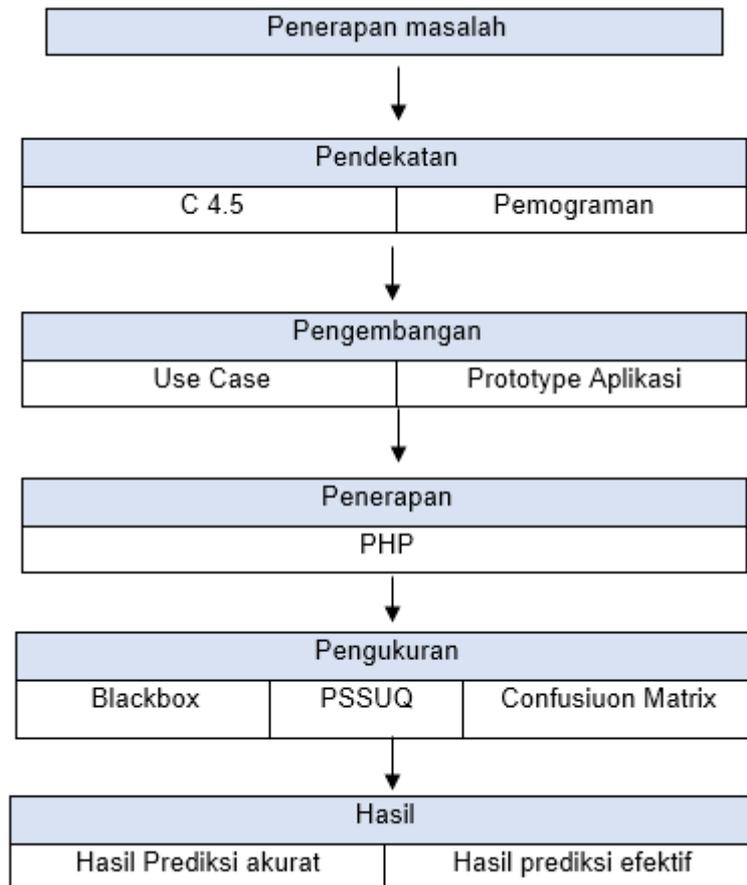
C. Kerangka Pemikiran

Penerapan Algoritma C 4.5 untuk memprediksi hasil nilai siswa.

Kerangka pemikiran pada gambar 2.4 dapat dijelaskan sebagai berikut :

1. Penerapan masalah mencakup fenomena yaitu rendahnya nilai bahasa inggris siswa dan belum adanya metode yang dipakai untuk memprediksi hasil nilai siswa
2. Pendekatan penelitian yaitu metode yang ditetapkan atau digunakan yaitu algoritma C4.5
3. Pengembangan yaitu tahap melakukan analisis dan perencanaan gambar untuk diagram activity, kemudian menentukan bahasa pemrograman dengan kontruksi PHP dan melakukan evaluasi serta validasi.
4. Penerapan yaitu memasukan rumus C4.5 dan data yang sudah kedalam website menggunakan php
5. Pengukuran yaitu melakukan pengujian terhadap metode C4.5 yang ditetapkan ke sistem menggunakan confusion matrix, blackbox, dan PSSUQ.

6. Hasil yaitu sistem informasi menampilkan hasil rekomendasi Prediksi hasil nilai bahasa Inggris siswa.



Gambar 2. 1 Kerangka Pemikiran

D. Hipotesis Penelitian

Dalam teori data mining, algoritma C4.5 adalah salah satu algoritma yang dapat melakukan pengklasifikasian untuk memprediksi suatu peluang di masa depan berdasarkan dari pengalaman dimasa sebelumnya. Algoritma C4.5 sudah banyak dikenal dan digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru di penelitian-penelitian sebelumnya dengan permasalahan yang serupa. Permasalahan yang dihadapi oleh peneliti yaitu belum efektif dan belum akuratnya proses dalam prediksi status ketuntasan nilai bahasa Inggris siswa di sekolah menengah atas Berdasarkan hal tersebut maka dapat ditetapkan. hipotesis pada penelitian ini adalah penerapan Algoritma C4.5 diduga dapat memprediksi nilai bahasa Inggris di sekolah menengah atas.