

BAB II KERANGKA TEORITIS

A. Landasan Teori

1. Pengertian *Data Mining*

Menurut (Hermawati, 2013:3), data mining adalah proses menganalisis dan mengekstraksi pengetahuan secara otomatis menggunakan satu atau lebih teknik pembelajaran komputer (machine learning). Menurut (Turban, dkk. 2005:5) Data Mining adalah suatu istilah yang digunakan dalam menguraikan penemuan pengetahuan di dalam suatu database. Data mining ialah suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dalam pengetahuan yang terkait di berbagai database besar. Menurut (Han, Kamber, dan Pei, 2011) Istilah data *mining* kadang disebut juga *Knowledge Discovery in Database* (KDD). Istilah data mining sering dipakai, mungkin karena istilah ini lebih pendek dari *Knowledge Discovery in Database*, Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain, *Data mining* dianggap hanya sebagai suatu langkah penting dalam KDD. Proses KDD secara garis besar dapat dijelaskan sebagai berikut:

- (a).Pembersihan data adalah untuk menghilangkan noise dan data yang tidak konsisten;
- (b).Integrasi data adalah dimana beberapa sumber data dapat dikombinasikan. Sebuah *tren* populer di industri informasi adalah untuk melakukan pembersihan dan integrasi data sebagai langkah *preprocessing*, dimana data yang dihasilkan akan disimpan dalam *data warehouse*;
- (c). Seleksi data adalah suatu data yang relevan dengan tugas analisis yang diambil dari *database*;
- (d).Data transformasi adalah dimana terdapat suatu data dapat diubah dan digabung ke dalam bentuk yang sesuai untuk pertambangan dengan melakukan ringkasan atau agregasi operasi;
- (e).Data mining adalah proses esensial dimana metode cerdas diaplikasikan untuk mengekstrak data pola;

- (f).Evaluasi Pola adalah untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan;
- (g).Presentasi pengetahuan adalah dimana visualisasi dan teknik representasi pengetahuan digunakan untuk menyajikan pengetahuan hasil data mining kepada pengguna.

Tahap ini merupakan bagian dari berbagai bentuk *processing data*, dimana data dipersiapkan untuk *data mining*. Hal ini, menunjukkan bahwa data mining sebagai salah satu langkah dalam proses KDD, karena dapat mengungkap pola-pola tersembunyi yang digunakan untuk sebuah evaluasi.

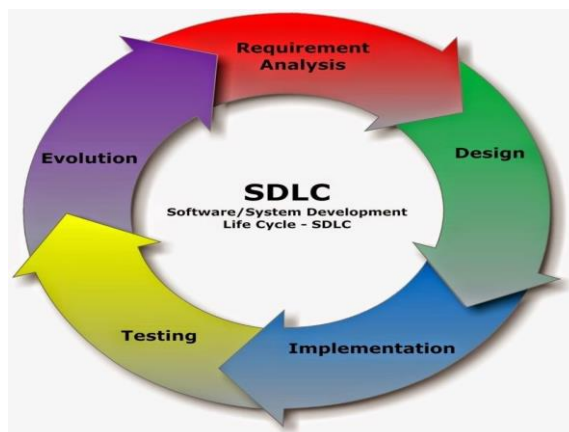
2. Pengertian Klasifikasi

Menurut (Kamber, 2006) Klasifikasi adalah sebuah proses untuk menemukan sekumpulan model yang dijelaskan kelas-kelas data, sehingga model tersebut bisa digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek. Maka jika ingin mendapatkan model, kita harus melakukan analisis terhadap data latih. Sedangkan data uji digunakan untuk mengetahui tingkat akurasi dan model yang telah dihasilkan. Klasifikasi dapat digunakan untuk memprediksi nama atau nilai dari suatu objek data. Klasifikasi adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui. Klasifikasi data terdiri dari dua langkah proses yaitu :

- (1). Pertama adalah *learning (fase training)*, dimana algoritma klasifikasi dibuat untuk menganalisis data training lalu dipresentasikan dalam bentuk rule klasifikasi.
- (2). Proses kedua adalah klasifikasi, dimana data tes digunakan untuk memperkirakan akurasi dari rule klasifikasi (Kamber, 2006). Proses klasifikasi didasarkan pada empat komponen (Gorunescu, 2011):
 - (a) *class*, variabel *dependen* yang berupa kategorikal yang merepresentasikan 'label' yang terdapat pada objek.
 - (b) *predictor*, variabel independen yang direpresentasikan oleh karakteristik (atribut) data.
 - (c) *training Dataset*, satu set data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.
 - (d) *testing Dataset*, berisi data baru yang akan diklasifikasikan oleh model yang telah dibuat dan akurasi klasifikasi dievaluasi.

3. Pengertian Pengembangan SDLC

Pendekatan sistem adalah sebuah metodologi. Metodologi merupakan satu cara yang direkomendasikan dalam melakukan sesuatu. Pendekatan sistem disebut juga metodologi dasar dalam memecahkan segala jenis masalah. Siklus hidup pengembangan sistem atau (System Development Life Cycle–SDLC) adalah aplikasi dari pendekatan sistem bagi pengembangan suatu sistem informasi (Raymond McLeod, 2007:199). SDLC (Siklus Hidup Pengembangan Sistem) disebut juga dalam rekayasa sistem dan rekayasa perangkat lunak, sebagai salah satu proses pembuatan dan perubahan sistem serta model dan metodologi yang digunakan untuk mengembangkan sistem-sistem tersebut. Maka Konsep ini umumnya lebih merujuk pada sistem komputer atau informasi. SDLC juga merupakan pola yang diambil untuk mengembangkan sistem perangkat lunak, yang terdiri dari tahap-tahap: rencana (*planning*), analisis (*analysis*), desain (*design*), implementasi (*implementation*), ujicoba (*testing*) dan pengelolaan (*maintenance*). System Development Life Cycle (SDLC) merupakan sebuah metodologi klasik yang digunakan untuk mengembangkan, memelihara dan menggunakan sistem informasi. Siklus hidup sistem itu sendiri disebut juga metodologi, tetapi polanya lebih dipengaruhi oleh kebutuhan untuk mengembangkan sistem yang lebih cepat. Sebuah pengembangan sistem yang lebih cepat dapat dicapai dengan peningkatan siklus hidup dan penggunaan peralatan pengembangan berbasis komputer.



**Gambar 2.1 Pola Melingkar dari Siklus Hidup Sistem
(Sumber: Raymond McLeod, Jr, 2007)**

Secara umum tahap-tahap dalam *System Development Life Cycle* (SDLC) terbagi dalam beberapa tahap :

(1). *Planning*

Merupakan tahap awal dari pengembangan sistem, tahap ini bertujuan untuk mengidentifikasi dan memprioritaskan sebuah sistem informasi apa yang akan dikembangkan, sasaran-sasaran yang ingin dicapai, jangka waktu pelaksanaan serta mempertimbangkan dana yang tersedia dan siapa yang melaksanakan.

(2). *Analysis*

Analisis system adalah penelitian atas sistem yang telah ada dengan tujuan untuk merancang suatu sistem baru atau memperbaharui sistem yang sudah ada.

(3). *Design*

Rancangan system adalah penentuan proses dan data yang diperlukan oleh sistem baru. Jika sistem ini berbasis komputer, rancangan dapat menyertakan spesifikasi jenis peralatan yang akan digunakan.

(4). *Implementation*

Penerapan merupakan kegiatan memperoleh dan mengintegrasikan sumber daya fisik dan konseptual yang menghasilkan suatu sistem yang bekerja. Pada tahapan ini dilakukan beberapa hal yaitu: *Coding, Testing, Instalasi*.

(5). *Operation and Support*

Tahapan penggunaan terdiri dari tiga langkah, yaitu :

- a. Menggunakan Sistem Pemakaian menggunakan sistem untuk mencapai tujuan yang diidentifikasi pada tahap suatu perancangan.
- b. Audit Sistem Setelah sistem baru mapan, peneliti formal dilakukan untuk menentukan seberapa baik sistem baru ini dapat memenuhi kriteria kinerja.
- c. Memelihara Sistem Selama manajer menggunakan sistem, berbagai modifikasi dibuat sehingga sistem terus memberikan dukungan yang diperlukan.

B. Pemahaman Dengan Teoritis

1. Pengertian *Metode Prototype*

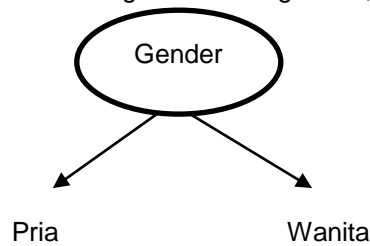
Metode *prototype* adalah metode yang dimulai dengan pengumpulan kebutuhan pengguna, dalam hal ini pengguna dari perangkat yang dikembangkan adalah prediksi prestasi belajar siswa. Kemudian membuat sebuah rancangan kilat yang selanjutnya akan dievaluasi kembali sebelum diproduksi secara benar. *Prototype* bukanlah merupakan sesuatu yang

lengkap, tetapi sesuatu yang harus dievaluasi dan dimodifikasi kembali. Segala perubahan dapat terjadi pada saat *prototype* dibuat untuk memenuhi kebutuhan pengguna dan pada saat yang sama memungkinkan pengembang untuk lebih memahami kebutuhan pengguna secara lebih baik (Roger S. Pressman, 2012). Pembuatan *prototype* dimulai dengan dilakukannya komunikasi antar tim pengembang perangkat lunak dengan para pelanggan. Tim pengembang perangkat lunak akan melakukan pertemuan-pertemuan dengan para stakeholder untuk mendefinisikan sasaran keseluruhan untuk perangkat lunak yang akan dikembangkan mengidentifikasi spesifikasi kebutuhan apapun yang saat ini diketahui dan menggambarkan dimana area-area definisi lebih jauh pada iterasi selanjutnya merupakan keharusan, *iterasi* pembuatan *prototype* direncanakan dengan cepat dan pemodelan (dalam bentuk "rancangan cepat") dilakukan. Suatu rancangan cepat berfokus pada representasi semua aspek perangkat lunak yang akan terlihat oleh pengguna akhir misalnya rancangan antar muka pengguna (*user interface*) atau (format tampilan) (Roger S. Pressman, 2012). Rancang cepat (*quick design*) akan memulai konstruksi pembuatan *prototype*, *prototype* kemudian akan diserahkan kepada para stakeholder dan kemudian akan melakukan evaluasi-evaluasi tertentu seperti *prototype* yang telah dibuat sebelumnya, kemudian akhirnya akan memberikan umpan balik yang akan digunakan untuk memperhalus spesifikasi kebutuhan. Sebuah iterasi akan terjadi saat *prototype* diperbaiki untuk memenuhi kebutuhan dari para stakeholder, sementara pada saat yang sama memungkinkan kita untuk lebih memahami kebutuhan apa yang kita kerjakan pada iterasi yang lalu.

2. Pengertian Metode Algoritma C4.5

Algoritma C4.5 merupakan sebuah pengembangan dari algoritma ID3 (*Iterative Dichotomiser 3*). Serangkaian perbaikan yang dilakukan pada algoritma ID3 mencapai puncaknya dengan menghasilkan sebuah system praktis dan berpengaruh untuk pembentukan pohon keputusan. Perbaikan tersebut meliputi metode untuk menangani *numeric attributes*, *missing values*, *noisy data*, dan aturan yang menghasilkan aturan dari *tree*. Menurut (Elisa, 2017:36) menyatakan bahwa Algoritma C4.5 merupakan sebuah algoritma yang sudah banyak dikenal dan digunakan untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. Maka, hasil dari proses klasifikasi yang berupa aturan-aturan dapat digunakan untuk memprediksi nilai atribut bertipe diskret dari record yang baru. Algoritma C4.5 sendiri merupakan pengembangan dari algoritma ID3, dimana pengembangan yang

dilakukan antara lain, bisa mengatasi missing value, bisa mengatasi continue data dan pruning.



Gambar 2. 2 Syarat pengujian fitur biner

Yang menjadi hal penting dalam induksi *decision tree* adalah bagaimana menyatakan syarat pengujian pada node. Ada 3 (tiga) kelompok penting dalam syarat pengujian node:

a. Fitur biner

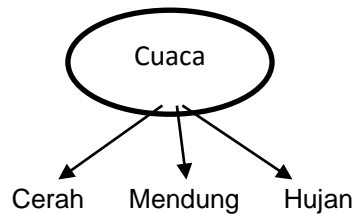
Fitur yang hanya mempunyai dua nilai berbeda disebut dengan fitur biner. Syarat pengujian ketika fitur ini menjadi node (akar maupun internal) hanya punya dua pilihan cabang. Contoh pemecahannya disajikan pada Gambar 2.2;

b. Fitur bertipe kategorikal

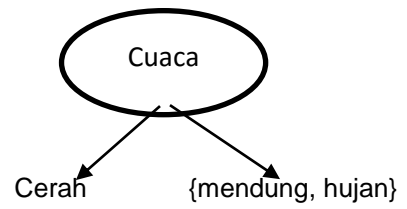
Untuk fitur yang nilainya bertipe kategorikal (nominal atau ordinal) bisa mempunyai beberapa nilai berbeda. Contohnya adalah fitur 'cuaca' mempunyai tiga nilai berbeda dan ini bisa mempunyai banyak kombinasi syarat pengujian pemecahan. Secara umum ada dua, yaitu pemecahan biner (*binary splitting*) dan (*multi splitting*). Kombinasinya disajikan seperti pada Gambar 2.3. Untuk pemecahan yang hanya membolehkan pemecahan biner, seperti algoritma CART, maka akan memberikan kemungkinan jumlah kombinasi pemecahan sebanyak 2^{k-1} , dimana k adalah jumlah nilai berbeda dalam fitur tersebut. Contoh pemecahannya disajikan pada Gambar 2.3. dan Gambar 2.4;

c. Fitur bertipe numerik

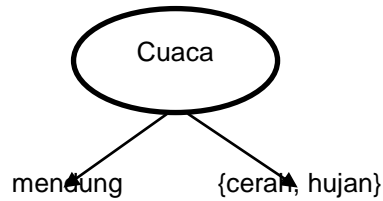
Untuk fitur bertipe numerik, syarat pengujian dalam node (akar maupun internal) dinyatakan dengan pengujian perbandingan ($A < v$) atau ($A \geq v$) dengan hasil biner, atau untuk multi dengan hasil berupa jangkauan nilai dalam bentuk $v_i \leq A < v_{i+1}$, untuk $i = 1, 2, \dots, k$. Untuk kasus pemecahan biner, maka algoritma akan memeriksa semua kemungkinan posisi pemecahan v dan memilih posisi v terbaik. Untuk cara multi, maka algoritma harus memeriksa semua kemungkinan jangkauan nilai *kontinyu*. Contoh pemecahan pada fitur numerik disajikan pada Gambar 2.5;



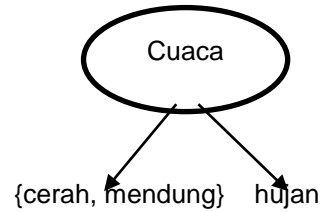
a. Multi splitting



b. Binary splitting

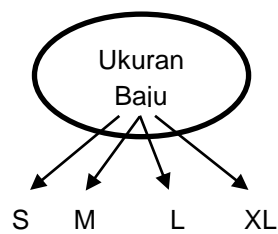


c. Binary splitting

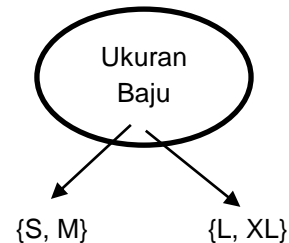


d. Binary splitting

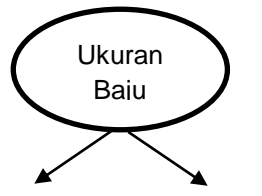
Gambar 2.3 Syarat pengujian fitur bertipe *nominal*



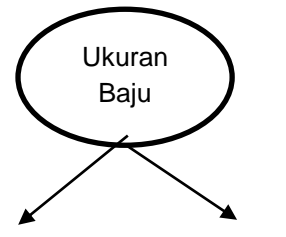
a. Multi splitting



b. Binary splitting

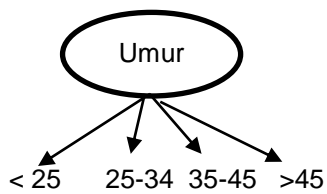


c. Binary splitting

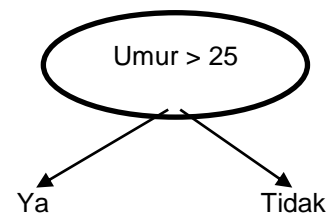


d. Binary splitting

Gambar 2.4 Syarat pengujian fitur bertipe *ordinal*



a. Multi splitting



b. Binary splitting

Gambar 2.5 Syarat pengujian fitur bertipe *numerik*

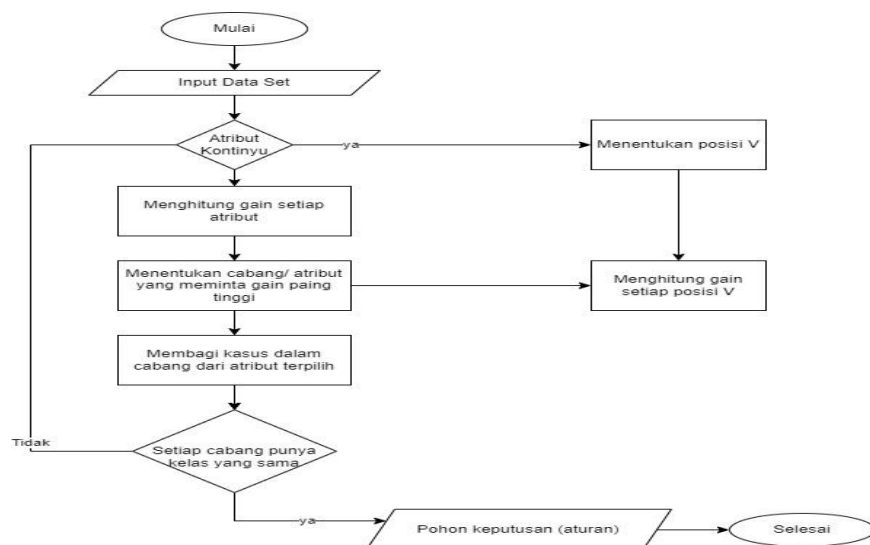
Kriteria yang paling banyak digunakan untuk memilih fitur sebagai pemecah dalam algoritma C4.5 adalah rasio gain, yang diformulasikan oleh persamaan berikut:

$$RasioGain(s,j) = \frac{Gain(s,j)}{SplitInfo(s,j)}$$

Persamaan diatas menyatakan nilai rasio gain pada fitur ke- j . $SplitInfo(s,j)$ didapat dari $SplitInfo(s,j) = -\sum_{i=1}^k p(v_i|s) \log_2 p(v_i|s)$ dimana k menyatakan jumlah pemecahan .

Berikut ini adalah algoritma C4.5 menggunakan flowcart yang disajikan pada Gambar 2.6 dan berserta penjelasannya :

- (a).Memasukan Data Set yang telah disediakan. Data Set adalah kumpulan objek dan atributnya.
- (a) Pada proses cek keputusan dilihat apakah data set tersebut termasuk Atribut Kontinu atau bukan, Atribut Kontinu adalah atribut yang mempunyai jangkauan real.
- (b) Jika iya maka akan menentukan posisi V, setelah itu menghitung gain setiap V. Jika tidak maka menghitung *gain* setiap atribut.
- (c) Setelah itu menentukan cabang/atribut yang memiliki *gain* paling tinggi.
- (d) Setelah itu membagi kasus dalam cabang dari atribut yang terpilih.
- (e) Jika setiap cabang mempunyai kelas yang sama, maka menghasilkan pohon keputusan. Jika tidak maka cabang atau atribut tersebut kembali melakukan cek atribut kontinu.



Gambar 2.6 Alur Algoritma C4.5

3. Pengertian *Unified Modeling Language* (UML)

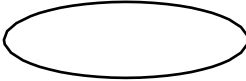
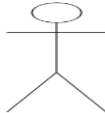
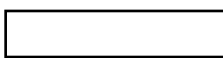

Menurut Sri Dharwiyanti, dari buku "Pengantar *Unified Modeling Language*, 2003, UML (*Unified Modeling Language*) adalah sebuah bahasa yang berdasarkan grafik/gambar untuk memvisualisasi, menspesifikasikan, membangun dari suatu sistem pengembangan software berbasis OO (*Object Oriented*). Pemodelan (modeling) adalah proses perancangan peranti lunak sebelum melakukan pengodean (coding). Peranti lunak tidak dapat dibuat secara asal melainkan banyak hal yang perlu diperhatikan semacam scalability, security, dan eksekusi yang robust pun dalam keadaan kurang efektif. Arsitektur perlu didefinisikan dengan jelas, agar *bug* mudah ditemukan dan diperbaiki. Keuntungan lain dari perencanaan arsitektur adalah dimungkinkannya pengguna kembali modul atau komponen untuk aplikasi peranti lunak lain yang membutuhkan fungsionalitas yang sama (Sugiarti, 2013:100).


Ada beberapa diagram dasar dalam UML, yaitu :

a. Model *Usecase Diagram*

Menurut (Sugiarti, 2018:111) Mengemukakan bahwa diagram usecase adalah sebuah interaksi antara satu atau lebih aktor dengan sistem yang akan dibuat, diagram usecase digunakan untuk mengetahui fungsi yang ada di dalam sistem dan yang berhak menggunakan kegunaannya. Dalam usecase diagram ada simbol-simbol yang sudah ditentukan :

Tabel 2.1 Simbol Use Case Diagram

Nama Komponen	Deskripsi	Gambar
<i>Use Case</i>	Menerangkan "apa" yang dikerjakan "sistem", bukan "bagaimana" sistem mengerjakannya	
<i>Actor</i>	Menggambarkan orang, sistem atau <i>external</i> entitas / <i>stakeholder</i> yang menyediakan atau menerima informasi dari system	
<i>Sistem Boundary</i>	Menggambarkan jangkauan system	
<i>Association</i>	Menggambarkan bagaimana <i>actor</i> terlibat dalam use case	

<i>Generalization</i>	Dibuat ketika ada sebuah keadaan yang lain / perlakuan khusus	
<i>Extend</i>	Perluasan dari <i>use case</i> lain jika kondisi atau syarat terpenuhi	<u><< extend >></u>


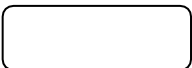
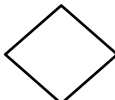
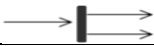

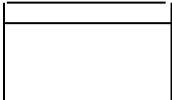
(Sumber : Sugiarti, 2018 : 111)

b. *Activity Diagram*

Menurut (Sugiarti, 2018:133) *diagram activity* adalah suatu dasar-dasar dalam pemrograman atau diagram yang menggambarkan *workflow* aktivitas dari sebuah sistem atau proses bisnis. Hal yang perlu diperhatikan bahwa diagram aktivitas sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem.

Berikut Simbol-Simbol di dalam *Activity Diagram* :

Tabel 2.2 Simbol Activity Diagram



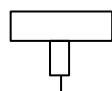
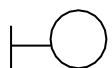
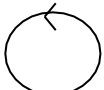
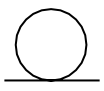
SIMBOL	NAMA	KETERANGAN
	<i>STATUS AWAL/INITIAL</i>	Status awal aktivitas sistem, sebuah diagram aktivitas memiliki sebuah status awal.
	<i>AKTIVITAS/ACTIVITY</i>	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja.
	<i>PERCABANGAN / DECISION</i>	Asosiasi percabangan dimana lebih dari satu aktivitas digabungkan menjadi satu.
	<i>PENGGABUNGAN/ JOIN</i>	Asosiasi penggabungan dimana lebih dari satu.
	<i>STATUS AKHIR/FINAL</i>	Status akhir yang dilakukan sistem, sebuah diagram aktivitas memiliki sebuah status baru.
	<i>SWIMLIE</i>	Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi.



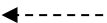
(Sumber : Sugiarti, 2018 : 133)

c. Diagram Interaksi (*Sequence Diagram*)

Diagram interaksi atau *interaction diagram* digunakan untuk memodelkan interaksi objek didalam sebuah *usecase* (proses). Diagram interaksi memperlihatkan interaksi yang memuat himpunan dari objek dan relasi yang terjadi antarobjek tersebut, termasuk juga bagaimana pesan mengalir diantara objek. Diagram sekuens atau *sequence diagram* merupakan gambaran *behavior* objek pada *usecase* dan message yang dikirimkan dan diterima antarobjek. Untuk menggambarkan diagram *sequence* maka harus diketahui objek-objek yang terlibat dalam sebuah *usecase* beserta metode-metode yang dimiliki class yang diinstansiasi menjadi objek itu (Sugiarti, 2018:131).

Tabel 2.3 Simbol Sequence Diagram

SIMBOL	NAMA	KETERANGAN
	<i>Actor</i>	Merepresentasikan entitas yang berada diluar sistem dan berinteraksi diluar sistem.
	<i>Lifeline</i>	Menghubungkan objek selama sequence (message dikirim atau diterima).
	<i>General</i>	Merepresentasikan entitas tunggal dalam sequence.
	<i>Boundary</i>	Berupa tepi dari sistem, seperti user interface dan alat yang berinteraksi dengan yang lain.
	<i>Control</i>	Elemen mengatur aliran dari informasi untuk sebuah skenario. Objek ini umumnya perilaku dan perilaku bisnis.
	<i>Entitas</i>	Elemen yang bertanggung jawab menyimpan atau informasi. Ini dapat berupa beans atau model object.

	<i>Activation</i>	Suatu titik dimana sebuah objek mulai berpartisipasi dalam sebuah sequence yang menunjukkan sebuah objek mengirim atau menerima objek.
	<i>Message Entry</i>	Berfungsi untuk menggambarkan pesan/hubungan antar objek yang menunjukkan urutan kejadian
	<i>Message to Self</i>	Simbol ini menggambarkan pesan/hubungan objek itu sendiri.


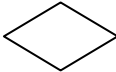
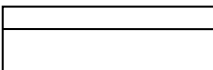
(Sumber : Sugiarti, 2018 : 131)


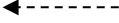
d. Diagram Kelas (*Class Diagram*)

Diagram kelas atau *class diagram* menggambarkan struktur sistem dari segi pendefinisian class – class yang akan dibuat untuk membangun sistem. Class diagram menggambarkan struktur dan deskripsi *class*, *package*, dan objek beserta hubungan satu sama lain seperti *containment*, pewarisan, asosiasi, dan lain–lain (Sugiarti,2018 :123).

Berikut simbol – simbol yang terdapat pada *class diagram* :

Tabel 2.4 Simbol Class Diagram

SIMBOL	NAMA	KETERANGAN
	<i>Generalization</i>	Hubungan dimana objek anak (descendent) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk (ancestor).
	<i>Nary Association</i>	Upaya untuk menghindari asosiasi dengan lebih dari 2 objek.
	<i>Class</i>	Himpunan dari objek-objek yang berbagi atribut serta operasi yang sama.

	<i>Collaboration</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu actor.
	<i>Realization</i>	Operasi yang benar-benar dilakukan oleh suatu objek

(Sumber : Sugiarti, 2018 : 123)

4. Pengertian Database

Menurut (Prayitno, 2015:2) "Database atau biasa disebut basis data merupakan kumpulan data yang saling berhubungan. Maka, data tersebut biasanya akan terdapat dalam tabel-tabel yang saling berhubungan satu sama lain, dengan menggunakan field/kolom pada tiap tabel yang ada". Menurut (Shalahuddin, 2015:43) menyatakan bahwa "basis data (*database*) merupakan salah satu bagian rekayasa perangkat lunak yang terkomputerisasi dan bertujuan utama sebagai media penyimpanan informasi agar dapat diakses dengan mudah dan cepat". Adapun kegunaan database yaitu sebagai alat untuk penyimpanan data yang saling berhubungan secara logikal dan deskripsi agar dapat diakses dengan lebih mudah dan cepat, dan terdapat tujuan utamanya yaitu untuk memelihara data yang sudah diolah dan membuat informasi tersedia pada saat dibutuhkan.

5. Pengertian Web Server (Apache)

Menurut (Firdaus, 2007:5) Web server adalah salah satu bentuk server yang khusus digunakan dalam menyimpan halaman website atau homepage. Apache disebut juga turunan dari web server yang dikeluarkan oleh NSCA yaitu NSCA HTTPd sekitar tahun 1995-an. Dan pada dasarnya, Apache adalah "APatCHy" (patch) dan pengganti dari NCSA HTTPd. Apache web server yaitu tulang punggung permintaan dari client yang menggunakan browser, seperti Netscape Navigator, Internet Explorer, Mozilla, lynx dan lain-lain. Cara web Server dalam berkomunikasi dengan kliennya menggunakan protokol HTTP (Hyper Text Transfer Protocol). Adapun kegunaan *web server* yaitu sebuah perangkat lunak (*software*) yang berfungsi menerima permintaan dari HTTP atau HTTPS pada client yang biasanya dikenal dengan nama *web browser* dan untuk mengirimkan kembali yang hasilnya dalam bentuk beberapa halaman web dan pada umumnya akan berbentuk dokumen HTML. Dalam bentuk sederhana *web server* juga akan mengirim data HTML kepada permintaan *web browser* sehingga akan terlihat sebuah tampilan website pada sebuah aplikasi pada biasanya. *web server* juga memiliki kegunaan untuk men-transfer atau memproses seluruh aspek pemberkasan berbagai data dalam sebuah halaman *web* yang kemudian memberikan hasil berupa teks, gambar dan beragam bentuk berkas lainnya.

6. Bahasa Pemrograman

a. Pengertian PHP

Menurut (Anhar, 2010:3) menyebutkan bahwa “PHP adalah script yang digunakan untuk membuat halaman website yang dinamis. Dinamis yaitu halaman yang akan ditampilkan pada saat halaman itu diminta oleh client, dan suatu bahasa pemrograman web server-side yang bersifat open source”. PHP (Hypertext Preprocessor) merupakan suatu bahasa pemrograman yang digunakan untuk menerjemahkan baris kode program menjadi kode suatu mesin yang dapat dimengerti oleh komputer yang berbasis server-side yang dapat ditambahkan ke dalam HTML. Bahasa PHP dapat dikatakan menggambarkan beberapa bahasa pemrograman seperti C, Java, dan perl serta mudah untuk dipelajari (Supono, 2016:3). Adapun sebuah pengertian lain PHP yaitu suatu akronim dari Hypertext Preprocessor, yaitu suatu bahasa pemrograman berbasis kode-kode(script) yang digunakan untuk mengolah suatu data dan mengirimkannya kembali ke web browser menjadi kode HTML, Sederhananya serverlah yang akan menerjemahkan skrip program, baru kemudian hasilnya akan dikirim kepada client yang melakukan permintaan. Sehingga kegunaan PHP dalam penelitian ini adalah suatu bentuk bahasa pemrograman untuk membuat/mengembangkan aplikasi berbasis web dan berbasis open source yang ditanamkan ke dalam script HTML.

b. Pengertian *Hypertext Markup Language* (HTML)

Menurut (Simarmata, 2010:52) Menjelaskan bahwa HTML adalah sebuah bahasa markup untuk menyebarkan informasi pada web. Maka, ketika merancang HTML, ide ini diambil dari Standard Generalized Markup Language (SGML). SGML merupakan suatu cara yang terstandarisasi dari pengorganisasian dan informasi yang terstruktur di dalam dokumen atau sekumpulan dokumen. Walaupun HTML tidak dengan mudah dipahami kebanyakan orang, ketika diterbitkan penggunaannya menjadi jelas pada penelitian ini. HTML adalah singkatan dari Hypertext Markup Language yaitu bahasa pemrograman standar atau salah satu format bahasa scripting yang digunakan dalam membuat sebuah halaman web yang kemudian dapat diakses untuk menampilkan berbagai informasi di dalam sebuah penjelajah web internet (browser) dan dapat juga digunakan sebagai link-link menuju halaman web yang lain dengan berdasarkan kode atau simbol tertentu yang sudah dimasukkan ke

dalam sebuah file atau dokumen, sehingga bisa ditampilkan pada layar komputer dan bisa dipahami oleh para pengguna internet.

C. Pengertian Prediksi Prestasi belajar

Menurut Winkel (1996:226) menjelaskan bahwa Prestasi belajar merupakan bukti keberhasilan yang telah dicapai oleh seseorang. Maka prestasi belajar merupakan hasil maksimum yang dicapai oleh setelah melaksanakan usaha-usaha belajar. Sedangkan Menurut (Suryabrata, 2006:297) mengemukakan bahwa “prestasi belajar sebagai nilai yang merupakan bentuk perumusan akhir yang diberikan oleh guru terkait dengan kemajuan prestasi siswa selama waktu tertentu”. Adapun komponen dalam prestasi belajar yaitu dengan bagaimana cara siswa melaksanakan proses belajar misalnya dengan cara mereka mempersiapkan belajar, mengikuti pelajaran, keaktifan disekolah, pola belajar, dan cara mengikuti ujian. Kualitas proses belajar mereka akan menentukan kualitas hasil belajar yang diperoleh, proses belajar yang baik akan menyebabkan berhasilnya belajar siswa, sedangkan proses belajar yang buruk akan menyebabkan kurang berhasilnya atau gagalnya belajar siswa dalam pembelajarannya. Prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan. Maka dari itu, prediksi prestasi belajar yaitu suatu saran yang dianjurkan bahwa adanya siswa yang mengalami penurunan atau kenaikan nilai dalam pembelajarannya, Sehingga perlunya prediksi prestasi belajar pada penelitian ini untuk membantu Kepala sekolah dan wali kelas agar melakukan tindakan dan langkah-langkah dalam meningkatkan kembali nilai prestasi belajar siswa di sekolah menengah pertama.

D. Tinjauan Pustaka

Penelitian rujukan merupakan acuan yang dibutuhkan seorang peneliti untuk melakukan penelitian. Penelitian rujukan pada penelitian ini diambil berdasarkan kesamaan metode yang digunakan yaitu Algoritma C4.5. Banyak penelitian yang menggunakan metode ini dalam berbagai kasus. Antara lain :

1. Menurut Khairunnisa Fanny Irnanda, Dedy Hartama, Agus Perdana Windarto (2019) didalam penelitiannya yang berjudul **“ANALISA KLASIFIKASI C4.5 TERHADAP FAKTOR PENYEBAB MENURUNNYA PRESTASI BELAJAR MAHASISWA PADA MASA PANDEMI”** menjelaskan tentang pada akhir tahun 2019 kita diguncangkan oleh suatu wabah yaitu virus covid-19 atau disebut juga dengan virus corona. Kementrian pendidikan menerapkan kebijakan dengan belajar dirumah saja dimana belajar dirumah saja ini dengan menggunakan sistem online atau daring.

Pada saat menggunakan sistem belajar online ini tidak selamanya berjalan dengan lancar pasti ada saja kendala-kendala yang dapat terjadi contohnya beberapa materi yang disampaikan kurang jelas dan mahasiswa juga kurang memahami apa yang sampaikan pengajar adanya kendala dalam proses pembelajaran menyebabkan terjadinya beberapa perubahan terhadap peserta didik yang pada akhirnya berpengaruh terhadap menurunnya motivasi dan prestasi belajarnya. Motivasi dan prestasi belajar merupakan dua aspek yang saling mempengaruhi. Prestasi belajar akan menjadi lebih optimal jika di dalamnya terkandung nilai motivasi yang tinggi, itulah kenapa jika seorang siswa berkurang motivasinya, maka prestasi belajarnya pun juga ikut terdampak, untuk itu peneliti menggunakan teknik data mining klasifikasi untuk menyelesaikan permasalahan menurunnya prestasi belajar mahasiswa. Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Teknik data mining yang digunakan adalah algoritma klasifikasi C4.5. Berdasarkan alasan tersebut peneliti menggunakan metode C4.5 untuk menganalisa faktor penyebab menurunnya prestasi belajar mahasiswa yang bertujuan untuk melihat atribut yang paling berpengaruh pada menurunnya prestasi belajar mahasiswa.

2. Menurut Yetti Fitriani, Sarjon Defit, Gunadi Widi Nurcahyo (2021) didalam penelitiannya yang berjudul **“PREDIKSI HASIL BELAJAR SISWA SECARA DARING PADA MASA PANDEMI COVID-19 MENGGUNAKAN METODE C4.5”** menjelaskan tentang penelitian dilakukan untuk memprediksi hasil belajar siswa pada masa pandemi COVID-19 sehingga hasil penelitian bisa digunakan untuk menjadi acuan dalam pengambilan kebijakan di sekolah. Metode C4.5 digunakan pada penelitian 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”.
3. Menurut Nurminton Silaban, Fevi Wira Citra, Mirna Yunita (2021) didalam penelitiannya yang berjudul **“PRESTASI BELAJAR MAHASISWA PENERIMA**

BIDIK MISI SELAMA MASA PANDEMI COVID-19” menjelaskan tentang penelitian ini bertujuan untuk melihat prestasi mahasiswa penerima bidik misi selama masa pandemi Covid-19 di Universitas Prof.Dr.Hazairin. Penelitian ini dilakukan menggunakan metode penelitian deskriptif. Populasi dan sampel yang digunakan pada penelitian ini adalah seluruh mahasiswa penerima bidik misi pada tahun akademik Ganjil 2019/2020 dan Genap 2019/2020 yang sedang menempuh perkuliahan sebanyak 219 orang mahasiswa. Teknik pengumpulan data adalah teknik dokumenter, alat yang digunakan berupa format isian yang dirancang sedemikian rupa sehingga dapat merekam semua data yang di perlukan. Data penelitian ini adalah data kuantitatif yang di analisis dengan menggunakan analisis secara deskriptif dan di kategorisasi berdasarkan Pedoman Akademik. Hasil penelitian, prestasi belajar mahasiswa penerima bidik misi pada tahun akademik 2019/2020 dilihat dari rata-rata IPK mahasiswa pada semester Ganjil adalah 3,31 dan rata-rata IPK mahasiswa pada semester Genap adalah 3,24. Nilai rata-rata prestasi akademik mahasiswa yang dilihat dari 219 mahasiswa penerima bidik misi terdapat penurunan indek prestasi sebesar 0,07 pada masa pandemik Covid-19. Penurunan indek prestasi belajar mahasiswa penerima bidik misi terlihat pada prodi Aquakultur di mana pada semester ganjil sebelum masa pandemik rata-rata prestasi akademiknya 3,8 dan rata-rata prestasi akademik selama masa pandemi Covid-19 adalah 2,7. Simpulan, persentase IPK rata-rata mahasiswa penerima bidik misik selama masa pandemi Covid-19 di Universitas Prof.Dr.Hazairin pada kategori sangat memuaskan atau di atas 2,75.

4. Menurut Ariansyah, Mira Kusmira (2021) didalam penelitiannya yang berjudul **“ANALISIS SENTIMEN PENGARUH PEMBELAJARAN DARING TERHADAP MOTIVASI BELAJAR DI MASA PANDEMI MENGGUNAKAN NAIVE BAYES DAN SVM”** menjelaskan tentang penelitian ini menggunakan metode machine learning untuk melakukan klasifikasi pendapat pengguna Twitter mengenai pengaruh pembelajaran daring terhadap motivasi belajar dan menentukan apakah pendapat pengguna Twitter tersebut memiliki sentimen positif atau negatif. Terdapat beberapa algoritma pada metode klasifikasi yaitu Naive bayes dan Support Vector Machine (SVM). Naïve Bayes adalah salah satu algoritma teknik data mining yang menerapkan teori Bayes dalam klasifikasi dimana algoritma ini mengasumsikan bahwa atribut suatu objek bersifat independen atau bebas. Metode Naïve Bayes telah banyak digunakan dalam penelitian mengenai text

mining, beberapa kelebihan Naïve Bayes diantaranya adalah algoritma sederhana tapi memiliki akurasi yang tinggi. Teknik SVM berakar pada teori pembelajaran statistik dan telah menunjukkan hasil empiris yang menjanjikan dalam berbagai aplikasi praktis dari pengenalan digit tulisan tangan sampai kategorisasi teks. SVM juga bekerja sangat baik pada data dengan banyak dimensi dan menghindari kesulitan dari permasalahan dimensionalitas. Berdasarkan penelitian sebelumnya, naïve bayes memiliki kelebihan yang dapat menangani data dengan jumlah yang besar dan tidak mempengaruhi irrelevant attribute. Naïve Bayes juga memiliki kelebihan dalam kesederhanaannya. SVM memiliki kelebihan dimana tingkat akurasi yang lebih baik dibandingkan Naïve Bayes. Fitur yang berbeda dari pendapat pengguna Twitter mengenai pengaruh pembelajaran daring terhadap motivasi belajar dimasa pandemi dengan algoritma Naive Bayes dan Support Vector Machine untuk memberikan hasil yang terbaik. Selanjutnya kedua algoritma tersebut dibandingkan untuk mendapatkan hasil berdasarkan accuracy, precision, recall, dan error. Penelitian terkait mengenai keberlanjutan pembelajaran daring dengan algoritma Support Vector Machine menghasilkan accuracy sebesar 86%, precision 87,38% dan recall 85,02%.

5. Menurut Andi Roy, Arga Ramadhana, (2021) didalam penelitiannya yang berjudul **“IMPLEMENTASI DAN ANALISIS ALGORITMA POHON KEPUTUSAN UNTUK MEMPREDIKSI PRESTASI SISWA”** menjelaskan tentang penelitian ini dilakukan dengan tujuan untuk membuat prediksi prestasi belajar siswa atribut yang digunakan dalam penelitian ini diantaranya status sosial ekonomi orang tua, motivasi, kedisiplinan siswa dan prestasi masa lalu menggunakan metode pohon keputusan dengan algoritma C4.5. Pendekatan penelitian yang digunakan adalah kuantitatif. Subyek penelitian ini adalah siswa tingkat X SMK Negeri 3 Parepare berjumlah 349 siswa. Teknik pengumpulan data yang digunakan adalah dokumentasi dan angket. Hasil penelitian ini menunjukkan bahwa evaluasi dari data test set, di dapatkan bahwa nilai keakuratan dalam memprediksi prestasi siswa menunjukkan hasil yaitu 98,85% dan error 1.15% tingginya tingkat akurasi dapat disimpulkan bahwa algoritma C4.5 memiliki kinerja yang cukup baik dalam membentuk Aturan-aturan. Dengan kata lain algoritma C4.5 bekerja dengan baik dan bisa di terapkan untuk memprediksi prestasi siswa.
6. Menurut N.Priya Dharsinni, Aisyah Hikmasari Sitepu, Rezz Youan Syuhada, Damanik Barasa, Andy Christanto Wijaya, (2021) didalam penelitiannya yang berjudul **“KENDALA PEMBELAJARAN MOODLE BERBASIS WEB TERHADAP MINAT**

BELAJAR SISWA MENGGUNAKAN ALGORITMA C4.5 SELAMA PANDEMI COVID-19

menjelaskan tentang penelitian ini bertujuan untuk membuat Sistem Pembelajaran pada masa pandemi Covid-19 beralih dari pembelajaran luar jaringan menjadi pembelajaran dalam jaringan yang membuat banyak kampus menggunakan berbagai platform E-Learning . Namun kebanyakan kampus menggunakan Pembelajaran Moodle Web-Based Learning dikarenakan menyediakan banyak fitur yang dapat mendukung dosen dan mahasiswa dalam proses pembelajaran online serta dapat diakses melalui laptop atau smartpone. Masalahnya, beberapa mahasiswa mengalami kendala dalam mengikuti model pembelajaran ini yang mempengaruhi naik turunnya minat belajar mahasiswa sehingga perlu ditemukan faktor kendala yang menjadi hambatan mahasiswa selama mengikuti pembelajaran Moodle Web-Based Learning. Algoritma C4.5 menghasilkan pohon keputusan yang dapat digunakan untuk memprediksi hasil yang bagus dan informasi yang tepat. Tujuan penelitian ini adalah menemukan hubungan kendala yang dialami mahasiswa selama mengikuti model Pembelajaran Moodle Web-Based Learning terhadap minat belajar mahasiswa menggunakan algoritma C4.5. Hasil penelitian menunjukkan kendala utama yang mempengaruhi menurunnya minat belajar mahasiswa dipengaruhi oleh fitur pembelajaran yang digunakan oleh dosen pada saat pembelajaran online kurang lengkap, kualitas jaringan kurang baik, mahasiswa menganggap pembelajaran Moodle Web-Based Learning kurang menarik sedangkan naiknya minat belajar mahasiswa dipengaruhi oleh fitur pembelajaran yang digunakan oleh dosen pada saat pembelajaran online sangat lengkap, kualitas jaringan yang baik, mahasiswa menggunakan laptop atau komputer dalam mengikuti pembelajaran Moodle Web-Based Learning dan mahasiswa menganggap pembelajaran Moodle Web-Based Learning menarik.

7. Menurut Nadiya Hijriana, Riadhul Muttaqin (2016) didalam penelitiannya yang berjudul **“PENERAPAN METODE DECISION TREE ALGORITMA C4.5 UNTUK KLASIFIKASI MAHASISWA BERPRESTASI”** menjelaskan tentang penelitian ini bertujuan untuk permasalahan yang sering muncul pada Mahasiswa yaitu dengan membludaknya jumlah mahasiswa yang ada di FTI, maka data-data akademik yang dimiliki juga semakin bertambah banyak. Apabila tidak diolah dengan baik dan benar hanya akan menjadi tumpukan data yang tidak berguna dan tidak bermanfaat. Berdasarkan permasalahan tersebut, maka peneliti akan mengklasifikasikan mahasiswa Prodi Teknik Informatika yang berprestasi dengan menggunakan metode Decision Tree Algoritma C4.5. Tujuan dari penelitian ini

adalah untuk mengklasifikasi mahasiswa berprestasi yang ada di Prodi Teknik Informatika FTI UNISKA dimana hasil penelitian tersebut dapat digunakan sebagai alat bantu bagi Fakultas untuk pengembangan akademis dalam hal pembimbingan dan pengarahan yang lebih intensif bagi mahasiswa IPK-nya kurang memuaskan.

8. Menurut Kamil Malik, Mochammad Faid (2016) didalam penelitiannya yang berjudul **“PREDIKSI PRESTASI SISWA SMP NURUL JADID MENGGUNAKAN ALGORITMA C4.5”** menjelaskan tentang penelitian ini bertujuan untuk mendapatkan data yang akurat sehingga dapat memudahkan dalam pengambilan keputusan untuk pembinaan siswa lebih dalam lagi terkait prestasi yang ada pada Lembaga SMP Nurul Jadid tersebut. Pengumpulan data primer dilakukan berdasarkan wawancara langsung dengan pihak Kepala Sekolah Lembaga Nurul Jadid. Analisis dilakukan dengan melakukan training pada data beberapa tahun 2015/2016 untuk membangun model C4.5. Karenanya, dapat disimpulkan bahwa data mining dengan teknik C4.5 dapat menghasilkan prediksi prestasi yang lebih tepat sehingga dapat membantu Lembaga Smp Nurul Jadid dalam pengambilan keputusan. Data mining dalam penelitian ini hanya berfungsi untuk mencari sebuah pola agar kita mengetahui apa yang menjadi kendala terhadap siswa untuk bisa berprestasi, hal ini dilakukan karena adanya data yang diljin/Kepentingan cukup untuk menjadi sampel dalam penelitian.
9. Menurut Kiki Aidi Saputra, Jaya Tata Hardinata, Muhammad Ridwan Lubis, Sundari Retno Andani, Ilham Syahputra Saragih, (2020) didalam penelitiannya yang berjudul **“KLASIFIKASI ALGORITMA C4.5 DALAM PENERAPAN TINGKAT KEPUASAN SISWA TERHADAP MEDIA PEMBELAJARAN ONLINE”** menjelaskan tentang penelitian ini bertujuan untuk mengetahui pengaruh tingkat kepuasan siswa terhadap media pembelajaran online di SMA YPI Swasta Dharma Budi. Media internet memiliki peranan penting untuk belajar bagi kalangan siswa. Populasi penelitian ini menerapkan Algoritma C4.5 dalam penerapan tingkat kepuasan siswa terhadap media pembelajaran online. Jenis penelitian ini yaitu Klasifikasi dengan konsep data mining dengan melibatkan sebanyak 100 data siswa di SMA YPI Swasta Dharma Budi yang dikategorikan dengan Puas dan Tidak Puas. Dalam pengambilan data menggunakan kuesioner dan variabel yang digunakan dalam penelitian ini ada empat yaitu minat belajar, cara belajar, sajian pembelajaran, media pembelajaran. Metode yang digunakan dalam penelitian ini yakni metode klasifikasi.

10. Menurut Astrid Noviriandini, Nurajijah (2019) didalam penelitiannya yang berjudul **“ANALISIS KINERJA ALGORITMA C4.5 dan NAIVE BAYES UNTUK MEMPREDIKSI PRESTASI SISWA SEKOLAH MENENGAH KEJURUAN”** , menjelaskan tentang penelitian ini menginformasikan kepada siswa dan guru untukantisipasi dini dalam mengikuti masa belajar agar mendapatkan hasil belajar yang maksimal. Metode yang digunakan adalah algoritma decision tree C4.5 dan algoritma Naïve Bayes. Tujuan dari penelitian ini adalah melakukan komparasi dan evaluasi model pohon keputusan C4.5 sebagai algoritma terpilih dan Naïve Bayes untuk mengetahui algoritma yang memiliki keakuratan lebih tinggi dalam memprediksi prestasi siswa. Prestasi belajar dapat diukur dari nilai raport. Setelah dilakukan komparasi dari kedua algoritma tersebut, maka didapatkan hasil prediksi prestasi belajar. Hasil penelitian menunjukkan bahwa algoritma Naïve Bayes memiliki nilai akurasi sebesar 95,67% dan nilai AUC sebesar 0,999 termasuk kedalam Excellent Clasification, untuk algoritma C4.5 memiliki nilai akurasi sebesar 90,91% dan nilai AUC sebesar 0,639 termasuk kedalam keadaan Poor Clasification. Dengan demikian algoritma Naïve Bayes dapat memprediksi prestasi belajar siswa dengan lebih baik.

Tabel 2.5 Tinjauan Pustaka

No.	Peneliti/Tahun	Judul	Sumber	Kontribusi / Kelemahan
1.	(Khairunnisa Fanny Irnanda,Dedy Hartama,Agus Perdana Windarto,2019)	Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi	Irnanda, K. F., Hartama, D., & Windarto, A. P. (2021). Analisa Klasifikasi C4. 5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi. https://ejurnal.stmik-budidarma.ac.id/ind	Kontribusi dalam penelitian ini adalah pengambilan variabel data angket untuk dijadikan variabel yang akan digunakan. Kelemahan dalam penelitian ini yaitu belum adanya perhitungan nilai akurasi.

			ex.php/mib/article/download/2763/1892	
2.	(Yetti Fitriani, Sarjon Defit, Gunadi Widi Nurcahyo, 2021)	Prediksi Hasil Belajar Siswa Secara Daring Pada Masa Pandemi Covid-19 Menggunakan Metode C4.5	Fitriani, Y., Defit, S., & Nurcahyo, G. W. (2021). Prediksi Hasil Belajar Siswa Secara Daring pada Masa Pandemi COVID-19 Menggunakan Metode C4. 5. https://jsisfotek.org/index.php/JSisfotek/article/view/54	Kontribusi dalam penelitian ini adalah Pengambilan data nilai PTS untuk dijadikan variabel yang akan digunakan. Kelemahan dalam penelitian ini yaitu hanya menggunakan 3 variabel.
3.	(Nurmintan Silaban, Fevi Wira Citra, Mirna Yunita, 2021).	Prestasi Belajar Mahasiswa Penerima Bidik Misi Selama Masa Pandemi Covid-19	Silaban, N., Citra, F. W., & Yunita, M. (2021). Prestasi Belajar Mahasiswa Penerima Bidik Misi Selama Masa Pandemi Covid-19. https://journal.ipm2kpe.or.id/index.php/JOEAI/article/view/2013	Kontribusi dalam penelitian ini adalah pengumpulan data yang menggunakan teknik dokumenter Metode penelitian yang digunakan yaitu metode penelitian deskriptif. Kelemahan dalam penelitian ini yaitu tidak terdapat cara perhitungan algoritma C4.5

4.	(A.Ariansyah, M Kusmira,2021)	ANALISIS SENTIMEN PENGARUH PEMBELAJARAN DARING TERHADAP MOTIVASI BELAJAR DI MASA PANDEMI MENGGUNAKAN NAIVE BAYES DAN SVM	Ariansyah, A., & Kusmira, M. (2021). ANALISIS SENTIMEN PENGARUH PEMBELAJARAN DARING TERHADAP MOTIVASI BELAJAR DI MASA PANDEMI MENGGUNAKAN NAIVE BAYES DAN SVM. https://journal.lppm.unindra.ac.id/index.php/Faktor_Exacta/article/view/10325	Kontribusi dalam penelitian ini adalah pengambilan uji hasil yang menggunakan confusion matrix. Kelemahan dalam penelitian ini yaitu tidak terdapat cara perhitungan algoritma C4.5
5.	(Andi Roy,Arga Ramadhana, 2021)	Implementasi Dan Analisis Algoritma Pohon Keputusan Untuk Memprediksi Prestasi Siswa	Roy, A., & Ramadhana, A. (2021). IMPLEMENTASI DAN ANALISIS ALGORITM APOHON KEPUTUSAN UNTUK MEMPREDIKSI PRESTASI SISWA. https://adoc.pub/jurnal-implementasi-data-mining-dengan-algoritma-c45-untuk-m.html	Kontribusi dalam penelitian ini adalah pengambilan variabel keadaan status sosial ekonomi orang tua untuk dijadikan variabel yang akan digunakan. Kelemahan dalam penelitian ini adalah proses output dari penentuan pola pada training data yang digunakan

				yaitu perangkat lunak bantu WEKA.
6.	(DHARSHINNI, N. Priya, Sitepu, A. H., Syuhada, R. Y., Barasa, D., & Wijaya, A. C. ,2021)	Kendala Pembelajaran Moodle Berbasis Web Terhadap Minat Belajar Siswa Menggunakan Algoritma C4.5 Selama Pandemi Covid-19	DHARSHINNI, N. P., Sitepu, A. H., Syuhada, R. Y., Barasa, D., & Wijaya, A. C. (2021). Kendala Pembelajaran Moodle Berbasis Web Terhadap Minat Belajar Siswa Menggunakan Algoritma C4.5 Selama Pandemi Covid-19. https://ojs.uma.ac.id/index.php/jite/article/download/5301/3416/20848	Kontribusi dalam penelitian ini adalah pengambilan uji hasil yang menggunakan kuesioner. Kelemahan dalam penelitian ini adalah nilai akurasi yang masih berada pada angka 74%, maka masih sangat dimungkinkan untuk dapat dilakukan penelitian selanjutnya untuk meningkatkan nilai akurasi.
7.	(Nadiya Hijriana,Riadhul Muttaqin, 2016)	Penerapan Metode Decision Tree Algoritma C4.5 Untuk Klasifikasi Mahasiswa Berprestasi	Hijriana, N., & Muttaqin, R. (2016). PENERAPAN METODE DECISION TREE ALGORITMA C4. 5 UNTUK KLASIFIKASI MAHASISWA BERPRESTASI. <i>htt</i>	Kontribusi dalam penelitian ini adalah pengambilan metode algoritma C4.5 dan cara perhitungan algoritma C4.5. Kelemahan dalam

			ps://ojs.uniska-bjm.ac.id/index.php/JST/article/view/651	penelitian ini adalah nilai akurasi yang masih berada pada angka 71%, maka masih sangat dimungkinkan untuk dapat dilakukan penelitian selanjutnya untuk meningkatkan nilai akurasi.
8	(Kamil Malik, Mochammad Faid, 2016)	Prediksi Prestasi Siswa Smp Nurul Jadid Menggunakan Algoritma C4.5	Malik, K., & Faid, M. (2016). Prediksi prestasi siswa smp nurul jadid menggunakan algoritma C4.5. https://repository.unuja.ac.id/id/eprint/7/9/Kamil_PREDIKSI_PRESTASI_SISWA_SMP.pdf .	Kontribusi dalam penelitian ini adalah mengetahui metode yang dapat digunakan dalam memprediksi siswa berprestasi dan bagaimana algoritma C4.5 dari data awal hingga menghasilkan akurasi dan pohon keputusan. Kelemahan dalam penelitian ini adalah belum adanya proses perhitungan nilai

				akurasi.
9.	(KA.Saputra, JT.Hardinata,MR.Lubis,Sundari R.A,Ilham S. 2020)	Klasifikasi Algoritma C4.5 Dalam Penerapan Tingkat Kepuasan Siswa Terhadap Media Pembelajaran Online	Saputra, K. A., Hardinata, J. T., Lubis, M. R., Andani, S. R., & Saragih, I. S. (2020). Klasifikasi Algoritma C4. 5 Dalam Penerapan Tingkat Kepuasan Siswa Terhadap Media Pembelajaran Online. https://djournal.com/klik/article/view/120	Kontribusi dalam penelitian ini adalah mengetahui metode yang efektif dalam menghitung tingkat kepuasan dan proses perhitungan yang dilakukan pada metode C4.5 Kelemahan dalam penelitian ini adalah tidak terdapatnya perhitungan nilai akurasi.

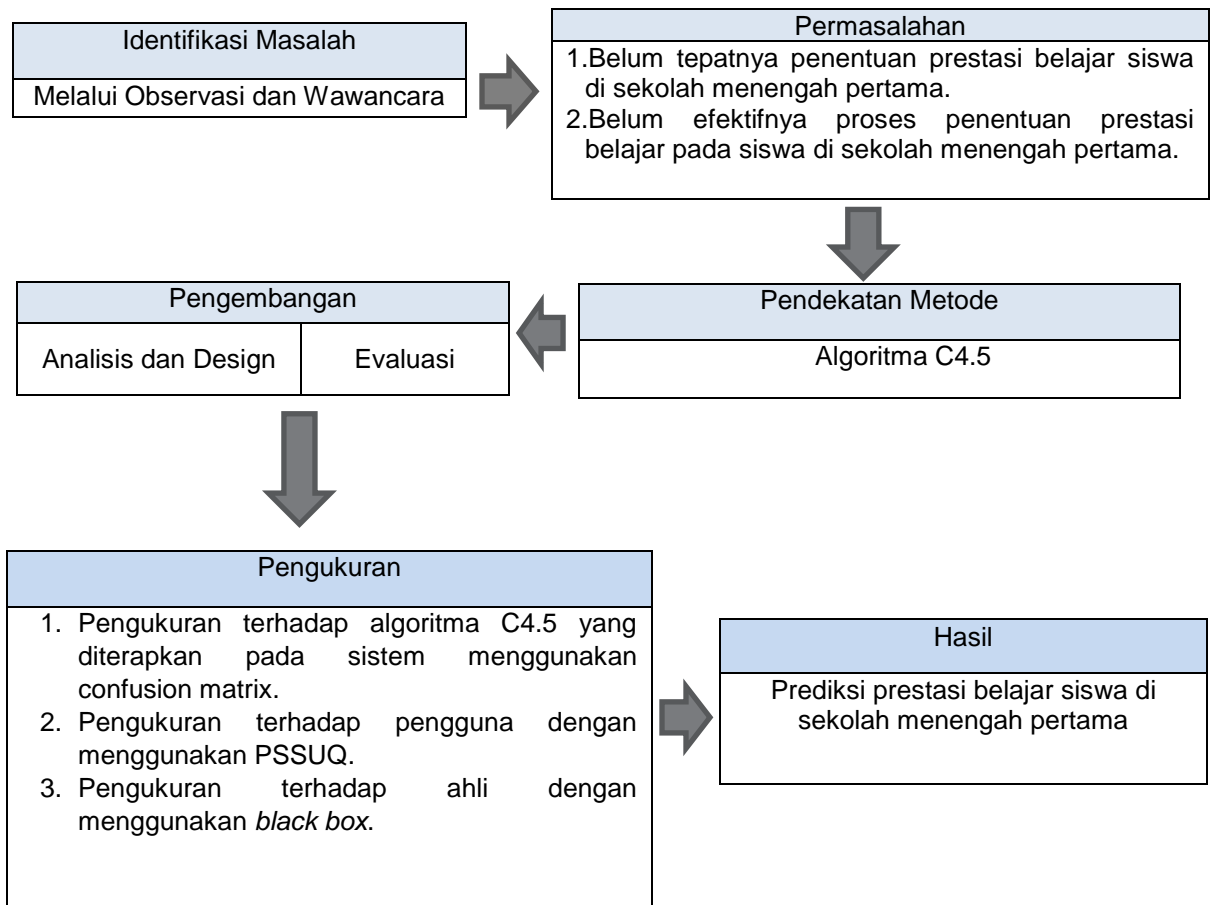
10.	(Astrid Noviriandini, Nurajijah, 2019)	Analisis Kinerja Algoritma C4.5 Dan Naïve Bayes Untuk Memprediksi Prestasi Siswa Sekolah Menengah Kejuruan	Noviriandini, A., & Nurajijah, N. (2019). ANALISIS KINERJA ALGORITMA C4. 5 DAN NAÏVE BAYES UNTUK MEMPREDIKSI PRESTASI SISWA SEKOLAH MENENGAH KEJURUAN. https://ejournal.nusamandiri.ac.id/index.php/jitk/article/view/607	Kontribusi dalam penelitian ini adalah pengambilan uji hasil yang menggunakan confusion matrix dan curva roc. Kelemahan dalam penelitian ini adalah belum adanya perhitungan nilai akurasi.
-----	--	--	--	---

Berdasarkan uraian rujukan penelitian yang sudah dilakukan sebelumnya dapat disimpulkan terdapat kesamaan metode yang digunakan yaitu metode algoritma C4.5. Sedangkan perbedaannya terletak pada penggunaan variabel, pada penelitian sebelumnya variabel yang digunakan meliputi cara belajar, lingkungan, ujian harian sedangkan pada penelitian yang akan dilakukan variabel yang digunakannya enam variabel yaitu Presensi, Pemahaman Materi, Pengerjaan Tugas, Hasil Ujian Akhir, Hasil Akhir. Selain itu juga pada penelitian sebelumnya belum diketahui tingkat akurasi dari metode C4.5 untuk prediksi prestasi belajar siswa, hal tersebut akan dilakukan pada penelitian yang akan dilakukan.

E. Kerangka Pemikiran

Identifikasi masalah di dapat dengan melalui wawancara, berdasarkan hasil wawancara komponen permasalahan mencakup fenomena yaitu belum tepat penentuan prestasi belajar siswa dan belum efektifnya proses penentuan prestasi belajar pada siswa. Dengan permasalahan yang ada pendekatan yang digunakan yaitu Algoritma C4.5 menjadi dasar hasil penelitian, dilanjutkan dengan pengembangan, pengembangan dibagi menjadi dua, yaitu tahap *analisis* dan *design*, dan tahap evaluasi. Tahap evaluasi tersebut sistem menguji ketepatan agar memperoleh hasil yang tepat, Hasil yang di dapatkan adalah sebuah prediksi prestasi belajar siswa di sekolah menengah pertama. Kemudian dilakukan uji hasil yaitu mengukur

ketepatan hasil perhitungan algoritma C4.5 Dengan menggunakan *confusion matrix*, pengukuran terhadap pengguna dengan menggunakan PSSUQ, dan pengukuran terhadap ahli dengan menggunakan *blackbox*. Berikut kerangka pemikiran pada gambar 2.7 untuk penelitian ini :



Gambar 2.7 Kerangka Pemikiran

F. Hipotesis Penelitian

Berdasarkan permasalahan yang dihadapi yaitu belum akurat dan belum efektif pada saat melakukan proses penentuan untuk prediksi hasil prestasi belajar siswa di sekolah menengah pertama, maka perlu adanya suatu cara untuk mengatasi permasalahan tersebut. Dalam teori prediksi ada beberapa metode yang dapat melakukan prediksi terhadap data yang diperoleh, salah satunya metode Algoritma C4.5. Metode Algoritma C4.5 akan menghasilkan hasil prediksi berdasarkan data yang telah ada. Serta pemberian hasil pohon keputusan yang telah dilakukan perhitungan terhadap data yang diperoleh. Penelitian yang dilakukan oleh Khairunnisa Fanny Irnanda, Dedy Hartama, Agus Perdana Windarto (2019) di dalam penelitiannya yang berjudul “ANALISA KLASIFIKASI C4.5 TERHADAP FAKTOR PENYEBAB MENURUNNYA PRESTASI BELAJAR MAHASISWA PADA MASA PANDEMI” menggunakan metode Algoritma C4.5. Maka dapat ditetapkan hipotesis pada penelitian ini penerapan Algoritma C4.5 diduga dapat memprediksi prestasi belajar siswa di sekolah menengah pertama dalam pembelajarannya