

BAB II KERANGKA TEORITIS

A. Landasan Teori

1. Data Mining

Menurut (Turban, Aronson and Liang, 2005, p. 263), Data mining adalah proses mengekstraksi dan mengidentifikasi informasi yang berguna dan pengetahuan yang terkait dari database besar menggunakan teknik statistik, matematika, kecerdasan buatan, dan pembelajaran mesin.

Istilah data mining dan *knowledge discovery in databases* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Padahal, konsep kedua istilah tersebut berbeda, namun saling berkaitan. Juga, salah satu tahapan dari seluruh proses KDD adalah data mining.

Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Han, Kamber and Pei, 2012, p. 6);

- a. pembersihan data, untuk menghilangkan noise dan data yang tidak konsisten.
- b. integrasi data, di mana beberapa sumber data dapat digabungkan.
- c. seleksi data, di mana data yang relevan dengan tugas analisis yang diambil dari database.
- d. data transformasi, dimana data diubah dan digabungkan ke dalam format yang cocok untuk penambangan dengan melakukan peringkasan atau agregasi operasi.
- e. data mining, merupakan proses penting untuk mengekstraksi pola data menggunakan metode cerdas.
- f. evaluasi Pola, Proses menafsirkan pola menjadi pengetahuan yang dapat digunakan untuk mendukung pengambilan keputusan.
- g. presentasi pengetahuan menggunakan teknik visualisasi dan representasi pengetahuan untuk menyajikan pengetahuan yang diekstraksi kepada pengguna.

Langkah 1 sampai 4 merupakan berbagai bentuk *pre-processing* data di mana data disiapkan untuk data mining. Hal ini menunjukkan bahwa data mining merupakan salah satu langkah dalam proses KDD karena dapat mengungkap pola tersembunyi yang digunakan untuk evaluasi..

2. Klasifikasi

Menurut (Prasetyo, 2013, p. 44), Klasifikasi dapat didefinisikan secara detail sebagai waktu pekerjaan yang melakukan pelatihan atau pembelajaran terhadap fungsi target yang memetakan setiap vektor (set fitur) x ke depan satu dari sejumlah table kelas y yang tersedia pekerjaan pelatihan tersebut akan menghasilkan suatu model yang kemudian disimpan sebagai memori.

Klasifikasi digunakan untuk pembuatan model yang dapat melakukan pemetaan dari setiap himpunan variabel kesetiap targetnya kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan variabel yang baru didapat. Dalam pembangunan model selama proses pelatihan tersebut diperlukan adanya suatu algoritma untuk membangunnya yang disebut dengan algoritma pelatihan.

3. Business Process Model and Notation (BPMN)

Menurut (Jeston and Nelis, 2006, p. 196), *Business Process Modeling Notation* (BPMN) adalah suatu notasi standar yang digunakan di dalam pemodelan proses bisnis yang berupa ikon atau gambar. Diusulkan oleh BPMI (*Business Process Management Initiative*) pada tahun 2004. BPMN dirancang bukan hanya mudah digunakan dan dipahami, tetapi juga memiliki kemampuan untuk memodelkan proses bisnis yang kompleks dan secara spesifik dirancang dengan mempertimbangkan *web services*. BPMN menyediakan notasi yang dapat dengan mudah dipahami oleh semua pengguna bisnis, mulai dari analis bisnis yang membuat draf awal proses hingga pengembang teknologi yang bertanggung jawab untuk menerapkan teknologi yang digunakan untuk menjalankan proses tersebut. Berikut adalah aturan dalam memodelkan suatu proses bisnis:

- a. Memodelkan kejadian-kejadian yang memulai proses, proses yang dilakukan dan hasil akhir dari aliran proses.
- b. Keputusan bisnis atau percabangan aliran dimodelkan dengan *gateways*. Sebuah *gateway* mirip dengan simbol keputusan dalam *flowchart*.
- c. Sebuah proses dalam aliran dapat mengandung *sub-process*, yang secara grafis dapat ditunjukkan dengan BPD (*Business Process Diagram*) lain yang tersambung melalui sebuah *hyperlink* ke simbol proses.
- d. Jika sebuah proses tidak didetilkan ke dalam sub proses, maka dianggap sebagai sebuah task – yaitu level proses paling rendah.
- e. Sebuah tanda '+' pada simbol proses menunjukkan bahwa proses ini didekomposisi, jika tidak ada tanda '+', maka proses ini disebut sebuah *task*.

BPD terdiri dari beberapa elemen grafis, dan kategori dasar elemen BPD adalah:

a. Flow Objects

Flow Object dibagi menjadi 3, yaitu *event*, *activity* dan *gateway*.

- 1) *Event* digambarkan dengan sebuah lingkaran dan merupakan sesuatu yang “terjadi” selama berlangsungnya proses bisnis. *Event-event* ini mempengaruhi aliran proses dan biasanya memiliki penyebab (*trigger*) atau hasil (*result*). *Event* adalah lingkaran dengan pusat terbuka untuk memungkinkan pembedaan *trigger* dan *result* yang berbeda. Terdapat 3 tipe event berdasarkan kapan mereka mempengaruhi aliran yaitu *Start*, *Intermediate*, dan *End*.



Gambar 2.1 Simbol *Event*

- 2) *Activity* ditunjukkan dengan persegi panjang dengan ujung-ujung bulat dan merupakan bentuk umum untuk pekerjaan yang dilakukan oleh perusahaan. Sebuah aktivitas dapat berdiri sendiri atau gabungan. Tipe dari aktivitas adalah *task* dan *sub-process* yang dibedakan dengan tanda + pada bagian tengah bawah dari bentuk tersebut.
- 3) *Gateway* digambarkan dengan bentuk seperti belah ketupat dan digunakan untuk mengontrol percabangan dan penggabungan *Sequence Flow*. Jadi, *gateway* menentukan keputusan tradisional, penggabungan, dan penggabungan aliran. *Internal Markers* akan menentukan perilaku dari kontrol.



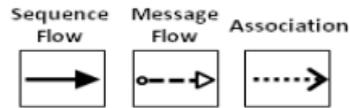
Gambar 2.2 Simbol *Activity* dan *Gateway*

b. Connecting Objects

Connecting Object adalah elemen yang menghubungkan *flow object*. *Connecting Object* juga memiliki 3 jenis elemen yaitu :

- 1) Alur Sequence (*Sequence Flow*) digunakan untuk menunjukkan urutan yang kegiatan akan yang dilakukan dalam sebuah proses.
- 2) Alur Pesan (*Messege Flow*) digunakan untuk menunjukkan aliran pesan antara dua entitas yang siap untuk mengirim dan menerima.

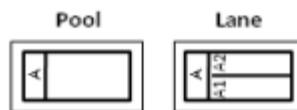
- 3) Asosiasi (*Association*) digunakan untuk asosiasi data, informasi dan artefak dengan aliran benda.



Gambar 2.3 Simbol *Sequence Flow*, *Message Flow* dan *Association*

c. Swimlanes

Swimlanes digambarkan dengan bentuk garis yang memisahkan dan mengelompokkan aktor (pelaku yang berinteraksi dengan sistem). Banyak metodologi pemodelan menggunakan konsep *swimlanes* sebagai mekanisme untuk membagi kategori visual yang menggambarkan kemampuan fungsional atau tanggung jawab yang berbeda. BPMN mendukung *swimlanes* dengan dua bentuk *swimlane objects* yaitu *pool* yang mewakili partisipasi dalam sebuah proses dan *lane* yaitu sub-bagian dalam sebuah *pool* dan akan menambah panjang dari *pool* baik vertikal ataupun horizontal. *Lanes* digunakan untuk mengatur dan mengkategorikan aktivitas.



Gambar 2.4 Simbol *Pool* dan *Lane* dalam *Swimlane*

d. Artifacts

Artifacts adalah elemen yang digunakan untuk memberikan informasi tambahan dari sebuah proses. BPMN dirancang untuk memungkinkan pemodelan dan alat pemodelan fleksibilitas untuk memperluas notasi dasar dan menyediakan kemampuan untuk konteks tambahan yang tepat untuk situasi pemodal tertentu. Berbagai *Artifacts* dapat ditambahkan ke dalam diagram sesuai dengan konteks dari proses bisnis yang dimodelkan. Versi BPMN saat ini memiliki 3 tipe *Artifacts*, yaitu

- 1) *Data Object*: mekanisme untuk menunjukkan bagaimana data dibutuhkan atau diproduksi oleh aktivitas. *Data object* dihubungkan dengan aktivitas melalui *Associations*.
- 2) *Group*: diwakili dengan persegi panjang dengan ujung bulat yang digambarkan dengan garis putus-putus. *Group* dapat digunakan untuk tujuan dokumentasi atau analisis, tetapi tidak mempengaruhi *Sequence Flow*.

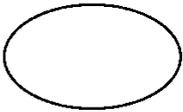
3) *Annotation*: mekanisme untuk pemodel memberikan informasi teks tambahan untuk pembaca dari diagram BPMN.

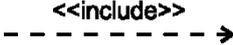
4. Unified Modeling Language (UML)

Menurut (Whitten and Bentley, 2007, p. 371), *Unified Modeling Language* (UML) adalah bahasa pemodelan yang digunakan untuk mendefinisikan atau menggambarkan sistem perangkat lunak berdasarkan objek dalam sistem. UML tidak menentukan metode yang akan digunakan saat mengembangkan sistem, hanya notasi standar yang biasa digunakan untuk pemodelan objek.

Menurut (Rosa and Shalahuddin, 2016, p. 155), *Use Case Diagram* adalah pemodelan untuk kelakuan (*behavior*) sistem informasi yang akan dibuat. *Use Case* mendeskripsikan sebuah interaksi antara satu atau lebih actor dengan sistem informasi yang akan dibuat. Berikut adalah simbol dari *use case diagram*.

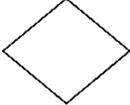
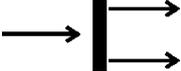
Tabel 2.1 Simbol *Use Case Diagram*

Simbol	Nama	Keterangan
	<i>Actor</i>	Orang Proses atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat diluar sistem informasi itu sendiri.
	<i>Usecase</i>	Fungsionalitas yang disediakan sistem sebagai unit-unit yang saling bertukar pesan antar unit atau <i>actor</i> .
	<i>Association</i>	Komunikasi antara <i>actor</i> dengan <i>usecase</i> yang berpartisipasi pada <i>usecase</i> atau <i>usecase</i> memiliki interaksi dengan <i>actor</i>
	<i>Extend</i>	Relasi <i>usecase</i> tambahan ke sebuah <i>usecase</i> dimana <i>usecase</i> yang ditambahkan dapat berdiri sendiri walau tanpa <i>usecase</i> tambahan memiliki nama depan yang sama dengan <i>usecase</i> yang ditambahkan

Simbol	Nama	Keterangan
	<i>Generalization</i>	Hubungan generalisasi dan spesialisasi (umum-khusus) antara dua buah <i>usecase</i> dimana fungsi yang satu adalah fungsi yang lebih umum dari lainnya
	<i>Include</i>	Relasi <i>usecase</i> tambahan ke sebuah <i>usecase</i> dimana <i>usecase</i> yang ditambahkan memerlukan <i>usecase</i> ini untuk menjalankan fungsional atau sebagai syarat dijalankan <i>usecase</i> ini.

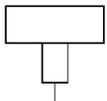
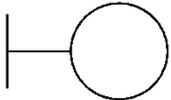
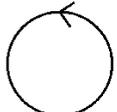
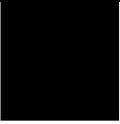
Menurut (Rosa and Shalahuddin, 2016, p. 161), *Activity Diagram* menggambarkan *workflow* (aliran kerja) atau aktifitas dari sebuah sistem atau proses bisnis atau menu yang ada pada perangkat lunak. Berikut adalah simbol *Activity Diagram*.

Tabel 2.2 Simbol *Activity Diagram*

Simbol	Nama	Keterangan
	Status Awal/Initial	Status awal aktivitas sistem
	<i>Activity</i>	Aktivitas yang dilakukan sistem, aktivitas biasanya diawali dengan kata kerja.
	<i>Decision</i>	Asosiasi percabangan dimana lebih dari satu aktivitas digabungkan menjadi satu.
	<i>Join</i>	Asosiasi penggabungan dimana lebih dari satu aktivitas lebih dari satu.
	Status Akhir/Final	Status akhir yang dilakukan sistem.
	<i>Swimlane</i>	Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi.

Menurut (Rosa and Shalahuddin, 2016, p. 165), *Sequence Diagram* menggambarkan kelakuan objek pada *usecase* dengan mendeskripsikan waktu hidup objek dan *message* yang dikirimkan dan diterima antar objek. Berikut adalah simbol dari *sequence diagram*.

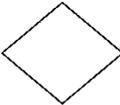
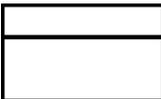
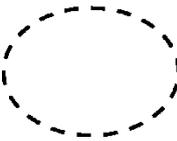
Tabel 2.3 Simbol *Sequence Diagram*

Simbol	Nama	Keterangan
	<i>Actor</i>	Mempresentasikan entitas yang berada diluar sistem dan berinteraksi diluar sistem.
	<i>Lifeline</i>	Menghubungkan objek selama <i>sequence</i> (<i>message</i> dikirim atau diterima).
	<i>General</i>	Merepresentasikan entitas tunggal dalam <i>sequence</i>
	<i>Boundary</i>	Berupa tepi dari sistem, seperti <i>user interface</i> dan alat yang berinteraksi dengan yang lain.
	<i>Control</i>	Elemen mengatur aliran dari informasi untuk sebuah skenario
	<i>Entity</i>	Elemen yang bertanggung jawab menyimpan informasi
	<i>Activation</i>	Suatu titik dimana sebuah objek mulai berpartisipasi dalam sebuah <i>sequence</i> yang menunjukkan sebuah objek atau mengirim atau menerima objek.
	<i>Message Entry</i>	Berfungsi untuk menggambarkan pesan/ hubungan antar objek yang menunjukkan urutan kejadian yang terjadi.

Simbol	Nama	Keterangan
	<i>Message to Self</i>	Simbol ini menggambarkan pesan/hubungan objek itu sendiri, yang menunjukkan urutan kejadian yang terjadi.
	<i>Message Return</i>	Menggambarkan hasil dari pengiriman <i>message</i> yang digambarkan dengan arah dari kanan ke kiri.

Menurut (Rosa and Shalahuddin, 2016, p. 141), *Class Diagram* menggambarkan struktur sistem dari segi pendefinisian kelas – kelas yang akan dibuat untuk membangun sistem. Diagram kelas dibuat agar pembuat program atau *programmer* membuat kelas – kelas sesuai rancangan di dalam diagram kelas agar antara dokumentasi perancangan dan perangkat lunak sinkron. Berikut adalah simbol dari *Class Diagram*.

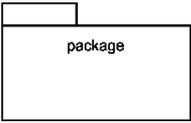
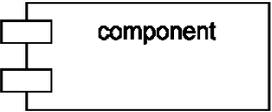
Tabel 2.4 Simbol *Class Diagram*

Simbol	Nama	Keterangan
	<i>Generalization</i>	Hubungan dimana objek anak (<i>descendent</i>) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk (<i>ancestor</i>)
	<i>Navy 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 <i>actor</i> .
	<i>Realization</i>	Operasi yang benar-benar dilakukan oleh suatu objek

Simbol	Nama	Keterangan
	<i>Dependency</i>	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (<i>independent</i>) akan mempengaruhi elemen yang bergantung pada elemen yang tidak mandiri.
	<i>Association</i>	Yang menghubungkan antara objek satu dengan objek lainnya.

Menurut (Rosa and Shalahuddin, 2016, p. 148), *Component Diagram* dibuat untuk menunjukkan organisasi dan ketergantungan diantara kumpulan komponen dalam sebuah sistem. *Component Diagram* fokus pada komponen sistem yang dibutuhkan dan ada di dalam sistem. Berikut adalah simbol dari *Component Diagram*.

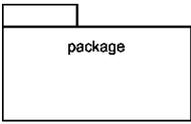
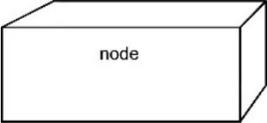
Tabel 2.5 Simbol Component Diagram

Simbol	Nama	Keterangan
	<i>Package</i>	<i>Package</i> merupakan sebuah bungkusan dari satu atau lebih komponen
	<i>Component</i>	Komponen sistem
	<i>Dependency</i>	Kebergantungan antar komponen, arah panah mengarah pada komponen yang dipakai.
	<i>Interface</i>	Sebagai antarmuka komponen agar tidak mengakses langsung komponen.
	<i>Link</i>	Relasi antar komponen

Menurut (Rosa and Shalahuddin, 2016, p. 154), *Deployment Diagram* menunjukkan konfigurasi komponen dalam proses eksekusi aplikasi. *Deployment Diagram* juga dapat digunakan untuk memodelkan sistem tambahan (*embedded*

system), sistem *client/server*, sistem terdistribusi murni dan rekayasa ulang aplikasi. Berikut adalah simbol dari *Deployment Diagram*.

Tabel 2.6 Simbol Deployment Diagram

Simbol	Nama	Keterangan
	Package	Package merupakan sebuah bungkus dari satu atau lebih node
	Node	Node biasanya mengacu pada perangkat keras (<i>hardware</i>), perangkat lunak yang tidak dibuat sendiri (<i>software</i>), jika di dalam <i>node</i> disertakan komponen untuk mengkonsistenkan rancangan maka komponen yang diikutsertakan harus sesuai dengan komponen yang telah didefinisikan sebelumnya pada diagram komponen.
	Dependency	Kebergantungan antar <i>node</i> , arah panah mengarah pada <i>node</i> yang dipakai.
	Link	Relasi antar <i>node</i> .

5. Database

a. Pengertian Database

Menurut (Sutabri, 2012, p. 47) menjelaskan bahwa Database merupakan kumpulan data yang saling berkaitan dan berhubungan satu sama lain yang disimpan pada perangkat keras komputer dan dikelola menggunakan perangkat lunak. Data perlu disimpan dalam basis data untuk keperluan penyediaan informasi lebih lanjut. Data yang ada di dalam database perlu diorganisasikan sedemikian rupa agar informasi yang dihasilkan berkualitas.

b. Pengertian MySQL

Menurut (Solichin, 2016, p. 85) menjelaskan bahwa MySQL adalah sebuah perangkat lunak DBMS yang multithread, multi-user, dengan sekitar 6 juta instalasi di seluruh dunia. Berikut adalah beberapa kelebihan MySQL:

- 1) Free (bebas didownload);
- 2) Stabil dan Tangguh;
- 3) Fleksibel dengan berbagai pemrograman;
- 4) Security yang baik;
- 5) Dukungan dari banyak komunitas;
- 6) Kemudahan management database;
- 7) Mendukung transaksi;
- 8) Perkembangan software yang cukup cepat.

6. Web Server

a. Web Server

Menurut (Sibero, 2013, p. 11), Web Server adalah sebuah komputer yang terdiri dari perangkat keras dan perangkat lunak. Secara bentuk fisik dan cara kerjanya, perangkat keras Web Server tidak berbeda dengan komputer rumah atau PC, yang membedakan adalah fasilitas dan kapabilitasnya.

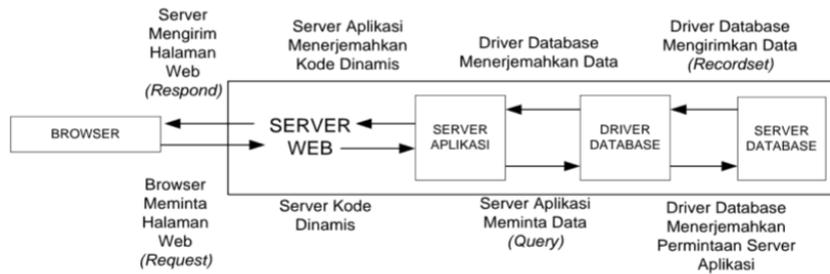
b. Apache

Menurut (Sadeli, 2013, p. 2), Apache (Server HTTP Apache atau ServerWeb/WWW Apache) adalah webserver yang dapat dijalankan di banyak sistem operasi (Unix, BSD, Linux, Microsoft Windows dan Novell Netware serta platform lainnya) yang berguna untuk melayani dan memfungsikan situs web.

7. Bahasa Pemrograman

a. PHP

Menurut (Supono and Putratama, 2018, p. 3), PHP (PHP: *Hypertext Preprocessor*) adalah bahasa pemrograman yang digunakan untuk menerjemahkan baris kode program menjadi kode mesin yang dapat dipahami oleh komputer server-side yang dapat ditambahkan ke dalam HTML. Dibawah ini adalah gambaran tentang cara kerja PHP.



Gambar 2.5 Prinsip Kerja PHP

Penjelasan Gambar:

- 1) Client/user mengirimkan file PHP (menggunakan browser) melalui Web Server (Seperti Internet Explorer, Mozilla Firefox, Google Chrome, dll);
- 2) Web server mendapatkan request atau permintaan dari user lalu meneruskan ke Server melalui jaringan internet;
- 3) Web Server lalu meneruskan permintaan file PHP tersebut ke PHP processor. PHP processor dapat berupa modul (bagian dari web-server) atau terpisah (sebagai CGI/ Fast-CGI);
- 4) Permintaan diproses oleh PHP dan diteruskan ke database (jika terdapat permintaan ke database), kemudian hasilnya. dikirim kembali ke web-server;
- 5) Web Server memaket kembali hasil tersebut dengan menambahkan HTTP header dan dikirim kembali ke browser melalui jaringan Internet;
- 6) Browser memproses HTTP paket dan menampilkannya kembali kepada user sebagai file HTML.

b. Hypertext Markup Language (HTML)

Menurut (Devi, 2020, pp. 1–2), *Hypertext Markup Language* (HTML) adalah sebuah bahasa markah yang digunakan untuk membuat sebuah halaman web, menampilkan berbagai informasi. dalam sebuah penjelajah web internet dan memformat hiperteks sederhana yang ditulis dalam berkas format ASCII agar dapat menghasilkan tampilan wujud yang terintegrasi. Dengan kata lain, berkas yang dibuat dalam perangkat lunak pengolah kata dan disimpan dalam format ASCII normal sehingga menjadi halaman web dengan perintah-perintah HTML.

B. Algoritma Naïve Bayes

Menurut (Nofriansyah, Erwansyah and Ramadhan, 2016, p. 87), Naïve Bayes merupakan pengklasifian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris bernama Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Klasifikasi Naïve Bayes diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya. Sehingga Naïve Bayesian Classifier dapat didefinisikan juga sebagai metode klasifikasi yang berdasarkan teori probabilitas dan teorema Bayes dengan asumsi bahwa setiap variabel atau parameter penentu keputusan bersifat bebas (independence) sehingga keberadaan setiap variabel tidak ada kaitannya dengan keberadaan atribut yang lain.

Menurut (Saleh, 2015, p. 3) prediksi Bayes didasarkan pada teorema Bayes dengan formula umum sebagai berikut :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{p(X)}$$

Dimana :

- X : Data dengan class yang belum diketahui
- H : Hipotesis data merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)
- P(H) : Probabilitas hipotesis H (prior probabilitas)
- P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) : Probabilitas X

Penentuan *class* dilakukan dengan cara membandingkan nilai probabilitas suatu sampel berada di *class* yang lain. Untuk menentukan *class* yang cocok dari suatu sampel dilakukan dengan cara membandingkan nilai posterior untuk masing-masing *class*, dan mengambil *class* dengan nilai posterior yang tertinggi.

Adapun algoritma penyelesaian dari Metode Naïve Bayes dapat dilihat pada gambar berikut :



Gambar 2.6 Algoritma Naive Bayesian Classifier

(Sumber : (Nofriansyah, 2015, p. 42))

Menurut (Suyanto, 2017) contoh kasus ini yaitu keputusan Ya dan Tidak dalam penerimaan bonus, data training yang digunakan yaitu 11 data dan data yang telah di klasifikasikan berdasarkan kartu, panggilan dan blok, adapun data training dibawah ini yaitu :

No.	Pelanggan	Kartu	Panggilan	Blok	Bonus
1.	Andi	Prabayar	Sedikit	Sedang	Tidak
2.	Budi	Pascabayar	Banyak	Sedang	Ya
3.	Citra	Prabayar	Banyak	Sedang	Ya
4.	Dedi	Prabayar	Banyak	Rendah	Tidak
5.	Evan	Pascabayar	Cukup	Tinggi	Ya
6.	Feni	Pascabayar	Cukup	Sedang	Ya
7.	Gito	Prabayar	Cukup	Sedang	Ya
8.	Hani	Prabayar	Cukup	Rendah	Tidak
9.	Jodi	Pascabayar	Sedikit	Tinggi	Ya
10.	Kafi	Pascabayar	Banyak	Tinggi	Ya
11.	Linda	Pascabayar	Sedikit	Rendah	Ya

Pelanggan	Kartu	Panggilan	Blok	Bonus
Mira	Pascabayar	Cukup	Rendah	?

Penyelesaian :

- 1) menghitung probabilitas prior untuk kelas pertama;

$$P(\text{Bonus} = \text{Ya}) = 8/11 = 0,7273$$

$$P(\text{Bonus} = \text{Tidak}) = 3/11 = 0,2727$$

- 2) menghitung probabilitas bersyarat untuk setiap kelas;

$$P(\text{Kartu} = \text{Pascabayar} | \text{Bonus} = \text{Ya}) = 6/8 = 0,75$$

$$P(\text{Kartu} = \text{Pascabayar} | \text{Bonus} = \text{Tidak}) = 0/3 = 0$$

$$P(\text{Panggilan} = \text{Cukup} | \text{Bonus} = \text{Ya}) = 3/8 = 0,375$$

$$P(\text{Panggilan} = \text{Cukup} | \text{Bonus} = \text{Tidak}) = 1/3 = 0,3333$$

$$P(\text{Blok} = \text{Rendah} | \text{Bonus} = \text{Ya}) = 1/8 = 0,125$$

$$P(\text{Blok} = \text{Rendah} | \text{Bonus} = \text{Tidak}) = 2/3 = 0,6667$$

- 3) menghitung probabilitas untuk setiap kelas;

$$P(X | \text{Bonus} = \text{Ya}) = 0,75 \times 0,375 \times 0,125 = 0,0352$$

$$P(X | \text{Bonus} = \text{Tidak}) = 0 \times 0,3333 \times 0,6667 = 0$$

- 4) menghitung probabilitas posterior;

$$P(X | \text{Bonus} = \text{Ya}) \times (P | \text{Bonus} = \text{Ya}) = 0,0352 \times 0,7273 = 0,256$$

$$P(X | \text{Bonus} = \text{Tidak}) \times (P | \text{Bonus} = \text{Tidak}) = 0 \times 0,2727 = 0$$

Pelanggan	Kartu	Panggilan	Blok	Bonus
Mira	Pascabayar	Cukup	Rendah	Ya

Mengingat $P(X | \text{Bonus} = \text{Ya}) \times (P | \text{Bonus} = \text{Ya}) = 0,256$ lebih besar dari $P(X | \text{Bonus} = \text{Tidak}) \times (P | \text{Bonus} = \text{Tidak}) = 0$, maka Naïve Bayes Classifier mengklasifikasikan pelanggan bernama Mira ke dalam kelas bonus = "Ya".

C. Prediksi Ketercapaian Nilai Kriteria Ketuntasan Minimal

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi (Herdianto, 2013, p. 8).

Kriteria Ketuntasan Minimal (KKM) merupakan kriteria paling rendah untuk menyatakan bahwa peserta didik telah mencapai ketuntasan belajar. Penetapan KKM dilakukan pada awal tahun pembelajaran berdasarkan hasil musyawarah guru mata pelajaran pada satuan pendidikannya atau beberapa satuan pendidikan yang memiliki karakteristik hampir sama.

Jadi, prediksi ketercapaian nilai kriteria ketuntasan minimal yaitu suatu proses dalam memperkirakan nilai mata pelajaran siswa yang tidak mencapai nilai kriteria paling rendah untuk membantu guru dalam mengetahui siapa saja siswa yang tidak tuntas nilai kriteria minimal di satu mata pelajaran.

D. Tinjauan Pustaka

Penelitian rujukan merupakan acuan yang dibutuhkan peneliti dalam melakukan sebuah penelitian. Pada penelitian ini, penelitian rujukan yang diambil berdasarkan persamaan metode yang digunakan yaitu Naïve Bayes. Banyak penelitian lain yang menggunakan metode ini dengan bermacam-macam kasus, antara lain:

1. **Penerapan Metode Naïve Bayes Pada Aplikasi Sistem Pakar Untuk Diagnosa Penyakit Kulit Pada Kucing** (Fadhilah, Andryana and Gunaryati, 2020)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat disimpulkan bahwa, sistem dapat memberikan informasi berupa diagnosa penyakit kulit kucing berdasarkan penerapan metode naïve bayes disertai penyebab dan solusi yang akurat. Dengan adanya sistem pakar yang telah dibuat maka dapat mempermudah pemilik atau pecinta kucing dalam mendiagnosa awal penyakit kulit pada kucing sehingga dapat dilakukan pengobatan awal sebelum dibawa ke dokter hewan atau rumah sakit hewan. Pengujian akurasi sistem yang dilakukan terhadap 50 data uji menggunakan kuisioner yang diberikan kepada dokter hewan MADE Pets Care Cinere menghasilkan tingkat akurasi sebesar 90%

2. **Analisis Penentuan Standar Nilai Kriteria Ketuntasan Minimal (KKM) Mata Pelajaran Fisika Kelas XI SMAN 17 Makassar** (Hidayat, et al., 2020)

Berdasarkan penelitian tentang klasifikasi pelanggan menggunakan metode naïve bayes dapat ditarik kesimpulan yaitu kriteria kompleksitas KKM mata pelajaran fisika kelas XI SMAN 17 Makassar dikategorikan sedang dengan nilai rata-rata 70, kriteria daya dukungnya dikategorikan tinggi dengan nilai 83 dan kriteria intakenya dikategorikan sedang dengan nilai 75. Selanjutnya, berdasarkan hasil analisis kuantitatif untuk penentuan standar nilai kriteria ketuntasan minimal menunjukkan adanya kesesuaian antara KKM yang ditetapkan sekolah dan KKM yang diukur oleh peneliti dengan persen beda sebesar 5%.

3. **Meningkatkan Ketuntasan Siswa Melalui Model Pembelajaran Inkuiri** (Masiah, 2020)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat disimpulkan bahwa Peningkatan ketuntasan pada siklus I belum mencapai ketuntasan klasikal yaitu $\geq 85\%$, sehingga perlu dilanjutkan ke siklus II. Tindakan pada siklus II dilakukan berdasarkan hasil refleksi pada siklus sebelumnya. Setelah adanya

perbaikan pada siklus II, terjadi peningkatan ketuntasan siswa yaitu sebanyak 22 orang siswa yang masuk kategori tuntas, atau sekitar 88% dan sisanya 12% belum tuntas. Pencapaian ketuntasan pada siklus I sebesar 72%, dan pada siklus II 88%, sehingga terjadi peningkatan sebesar 12%. Berdasarkan pencapaian pada siklus II, dapat disimpulkan bahwa pembelajaran inkuiri dapat meningkatkan ketuntasan siswa.

4. **Penerapan Algoritma Naive Bayes pada Penentuan Kelayakan Calon Tenaga Kerja Indonesia** (Purba *et al.*, 2019)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat disimpulkan bahwa, Penerapan Data Mining menggunakan algoritma naïve bayes untuk penentuan kelayakan tenaga kerja Indonesia dengan data uji sebanyak 17 orang dengan menggunakan dua kelas. klasifikasi dengan kelas Layak sebanyak 10 orang dan kelas Tidak Layak sebanyak 7orang. Pengujian data pada Rapiminer 5.3 dengan menggunakan metode Naive Bayes berhasil menampilkan dua kelas dari hasil klasifikasi dengan persentase keakuratan sebesar 90,00%. Dengan class precision pada prediksi Layak memiliki nilai 90,91%, sedangkan pada prediksi Tidak Layak memiliki nilai 100 %. Class recall pada true Layak memiliki nilai 100%, sedangkan pada true Tidak Layak memiliki nilai 85,71%.

5. **Penerapan Metode Naïve Bayes untuk Prediksi Minat Baca Berdasarkan Usia** (Utari and Ulfah, 2021)

Berdasarkan hasil penelitian yang dilakukan, kesimpulan yang bisa diuraikan bahwa dengan menerapkan metode Naive Bayes dapat dijadikan sebagai sistem pendukung keputusan untuk melakukan prediksi minat baca berdasarkan usia. Dilakukan pengukuran nilai akurasi dalam penerapan metode Naive Bayes untuk uji hasil dengan hasil 80% yang disimpulkan layak dan bagus untuk digunakan.

6. **Penerapan Metode Naïve Bayes Untuk Rekomendasi Topik Tugas Akhir Berdasarkan Daftar Hasil Studi Mahasiswa di Perguruan Tinggi** (Triawan and Lintang Melinda, 2020)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat disimpulkan bahwa, Proses penentuan topik menjadi lebih efektif karena adanya perangkat aplikasi rekomendasi topik tugas akhir, sehingga mahasiswa mendapatkan rekomendasi topik tugas akhir dengan acuan yang tepat tanpa harus mengikuti pilihan teman. Dengan menerapkan metode Naive Bayes dibuat sebuah aplikasi rekomendasi topik tugas akhir berdasarkan daftar hasil studi. Aplikasi membantu memberi rekomendasi kepada mahasiswa dalam menentukan topik untuk tugas akhir. Telah dilakukan uji kelayakan terhadap aplikasi yang dibuat dengan nilai persentase kelayakan secara keseluruhan dari tiga indikator yaitu Efficiency,

Effectiveness Satisfaction adalah sebesar 79% yang bermakna aplikasi masuk dalam kategori layak digunakan, hal tersebut didasarkan pada tabel kategori kelayakan skala likert.

7. **Penerapan Naive Bayes dalam Prediksi Ketercapaian Nilai Ketuntasan Siswa** (Hartani, et al., 2018)

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa dengan menggunakan naïve bayes dapat melakukan prediksi terhadap ketercapaian nilai KKM siswa. Dengan menggunakan naïve bayes dengan data sebanyak 473 data record untuk data latih dan 157record untuk data uji. Dari pengujian didapatkan nilai accuracy sebesar 78%, precion sebesar 92% dan recall sebesar 48%.

8. **Perbandingan Tingkat Akurasi Klasifikasi Penerimaan Dosen Tetap Menggunakan Metode Naive Bayes Classifier dan C4.5** (Sadikin *et al.*, 2020)

Berdasarkan penelitian yang telah dilakukan tentang perbandingan tingkat akurasi dalam Penerimaan Dosen Tetap pada Universitas Potensi Utama dengan menerapkan metode Naïve Bayes Classification dan C4.5 dapat ditarik kesimpulan bahwa berdasarkan 9 kriteria yang diuji dalam penerimaan dosen tetap yaitu nilai IPK, usia, jenis kelamin, kesehatan, pengalaman mengajar, nilai tes tertulis, wawancara, dan homebase, kriteria homebase merupakan kriteria yang sangat dominan dalam menentukan apakah dosen tersebut dapat diterima atau tidak. Hasil percobaan yang telah dilakukan terhadap data penerimaan dosen tetap yang digunakan sebagai data latih, metode naïve bayes memiliki tingkat akurasi sebesar 83,7838 % sedangkan C4.5 sebesar 91,891 %. Sehingga dapat diperoleh kesimpulan dalam penelitian ini bahwa metode C4.5 merupakan metode yang memiliki tingkat akurasi yang baik dan efisien dari pada naïve bayes.

9. **Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga** (Saleh, 2015)

Berdasarkan hasil dari penelitian yang telah dilakukan maka dapat disimpulkan bahwa, Metode Naive Bayes memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk class yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi penggunaan listrik berdasarkan proses klasifikasi yang dilakukan oleh metode Naive Bayes itu sendiri. Berdasarkan data rumah tangga yang dijadikan data training, metode Naive Bayes berhasil mengklasifikasikan 47 data dari 60 data yang diuji. Sehingga metode Naive Bayes berhasil memprediksi besarnya penggunaan listrik rumah tangga dengan persentase keakuratan sebesar 78,3333%.

10. Penerapan Metode Naïve Bayes Classifier Untuk Penentuan Topik Tugas Akhir Pada Website Perpustakaan STIKOM Binaniaga (Ghaniy and Sihotang, 2019)

Berdasarkan hasil penelitian yang telah diuraikan, maka dapat ditarik kesimpulan bahwa, Metode Naive Classifier dapat digunakan untuk penentuan dan menampilkan topik jurusan TI pada judul yang di ajukan. Ketepatan dan menemukan dalam menentukan topik pada data judul skripsi jurusan TI baru dipengaruhi oleh data pembelajaran atau data latih pada setiap kategori. Data latih ini berisi kata – kata yang sering muncul pada masing – masing kategori atau kata – kata yang dapat mewakili kategori tertentu.

Tabel 2.7 Tinjauan Pustaka

No.	Nama Peneliti	Judul Penelitian	Jurnal	Kontribusi
1.	Furqon Fadhilah, Septi Andryana, Aris Gunaryati	Penerapan Metode Naïve Bayes Pada Aplikasi Sistem Pakar Untuk Diagnosa Penyakit Kulit Pada Kucing	Jurnal Infomedia Vol. 5 No. 1 Juni 2020 http://e-jurnal.pnl.ac.id/index.php/infomedia/article/view/1602	Kontribusi dalam penelitian ini adalah menggunakan metode Naïve Bayes pada aplikasi sistem pakar untuk diagnose penyakit kulit pada kucing.
2.	Muhammad Yusuf Hidayat, A. Nenyhindrawaty, Fitriani Nur	Analisis Penentuan Standar Nilai Kriteria Ketuntasan Minimal (KKM) Mata Pelajaran Fisika Kelas XI SMAN 17 Makassar	Jurnal Pendidikan Fisika Vol. 8 No. 1 Maret 2020 http://journal.uin-alauddin.ac.id/index.php/PendidikanFisika/article/view/7837	Kontribusi dalam penelitian ini adalah menganalisis data dengan analisis kualitatif untuk penentuan standar nilai KKM pada mata pelajaran Fisika.
3.	Masih	Meningkatkan Ketuntasan Siswa Melalui Model Pembelajaran Inkuiri	Bioscientist : Jurnal Ilmiah Biologi Vol. 8, No. 1; 2020 https://e-journal.undikma.ac.id/index.php/bioscientist/article/view/2657	Kontribusi dalam penelitian ini adalah memakai model pembelajaran inkuiri untuk meningkatkan ketuntasan belajar siswa pada bidang studi IPA
4.	Winda Hana Purba,	Penerapan Algoritma Naive	Prosiding Seminar Nasional Riset	Kontribusi dalam penelitian ini adalah menggunakan

No.	Nama Peneliti	Judul Penelitian	Jurnal	Kontribusi
	Poningsih, Dedi Suhendro, Irfan Sudahri Damanik, Ilham Syahputra Saragih	Bayes pada Penentuan Kelayakan Calon Tenaga Kerja Indonesia	Information Science (SENARIS) https://tunasbangsa.ac.id/seminar/index.php/senaris/article/view/83	Metode Naïve Bayes untuk menentukan kelayakan calon tenaga kerja Indonesia berdasarkan kriteria-kriteria seperti pendidikan, usia, pengalaman kerja, test kesehatan dan tinggi badan. sehingga bisa menjadi bahan pertimbangan untuk penerimaan calon tenaga kerja yang melakukan pendaftaran tenaga kerja.
5.	Lis Utari , Yawmil Ulfah	Penerapan Metode Naïve Bayes untuk Prediksi Minat Baca Berdasarkan Usia	Volume 11 Number 1 May 2021 Page. 67- 74 https://teknois.stikombinaniaga.ac.id/index.php/JBS/article/view/104	Kontribusi dalam penelitian ini adalah menggunakan Metode Naïve Bayes untuk memprediksi minat baca berdasarkan usia di perpustakaan GOR Pajajaran.
6.	Anggra Triawan, Della Lintang Melinda	Penerapan Metode Naïve Bayes Untuk Rekomendasi Topik Tugas Akhir Berdasarkan Daftar Hasil Studi Mahasiswa di Perguruan Tinggi	Volume 10 Number 2 November 2020 Page. 58-70 https://teknois.stikombinaniaga.ac.id/index.php/JBS/article/view/91	Kontribusi dalam penelitian ini adalah menggunakan Metode Naïve Bayes untuk merekomendasikan topik untuk tugas akhir mahasiswa dengan menggunakan daftar hasil studi mahasiswa sebagai dasar atau acuan yang jelas.
7.	Dwi Hartanti, Kusrini, Emha Luthfi Taufiq	Penerapan Naïve Bayes dalam Prediksi Ketercapaian Nilai Ketuntasan Minimal Siswa	JUSIKOM PRIMA (Jurnal Sistem Informasi Ilmu Komputer Prima) Vol. 2 No. 1, Juli 2018	Kontribusi dalam penelitian ini adalah menggunakan Metode Naïve Bayes untuk memprediksi Ketercapaian Nilai Ketuntasan Minimal Siswa

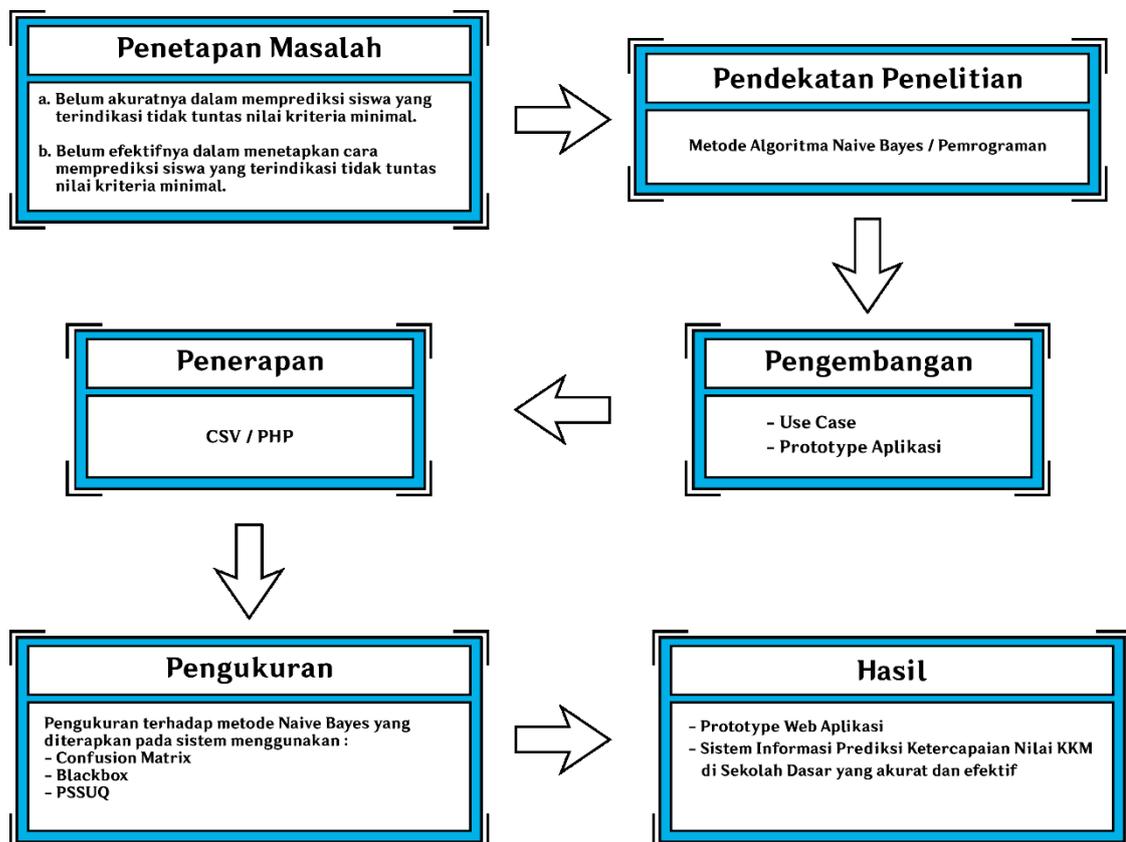
No.	Nama Peneliti	Judul Penelitian	Jurnal	Kontribusi
			http://jurnal.unprimdn.ac.id/index.php/JUSIKOM/article/view/147	
8.	Muhammad Sadikin, Rika Rosnelly, Roslina, Teddy Surya Gunawan, Wanayumini	Perbandingan Tingkat Akurasi Klasifikasi Penerimaan Dosen Tetap Menggunakan Metode Naive Bayes Classifier dan C4.5	JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 4, Nomor 4, Oktober 2020, Page 1100-1109 https://ejurnal.stmik-budidarma.ac.id/index.php/mib/article/download/2434/1787	Kontribusi dalam penelitian ini adalah membandingkan metode naive bayes dan C4.5 untuk dapat menganalisa dan mengklasifikasikan penerimaan dosen tetap dengan cara mengambil sample data pada penerimaan dosen
9.	Alfa Saleh	Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga	Vol. 2, No. 3, Mei 2015 – Juli 2015 https://citec.amikom.ac.id/main/index.php/citec/article/view/49	Kontribusi dalam penelitian ini adalah menggunakan metode Naive Bayes untuk mengetahui besaran penggunaan listrik. Sebagai upaya memahami masyarakat dalam menggunakan listrik dengan bijak.
10.	Kepin Sihotang, Rajib Ghaniy	Penerapan Metode Naive Bayes Classifier Untuk Penentuan Topik Tugas Akhir Pada Website Perpustakaan STIKOM Binaniaga	Jurnal Ilmiah Teknologi - Informasi dan Sains (TeknoIS) Volume 9 Nomor 1 Bulan Mei 2019 Hal. 63-72 https://teknois.stikombinaniaga.ac.id/index.php/JBS/article/view/7	Kontribusi dalam penelitian ini adalah menggunakan metode Naive Bayes untuk menentukan topik tugas akhir pada website perpustakaan STIKOM Binaniaga.

Dari penjelasan yang sesuai dengan tabel 2.9 diatas, terdapat beberapa kekurangan yang menjadi salah satu dasar penelitian ini yang dilakukan yaitu

kebanyakan jurnal hanya membahas tentang metode yang sama dan beberapa jurnal membahas tentang kasus yang sama. Selain itu, variabel-variabel yang digunakan juga berbeda dengan penelitian sebelumnya dimana pada penelitian ini menggunakan 6 variabel yaitu nama, kelas, nilai tugas, nilai UTS, absensi dan keterangan. Output yang dihasilkan dari penelitian sebelumnya dalam mengukur evaluasi uji hasil juga berbeda-beda, sedangkan pada penelitian ini pengukuran evaluasi uji hasil menggunakan confusion matrix.

E. Kerangka Pemikiran

Berikut adalah kerangka pemikiran yang dipakai untuk memecahkan masalah penelitian ini :



Gambar 2.7 Kerangka Pemikiran

Penelitian diawali dengan munculnya permasalahan terkait belum adanya metode untuk memprediksi nilai ketercapaian ketuntasan minimum siswa agar mengetahui apakah nilai siswa mencapai minimum ketuntasan atau tidak.

Kerangka pemikiran pada gambar diatas dapat dijelaskan sebagai berikut :

1. Penetapan masalah terkait yaitu belum akurat dalam memprediksi siswa yang terindikasi tidak lulus nilai kriteria minimum dan belum efektif dalam menetapkan cara memprediksi siswa yang terindikasi tidak tuntas nilai kriteria minimum;
2. Pendekatan penelitian menggunakan metode Naïve Bayes dan pemrograman;
3. Pengembangan yaitu tahap melakukan analisi dan perancangan gambar untuk use case diagram kemudian menerapkan metode algoritma naïve bayes ke dalam prototype aplikasi yang ingin dibuat;
4. Penerapan sebagai dasar tahap evaluasi dan validasi system prototype yang akan dibuat menggunakan PHP / CSV;
5. Pengukuran yaitu menguji ketepatan hasil perhitungan algoritma Naïve Bayes dengan menggunakan confusion matrix, pengujian sistem kepada ahli sistem informasi dan ahli materi menggunakan blackbox, dan uji kebergunaan kepada pengguna menggunakan PSSUQ;
6. Hasil yaitu prototype aplikasi dan sistem informasi yang menampilkan hasil dari prediksi nilai ketercapaian kriteria minimum siswa.

F. Hipotesis Penelitian

Hipotesis adalah jawaban sementara terhadap masalah yang masih bersifat praduga karena masih harus dibuktikan kebenarannya. Teori yang digunakan dalam penelitian ini ialah Algoritma Naïve Bayes dan diduga dapat Memprediksi Ketercapaian Nilai Kriteria Ketuntasan Minimal (KKM) Siswa di Sekolah Dasar secara tepat dan efektif. Algoritma Naïve Bayes, yaitu algoritma pengklasifikasian dengan menggunakan metode probabilitas dan statistik. Penelitian yang dilakukan sebelumnya oleh **(Dwi Hartanti, Kusri, Emha Luthfi Taufiq) Penerapan Naïve Bayes dalam Prediksi Ketercapaian Nilai Ketuntasan Minimal Siswa**, dapat ditetapkan hipotesis pada penelitian ini yaitu metode algoritma Naïve Bayes diduga dapat mengklasifikasikan siswa yang terindikasi tidak mencapai nilai kkm di Sekolah Dasar.