

## **BAB II KERANGKA TEORITIS**

### **A. Tinjauan Objek Penelitian**

Institut Pertanian Bogor merupakan Perguruan Tinggi Negeri yang berlokasi di Bogor, Jawa Barat. IPB diresmikan pada tahun 1963, yang mana sebelum itu IPB merupakan sebuah fakultas pertanian di Universitas Indonesia. Ada banyak sekali program kreativitas mahasiswa(PKM) yang ada di Institut Pertanian Bogor dimana puncak kegiatan PKM Ini yaitu Pekan Ilmiah Mahasiswa Nasional(PIMNAS). Namun pengolahan Dan Penyeleksian data mahasiswa yang mengikuti program ini pihak Perguruan Tinggi masih kesulitan dalam menentukan pilihan dan kompleksitas penilaian yang tinggi, sehingga menimbulkan kecemburuan social. Ketidaktransparan dan ketidakjelasan metode yang digunakan yang menimbulkan ketidakadilan dalam melakukan penilaian dalam menentukan Mahasiswa yang layak menjadi peserta PIMNAS , sehingga terdapat masalah yang terjadi dalam sistem ini. Permasalahan yang sering muncul yaitu belum tepatnya Penyeleksian Mahasiswa yang mengikuti Program PIMNAS, misalnya Mahasiswa yang tidak layak Menjadi Peserta PIMNAS namun layak menjadi Peserta PIMNAS begitupun sebaliknya . Masalah tersebut muncul dikarenakan kurang ketelitian dari para penyeleksi dalam menentukan Mahasiswa Yang layak dan Tidak untuk Menjadi Peserta PIMNAS.

### **B. Landasan Teori**

Dalam rangka memperoleh acuan penelitian, maka perlu dikemukakan sebuah landasan teori yang bersifat ilmiah. Dalam landasan teori dikemukakan teori yang berhubungan dengan materi yang akan digunakan untuk memecahkan masalah pada penelitian ini.

#### **1. Data Mining**

Menurut (Suryanto, 2017, p. 15) Data mining dalam konteks algoritma Naive Bayes adalah tentang penggunaan teknik analisis data untuk mempersiapkan data dan mengidentifikasi pola-pola yang relevan dalam dataset. salah satu metode dalam data mining yang digunakan untuk klasifikasi dan prediksi berdasarkan probabilitas. Di bawah ini adalah langkah-langkah utama dalam proses data mining menggunakan algoritma Naive Bayes:

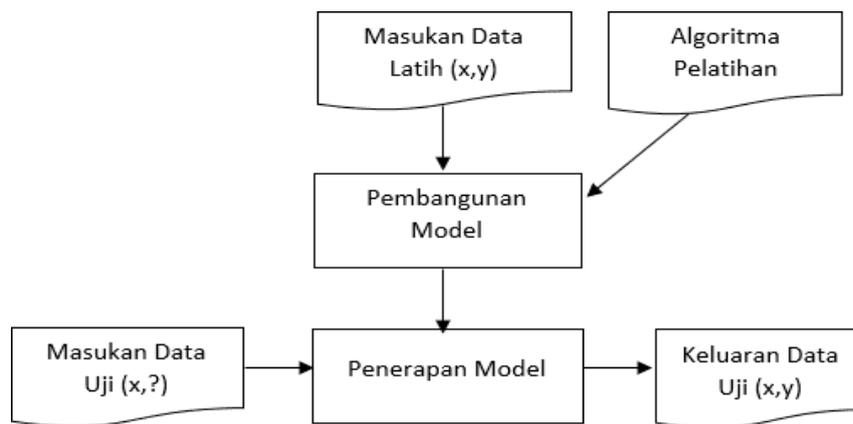
- (a) Persiapan data
- (b) Ekstraksi atribut pada dataset
- (c) Penerapan model
- (d) Pengujian model
- (e) Validasi model
- (f) Evaluasi dan penyesuaian

Data mining menggunakan algoritma Naive Bayes merupakan pendekatan yang kuat untuk klasifikasi dan prediksi berdasarkan probabilitas yang diambil dari data Tahapan ini adalah sepenggal dari proses untuk menemukan pengetahuan yang terdiri dari pengecekan

pola yang sangat berbeda dari teori sebelumnya. mengidentifikasi pola-pola yang relevan dalam dataset.

## 2. Klasifikasi

Menurut (Bambang, 2017, p. 23) menyatakan bahwa klasifikasi adalah aktivitas memperkirakan data untuk menggolongkan ke dalam golongan yang telah ditetapkan dengan sejumlah golongan yang sudah ada. Rancangan kerja pada klasifikasi ditunjukkan pada Gambar 2.1. Dimana memiliki beberapa data latih  $(x,y)$  berfungsi untuk data pembangun pola. Pola berdasarkan data latih kemudian berfungsi untuk memperkirakan kelas dari data uji  $(x,?)$  sehingga diketahui kelas  $y$  yang sebenarnya



Gambar 2. 1 Proses Klasifikasi

(Bambang, 2017)

Pola yang sudah dibuat ketika pelatihan kemudian dapat digunakan untuk memprediksi label kelas data baru yang belum diketahui. Dalam pembangunan pola selama proses pelatihan memerlukan suatu algoritma untuk membangunnya yaitu Algoritma pelatihan. Ada berbagai macam algoritma pelatihan yang sudah dikembangkan oleh para peneliti yaitu C4.5, Nearest Neighbor, Bayesian Classification, Neural Network, dan sebagainya (Bambang, 2017, p. 23)

## 3. Sistem Pendukung Keputusan

Konsep sistem pendukung keputusan pertama kali oleh Michael S. Scott Morton pada tahun 1970-an dengan istilah Management Decision System, SPK dirancang untuk mendukung seluruh tahap pengambilan keputusan mulai dari mengidentifikasi masalah, memilih data dan menentukan pendekatan yang digunakan untuk pengambilan keputusan dan mengevaluasi pemilihan alternatif.

Menurut (Andi, 2016, p. 3) Sistem Pendukung Keputusan adalah sistem penghasil informasi yang ditujukan pada suatu masalah tertentu yang harus dipecahkan oleh manager dan dapat membantu manager dalam pengambilan keputusan.

Proses pengambilan keputusan dimulai dari fase inteligensi, realitas diuji, masalah diidentifikasi dan ditentukan. Kepemilikan masalah juga ditetapkan. Selanjutnya pada fase desain akan dikonstruksi sebuah model yang merepresentasikan sistem. Hal ini dilakukan

dengan membuat asumsi-asumsi yang menyederhanakan realitas dan menuliskan hubungan di antara semua variabel. Model ini kemudian di validasi dan ditentukanlah kriteria dengan menggunakan prinsip memilih untuk mengevaluasi alternatif tindakan yang telah diidentifikasi. Proses pengembangan model sering mengidentifikasi solusi-solusi alternatif dan demikian sebaliknya. Selanjutnya adalah fase pilihan, dalam tahap ini pengambil keputusan memilih salah satu alternatif pemecahan yang dibuat pada tahap Design yang dipandang sebagai aksi yang paling tepat untuk mengatasi masalah yang sedang dihadapi. Dari tahap ini didapatkan keluaran berupa dokumen Solusi dan Rencana Implementasinya. Begitu solusi yang diusulkan tampak masuk akal, maka kita siap untuk masuk kepada fase terakhir yakni fase implementasi keputusan. Hasil implementasi yang berhasil adalah dapat dipecahkannya masalah riil. Sedangkan kegagalan implementasi mengharuskan kita kembali fase sebelumnya (Andi, 2016, p. 3)

- (a) Fase Intelejensi, fase ini merupakan kegiatan mengamati lingkungan untuk mengetahui kondisi-kondisi yang perlu diperbaiki. Kegiatan ini merupakan tahapan dalam perkembangan cara berfikir. Untuk melakukan kegiatan intelijensi ini diperlukan sebuah sistem informasi, dimana informasi yang diperlukan ini didapatkan dari kondisi internal maupun eksternal sehingga seorang manajer dapat mengambil sebuah keputusan dengan tepat.
- (b) Fase Desain, fase ini merupakan sebuah kegiatan untuk menemukan, mengembangkan, dan menganalisis semua pemecahan yang mungkin, yaitu melalui pembuatan model yang bisa mewakili kondisi nyata masalah. Dari tahap ini didapatkan keluaran berupa dokumen Alternatif Solusi.
- (c) Fase Memilih, fase memilih dan menelaah ini digunakan untuk memilih satu rangkaian tindakan tertentu dari beberapa yang tersedia dan melakukan penilaian terhadap tindakan yang telah dipilih.

#### **4. Metode Prototype**

Menurut (Dwi Purnomo, 2017, p. 55) menjelaskan *prototyping* adalah cara mengembangkan sistem, yaitu model alur proses pada sebuah sistem dan berfungsi sebagai tampilan awal pada sistem yang akan dikembangkan.

Ada 4 hal utama dalam metodologi *prototyping*, yaitu :

- (a) Illustrative, mengeluarkan tampilan antarmuka sistem.
- (b) Simulated, penggambaran beberapa alur kerja sistem.
- (c) Functional, penggambaran beberapa alur kerja sistem sesuai data real
- (d) Evolutionary, memperoleh pola yang digunakan sebagai elemen dari alur sistem.

Berikut adalah manfaat dari penggunaan *prototyping* :

- (1) Menjadikan sistem sebenarnya dalam tiruan sistem yang akan dikembangkan, serta demi kesempurnaan sistem maka perlukan juga saran dari pengguna.

- (2) Setiap perubahan sampai hasil akhir pada sistem yang akan dikembangkan yang disesuaikan dengan *prototype*, pengguna akan lebih siap menerima perubahan-perubahan yang terjadi.
- (3) Pengembangan *prototype* disesuaikan dengan pengguna, kemungkinan yang terjadi yaitu *prototype* akan mengalami penambahan ataupun pengurangan dalam proses pengembangannya.
- (4) Demi menghasilkan sistem informasi yang tepat dan lebih baik bagi pengguna maka meminimalisir penggunaan sumber daya dan waktu.

## 5. Bahasa Pemrograman

### a. Hypertext Markup Language (HTML)

Menurut (Shalahuddin, 2017, p. 34) HTML atau *Hypertext Markup Language* adalah suatu format data yang digunakan untuk membuat dokumen *hypertext* (teks pada komputer yang memungkinkan *user* saling mengirimkan informasi)". Dokumen HTML harus disimpan dengan ekstensi *.htm* atau *.html*. HTML memiliki *tag* yang telah didefinisikan untuk membuat halaman *web*. Penulisan *tag-tag* HTML dapat menggunakan huruf besar atau huruf kecil, karena HTML tidak *case sensitive* (membedakan huruf besar dan huruf kecil memiliki maksud berbeda).

### b. PHP

Menurut (Arief, 2011, p. 43) PHP adalah bahasa *server-side-scripting* yang menyatu dengan HTML untuk membuat halaman *web* yang dinamis". Dengan menggunakan program PHP, sebuah *website* akan lebih interaktif dan dinamis. Kelebihan-kelebihan dari PHP yaitu:

- (1) PHP merupakan sebuah bahasa *script* yang tidak melakukan sebuah kompilasi dalam penggunaannya. Tidak seperti halnya bahasa pemrograman aplikasi yang lainnya.
- (2) PHP dapat berjalan pada *web server* yang dirilis oleh Microsoft, seperti IIS atau PWS juga pada *apache* yang bersifat *open source*.
- (3) Karena sifatnya yang *open source*, maka perubahan lebih cepat dan mudah

## 6. Database

Menurut (Krisbiantoro, 2021, p. 29) menyatakan arti database yaitu sekumpulan data yang dikumpulkan dalam tabel dan saling berintegrasi kemudian disimpan pada perangkat keras komputer dan untuk mengubahnya membutuhkan perangkat lunak.

## 7. Webserver (Apache)

Menurut (Krisbiantoro, 2021, p. 41) menjelaskan bahwa tugas utama dari webserver adalah menerima (response) permintaan (request) dari client (browser). Request tersebut merupakan konten statis berupa halaman HTML yang dikirimkan oleh browser melalui

Hypertext Transfer Protocol (HTTP). Webserver menanggapi request dengan mengirimkan kembali jawaban ke browser melalui HTTP.

## 8. Pengembangan Sistem SDLC

SDLC (*Systems Development Life Cycle*) merupakan siklus hidup pengembangan system. Dalam rekayasa sistem dan rekayasa perangkat lunak, SDLC berupa suatu proses pembuatan dan perubahan sistem serta model dan metodologi yang digunakan untuk mengembangkan sistem-sistem tersebut. Dalam rekayasa perangkat lunak, konsep SDLC mendasari berbagai jenis metodologi pengembangan perangkat lunak. Metodologi ini membentuk suatu kerangka kerja untuk perencanaan dan pengendalian pembuatan sistem informasi, yaitu proses pengembangan perangkat lunak. Pendekatan sistem merupakan sebuah metodologi. Metodologi adalah satu cara yang direkomendasikan dalam melakukan sesuatu. Pendekatan sistem adalah metodologi dasar dalam memecahkan segala jenis masalah. Kemudian SDLC juga Untuk menghasilkan produk dan sistem perangkat lunak yang terbaik, diperlukan perancangan dan rencana kerja yang maksimal. Penggunaan metode SDLC membantu memenuhi keperluan tersebut. Dengan SDLC durasi proses produksi dapat dipersingkat, mengurangi biaya operasional, dan meningkatkan dan menjaga kualitas produk.

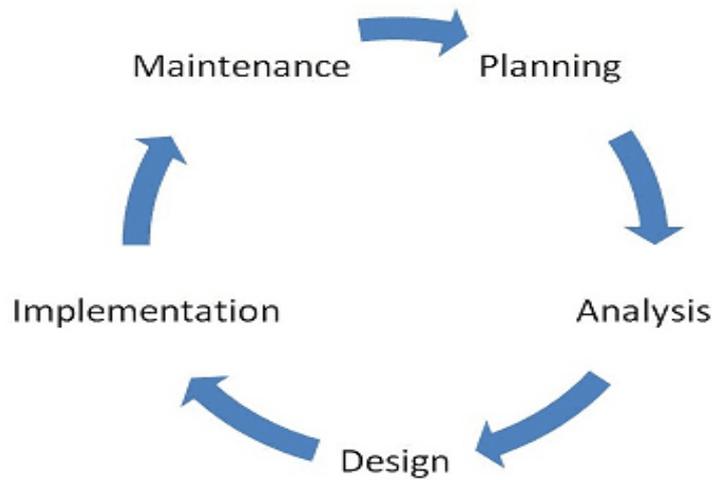
Siklus hidup pengembangan sistem (*System Development Life Cycle – SDLC*) adalah aplikasi dari pendekatan sistem bagi pengembangan suatu sistem informasi (Setiawan, 2022, p. 7) Model siklus hidup digunakan untuk merencanakan kegiatan pengembangan, merancang, membangun, hingga memelihara aplikasi dari kesalahan yang menyebabkan tingginya biaya dan perbaikan ulang. Pengembangan SDLC adalah proses yang digunakan oleh analis sistem untuk mengembangkan sistem informasi, termasuk persyaratan, validasi, pelatihan, dan pengguna (*stakeholder*) kepemilikan. Tahapan-tahapan dalam SDLC bisa dilihat pada gambar 2.2.

Ada berbagai macam urutan dalam pengembangan apabila proyek tersebut kemungkinan akan berhasil. Urutan yang dapat dilakukan sebagai berikut :

- (1) Merencanakan
- (2) Melakukan analisis
- (3) Pembuatan desain
- (4) Melakukan implementasi
- (5) Pemakaian

Sebelum melaksanakan proyek, maka perlu merencanakan sumberdaya apa saja yang diperlukan. Melakukan analisis terhadap sistem yang sudah serta menentukan permasalahan dan kebutuhan pada sistem baru yang akan dikembangkan. Kemudian merancang dan menerapkannya terhadap sistem baru. Kemudian sistem baru tersebut digunakan dan diharapkan dapat bertahan dalam waktu yang cukup panjang. Karena

urutandi atas mengikuti struktur yang sistematis dan dikerjakan dengan cara *top-down*, umumnya proses air terjun (*waterfall approach*) sering disebut sebagai SDLC tradisional. Proses ini mempunyai aliran satu arah dalam menyelesaikan proyek.



Gambar 2. 2 Konsep SDLC

(Setiawan, 2022)

Pada Gambar 2.2 di atas menggambarkan sebuah siklus dengan sifat melingkar. untuk membangun sistem dan memberikannya kepada pengguna melalui tahapan perencanaan, analisa, perancangan dan implementasi (penerapan) dengan cara memahami dan menyeleksi keadaan dan proses yang dilakukan pengguna untuk dapat mendukung kebutuhan pengguna. Untuk menggunakan SDLC maka dibutuhkan sumber data awal dari pengguna yang dijadikan acuan dalam perencanaan, analisa, perancangan dan implementasi. Penggunaan acuan ini dimaksudkan agar sistem yang dibangun bisa menjembatani kebutuhan pengguna dari permasalahan yang dihadapinya. Setelah sistem tersebut telah berhasil diimplementasikan maka untuk langkah selanjutnya melakukan perawatan terhadap sistem agar dapat berjalan sebagaimana mestinya.

## 9. Metode Naive Bayes

Algoritme Naive Bayes yaitu cara memperkirakan sesuatu yang berbasis probabilistik sederhana dengan mengacu pada teorema Bayes (atau aturan Bayes) yang memiliki dugaan ketidaktergantungan yang kuat. Naive Bayes model menggunakan “model fitur independent”. Pada Naive Bayes arti ketidaktergantungan yaitu bahwa sebuah karakteristik data tidak saling berhubungan meskipun terdapat atau tidaknya karakteristik lain pada data yang serupa. (Prasetyo, 2021, p. 7)

Teorema keputusan bayes adalah metode statistik dasar pengenalan pola. Metode ini didasarkan pada penggunaan probabilitas dan biaya yang dihasilkan oleh berbagai keputusan untuk mengukur trade-off antara berbagai keputusan klasifikasi.

Adapun Langkah Langkah perhitungan Naive, yaitu :

- (a) Persiapan data: Pertama-tama, kumpulkan dan bagi data menjadi data latih dan data uji.
- (b) Hitung probabilitas kelas: Hitung jumlah kelas yang ada pada data latih dan hitung frekuensi kelas yang muncul.
- (c) Hitung probabilitas fitur: Setelah itu, hitung probabilitas setiap fitur untuk setiap kelas.
- (d) Prediksi kelas: Setelah memiliki probabilitas setiap kelas dan fitur, selanjutnya adalah memprediksi kelas dari data uji

Tentukan kelas dengan probabilitas terbesar: Kemudian, dari hasil prediksi kelas, tentukan kelas dengan probabilitas terbesar sebagai kelas akhir dari data uji tersebut

Menurut (Prasetyo, 2021, p. 7) prediksi Bayes didasarkan pada teorema Bayes dengan Bentuk sebagai berikut

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

Penjelasan :

X : data kelas yang belum diketahui

H : data hipotesis

$P(H|X)$  : jumlah probabilitas hipotesis H berdasarkan kondisi X

$P(X|H)$  : jumlah probabilitas posterior X berdasarkan kondisi hipotesis H

$P(H)$  : jumlah probabilitas prior hipotesis H

$P(X)$  : jumlah probabilitas prior bukti X

Artinya bahwa Penentuan class dilakukan dengan membandingkan nilai probabilitas dari suatu sampel berada di class yang satu dengan nilai probabilitas suatu sampel berada di class yang lain. Untuk menentukan class yang cocok dari suatu sampel yaitu melakukan perbandingan antara nilai posterior pada setiap class, pada nilai posterior tertinggi maka class tersebut yang menjadi hasil akhir.

Menurut (Khanif, 2022, p. 67) untuk menentukan Kelayakan Pengajuan Kredit Sepeda Motor dengan menerapkan metode Naive Bayes terdapat 15 data training dan data yang telah di klasifikan berdasarkan status pemilik rumah, Umur, Karakter dan Penghasilan dengan keputusan Layak dan Tidak dalam pengajuan kredit setelah dilakukan dalam mengambil keputusan atau prediksi dalam pengklasifikasian akan menghasilkan output Layak atau Tidak. Adapun data training dalam tabel dibawah ini :

Tabel 2. 1 Data Training Contoh Kasus

No	Nama	Status Rumah	Umur	Penghasilan	Karakter	Kredit
1	Farhan	MILIK SENDIRI	TUA	RENDAH	Baik	Ya
2	Muhamad Fauji	SEWA	MUDA	RENDAH	Cukup	Ya
3	Adam	SEWA	TUA	RENDAH	Baik	Tidak
4	Budi	MILIK SENDIRI	TUA	TINGGI	Cukup	Ya
5	Dedi	SEWA	TUA	RENDAH	Cukup	Tidak
6	Evan	SEWA	MUDA	TINGGI	Cukup	Ya
7	Fany	MILIK SENDIRI	TUA	SEDANG	baik	ya
8	Heny intan	MILIK SENDIRI	MUDA	TINGGI	baik	ya
9	Faqih	MILIK SENDIRI	TUA	TINGGI	Cukup	ya
10	Citra	SEWA	TUA	SEDANG	Cukup	ya
11	Andi	SEWA	MUDA	TINGGI	Baik	Tidak
12	Nisa	MILIK SENDIRI	MUDA	TINGGI	Baik	ya
13	Khoffifah	SEWA	TUA	RENDAH	Cukup	Tidak
14	Ulfah	MILIK SENDIRI	MUDA	SEDANG	Cukup	ya
15	Darul	MILIK SENDIRI	TUA	TINGGI	Baik	ya

No	Nama	Status Rumah	Umur	Penghasilan	Karakter	Kredit
16	Muhamad Fadil	Milik Sendiri	Muda	Sedang	Cukup	?

**Penyelesaian :**

1. Menghitung probabilitas prior untuk kelas pertama

$$P(\text{Kredit} = \text{Ya}) = 11/15 = 0.7333$$

$$P(\text{Kredit} = \text{Tidak}) = 4/15 = 0,2667$$

- b. Menghitung probabilitas bersyarat untuk setiap kelas :

$$P(\text{status rumah} = \text{Milik sendiri} \mid \text{Layak} = \text{Ya}) = 8/11 = 0,0485$$

$$P(\text{status rumah} = \text{Milik sendiri} \mid \text{Layak} = \text{Tidak}) = 0/4 = 0$$

$$P(\text{Umur} = \text{Muda} \mid \text{Layak} = \text{Ya}) = 5/11 = 0.4545$$

$$P(\text{Umur} = \text{Muda} \mid \text{Layak} = \text{Tidak}) = 1/4 = 0,2525$$

$$P(\text{Penghasilan} = \text{Sedang} \mid \text{Layak} = \text{Ya}) = 3/11 = 0,2727$$

$$P(\text{Penghasilan} = \text{Sedang} \mid \text{Layak} = \text{Tidak}) = 0/4 = 0$$

$$P(\text{Karakter} = \text{Cukup} \mid \text{Layak} = \text{Ya}) = 6/11 = 0,5455$$

$$P(\text{Karakter} = \text{Cukup} \mid \text{Layak} = \text{Tidak}) = 2/4 = 0.5$$

- c. Menghitung probabilitas untuk setiap kelas

$$P(X \mid \text{Layak} = \text{Ya}) = 0,0485 \times 0,4545 \times 0,2727 \times 0,5455 = 0,0385$$

$$P(X \mid \text{Layak} = \text{Tidak}) = 0 \times 0,2525 \times 0 \times 0.5 = 0$$

- d. Menghitung probabilitas posterior

$$P(X \mid \text{Layak} = \text{Ya}) \times (P \mid \text{Bonus} = \text{Ya}) = 0,0385 \times 0.7333 = 0,282$$

$$P(X \mid \text{Layak} = \text{Tidak}) \times (P \mid \text{Bonus} = \text{Tidak}) = 0 \times 0,2667 = 0$$

Tabel 2. 2 Data Sampel Contoh Kasus

<b>Nama</b>	<b>Status Rumah</b>	<b>Umur</b>	<b>Penghasilan</b>	<b>Karakter</b>	<b>Layak</b>
Muhamad Fadil	Milik Sendiri	Muda	Sedang	Cukup	YA

Pada  $P(X|Layak =Ya) \times (P|Layak=Ya) = 0,282$  lebih besar dibanding  $P(X|Layak=Tidak) \times (P|Layak =Tidak) = 0$ , maka Naive Bayes Classifier mengklasifikasikan bahwa pelanggan yang bernama Muhamad Fadil ke dalam kelas Layak = “Ya”.

### C. Tinjauan Studi

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

- (1) **Wijayanti ,Sulastri 2021 yang berjudul “Analisa klasifikasi kartu kredit menggunakan algoritma naive bayes”** (Wijayanti & Sulastri , 2021)

Berdasarkan analisa pada penelitian ini, ketika menggunakan 23 variabel didapatkan akurasi klasifikasi untuk 18001 data yaitu 69.11%. akurasi ini masih minim sehingga penulis mencoba mengkomputasikan menggunakan 11 variabel dan mendapatkan tingkat akurasi data sebanyak 79.7%. Untuk mendapatkan keakurasian data yang lebih, peneliti mencoba mengkomputasikan menggunakan 6 variabel dan mendapat akurasi sebanyak 80.47%. Sedangkan jika menggunakan 4 variabel akurasi data sebanyak 80.52%. b. Dan terakhir menggunakan 3 variabel memiliki akurasi data sebanyak 80.59%. Akurasi tersebut lebih baik jika dibandingkan klasifikasi terhadap semua variabel. Dimana variabel tersebut adalah X6, X7, X10 yaitu status pembayaran selama bulan September, Agustus, dan Mei 2005 dan terdapat 10 rule Hal ini disebabkan karena semua variabel tidak digunakan, sehingga mempengaruhi proses klasifikasi data.

- (2) **Annur, 2020 yang berjudul “Klasifikasi Masyarakat Miskin Menggunakan**

### **Metode Naïve Bayes” (Annur, 2020)**

Permasalahan utama dalam upaya pengurangan kemiskinan saat ini terkait dengan adanya fakta bahwa pertumbuhan ekonomi tidak tersebar secara merata. Penelitian akan melakukan klasifikasi berdasarkan data penduduk miskin yang diperoleh dari Kecamatan Tibawa dengan menggunakan teknik data mining. Atribut yang akan digunakan dalam melakukan klasifikasi penduduk adalah Umur, Pendidikan, Pekerjaan, Penghasilan, Tanggungan, Status (Kawin/Belum Kawin). Metode yang akan digunakan adalah metode Naïve Bayes Classifier, yang merupakan salah satu teknik pengklasifikasian dalam data mining. Berdasarkan penelitian yang dilakukan dihasilkan kesimpulan bahwa, Sistem klasifikasi masyarakat miskin di wilayah pemerintahan Kecamatan Tibawa Kab.Gorontalo dapat direkayasa dan Berdasarkan hasil pengujian confusion matrix dengan teknik split validasi, penggunaan metode klasifikasi naïve bayes terhadap dataset yang telah diambil pada objek penelitian diperoleh tingkat akurasi sebesar 73% atau termasuk dalam kategori Good. Sementara nilai Precision sebesar 92% dan Recall sebesar 86%.

### **(3) Saputra,2020 yang berjudul “Penerapan Algoritma Naïve Bayes Untuk Penentuan Calon Penerima Beasiswa Pada Smk Pasim Plus Sukabumi”**

(Saputra, 2020)

Kriteria yang ditetapkan antara lain indeks prestasi akademik, penghasilan orangtua, jumlah saudara kandung, jumlah tanggungan orangtua, semester dan lain lain. Dalam penelitian ini dilakukan pengujian model algoritma *Naïve Bayes*, hasil evaluasi dan validasi, diketahui bahwa *Naïve Bayes* memiliki nilai *accuracy* dan AUC cukup tinggi yaitu sebesar 96,67%. Dengan demikian, metode *Naïve Bayes* merupakan metode yang cukup baik dalam menentukan calon penerima beasiswa hasil evaluasi dan validasi, diketahui bahwa *Naïve Bayes* memiliki nilai *accuracy* dan AUC cukup tinggi yaitu sebesar 96,67%. Dengan demikian, metode *Naïve Bayes* merupakan metode yang cukup baik dalam menentukan calon penerima beasiswa secara lebih efektif dan efisien

### **(4) Bustami , 2019 yang berjudul “Penerapan Algoritma Naïve Bayes Untuk Mengklasifikasi Data Nasabah Asuransi” (Bustami, 2019)**

Sistem klasifikasi data nasabah ini digunakan untuk menampilkan informasi klasifikasi lancar, kurang lancar atau tidak lancarnya calon nasabah dalam membayar premi asuransi dengan menggunakan algoritma *Naïve Bayes*. Dengan adanya sistem ini maka mempermudah pihak asuransi dalam memperkirakan nasabah yang bergabung, sehingga perusahaan bisa mengambil keputusan untuk menerima atau menolak calon nasabah tersebut. Algoritma *Naïve Bayes* di dukung oleh ilmu Probabilistik dan ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian. Pada algoritma *Naïve Bayes*, semua atribut akan memberikan

kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain. Variabel penentu yang digunakan dalam penelitian ini adalah jenis kelamin, usia, status, pekerjaan, penghasilan per tahun, masa pembayaran asuransi, dan cara pembayaran asuransi.

- (5) **Rahman & Firdaus, 2022 yang berjudul “Penerapan Data Mining Metode Naïve Bayes Untuk Prediksi Hasil Belajar Siswa Sekolah Menengah Pertama (Smp)”** (Rahman & Firdaus, 2022)

Berdasarkan hasil pengujian klasifikasi hasil belajar siswa pada penelitian ini disimpulkan bahwa algoritma Naïve Bayes memiliki kinerja yang baik dari segi akurasi, presisi dan recall walaupun dari jumlah akurasi yang hanya 56,79% masih tergolong hasil yang rendah dibandingkan beberapa model data mining yang lainnya. Hal ini karena banyaknya input tidak terlalu relevan algoritma yang diujikan tidak terlalu sensitif dengan input yang tidak terlalu relevan tersebut.

- (6) **Khadhamsyah, 2022 yang berjudul “Sistem Pendukung Keputusan Penetapan Peserta Sertifikasi Guru Menggunakan Naïve Bayes”** (Kadhamsyah, 2022)

kriteria yang digunakan dalam penetapan peserta sertifikasi guru di dinas pendidikan kabupaten jombang diantaranya : masa kerja, usia, golongan, beban kerja, tugas tambahan, dan prestasi kerja output yang di keluarkan adalah dapat memilih guru yang memenuhi kriteria yang telah di uji dalam sistem. Sertifikasi guru adalah proses pemberian sertifikasi kepada guru yang telah memenuhi standar profesional guru. Guru profesional merupakan syarat mutlak untuk menciptakan sistem dan praktik pendidikan yang bermutu. Guru profesional juga harus memiliki kualifikasi akademik minimum sarjana (S-1) atau diploma empat (D-IV), menguasai kompetensi (pedagogic, profesional, social dan kepribadian), memiliki sertifikat pendidik, sehat jasmani dan rohani dengan nilai akurasi 87 %.

- (7) **Buani, 2020 yang berjudul “Prediksi Penyakit Hepatitis Menggunakan Algoritma Naïve Bayes Dengan Seleksi Fitur Algoritma Genetika”** (Buani, 2020)

Penelitian ini dilakukan untuk menguji hasil prediksi dari algoritma Naïve Bayes dengan seleksi fitur algoritma genetika, dah hasil prediksi yang didapatkan dalam pengujian ini adalah 96,77% hasil ini meningkat dari penelitian yang sebelumnya menggunakan data yang sama dan algoritma yang sama yaitu algoritma naïve bayes hasil prediksinya adalah 83, 71%, selisih dari penelitian sebelumnya dengan penelitian ini adalah 13.06%, selisih ini membuktikan bahwa tingkat dari algoritma naïve bayes setelah dilakukan seleksi fitur menggunakan algoritma genetika tingkat akurasinya lebih baik.

- (8) **Sugianto, 2019 yang berjudul “Kombinasi Metode Simple Additive Weigthing dan Weigthed Product Untuk Seleksi Proposal Program Kreatifitas Mahasiswa”** (Sugianto, 2019)

Dari uraian penelitian ini, maka dapat diambil kesimpulan, berdasarkan dari hasil yang telah dilakukan bahwa pemilihan kriteria untuk pengambilan rekomendasi, memerlukan beberapa kriteria sebagai pertimbangan. Semakin banyak kriteria yang digunakan maka akan menghasilkan pilihan rekomendasi yang tepat sebagai acuan pemilihan kelayakan proposal PKM. Dari penerapan metode Modified Simple Additive Weigthing (MOD SAW), metode SAW digunakan sebagai proses pembobotan dan normalisasi, sedangkan dalam metode Weigthed Product (WP) digunakan untuk perbaikan nilai bobot, menentukan preferensi alternatif dan perankingan. Hasil rekomendasi yang terpilih adalah nilai V yang paling besar yang akan menjadi alternatif terpilih berdasarkan kriteria yang dipilih oleh tim seleksi penerimaan proposal PKM. Dari hasil rekomendasi diatas, diperoleh hasil bahwa dari 5 proposal mahasiswa, diperoleh nilai perankingan yang memiliki angka paling tinggi adalah V5 dengan nilai 0,0717.

- (9) **Robi Yanto & Deni Apriadi, 2022 yang berjudul “PENERAPAN METODE WEIGHTED PRODUCT UNTUK SELEKSI KELAYAKAN PROPOSAL PROGRAM KREATIVITAS MAHASISWA”** (Robi Yanto & Deni Apriadi, 2022)

Berdasarkan hasil analisis menggunakan metode weighted product dan sistem yang dibangun dapat diambil kesimpulan bahwa dari data tahun 2017 usulan proposal PKM-Pengabdian kepada Masyarakat, yang layak untuk direkomendasikan mengikuti seleksi hibah DIKTI berdasarkan hasil perankingan adalah PKM Sosialisasi Masyarakat Melati (Melek Teknologi dan Informasi) dengan nilai 0,27308.

- (10) **Adi, 2020 yang berjudul “IMPLEMENTASI ALGORITMA NAÏVE BAYES CLASSIFIER UNTUK KLASIFIKASI PENERIMA BEASISWA PPA DI UNIVERSITAS AMIKOM YOGYAKARTA”** (Adi, 2020)

Setiap tahun Universitas menawarkan banyak jenis beasiswa kepada mahasiswanya, tidak terkecuali Amikom Universitas Yogyakarta juga menawarkan jenis beasiswa kurang mampu (PPA) untuk mahasiswa di membutuhkan. Setiap tahun jumlah penerima beasiswa terus meningkat, namun jumlahnya diterima setiap tahun tetap. Oleh karena itu, perlu dikembangkan suatu sistem untuk melakukan data mining dari tumpukan data yang akan digunakan untuk tujuan tertentu, salah satunya untuk menganalisis kelayakan penerima beasiswa agar tepat sasaran. Pengklasifikasi Naive Bayes adalah pendekatan yang mengacu pada Teorema Bayes yang menggabungkan pengetahuan sebelumnya dengan pengetahuan baru. Jadi itu salah satunya algoritma klasifikasi yang sederhana namun memiliki akurasi yang tinggi. Oleh karena itu, dalam penelitian ini akan membuktikan kemampuan Naive Bayes

Classifier untuk mengklasifikasikan data pelamar beasiswa yang menginformasikan kelayakan pemberian beasiswa PPA. Data beasiswa dari beasiswa adalah dilakukan preprocessing, agar data menjadi “clean”, sehingga layak untuk dilakukan proses selanjutnya. Kemudian preprocessing diklasifikasikan dengan Naive Bayes Classifier, menghasilkan klasifikasi model probabilitas untuk penentuan kelas pelamar beasiswa berikutnya. Dari hasil pengujian training yang dilakukan, dapat disimpulkan bahwa semakin banyak data yang dapat diolah maka semakin besar nilai akurasi kelas.

Tabel 2. 3 Penelitian Rujukan

No		Judul Penelitian	Deskripsi	Link	Kontribusi
1	(Wijayanti & Sulastri , 2021)	Analisa Klasifikasi Kartu Kredit Menggunakan Algoritma Naive Bayes	Kredit yang bermasalah atau kredit acet Sering terjadi akibat analisis kredit kurang cermat dalam proses pemberiankredit.	<a href="https://www.unisbank.ac.id/ojs/index/fti2/article/view/8137">https://www.unisbank.ac.id/ojs/index/fti2/article/view/8137</a>	penilaian beberapa atribut untuk ditambahkan divaiabel
2	(Annur, 2020)	Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes	Permasalahan utamadalam upaya pengurangan kemiskinan ini terkait dengan adanya fakta bahwa pertumbuhan ekonomitidak tersebarsecara merata.	<a href="https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/303">https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/303</a>	Penilaian beberapa atribut untuk ditambahkan divaiabel yang di ambil

3	(Saputra, 2020)	<p>Penerapan Algoritma NaïveBayes Untuk Penentuan Calon Penerima Beasiswa Pada Smk Pasim sukabumi</p>	<p>tidakadanya sistem yang dapat membantu untuk melakukan penyeleksian atau penentuan penerima beasiswa.</p>	<p><a href="https://ejournal.bsi.ac.id/ejournal/index.php/swabumi/article/view/5401">https://ejournal.bsi.ac.id/ejournal/index.php/swabumi/article/view/5401</a></p>	<p>Penilaian dalam pengambilan atribut untuk ditambahkan di variabel</p>
4	(Bustami, 2019)	<p>Penerapan Algoritma NaiveBayes Untuk Mengklasifikasi Data Nasabah Asuransi</p>	<p>Perhitungan Masih secara Manual Sehingga diperlukan sebuah sistem yang dapat mengklasifikasi nasabah mana yang masuk dalam kelompok lancar, kelompok kurang lancar dan nasabah mana yang masuk ke dalam kelompok tidak lancar dalam membayar iuran</p>	<p><a href="https://ojs.unimal.ac.id/techsi/article/view/154">https://ojs.unimal.ac.id/techsi/article/view/154</a></p>	<p>Kontribusi: pengambilan atribut untuk ditambahkan di variabel</p>

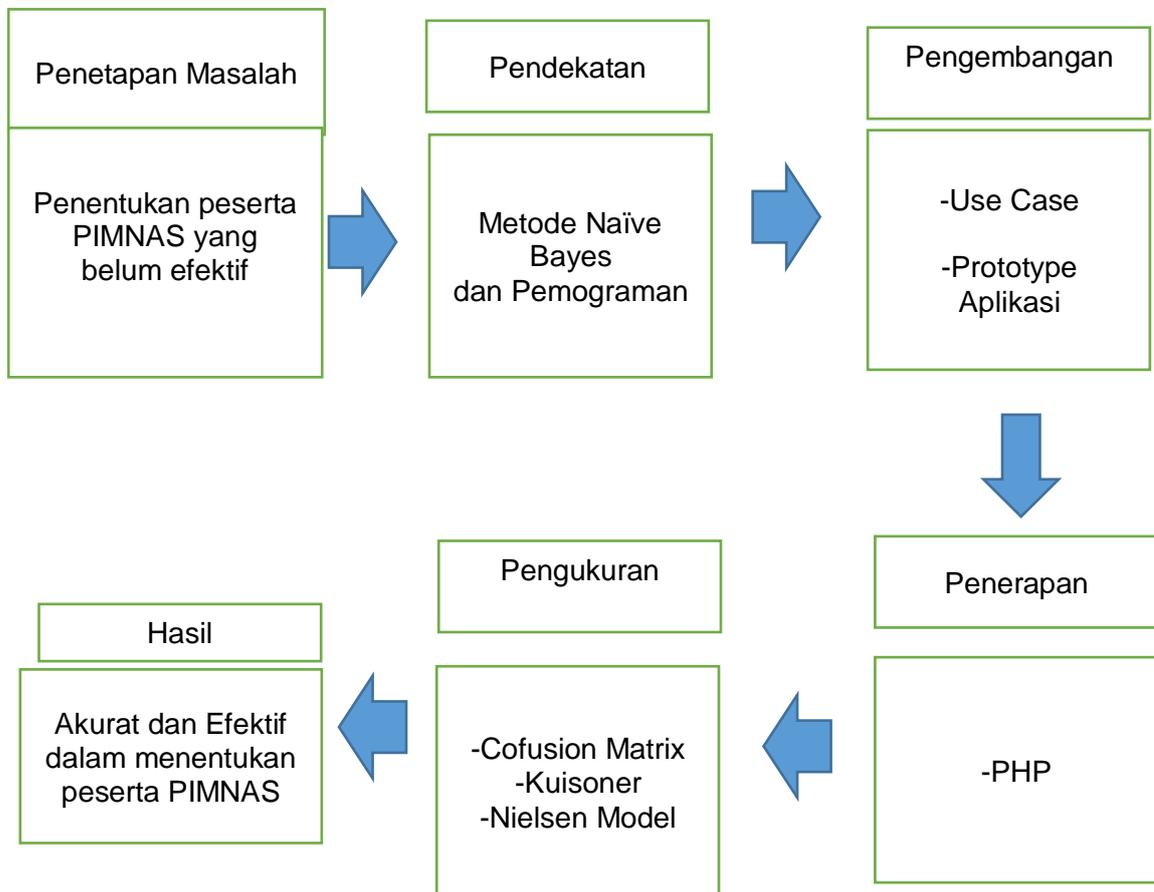
5	(Rahman & Firdaus, 2022)	Penerapan DataMining Metode Naïve Bayes Untuk Prediksi Hasil Belajar Siswa Sekolah	Sulitnya menentukan faktoratau variabel yang mempengaruhi hasil belajar siswa.	<a href="https://ojs.uniska-bjm.ac.id/index.php/JST/article/view/436">https://ojs.uniska-bjm.ac.id/index.php/JST/article/view/436</a>	Penilaian beberapa atribut untuk ditambahkan kevariabel-variabel yang di angkat .
6	(Kadhamsyah, 2022)	Sistem Pendukung Keputusan Penerapan Peserta Sertifikasi Guru Menggunakan Naive bayes	Masalah efektifitas, efisiensi dan standarisasi pengajaryang mempengaruhi kualitas guru.	<a href="http://etheses.uin-malang.ac.id/50008/">http://etheses.uin-malang.ac.id/50008/</a>	pengambilan atribut untuk ditambahkan divariabel Kelemahan: tidak ada perhitungan
7	(Buani, 2020)	Prediksi Penyakit Hepatitis Menggunakan Algoritma Naïve Bayes Dengan Seleksi Fitur Algoritma Genetika	Salah satu yang dapat diaplikasikan untuk memprediksi penyakit hepatitis	<a href="https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/4381">https://ejournal.bsi.ac.id/ejurnal/index.php/evolusi/article/view/4381</a>	penilaian dalam pengambilan atribut untuk ditambahkan divariabel

8	(Sugianto, 2019)	Kombinasi Metode Simple Additive Weigthing dan Weigthed Product Untuk Seleksi Proposal Program Kreatifitas Mahasiswa	Sulitnya menentukan Mahasiswa atau variabel yang mempengaruhi hasil dalam Pengajuan Proposal Pimnas.	<a href="https://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/2929">https://www.ejurnal.stmik-budidarma.ac.id/index.php/mib/article/view/2929</a>	Penilaian beberapa atribut untuk ditambahkan kevariabel-variabel yang di angkat
9	(Robi Yanto & Deni Apriadi , 2022)	PENERAPAN METODE WEIGHTED PRODUCT UNTUK SELEKSI KELAYAKAN PROPOSAL PROGRAM KREATIVITAS MAHASISWA	Tidak aja uji validitas terhadap kriteria yang digunakan untuk seleksi proposal PKM-Pengabdian kepada Masyarakat sehingga dapat diperoleh kriteria yang sesuai dengan kebutuhan seleksi	<a href="https://pdfs.semanticscholar.org/7ea77b4da0934db6ce4ce29b9e6cdf5419ee1547.pdf">https://pdfs.semanticscholar.org/7ea77b4da0934db6ce4ce29b9e6cdf5419ee1547.pdf</a>	penelitian ini adalah pengambilan uji hasil menggunakan Weighted Product.

10	(Adi, 2020)	IMPLEMENTASI ALGORITMA NAÏVE BAYES CLASSIFIER UNTUK KLASIFIKASI PENERIMA BEASISWA PPA DI UNIVERSITAS AMIKOM YOGYAKARTA	<p>Setiap tahun Univeritas menawarkan banyak jenis beasiswa ke mahasiswanya, tidak terkecuali Universitas Amikom Yogyakarta juga menawarkan jenis beasiswa kurang mampu (PPA) bagi mahasiswa yang membutuhkan.</p> <p>Setiap tahun jumlah pendaftar beasiswa ini terus mengalami peningkatan, namun jumlah yang diterima setiap tahunnya tetap. Dengan demikian, perlu dikembangkan sistem untuk melakukan data mining dari tumpukan data tersebut yang akan digunakan</p>	<a href="https://e-jurnal.pelitanusanara.ac.id/index.php/antik/article/view/342">https://e-jurnal.pelitanusanara.ac.id/index.php/antik/article/view/342</a>	<p>Data pendaftar n beasiswa dilakukan preprocessing, agar data tersebut menjadi “bersih”, sehingga layak untuk penentuan kelas pada pendaftar beasiswa selanjutnya</p> <p>Beberapa kriteria antara lain: IPK, Keaktifan Organisasi, Prodi, Fakultas, Semester, Penghasilan orang tua, Jumlah tanggungan orang tua, pekerjaan orang tua, dan status penerimaan beasiswa</p>
----	-------------	--	--	---	---

#### D. Kerangka Pemikiran

Berikut adalah kerangka pemikiran untuk memecahkan masalah penelitian ini yang dapat dilihat pada Gambar 2.3



Gambar 2. 3 Kerangka Pemikiran

Kerangka pemikiran pada gambar 2.3 dapat dijelaskan sebagai berikut :

- (1) Penetapan Masalah Fenomena yaitu .Belum Akurat dalam Menentukan Kelayakan peserta pekan Ilmiah Mahasiswa Nasional
- (2) Metode yang digunakan yaitu Naïve Bayes.
- (3) Dalam pengembangan yaitu tahap analisis dan tahap pembuatan sistem informasi.
- (4) Hasilnya yang di dapatkan adalah sebuah sistem informasi untuk menentukan Kelayakan peserta pekan Ilmiah Mahasiswa Nasional berdasarkan Nilai Akhir(NA), Gagasan ,Inovasi , Keunggulan Produk / jasa dan Kemampuan Teknis dengan menerapkan metode naïve bayes.

## E. Hipotesis Penelitian

Berdasarkan permasalahan yang dihadapi yaitu belum efektif dan belum akurat pada saat melakukan proses pemilihan peserta pimnas, maka perlu adanya cara dalam memecahkan masalah tersebut. Dalam teori data mining ada berbagai macam metode untuk melakukan pengklasifikasian dalam memperkirakan kejadian di masa depan berdasarkan pengalaman di masa lampau, salah satunya adalah metode *Bayesian Classification*. Naive Bayes yaitu akan menghasilkan hasil prediksi yang telah ditentukan nilai probabilitas akhir dari setiap kelas atau label dan melihat hasil akhir yang terbesar dari setiap kelas. Penelitian yang telah dilakukan oleh (Alfiah, 2018) berjudul "Analisis sistem pendukung keputusan dalam menyeleksi proposal Program Kreativitas Mahasiswa Penelitian Terbaik di AMIK Tunas Bangsa" atribut yang digunakan oleh Alfiah yaitu Nilai Akhir, Kreativitas, dan kemanfaatan. Berdasarkan hal tersebut, maka dapat ditetapkan hipotesis adalah metode Naive Bayes diduga dapat memberikan rekomendasi pemilihan peserta pekan ilmiah mahasiswa nasional.