

## BAB II KERANGKA TEORITIS

### A. Landasan Teori

#### 1. Data Mining

Menurut Larose (2014), dikutip dari buku dengan judul "*Discovering Knowledge In Data: An Introduction To data Mining*" menjelaskan bahwa data mining adalah kegiatan menganalisis sejumlah data yang besar. Setelah data diproses, korelasi, dan pola baru ditemukan dengan cara yang berbeda dari masa lalu, yaitu penggunaan teknik pengenalan pola dan teknik statistik dan analisis untuk memfilter sejumlah data yang besar yang disimpan dalam data repositori. Dengan teknik matematika yang bisa dimengerti dan berguna untuk pemilik data.

Menurut Tan (2006), serangkaian proses dibagi menjadi beberapa tahapan sebagai berikut:

- a. Pembersihan data yaitu untuk menghilangkan data dan menghapus data yang tidak sesuai
- b. Integrasi data yaitu kombinasi data dari berbagai sumber
- c. Transformasi data yaitu mengubah menjadi bentuk yang dapat di proses mining
- d. Aplikasi teknik data mining yaitu proses mengestrak pola dari data yang sudah ada
- e. Evaluasi pola yang didapat yaitu proses menafsirkan pola sebagai pengetahuan yang digunakan untuk pengambilan pendukung keputusan.
- f. Presentasi pengetahuan yaitu dengan memberikan pengetahuan.

Menurut Tan (2006), tahapan ini merupakan bagian proses menemukan pengetahuan, yang meliputi pengecekan pola atau informasi yang ditemukan bertentangan dengan fakta atau asumsi sebelumnya. Langkah terakhir adalah menyajikan pengetahuan dalam bentuk yang mudah dipahami.

#### 1. Guru Berprestasi

Guru yang merupakan inti dari peningkatan pendidikan nasional harus diberi penghargaan atas prestasi dan dedikasinya. Hal ini sesuai amanat Undang-Undang Republik Indonesia Nomor 14 Tahun 200

tentang Guru dan Dosen, Pasal 36 ayat (1) bahwa "Guru yang berprestasi, berdedikasi luar biasa, dan/atau bertugas di daerah khusus berhak memperoleh penghargaan" dan berdasarkan Peraturan Pemerintah Republik Indo

Nomor 74 Tahun 2008 Pasal 30 ayat (1) bahwa “Guru memiliki hak untuk mendapatkan penghargaan sesuai dengan prestasi kerja, dedikasi luar biasa, dan/atau bertugas di Daerah Khusus”.

## 2. Metode Naïve Bayes

Algoritme Naive Bayes adalah klasifikasi probabilitas sederhana, yang menghitung sekumpulan probabilitas dengan menambahkan kombinasi frekuensi dan nilai dalam kumpulan data tertentu. Algoritma ini menggunakan teorema Bayes dan mengasumsikan bahwa semua atribut adalah independen atau tidak bergantung satu sama lain, diberikan oleh nilai variabel kelas. Naive Bayes didasarkan pada asumsi yang disederhanakan bahwa nilai atribut tidak bergantung secara kondisional satu sama lain ketika nilai keluaran diberikan. Dengan kata lain, mengingat nilai keluaran, probabilitas pengamatan kolektif adalah produk dari probabilitas individu. Keuntungan menggunakan Naive Bayes adalah metode ini hanya membutuhkan sedikit data pelatihan untuk menentukan estimasi parameter yang diperlukan dalam proses klasifikasi. Naive Bayes biasanya berkinerja jauh lebih baik dari yang diharapkan dalam situasi dunia nyata yang paling kompleks (Karthika & Sairam, 2015).

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.

Menurut Syarli metode Naive Bayes ini memiliki kekuatan dan kelemahan, yaitu :

### a) Kelebihan

Mudah diimplementasikan dan Memberikan hasil yang baik untuk banyak kasus.

### b) Kekurangan

Harus mengasumsi bahwa antara fitur tidak terkait (*independent*). Dalam realita keterkaitan itu ada. Keterkaitan itu tidak dapat dimodelkan oleh Naive Bayes Classifier.

### a. Persamaan Naive Bayes

Menurut Alfa Saleh (2015) 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 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  $X$

Penjelasan metode Naive Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apa yang cocok bagi sampel yang dianalisis tersebut. Karena itu, metode Naive Bayes di atas disesuaikan sebagai berikut :

$$P(C|F1...Fn) = \frac{P(C)P(F1...Fn|C)}{(F1...Fn)}$$

Variabel  $C$  mewakili kelas, dan variabel  $F1 \dots Fn$  mewakili karakteristik dari instruksi yang diperlukan untuk melakukan klasifikasi. Kemudian, rumus tersebut diartikan bahwa peluang suatu sampel fitur jenis tertentu memasuki kelas  $C$  (posterior) adalah peluang munculnya kelas  $C$  (sebelum sampel masuk, biasa disebut prior) dikalikan dengan kemunculan fitur sampel di kelas  $C$  (juga dikenal sebagai *Is the likelihood*) dibagi dengan kemungkinan kemunculan fitur sampel global (disebut juga *evidence*). Oleh karena itu, rumus di atas juga dapat secara sederhana dituliskan dalam bentuk berikut:

$$\text{Posterior} = \frac{\text{prior} \times \text{likelihod}}{\text{evidence}}$$

Nilai bukti selalu ditetapkan untuk setiap kategori dalam sampel. Nilai posterior nantinya akan dibandingkan dengan nilai kategori posterior lainnya untuk menentukan kategori sampel yang akan diklasifikasikan. Rumus Bayesian dijabarkan lebih lanjut dengan menggunakan uraian aturan perkalian berikut ( $C|F1, \dots, Fn$ ):

$$\begin{aligned} P(C|F1, \dots, Fn) &= P(C)P(F1, \dots, Fn|C) \\ &= P(C)P(F1|C)P(F2, \dots, Fn|C, F1) \\ &= P(C)P(F1|C)P(F2|C, F1)P(F3, \dots, Fn|C, F1, F2) \end{aligned}$$

$$\begin{aligned}
&= P(C)P(F_1|C)P(F_2|C,F_1)P(F_3|C,F_1,F_2)P(F_4,\dots,F_n|C,F_1,F_2,F_3) \\
&= P(C)P(F_1|C)P(F_2|C,F_1)P(F_3|C,F_1,F_2)\dots P(F_n|C,F_1,F_2,F_3,\dots,F_{n-1})
\end{aligned}$$

Dapat dilihat bahwa hasil yang dijelaskan telah menyebabkan kondisi-faktor yang semakin kompleks yang mempengaruhi nilai probabilitas, hampir tidak mungkin untuk menganalisis satu per satu. Akibatnya, kalkulasi tersebut menjadi sulit dilakukan. Diasumsikan di sini bahwa digunakan independensi yang sangat tinggi, yaitu, setiap indikator ( $F_1, F_2 \dots F_n$ ) tidak bergantung satu sama lain. Berdasarkan asumsi ini, persamaan berikut berlaku:

$$P(F_i|F_j) = \frac{P(F_i \cap F_j)}{P(F_j)} = \frac{P(F_i)P(F_j)}{P(F_j)} = P(F_i)$$

Untuk  $i \neq j$ , sehingga :

$$P(F_i|C, F_j) = P(F_i|C)$$

Persamaan di atas merupakan model dari teorema Naive Bayes yang kemudian digunakan dalam proses klasifikasi. Untuk klasifikasi dengan data kontinyu digunakan rumus Densitas Gauss :

$$P(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi\sigma^2_{ij}}} e^{-\frac{(x_i - \mu_{ij})^2}{2\sigma^2_{ij}}}$$

Dimana :

- P : Peluang
- $X_i$  : Atribut ke  $i$
- $x_i$  : Nilai atribut ke  $i$
- Y : Kelas yang dicari
- $y_i$  : Sub kelas Y yang dicari
- $\mu$  : Mean, menyatakan rata-rata dari seluruh atribut
- $\sigma$  : Deviasi standar, menyatakan varian dari seluruh atribut

### b. Studi Kasus

Contoh kasus ini diambil dari buku yang berjudul “Data Mining Teknik Pemanfaatan Data Untuk keperluan Bisnis” oleh Budi Santoso. metode naïve bayes ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Naive bayes sering bekerja jauh lebih baik dalam kebanyakan situasi dunia nyata yang kompleks dari pada yang diharapkan.

Contoh kasus ini diambil dari buku yang berjudul “Data Mining Teknik Pemanfaatan Data Untuk keperluan Bisnis” oleh Budi Santoso. Dalam contoh kasus ini yaitu keputusan bermain Sepak Bola atau tidak. Data training yang digunakan ada 14 data dan data yang telah di klasifikasikan berdasarkan cuaca, temperatur, kelembaban, dan angin. Setelah dilakukan dalam mengambil keputusan atau prediksi dalam pengklasifikasian akan menghasilkan output main atau tidak. Adapun data training dalam tabel dibawah ini :

Tabel 2.1 Data Training Contoh Kasus

<b>Cuaca (X1)</b>	<b>Temperatur (X2)</b>	<b>Kelembaban (X3)</b>	<b>Angin (X4)</b>	<b>Main atau Tidak (Y)</b>
Cerah	Panas	Tinggi	Kecil	Tidak
Cerah	Panas	Tinggi	Besar	Tidak
Mendung	Panas	Tinggi	Kecil	Ya
Hujan	Sedang	Tinggi	Kecil	Ya
Hujan	Dingin	Normal	Kecil	Ya
Hujan	Dingin	Normal	Besar	Tidak
Mendung	Dingin	Normal	Besar	Ya
Cerah	Sedang	Tinggi	Kecil	Tidak
Cerah	Dingin	Normal	Kecil	Ya
Hujan	Sedang	Normal	Kecil	Ya
Cerah	Sedang	Normal	Besar	Ya
Mendung	Sedang	Tinggi	Besar	Ya
Mendung	Panas	Normal	Kecil	Ya
Hujan	Sedang	Tinggi	Besar	Tidak

Tabel 2.2 Data Sampel Contoh Kasus

<b>Cuaca (X1)</b>	<b>Temperatur (X2)</b>	<b>Kelembaban (X3)</b>	<b>Angin (X4)</b>	<b>Main atau Tidak (Y)</b>
Cerah	Dingin	Tinggi	Besar	???

### c. Hasil dan Pembahasan

Menggunakan rumus :

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

$$\begin{aligned}
P(\text{main}) &= \frac{9}{14} = 0,64 \\
P(\text{tidak}) &= \frac{5}{14} = 0,36 \\
P(\text{Angin=besar} \mid \text{main}) &= \frac{3}{9} = 0,33 \\
P(\text{Angin=besar} \mid \text{tidak}) &= \frac{3}{5} = 0,60
\end{aligned}$$

Kemudian menghitung jumlah probabilitasnya dengan menggunakan rumus :

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

$$P(\text{main}) P(\text{cerah} \mid \text{main}) P(\text{dingin} \mid \text{main}) P(\text{tinggi} \mid \text{main}) P(\text{besar} \mid \text{main}) =$$

$$= \frac{9}{14} \times \frac{2}{9} \times \frac{3}{9} \times \frac{3}{9} \times \frac{3}{9} = 0,0053$$

$$P(\text{tidak}) P(\text{cerah} \mid \text{tidak}) P(\text{dingin} \mid \text{tidak}) P(\text{tinggi} \mid \text{tidak}) P(\text{besar} \mid \text{tidak}) =$$

$$= \frac{5}{14} \times \frac{3}{5} \times \frac{1}{5} \times \frac{4}{5} \times \frac{3}{5} = 0,0206$$

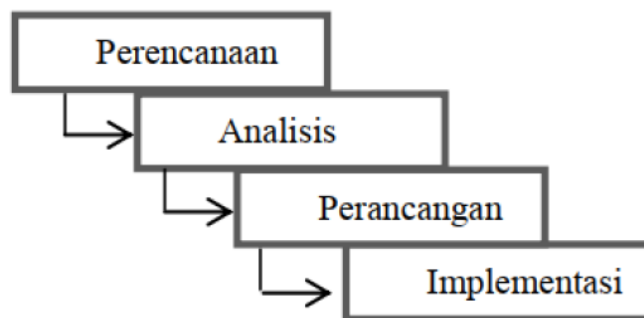
Sehingga dengan Naive Bayes dapat disimpulkan **“Tidak Main”** untuk data input ini, berdasarkan estimasi probabilitas yang dipelajari dari data training. Dengan normalisasi, agar jumlah probabilitas sama dengan 1, bisa menghitung *probabilitas conditional* untuk pilihan *tidak main* jika diberikan nilai-nilai atribut. Untuk contoh ini, probabilitasnya adalah :

$$\text{Probabilitas} = \frac{0,0206}{0,0206+0,0053} = 0,795$$

### 3. Pengembangan Sistem SDLC

Menurut M.Muslihudin 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. 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 2.1:



Gambar 2.1. Tahapan Dalam Metode SDLC

(Sumber : Raymond McLeod 2007: 188)

Pendekatan sistem merupakan sebuah metodologi. Metodologi adalah satu cara yang direkomendasikan dalam melakukan sesuatu. Pendekatan sistem adalah metodologi dasar dalam memecahkan segala jenis masalah. Siklus hidup pengembangan sistem (*System Development Life Cycle – SDLC*) adalah aplikasi dari pendekatan sistem bagi pengembangan suatu sistem informasi (Raymond McLeod 2007: 188).

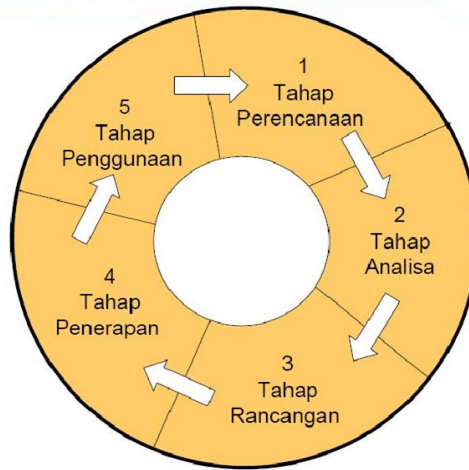
Terdapat beberapa tahapan pekerjaan pengembangan yang perlu dilakukan jika suatu proyek ingin memiliki kemungkinan berhasil yang besar. Tahapan-tahapan tersebut adalah:

- a. Perencanaan
- b. Analisis
- c. Desain
- d. Implementasi
- e. Penggunaan

Kemudian rencanakan sumber daya yang dibutuhkan untuk proyek, dan menyelesaikan proyek. Menganalisis sistem untuk dapat memahami masalah dan menentukan sistem baru sesuai kebutuhan fungsional. Kemudian dirancang sistem baru tersebut dan diterapkan. Setelah diterapkan, idealnya adalah menggunakan dengan jangka waktu yang lama. Karena proyek di atas memakai pola konvensional serta dilakukan dengan model top-down, biasanya

sebutan metode air terjun pada SDLC tradisional. Kegiatan ini akan mengalir satu arah untuk proyek diselesaikan.

Gambar 2.2 di bawah ini menggambarkan siklus hidup dengan sifat melingkar. Ketika sistem perlu diganti maka telah melampaui batas manfaatnya, dimulai dengan tahap perencanaan untuk membuat siklus hidup baru. Cukup mudah untuk SDLC tradisional dianggap sebagai aplikasi dengan pendekatan sistem. Selama tahap perencanaan dan analisis masalah dapat ditentukan. Identifikasi dan evaluasi salah satu alternatif selama fase desain. Kemudian, terapkan dan gunakan solusi terbaik. Dalam fase penggunaan, kumpulkan masukan untuk melihat kemampuan sistem dalam memecahkan masalah yang telah ditentukan sebelumnya (Raymond McLeod, 2007:188).



Gambar 2.2 SDLC Konsep

(Sumber : Raymond Mcleod,Schell George,2007:188)

#### 4. Pemrograman

##### a. Hypertext Preprocessor (PHP)

Menurut (Anhar 2010: 3) Singkatan dari PHP atau Hypertext Preprocessor, merupakan Bahasa pemrograman open source yang sangat cocok atau didedikasikan untuk pengembangan web dan dapat disematkan dalam makalah HTML. Dapat dikatakan bahwa bahasa PHP menjelaskan beberapa bahasa pemrograman, seperti C, Java dan Perl, serta mudah dipelajari. PHP adalah bahasa skrip sisi server di mana pemrosesan data dilakukan di sisi server. Singkatnya, server akan menerjemahkan skrip program dan kemudian mengirimkan hasilnya ke klien yang meminta. Definisi lain dari PHP adalah singkatan dari Hypertext Preprocessor. Hypertext Preprocessor adalah sebuah bahasa pemrograman berbasis kode (script)



yang digunakan untuk mengolah data dan mengirimkannya kembali ke web browser sebagai HTML Code.

b. Hypertext Markup Language (HTML)

Menurut (Simarmata 2010: 52), HTML adalah bahasa markup untuk menyebarkan informasi pada web. Ketika merancang HTML, ide ini diambil dari *Standart Generalized Markup Language* (SGML). SGML adalah cara yang terstandarisasi dari pengorganisasian dan informasi yang terstruktur di dalam dokumen atau sekumpulan dokumen. Walaupun HTML tidak dengan mudah dapat dipahami kebanyakan orang, ketika diterbitkan penggunaannya menjadi jelas.

HTML adalah singkatan dari *Hypertext Markup Language* yaitu bahasa pemrograman standar yang digunakan untuk membuat sebuah halaman web, yang kemudian dapat diakses untuk menampilkan berbagai informasi di dalam sebuah penjelajah web internet (*browser*).

## 5. Web Server

### Apache

Menurut Robert McCool (1995), Apache adalah sebuah nama web server yang bertanggung jawab pada *request-response* HTTP dan *logging* informasi secara detail (kegunaan basicnya). Selain itu, Apache juga diartikan sebagai suatu web server yang kompak, modular, mengikuti standar protokol HTTP, dan tentu saja sangat digemari.

## B. Tinjauan Studi

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

1. **Wijayanti, Sulastri (2018) yang berjudul “Analisa klasifikasi kartu kredit menggunakan algoritma naive bayes”** 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 (2018) yang berjudul “Klasifikasi Masyarakat Miskin Menggunakan Metode *Naïve Bayes*”** Mengingat nilai akurasi masih berada pada angka 73%, maka masih sangat dimungkin untuk dapat dilakukan penelitian selanjutnya untuk meningkatkan nilai akurasi dengan menambahkan fitur seleksi atau penggunaan Algoritma komputer yang lain. Pada penelitian lain diharapkan dapat digunakan dataset dalam jumlah yang lebih besar atau dengan sejumlah variabel lainnya guna meningkatkan performa dari metode yang digunakan.
3. **Saputra, Ayuningtias (2016) yang berjudul “Penerapan Algoritma *Naïve Bayes* Untuk Penentuan Calon Penerima Beasiswa Pada Smk Pasim Plus Sukabumi”** 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 secara lebih efektif dan efisien.
4. **Bustami (2014) yang berjudul “Penerapan Algoritma *Naive Bayes* Untuk Mengklasifikasi Data Nasabah Asuransi”** 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 *Naive 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 *Naive Bayes* di dukung oleh ilmu Probabilistik dan ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian. Pada algoritma *Naive 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. **Saleh (2015) yang berjudul “Implementasi Metode Klasifikasi *Naïve Bayes* Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga”** kriterian yang di gunakan adalah jumlah tanggungan keluarga, luas rumah, pendapatan/bulan, daya listrik perlengkapan yang dimiliki, dan penggunaan listrik Berdasarkan prediksi konsumsi listrik rumah tangga maka dapat diambil

kesimpulan sebagai berikut: Metode Naive Bayes menggunakan data latih untuk menghasilkan probabilitas masing-masing kriteria dalam kategori yang berbeda, sehingga nilai probabilitas kriteria tersebut dapat dioptimalkan untuk memprediksi kelistrikan sesuai ke metode naïve bayes Proses klasifikasi yang dilakukan oleh metode Sri Lanka itu sendiri digunakan. Berdasarkan data rumah tangga yang digunakan sebagai data pelatihan, metode Naive Bayes berhasil mengklasifikasikan 47 data dari 60 data uji. Dengan demikian metode Naive Bayesian berhasil memprediksi konsumsi listrik rumah tangga dengan akurasi sebesar 78.3333%.

6. **Harimurti (2017) yang berjudul “Klasifikasi Penerimaan Beasiswa Menggunakan Metode Naïve Bayes Classifier (Studi Kasus Universitas Trunojoyo Madura)”** Kriteria seleksi beasiswa Peningkatan Prestasi Akademik (PPA) adalah Indeks Prestasi Kumulatif (IPK) minimal 3,00 dan sertifikat atau prestasi non akademik. Proses klasifikasi dalam data rencana seleksi beasiswa yang berhasil ditetapkan meliputi dua kategori yaitu penerimaan dan penolakan. Jumlah sampel data yang digunakan terbagi menjadi dua jenis, data pelatihan untuk setiap program beasiswa adalah 2/3, dan data tes 1/3. Hasil akurasi program beasiswa PPA 83,05%, sedangkan akurasi program beasiswa BBM 75%.
7. **Rahman, Firdaus (2016) yang berjudul “Penerapan Data Mining Metode Naive Bayes Untuk Prediksi Hasil Belajar Siswa Sekolah Menengah Pertama (Smp)”** 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 dan algoritma yang diujikan tidak terlalu sensitif dengan input yang tidak terlalu relevan tersebut.
8. **Buani (2018) yang berjudul “Prediksi Penyakit Hepatitis Menggunakan Algoritma Naïve Bayes Dengan Seleksi Fitur Algoritma Genetika”** 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 akurasi dari algoritma naïve bayes setelah dilakukan seleksi fitur menggunakan algoritma genetika tingkat akurasinya lebih baik.

9. **Kadhamsyah (2013) yang berjudul “Sistem Pendukung Keputusan Penetapan Peserta Sertifikasi Guru Menggunakan Naïve Bayes”** 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.
10. **Fatimah, Yudihartanti (2016) yang berjudul “Penerapan Metode Naive Bayes Dalam Penentuan Bantuan Padat Karya”** Aplikasi untuk menentukan penerima bantuan Padat Karya Pada Dinas Tenaga Kerja dan Transmigrasi Kabupaten Banjar dengan metode *Naive Bayes* yang dibangun dapat membantu pihak Dinas Tenaga Kerja Dan Transmigrasi Kabupaten Banjar bagian Pembinaan Pelatihan dan Penempatan Tenaga Kerja (P3TK) yang menangani langsung bantuan padat karya.

Dari penelitian diatas yang menjadi acuan dalam penelitian ini adalah jurnal yang berjudul “Sistem Pendukung Keputusan Penetapan Peserta Sertifikasi Guru Menggunakan Naïve Bayes”. Maka penelitian ini diberikan judul “Penerapan Metode Naïve Bayes untuk penentuan Guru Berprestasi”.

Dapat disimpulkan bahwa metode dan hasil dari 10 jurnal tersebut dapat digunakan untuk prediksi, dan dapat dijadikan sebagai menentukan permasalahan yang di analisis.

Sedangkan dengan penelitian ini memiliki persamaan dan perbedaannya. Persamaan masalah sama. Sedangkan perbedaannya yaitu dari kriteria yang di gunakan dan metode naïve bayes digunakan untuk prediksi kelulusan, prediksi stok barang, penentuan kelayakan calon tenaga kerja indonesia, prediksi jumlah produksi, untuk mengetahui minat beli pelanggan, evaluasi kinerja akademik mahasiswa, untuk menganalisis tingkat mutu pelayanan, prediksi minat konsumen, dan klasifikasi peminatan siswa.

Tabel 2.3 Tinjauan Studi

No	Peneliti/ Tahun	Judul Penelitian	Permasalahan	Jurnal	Kontribusi
1	Rahman dan Firdaus (2016)	Penerapan Data Mining Metode Naïve Bayes Untuk Prediksi Hasil Belajar Siswa Sekolah	Sulitnya menentukan faktor atau variabel yang mempengaruhi hasil belajar siswa.	AI Ulum Sains dan Teknologi Vol.1 No.2 Mei 2016	Kontribusi: Penilaian beberapa atribut untuk ditambahkan ke variabel- variabel yang

No	Peneliti/ Tahun	Judul Penelitian	Permasalahan	Jurnal	Kontribusi
		Menengah Pertama(SMP )			di angkat Kelemahan: masih menggunakan aplikasi Rapidminer
2	Wijayanti dan Sulastri (2018)	Analisa Klasifikasi Kartu Kredit Menggunakan Algoritma Naive Bayes	kegiatan yang bermasalah atau kredit macet sering terjadi akibat analisis kredit kurang cermat dalam proses pemberian kredit.	Prosiding SINTAK 2018 ISBN 978- 602-8557- 20-7	Kontribusi: penilaian beberapa atribut untuk di tambahkan di variabel Kelemahan: hanya satu tempat tidak umum
3	Annur (2018)	Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes	Permasalahan utama dalam upaya pengurangan kemiskinan saat ini terkait dengan adanya fakta bahwa pertumbuhan ekonomi tidak tersebar secara merata.	ILKOM Jurnal Ilmiah Volume 10 Nomor 2 Agustus 2018	Kontribusi: Penilaian beberapa atribut untuk ditambahkan di variabel yang di ambil Kelemahan: tidak adanya perhitungan
4	Buani (2018)	Prediksi Penyakit Hepatitis Menggunakan Algoritma Naïve Bayes Dengan Seleksi Fitur	Salah satu yang dapat diaplikasikan untuk memprediksi penyakit hepatitis	Jurnal Evolusi Volume 6 No 2 - 2018   evolusi.we b.id	Kontribusi: penilaian dalam pengambilan atribut untuk ditambahkan di variabel Kelemahan:

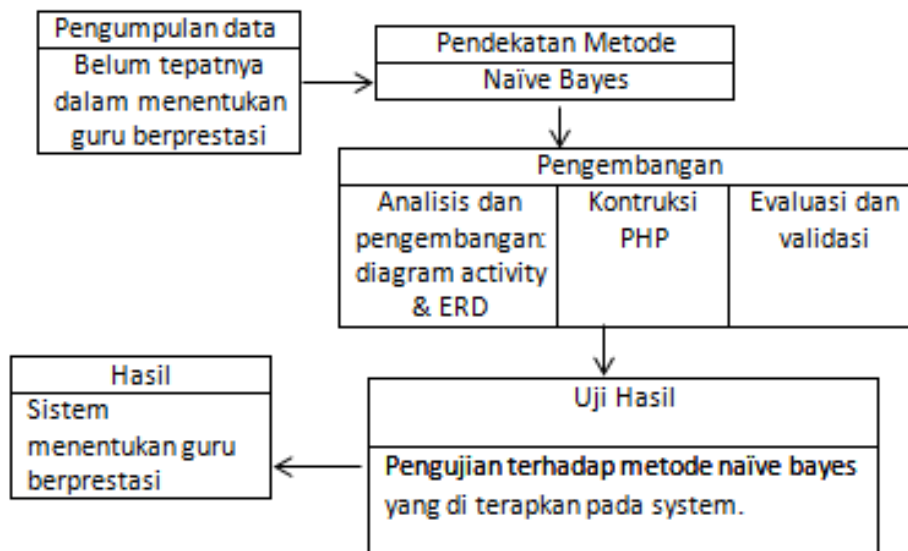
No	Peneliti/ Tahun	Judul Penelitian	Permasalahan	Jurnal	Kontribusi
		Algoritma Genetika			tidak adanya perhitungan
5	Saputra dan Ayuningti as (2016)	Penerapan Algoritma Naïve Bayes Untuk Penentuan Calon Penerima Beasiswa Pada Smk Pasim Plus Sukabumi	tidak adanya sistem yang dapat membantu untuk melakukan penyeleksian atau penentuan penerima beasiswa.	SWABUM I VOL IV No.2, Septembe r 2016	Kontribusi: Penilaian dalam pengambilan atribut untuk ditambahkan di variabel Kelemahan: Tidak ada perhitungan
6	Bustami (2014)	Penerapan Algoritma Naive Bayes Untuk Mengklasifika si Data Nasabah Asuransi	Sehingga diperlukan sebuah sistem yang dapat mengklasifikasika n nasabah mana yang masuk ke dalam kelompok lancar, kelompok kurang lancar dan nasabah mana yang masuk kedalam kelompok tidak lancar dalam membayar iuran premi.	JURNAL INFORMA TIKA Vol. 8, No. 1, Januari 2014	Kontribusi: pengambilan atribut untuk ditambahkan di variabel Kelemahan: tidak ada perhitungan
7	Kadhams yah (2013)	Sistem Pendukung Keputusan Penerapan Peserta Sertifikasi Guru Menggunakan	Masalah efektifitas, efisiensi dan standardisasi pengajar yang mempengaruhi kualitas guru.	Fakultas Sains Dan Teknologi Universita s Islam Negeri Maulana Malik	Kontribusi: pengambilan atribut untuk ditambahkan di variabel Kelemahan: tidak ada perhitungan

No	Peneliti/ Tahun	Judul Penelitian	Permasalahan	Jurnal	Kontribusi
		Naïve Bayes		Ibrahim Malang 18 Juni 2013	
8	Saleh (2015)	Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga	memprediksi besarnya penggunaan listrik tiap rumah tangga agar lebih mudah mengatur penggunaan listrik.	Citec Journal, Vol. 2, No. 3, Mei 2015 – Juli 2015	Kontribusi: pengambilan atribut untuk ditambahkan di variabel Kelemahan: tidak ada perhitungan
9	Harimurti (2017)	Klasifikasi Penerimaan Beasiswa Menggunakan Metode Naïve Bayes Classifier (Studi Kasus Universitas Trunojoyo Madura)	proses penyeleksian untuk penerima beasiswa di Perguruan Tinggi yang masih menggunakan aplikasi Excel.	Informatio n System Departme nt Informatio n Technolog y Faculty 10 November Institut Of Technolog y Surabaya 2017	Kontribusi: acuan pemberian penghargaan yang dipakai hasil dari proses naïve bayes Kelemahan: masih belum maksimal
10	Fatimah dan Yudiharta nti (2016)	Penerapan Metode Naive Bayes Dalam Penentuan Bantuan Padat Karya	Hasil penilaian pemberian bantuan padat karya tidak sesuai dengan pedoman pemberian bantuan padat karya yang	PROGRE SIF Vol 12, No. 2, Agustus 2016: 1387- 1524	Kontribusi: acuan dalam keputusan pemilihan guru berprestasi Kelemahan: tidak ada

No	Peneliti/ Tahun	Judul Penelitian	Permasalahan	Jurnal	Kontribusi
			berdasarkan pada lima kriteria dalam penentuan pemberian padat karya.		perhitungan

### C. Kerangka Pemikiran

Berikut adalah kerangka pemikiran untuk memecahkan masalah penelitian ini yang digambarkan pada gambar 2.3:



Gambar 2.3 Kerangka Pemikiran

Kerangka pemikiran pada gambar 2.3 dapat dijelaskan sebagai berikut :

1. Pengumpulan data di dapat melalui observasi untuk menangkap masalah awal dan untuk uji kelayakan sistem.
2. Komponen identifikasi masalah mencakup fenomena yaitu menentukan guru berprestasi masih sulit dan belum adanya sistem untuk menentukan guru berprestasi.
3. Metode yang digunakan yaitu Naïve Bayes.
4. Dalam pengembangan yaitu tahap analisis dan tahap pembuatan sistem informasi.



5. Hasilnya yang di dapatkan adalah sebuah sistem informasi untuk menentukan guru berprestasi berdasarkan kehadiran, penilaian kinerja dan lama kerja dengan menerapkan metode naïve bayes.

#### **D. Hipotesis Penelitian**

Hipotesis yang dapat ditetapkan dalam penelitian ini adalah penerapan metode Naive Bayes diduga dapat menentukan guru berprestasi.