

## **BAB II**

### **TINJAUAN PUSTAKA**

Bab ini berisi dan menjelaskan mengenai penelitian sistem prediksi penyakit memakai metode *Naïve Bayes*. Bab ini juga menjelaskan mengenai teori-teori dan bahan penelitian yang dipakai dalam penyusunan penelitian sistem prediksi penyakit paru-paru yang akan dilakukan.

#### **3.1 Studi Literatur**

Ada beberapa penelitian sejenis yang pernah dilakukan: Penelitian dengan judul “Aplikasi Web Untuk Pendeteksi Penyakit Paru – Paru Menggunakan Metode *Certainty Factor*” Tujuan dari penelitian ini adalah untuk mendiagnosa penyakit paru-paru dari gejala yang menyertai dengan memilih menu diagnose penyakit yang kemudian ditampilkan beberapa gejala untuk dipilih. Pengguna memilih gejala apa yang mereka rasakan selanjutnya diproses oleh sistem. Sistem memulai perhitungan berdasarkan gejala dipilih oleh pengguna dengan perhitungan *Certainty Factor*. Kemudian sistem akan menampilkan hasil perhitungan untuk setiap gejala dari setiap penyakit dan gejalanya yang memiliki skor *Certainty Factor* tertinggi merupakan solusi atau hasil untuk menentukan kemungkinan terkena penyakit paru-paru (Karimah, Nikmah, Aditya, & Wahyuni, 2019).

Penelitian yang berjudul “Diagnosis Penyakit Paru Pada Perokok Pasif Menggunakan Metode *Certainty Factor*” Tujuan dari penelitian ini adalah membuat sistem pakar yang bisa dipakai untuk mendiagnosa gejala perokok pasif. Sistem pakar hanya memprediksi (tidak menilai) melalui perhitungan matematis berdasarkan gejala yang dialami dengan memakai metode *Certainty Factor*. Sistem ini bisa digunakan untuk membantu masyarakat mengidentifikasi penyakit paru apa yang mungkin mereka derita berdasarkan gejala yang dialami pengguna. Selain itu, sistem pakar ini bisa mengedukasi dan menginformasikan kepada masyarakat terkait penyakit paru-paru yang dapat terjadi bila menjadi perokok pasif dan bagaimana penyakit tersebut dapat diobati sejak dini (Novianti, Jendra, & Wibawa, 2021).

Penelitian dengan judul “Identifikasi Penyakit Paru-Paru Menggunakan Metode *K Nearest Neighbor* (KNN)” Tujuan dari penelitian ini untuk prediksi serta diagnosis penyakit paru-paru dengan derajat kemiripannya. Semakin dekat nilai jarak bermakna semakin besar derajat kemiripannya, yang selanjutnya dijadikan data training dan testing menggunakan software rapidminer. Diawali seleksi data, pembersihan data, transformasi data dan evaluasi data. Dan mendapatkan hasil akurasi 96.96% sehingga *K-Nearest Neighbor* (KNN) bisa dipakai untuk memprediksi dalam mengidentifikasi penyakit Paru-Paru berdasarkan tingkat kemiripan. (Putri, Diana, & Purnamasari, 2020).

Penelitian dengan judul “Sistem Klasifikasi pada Penyakit *Breast Cancer* dengan Menggunakan Metode *Naïve Bayes*” Tujuan dari penelitian ini adalah memberikan pengetahuan tentang bahaya kanker payudara juga gejala awal penyakit ini sehingga terjadi peningkatan angka kematian pada penderita kanker payudara, sehingga peneliti membuat sistem informasi yang berisi tentang penyakit kanker payudara dengan mendiagnosa sejak dini dan memuat informasi cara penanganannya. Terdapat beberapa sistem klasifikasi yang dapat memberikan informasi seperti diagnosa, salah satu metode yang bisa dipakai yaitu *Naïve Bayes*, metode ini sangat baik dalam melakukan klasifikasi dengan acuan kejadian sebelumnya, pada penelitian ini mendapatkan hasil pengujian *Confusion Matrix* didapat akurasi terbaik sebesar 80% pada jumlah 116 dataset (Mubarog, Setyanto, & Sismoro, 2021).

Penelitian dengan judul “Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma *Naïve Bayes Classifier* Berbasis *Particle Swarm Optimization*” Tujuan dari penelitian ini melakukan prediksi pada penyakit gagal ginjal kronis dengan metode untuk prediksi, *Naïve Bayes Classifier* dipilih karena ini mempunyai beberapa keunggulan, yakni perhitungan cepat, algoritma sederhana dan presisi tinggi, membuatnya lebih cocok untuk aplikasi ke kumpulan data besar, mampu menangani data yang tidak lengkap (*missing value*), dan kuat terhadap atribut data yang tidak relevan dan kebisingan. Untuk meningkatkan akurasi maka dipakai *Particle Swarm Optimization* untuk pembobotan atribut. Dari hasil penelitian *Naïve Bayes Classification* berbasis *Particle Swarm Optimization* mempunyai akurasi

confusion matrix sebesar 98,75% dan AUC sebesar 99%. sedangkan *Naïve Bayes* mempunyai akurasi confusion matrix 97.00% dan AUC sebesar 99.8% (Arifin & Ariesta, 2019).

Pada penelitian Komparasi Algoritma *Naïve Bayes* dan SVM Untuk Memprediksi Keberhasilan Imunoterapi Pada Penyakit Kutil membandingkan akurasi antara dua metode klasifikasi, untuk prediksi keberhasilan metode pengobatan imunoterapi pada pengobatan penyakit kutil. Pengujian yang dilakukan memakai metode *Naïve Bayes* dan *Support Vector Machine (SVM)* memakai bahasa pemrograman R. Hasil dari penelitian ini dinyatakan bahwa metode *Naïve Bayes* mempunyai kemampuan prediksi lebih unggul dibandingkan dengan *Support Vector Machine (SVM)* dikarenakan *Naïve Bayes* bisa melakukan prediksi terhadap semua *instance class* secara tepat dengan tingkat akurasi yaitu 1, sedangkan *Support Vector Machine (SVM)* masih terdapat prediksi yang tidak tepat dengan tingkat akurasi 0.8. sehingga metode *Naïve Bayes* lebih unggul dibandingkan dengan *Support Vector Machine (SVM)* (Supriyatna & Mustika, 2018).

Dari ringkasan penelitian diatas terdapat beberapa metode yang digunakan untuk melakukan prediksi serta klasifikasi penyakit. Terdapat kesamaan dan perbedaan yang ditemukan diantaranya perbedaan metode yang digunakan dan kesamaan dalam prediksi suatu penyakit. Pada penelitian yang sedang dilakukan memakai metode *Naïve Bayes* dengan perbedaan melakukan prediksi penyakit paru-paru pada perokok aktif dan pasif dengan data yang digunakan diambil dari *Kaggle* dengan jumlah 113 data yang dibagi menjadi 100 data training dan data testing 13.

## **3.2 Tinjauan Teori**

Berikut yaitu teori-teori relevan yang akan menjadi acuan penulis untuk penelitian dan pembuatan sistem.

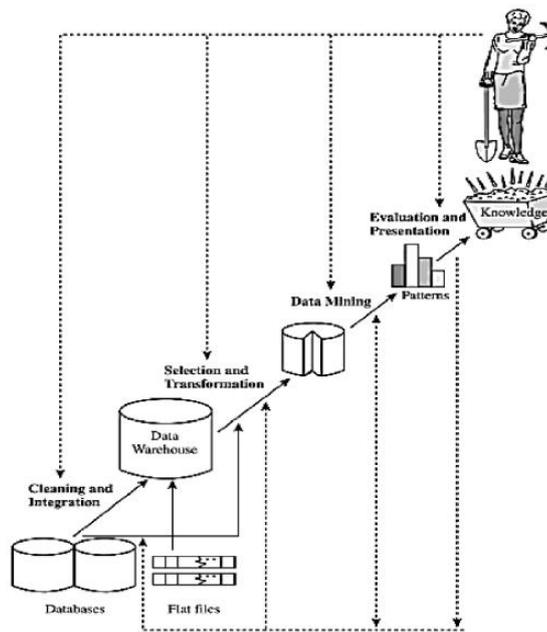
### **2.2.1 Data Mining**

Terdapat istilah Data mining dan *Knowledge Discovery Data (KDD)* untuk menjelaskan proses penggalian informasi tersembunyi dari kumpulan data besar.

Namun, kedua istilah ini mempunyai konsep yang berbeda namun terkait dan salah satu langkah dari proses KDD yaitu data mining. Data mining yaitu langkah dalam proses KDD secara keseluruhan dan sering digunakan oleh banyak peneliti sebagai sinonim untuk proses KDD (Marisa, Maukar, & Akhriza, 2021).

Data mining yaitu proses untuk mendapatkan pengetahuan dari data dalam jumlah besar dengan menggabungkan teknik dari berbagai bidang seperti pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi informasi. (Harani & Nugraha, 2020).

Data mining termasuk *knowledge discovery* dalam *database* (KDD). Yang mampu mengurai sekumpulan data menjadi informasi yang implisit dan belum diketahui sebelumnya. Data mining melakukan penggalian serta analisis, secara otomatis maupun semi otomatis dari kumpulan data yang besar untuk menemukan pola yang memiliki arti (Muflikhah, Ratnawati, & Putri, 2018).



**Gambar 2.1 Knowledge Discovery Data (KDD)**

Pada Gambar 2.1 merupakan langkah penting dalam proses KDD yang terdiri dari tahapan berikut :

- a. Pembersihan data (*data cleaning*)

Proses menghilangkan, memeriksa data ganda yang tidak konsisten.

- b. Integrasi data (*data integration*)  
Merupakan proses menambah serta menggabungkan data dari berbagai data base ke dalam satu *database*.
- c. Seleksi data (*data selection*)  
Merupakan proses pemilihan data yang relevan serta melakukan analisis dari data operasional.
- d. Transformasi data (*data transformation*)  
Merupakan proses merubah atau menggabung data dalam format tertentu untuk dilakukan proses data mining.
- e. Data mining  
Merupakan proses utama untuk mengekstraksi pola data yang penting dan tersembunyi pada *database* menggunakan metode atau algoritma tertentu.
- f. Evaluasi pola (*pattern evaluation*)  
Merupakan proses identifikasi pola yang benar dan menarik yang merepresentasikan informasi yang sesuai.
- g. *Knowledge presentation*  
Merupakan proses visualisasi yang menyajikan informasi dari data mining sehingga mudah dimengerti.

### 2.2.2 Prediksi

Dalam konteks data mining, prediksi merujuk pada proses menggunakan algoritma dan teknik statistik untuk memperkirakan atau meramalkan nilai atau peristiwa di masa yang akan datang berdasarkan pola dan tren yang terlihat pada data masa lalu menurut Harani & Nugraha, (2020) untuk melakukan prediksi, kita dapat menggunakan beberapa teknik dan metode yang sama dengan yang dipakai untuk klasifikasi dan estimasi (jika sesuai dengan kondisinya).

Prediksi mirip dengan klasifikasi dan estimasi, namun untuk prediksi terdapat kesalahan hasil ke depannya. Menggunakan banyak variabel untuk memprediksi nilai variabel lain yang tidak diketahui atau yang akan datang (Muflikhah, Ratnawati, & Putri, 2018).

Prediksi dalam data mining dapat dilakukan dalam berbagai bidang, seperti bisnis, keuangan, pemasaran, kesehatan, dan ilmu sosial. Contoh-contoh prediksi

dalam data mining meliputi perkiraan penjualan, prediksi harga saham, prediksi *churn* pelanggan dan memberikan manfaat besar bagi kesehatan, dengan menganalisis data tentang praktik terbaik, kita dapat meningkatkan kualitas perawatan dan mengurangi biaya. Data mining juga dapat diterapkan untuk mendiagnosis penyakit dan memperkirakan jumlah pasien dari tiap kategori (Harani & Nugraha, 2020).

### **2.2.3 Klasifikasi**

Dalam data mining, klasifikasi merupakan suatu metode yang diterapkan untuk mengelompokkan atau mengatur data ke dalam kategori atau kelas yang telah ditentukan sebelumnya berdasarkan pada atribut-atribut yang ada. Hal ini bertujuan untuk merancang model atau aturan yang dapat dipergunakan untuk memperkirakan kelas atau label yang sesuai untuk data baru yang belum diketahui (Harani & Nugraha, 2020).

Klasifikasi yaitu peran utama dalam data mining yang memakai metode pendekatan prediktif, adanya variabel target yang bertipe kategori. Model data mining mengujikan sejumlah record, dan setiap recordnya berisikan variabel target dan sekumpulan variabel input atau pemrediksi (Muflikhah, Ratnawati, & Putri, 2018).

Klasifikasi pada data mining dapat diimplementasikan pada berbagai bidang dan memiliki beragam pemanfaatan, seperti pengelompokan email spam, pengelompokan gambar, pengelompokan pelanggan, dan pengelompokan diagnosis medis. Seleksi model pengelompokan yang sesuai bergantung pada karakteristik data dan tujuan analisis yang dikehendaki.

### **2.2.4 Algoritma Naïve Bayes**

*Naïve Bayes Classifier* adalah metode klasifikasi berdasarkan teorema Bayes. Metode klasifikasi memakai metode probabilitas dan statistik yang diusulkan oleh peneliti Inggris Thomas Bayes, yakni memprediksi kemungkinan masa depan berdasarkan pengalaman masa lalu. Dan populer dengan nama Teorema Bayes dengan ciri asumsi yang sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian (Sari, 2021).

Ada beberapa keuntungan menggunakan metode ini yaitu anda tidak memerlukan data training yang banyak untuk menentukan estimasi parameter yang dibutuhkan dalam proses klasifikasi. Karena yang diasumsikan menjadi *variable independent*, maka hanya varians dari suatu *variable* pada sebuah kelas yang diperlukan untuk memilih klasifikasi, bukan holistik dari matriks kovarians (Sari, 2021).

Algoritma *Naïve Bayes* sangat cocok untuk mengklasifikasikan dataset bertipe kategorikal. Perhitungan menggunakan rumus 2.1

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)} \dots\dots\dots (2.1)$$

### **Rumus 2.1 *Naïve Bayes***

Keterangan :

- X : Data testing yang kelasnya belum diketahui
- H : Hipotesis data X yang adalah satu kelas spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (*posteriori probability*)
- P(H) : Probabilitas hipotesis H (*prior probability*)
- P(X|H) : Probabilitas X berdasarkan hipotesis H
- P(X) : Probabilitas X

Langkah-langkah algoritma *Naïve Bayes*:

1. Siapkan dataset
2. Hitung jumlah kelas pada data training.
3. Hitung jumlah kasus yang sama dengan kelas yang sama.
4. Kalikan semua hasil sesuai dengan data testing yang akan dicari kelasnya.
5. Bandingkan hasil per kelas, nilai tertinggi ditetapkan sebagai kelas baru.

### **2.2.5 Perhitungan Manual Metode *Naïve Bayes***

Dataset yang perhitungan manual adalah data contoh prediksi suatu penyakit selanjutnya data dibagi dua, yaitu data training pada Tabel 2.1 dan data testing pada Tabel 2.2. Dataset prediksi ini terdiri dari 7 (Sembilan) atribut dan 1 (satu) hasil atau kelas. Berikut langkah-langkah perhitungan manual algoritma *Naïve Bayes*:

1. Siapkan dataset

Seperti pada penjelasan diatas, dataset yang dipakai pada perhitungan manual ini bisa dilihat pada tabel berikut :

**Tabel 2.1 Data Training**

Jenis Kelamin	Batuk	Kelelahan	Sulit Bernafas	Demam	Tekanan Darah	Kolesterol	Outcome
Female	No	Yes	Yes	Yes	High	Normal	Positive
Female	Yes	Yes	No	No	Normal	Normal	Negative
Female	Yes	Yes	No	No	Normal	Normal	Negative
Male	Yes	No	Yes	Yes	High	High	Positive
Male	Yes	No	Yes	Yes	Normal	Normal	Positive
Female	No	No	No	Yes	Normal	High	Positive

Pada Tabel 2.1 merupakan dataset yang digunakan untuk acuan dalam contoh perhitungan manual metode *Naïve Bayes* pada sistem prediksi.

**Tabel 2.2 Data Testing**

Jenis Kelamin	Batuk	Kelelahan	Sulit Bernafas	Demam	Tekanan Darah	Kolesterol	Outcome
Female	Yes	Yes	No	No	Normal	Normal	?

Pada Tabel 2.2 merupakan dataset yang akan digunakan untuk perhitungan yang belum diketahui hasil *outcome*.

2. Hitung jumlah probabilitas *outcome* pada data training

Kelas pada data training terdiri dari dua kategori, yakni *Positive* dan *Negative* sehingga probabilitas untuk *Positive* terkena penyakit dan *Negative* tidak terkena suatu penyakit sebagai berikut:

Jumlah hasil *Positive* = 4

Jumlah hasil *Negative* = 2

$$P(C = \text{"Positive"}) = \frac{4}{6} = 0,7$$

$$P(C = \text{"Negative"}) = \frac{2}{6} = 0,3$$

3. Hitung jumlah kasus yang sama dengan hasil *outcome* yang sama serta menentukan nilai probabilitas dari setiap variable terhadap *outcome*.

a. Probabilitas dari jenis kelamin pada setiap kategori *outcome*.

**Tabel 2.3 Probabilitas Jenis Kelamin**

Jenis Kelamin	<i>Outcome</i>		Probabilitas	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Female</i>	2	2	0.5	1.0
<i>Male</i>	2	0	0.5	0
Jumlah	4	2	1.00	1.00

b. Probabilitas dari batuk pada setiap kategori *outcome*.

**Tabel 2.4 Probabilitas Batuk**

Batuk	<i>Outcome</i>		Probabilitas	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	2	2	0.5	1.0
<i>No</i>	2	0	0.5	0
Jumlah	4	2	1.00	1.00

c. Probabilitas dari kelelahan pada setiap kategori *outcome*.

**Tabel 2.5 Probabilitas Kelelahan**

Kelelahan	<i>Outcome</i>		Probabilitas	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	1	2	0.25	1.0
<i>No</i>	3	0	0.75	0
Jumlah	4	2	1.00	1.00

d. Probabilitas dari sulit bernafas pada setiap kategori *outcome*.

**Tabel 2.6 Probabilitas Sulit Bernafas**

Sulit Bernafas	<i>Outcome</i>		Probabilitas	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Yes</i>	3	0	0.75	0
<i>No</i>	1	2	0.25	1.0
Jumlah	4	2	1.00	1.00

- e. Probabilitas dari demam pada setiap kategori *outcome*.

**Tabel 2.7 Probabilitas Demam**

Demam	Outcome		Probabilitas	
	Positive	Negative	Positive	Negative
Yes	1	0	0.25	0
No	3	2	0.75	1.0
Jumlah	4	2	1.00	1.00

- f. Probabilitas dari tekanan darah pada setiap kategori *outcome*.

**Tabel 2.8 Probabilitas Tekanan Darah**

Tekanan Darah	Outcome		Probabilitas	
	Positive	Negative	Positive	Negative
Normal	2	2	0.5	1.0
High	2	0	0.5	0
Jumlah	4	2	1.00	1.00

- g. Probabilitas dari kolesterol pada setiap kategori *outcome*

**Tabel 2.9 Probabilitas Kolesterol**

Kolesterol	Outcome		Probabilitas	
	Positive	Negative	Positive	Negative
Normal	2	2	0.5	1.0
High	2	0	0.5	0
Jumlah	4	2	1.00	1.00

4. Kalikan semua hasil sesuai dengan data testing yang akan dicari hasilnya

Contoh perhitungan user menginputkan data prediksi dengan rincian yang terdapat pada Tabel 2.2 data testing diatas.

$$\begin{aligned}
 P(X|C = \text{"Positive"}) &= 0,5 * 0,5 * 0,25 * 0,25 * 0,75 * 0,5 * 0,5 \\
 &= 0,00293
 \end{aligned}$$

$$\begin{aligned}
 P(X|C = \text{"Negative"}) &= 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 \\
 &= 1,0
 \end{aligned}$$

$$P(C = \text{"Positive"}|X) = 0,00293 * 0,7 = 0,00195$$

$$P(C = \text{"Negative"}|X) = 1,0 * 0,3 = 0,3$$

## 5. Bandingkan kedua hasil

Dari perhitungan probabilitas “*Positive*” dan “*Negative*” pada langkah sebelumnya, bisa disimpulkan bahwa hasilnya *Negative* karena probabilitas “*Negative*” (0,3) lebih tinggi dari pada probabilitas “*Positive*” (0.00195), sehingga seseorang tersebut dinyatakan tidak terkena suatu penyakit.

### 2.2.6 Website

Website adalah rangkaian halaman web yang berisi informasi yang saling berhubungan dan dapat diakses melalui internet. Website yaitu kumpulan halaman situs web yang berada di bawah domain atau subdomain tertentu, dan dapat diakses melalui World Wide Web (WWW) di Internet. Halaman web adalah dokumen yang menggunakan format HTML (Hyper Text Markup Language), yang biasanya bisa diambil dengan HTTP, yakni protokol yang mengirim informasi dari server website untuk ditampilkan (Marisa, 2017).

Website terdiri dari dua elemen utama, yakni *client-side* dan *server-side*. *Client-side* website adalah bagian situs yang terlihat oleh pengguna melalui browser, seperti *Google Chrome* atau *Internet Explorer*. Bagian ini tersusun atas HTML, CSS, dan *JavaScript* untuk merancang dan menampilkan halaman web. Sebaliknya, *server-side* website adalah bagian website yang tersembunyi oleh pengguna, dan berisi file dan data yang diolah oleh web server. Bagian *server-side* ini terdiri dari bahasa pemrograman, seperti PHP, *Python*, atau *Ruby on Rails*.

### 2.2.7 Pengertian PHP

PHP yaitu bahasa skrip yang tertanam di sisi server yang berarti semua sintaks dan perintah yang kita masukkan akan dieksekusi oleh server namun dimasukkan ke dalam halaman HTML biasa. Hasil dari aplikasi-aplikasi yang dibuat dengan PHP biasanya akan ditampilkan di web browser, namun prosesnya secara keseluruhan berlangsung di server (Marisa, 2017).

Menurut Rusli et al. (2019) PHP yaitu sebuah bahasa scripting yang terintegrasi tag- tag HTML, bersifat open source dan gratis yang berjalan di sisi server, Bahasa ini tidak dapat dipisahkan dari data base MySQL sehingga untuk membuat halaman web yang dinamis seperti *guestbook*, statistik pengunjung,

polling, email, dan lain-lain membutuhkan web server. Bahasa ini memiliki beberapa keunggulan diantaranya:

- a. *Life Cycle* yang sangat singkat , sehingga PHP selalu up to date mengikuti perkembangan teknologi internet.
- b. *Cross Platform*, sehingga dapat digunakan di hampir semua web server yang ada di pasaran (terutama *Apache* dan *Microsoft IIS*) juga dapat dijalankan pada berbagai sistem operasi (*Linux, Windows, FreeBSD*).
- a. Serta mendukung koneksi ke banyak *database* baik yang gratis maupun komersil, seperti *MySQL, Oracle, Microsoft SQL Server, Interbase*, dan banyak lagi.

### 2.2.8 Web Server

Sebuah perangkat lunak dengan layanan dan fungsi berbasis data untuk menerima permintaan *HTTP* atau *HTTPS* untuk klien yang dikenal yang biasanya disebut dengan browser (*Mozilla Firefox, Google Chrome*) untuk mengirim kembali hasil dalam beberapa bentuk halaman web yang umumnya berbentuk *HTML* (Agustinus Budi Santoso, 2022).

Server yaitu sistem komputer yang menyediakan semacam layanan kepada klien di jaringan komputer. Server dilengkapi dengan sistem operasi khusus untuk mengontrol hak akses dan sumber daya di dalamnya. Sistem operasi khusus tersebut disebut sistem operasi jaringan atau *Network Operating System*. Selain itu server didukung dengan RAM yang besar dan prosessor yang bersifat *scalable* (Setiawan, 2017).

Fungsi utama web server digunakan untuk mengerjakan dan mengirimkan file permintaan pengguna melalui protokol komunikasi yang diputuskan, penggunaan web server berfungsi untuk menyampaikan semua aspek berkas meliputi halaman web termasuk berisi teks, video, gambar, dan lainnya (Agustinus Budi Santoso, 2022).

### 2.2.9 MySQL

Dengan kemajuan zaman, teknologi pun semakin maju termasuk perangkat lunak. MySQL adalah salah satu perangkat lunak yang terus diperbarui oleh

produsen-produsennya. MySQL yaitu DBMS yang bersifat *open source* dengan dua jenis lisensi, yaitu lisensi gratis (*Free Software*) dan lisensi berbayar (*Shareware*) yang memiliki batasan penggunaan. MySQL adalah server *database* yang dapat Anda gunakan secara gratis dengan lisensi GNU *General Public License* (GPL) baik untuk keperluan pribadi maupun komersial tanpa harus membayar lisensi apapun (Fitri, 2020).

MySQL yaitu sebuah DBMS (*Database Management System*) memakai perintah SQL (*Structured Query Language*) yang biasa dipakai saat ini pada pembuatan aplikasi berbasis website. MySQL dibagi menjadi 2 lisensi, pertama merupakan *software* gratis dimana aplikasi bisa diakses siapa saja. Dan ke 2 merupakan *shareware* dimana aplikasi berpelelik mempunyai batasan pada penggunaannya. MySQL termasuk ke dalam RDBMS (*Relational Database Management System*). Sehingga, memakai tabel, kolom, baris, pada pada struktur *database*-nya. Jadi, pada proses pengambilan data memakai metode relational *database* serta sebagai penghubung antara aplikasi & *database* server (Remawati, 2021).

MySQL yaitu suatu *software* yang bersifat *open source* sehingga dimungkinkan untuk memodifikasinya dirancang untuk menyimpan, mengatur, dan mengelola data dalam sebuah basis data.