

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Penelitian Terdahulu**

Bagian ini bertujuan untuk memberikan landasan teoritis dan konteks bagi penelitian yang sedang dilakukan, serta membantu mengidentifikasi celah atau kekurangan dalam penelitian sebelumnya yang dapat diisi oleh penelitian baru.

Tabel 2. 1 Tabel Jurnal Sebelumnya

No	Judul	Masalah	Hasil	Kelebihan dan Kekurangan
1	Identifikasi Tingkat Stress Pada Manusia Menggunakan Metode K-NN ( <i>K-Nearest Neighbour</i> )(Wirayudha dkk., 2020)	Stres yang berkepanjangan berdampak negatif pada jasmani dan rohani, sehingga diperlukan media untuk mengetahui tingkat stress guna mencegah dampak yang tidak diinginkan.	Pengujian menggunakan nilai K serta menggunaan data latih yang beragam menghasilkan tingkat akurasi maksimum sebesar 82%.	<b>Kelebihan :</b> Pengujian menggunakan nilai K yang bervariasi dan sejumlah dataset pelatihan yang berbeda-beda. <b>Kekurangan :</b> Akurasi menurun dengan meningkatnya nilai K.
2	Penerapan Metode <i>K-Nearest Neighbor</i> untuk Sistem Pendukung Keputusan Identifikasi Penyakit <i>Diabetes</i>	<i>Diabetes Melitus</i> yang tidak terkendali dapat mengakibatkan permasalahan serius hingga kematian,	Implementasi metode K-NN untuk SPK identifikasi <i>diabetes melitus</i> dengan akurasi	<b>Kelebihan :</b> Akurasi tinggi (96%). <b>Kekurangan :</b> Sistem hanya berfokus pada

No	Judul	Masalah	Hasil	Kelebihan dan Kekurangan
3	<i>Melitus(Lestari dkk., 2021)</i>	sehingga diperlukan sistem pendukung keputusan untuk mengidentifikasi penyakit ini.	sebesar 96% (k=23).	perhitungan klasifikasi saja.
	Perbandingan Algoritma <i>K-Nearest Neighbor</i> dan <i>Naïve Bayes Classifier</i> untuk Klasifikasi Status Gizi Pada Balita(Loka & Marsal, 2023)	Banyak balita di Indonesia mengalami masalah gizi ganda, seperti obesitas dan stunting sekaligus, sehingga diperlukan klasifikasi untuk memprediksi status gizi pada balita.	Algoritma KNN memberikan akurasi sebesar 96,10%, sementara NBC mencapai 90,94%.	<b>Kelebihan :</b> KNN memberikan akurasi tinggi (96,10%). <b>Kekurangan :</b> Tidak menjelaskan tentang sistem yang digunakan dan hanya menampilkan perhitungan klasifikasi saja.
	<i>Supervised Learning Metode K-Nearest Neighbor Untuk Prediksi Diabetes Pada Wanita(Silalahi dkk., 2023)</i>	<i>Diabetes Mellitus</i> diperkirakan akan mempengaruhi 300 juta orang pada 2025 menurut WHO, sehingga diperlukan penelitian untuk memprediksi penyakit ini.	Metode KNN berhasil diterapkan untuk memprediksi diabetes pada wanita. Pengujian dengan dataset 768 data menunjukkan tingkat akurasi 82% hingga 84%.	<b>Kelebihan :</b> Dataset yang digunakan mencapai 768 data. <b>Kekurangan :</b> Hasil prediksi dapat terpengaruh oleh jumlah tetangga yang dipilih (nilai K), serta keterbatasan

No	Judul	Masalah	Hasil	Kelebihan dan Kekurangan
				dataset yang dapat mempengaruhi generalisasi model pada populasi yang lebih luas.
5	Prediksi Tingkat Stres Pada Mahasiswa UNUGHA Cilacap Menggunakan Algoritma K-Nearest Neighbor( <i>Ulul Azmi Wafiqi dkk., 2024</i> )	Menurut Wirayudha, mahasiswa kerap mengalami emosi negatif yang dapat berkembang menjadi stress dan gangguan kesehatan, sehingga dibutuhkan pendekatan prediktif berbasis kecerdasan buatan.	KNN mampu memprediksi tingkat stres mahasiswa dengan akurasi 83,33%.	<b>Kelebihan :</b> Perhitungan Algoritma KNN cukup detail. <b>Kekurangan :</b> Jumlah data uji yang terbatas (6 data) dan data pelatihan (84 data) membuat hasil penelitian kurang representatif untuk populasi yang lebih besar

### 2.1.1 *Research Gap*

1. Hasil penelitian sebelumnya yang dilakukan oleh Wirayudha (2020), Loka dan Marsal (2023) serta Ulul Azmi Wafiqi (2024) cenderung hanya pada perhitungan klasifikasi saja tanpa adanya aplikasi.
2. Penelitian yang dilakukan oleh Lestari (2021) dan Silalahi (2023) aplikasi hanya berfungsi untuk melakukan klasifikasi saja tanpa ada fitur lain.

3. Pada penelitian ini ada dua pengguna yaitu user dan admin. Admin dapat mengatur *data training* dan user dapat menginput data tidur dan mengubah profile.

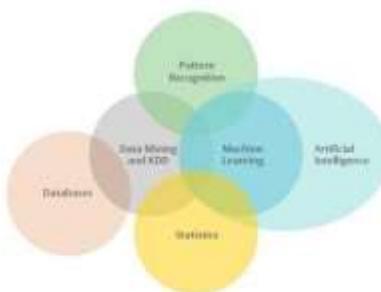
## 2.2 Teori Terkait

### 2.2.1 *Artificial Intelligence*

Menurut Rifky dan rekan-rekannya (2024) *Artificial Intelligence* atau kecerdasan buatan merupakan disiplin ilmu dari ranah komputer yang berfokus pada penciptaan sistem cerdas untuk menangani tugas-tugas yang biasa dilakukan oleh manusia. Dalam referensi lainnya, disebutkan bahwa kecerdasan buatan merupakan disiplin ilmu yang berfokus pada pengembangan komputer yang mampu meniru bahkan melampaui kemampuan manusia dalam menyelesaikan berbagai permasalahan (Al-Khowarizmi & Lubis, 2023). Terdapat beberapa metode dan pendekatan utama yang umum digunakan dalam *Artificial Intelligence*, salah satu metode tersebut adalah *Machine Learning*.

### 2.2.2 *Machine Learning*

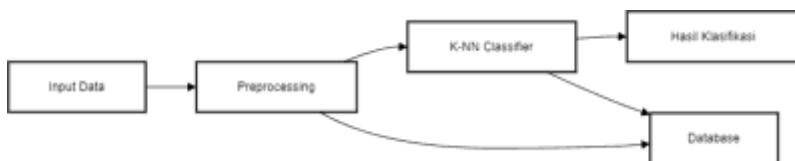
Menurut (Kurniawan, 2022) *Machine Learning* merupakan ranah komputer yang mempelajari ilmu tentang algoritma dan model komputer dimana komputer tersebut dapat belajar mengenali pola-pola di dalam data yang ada. Dalam referensi lainnya, mengemukakan *Machine Learning* didefinisikan sebagai kemampuan sistem komputer untuk mempelajari pola dan membuat keputusan secara mandiri tanpa perlu diprogram secara langsung (Id, 2021).



Gambar 2. 1 Cabang-Cabang Ilmu Data

Selain *Machine Learning*, terdapat disiplin ilmu yang mirip, yaitu *Data Mining*. Meskipun mirip, *Machine Learning* tidak sama dengan *Data Mining*. Perbedaan utamanya adalah *Data Mining* berfokus pada penggalian informasi dari kumpulan data berskala besar, sementara *Machine Learning* diarahkan untuk mengembangkan kemampuan komputer dalam mengenali pola dan belajar dari data secara otomatis (Id, 2021).

### 2.2.3 Klasifikasi



Gambar 2. 2 Arsitektur Sistem Klasifikasi

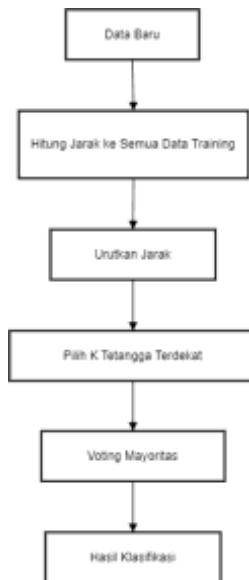
Menurut Ibnu Daqiqil (2021), klasifikasi merupakan proses pengelompokan data yang melibatkan data berlabel, di mana setiap data memiliki target atau kategori tertentu yang menjadi acuan dalam proses pengklasifikasian. Gambar 2.2 menunjukkan alur data dan komponen utama dalam sistem klasifikasi. Proses dimulai dengan melakukan *input data* yang kemudian dilanjutkan dengan tahap *preprocessing*, yakni tahap persiapan data agar data yang diproses bisa sesuai dengan format yang dibutuhkan

oleh algoritma. Setelah melalui tahap *preprocessing*, data akan diproses oleh algoritma K-NN untuk pengklasifikasian. *Output* dari proses klasifikasi tersebut yang merupakan hasil klasifikasi akan disimpan dalam *database* agar dapat digunakan untuk keperluan lebih lanjut atau analisis data.

#### **2.2.4    *K-Nearest Neighbor***

Dalam ranah *Machine Learning*, algoritma *K-Nearest Neighbor* (KNN) dikenal sebagai metode *supervised learning* yang digunakan untuk menyelesaikan tugas-tugas regresi maupun klasifikasi. Algoritma ini melakukan prediksi dengan cara mengukur tingkat kemiripan antara data uji dan sejumlah data pelatihan yang ada. Tingkat kemiripan tersebut dihitung menggunakan fungsi jarak tertentu, seperti *Euclidean Distance*, untuk menentukan k tetangga terdekat yang akan menjadi dasar pengambilan keputusan (Id, 2021).

Berikut merupakan tahapan algoritma K-NN yang bisa digunakan dalam pelaksanaannya:



Gambar 2. 3 Proses K-Nearest Neighbor

Proses kerja algoritma K-NN diawali dengan memasukkan data baru, kemudian dilakukan perhitungan jarak antara data tersebut dengan seluruh data pada himpunan data pelatihan (*data training*). Setelah semua jarak dihitung, data pelatihan diurutkan berdasarkan jarak terdekat hingga terjauh. Selanjutnya, K-NN akan memilih sejumlah  $k$  data dengan jarak terdekat sebagai tetangga terdekat. Proses klasifikasi dilakukan melalui mekanisme voting mayoritas, di mana kelas yang paling sering muncul di antara  $k$  tetangga tersebut akan ditetapkan sebagai kelas prediksi untuk data baru. Hasil akhir dari proses ini kemudian ditampilkan sebagai hasil klasifikasi.

#### 2.2.5 Penyakit Gangguan Tidur

Gangguan tidur merupakan kondisi abnormal yang dapat mengakibatkan permasalahan dalam siklus tidur seseorang (Okvitasisari,

2016). Dalam referensi lainnya, mengemukakan gangguan tidur dapat diartikan sebagai kondisi yang mencerminkan ketidakteraturan dalam durasi, kualitas, maupun pola waktu tidur individu (Sari dkk., 2020). Berikut beberapa jenis gangguan tidur :

a. *Insomnia*

Tanda seorang menderita *insomnia* yakni ketidakmampuan untuk memulai tidur, tidak mampu untuk tetap tertidur dalam jangka waktu yang cukup, ataupun bangun tidur tanpa merasa segar dan pulih. Gejala yang umum terjadi pada penderita *insomnia* antara lain kesulitan tidur pada malam hari, terbangun terlalu pagi, kelelahan sepanjang hari, kesulitan dalam berkonsentrasi, serta meningkatnya tingkat iritabilitas. Faktor-faktor penyebab *insomnia* mencakup stres, gangguan kecemasan, depresi, pola tidur yang tidak teratur, dan kondisi kesehatan tertentu yang mempengaruhi kualitas tidur.

b. *Sleep Apnea*

*Sleep apnea* merupakan gangguan pernapasan saat tidur yang dapat menyebabkan kualitas tidur menurun. Gejala *sleep apnea* antara lain mendengkur dengan keras, terjadinya jeda napas selama tidur, serta perasaan lelah atau kurang bugar saat bangun di pagi hari. Faktor-faktor yang berkontribusi terhadap kondisi ini meliputi kelebihan berat badan, kelainan anatomi saluran napas atas, serta adanya riwayat *sleep apnea* dalam keluarga.

## 2.2.6 Variabel Klasifikasi Gangguan Tidur

### 1. Jenis Kelamin

Penelitian menunjukkan bahwa jenis kelamin dapat memengaruhi pola tidur dan prevalensi gangguan tidur. *Insomnia* sering dilaporkan lebih umum pada wanita, sementara *Sleep Apnea* menunjukkan prevalensi yang lebih tinggi pada pria. Prevalensi *insomnia* pada wanita mencapai 61,7%, secara

signifikan lebih tinggi dibandingkan 52,7% pada pria, dan perbedaan ini semakin diperkuat setelah disesuaikan dengan karakteristik pekerjaan dan tanggung jawab keluarga (Tsou, 2022). Sedangkan pria tiga kali lebih mungkin didiagnosis dengan Sleep Apnea. Rasio perbandingan penderita Sleep Apnea pria dengan wanita berkisar antara 3:1 hingga 5:1 di populasi umum, dan bahkan 8:1 hingga 10:1 di populasi klinis tertentu (Bonsignore dkk., 2019).

## 2. Usia

Banyak gangguan tidur seperti Insomnia dan Sleep Apnea, umumnya meningkat seiring bertambahnya usia. Lansia merupakan kelompok usia yang sangat rentan terhadap insomnia. Prevalensi gejala insomnia pada lansia berkisar antara 30% hingga 48% (Patel dkk., 2018). Sedangkan Sleep Apnea seringkali terjadi pada usia 50 - 60 tahun (Annisarahma & Karima, 2024).

## 3. Durasi Tidur

Tidur yang tidak cukup atau kurang dari 6 jam maupun durasi tidur yang berlebihan melebihi 10 jam dikaitkan dengan hasil kesehatan yang negatif, termasuk kelelahan dan tidur yang tidak menyegarkan. Penderita insomnia seringkali memiliki durasi tidur yang pendek dikarenakan kesulitan memulai tidur, kesulitan mempertahankan tidur (sering terbangun di malam hari atau bangun terlalu dini), atau tidur yang tidak restoratif (tidak merasa segar setelah tidur), meskipun individu memiliki kesempatan yang memadai untuk tidur.

## 4. Kualitas Tidur

Kualitas tidur dapat didefinisikan sebagai tingkat kemudahan seseorang dalam memulai dan mempertahankan tidur. Ini mencerminkan kepuasan individu terhadap tidurnya,

yang bermanifestasi dalam perasaan segar, tidak lelah, dan bebas dari kegelisahan saat bangun di pagi hari. Penderita insomnia seringkali memiliki kualitas tidur yang buruk dikarenakan kesulitan memulai atau mempertahankan tidur. Sedangkan Sleep Apnea adalah gangguan tidur yang ditandai oleh episode berulang henti napas atau penurunan pernapasan yang signifikan selama tidur. Jeda pernapasan ini dapat berlangsung sekitar 10 detik dan terjadi ratusan kali dalam semalam. Episode ini menyebabkan tubuh kekurangan oksigen dan seringkali memicu terbangun singkat dari tidur, yang mengganggu kualitas tidur.

#### 5. Tingkat Stres

Stres, kecemasan, dan depresi adalah pemicu utama atau memperparah insomnia. Tingkat stress yang tinggi merupakan faktor signifikan yang menyebabkan kualitas tidur yang buruk.

#### 6. Kategori BMI

Obesitas merupakan faktor risiko utama untuk Sleep Apnea. Peningkatan indeks massa tubuh (BMI) berkorelasi kuat dengan peningkatan prevalensi dan keparahan Sleep Apnea. Sebuah penelitian menunjukkan bahwa prevalensi Sleep Apnea mencapai 24-61% pada anak-anak dan remaja dengan obesitas.

#### 7. Denyut Jantung

Rata-rata denyut jantung adalah variabel fisiologis numerik yang termasuk dalam kumpulan data tidur komprehensif. Variasi denyut jantung selama tidur dapat mengindikasikan gangguan tidur yang mendasari, terutama gangguan pernapasan terkait tidur, atau mencerminkan masalah kesehatan kardiovaskular yang seringkali terjadi bersamaan dengan gangguan tidur.

#### 8. Langkah Harian

Jumlah langkah yang diambil setiap hari adalah variabel numerik yang mencerminkan tingkat aktivitas fisik seseorang. Aktivitas fisik yang teratur dapat membantu seseorang tertidur lebih cepat dan nyenyak. Sebaliknya, kurangnya aktivitas fisik yang baik dikaitkan dengan kualitas tidur yang buruk.

#### 9. Tekanan Darah Sistolik dan Diastolik

Tekanan darah adalah metrik kesehatan penting yang termasuk dalam kumpulan data kesehatan tidur. Tidur yang buruk sangat terkait dengan peningkatan risiko hipertensi dan penyakit kardiovaskular, menjadikan tekanan darah sebagai indikator fisiologis yang relevan.

### **2.2.7 Kaggle**

Kaggle merupakan sebuah platform berbasis komunitas *data science* yang menyediakan berbagai dataset dari beragam bidang, seperti kesehatan, ekonomi, sosial, hingga lingkungan. Platform ini memungkinkan pengguna untuk mempelajari, mengembangkan, serta mengimplementasikan algoritma *Machine Learning* secara praktis. Dalam konteks penelitian akademik, Kaggle berperan sebagai sumber penyedia dataset sekaligus media untuk menguji dan membandingkan kinerja berbagai pendekatan algoritma yang dirancang oleh peneliti.

### **2.2.8 Python**

Bahasa pemrograman Python diperkenalkan pada tahun 1991 oleh Guido van Rossum sebagai bahasa tingkat tinggi yang dirancang untuk mendukung kemudahan penulisan kode dan fleksibilitas dalam pengembangan aplikasi. Bahasa pemrograman Python dapat digunakan untuk berbagai jenis proyek, mulai dari pengembangan web hingga analisis data dan penerapan algoritma kecerdasan buatan. Salah satu keunggulan utama Python terletak pada ketersediaan pustaka dan modul, seperti

NumPy untuk komputasi numerik, Pandas untuk pengolahan data, serta TensorFlow yang umum digunakan dalam pengembangan model pembelajaran mesin.

#### **2.2.9 *Unified Modeling Language (UML)***

UML merupakan bahasa yang diterapkan secara luas dalam pengembangan perangkat lunak untuk menggambarkan kebutuhan sistem, mendukung proses analisis dan perancangan, serta mempresentasikan struktur arsitektur dalam konteks pemrograman berbasis objek. UML hadir sebagai solusi atas kebutuhan pemodelan visual yang berfungsi untuk menspesifikasi, memvisualisasikan, mengkontruksi, dan mendokumentasikan berbagai komponen dalam pembangunan sistem perangkat lunak (Fatimah & Nuryaningsih, 2018).

#### **2.2.10 *Flowchart***

*Flowchart* merupakan salah satu alat bantu visual untuk mempresentasikan alur kerja atau tahapan proses dalam suatu sistem atau metode secara sistematis. Diagram ini terdiri dari sejumlah simbol satndar, seperti persegi panjang, belah ketupat, lingkaran, dan panah, yang masing-masing melambangkan aktivitas, keputusan, titik awal atau akhir serta aliran proses. Penggunaan *flowchart* dapat mempermudah dalam memahami logika prosedural sistem, karena menyajikan urutan langkah dan interaksi antar komponen secara terstruktur dan mudah dipahami.

No	Simbol	Nama	Keterangan
1		<i>Terminator</i>	Permulaan atau akhiran program.
2		<i>Input / Output data</i>	<i>input</i> atau <i>output</i> data
3		<i>Decision</i>	Pemilihan proses berdasarkan kondisi yang ada.
4		<i>Process</i>	Pengolahan data yang dilakukan oleh komputer.
5		<i>On Page Connector</i>	Simbol masuk atau keluar proses dalam lembar yang sama
6		<i>Off Page Connector</i>	Simbol masuk atau keluar proses dalam lembar yang berbeda.
7		<i>Predefined Process</i>	Menandai awal sub program dalam alur proses.
8		<i>Manual Operation</i>	Pengolahan data yang dilakukan secara manual.
9		<i>Document</i>	Proses yang melibatkan dokumen kertas, baik sebagai input atau output

Gambar 2. 4 Penjelasan Flowchart

### **2.2.11 *Use Case Diagram***

Dalam UML, diagram *use case* dipakai pengembang untuk mengidentifikasi serta memvisualisasikan kebutuhan sistem sehingga membantu memastikan bahwa seluruh fungsionalitas utama telah tercakup dan terdokumentasi secara sistematis. Elemen-elemen utama dalam *use case diagram* meliputi aktor, *use case*, serta relasi antar keduanya, yang direpresentasikan melalui garis atau panah sebagai bentuk hubungan.

No	Simbol	Nama	Keterangan
1		<i>Actor</i>	Simbol untuk pengguna yang berinteraksi langsung dengan sistem
2		<i>Use Case</i>	Menjelaskan bagian utama dari fungsionalitas sistem
3		<i>System</i>	Menguraikan paket untuk membagi sistem ke dalam bagian-bagian terbatas.
4		<i>Collaboration</i>	Sinergi dari aturan dan elemen yang berinteraksi menghasilkan perilaku melebihi jumlah komponennya.
5		<i>Note</i>	Elemen yang berperan saat sistem berlangsung dan mempresentasikan sumber daya komputasi
6		<i>Dependency</i>	Hubungan elemen independent mempengaruhi elemen yang bergantung padanya.
7		<i>Generalization</i>	Hubungan <i>descendent</i> mewarisi perilaku dari <i>ancestor</i>
8	<<include>>	<i>Include</i>	<i>use case</i> berasal dari sumber tertentu.
9	<<extend>>	<i>Extend</i>	<i>use case</i> yang dituju melakukan perluasan tindakan
10		<i>Association</i>	Mengaitkan satu objek dengan objek yang lain.

Gambar 2. 5 Penjelasan Diagram *Use Case*

### 2.2.12 Activity Diagram

Dalam UML, *activity diagram* digunakan untuk menyajikan urutan kegiatan secara sistematis, mencakup pengambilan keputusan, aliran kontrol, serta kondisi-kondisi yang mempengaruhi jalannya proses. Elemen-elemen utama yang terdapat dalam *activity diagram* meliputi aktivitas, keputusan (*decision node*), alur kontrol (*control flow*), serta objek yang terlibat dalam proses.

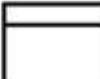
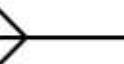
No	Simbol	Nama	Keterangan
1		<i>Initial Node</i>	Mempresentasikan titik awal dari alur tindakan dalam <i>activity diagram</i> .
2		<i>Activity Final Node</i>	Simbol yang menunjukkan akhir dari suatu proses atau alur kerja.
3		<i>Activity</i>	Aktivitas yang dijalankan oleh sistem
4		<i>Decision</i>	Simbol yang menandai percabangan proses dengan melakukan pemilihan salah satu aktivitas.
5		<i>Fork Node</i>	Digunakan untuk membagi aliran menjadi beberapa aliran bersamaan.

Gambar 2. 6 Penjelasan *Activity Diagram*

### 2.2.13 Conceptual Data Model (CDM)

CDM merupakan representasi abstrak yang menggambarkan struktur informasi yang dibutuhkan oleh suatu sistem tanpa mengacu pada

aspek teknis implementasi. Fokus utama dari model ini adalah identifikasi entitas, atribut, serta hubungan antar entitas yang relevan untuk mendukung kebutuhan fungsionalitas sistem.

No	Simbol	Nama	Keterangan
1		<i>Entity</i>	Menyatakan objek atau konsep yang penting dan memiliki data yang disimpan dalam sistem.
2		<i>One-to-One Relationship</i>	Menggambarkan relasi <i>one-to-one</i> antara dua entitas dalam model data.
3		<i>One-to-Many Relationship</i>	Menunjukkan hubungan di mana satu entitas terkait langsung dengan banyak entitas.
4		<i>Many-to-One Relationship</i>	Menunjukkan hubungan di mana banyak entitas terkait langsung dengan satu entitas.
5		<i>Many-to-Many Relationship</i>	Menunjukkan hubungan di mana banyak entitas terkait dengan banyak entitas.

Gambar 2. 7 Penjelasan *Conceptual Data Model*

#### 2.2.14 Physical Data Model (PDM)

PDM biasa dipakai untuk menggambarkan bagaimana data disimpan, diatur, dan diakses dalam sistem basis data. Model ini mencakup struktur tabel, kolom, tipe data, *index*, *primary key*, dan *foreign key*, serta berbagai batasan yang dirancang untuk menjamin integritas, keamanan, dan efisiensi pengelolaan data. Sebagai pedoman implementasi, *physical data model* memastikan bahwa sistem basis data dapat beroperasi secara optimal, andal, dan sesuai dengan kebutuhan fungsional sistem.

No	Simbol	Nama	Keterangan
1		<i>Table</i>	Menyatakan objek <i>database</i> yang menyimpan data dalam bentuk baris dan kolom.
2		<i>View</i>	Menyatakan objek <i>database</i> yang menyajikan hasil <i>query</i> sebagai tabel virtual.
3		<i>Relationship (Solid Line)</i>	Menyatakan hubungan antara tabel melalui <i>foreign key</i> , digambarkan dengan garis <i>solid</i> .

Gambar 2. 8 Penjelasan *Physical Data Model*