

BAB II

TINJAUAN PUSTAKA

2.1 Data Mining

Data Mining adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data (database) yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui. Data mining merupakan suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting. (Reza Noviansyah et al. 2018)

Data mining biasa juga disebut dengan “*Data atau Knowledge Discovery*” atau menemukan pola tersembunyi pada data. Data mining adalah proses dari menganalisa data dari prespektif yang berbeda dan menyimpulkannya ke dalam informasi yang berguna.

Menurut David Hand, Heikki Mannila, dan Padhraic Smyth dari MIT Data mining adalah Analisa terhadap data (biasanya data yang berukuran besar) untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut (Darmawan et al. 2018).

Data mining adalah proses yang menggunakan Teknik statistic, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Data mining merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual.

Pada prosesnya data mining akan mengekstrak informasi yang berharga dengan cara menganalisis adanya pola-pola ataupun hubungan keterkaitan tertentu dari data-data yang berukuran besar. Data mining berkaitan dengan bidang ilmu-ilmu lain seperti *Database Sistem, Data Warehousing, Statistic, Mechine Learning, Information Retrieval*, dan komputasi Tingkat Tinggi. Selain itu data mining

didukung oleh ilmu lain seperti *Neural Network*, Pengenalan Pola, Spatial Data Analysis, Image Database, Signal Processing.

2.1.1 Proses Data Mining

Proses data mining atau dapat disebut *Knowledge Discovery in Database* (KDD) secara garis besar bisa dijelaskan sebagai berikut:(Mardi 2017)

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery in Database* (KDD) dimulai. Data hasil seleksi akan digunakan untuk proses data mining, disimpan dalam suatu berkas terpisah dari basis data operasional.

2. *Pre-processing / Cleaning*

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *Knowledge Discovery in Database* (KDD). Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak.

3. *Transformation*

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai dengan proses data mining. Proses *coding* dalam *Knowledge Discovery in Database* (KDD) merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan Teknik atau metode tertentu.

5. *Interpretation / Evaluation*

Tahapan ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

2.1.2 Karakteristik Data Mining

Data mining memiliki karakteristik sebagai berikut:

1. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
2. Data mining menggunakan data yang berukuran besar yang bertujuan agar hasil atau keakuratan data dapat lebih dipercaya.
3. Data mining digunakan sebagai Teknik dalam mengambil keputusan di masa depan berdasarkan informasi/ data yang diperoleh melalui data masa lalu.

2.1.3 Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu : (Reza Noviansyah et al. 2018)

1. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menentukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun menggunakan *record* lengkap dengan menyediakan nilai dari variabel target sebagai nilai prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam bentuk lain akan bernilai minimal.

6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.2 Prediksi

Prediksi adalah memperkecil kesalahan, sehingga perbedaan yang diperkirakan antara kejadian yang sebenarnya diminimalkan. Prediksi/forecasting menentukan kebutuhan bulan yang akan datang terkait dengan dukungan jumlah data (historical data) atau perkiraan waktu/periode yang dianalisis sehingga dapat diperhitungkan untuk memprediksi jumlah permintaan pada bulan mendatang. (Muhammad et al. 2021)

Forecasting/ prediksi merupakan suatu usaha untuk meramalkan keadaan di masa mendatang melalui pengujian keadaan di masa lalu. Dalam kehidupan sosial

segala sesuatu itu serba tidak pasti, sukar untuk diperkirakan secara tepat. Dalam hal ini perlu diadakan forecast. Forecasting yang dibuat selalu diupayakan agar dapat meminimumkan pengaruh ketidakpastian ini terhadap perusahaan. Dengan kata lain forecasting bertujuan mendapatkan forecast yang bisa meminimumkan kesalahan meramal (forecast error) yang biasanya diukur dengan mean squared error, mean absolute error, dan sebagainya. (Neighbor et al. 2019)

2.3 Penjualan

Penjualan merupakan kegiatan ekonomi normal dimana perusahaan menerima hasil/manfaat yang direncanakan dari penjualan atau pengembalian biaya yang dikeluarkan. Penjualan produk perusahaan ditampilkn setelah dikurangi diskon dan return. Penjualan adalah pembelian sesuatu (barang atau jasa) dari satu pihak ke pihak lain dengan imbalan uang dari suatu pihak. Berdasarkan pengertian penjualan tersebut dapat disimpulkan bahwa penjualan adalah suatu kegiatan ekonomi yang dilakukan antara pembeli dengan penjual melakukan transaksi berupa pertukaran barang atau jasa dengan uang tunai.(Dewi et al. 2022)

2.4 K-Nearest Neighbor

Metode K-Nearest Neighbor (KNN) menjadi salah satu metode NN yang paling tua dan populer. Nilai K yang digunakan disini menyatakan jumlah tetangga terdekat yang dilibatkan dalam penentuan prediksi label kelas pada data uji. Dari K tetangga terdekat yang terpilih kemudian dilakukan voting kelas K tetangga terdekat tersebut. Kelas dengan jumlah tetangga terbanyaklah yang diberikan sebagai label kelas prediksi pada data uji tersebut. (Muhammad et al. 2021)

K-Nearest Neighbor (K-NN) merupakan salah satu algoritma yang digunakan dalam masalah pengklasifikasian. Prinsip kerja K-NN ialah mencari jarak terdekat antar data yang akan dievaluasidengan tetangga terdekat dalam data pelatihan. Di KNN, K adalah jumlah tetangga terdekat. Jumlah tetangga adalah factor penentu inti. K umumnya angka ganjil jika jumlah kelasnya adalah 2.(T. Rismawan, A. W. Irawan, W. Prabowo 2017)

Algoritma *K-Nearest Neighbor* merupakan metode klasifikasi dengan berdasarkan jarak data bar uke beberapa data atau tetangga terdekat. Pendekatan

Nearest Neighbor adalah pendekatan untuk mencari kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. Tujuan dari algoritma *K-Nearest Neighbor* jarak antara stu data dengan data yang lain dapat dihitung. Nilai hasil jarak perhitungan dapat digunakan sebagai nilai kedekatan atau kemiripan antara data uji dengan data latih.(Sukamto et al. 2020)

2.4.1 Tahapan Langkah Algoritma *K-Nearest Neighbor*

1. Menentukan parameter k (jumlah tetangga paling dekat).
2. Menghitung kuadrat jarak *eucliden* objek terhadap data *training* yang diberikan.
3. Mengurutkan hasil poin b secara *ascending* (berurutan dari nilai tinggi ke rendah).
4. Mengumpulkan kategori Y (Klasifikasi *Nearest Neighbor* berdasarkan nilai k).
5. Dengan menggunakan kategori *nearest neighbor* yang paling mayoritas maka dapat diprediksikan kategori objek.

K pada algoritma *k-nearest neighbor* adalah banyaknya tetangga terdekat yang akan digunakan sebagai titik untuk melakukan klasifikasi pada data atau objek baru. Dalam menentukan jumlah nilai k, sebaiknya menggunakan angka ganjil. Untuk menghitung jarak antar objek data pada algoritma *k-nearest neighbor* dapat dilakukan dengan beberapa cara, salah satunya dengan *Ecludiean Distance*. Berikut adalah rumus dari *Euclidean Distance*:

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2} \dots\dots\dots(2.1)$$

Keterangan:

- X_{1i} : Data Uji/Testing
- X_{2i} : Data Training
- i : Variabel Data
- d : Jarak
- n : jumlah fitur

2.4.2 Contoh Perhitungan Algoritma *K-Nearest Neighbor*

Berikut adalah contoh perhitungan sederhana penerapan algoritma *K-Nearest Neighbor* yang dapat kita lihat dibawah ini.

Tabel 2. 1 Data Training

Age	Income	Class
22	350	A
43	290	A
26	355	B
30	340	A
50	410	B
35	370	B
41	520	A
28	155	A

Tabel 2. 2 Data Testing

Age	Income	Class
37	250	?

1. Menentukan nilai k (tetangga terdekat). Misalkan $K=3$.
2. Menghitung jarak antara data uji dengan data training dengan *Euclidean Distance*.

Data 1

$$Dis = \sqrt{(37 - 22)^2 + (250 - 350)^2} = 101.11$$

Data 2

$$Dis = \sqrt{(37 - 43)^2 + (250 - 290)^2} = 40.44$$

Data 3

$$Dis = \sqrt{(37 - 26)^2 + (250 - 355)^2} = 105.05$$

Data 4

$$Dis = \sqrt{(37 - 30)^2 + (250 - 340)^2} = 90.27$$

Data 5

$$Dis = \sqrt{(37 - 50)^2 + (250 - 410)^2} = 160.52$$

Data 6

$$Dis = \sqrt{(37 - 35)^2 + (250 - 370)^2} = 120.01$$

Data 7

$$Dis = \sqrt{(37 - 41)^2 + (250 - 520)^2} = 270.02$$

Data 8

$$Dis = \sqrt{(37 - 28)^2 + (250 - 155)^2} = 95.26$$

Tabel 2. 3 Hasil Perhitungan Jarak

Age	Income	Euclidean Distance	Ranking	Class
22	350	101.11	5	A
43	290	40.44	8	A
26	355	105.05	4	B
30	340	90.27	7	A
50	410	160.52	2	B
35	370	120.01	3	B
41	520	270.02	1	A
28	155	95.26	6	A

3. Setelah menghitung jarak satu persatu data uji ke data training, selanjutnya mencari jarak yang terkecil dengan mengurutkan hasil perhitungan jarak secara *ascending* (berurutan dari jarak terkecil ke jarak terbesar).

Tabel 2. 4 Jarak Urut Hasil Perhitungan Jarak

Age	Income	Euclidean Distance	Ranking	Class
41	520	270.02	1	A
50	410	160.52	2	B
35	370	120.01	3	B
26	355	105.05	4	B

22	350	101.11	5	A
28	155	95.26	6	A
30	340	90.27	7	A
43	290	40.44	8	A

4. Selanjutnya mengumpulkan jarak terkecil yang telah diurutkan berdasarkan inputan nilai k (tetangga terdekat). Karena nilai k = 3, maka yang diambil 3 terkecil.

Tabel 2. 5 Tetangga Terdekat k=3

Age	Income	Euclidean Distance	Ranking	Class
41	520	270.02	1	A
50	410	160.52	2	B
35	370	120.01	3	B

5. Berdasarkan tabel diatas label yang banyak muncul adalah B, maka hasil prediksi adalah **B**.

2.5 Gain Ratio

Gain Ratio adalah bagian dari Decision tree yang mana merupakan tahapan untuk menentukan nilai atribut yang lebih tepat, dengan begitu nilai atribut yang dihasilkan lebih akurat, dengan adanya *Gain Ratio* dapat mengurangi bias atribut yang memiliki banyak sekali cabang. *Gain Ratio* memiliki sifat yaitu sebagai berikut:

1. Bernilai lebih besar ketika data berubah.
2. Akan kecil jika data diletakkan di satu tempat.

Sebelum menghitung nilai *Gain Ratio*, terlebih dahulu menghitung nilai *entropy* terdahulu, kemudian menghitung nilai *split info* sehingga mendapatkan nilai *Gain Ratio*.

$$Entropy (S) = \sum_{i=1}^n -P_i * \log_2 P_i \dots\dots\dots(2.2)$$

Dimana:

S : (Himpunan kasus)

n : (Jumlah Partisi s)

Pi : (Proporsi dari Si terhadap S)

Selanjutnya melakukan perhitungan Information Gain pada masing-masing atribut dengan persamaan:

$$Information\ Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{S_i}{S} * Entropy(S_i) \dots\dots\dots(2.3)$$

Dengan begitu *Gain Ratio* memiliki rumus sebagai berikut:

$$Gain\ Ratio(S, A) = \frac{Gain(S,A)}{SplitInformation(S,A)} \dots\dots\dots(2.4)$$

Akan tetapi, sebelum melakukan penentuan nilai atribut dari *Gain Ratio*, harus melakukan perhitungan split information, split information adalah bagian dari formula yang digunakan untuk memilih sejumlah atribut, Adapun rumus nya adalah sebagai berikut: (Nur et al. 2022)

$$SplitInfo(S, A) = \sum_i^c = 1 \frac{S_i}{S} \log_2 \frac{S_i}{S} \dots\dots\dots(2.5)$$

S = Jumlah data sampel

Si = Jumlah masing-masing dari setiap atribut

Selanjutnya menghitung *Gain Ratio* berdasarkan formula rumus yang ada diatas.

Gain Ratio akan digunakan sebagai alat untuk memeriksa kolerasi antara atribut dalam dataset, dimana *Gain Ratio* akan digunakan sebagai pembobotan untuk atribut dengan rasio gain tinggi. Dengan memasukkan bobot kedalam atribut, diharapkan ini akan mengurangi berat atribut yang tidak relevan dengan hasil klasifikasi *K-Nearest Neighbor*, memungkinkan peningkatan kecepatan klasifikasi.

2.6 MySQL (My Structure Query Language)

MySQL (My Structure Query Language) adalah salah satu DataBase Management System (DBMS dari sekian banyak DBMS seperti Oracle, MS SQL, Postagre SQL, dan lainnya. MySQL berfungsi untuk mengolah database menggunakan Bahasa SQL. MySQL bersifat open source sehingga kita bisa menggunakannya secara gratis. Pemrograman PHP juga sangat mendukung/support dengan database MySQL. (Annisa 2019)

MySQL juga merupakan sebuah perangkat lunak yang mengatur basis data relasio (RDBMS) yang di distribusikan gratis dengan lesensi GPL (General Public License). Pengguna dapat menggunakan MySQL secara bebas namun dengan Batasan perangkat lunak tersebut tidak boleh dijadikan produk turunan bersifat komersial. MySQL adalah server yang melayani database. Untuk membuat dan mengolah database, pada MySQL ini dapat mempelajari pemrograman khusus yang disebut query (perintah) SQL. Database diperlukan jika akan menginput data dari user menggunakan form HTML untuk diolah PHP agar dapat disimpan kedalam database MySQL.



Gambar 2. 1 Logo MySQL

(Sumber: www.mysql.com)

2.6.1 Kelebihan MySQL

MySQL memiliki beberapa kelebihan sebagai *server database* dengan konsep *database* modern, yaitu:

1. Open Source

MySQL merupakan sebuah bisnis data yang berada dibawah lesensi GPL sehingga dapat digunakan secara gratis atau tanpa mengenakan biaya.

2. Portability

Database MySQL memiliki fungsi yaitu menstabilkan berbagai sistem operasi seperti linux, windows, Mac OS X Server, FreeBSD, Solaris, HP-Linux, Amigo dan lain nya yang bertujuan agar tidak menjadi kendala.

3. Multiuser

MySQL adalah *database* yang dapat digunakan untuk menangani beberapa *user* dalam waktu bersamaan tanpa harus mengalami masalah maupun konflik.

4. Column Type

MySQL didukung oleh tipe data yang cukup kompleks, seperti *signed/unsigned integer, float, double, char, varchar, text, time, data, year, set, enum, dan date time.*

5. Security

Sistem *security* pada MySQL memiliki beberapa lapisan sekuritas seperti nama *host, subnetmask* dan izin akses *user* dengan sistem perizinan secara mendetail dan *password* yang terenkripsi.

6. Performance Tuning

MySQL memiliki kecepatan yang cukup bagus dalam menangani *query-query* sederhana dan mampu memproses lebih banyak MySQL pada persatuan waktu.

7. Command dan Functions

MySQL *server* mempunyai operator serta fungsi secara penuh yang mendukung perintah *SELECT* dan *WHERE* dalam *query.*

8. Connectivity

Kemampuan MySQL dalam melakukan koneksi dengan *client* menggunakan *protocol* TCP/IP, Unix soket (Unix) maupun Named Pipes (NT).

2.7 Pengertian PHP

PHP merupakan singkatan dari “PHP: Hypertext Preprocessor”, yang merupakan sebuah Bahasa scripting (kode untuk membangun suatu program) yang

dikombinasikan pada HyperText Markup Language (HTML). Tujuan utama penggunaan Bahasa ini adalah untuk memungkinkan perancang web menulis halaman web dinamik dengan cepat. (Putratama 2018)

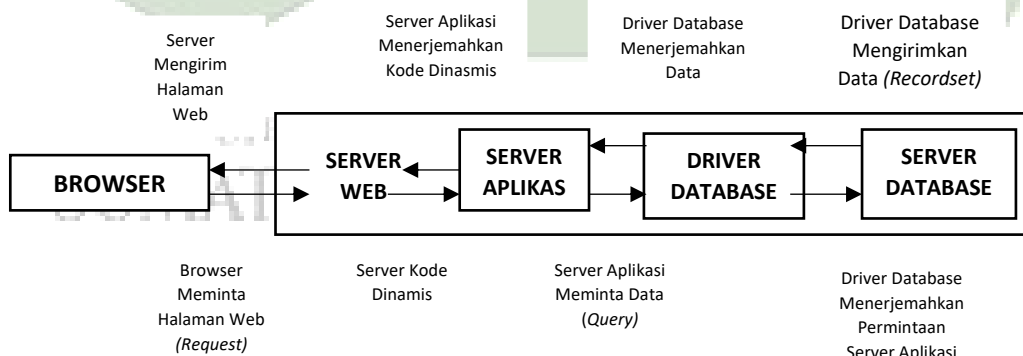
PHP adalah salah satu bahasa pemrograman yang berjalan dalam sebuah webserver dan berfungsi sebagai pengolah data pada sebuah server. Dengan menggunakan PHP, sebuah website akan lebih interaktif dan dinamis. (Annisa 2019)



Gambar 2. 2 Logo PHP
(Sumber: www.php.net)

2.7.1 Prinsip Kerja PHP

Bahasa pemrograman PHP merupakan bahasa pemrograman yang dikategorikan kepada *ServerSide Programming*, yang artinya Bahasa pemrograman ini memerlukan penerjemah dalam hal ini *web server* untuk menjalankannya. Berikut ini diberikan gambaran tentang cara kerja bahasa pemrograman PHP:



Gambar 2. 4 Prinsip Kerja PHP
(Sumber: Putratama, 2018)

2.7.2 Kelebihan PHP

Bahasa pemrograman PHP merupakan bahasa pemrograman yang paling banyak digunakan, tentu karena berbagai alasan, salah satunya adalah mempunyai beberapa kelebihan dibandingkan dengan bahasa pemrograman lainnya yang sejenis. Berikut ini kelebihan bahasa pemrograman PHP (Putratama 2018):

1. PHP adalah bahasa multiplatform yang artinya dapat berjalan di berbagai mesin dan sistem operasi (Linux, Unix, Macintosh, Windows) dan dapat dijalankan secara rutime melalui console serta juga dapat menjalankan perintah-perintah sistem lainnya.
2. PHP bersifat *OpenSource* yang berarti dapat digunakan oleh siapa saja secara gratis.
3. Web Server yang mendukung PHP dapat ditemukan dimana-mana dari mulai apache, IIS, Lighttpd, Nginx, hingga Xitami dengan konfigurasi yang relative mudah dan tidak berbelit-belit, bahkan banyak yang membuat dalam bentuk paket atau package (PHP, MySQL dan Web Server).
4. Dalam sisi pengembangan lebih mudah, karena banyaknya milis-milis, komunitas dan *developer* yang siap membantu dalam pengembangan.
5. Dalam sisi pemahaman, PHP adalah bahasa *scripting* yang paling mudah karena memiliki referensi yang banyak.
6. Banyak bertebaran aplikasi dan Program PHP yang gratis dan siap pakai seperti WordPress, PrestaShop, dan lain-lain.
7. Dapat mendukung banyak *database*, seperti MySQL, Oracle, MS-SQL dst.

2.8 Website

Website merupakan kumpulan halaman-halaman yang digunakan untuk menampilkan informasi teks, gambar diam atau gerak, animasi, suara, dan atau gabungan dari semuanya, baik yang bersifat statis maupun dinamis yang membentuk satu rangkaian bangunan yang saling terkait, yang masing-masing masing dihubungkan dengan jaringan-jaringan halaman. (Riko Rivanthio 2020)

2.9 Sublime Text

SublimeText adalah aplikasi editor untuk kode dan teks yang dapat berjalan diberbagai plat form operating system dengan menggunakan teknologi Phyton API. Terciptanya aplikasi ini terinspirasi dari aplikasi Vim, Aplikasi ini sangatlah fleksibel dan powerfull. Fungsionalitas dari aplikasi ini dapat dikembangkan dengan menggunakan sublime-packages. SublimeText bukanlah aplikasi open source dan juga aplikasi yang dapat digunakan dan didapatkan secara gratis, akan tetapi beberapa fitur pengembangan fungsionalitas (packages) dari aplikasi ini merupakan hasil dari temuan dan mendapat dukungan penuh dari komunitas serta memiliki linsensi aplikasi gratis.(Subowo and Saputra 2019)

2.10 Flowchart

Flowchart atau sering disebut dengan diagram alir merupakan suatu jenis diagram yang merepresentasikan algoritma atau langkah-langkah instruksi yang berurutan dalam sistem. seorang analis sistem menggunakan flowchart sebagai bukti dokumentasi untuk menjelaskan gambaran logis sebuah sistem yang akan dibangun kepada programmer. Dengan begitu, flowchart dapat membantu untuk memberikan solusi terhadap masalah yang bisa saja terjadi dalam membangun sistem. Pada dasarnya, flowchart digambarkan dengan menggunakan simbol-simbol. Setiap simbol mewakili suatu proses tertentu. Sedangkan untuk menghubungkan satu proses ke proses selanjutnya digambarkan dengan menggunakan garis penghubung. (Rosaly and Prasetyo 2019)

Dengan adanya flowchart, setiap urutan proses dapat digambarkan menjadi lebih jelas. Selain itu, ketika ada penambahan proses baru dapat dilakukan dengan mudah menggunakan flowchart ini. Setelah proses membuat flowchart selesai, maka giliran programmer yang akan menerjemahkan desain logis tersebut kedalam bentuk program dengan berbagai bahasa pemrograman yang telah disepakati.

2.8.1 Fungsi Flowchart

Setelah mengetahui definisi dari flowchart, mungkin Anda sudah dapat memahami untuk apa tujuan flowchart dibuat. Pada dasarnya, flowchart dibuat dengan tujuan untuk menunjukkan setiap proses yang harus dilalui dalam suatu

sistem. Namun, sebenarnya masih banyak lagi tujuan dari flowchart yang mungkin belum Anda ketahui. Berikut ini akan dijelaskan mengenai beberapa fungsi flowchart yang bisa Anda simak dibawah ini.

1. Merancang Proyek Baru

Ketika Anda akan merancang suatu proyek, maka hal selanjutnya yang dapat Anda lakukan adalah memetakan proyek tersebut ke bentuk flowchart. Itu dapat membantu Anda untuk merancang serangkaian langkah-langkah yang melibatkan keputusan bersama.

2. Mengelola Alur Kerja

Ketika Anda akan merancang suatu proyek, maka hal selanjutnya yang dapat Anda lakukan adalah memetakan proyek tersebut ke bentuk flowchart. Itu dapat membantu Anda untuk merancang serangkaian langkah-langkah yang melibatkan keputusan bersama.

3. Memodelkan Proses Bisnis

Proses bisnis yang dimaksud bukan hanya berkaitan dengan keuntungan, melainkan serangkaian tugas baik itu yang sederhana sampai yang rumit juga termasuk kedalam proses bisnis. Tujuan pemodelan flowchart dapat dilakukan untuk memberikan hasil yang konsisten dan dapat juga diprediksi.

4. Mendokumentasikan Setiap Proses

Dalam menyelesaikan suatu proyek perlu adanya dokumentasi proses. Dengan begitu, flowchart menjadi media yang bagus untuk memenuhi tujuan tersebut. Jika dibandingkan dengan membuat dokumentasi setiap proses melalui narasi dengan memetakannya ke dalam bentuk flowchart, maka menggunakan flowchart merupakan hal yang lebih efisien dibanding dengan narasi.

5. Merepresentasikan Algoritma

Biasanya, sebelum menuangkan proyek tersebut kedalam bentuk program, para perancang sistem terlebih dahulu menentukan algoritma untuk menyelesaikan proyek tersebut menggunakan SDL. SDL (Specification and Description Language) merupakan suatu spesifikasi





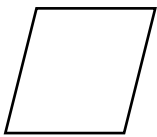
bahasa yang digunakan untuk mendeskripsikan perilaku dari sistem tersebut. Nah, flowchart dapat memenuhi kebutuhan tersebut. Sebab, flowchart menawarkan berbagai simbol yang unik yang dapat digunakan untuk memetakan sistem yang akan dirancang. Selain itu, flowchart juga menawarkan sumber daya yang dapat digunakan untuk menyelesaikan suatu masalah.

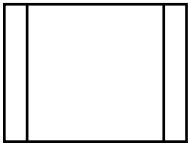
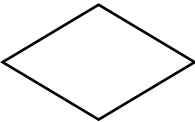
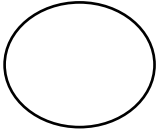
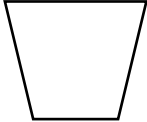
6. Mengaudit Proses

Secara umum, flowchart dapat digunakan untuk mendeteksi kerusakan yang terjadi dalam setiap proses. Flowchart dapat membantu Anda menyelesaikan permasalahan dengan cara membagi setiap langkah dari proses itu kedalam segmen-segmen yang lebih kecil, kemudian memeriksa bagian mana yang tidak berfungsi atau perlu diadakan perbaikan.

2.8.2 Simbol-Simbol Flowchat

Tabel 2. 8 Simbol-Simbol *Flowchart*

No	Simbol	Nama	Keterangan
1.		Terminator	Terminator untuk permulaan program dan akhir program.
2.		Garis Alur	Garis alur untuk aliran program
3.		Simbol Persiapan Program	Yaitu untuk proses instalasi dan pemberian nilai awal
4.		Simbol Proses	Proses perhitungan atau proses pengolahan data.
5.		<i>Input</i> Dan <i>Output</i>	Untuk proses input dan output data, parameter dan informasi

6.		Simbol Sub Proses	Untuk permulaan sub-program dan proses menjalankan sub-program.
7.		Seleksi	Untuk memilih aliran sesuai syarat
8.		<i>Connector</i>	Untuk menghubungkan bagian Flowchart yang berada pada satu halaman.
9.		Simbol Penghubung	Untuk menghubungkan Flowchart yang berbeda.

2.11 Xampp

Pada laptop atau PC umumnya sering digunakan XAMPP sebagai paket aplikasi web server. XAMPP merupakan singkatan dari X Apache MySQL PHP Perl. Simbol X (cross) menunjukkan bahwa XAMPP dapat berjalan cross (lintas) platform, yaitu pada 4 sistem operasi yang berbeda: Windows, Linux, Mac OS, dan Solaris. Sedangkan Apache merupakan modul dalam perangkat lunak XAMPP yang berfungsi sebagai web server. Jadi dalam hal ini XAMPP berfungsi sebagai server lokal untuk mengatur berbagai jenis data website yang sedang dikembangkan. (Fitri 2021)

XAMPP dapat menggantikan peran web hosting. Caranya adalah dengan menyimpan file website ke dalam localhost. File ini kemudian dapat dipanggil melalui browser. XAMPP bekerja secara offline seperti web hosting pada umumnya, tetapi tidak dapat diakses oleh banyak orang.

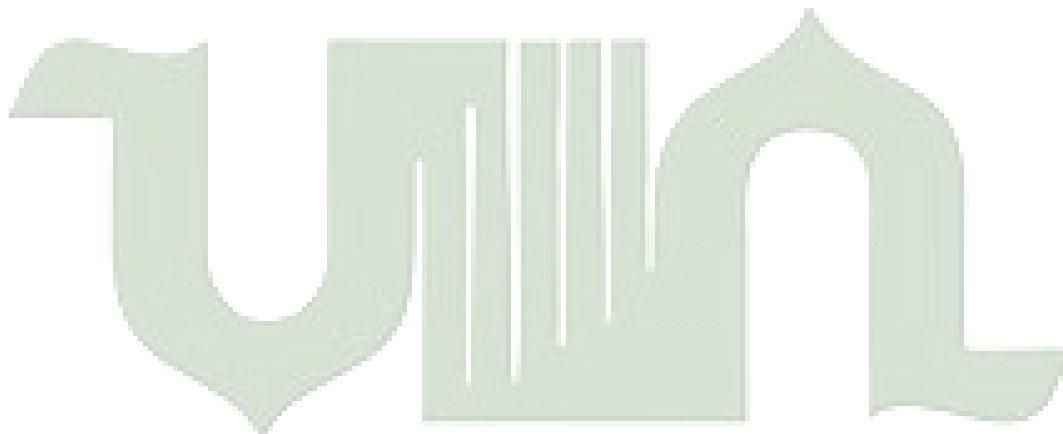
2.12 Penelitian Relevan

No.	Peneliti	Judul	Kesimpulan
1.	(Kurniabudi et al. 2021)	Komparasi <i>Information Gain, Gain Ratio, CFs-Bestfirst</i> dan <i>CFs-PSO Search</i> Terhadap Performa Deteksi Anomali	Hasil penelitian ini menunjukkan Teknik seleksi fitur yang penulis gunakan seperti <i>IG, GR, CB</i> dan <i>CF-PSO</i> mampu meningkatkan algoritma klasifikasi <i>NB, k-NN</i> dan <i>J48</i> untuk mengklasifikasikan trafik normal dan serangan tertentu, meskipun belum maksimal. Hal ini dapat dipengaruhi oleh jenis fitur yang terpilih. Hasil pengujian juga memperlihatkan bahwa jumlah fitur yang dianalisis sangat mempengaruhi waktu proses/komputasi.
2.	Muhammad Rivki dan Adam Mukharil Bachtiar (Bachtiar and Rivki 2017)	Implementasi Algoritma <i>K-Nearest Neighbor</i> Dalam Pengklasifikasian <i>Follower Twitter</i> Yang Menggunakan Bahasa Indonesia	penelitian ini menghasilkan sebuah program aplikasi yang dapat membantu memudahkan pengguna <i>Twitter</i> yang menggunakannya sebagai media pemasaran atau promosi untuk melakukan promosi dengan cara melakukan <i>twitt</i> promosi terhadap <i>follower</i> yang sudah diklasifikasikan, aplikasi ini juga dapat membantu memudahkan pengguna <i>Twitter</i> yang menggunakannya sebagai media pemasaran atau promosi untuk menentukan waktu <i>twitt</i> promosi yang tepat yang sudah dianalisa dari waktu aktif <i>follower</i> pengguna.

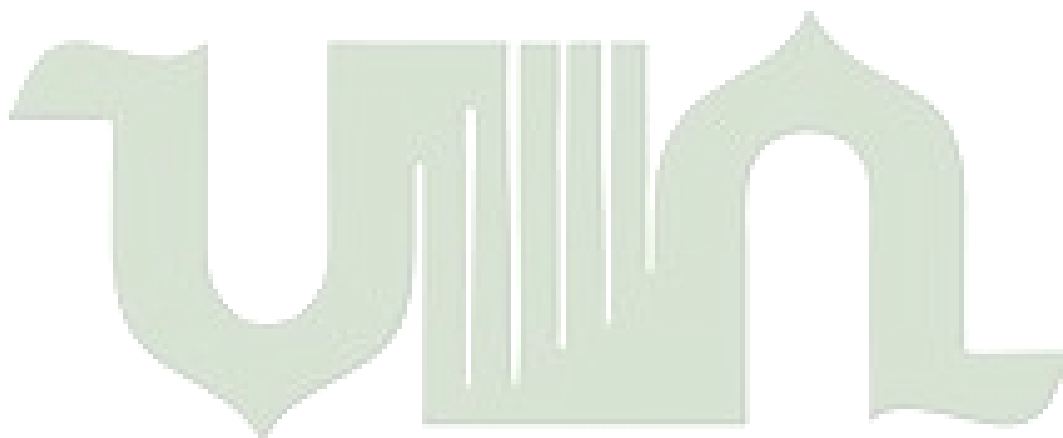
3.	(Setiyorini and Asmono 2020)	Implementation Of Gain Ratio And K-Nearest Neighbor For Classification Of Student Performance	Percobaan telah dilakukan 10 kali dengan nilai k adalah 1 sampai 10 menggunakan kumpulan data kinerja siswa. Hasil ini percobaan diperoleh ketelitian rata-rata sebesar 74.068 dengan K-Nearest Neighbor, dan diperoleh akurasi rata-rata 75,105 dengan Rasio Keuntungan dan K-Nearest Neighbor. Hasil eksperimen menunjukkan bahwa Gain Rasio mampu mengurangi dimensi tinggi fitur yang menjadi kelemahan K-Nearest Tetangga, sehingga penerapan Rasio Keuntungan dan K-Nearest Neighbor dapat meningkatkan akurasi klasifikasi prestasi siswa dibandingkan dengan menggunakan K-Nearest Neighbor saja.
4.	Fajriyan Nur, Moh. Ahsan, Wahyudi Harianto (Nur et al. 2022)	Komparasi Tingkat Akurasi Information Gain Dan Gain Ratio Pada Metode K-Nearest Neighbor	pada penelitian ini didapatkan kesimpulan bahwa ketika menggunakan seleksi fitur information gain dan gain ratio terbukti dapat meningkatkan nilai akurasi dibanding dengan menggunakan K-Nearest Neighbor konvensional. Berdasarkan dari pengujian menggunakan dataset glass dengan berbagai macam kondisi nilai k yang berbeda, sehingga dapat

			<p>disimpulkan penggunaan seleksi fitur gain ratio dianggap lebih optimal dibanding dengan information gain dikarenakan dalam pembobotan seleksi fitur gain ratio dianggap lebih baik dengan contoh perolehan hasil pembobotan pada atribut "Ba" mendapatkan nilai 0,7204 pada gain ratio, sedangkan pada information gain hanya mendapatkan nilai 0,41235 sehingga didapatkan selisih 0,3081</p>
5.	(Neighbor et al. 2019)	Aplikasi prediksi penjualan barang menggunakan metode <i>k-nearest neighbor</i> (knn) (studi kasus tumaka mart)	Berdasarkan hasil pengujian yang dilakukan, Aplikasi prediksi penjualan barang menggunakan metode <i>K-Nearest Neighbor</i> (KNN) (studi kasus Tumaka Mart) mampu melakukan prediksi penjualan dengan nilai error terkecil sebesar 0,001 % dan nilai error tertinggi adalah sebesar 1,231 %. Rata-rata nilai K yang menghasilkan error terkecil adalah nilai $k = 2$.
6.	(Muhammad et al. 2021)	Implementasi Metode K-Nearest Neighbor Untuk Prediksi Penjualan Kemasan Skincare Pada	Dari hasil penelitian ini dilakukan pemodelan menggunakan algoritma K-Nearest Neighbor dengan menggunakan data yang telah diolah berdasarkan tahapan Knowledge Discovery in Database (KDD). Hasil perhitungan data mining menggunakan teknik klasifikasi dan algoritma K-

		PT. Universal Jaya Perkasa	Nearest Neighbor, didapatkan hasil prediksi penjualan produk kemasan skincare dengan 7 kategori produk menggunakan data pada Januari 2019 hingga Juni 2021. Atribut terdiri dari kategori produk, kuantitas dan bulan dengan hasil pengujian perhitungan akurasi menggunakan tools Rapidminer Studio untuk mengetahui penjualan dari setiap produk beberapa bulan mendatang dan diperoleh hasil nilai akurasi sebesar 80%.
--	--	-------------------------------	--



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN