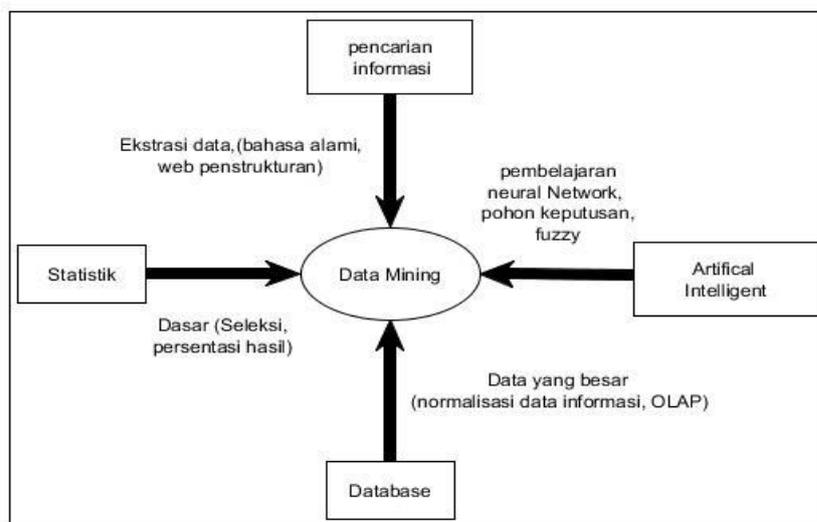


BAB II

LANDASAN TEORI

2.1 Data Mining

Menurut penelitian sebelumnya (Ochthario & Karnila, 2013), menjelaskan bahwa data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata mining sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu data mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (artificial intelligent), machine learning, statistik dan database.

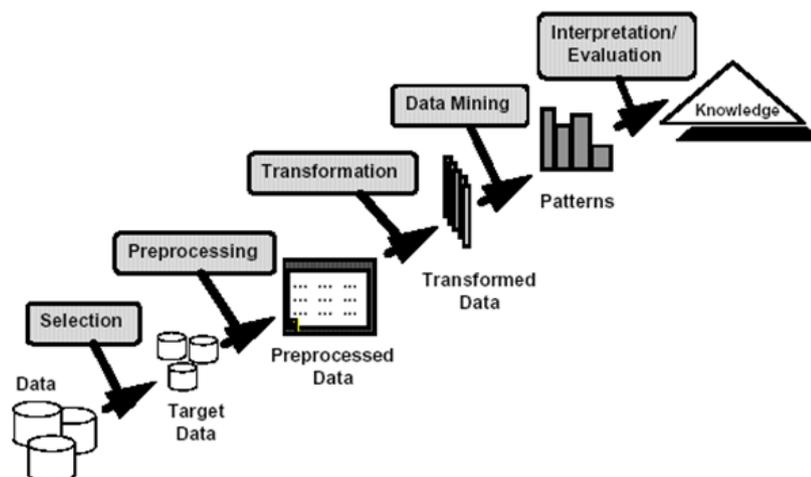


Gambar 2.1 Bidang Ilmu Data Mining

Menurut penelitian sebelumnya (Nurjoko & Kurniawan, 2016), Data mining membantu perusahaan untuk mendapatkan pola dari data-data yang tersimpan di dalam basis data perusahaan. Pengetahuan yang diperoleh tersebut akan menjadi pedoman dalam mengambil tindakan-tindakan bisnis sebagai upaya pemeliharaan dan meningkatkan tingkat kompetitif bisnis perusahaan. Data mining dapat menjalankan fungsi-fungsi berikut: deskripsi, estimasi, prediksi, klasifikasi, clustering, dan asosiasi.

2.2 Operasi Data Mining

Menurut (Hermawati, 2013), Operasi data mining menurut sifatnya dibedakan menjadi dua, yaitu bersifat (1) Prediksi (prediction driven) untuk menjawab pertanyaan apa dan sesuatu yang bersifat transparan. Operasi prediksi digunakan untuk validasi hipotesis, querying dan pelaporan (missal : spreadsheet dan pivot table), analisis multidimensi (dimensional summary); OLAP (Online Analytic Processing) serta analisis statistic. (2) Penemuan (discovery driven) bersifat transparan dan untuk menjawab pertanyaan “mengapa?”. Operasi penemuan digunakan untuk analisis data eksplorasi, pemodelan prediktif, segmentasi database, analisis keterkaitan (link analysis) dan deteksi deviasi.



Gambar 2.2 Proses KDD

Tahapan proses dalam penggunaan data mining yang merupakan proses knowledge discovery in database (KDD) seperti yang terlihat pada Gambar 2.2 dapat diuraikan sebagai berikut :

1. Memahami domain aplikasi untuk mengetahui dan menggali pengetahuan awal serta apa sasaran pengguna.
2. Membuat target data-set yang meliputi pemilihan data dan fokus pada sub-set data.
3. Pembersihan dan transformasi data meliputi eliminasi drau, outliers, missing value serta pemilihan fitur dan reduksi dimensi.

4. Penggunaan algoritma data mining yang terdiri dari asosiasi, sekuensial, klasifikasi, dan lain-lain.
5. Interpretasi, evaluasi dan visualisasi pola untuk melihat apakah ada sesuatu baru dan menarik dan dilakukan iterasi jika diperlukan

2.3 Clustering

Menurut penelitian sebelumnya (Satria & Aziz, 2016), menjelaskan bahwa clustering adalah suatu teknik analisis yang bertujuan untuk memilah obyek ke dalam beberapa kelompok yang mempunyai sifat berbeda antara kelompok satu dengan yang lain. Pada analisis ini tiap-tiap kelompok bersifat sejenis antar anggota dalam kelompok atau variasi obyek dalam kelompok yang terbentuk sekecil mungkin.

2.4 Algoritma K-Means

Menurut penelitian sebelumnya (Fahmi & Suprpto, 2013), menjelaskan bahwa Algoritma K-Means Clustering merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah K cluster yang sudah ditetapkan di awal. Algoritma K-Means Clustering sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi serta umum penggunaannya dalam paraktek. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster/kelompok yang lain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Jarak paling dekat antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Perhitungan jarak semua data ke setiap titik pusat cluster menggunakan teori jarak euclidean yang dirumuskan :

$$d(x, y) = \sqrt{\sum_{i=1}^r (x_i - y_i)^2} \dots (1)$$

Dimana :

$$\begin{aligned} d(x, y) &= \text{Jarak data } x \text{ ke } y \\ x_i &= \text{Nilai fitur ke-}i \text{ dari } x \\ y_i &= \text{Nilai fitur ke-}i \text{ dari } y \\ r &= \text{Jumlah fitur dalam vektor} \end{aligned}$$

Suatu data akan menjadi anggota dari cluster ke- j apabila jarak data tersebut ke pusat cluster ke- j bernilai paling kecil jika dibandingkan dengan jarak ke pusat cluster lainnya. Selanjutnya, kelompokkan data-data yang menjadi anggota pada setiap cluster. Pusat cluster yang baru didapatkan dengan menghitung rata-rata setiap fitur dari semua data yang tergabung dalam setiap cluster. Centroid dari semua data dalam cluster dinyatakan dengan :

$$c_j = \frac{1}{Nk} \sum_{l=1}^{Nk} x_{jl}$$

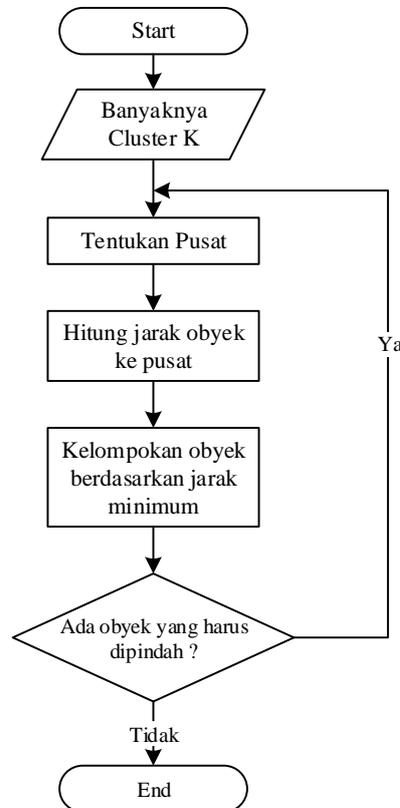
Dimana :

$$\begin{aligned} c_j &= \text{Centroid baru} \\ x_{jl} &= \text{Anggota cluster } l \text{ pada atribut ke-}j \\ Nk &= \text{Jumlah data dalam cluster} \end{aligned}$$

Nilai X_{jl} adalah anggota cluster l pada atribut ke- j setelah dihitung dengan teori jarak euclidean. dan Nk adalah jumlah data yang tergabung dalam sebuah cluster setelah dihitung dengan teori jarak euclidean. Langkah langkah melakukan Algoritma *K-Means Clustering* adalah sebagai berikut :

1. Tentukan nilai k sebagai jumlah *cluster*.
2. Pilih k dari dataset x sebagai *centroid*.
3. Alokasikan semua data ke centroid dengan metrik jarak menggunakan teori jarak euclidean.
4. Hitung kembali *centroid* c berdasarkan data yang mengikuti cluster masing-masing menggunakan rumus mencari centroid baru.
5. Ulangi langkah 3 dan 4 hingga kondisi konvergen tercapai, yaitu (a) perubahan fungsi objektif sudah dibawah ambang batas yang diinginkan atau (b) tidak ada data yang berpindah cluster atau perubahan posisi centroid sudah dibawah ambang batas yang ditetapkan.

Alur Algoritma *K-Means Clustering* ditunjukkan pada gambar dibawah ini :



Gambar 2.3 Flowchart Algoritma K-Means

Menurut penelitian sebelumnya (Alfina et al., 2012), menjelaskan bahwa metode k-means merupakan metode clustering yang paling sederhana dan umum. Hal ini dikarenakan k-means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Namun, k-means mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal cluster. Hasil cluster yang terbentuk dari metode k-means ini sangatlah tergantung pada inisiasi nilai pusat awal cluster yang diberikan. Hal ini menyebabkan hasil clusternya berupa solusi yang sifatnya *local optimal*.

2.5 Prioritas

Menurut Kamus Besar Bahasa Indonesia (KBBI), prioritas adalah yang didahulukan dan diutamakan daripada yang lain. Sedangkan memprioritaskan diartikan sebagai mendahulukan atau mengutamakan sesuatu daripada yang lain.

2.6 Distribusi

Menurut Kamus Besar Bahasa Indonesia (KBBI), distribusi adalah penyaluran (pembagian, pengiriman) kepada beberapa orang atau ke beberapa tempat. Sedangkan mendistribusikan diartikan sebagai menyalurkan (membagikan, mengirimkan) kepada beberapa orang atau ke beberapa tempat (seperti pasar, toko). Orang atau badan yang bertugas mendistribusikan sesuatu, atau penyalur disebut dengan distributor.

Menurut penelitian sebelumnya (Inma et al., 2013), salah satu prioritas dalam kegiatan pemasaran perusahaan adalah aspek penentuan saluran distribusi pemasaran produk. Karena organisasi atau perusahaan yang mampu menerapkan dan mengembangkan strategi saluran distribusi pemasaran yang tepat, dapat menghasilkan volume penjualan produk yang optimal, dapat mencapai suatu jangkauan yang lebih luas, barang atau jasa yang dihasilkan dapat terarah dengan baik secara langsung maupun tidak langsung. Saluran distribusi yang baik dan terarah dapat menjamin ketersediaan produk yang dibutuhkan oleh konsumen. Tanpa adanya saluran distribusi, produsen akan mengalami kesulitan untuk memasarkan produknya dan konsumen juga harus bersusahpayah untuk memperoleh produk yang diinginkannya.

2.7 Sayur Mayur

Menurut Kamus Besar Bahasa Indonesia (KBBI), sayur adalah daun-daunan atau tumbuh-tumbuhan yang boleh dimasak (seperti bayam, kubis, dan sebagainya). Sedangkan sayur-mayur adalah berbagai-bagai sayur.

2.8 PHP (Hypertext Preprocessor)

Menurut (Ardhama, 2012), PHP Hypertext Preprocessor atau sering disebut PHP merupakan bahasa pemrograman berbasis server-side yang dapat melakukan parsing script php menjadi menjadi script web sehingga dari sisi client menghasilkan suatu tampilan yang menarik. PHP merupakan pengembangan dari FI atau Form Interface yang dibuat oleh Rasmus Lerdoff pada tahun 1995.

Berbeda dengan HTML, kode PHP tidak diberikan secara langsung oleh server ketika ada permintaan atau request dari sisi client namun dengan cara pemrosesan dari sisi server. Kode PHP disisipkan pada kode HTML. Perbedaan dari kode (script) HTML dan PHP yaitu setiap kode PHP ditulis selalu diberi tag pembuka yaitu `<?php` dan pada ahir kode PHP diberi tag penutup yaitu `?>`. Adapun kelebihan dan kekurangan bahasa PHP antara lain sebagai berikut:

Tabel 2.1 Kelebihan Dan Kekurangan PHP

NO	KELEBIHAN	KEKURANGAN
1.	PHP menjadi populer karena kesederhanaannya dan kemampuannya dalam menghasilkan berbagai aplikasi web seperti <i>counter</i> , sistem artikel/CMS, <i>e-commerce</i> , <i>bulletin board</i> , dll.	Tidak detail untuk pengembangan skala besar.
2	PHP adalah salah satu bahasa <i>server-side</i> yang didesain khusus untuk aplikasi web.	Tidak detail untuk pengembangan skala besar.
3	PHP termasuk dalam <i>Open Source Product</i>	Tidak memiliki system pemrograman berorientasi objek yang sesungguhnya.
4	Aplikasi PHP cukup cepat dibandingkan dengan aplikasi CGI dengan <i>Perl / Python</i> bahkan lebih cepat dibanding dengan ASP maupun Java dalam berbagai aplikasi web.	Tidak bisa memisahkan antara tampilan dengan <i>logic</i> dengan baik.

2.9 HTML 5 (*Hypertext Markup Language 5*)

Menurut (Sianipar, 2015), HTML5 adalah sebuah markup untuk menstrukturkan dan menampilkan isi dari halaman web. HTML (yang pertama kali diciptakan pada tahun 1990 dan versi keempatnya, HTML4, pada tahun 1997) dan hingga bulan juni 2011 tetap dalam proses pengembangan. Tujuan utama pengembangan HTML5 adalah untuk memperbaiki teknologi HTML agar mendukung teknologi multimedia terbaru, mudah dibaca manusia dan juga mudah dimengerti oleh mesin. HTML5 merupakan salah satu karya World Wide Web Consortium, W3C untuk mendefinisikan sebuah bahasa markup tunggal yang dapat ditulis dengan cara HTML ataupun XHTML. HTML5 merupakan jawaban atas pengembangan HTML 4.01 dan XHTML 1.1 yang selama ini berjalan terpisah, dan diimplementasikan secara berbeda-beda oleh banyak perangkat lunak pembuat web.

2.10 MySQL

Menurut (M.Rudianto, 2011), MySQL dikembangkan oleh sebuah perusahaan Swedia bernama MySQL AB yang pada saat itu bernama TcX DataKonsult AB sekitar tahun 1994-1995, namun cikal bakal kodenya sudah ada sejak 1979. Awalnya TcX membuat MySQL dengan tujuan mengembangkan aplikasi web untuk klien. TcX merupakan perusahaan pengembang software dan konsultan database. Saat ini MySQL sudah diakuisisi oleh Oracle Crop. MySQL adalah salah satu jenis database server yang sangat terkenal dan banyak digunakan untuk membangun aplikasi web yang mana database sebagai sumber dan pengelolaan datanya.

Kepopuleran MySQL antara lain karena MySQL menggunakan SQL sebagai bahasa dasar untuk mengakses database sehingga mudah untuk digunakan. MySQL juga bersifat open source dan free pada berbagai platform kecuali pada windows yang bersifat shareware. MySQL didistribusikan dengan lisensi open source GPL (General Public License) mulai versi 3.23, pada bulan Juni 2000. Software MySQL bisa diunduh di <http://mysql.org> atau <http://www.mysql.com>.

2.11 Bootstrap

Menurut (Alatas, 2015), Bootstrap adalah framework ataupun tools untuk membuat aplikasi web ataupun website yang bersifat responsive secara cepat, mudah dan gratis. Kata 'responsive' disini berarti bahwa tampilan web (lebar dan susunan isinya) dapat berubah secara otomatis sesuai dengan lebar layar yang menampilkannya. Bootstrap terdiri dari CSS dan HTML untuk menghasilkan Grid, Layout, Tipografi, Tabel, Form, Navigasi, dan lain-lain. Didalam bootstrap juga sudah terdapat Javascript (jQuery Plugins) untuk menghasilkan komponen UI yang cantik seperti Transitions, Modal, DropDown, ScroollSpy, ToolTip, PopOver, Tab, Alert, Button, Carousel, dan lain-lain.

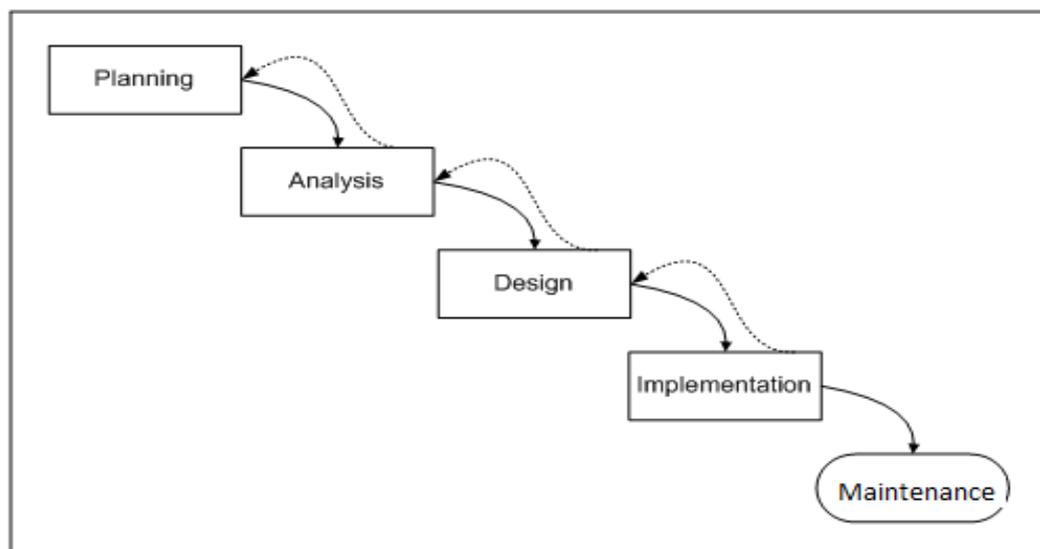
Dengan bootstrap kita dapat membuat responsive website dengan cepat dan mudah dan dapat berjalan sempurna pada browser-browser modern seperti Chrome, Firefox, Safari, Opera, dan Internet Explorer.

2.12 Sublime Text 3

Menurut (Faridl, 2015), Sublime text adalah teks editor berbasis Python, sebuah teks editor yang elegan, kaya fitur, cross platform, mudah dan simpel yang cukup terkenal di kalangan developer (pengembang), penulis dan desainer. Para programmer biasanya menggunakan sublime text untuk menyunting source code yang sedang ia kerjakan. Sampai saat ini sublime text sudah mencapai versi 3.

2.13 Metode Pengembangan Sistem *Waterfall*

Menurut (Shalahuddin & A.S, 2015), Model SDLC air terjun (*waterfall*) sering juga disebut model eksekusial linier (sequential linear) atau alur hidup klasik (classiclifecycle). Model air terjun menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau terurut dimulai dari planning, analisis, desain, pengkodean, pengujian dan tahap pendukung (support). Berikut ini adalah gambar model air terjun :



Gambar 2.4 Metode *Waterfall*

2.13.1 Planning (Perencanaan)

Tahap perencanaan merupakan proses penting untuk mengetahui mengapa sistem harus dibuat dan menentukan bagaimana cara membangun sistem tersebut. Langkah pertama dari proses tersebut adalah dengan

mengidentifikasi peluang apakah dapat memberikan kemungkinan biaya rendah tetapi menghasilkan keuntungan.

2.13.2 Analisis (Analisa Kebutuhan)

Proses pengumpulan kebutuhan dilakukan secara intensif untuk mespesifikasikan kebutuhan perangkat lunak agar dapat dipahami perangkat lunak seperti apa yang dibutuhkan oleh user. Spesifikasi kebutuhan perangkat lunak pada tahap ini perlu untuk didokumentasi.

2.13.3 Design (Desain)

Desain perangkat lunak adalah proses multi langkah yang focus pada desain pembuatan program perangkat lunak termasuk struktur data, arsitektur perangkat lunak, representasi antar muka, dan prosedur pengkodean.

2.13.4 Implementation (Pengujian)

Merupakan tahap berikutnya untuk menerjemahkan data atau pemecahan masalah yang telah dirancang ke dalam bahasa pemrograman komputer yang telah ditentukan. Semua tahap ini desain perangkat lunak sebagai sebuah program lengkap atau unit program.

2.13.5 Maintenance (Pemeliharaan)

Tidak menutup kemungkinan sebuah perangkat lunak mengalami perubahan ketika sudah dikirim ke user. Perubahan bisa terjadi karena adanya kesalahan yang muncul dan tidak terdeteksi saat pengujian atau perangkat lunak harus beradaptasi dengan lingkungan baru. Tahap pendukung dan pemeliharaan dapat mengulangi proses pengembangan mulai dari analisis spesifikasi untuk perubahan perangkat lunak yang sudah ada, tapi tidak untuk membuat perangkat lunak baru.

2.14 Diagram Flowchart

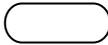
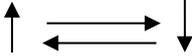
Menurut (Shalahuddin & A.S, 2015) menjelaskan bahwa Bagan Alir merupakan alat pengembangan sistem yang akan membantu menggambarkan sistem yang sedang berjalan pada perusahaan. Bagan alir (*flowchart*) adalah bagan (*chart*) yang menunjukkan alir (*flow*) di dalam program atau prosedur sistem secara logika. Bagan

alir digunakan terutama untuk alat bantu komunikasi dan untuk dokumentasi. Terdapat beberapa jenis bagan alir yang biasa digunakan, yaitu sebagai berikut :

2.14.1 Bagan Alir Dokumen (*Document Flowchart*)

Bagan alir dokumen (*document flowchart*) atau disebut dengan bagan alir formulir (*form flowchart*) atau *paperwork flowchart* merupakan bagan alir yang menunjukkan arus dari laporan dan formulir termasuk tembusan-tembusannya. Adapun simbol bagan alir dokumen dapat dilihat pada tabel berikut :

Tabel 2.2 Bagan Alir Dokumen (*Document Flowchart*)

SIMBOL	KETERANGAN
Dokumen 	Menunjukkan dokumen yang digunakan untuk input dan output baik secara manual maupun komputerisasi
Proses Manual 	Menunjukkan pekerjaan yang dilakukan secara manual.
Proses Komputerisasi 	Menunjukkan proses dari operasi program komputer.
Simpanan 	Menunjukkan arsip atau Menunjukkan bahwa data dalam simbol ini akan disimpan ke suatu media tertentu
Terminator 	Menyatakan permulaan atau akhir suatu program
Garis alir 	Menyatakan jalannya arus suatu proses
Decision 	Digunakan untuk suatu penyeleksian kondisi di dalam program.
Keyboard 	Memasukkan data secara manual dengan menggunakan keyboard

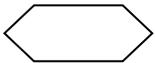
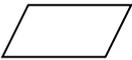
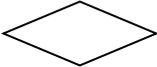
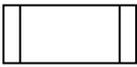
Tabel Lanjutan 2.2 Bagan Alir Dokumen (*Document Flowchart*)

<p>Hard disk</p> 	<p>Media penyimpanan, menggunakan perangkat hard disk.</p>
<p>Keterangan</p> 	<p>Digunakan untuk memberikan keterangan yang lainnya.</p>
<p>Penghubung</p> 	<p>Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang sama Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang berbeda</p>

2.14.2 Bagan Alir Program (*Program Flowchart*)

Bagan alir program (*program flowchart*) merupakan bagan yang menjelaskan secara rinci langkah-langkah dari proses program. Bagan alir program terdiri dari dua macam, yaitu baganalir logika program (*program logic flowchart*) dan bagan alir program komputer terinci (*detailed computer program flowchart*). Bagan alir logika program digunakan untuk menggambarkan tiap-tiap langkah didalam program komputer secara logika, sedangkan bagan alir program komputer terperinci digunakan untuk menggambarkan instruksi-instruksi program komputer secara terperinci. Bagan alir logika program dipersiapkan oleh analis sistem. sedangkan bagan alir program komputer terperinci dipersiapkan oleh pemrogram. Berdasarkan penjelasan diatas penulis dapat menarik kesimpulan bahwa baganalir (*flowchart*) adalah suatu gambaran umum tentang sistem yang berjalan yang berfungsi sebagai alat bantu komunikasi dan untuk dokumentasi serta menyajikan kegiatan baik secara manual maupun komputerisasi. Adapun simbol bagan alir program dapat dilihat pada table berikut:

Tabel 2.3 Bagan Alir Program (*Program Flowchart*)

SIMBOL	KETERANGAN
Terminator 	Digunakan untuk memberikan awal dan akhir suatu proses.
Proses 	Menunjukkan proses dari operasi program komputer.
<i>Preparation</i> 	Proses inisialisasi/pemberian harga awal.
Input/Output Data 	Proses input/output data, parameter, informasi.
Garis alir 	Digunakan untuk menunjukkan arus dari proses.
<i>Decision</i> 	Digunakan untuk suatu penyeleksian kondisi di dalam program.
Proses terdefinisi 	Simbol yang digunakan untuk menunjukkan suatu operasi yang rinciannya ditunjukkan ditempat lain.
Penghubung 	Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang sama, Menyatakan sambungan dari proses ke proses lainnya dalam halaman yang berbeda

2.15 DFD (*Data Flow Diagram*)

Menurut (Shalahuddin & A.S, 2015) *Data Flow Diagram* adalah representasi grafik yang menggambarkan aliran informasi dan tranformasi informasi yang diaplikasikan sebagai data yang mengalir dari masukan(input) dan keluaran(output). DFD dapat digunakan untuk merepresentasikan sebuah sistem atau perangkat lunak pada beberapa level abstraksi. DFD dapat dibagi menjadi beberapa level yang lebih detail untuk merepresentasikan aliran informasi atau fungsi yang lebih detail. DFD menyediakan mekanisme untuk pemodelan fungsional ataupun pemodelan aliran informasi. Notasi pada DFD adalah sebagai berikut :

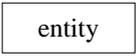
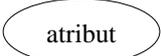
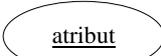
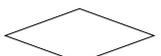
Tabel 2.4 Notasi *Data Flow Diagram*

NOTASI	KETERANGAN
	Proses atau fungsi atau prosedur; pada pemodelan perangkat lunak yang akan diimplementasikan dengan pemrograman terstruktur, maka pemodelan notasi inilah yang harus menjadi fungsi atau prosedur didalam kode program
	<i>File</i> atau basis data atau penyimpanan (<i>storage</i>) pada pemodelan perangkat lunak yang akan diimplementasikan dengan pemrograman terstruktur, maka pemodelan notasi inilah yang harusnya dibuat mejadi table-tabel basis data yang dibutuhkan, table-tabel ini juga harus sesuai dengan perancangan table-tabel pada basis data (ERD)
	Entitas luar (<i>external entity</i>) atau masukan atau keluaran atau orang-orang yang memakai atau berinteraksi dengan perangkat lunak yang dmodelkan atau system lain yang terkait dengan aliran data dari system yang dimodelkan
	Aliran data; merupakan data yang dikirim antar proses, dari penyimpanan keproses, atau dari proses kemasukan (<i>input</i>) atau keluaran (<i>ouput</i>)

2.16 ERD (*Entity Relational Diagram*)

Menurut (Shalahuddin & A.S, 2015) pemodelan awal basis data yang paling banyak digunakan adalah *Entity Relationship Diagram* (ERD). ERD dikembangkan berdasarkan teori himpunan dalam bidang matematika. ERD digunakan untuk pemodelan basis data relasional. Notasi-notasi pada ERD adalah seperti pada tabel berikut ini :

Tabel 2.5 Notasi Pada Model ERD

SIMBOL	KETERANGAN
Entitas /entity 	Entitas/entity merupakan data inti yang akan disimpan; bakal table pada basis data
Atribut 	Atribut/field; atau kolom data yang butuh disimpan dalam suatu entitas
Atribut Kunci Primer 	<i>Field</i> atau kolom yang butuh disimpan dalam suatu entitas dan digunakan sebagai kunci akses record yang diinginkan; biasanya berupa id
Atribut Multi nilai 	<i>Field</i> atau kolom data yang butuh disimpan dalam suatu entitas yang dapat memiliki nilai lebih dari satu.
Relasi 	Relasi yang menghubungkan antar entitas; biasanya diawali dengan kata kerja.
Asosiasi / Association 	Penghubung adalah penghubung antar relasi dan entitas dimana di kedua ujungnya memiliki <i>multiplicity</i> kemungkinan jumlah pemakaian.

2.17 Penelitian Terdahulu

Pada bagian ini akan membahas mengenai tabel penelitian terdahulu yang berhubungan dengan karya ilmiah penulis, yang terdapat pada tabel berikut :

Tabel 2.6 Penelitian Terdahulu

NO	NAMA	JUDUL	DESKRIPSI
1	Fiqih Satria, RZ. Abdul Aziz (2016)	Perbandingan Kinerja Metode Ward dan K-Means Dalam Menentukan Cluster Data Mahasiswa Pemohon Beasiswa (Studi Kasus : Stmik Pringsewu)	Penelitian ini bertujuan untuk mengetahui langkah-langkah analisis cluster dengan metode metode Ward dan metode K-Means, serta membandingkan hasil analisis kedua metode tersebut untuk mengclusterkan data mahasiswa terkait pengambilan keputusan untuk menentukan mahasiswa yang berhak menerima beasiswa di STMIK Pringsewu.

Tabel Lanjutan 2.6 Penelitian Terdahulu

2	Raymond Othario, Sri Karnila (2013)	Sistem Data Mining Untuk Mengetahui Tingkat Kecenderungan Memilih Menu Makanan Dengan Metode Association Rule Mining (Studi Kasus : Kedai Kemangi)	Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar. Salah satu metode yang ada dalam data mining adalah metode association rule dimana metode ini mencari sekumpulan items yang sering muncul bersamaan.
3	Nurjoko, Hendra Kurniawan (2016)	Aplikasi Datamining Untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Apriori di IBI Darmajaya Bandar Lampung	Dengan memanfaatkan data induk mahasiswa dan data kelulusan mahasiswa, diharapkan dapat menghasilkan informasi tentang tingkat kelulusan dengan data induk mahasiswa melalui teknik data mining. Kategori tingkat kelulusan di ukur dari lama studi dan IPK. Algoritma yang digunakan adalah algoritma apriori, informasi yang ditampilkan berupa nilai support dan confidence dari masing-masing kategori tingkat kelulusan.
4	Shindy Inma, Edy Wahyudi, Yuslinda Dwi Handini (2013)	Identifikasi Saluran Distribusi dan Keterbatasan Daya Saing Kelompok Tani Sayuran di Desa Argosari	Penelitian ini bertujuan untuk mengetahui dan mendeskripsikan identifikasi saluran distribusi sayuran hasil panen, serta untuk mengetahui penyebab kelompok tani sayuran di Desa Argosari mengalami keterbatasan daya saing. Hasil penelitian menemukan bahwa kegiatan distribusi langsung hanya akan terjadi jika ada konsumen yang mengunjungi Desa Argosari. Sedangkan kegiatan distribusi tak langsung terjadi melalui petanipedagang ritel-konsumen atau petani-pedagang besar-pedagang ritel-konsumen.

Tabel Lanjutan 2.6 Penelitian Terdahulu

5	Muhamma d Farid Fahmi, Yoyon K. Suprpto (2013)	Penentuan Prioritas Rehabilitasi DAS Menggunakan Algoritma K-Means Clustering	Ketersediaan informasi mengenai tingkat kekeritisan lahan yang akurat mempunyai arti yang sangat penting dalam program RHL sehingga Priorotas DAS mana yang akan dilakukan rehabilitasi bisa diketahui. Dari permasalahan diatas diperlukan sebuah cara untuk menentukan prioritas DAS yang direhabilitasi. Adapun metode yang digunakan dalam penelitan ini adalah K-Means Clustering. KMeans Clustering memodelkan dataset menjadi klaster-klaster dimana data pada satu klaster yang memiliki karakteristik sama dan memiliki karakteristik yang berbeda dari klaster lain berdasarkan parameter tingkat kekeritisan lahan.
6	Tahta Alfina, Budi Santosa, dan Ali Ridho Barakbah (2012)	Analisa Perbandingan Metode Hierarchical Clustering, K- means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)	Salah satu teknik yang dikenal dalam data mining adalah clustering, berupa proses pengelompokan sejumlah data atau objek ke dalam cluster (group) sehingga setiap dalam cluster tersebut akan berisi data yang semirip mungkin dan berbeda dengan objek dalam cluster yang lainnya. Clustering memiliki dua metode, yaitu partisi dan hierarki. Dua metode ini memiliki kelebihan dan kekurangan masing-masing, dan dengan menggabungkan keduanya dapat diperoleh hasil cluster yang lebih baik. Dari hasil cluster dengan menggunakan data problem Kerja Praktek Jurusan Teknik Industri ITS, maka diperoleh hasil bahwa gabungan metode Single Linkage Clustering dan K-means memberikan hasil cluster yang lebih baik dengan parameter uji cluster variance dan metode silhouette coefisien.

