

BAB II

LANDASAN TEORI

2.1 DATA MINING

Data mining sering disebut juga *knowledge discovery in database*, adalah kegiatan yang meliputi pengumpulan dengan pemakaian data *history* untuk menemukan keteraturan pola atau hubungan dalam set data yang besar. Istilah *data mining* memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki .

Menurut (Kamber et al,2012) Mengatakan bahwa :

“Mining adalah istilah yang sangat jelas untuk menggambarkan proses yang menemukan sejumlah kecil nominal berharga dari banyaknya bahan mentah”.

Menurut Retno Tri Vulandari (2017:1) menyatakan bahwa:

“*Data Mining* adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstrasi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Databases (KDD)*.”

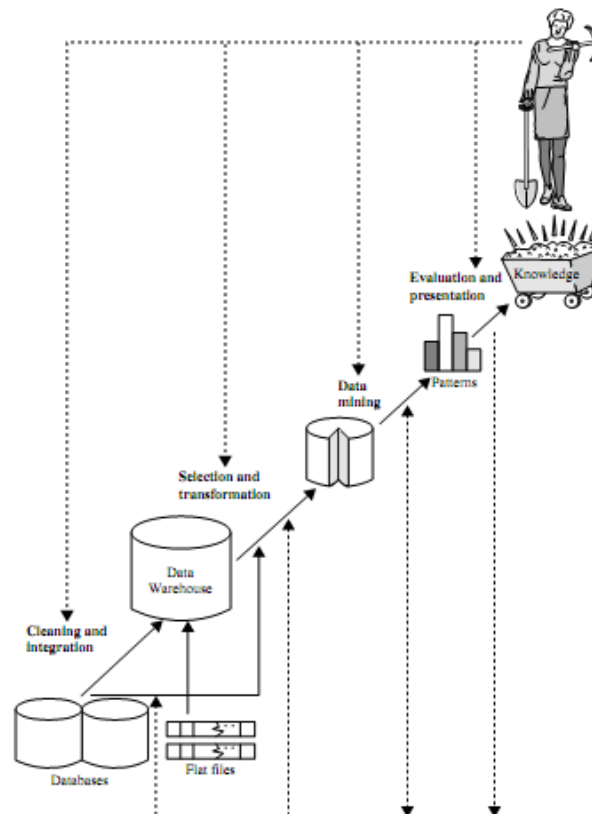
Menurut Turban dalam jurnal Dicky Nofriansyah (2016 : 84)

“Data Mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik,database dan visualisasi untuk pengenalan permasalahan pengambilan informasi dari database yang besar”.

Dari beberapa pendapat diatas , maka dapat disimpulkan bahwa *data mining* adalah merupakan sebuah teknik dari beberapa bidang ilmu yang dapat mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar.

2.2 PROSES TAHAPAN *DATA MINING*

Istilah *data mining* dan *knowledge discovery in databases* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar .Beberapa tahapan yang dimiliki data mining berdasarkan tugas yang bisa dilakukan, (Kamber et al.) yaitu :



Gambar 2.1 Tahap-tahap data mining

(Kamber et al,2012)

Tahap-tahap *data mining* adalah sebagai berikut:

1. Pembersihan data (*data cleaning*)

Untuk menghilangkan kebisingan dan data yang tidak konsisten.

2. Integrasi data (*data integration*)

Melakukan pembersihan data dan integrasi data sebagai langkah preprocessing, di mana data yang dihasilkan disimpan dalam data warehouse.

3. Seleksi data (*data selection*)

Di mana data yang relevan dengan tugas analisis diambil dari database.

4. Transformasi dari *database*

Di mana data ditransformasikan dan dikonsolidasikan ke dalam bentuk yang sesuai untuk penambangan dengan melakukan operasi ringkasan atau agregasi.

5. Proses mining

Proses penting di mana metode cerdas diterapkan untuk mengekstraksi pola data

6. Evaluasi pola (*patren evaluation*)

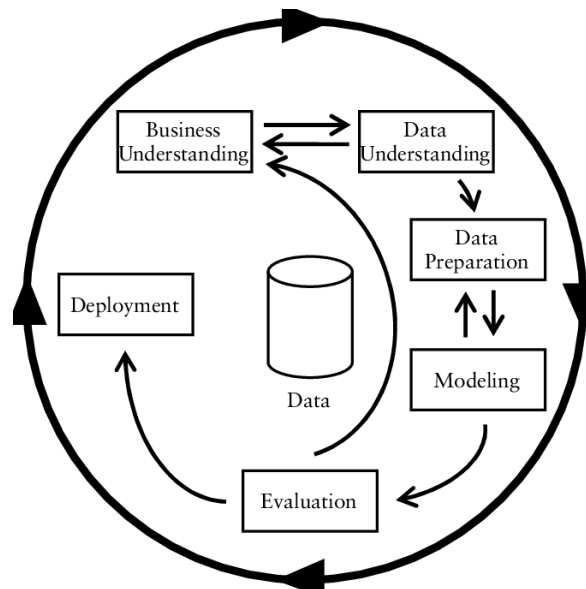
Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan.

7. Presentasi pengetahuan (*knowledge presentation*)

Di mana visualisasi dan teknik representasi pengetahuan digunakan untuk menyajikan pengetahuan yang ditambang kepada pengguna

2.3 CRIPS-DM

CRISP-DM adalah metodologi yang dibentuk oleh komisi Eropa pada tahun 1996 yang menerapkan standar dalam proses data mining. Dalam CRISP-DM terdapat enam fase yang akan dilakukan dalam penelitian pengembangan data mining yaitu (Defiyanti Sofi, 2013) dapat di jelaskan dalam gambar berikut :



Gambar 2.2 Fase CRISP-DM

1. Fase Business Understanding (Pemahaman Bisnis)

Pada fase pemahaman bisnis ini adalah tahapan pertama dalam pengembangan data mining. Pada fase ini adalah fase dimana memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian menterjemahkan pemahaman tersebut kedalam pendefinisian data mining, yang kemudian akan diterjemahkan menjadi rencana dan strategi untuk mencapai tujuan tersebut.

2. Fase Data understanding (Pemahaman Data)

Pemahaman data dilakukan pada tahap kedua setelah pemahaman bisnis selesai dilakukan. Dalam tahap pemahaman data dimulai dengan pengumpulan data awal dan hasil yang menarik dari data yang dapat digunakan. Kemudian akan dilanjutkan dengan proses mendapatkan pemahaman yang mendalam tentang data, mengidentifikasi masalah kualitas data, atau untuk mendeteksi adanya bagian yang menarik dari data yang dapat digunakan untuk hipotesa informasi yang tersembunyi.

3. Fase Data Preparation (Persiapan Data)

Tahap persiapan data dilakukan setelah tahap pemahaman data selesai dilakukan. Tahap persiapan data adalah tahapan yang mencakup semua kegiatan yang diperlukan untuk membangun dataset akhir atau data yang akan dimasukkan kedalam alat pemodelan yang berasal dari data mentah awal. Tahap persiapan data kemungkinan akan dilakukan beberapa kali dan tidak dalam urutan yang ditentukan. Pada tahap ini mencakup pemilihan tabel, record, atribut-atribut data, termasuk proses pembersihan dan transformasi data untuk tahap pemodelan.

4. Fase Modeling (Pemodelan)

Tahapan Pemodelan dilakukan setelah tahap persiapan data selesai dilakukan. Pada tahap pemodelan dilakukan pemilihan dan penerapan teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapatkan nilai yang optimal. Terdapat beberapa pemodelan untuk menyelesaikan masalah yang sama dalam data mining. Beberapa pemodelan

tersebut memiliki spesifikasi dan format data yang khusus. Oleh karena itu pada tahap pemodelan akan sangat memungkinkan untuk mebalik ke tahap sebelumnya yaitu tahap persiapan data.

5. Fase Evaluation (Evaluasi)

Setelah tahap pemodelan selesai dilakukan maka sebuah model atau beberapa model sudah terbentuk. Sebelum masuk ke dalam tahap penyebaran akhir dari model, maka perlu diadakan evaluasi dan meninjau apakah model yang tercipta sudah benar-benar mencapai tujuan bisnis yang telah ditetapkan pada tahap pemahaman bisnis.

6. Fase Deployment (penempatan)

Setelah fase evaluasi selesai dilakukan dan didapat hasil yang memuaskan dalam tahap evaluasi, bahwa semua tujuan bisnis sudah tercapai. Maka masuk ke dalam fase deployment. Pada fase deployment ini, pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna.

2.4 PENGELOMPOKAN DATA MINING

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Kusrini dan Emha Taufiq Luthfi, 2009) :

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menemukan 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 yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

Contoh lain yaitu estimasi nilai indeks prestasi kumulatif mahasiswa program pascasarjana dengan melihat nilai indeks prestasi mahasiswa tersebut pada saat mengikuti program sarjana

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

Contoh prediksi dalam bisnis dan penelitian adalah:

- a. Prediksi harga beras dalam tiga bulan yang akan datang .

- b. Prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan.

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.

Contoh lain klasifikasi dalam bisnis dan penelitian adalah:

- a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
- b. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- c. Mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa.

5. Pengklasteran

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 *recond* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

Contoh pengklusteran dalam bisnis dan penelitian adalah:

- a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
- c. Melakukan pengklusteran terhadap ekspresi dari gen, untuk mendapatkan kemiripan perilaku dari gen dalam jumlah besar.

6. Asosiasi

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

Contoh asosiasi dalam bisnis dan penelitian adalah:

- a. Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respons positif terhadap penawaran *upgrade* layanan yang diberikan.

- b. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

2.5 ATURAN ASOSIASI

Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi item. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien. Dalam dunia bisnis, sering disebut dengan *affinity Analysis* atau *Market Basket Analysis*, Fungsi *Association rules* sering kali disebut dengan *Market Basket Analysis*, yang digunakan untuk menentukan kombinasi diantara himpunan item-item. *Market Basket Analysis* adalah kebiasaan membeli konsumen dengan mencari asosiasi dan kombinasi antara item-item berbeda (Mateus,2018)

Aturan asosiasi (*association rule*) adalah salah satu konsep menarik dalam data mining yang berusaha menemukan asosiasi atau keterkaitan data. Diberi nama *market basket analysis* karena pada awal ditemukannya, konsep ini berkaitan dengan barang – barang yang berada di dalam keranjang belanjaan yang secara langsung menunjukkan tingkah laku konsumen berbelanja (Fajar Astuti Hermawati,2013)

Menurut (Kamber et al,2012) Mengatakan bahwa :

“Aturan asosiasi yang di hasilkan dari data mining dari beberapa tingkat abstraksi yang disebut dengan multi tingkat atau bertingkat aturan asosiasi. ”.

Menurut Retno Tri Vlandari (2017:65) menyatakan bahwa:

“Analisa asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan asosiasi antara suatu kombinasi item ”

Menurut Michael J.A Berry & Gordon S. Linoff di dalam jurnal Eka Fitria Wulansari Analisis asosiasi atau *association rule mining* diartikan sebagai :

“Teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item dalam suatu basis data relasional. Karena kegunaannya secara luas dapat dipergunakan untuk kepentingan bisnis dan menganalisa data transaksi, analisis asosiasi sering juga disebut dengan Market Basket analysis”

Dari beberapa pendapat di atas, maka dapat di simpulkan bahwa asosiasi didefinisikan dari suatu proses untuk menemukan semua aturan asosiasi yang memenuhi syarat minimum untuk *support (minimum support)* dan syarat minimum untuk *confidence (minimum confidence)*.

Di dalam jurnal Efori Buulolo (2013) metodologi dasar analisis asosiasi terbagi menjadi dua tahap :

1. Analisis Pola Frekuensi Tinggi

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* sebuah *item* diperoleh dengan rumus berikut :

$$Support (A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ transaksi} \dots\dots\dots (2.1)$$

Sementara itu, nilai support dari 2 item diperoleh dari rumus 2 berikut :

$$Support (A, B) = P(A \cap B)$$

$$Support (A, B) = \frac{\sum Transaksi \text{ mengandung } A \text{ dan } B}{\sum Transaksi} \dots\dots\dots (2.2)$$

2. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus berikut (Umami, 2015):

$$Confidence = P(B|A) = \frac{\sum Transaksi \text{ mengandung } A \text{ dan } B}{\sum Transaksi \text{ mengandung } A} \dots\dots\dots (2.3)$$

Ada beberapa metode untuk menyelesaikan dalam aturan asosiasi, yaitu :

1. Algoritma *Hash Based*

Algoritma *hash based* menggunakan teknik *hashing* untuk menyaring keluar *itemset* yang seluruh kemungkinan $(k+1)$ -*itemset* dihash ke dalam hash table dengan menggunakan hash (yang menggunakan sebuah bilangan prima untuk operasi modulo).

2. Algoritma *Generalized Rule Induction*

Generalized Rule Induction menggali *rule* yang memberikan informasi tertinggi berdasarkan *support* dan *confidence*. Memiliki batasan jumlah *rule* yang dipertahankan. GRI dapat mengatasi input data numerik dan kategorikal, namun target harus kategorikal.

3. Algoritma *FP-Growth*

Frequent Pattern Growth (FP-Growth) adalah algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. Dalam

Algoritma *Frequent Pattern Growth (FP-Growth)* menyimpan informasi mengenai frequent itemset dalam bentuk struktur prefix-tree atau sering di sebut *FP-Tree*.

4. Algoritma Apriori

Algoritma apriori diusulkan oleh R. Agrawal dan R Srikant pada tahun 1994 untuk menggali *frequent item set* dalam memperoleh *association rule* yang paling kuat hubungan antar atributnya. *Frequent item set* adalah himpunan transaksi yang terjadi berdasarkan *minimum support*. *Rule* yang kuat adalah yang memenuhi syarat *minimum support* dan *minimum confidence*. Algoritma apriori menggunakan pencarian berulang berdasarkan level *item set*, dimana *k-itemsets* digunakan untuk menemukan *item set* pada $k+1$. Metode dasar pada algoritma apriori adalah menemukan himpunan *frequent* pada *1-itemsets* ($k=1$), misalkan dinotasikan L1. L1 kemudian digunakan untuk menemukan himpunan *frequent 2-itemsets*, L2. Selanjutnya L2 akan digunakan untuk menemukan L3 begitu seterusnya sampai tidak ada *frequent k-itemsets* ditemukan (Wardah, 2017)

2.6 ALGORITMA APRIORI

Dalam penelitian ini menggunakan algoritma apriori karena memiliki akurasi yang baik dalam keterhubungan antar item dengan melihat nilai *minimum support* dan nilai *minimum confidence*, yang hasilnya nanti dapat digunakan untuk mengetahui media promosi dan pemasaran apa yang paling cocok untuk pelanggan (Fitriati, 2016)

Algoritma apriori adalah satu algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk menemukan *frequent itemsets* pada aturan asosiasi Boolean. Ide utama pada algoritma apriori adalah : pertama, mencari *frequent itemset* (himpunan item-item yang memenuhi *minimum support*) dari basis data transaksi. Kedua, menghilangkan *itemset* dengan frekuensi yang rendah berdasarkan level *minimum support* yang telah ditentukan sebelumnya. Selanjutnya membangun aturan asosiasi dari *itemset* yang memenuhi nilai *minimum confidence* dalam basis data. Algoritma apriori menggunakan *knowledge* mengenai *frequent itemset* yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya. Pola frekuensi tinggi adalah pola-pola *item* dalam *database* yang memiliki frekuensi atau *support* di atas ambang batas tertentu yang disebut istilah *minimum support*. Pola frekuensi tinggi ini digunakan untuk menyusun aturan asosiatif dan juga beberapa teknik *data mining* lainnya (Agrawal & Srikant dalam jurnal Putria, 2018)

Menurut (Kamber et al,2012) Mengatakan bahwa :

“Apriori menggunakan pendekatan berulang yang dikenal sebagai pencarian level-bijaksana, di mana k-itemset digunakan untuk mengeksplorasi (k + 1) -itemset. Set 1-itemset sering ditemukan dengan memindai basis data untuk mengakumulasi jumlah untuk setiap item, dan mengumpulkan item-item yang memenuhi dukungan minimum”

Menurut Erwin di dalam jurnal Eka Fitria Wulansari (2014) Algoritma apriori diartikan sebagai berikut :

“Suatu bentuk algoritma dalam data mining yang akan memberikan informasi tentang hubungan antar item dalam database yang dapat di manfaatkan secara luas dalam proses bisnis khususnya dalam proses penjualan. Knowledge algoritma apriori terletak pada frequent itemset yang telah diketahui sebelumnya, untuk memproses informasi selanjutnya”

Untuk membentuk kandidat *itemset* ada dua proses utama yang dilakukan algoritma apriori (Han & Kamber dalam artikel Dewi dkk, 2016) :

1. *Join Step* (Penggabungan)

Pada proses ini setiap item dikombinasikan dengan item lainnya sampai tidak terbentuk kombinasi lagi.

2. *Prune Step* (Pemangkasan)

Pada proses ini, hasil dari item yang dikombinasikan tadi kemudian dipangkas dengan menggunakan *minimum support* yang telah ditentukan oleh user.

2.7 DATA MINING TOOLS

Ada banyak *tools* yang tersedia untuk *data mining*. Tujuan utama *data mining tools* adalah untuk menemukan data, mengekstrak data, menyaring data, mendistribusikan informasi dan memonetisasinya. Berikut beberapa penjelasan *data mining tools* antara lain:

1. *RapidMiner*

RapidMiner adalah sebuah lingkungan *machine learning data mining*, *text mining* dan *predictive analytics* (Han, J. and Kamber, M dalam jurnal Rima Dias Ramadhani 2013: 3). *RapidMiner* adalah salah satu *software* untuk pengolahan *data mining*. Pekerjaan yang dilakukan oleh *RapidMiner text mining* adalah berkisar dengan analisis teks, mengekstrak pola-pola dari data set yang besar dan mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan *database*. Tujuan dari analisis teks ini adalah untuk mendapatkan informasi bermutu tertinggi dari teks yang diolah. *RapidMiner* menyediakan prosedur *data mining* dan *machine learning*, di dalamnya termasuk: ETL (*extraction*,

transformation, loading), data *preprocessing*, visualisasi, *modelling* dan evaluasi. Proses *data mining* tersusun atas operator-operator yang *nestable*, dideskripsikan dengan XML, dan dibuat dengan GUI. Penyajiannya dituliskan dalam bahasa pemrograman Jawa.

2. WEKA

WEKA adalah sebuah paket tools machine learning praktis. “*Waikato Environment for Knowledge Analysis (WEKA)* merupakan perangkat lunak pembelajaran mesin yang populer yang ditulis dalam bahasa pemrograman *java*. *WEKA* dikembangkan di Universitas Waikato, Selandia Baru. *WEKA* berisikan kumpulan algoritma beserta visualisasinya untuk analisis data dan pemodelan prediktif. Algoritma-algoritma pembelajaran mesin pada *WEKA* dapat dimanfaatkan untuk pemecahan masalah dibidang *data mining*. *WEKA* versi asli awalnya dirancang untuk menganalisis data dari domain pertanian, tetapi *WEKA* versi lengkap berbasis *java* (versi 3), yang mulai dibangun pada tahun 1997, yang sekarang dapat digunakan untuk menganalisis data dari berbagai domain, khususnya untuk pendidikan dan penelitian (Retno Tri Vulandari,2017)

3. SPSS

Menurut Priyatno, dalam buku Haryadi Sarjono dan Winda Julianita (2011: 113) mengatakan bahwa *SPSS* adalah program atau *software* yang digunakan untuk mengolah data statistik. Dari berbagai program olah data statistik lainnya, *SPSS* merupakan Program yang paling banyak digunakan. Umumnya, *SPSS* dulu digunakan untuk mengolah data statistik pada ilmu sosial saja. Namun demikian seiring berjalannya waktu, *SPSS* mengalami perkembangan dan penggunaanya

semakin kompleks untuk berbagi ilmu pengetahuan. Selain ilmu sosial, *SPSS* juga sering digunakan untuk ilmu ekonomi, psikologi, pertanian, teknologi, industri, dan lain-lain.

4. Rattle

Rattle adalah sebuah aplikasi penggalian data grafis yang dibangun di atas bahasa statistik pemahaman R. R tidak diperlukan untuk menggunakan Rattle. Rattle ini mudah digunakan, cepat untuk menyebarkan, dan memungkinkan kita untuk bekerja dengan cepat melalui pengolahan data, permodelan, dan evaluasi dari proyek *data mining*. Di sisi lain, R menyediakan bahasa yang sangat kuat untuk melakukan penggalian data jauh melampaui keterbatasan yang harus terkandung dalam antarmuka pengguna grafis dan pendekatan sebagai konsekuensinya *data mining*. (Feri Sulianta dan Dominikus Juju, 2010 : 103)

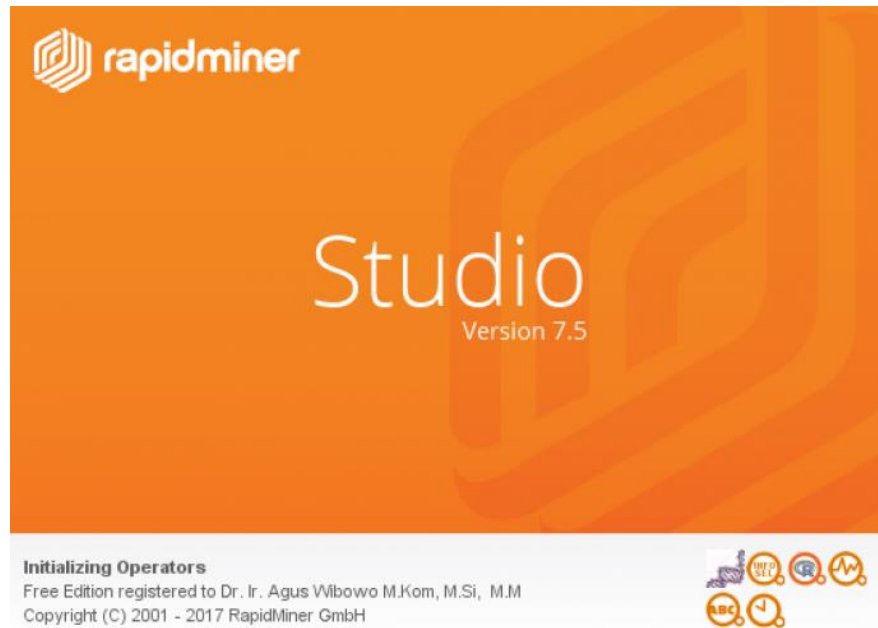
5. Orange Ailab

Orange Ailab adalah perangkat lunak open *source* yang memungkinkan pengguna yang tidak memahami sedikitpun tentang pemrograman dapat melakukan visualisasi dan analisis data. Fitur-fitur yang dimiliki di antaranya *scatterplots*, *bar charts*, *trees*, *dendrograms*, *networks* dan *heatmaps*. (Subekti Mujiasih 2015 : 191).

2.8 RAPIDMINER

Dalam penelitian ini menggunakan *tools weka* karena mampu menyelesaikan masalah-masalah data mining di dunia nyata serta dapat

menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik.



Gambar 2.3 Tampilan Utama RapidMiner

2.9 PROMOSI

Menurut Sigit Setyo, dkk dalam jurnalnya (2014) mengatakan bahwa Strategi promosi adalah perencanaan, implementasi, dan pengendalian komunikasi dari suatu organisasi kepada para konsumen dari sasaran lainnya.

Menurut Winardi dalam jurnal Sigit Setyo, dkk (2014:2) mengungkapkan bahwa Promosi adalah upaya-upaya suatu perusahaan untuk mempengaruhi para pembeli agar mereka mau membeli produk yang dihasilkan perusahaan atau menggunakan jasa yang disediakan.

Menurut Buchari Alma dalam jurnal Tulandi Riry, dkk (2015) mengatakan bahwa:

“Promosi adalah suatu bentuk komunikasi pemasaran yang merupakan aktivitas pemasaran yang berusaha menyebarkan informasi, mempengaruhi, dan mengingatkan pasar sasaran atas perusahaan dan produknya agar bersedia menerima, membeli dan loyal pada produk yang ditawarkan perusahaan yang bersangkutan.”

Dari beberapa pendapat diatas , maka dapat disimpulkan bahwa definisi tersebut mempunyai pengertian bahwa promosi meliputi semua alat yang terdapat dalam bauran promosi yang peranan utamanya adalah mengadakan komunikasi yang bersifat membujuk.

2.10 TINJAUAN STUDI

Penulis memulai penelitian ini dengan terlebih dahulu melakukan studi kepustakaan dari penelitian-penelitian dan sumber-sumber lain. Beberapa penelitian telah banyak dilakukan dengan menggunakan teknik *data mining* untuk menggali berbagai informasi dari sebuah *database*. Penelitian tersebut membahas tentang topik yang terkait dengan penelitian penulis, antara lain adalah penelitian mengenai algoritma yang akan digunakan penulis.

Tabel 2.1 Tinjauan Studi

No	Penulis dan Tahun	Masalah	Metode	Hasil
1	Laela Kurniawati, Aldo eka kusuma, Benti dewansyah 2019	Dalam Permasalahan yang saat ini di hadapi adalah dalam menentukan promosi sparepart, tidak mengetahui pola pelanggan dalam membeli sparepart mana yang di beli secara bersamaan	Algoritma apriori	Metode algoritma apriori dalam penentuan pola penjualan dapat di lakukan dengan melihat hasil kecenderungan konsumen dalam membeli sparepart, berdasarkan pola pembelian sparepart yang di olah dari data transaksi-transaksi sebelumnya, kombinasi 2 <i>itemset</i> menghasilkan sparepart yang sering di beli bersamaan dengan nilai <i>support</i> di atas nilai minimum <i>support</i>
2	Sri Rahayu Siregar 2014	Sering sekali pemesanan tiket pesawat yang di dinginkan konsumen tidak ada atau habis karena mereka tidak mengamati transaksi yang ada. Hal ini tentu sangat mengecewakan konsumen yang hendak memesan tiket pesawat , karena persediaan tiket pesawat di PT.Jumbo Travel tidak terkontrol dengan baik	Algoritma Apriori	Dapat di simpulkan bahwa data mining sangat berguna untuk mengetahui hubungan pola <i>frekuensi</i> penjualan tiket pesawat yang paling sering di beli oleh konsumen

3	Chavid syukri fatoni, Ema utami, Ferry wahyu wibowo 2018	Permasalahannya terdapat pada pelanggan yang bingung ketika membeli atau memilih produk yang di inginkan secara versamaan, bahkan ada pelanggan yang membeli bermacam – macam produk sekali transaksi	Apriori	Dengan menerapkan algoritma system memberikan rekoemndasi produk kepada pelanggan online store berdasarkan nilai <i>confidence</i> kombinasi produk yang dibeli secara periode waktu tertentu
4	Dewi Listriani, Anif hanifa setyaningrum 2016	Setiap harinya harus memenuhi kebutuhan konsumen dan di tuntutan untuk mengambil keputusan yang tepat dalam menentukan strategi penjualan	Algoritma apriori	Analisis pola yang di hasilkan di aplikasi ini dapat digunakan oleh pihak gramedia untuk pengaturan tata letak buku, Rekomendasi pencarian buku, pada gramedia online dan lain-lainnya yang mana setiap bulannya pola yang di hasilkan bisa berbeda-beda sesuai data transaksi yang di analisis
5	Khairul Umami 2015	Tidak mengetahui pasangan barang suku cadang yang di beli secara bersamaan , serta kesulitan dalam penyetokan barang karena banyaknya jenis dari suku cadang mobil.	Algoritma apriori	Dapat mengetahui pola penjualan yang dilakukan dengan melihat hasil dari kecenderungan konsumen berdasarkan 2 itemset. Dan perusahaan harus mengatur letak posisi dari sparepart secara berdekatan untuk memudahkan keberadaan sparepart.

Berdasarkan penelitian terkait diatas, dapat dijelaskan bahwa metode Algoritma Apriori dapat menemukan keterhubungan antar item dengan melihat nilai minimum *support* dan nilai minimum *confidence*, yang hasilnya nanti dapat merekomendasikan kepada perusahaan agar mengetahui barang sparepart yang kemungkinan besar di beli oleh *customer* secara bersamaan, rekomendasi tersebut didapatkan berdasarkan hasil dari penelitian yang sedang dilakukan. Pada penelitian ini, penulis menggunakan metode Algoritma Apriori untuk menemukan pola transaksi penjualan pada sentral Yamaha murni.

