

**KAJIAN *DATA MINING* FITUR *VIDEO GAME* PADA *STEAM*
MENGUNAKAN ALGORITMA APRIORI
DAN FP-GROWTH**

TESIS

Disusun sebagai salah satu syarat untuk
memperoleh gelar Magister Komputer
dari Sekolah Tinggi Manajemen Informatika dan Komputer LIKMI

Oleh:

YAKASETAVIRA MULYADI WIRYA

NPM : 2017210041



**PROGRAM STUDI PASCASARJANA
MAGISTER SISTEM INFORMASI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA & KOMPUTER LIKMI
BANDUNG
2019**

**KAJIAN *DATA MINING* FITUR *VIDEO GAME* PADA *STEAM*
MENGUNAKAN ALGORITMA APRIORI
DAN FP-GROWTH**

Oleh:

YAKASETAVIRA MULYADI WIRYA

NPM : 2017210041

**Bandung, 11 Oktober 2018
Menyetujui,**

**Dr. Hery Heryanto, S.Kom., M.Kom
Pembimbing**

**PROGRAM STUDI PASCASARJANA
MAGISTER SISTEM INFORMASI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA & KOMPUTER LIKMI
BANDUNG
2019**

“Sebab itu janganlah kamu kuatir akan hari besok, karena hari besok mempunyai kesusahannya sendiri. Kesusahan sehari cukuplah untuk sehari.”

Matius 6:34

Tesis ini dipersembahkan untuk keluarga tercinta:

Almarhum Opa: Eppy Adhistanavira

Ibu: Diana Devi

Ayah: Rudy Saputra

Bibi: Susi Susilawati

Bibi: Lily Rimsa

Oma: Werry Ong

ABSTRAK

KAJIAN DATA MINING FITUR VIDEO GAME PADA STEAM MENGUNAKAN ALGORITMA APRIORI DAN FP-GROWTH

Oleh:

Yakasetavira Mulyadi Wirya
NPM: 2017210041

Berdasarkan *Global Games Market Report* oleh Newzoo Games, *Video game* adalah komoditas penting pada beberapa negara maju yang memiliki genre dan fitur tertentu yang unik pada masing-masing judul. Keunikan pada masing-masing *video game* memberikan pola pasangan tertentu yang dapat digunakan untuk mencari tahu tren yang ada pada periode waktu tertentu dengan menggunakan salah satu teknik *data mining*, yaitu *association rule*.

Association rule yang menjadi teknik yang dipilih, memiliki beberapa algoritma yang digunakan, dua diantaranya yaitu algoritma Apriori dan FP-Growth. Meskipun kedua algoritma memiliki tujuan yang sama dalam mencari pola keterkaitan antar fitur dan genre, masing-masing algoritma memiliki cara kerja, performa, serta efisiensi waktu yang berbeda. Perbedaan tersebut yang akan digunakan penulis untuk mencari tahu algoritma yang sebaiknya digunakan pada penelitian selanjutnya.

Salah satu penyedia layanan distribusi *video game* terbesar adalah Steam. Steam menggunakan situs web serta platform untuk mempromosikan dan memberikan informasi mengenai *video game* yang hendak dibeli dan/atau dimainkan. Kumpulan judul daripada *video game* yang diambil merupakan sampel acak yang tercatat sepanjang tahun 2017. Setiap judul yang tercatat akan melalui *data preprocessing* sebelum melalui proses *data mining*. Hasil yang diperoleh nantinya akan digunakan penulis sebagai acuan untuk mengetahui spesifikasi pasangan fitur dan genre *video game* yang mendominasi Steam pada tahun 2017.

Pada bagian akhir dari penelitian ini, dapat disimpulkan bahwa *video game* dengan genre MOBA (*Multiplayer Online Battle Arena*) adalah genre yang mendominasi Steam, serta pemilihan algoritma FP-Growth sebagai rekomendasi dibandingkan dengan algoritma Apriori karena efisiensi dan efektivitas yang lebih baik.

Kata kunci: *Association rule*, Steam, *Video game*, Fitur, Genre, Algoritma Apriori, Algoritma FP-Growth

ABSTRACT

DATA MINING RESEARCH ABOUT STEAM VIDEO GAME FEATURES USING APRIORI AND FP-GROWTH ALGORITHM

By:

**Yakasetavira Mulyadi Wirya
NPM: 2017210041**

According to Global Games Market Report from Newzoo Games, Video game is an important commodity in developed countries which has different genres and features to each video game. The differences of each video game give a certain pairing pattern which can be used to give insight about trend that is occurring at certain period of time by using one of the data mining techniques, the association rule.

Association rule, which is the technique of selection, has several algorithms, which two of them are Apriori algorithm and FP-Growth algorithm. Both algorithm have different workflow, performance, and efficiency, and having the similar purpose to discover pairing pattern between genres and features. The differences will become a recommendation about which algorithm should be used for the next research.

One of the largest digital video game distribution service provider is Steam. Steam uses website and platform to promote and give information about video games which can be bought or played. Video game records taken are random samples from video games that were released in 2017. Each records will be preprocessed before go to data mining process. The data mining rules will be a reference to discover pairings of genres and features that saturated Steam in 2017.

Finally, it is concluded that MOBA (Multiplayer Online Battle Arena) is a particular genre that dominates Steam in 2017. On the other hand, FP-Growth algorithm is recommended for having a better efficiency and effectivity than Apriori algorithm.

Keywords: Association rule, Steam, Video game, Feature, Genre, Apriori algorithm, FP-Growth algorithm

KATA PENGANTAR

Puji dan syukur senantiasa penulis panjatkan kepada Tuhan Yang Maha Esa, atas segala karunia-Nya, sehingga tesis dengan judul “Kajian *Data Mining* Fitur *Video Game* Pada *Steam* Menggunakan Algoritma Apriori Dan FP-Growth” dapat diselesaikan sesuai dengan rencana.

Ucapan terima kasih yang setinggi-tingginya penulis sampaikan kepada semua pihak yang telah memberikan dukungan selama penyelesaian tesis ini, antara lain:

1. Bapak Dr. Hery Heryanto, S. Kom., M. Kom., sebagai pembimbing yang selalu memberikan bimbingan dan arahan selama pembuatan tesis ini sehingga dapat penulis selesaikan dengan baik.
2. Segenap dosen Program Pascasarjana Sekolah Tinggi Manajemen Informatika dan Komputer LIKMI atas ilmu yang diberikan selama penulis menempuh perkuliahan.
3. Staf administrasi dan rekan-rekan satu angkatan di Program Magister Sistem Informasi yang selalu memberikan dorongan dan semangat pada penulis.
4. Almarhum opa tercinta Eppy Adhistanavira yang telah memberikan bekal semangat dan keyakinan akan pertolongan-Nya.
5. Rekan kerja Adhi Darmawan Sutjiadi yang telah menyediakan waktunya untuk membantu penulis mencari jurnal yang dibutuhkan serta memberikan dan/atau meminjamkan sumber daya bagi penulis untuk digunakan.
6. Anggota-anggota keluarga tercinta: ibu Diana Devi, ayah Rudy Saputra, bibi Susi Susilawati, bibi Lily Rimsa, dan anggota keluarga lainnya. Terima kasih atas dukungan moral maupun material yang telah diberikan.

Seluruh pihak yang tidak dapat disebutkan Namanya satu per satu, semoga Tuhan Yang Maha Esa membalasnya dengan keberkahan. Amin.

DAFTAR ISI

LEMBAR PENGESAHAN	i
ABSTRAK.....	ii
ABSTRACT	iii
KATA PENGANTAR	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	viii
DAFTAR TABEL	x
DAFTAR RUMUS.....	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Metode Penelitian	3
1.6 Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Data Mining	6
2.1.1 Karakteristik Data Mining.....	6
2.1.2 Data Preprocessing	8
2.1.3 Cross-Industry Standard Process for Data Mining (CRISP-DM).....	10
2.1.4 Tugas-tugas Data Mining.....	12
2.2 Association Rule	14

2.2.1	Apriori.....	14
2.2.2	FP-Growth.....	19
2.2.3	Perbandingan Algoritma	21
2.3	Game.....	23
2.3.1	Evolusi Video Game	23
2.3.2	Jenis-jenis Video Game.....	25
2.4	Distribusi Digital	28
2.5	Steam	31
2.5.1	Fitur-fitur pada Steam	31
2.5.2	Informasi Video Game pada Steam.....	32
2.6	Weka	35
2.7	Penelitian Terdahulu	37
BAB III OBJEK DAN METODOLOGI PENELITIAN		39
3.1	Profil Steam.....	39
3.2	Metodologi Penelitian.....	47
3.3	Alat Penelitian	50
BAB IV HASIL DAN PEMBAHASAN		52
4.1	Business Understanding Phase.....	52
4.2	Data Understanding Phase.....	54
4.2.1	Informasi Dasar.....	54
4.2.2	Genre	56
4.2.3	Tipe Keterlibatan Pemain	59
4.2.4	Fitur.....	61
4.3	Data Preparation Phase.....	64

4.4	Modelling Phase.....	68
4.4.1	Penentuan Nilai Minimum <i>Support</i> dan <i>Confidence</i>	68
4.4.2	Metode 1 – Algoritma Apriori	70
4.4.3	Metode 2 – Algoritma FP-Growth	77
4.5	Evaluation Phase	84
4.6	Perbandingan Algoritma.....	89
BAB V KESIMPULAN DAN SARAN		93
5.1	Kesimpulan.....	93
5.2	Saran.....	94
DAFTAR PUSTAKA		95
LAMPIRAN.....		98

DAFTAR GAMBAR

Gambar 2.1 Tabel Transaksi Pembelian Sayuran	9
Gambar 2.2 Data Tabular daripada Tabel Transaksi Pembelian Sayuran	10
Gambar 2.3 Enam Fase CRISP-DM	11
Gambar 2.4 Tabel Transaksi.....	16
Gambar 2.5 Langkah-langkah Algoritma Apriori.....	18
Gambar 2.6 Tabel Perbedaan Algoritma FIM	21
Gambar 2.7 Sistem Distribusi	29
Gambar 2.8 Antarmuka Awal Weka 3.8.2.....	36
Gambar 3.1 DFD Steam	41
Gambar 3.2 Grafik <i>Video Game Steam</i>	42
Gambar 3.3 Metodologi Penelitian.....	47
Gambar 4.1 Informasi Dasar <i>Cuphead</i>	54
Gambar 4.2 Genre <i>Cuphead</i>	56
Gambar 4.3 Tipe Keterlibatan Pemain <i>Cuphead</i>	59
Gambar 4.4 Fitur <i>Cuphead</i>	61
Gambar 4.5 Perbedaan Genre dengan <i>Tag</i>	64
Gambar 4.6 Data Sebelum Mengalami <i>Data Preprocessing</i>	65
Gambar 4.7 Data Setelah Mengalami <i>Data Preprocessing</i>	65
Gambar 4.8 Kesalahan Format pada Weka	66
Gambar 4.9 Weka Explorer.....	67
Gambar 4.10 Tab <i>Associate</i>	68
Gambar 4.11 Pengaturan Algoritma Apriori.....	70
Gambar 4.12 <i>Log</i> Eksekusi Algoritma	71
Gambar 4.13 <i>Rule</i> yang Dihasilkan	72
Gambar 4.14 Pengaturan Algoritma FP-Growth.....	77
Gambar 4.15 <i>Log</i> Eksekusi Algoritma	78
Gambar 4.16 <i>Rule</i> yang Dihasilkan	79
Gambar 4.17 Perbandingan Waktu Eksekusi Kedua Algoritma	89

Gambar 4.18 Log Eksekusi Algoritma Apriori	90
Gambar 4.19 Log Eksekusi Algoritma FP-Growth	91

DAFTAR TABEL

Tabel 2.1 Tabel Penelitian dan Jurnal	37
Tabel 3.1 Tabel Parameter <i>Video Game</i> Steam.....	43
Tabel 3.2 Tabel Parameter <i>Video Game</i> Steam, Setelah <i>Data Preprocessing</i>	45
Tabel 4.1 Tabel Kategori <i>Itemset</i>	54
Tabel 4.2 Tabel Jumlah Kemunculan Genre	59
Tabel 4.3 Tabel Jumlah Kemunculan Tipe Keterlibatan Pemain.....	61
Tabel 4.4 Tabel Jumlah Kemunculan Fitur	63
Tabel 4.5 Tabel Jumlah Kemunculan Atribut yang Digunakan.....	68
Tabel 4.6 Tabel Pengaturan Nilai Parameter Algoritma Apriori.....	71
Tabel 4.7 <i>Rule</i> yang Dihasilkan	72
Tabel 4.8 Tabel Pengaturan Nilai Parameter Algoritma FP-Growth.....	78
Tabel 4.9 <i>Rule</i> yang Dihasilkan	79
Tabel 4.10 <i>Rule</i> dengan Nilai Confidence 100%.....	84
Tabel 4.11 <i>Rule</i> dengan Nilai Support melebihi 3%	86
Tabel 4.12 Tabel Rekomendasi Komposisi <i>Video Game</i>	88
Tabel 4.13 Tabel Perbandingan Waktu	91

DAFTAR RUMUS

Rumus 2.1 Nilai <i>Support</i> Satu <i>Item</i>	17
Rumus 2.2 Nilai <i>Support</i> Dua <i>Item</i>	17
Rumus 2.3 Nilai <i>Confidence</i> Dua <i>Item</i>	17
Rumus 2.4 Batasan Nilai <i>Support</i> Minimum	17
Rumus 2.5 Batasan Nilai <i>Confidence</i> Minimum	17

DAFTAR LAMPIRAN

Lampiran 1 Halaman Utama Steam.....	103
Lampiran 2 Halaman Penawaran Steam	103
Lampiran 3 Halaman Informasi <i>Video Game</i> pada Steam	104
Lampiran 4 Halaman Komunitas pada Steam	104
Lampiran 5 Halaman Tentang pada Steam	105
Lampiran 6 Halaman Bantuan pada Steam.....	105

BAB I

PENDAHULUAN

1.1 Latar Belakang

Video game adalah komoditas yang penting pada beberapa negara maju, hal tersebut disebabkan karena tingginya permintaan dan profit yang dihasilkan. Tren pada *video game* perlu diprediksi dengan akurat, karena hasil prediksi yang akurat sangat penting untuk menentukan tren *video game* di masa mendatang (Groeneveld, 2016). Dalam konteks *video game*, tren dapat didefinisikan sebagai bentuk perilaku kelompok sosial, khususnya dalam komunitas pemain *video game*, yang mana sekelompok orang secara antusias memainkan genre atau fitur tertentu yang dimiliki sebuah *video game* dalam jangka waktu tertentu. Tren *video game* muncul sebagai sebuah persepsi pemain yang melihat bahwa fitur tertentu daripada sebuah *video game* dikatakan seru, mengasyikkan, atau menantang (Griffith, 2013).

Tren tersebut dapat dicari menggunakan salah satu teknik *data mining*, yaitu asosiasi. Beberapa metode yang menjadi pilihan daripada teknik asosiasi adalah algoritma Apriori (Hegland, 2005) dan algoritma FP-Growth (Han, 2004). Meskipun algoritma Apriori tidak memiliki kendala secara logika maupun urutan, yang menjadi masalah adalah efisiensi waktu yang digunakan. Algoritma Apriori dapat menyebabkan *bottleneck* dan berjalan sangat lambat, karena basis data atau tabel harus diperiksa pada setiap tingkat. Jumlah pemeriksaan adalah $(n + 1)$, dimana n adalah panjang dari pola terpanjang (INSOFE, 2014).

Alasan mengapa algoritma Apriori dipilih adalah untuk membandingkan performa metode dibandingkan dengan algoritma turunannya, yaitu Algoritma FP-Growth. Meskipun secara teori, kedua algoritma memiliki metode yang berbeda dalam melakukan *association rule*, terdapat beberapa perbedaan dan kesamaan lainnya yang hendak ditelusuri dalam penelitian ini. Dewasa ini, frekuensi penggunaan kedua algoritma relatif banyak, khususnya pada perusahaan yang mulai menerapkan *machine learning* pada perusahaannya, atau pada *search engine* pada internet dan aplikasi berbasis teknologi lainnya.

Pada era globalisasi saat ini, penggunaan internet dan aplikasi berbasis teknologi informasi, sudah menjadi kebiasaan yang melekat dalam kehidupan masyarakat. Seiring berkembangnya populasi masyarakat sebagai pengguna, semakin berkembang pula kecanggihan teknologi pada aplikasi yang digunakan, khususnya video game (Kelly, 2014). Dewasa ini, para pemain *video game* menggunakan Steam sebagai media bermain secara *online* maupun *offline*. Steam adalah platform digital terdistribusi yang dikembangkan oleh Valve Corporation, yang memiliki kelebihan untuk membantu penggunanya melakukan instalasi *video game* dan pembaharuan *game* dan platform secara otomatis dan berkala. Pada tahun 2017, jumlah *video game* yang dikembangkan para *developer* dan *publisher* melebihi angka 7000.

Pada penelitian ini, beberapa metode asosiasi akan diterapkan untuk melihat perbedaan antara dua algoritma: algoritma Apriori dan algoritma FP-Growth, dan pada akhirnya memunculkan pengetahuan yang didapat atas hasil metode tersebut, dimana data yang digunakan bersumber dari Steam. Melihat jumlah data *video game* yang besar pada Steam dan sifatnya yang *up-to-date*, menjadikan data-data tersebut ideal untuk diambil pengetahuan dan intisarinya untuk mengetahui tren *video game* pada tahun 2017. Berdasarkan penjelasan yang dipaparkan, penulis memilih judul tesis: **“KAJIAN KECENDERUNGAN FITUR VIDEO GAME PADA STEAM MENGGUNAKAN ALGORITMA APRIORI DAN FP-GROWTH”**.

1.2 Rumusan Masalah

Berdasarkan uraian diatas, maka terdapat beberapa masalah yang diidentifikasi dalam penelitian ini, yaitu:

1. Pengetahuan apa yang diperoleh dengan mengaplikasikan teknik *data mining* asosiasi pada kumpulan *record* yang diperoleh dari Steam?
2. Perbedaan apa yang muncul dalam pengaplikasian beberapa metode berbeda dalam teknik asosiasi pada kumpulan *record* Steam?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk mengetahui kecenderungan kombinasi dan tren *genre* yang muncul pada tahun 2017.
2. Untuk mengetahui hasil yang diperoleh menggunakan beberapa metode dalam teknik asosiasi.

1.4 Batasan Masalah

Penelitian ini dibatasi dengan beberapa kondisi, yaitu:

1. Data yang diteliti adalah data mentah *video game* yang terdaftar dalam *Steam*, jatuh dalam kategori *Top Sellers*, dan hanya berjalan pada sistem operasi Windows. Pemilihan kategori *Top Sellers* didasarkan pada kelengkapan informasi yang diberikan pada judul *video game* yang bersangkutan.
2. Data yang diambil merupakan sampel acak terhitung dari tanggal 1 Januari 2017 hingga 31 Desember 2017.
3. Jumlah data yang diambil setidaknya mencapai 2000 *record*.
4. Algoritma *association rule* yang digunakan adalah Apriori dan FP-Growth. Kedua metode tersebut akan diimplementasikan menggunakan Weka.

1.5 Metode Penelitian

Penelitian ini menggunakan metode deskriptif. Metode deskriptif adalah sebuah metode penelitian yang dapat diartikan sebagai prosedur pemecahan masalah yang diselidiki dengan menggambarkan keadaan subjek atau objek dalam penelitian, dapat berupa orang, lembaga, masyarakat dan yang lainnya pada saat ini berdasarkan fakta-fakta yang tampak atau apa adanya.

Metode deskriptif merupakan suatu metode dalam meneliti status sekelompok manusia, suatu objek, suatu set kondisi, suatu sistem pemikiran ataupun suatu kelas peristiwa pada masa sekarang. Tujuan dari penelitian deskriptif ini adalah untuk membuat deskripsi, gambaran, atau lukisan secara sistematis, faktual, dan akurat mengenai fakta-fakta, sifat-sifat serta hubungan antar fenomena yang diselidiki.

Kesimpulan yang dapat diambil dari definisi metode deskriptif adalah penelitian deskriptif merupakan penelitian yang berusaha mendeskripsikan suatu gejala, peristiwa yang terjadi pada saat sekarang atau masalah faktual. Pada penelitian ini, analisis dilakukan hanya sampai taraf deskripsi yaitu menganalisis dan menyajikan fakta secara sistematis sehingga dapat lebih mudah untuk dipahami dan disimpulkan. Penelitian deskriptif bertujuan untuk mendapatkan gambaran yang benar mengenai subjek yang diteliti. Pengolahan data akan dilakukan dengan menganalisis persentase dan menganalisis tren.

1.6 Sistematika Penulisan

Penulisan tesis ini menggunakan sistematika penulisan tertentu yang bertujuan untuk memberikan gambaran umum mengenai penelitian yang dilakukan. Sistem penulisan yang digunakan adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini dijelaskan tentang latar belakang tentang perkembangan usaha ritel dan peluangnya ke depan, identifikasi masalah, tujuan penelitian dan pembatasan masalah serta metodologi penelitian yang digunakan.

BAB II TINJAUAN PUSTAKA

Pada bab ini dijelaskan mengenai dasar-dasar teori, rujukan dan teknik *data mining* beserta algoritma yang digunakan

BAB III METODOLOGI PENELITIAN

Pada bab ini dijelaskan mengenai analisis dan persiapan data dengan mengambil data secara langsung pada situs resmi Steam, lalu melakukan *data preprocessing* sebelum nantinya menjadi kumpulan data yang ideal untuk dilakukan teknik asosiasi.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini dijelaskan mengenai implementasi metode-metode *data mining* terhadap kumpulan data yang sudah diperoleh, dan memberitahu pengetahuan atau informasi yang didapat atas hasil dari metode-metode tersebut, berikut dengan perbandingannya.

BAB V PENUTUP

Pada bab ini disimpulkan hasil dari penelitian, penjelasan mengenai hasil yang diperoleh dari *data mining*, dan saran bagi penelitian sejenis pada masa yang akan datang.

BAB II

TINJAUAN PUSTAKA

2.1 Data Mining

Jiawei Han memberikan definisi pada bukunya yang berjudul "*Data Mining : Concept and Techniques Third Edition*", menjelaskan bahwa : "*Data mining is the process of discovering interesting patterns and knowledge from large amount of data. The data sources can include databases, data warehouses, the Web, other information repositories, or data that are streamed into the system dynamically.*" (Han, 2012:8)

Daniel T. Larose memberikan definisi pada bukunya yang berjudul "*Discovering Knowledge in Data*", menjelaskan bahwa : "*Data mining is the process of discovering meaningful new correlations, patterns and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and mathematical techniques.*" (Larose, 2005:21)

Ian H. Witten memberikan definisi pada bukunya yang berjudul "*Data Mining – Practical Machine Learning Tools and Techniques Second Edition*" (Witten, 2005:5), menjelaskan bahwa :

Data mining is defined as the process of discovering patterns in data. The process must be automatic or (more usually) semiautomatic. The patterns discovered must be meaningful in that they lead to some advantage, usually an economic advantage. The data is invariably present in substantial quantity.

Berdasarkan definisi, disimpulkan bahwa *data mining* adalah proses mencari pengetahuan daripada sekumpulan data dengan jumlah besar, dimana data-data tersebut dapat didapatkan dari *database*, arsip di Web, data transaksi, dan lain sebagainya.

2.1.1 Karakteristik Data Mining

Ian H. Witten menjabarkan beberapa karakteristik yang muncul pada setiap proses *data mining*:

1. Ekstraksi Informasi

Data mining bertujuan untuk mengekstraksi informasi yang implisit dan sebelumnya tidak diketahui dari sebuah data. Gagasan daripada *data mining* adalah untuk mengembangkan

program komputer yang dapat menganalisis basis data secara otomatis, mencari kesamaan dan pola. Apabila terdapat pola-pola tertentu yang kuat, maka pola tersebut akan diambil untuk membuat prediksi akurat terdapat data di masa mendatang. (Witten, 2005:xxiii)

2. Tren

Penggunaan *data mining* dikelilingi oleh banyak tekanan terhadap tren dalam hal teknis, dan terkadang pres yang populer. Laporan-laporan yang bersifat membesar-besarkan yang selama ini dirahasiakan, dapat dikuak dengan mengatur algoritma pembelajaran pada lautan data. (Witten, 2005:xxiii)

3. Penyelesaian Masalah

Data mining ditujukan untuk menyelesaikan masalah yang muncul dengan cara menganalisis data yang sudah tersedia dalam basis data. Diasumsikan bahwa terdapat permasalahan mengenai ketidakpastian kesetiaan konsumen pada pasar. Basis data mengenai pilihan konsumen berikut dengan profil mereka, adalah kunci dari masalah tersebut. Pola-pola perilaku konsumen dapat dianalisis untuk menentukan karakteristik konsumen yang kemungkinan besar akan berganti produk atau tetap setia. Apabila karakteristik tersebut telah ditemukan, maka mereka dapat menentukan konsumen mana yang memiliki kemungkinan untuk berpindah tempat. Kelompok konsumen ini, nantinya akan diberi perlakuan khusus (seperti diskon, paket, dan lain sebagainya), dibandingkan dengan memberi seluruh konsumen perlakuan khusus. (Witten, 2005:5)

4. Praktis

Data mining adalah topik yang bersifat praktis, sehingga dalam pembelajarannya membutuhkan lebih banyak praktik dibandingkan teori. Topik *data mining* lebih berfokus pada penemuan dan penjelasan pola terstruktur daripada data sebagai sebuah alat, untuk membantu menjelaskan dan membuat prediksi dari data tersebut.

5. Pengambilan Keputusan

Data mining adalah sebuah alat yang digunakan orang-orang untuk menerima hasil bersamaan dengan pengetahuan lainnya, lalu menggunakannya untuk mengambil langkah apa di masa mendatang.

6. Pola

Data mining secara keseluruhan membahas hal mengenai pola pada data. Seperti pada *text mining*, yang bertugas untuk mencari pola-pola pada teks yang nantinya dianalisis untuk mengekstraksi informasi penting untuk tujuan tertentu.

2.1.2 Data Preprocessing

Data preprocessing adalah sebuah langkah penting pada proses ditemukannya pengetahuan, hal tersebut disebabkan pilihan-pilihan yang baik harus didasarkan pada data-data yang baik pula. Di dunia nyata, data memiliki kecenderungan bersifat kotor, tidak lengkap, dan tidak konsisten. *Data preprocessing* dilakukan untuk meningkatkan kualitas data, yang mana nantinya membantu meningkatkan proses akurasi dan efisiensi dari *data mining* (Han, 2012:87). Secara garis besar, data preprocessing dibagi menjadi empat tahap, yaitu : data cleaning, data integration, data reduction, dan data transformation and data discretization (Han, 2012:120).

1. Data Cleaning

Data cleaning adalah langkah yang diambil dengan tujuan mengisi nilai-nilai yang kosong, menghilangkan *noise* dan mengidentifikasi *outlier*, dan memperbaiki inkonsistensi pada data. *Data cleaning* biasanya dilakukan secara iteratif dan dibagi menjadi dua proses yang terdiri atas *discrepancy detection* dan *data transformation*.

2. Data Integration

Data integration adalah langkah penggabungan data dari beberapa sumber untuk membentuk sumber data yang koheren. Hasil-hasil dari keberagaman semantik, *metadata*, analisis korelasi, deteksi duplikasi *tuple*, dan deteksi konflik data membantu melancarkan proses integrasi data.

3. Data Reduction

Data reduction adalah teknik yang digunakan untuk menghasilkan representasi data yang lebih sedikit tanpa mengurangi informasi yang terkandung dalam data tersebut. Metode-metode yang dicapai dalam *data reduction* adalah *dimensionality reduction*, *numerosity reduction*, dan *data compression*.

4. Data Transformation

Data transformation adalah langkah mengkonversi data menjadi bentuk yang layak digunakan dalam *data mining*. Sebagai contoh, dalam proses normalisasi, atribut data dimasukkan ke dalam skala 0.0 hingga 1.0. Contoh lain adalah *data discretization* dan *concept hierarchy generation*.

5. Data Discretization

Data discretization adalah langkah mengubah data numerik dengan memetakan nilai-nilai yang tersedia menjadi interval atau label. Metode ini secara tidak langsung menghasilkan konsep hirarki pada data, yang memungkinkan proses *mining* dilakukan pada berbagai tingkatan.

Transaction	Items Purchased
1	Broccoli, green peppers, corn
2	Asparagus, squash, corn
3	Corn, tomatoes, beans, squash
4	Green peppers, corn, tomatoes, beans
5	Beans, asparagus, broccoli
6	Squash, asparagus, beans, tomatoes
7	Tomatoes, corn
8	Broccoli, tomatoes, green peppers
9	Squash, asparagus, beans
10	Beans, corn
11	Green peppers, broccoli, beans, squash
12	Asparagus, beans, squash
13	Squash, corn, asparagus, beans
14	Corn, green peppers, tomatoes, beans, broccoli

Gambar 2.1
Tabel Transaksi Pembelian Sayuran
(Larose, 2005:182)

Gambar 2.1 menunjukkan nomor transaksi beserta dengan *tuple* yang mewakili *item* yang dibeli dalam sebuah transaksi, dengan mengabaikan jumlah *item* yang dibeli. Untuk

dapat dibaca dengan lebih mudah dan diproses menggunakan aplikasi *data mining*, maka tabel diatas harus diubah menjadi bentuk tabular, yaitu perubahan *tuple* menjadi angka biner.

Transaction	Asparagus	Beans	Broccoli	Corn	Green Peppers	Squash	Tomatoes
1	0	0	1	1	1	0	0
2	1	0	0	1	0	1	0
3	0	1	0	1	0	1	1
4	0	1	0	1	1	0	1
5	1	1	1	0	0	0	0
6	1	1	0	0	0	1	1
7	0	0	0	1	0	0	1
8	0	0	1	0	1	0	1
9	1	1	0	0	0	1	0
10	0	1	0	1	0	0	0
11	0	1	1	0	1	1	0
12	1	1	0	0	0	1	0
13	1	1	0	1	0	1	0
14	0	1	1	1	1	0	1

Gambar 2.2
Data Tabular daripada Tabel Transaksi Pembelian Sayuran
(Larose, 2005:183)

Pada beberapa kasus, angka biner 1 dan 0 dapat diubah menjadi TRUE dan FALSE. Penentuan nilai *support* dan *confidence* daripada kemunculan *item* akan dihitung berdasarkan kemunculan 1 atau TRUE daripada tabel data tabular.

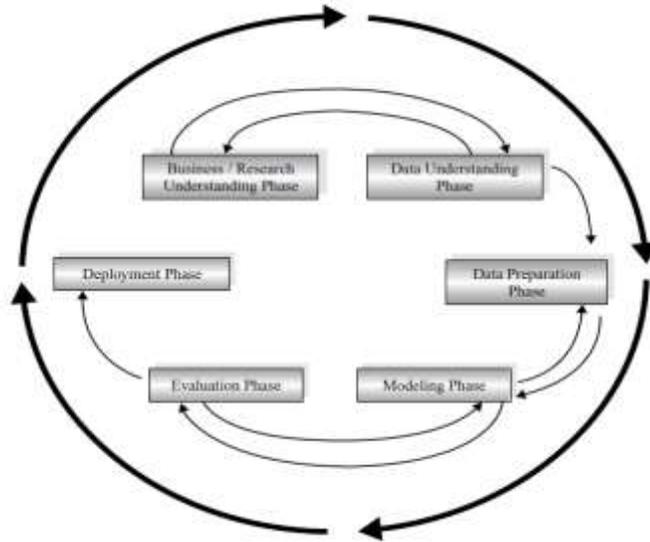
2.1.3 Cross-Industry Standard Process for Data Mining (CRISP-DM)

Daniel T. Larose memberikan definisi pada bukunya yang berjudul "*Discovering Knowledge in Data*", menjelaskan bahwa : "*The particular standard process used is the CRISP-DM framework: the Cross-Industry Standard Process for Data Mining. CRISP-DM demands that data mining be seen as an entire process, from communication to business problem through data collection and management, data preprocessing, model building, model evaluation, and finally, model deployment.*" (Larose, 2005)

CRISP-DM dikembangkan pada akhir tahun 1996 oleh tiga pasar *data mining*, yaitu DaimlerChrysler, SPSS, dan NCR. DaimlerChrysler (sekarang Daimler AG) adalah organisasi komersil dan industrial yang sangat maju dalam mengaplikasikan *data mining* pada operasi bisnisnya. SPSS (dulu dikenal sebagai ISL), adalah penyedia jasa *data mining* sejak tahun

1990 dan telah meluncurkan Clementine sebagai *data mining workbench* pertama pada tahun 1994.

NCR adalah sebuah organisasi yang terdiri atas konsultan-konsultan *data mining* dan para spesialis teknologi yang menyediakan jasa untuk memenuhi kebutuhan para klien nya. (Chapman, et al., 2000:1)



Gambar 2.3
Enam Fase CRISP-DM
(Larose, 2005:6)

Berdasarkan gambar 2.3, terdapat enam fase daripada CRISP-DM:

1. *Business understanding phase*

Fase pertama dari CRISP-DM adalah mengerti tujuan dan kebutuhan projek secara jelas dalam konteks bisnis, mengubah kebutuhan bisnis tersebut menjadi formulasi *data mining*, dan menyiapkan strategi untuk mencapai tujuan tersebut.

2. *Data understanding phase*

Fase kedua meliputi pengumpulan data, analisis informasi dan kualitas data, dan pemilihan subset yang memiliki pola-pola tertentu.

3. *Data preparation phase*

Fase ketiga meliputi persiapan data mentah, pemilihan variable, dan melakukan transformasi data untuk siap digunakan oleh *modelling tool*.

4. *Modelling phase*

Fase keempat meliputi pemilihan teknik, dimana memungkinkan untuk memilih lebih dari satu, lalu sesuaikan pengaturan untuk menunjukkan hasil yang optimal. Apabila diperlukan, kembali ke fase sebelumnya (*data preparation phase*) untuk memulai kembali menggunakan teknik *data mining* yang lain.

5. *Evaluation phase*

Fase kelima meliputi mengevaluasi beberapa hasil yang ditampilkan fase-fase pemodelan untuk menguji kualitas dan efektivitas sebelum digunakan di lapangan, menguji kedekatan dengan tujuan proyek, mengkaji ulang apakah terdapat masalah riset maupun bisnis yang belum terselesaikan, dan memilih penggunaan dari hasil *data mining*.

6. *Deployment phase*

Fase terakhir meliputi penggunaan model, pembuatan laporan, implementasi proses *data mining* secara parallel pada departemen lainnya.

2.1.4 Tugas-tugas Data Mining

Secara umum, terdapat enam tugas yang dilakukan dalam *data mining*. Tugas-tugas tersebut adalah *description*, *estimation*, *prediction*, *classification*, *clustering*, dan *association*.

(Larose, 2005:11)

1. *Description*

Description adalah tugas mendeskripsikan pola-pola dan tren-tren yang memberikan penjelasan terhadap pola maupun tren tersebut. Penjelasan tersebut harus dijabarkan secara transparan, jelas, dan tidak berbelit-belit.

2. *Estimation*

Estimation memiliki kesamaan dengan *classification*, namun target variable yang digunakan bersifat numerik dibandingkan kategori. Model-model yang dibangun

menggunakan *record* penuh, yang mana menyediakan nilai target variable dan nilai predictornya. Lalu, untuk observasi yang baru, perkiraan/estimasi nilai dari target variable dibuat, berdasarkan nilai dari predictor.

3. *Prediction*

Prediction memiliki kesamaan dengan *classification* dan *estimation*, namun pada *prediction*, sesuai dengan artinya, hasilnya baru bisa terlihat di masa depan.

4. *Classification*

Classification adalah tugas mengklasifikasikan sekumpulan data berdasarkan karakteristik lainnya yang terdapat pada data tersebut.

5. *Clustering*

Clustering mengacu pada pengelompokan *record*, hasil pengamatan, atau kasus-kasus lainnya menjadi beberapa kelas daripada objek-objek yang sama, Sebuah *cluster* adalah sekumpulan *record* yang memiliki kesamaan antara satu data dengan lainnya, dan memiliki perbedaan dengan *record* pada *cluster* lainnya. *Clustering* bertujuan untuk mengelompokkan data menjadi *cluster* tertentu.

6. *Association*

Association adalah tugas menemukan atribut yang bersifat “saling berhubungan” atau “muncul bersamaan”. Pada dunia bisnis, istilah ini lebih dikenal dengan analisis keranjang belanja. Tugas dari *association* adalah membuka aturan-aturan untuk menemukan hubungan antara dua atau lebih atribut. *Association* mengacu pada prinsip “jika antisenden, maka konsekuen.”

2.2 Association Rule

Daniel T. Larose memberikan definisi pada bukunya yang berjudul “Discovering Knowledge in Data” (Larose, 2005:17), menjelaskan bahwa :

The association task for data mining is the job of finding which attributes ‘go together.’ Most prevalent in the business world, where it is known as affinity analysis or market basket analysis, the task of association seeks to uncover rules for quantifying the relationship between two or more attributes.

Ian H. Witten memberikan definisi pada bukunya yang berjudul “Data Mining – Practical Machine Learning Tools and Techniques Second Edition”, menjelaskan bahwa : “*It is equally possible to disregard the classification and just look for any rules that strongly associate different attribute values. These are called association rules.*” (Witten, 2005:12)

Jiawei Han memberikan definisi pada bukunya yang berjudul “Data Mining : Concept and Techniques Third Edition” (Han, 2012:244), menjelaskan bahwa :

A typical example of frequent itemset mining is market basket analysis. This process analyzes customer buying habits by finding associations between the different items that customer place in their ‘shopping baskets’. The discovery of these associations can help retailers develop marketing strategies by gaining insight into which items are frequently purchased together by customers.

Berdasarkan definisi, disimpulkan bahwa *association task* adalah tugas menemukan beberapa atribut yang muncul bersamaan dalam sebuah *record* atau *item*. *Association* biasa dipakai pada tabel transaksi untuk mengetahui tren dari konsumen. Terdapat beberapa metode yang dapat dipilih pada teknik asosiasi, seperti algoritma Apriori dan FP-Growth.

2.2.1 Apriori

Algoritma Apriori pertama kali ditemukan oleh Agrawal, et al., yang mana merupakan salah satu algoritma penting dari *data mining*. Prinsip algoritma ini adalah mencari lebar data, dimulai dari kemunculan 1 *itemset*, lalu 2 *itemset*, dan seterusnya hingga menuju yang terbesar, hingga akhirnya memunculkan prediksi atas kemunculan *itemset* terbesar. (Hegland, 2005:22)

Nama dari algoritma ini didasarkan pada penggunaan “pengetahuan sebelumnya” daripada kemunculan *itemset*. Apriori menggunakan pendekatan iteratif, yang dikenal sebagai pencarian berdasarkan tingkatan, dimana *itemset* berjumlah (k) digunakan untuk menelusuri

itemset berjumlah $(k+1)$. Langkah pertama yang dilakukan adalah pemindaian basis data yang bertujuan untuk mencari kemunculan 1 *itemset*, dan mencari jumlah kemunculan setiap *item*, yang nantinya digunakan untuk mencari nilai *support* minimum, atau disebut dengan L_1 . L_1 nantinya digunakan sebagai dasar untuk mencari L_2 , dan seterusnya hingga tidak ada lagi *itemset* yang sering muncul. Pencarian L_k membutuhkan satu kali pemindaian basis data. Untuk meningkatkan efisiensi daripada pencarian *itemset* yang sering muncul, maka terdapat prinsip yang dikenal dengan *Apriori Property* yang digunakan untuk mengurangi ruangan pencarian.

Apriori Property berbunyi: “Seluruh *subset* yang tidak kosong daripada *itemset* yang sering muncul, juga harus sering muncul.” Apabila sebuah *itemset* I tidak memenuhi nilai minimum *support* (min_sup), maka I dikatakan “tidak sering muncul”, dimana $P(I) < min_sup$. Apabila *item* A ditambahkan pada *itemset* I , maka hasil dari *itemset* tidak dapat muncul lebih banyak dari I . Maka, $I \cup A$ juga dikatakan “tidak sering muncul”, karena $P(I \cup A) < min_sup$. Prinsip ini dimiliki prinsip *antimonotonicity*, yang berbunyi: “Apabila sebuah *set* tidak lulus uji, maka seluruh *superset* yang dimilikinya akan tidak lulus uji pula.” (Han, 2012:249)

Apabila *Apriori Property* diaplikasikan pada algoritma, terdapat dua proses yang akan muncul, yaitu *join* dan *prune*.

1. *Join*

Untuk menentukan L_k , kandidat set dari k -*itemset* dibuat dengan menggabungkan (*join*) L_{k-1} dengan kandidat set itu sendiri, dimana kandidat set disimbolkan dengan C_k . Biarkan l_1 dan l_2 menjadi *itemset* daripada L_{k-1} . Notasi $l_1[j]$ mengacu pada *item* ke- (j) daripada l_1 (sebagai contoh, $l_1[k-2]$ mengacu pada *item* terakhir dari l_1). Untuk implementasi yang efisien, algoritma Apriori mengasumsikan bahwa *item* pada tabel transaksi atau *itemset* telah diurutkan berdasarkan urutan alfabetis. Untuk $(k-1)$ -*itemset*, l_1 , ini menunjukkan bahwa *item* diurutkan sehingga $l_1[1] < l_1[2] < \dots < l_1[k-1]$. Langkah *join*, $L_{k-1} \bowtie L_{k-1}$, dilakukan dimana anggota dari L_{k-1} dapat digabungkan apabila *item* $(k-2)$ pertama adalah umum. Anggota l_1 dan l_2 daripada L_{k-1} dilakukan *join* apabila $(l_1[1] = l_2[1]) \wedge (l_1[2] = l_2[2]) \wedge \dots \wedge (l_1[k-$

$2] = l_2[k-2] \wedge (l_1[k-1] < l_2[k-1])$. Kondisi $l_1[k-1] < l_2[k-1]$ muncul untuk memastikan bahwa tidak ada duplikasi yang dibuat. Hasil *itemset* dengan melakukan *join* l_1 dengan l_2 adalah $\{l_1[1], l_1[2], \dots, l_1[k-2], l_1[k-1], l_2[k-1]\}$.

2. Prune

C_k adalah *superset* daripada L_k , dimana anggota-anggotanya memiliki kemungkinan sering atau tidak sering muncul, namun seluruh k -*itemset* yang sering muncul tetap termasuk pada C_k . Pemindaian basis data dilakukan untuk menemukan jumlah setiap kandidat di C_k , dan menentukan L_k . Dengan ukuran C_k yang besar, maka proses terdapat proses komputasi yang berat dalam penentuan tersebut. Untuk mengurangi ukuran dari C_k , digunakan *Apriori Property*, dimana $(k - 1)$ -*itemset* yang tidak sering muncul, tidak dapat menjadi *subset* daripada k -*itemset* yang sering muncul. Apabila kandidat dikatakan tidak sering muncul, maka akan dihilangkan dari C_k .

<i>TID</i>	<i>List of item_IDs</i>
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3

Gambar 2.4
Tabel Transaksi
(Han, 2012:250)

Berdasarkan gambar 2.4 diatas, terdapat sembilan transaksi dan beberapa definisi penting yang nantinya akan digunakan sebagai panduan metode asosiasi, yaitu:

1. *Itemset*

Itemset adalah *subset* dari *item* yang ditandai dengan indeks. Contoh: (I1, I2, I5)

2. Support

Support adalah nilai kemunculan sebuah *item* pada kumpulan transaksi, dibagi dengan jumlah seluruh transaksi. Rumus dibawah menunjukkan berapa nilai *support* dari sebuah *itemset*.

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

Apabila terdapat dua atau lebih *itemset*, maka rumus berubah menjadi $A \cup B$, tergantung dari seberapa banyak *itemset* yang akan dicari nilai *support* nya.

$$Support(A, B) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{Total\ transaksi} \quad (Rumus\ 2.2)$$

Berdasarkan tabel, nilai *support* daripada I1 adalah $6 / 9 = 0.67$

3. Confidence

Confidence adalah nilai persentase atau pecahan yang didasarkan pada aturan $a \rightarrow b$ (jika a, maka b), yang menandakan kemunculan b pada saat a muncul.

$$Confidence(B|A) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{Total\ transaksi\ mengandung\ A} \quad (Rumus\ 2.3)$$

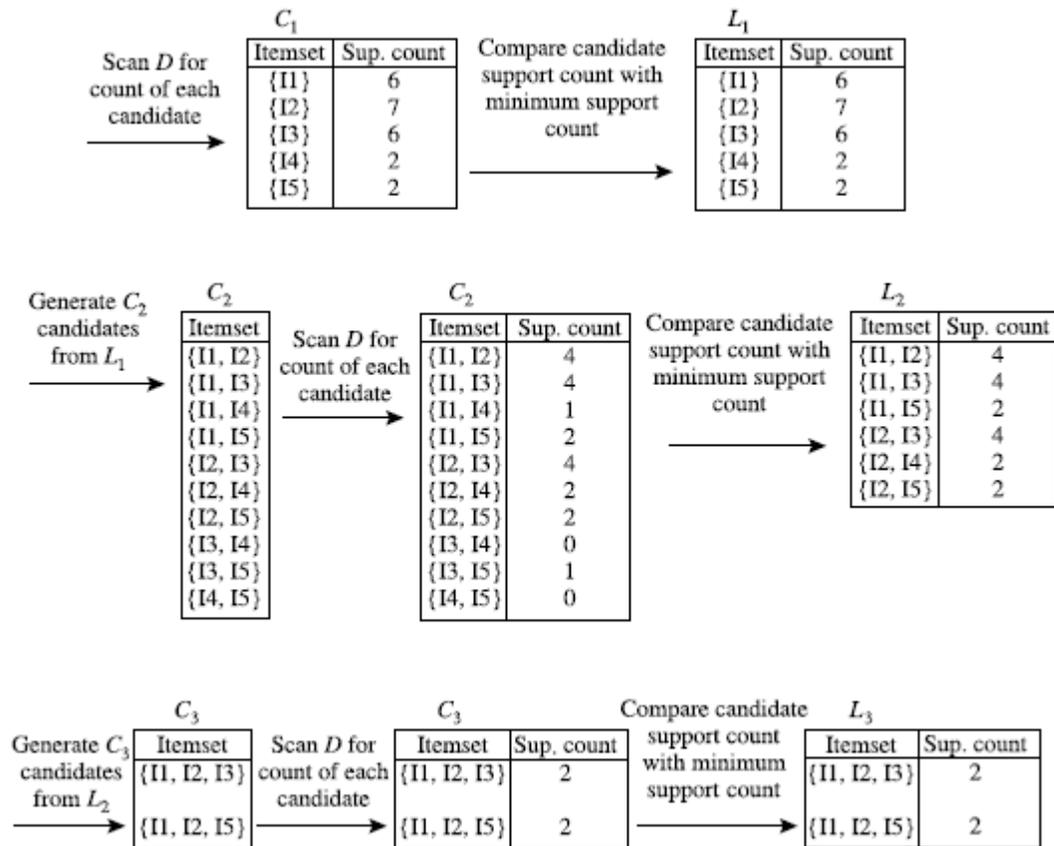
Berdasarkan tabel, nilai *confidence* I2 terhadap I1 adalah $4 / 6 = 0.67$

Dalam menentukan nilai *support* minimum, maka nilai rata-rata daripada jumlah kemunculan harus diketahui. Setelah nilai rata-rata daripada 2000 *record* diketahui, maka selanjutnya adalah membuat persentase nilai yang membandingkan antara nilai rata-rata dengan jumlah *record* keseluruhan.

$$Nilai\ support\ minimum = \frac{Rata - rata\ nilai\ kemunculan}{Jumlah\ record\ keseluruhan} \quad (Rumus\ 2.4)$$

Dalam menentukan nilai *confidence* minimum, maka nilai tertinggi daripada jumlah kemunculan harus diketahui. Setelah nilai rata-rata daripada 2000 *record* diketahui, maka selanjutnya adalah membuat persentase nilai yang membandingkan antara nilai tertinggi dengan jumlah *record* keseluruhan.

$$Nilai\ confidence\ minimum = \frac{Nilai\ tertinggi\ kemunculan}{Jumlah\ record\ keseluruhan} \quad (Rumus\ 2.5)$$



Gambar 2.5
Langkah-langkah Algoritma Apriori
(Han, 2012:251)

Adapun langkah-langkah yang dilakukan dalam algoritma Apriori, berdasarkan gambar 2.5:

1. Pada iterasi pertama, setiap *item* adalah bagian dari set kandidat dari 1 *itemset*, C_1 . Algoritma lalu memindai seluruh transaksi untuk menghitung jumlah kemunculan dari setiap *item* (nilai *support*).
2. Diasumsikan bahwa nilai *support* minimum adalah 2, set dari 1 *itemset* yang sering muncul, dapat ditentukan. Isi dari set ini adalah seluruh 1 *itemset* yang memenuhi syarat nilai *support* minimum.
3. Untuk menemukan set dari 2 *itemset* yang sering muncul, algoritma ini menggunakan metode *join* untuk menghasilkan set kandidat dari 2 *itemset*, C_2 .

4. Selanjutnya, transaksi D dipindai dan nilai *support* dari setiap *itemset* kandidat C_2 diakumulasikan.
5. Set dari 2 *itemset* yang sering muncul, L_2 , ditentukan, dimana isinya adalah kumpulan 2 *itemset* yang memenuhi syarat nilai *support* minimum.
6. Pembuatan set kandidat dari 3 *itemset*, C_3 . L_3 ditentukan.
7. Transaksi D dipindai untuk menentukan L_3 , dimana isinya adalah kumpulan 3 *itemset* yang memenuhi syarat nilai *support* minimum.
8. Pembuatan set kandidat dari 4 *itemset*. Apabila seluruh *itemset* tidak memenuhi syarat nilai *support* minimum, maka algoritma dihentikan, dan menghasilkan seluruh *itemset* yang sering muncul.

2.2.2 FP-Growth

Pada algoritma Apriori, dapat terlihat bahwa metode kemunculan dan pengujian kandidat akan secara signifikan mengurangi ukuran dari set kandidat, yang menyebabkan performa yang baik. Namun, terdapat beberapa kelemahan yang dimiliki algoritma Apriori:

1. Algoritma Apriori masih harus menghasilkan sejumlah besar set kandidat. Apabila terdapat 10^4 kemunculan pada 1 *itemset*, maka kandidat 2 *itemset* akan berjumlah lebih dari 10^7 .
2. Algoritma Apriori masih harus memeriksa seluruh basis data berulang kali dan memeriksa sejumlah besar kandidat melalui kesamaan pola. Dibutuhkan sumber daya yang besar untuk memeriksa setiap transaksi dari basis data untuk memeriksa nilai *support* daripada kandidat *itemset*.

Jiawei Han memberikan definisi pada bukunya yang berjudul "*Data Mining : Concept and Techniques Third Edition*", menjelaskan bahwa : "*Can we design a method that mines the complete set of frequent itemsets without such a costly candidate generation process? An interesting method in this attempt is called frequent pattern growth, or simply FP-Growth*" (Han, 2012:257)

FP-Growth bekerja dengan cara “membagi dan menakhlukkan”. Pada prosesnya, dibutuhkan dua pemeriksaan pada basis data. Pertama-tama, FP-Growth menghitung jumlah daftar dari *item* yang sering muncul, lalu diurutkan berdasarkan frekuensi terkecil atau *descending order* saat pertama kali basis data dipindai. Pada pemindaian kedua, basis data dilakukan kompresi menjadi pohon FP. Lalu, algoritma FP-Growth dilakukan untuk menambang setiap *item* yang nilai *support* nya lebih tinggi dari ξ secara rekursif. Metode penemuan masalah terhadap besarnya kemunculan *item* juga diubah menjadi pohon dan proses rekursif dilakukan kembali. (Li et al., 2008:2)

Meskipun algoritma FP-Growth terbilang lebih efisien dibandingkan dengan Apriori, terdapat beberapa tantangan mengenai manajemen *resource* yang dihadapi algoritma FP-Growth:

1. Kapasitas penyimpanan

Untuk basis data yang memiliki ukuran besar, maka pohon FP yang dihasilkan pun besar, dan tidak memungkinkan cukup untuk diproses pada memori utama. Maka diperlukan untuk membuat beberapa basis data kecil untuk merepresentasikan sebuah basis data yang utuh. Sebagai hasil, setiap basis data kecil yang baru, dapat dimuat dalam memori dan menghasilkan masing-masing pohon FP lokal.

2. Distribusi komputasi

Seluruh tahapan pada algoritma FP-Growth dapat dilakukan secara parallel, khususnya pada saat melakukan panggilan rekursif.

3. Harga komunikasi yang tinggi

Algoritma FP-Growth parallel sebelumnya membagi-bagi basis data menjadi beberapa kelompok transaksi. Pohon FP yang terdistribusi dapat bersifat inter-dependen, dan dapat memunculkan sinkronisasi diantara eksekusi *thread* secara parallel.

4. Nilai *support*

Nilai *support* ξ memegang peranan penting pada algoritma FP-Growth. Semakin besar nilai ξ , maka lebih sedikit hasil pola yang dimunculkan, dan lebih sedikit harga daripada

komputasi dan penyimpanan. Umumnya, pada basis data berskala besar, nilai ξ harus diatur cukup tinggi untuk dapat memenuhi kapasitas penyimpanan. Namun, untuk tugas *web mining*, nilai ξ diatur sangat rendah untuk menerima *itemset* yang panjang. Nilai rendah ini mungkin membutuhkan waktu komputasi yang sulit diterima.

2.2.3 Perbandingan Algoritma

Berdasarkan riset terdahulu yang dilakukan oleh Philippe Fournier-Viger, et al., yang berjudul “A Survey of Itemset Mining”, terdapat beberapa perbedaan algoritma FIM (*Frequent Item Mining*) berdasarkan metode pencarian dan representasi basis data. Dari keenam algoritma FIM yang saat ini telah dikembangkan, terdapat Apriori dan FP-Growth.

Algorithm	Type of search	Database representation
<i>Apriori</i> ³	breadth-first (candidate generation)	Horizontal
<i>Apriori – TID</i> ³	breadth-first (candidate generation)	Vertical (TID-lists)
<i>Eclat</i> ²⁸	depth-first (candidate generation)	Vertical (TID-lists, diffsets)
<i>FP – Growth</i> ²⁷	depth-first (pattern-growth)	Horizontal (prefix-tree)
<i>H – Mine</i> ²⁹	depth-first (pattern-growth)	Horizontal (hyperlink structure)
<i>LCM</i> ³⁰	depth-first (pattern-growth)	Horizontal (with transaction merging)

Gambar 2.6

Tabel Perbedaan Algoritma FIM (Fournier-Viger, et al., 2017:15).

Berdasarkan gambar 2.6, dapat dilihat bahwa terdapat perbedaan yang nampak dalam cara kerja dan metode yang dilakukan kedua algoritma. *Pseudocode* pada masing-masing algoritma dibahas lebih rinci pada bab 2. Algoritma Apriori adalah salah satu algoritma yang penting karena menjadi dasar daripada algoritma-algoritma lainnya. Namun terdapat beberapa keterbatasan yang dimiliki daripada algoritma ini.

1. Algoritma Apriori menciptakan *candidate* dengan menggabungkan *itemset* tanpa melihat basis data. Maka itu, pola-pola yang tidak terbentuk dalam basis data akan diciptakan pula. Hal ini menyebabkan besarnya waktu pemrosesan yang dibutuhkan sebuah kandidat yang tidak terdapat dalam basis data.
2. Algoritma Apriori harus berulang kali memeriksa basis data untuk mengetahui nilai *support* dari masing-masing *candidate*, yang mana merupakan tugas yang berat.

3. Algoritma Apriori memakan sejumlah besar memori dengan metode pencarian *breadth-first* karena membutuhkan pemeriksaan *itemset* berjumlah (k) dan (k-1) pada memori, dimana $k > 1$. Dari segi kompleksitas, Hegland telah melakukan analisis kompleksitas terhadap algoritma Apriori. Secara singkat, kompleksitas waktu daripada algoritma ini adalah $O(m^2n)$, dimana (m) adalah jumlah daripada *item* yang terlibat, dan (n) adalah jumlah transaksi yang terlibat.

Terdapat kelebihan yang dimiliki oleh algoritma yang memiliki metode *pattern-growth*. Kelebihan tersebut adalah algoritma ini hanya menelusuri *itemset* yang sering muncul dan melewati *itemset* yang tidak muncul dalam basis data. Algoritma ini menggunakan konsep *projected database*, yang mana hanya *field* tertentu saja yang dimunculkan dan/atau digunakan dengan tujuan untuk mengurangi waktu dan harga yang dibutuhkan dalam memeriksa basis data. Algoritma FP-Growth menggunakan *tree* untuk merepresentasikan basis data sebuah projek untuk mengurangi penggunaan memori. Satu-satunya masalah yang memiliki kemungkinan untuk muncul adalah interdependensi data yang disebabkan karena struktur *tree* yang dihasilkan algoritma ini. Interdependensi data ini biasanya menyebabkan *bottleneck* pada memori.

2.3 Game

Terdapat dua definisi yang terkandung dalam Kamus Besar Bahasa Indonesia (KBBI), yaitu :

1. Sesuatu yang dilakukan dengan tujuan untuk menyenangkan hati (dengan menggunakan alat-alat tertentu atau tidak)
2. Melakukan perbuatan untuk bersenang-senang (dengan alat-alat tertentu atau tidak)
3. Menurut Wolfgang Kramer, seorang desainer permainan papan asal Jerman pada jurnalnya yang berjudul "*What Is A Game?*" mendefinisikan bahwa *game* merupakan aktivitas yang dilakukan hanya untuk tujuan kesenangan dan tanpa adanya keterpaksaan. Jadi, bentuk apapun dari sebuah aktivitas yang memberikan kesenangan dapat disebut sebagai *game*, seperti berdansa, memainkan alat musik, bermain boneka, dan lain sebagainya.

Ernest Adams memberikan definisi pada bukunya yang berjudul "Fundamentals of Game Design", menjelaskan bahwa : "A game is a type of play activity, conducted in the context of a pretended reality, in which participant(s) try to achieve at least one arbitrary, nontrivial goal by acting in accordance with rules." (Adams,2010:3)

Berdasarkan pengertian diatas, dapat disimpulkan bahwa *game* atau permainan adalah aktivitas yang dilakukan oleh satu atau lebih pemain dengan beberapa aturan yang berlaku. Aktivitas tersebut pada umumnya dilakukan untuk memberikan kesenangan bagi para pemainnya.

2.3.1 Evolusi Video Game

Menurut Steven L. Kent pada bukunya yang berjudul "*The Ultimate History of Video Games*", desain dari permainan komputer mengalami perubahan dari tahun 1889 hingga saat ini, yang mana memiliki interval waktu lebih dari 100 tahun lamanya.

Adapun tujuh elemen yang mengalami perubahan, dimana saling mempengaruhi satu sama lain, yaitu :

1. Perangkat keras

Dengan berkembangnya teknologi setiap tahunnya, perkembangan *game* pun mengalami perubahan dalam penggunaan perangkat keras yang lebih canggih, seperti *3D graphic card*, *DVD*, *internet*, dan lain sebagainya. Penggunaan beberapa perangkat keras tersebut memungkinkan para *game designer* untuk membuat *game* yang lebih berbeda dan menarik.

2. Alat interaksi

Penggunaan alat interaksi yang lebih canggih seperti *keyboard*, *joystick*, *oculus*, *Wii controller*, dan *kinect system* memberikan dampak yang signifikan pada *gameplay*. Dengan adanya peralatan tersebut, pergerakan dan interaksi pemain terhadap *game* semakin bertambah luas.

3. Perangkat lunak yang tersedia

Pada awalnya, para pengembang *game* menulis baris kode dan membuat aset secara manual. Namun dengan adanya *level editor*, *game engine*, *image editor*, dan perangkat lunak lainnya, proses pengembangan *game* dapat dilakukan dengan efektif dan efisien tanpa harus menambah waktu pada proses pembuatannya.

4. Bisnis

Pada awalnya, *game* hanya dikembangkan oleh kelompok kecil atau orang-orang tertentu. Namun dewasa ini, banyak perusahaan *game* yang menerima dana pengembangan dengan kisaran ribuan hingga jutaan dollar per *game*. Bahkan, muncul beberapa program pendidikan khusus yang melatih para profesional untuk masuk pada industri pembuatan *game*.

5. Demografi pemain

Pada beberapa tahun terakhir, jumlah wanita sebagai pemain *game* hampir mendekati jumlah pria. Dewasa ini, bahkan umur dari pemain berkisar antara 4 hingga 100 tahun. Hal

ini merupakan perubahan yang signifikan apabila dibandingkan dengan para remaja pria sebagai pemain *game* pada awalnya.

6. Diversifikasi

Pada awalnya, *game* komputer hanya dimainkan pada mesin arkade. Namun dengan perkembangan selama bertahun-tahun pada jenis-jenis platform, banyak orang menggunakan konsol, PC (*Personal Computer*), alat genggam, telepon, televisi, dan lain sebagainya sebagai platform pengganti. Jenis-jenis pemain pun sudah bervariasi, dimulai dari pemain kasual, *online*, hingga *hardcore*. Segala perubahan tersebut membawa pada perubahan pada model bisnis dan biaya pengembangan *game*.

7. Desain video game

Desain daripada *game* dapat berubah, tergantung dari penggunaan perangkat keras, tipe-tipe pemain, dan lain sebagainya. Dengan munculnya berbagai desain yang baru, para pengembang *game* memahami lebih baik tentang cara membuat *game* yang menarik.

Berdasarkan tujuh poin diatas, Steven L. Kent menarik kesimpulan bahwa perkembangan *video game* merupakan sebuah proyek tidak akan menemui batas akhir, karena perkembangan teknologi yang pesat dari hari ke hari akan membuat *game* yang dihasilkan menjadi semakin menarik. (Kent, 2000:605)

2.3.2 Jenis-jenis Video Game

Video game dapat diklasifikasikan menjadi beberapa genre berdasarkan interaksi pemain dengan *video game*, atau dapat disebut *gameplay*. (Wiley, 2014:5)

1. Educational

Dalam *educational game*, pemain dihadapkan pada permainan yang menyenangkan, dimana mereka dapat belajar ilmu pengetahuan dasar seperti perhitungan maupun pengejaan kata. Contoh dari *educational game* adalah *The Magic School Bus in the Time of Dinosaurs*.

2. Action

Dalam *action game*, *gameplay* lebih berfokus pada pertarungan. *Game* ini memberikan instruksi pada pemain untuk melakukan gerakan cepat untuk menyerang lawan. Ada dua sub tipe dari genre ini, yaitu *shooter game* dan *fighting game*. *Shooter game* melibatkan perlawanan yang menggunakan senjata tertentu, biasanya berupa senjata api atau benda lainnya yang biasa digunakan untuk menembak. Sasaran dari *game* ini adalah memastikan bahwa pemain tetap bertahan hingga akhir dari misi yang diberikan. Beberapa contoh dari *shooter game* adalah *Time Crisis*, *Halo 4 : Forward Unto Dawn*, dan *Gunstinger*. Sedangkan *fighting game* melibatkan serangan fisik pada lawan hingga lawan mati atau kalah. Beberapa contoh dari *fighting game* adalah *Street Fighter*, *Tekken*, dan *Mortal Kombat*.

3. Adventure

Adventure game lebih berfokus pada pemecahan teka-teki dibandingkan dengan aksi maupun tantangan. Biasanya pemecahan teka-teki tersebut dapat dilakukan dengan berinteraksi dengan karakter lainnya pada lingkungan *game*. Beberapa contoh dari *adventure game* adalah *Myst*, seri *game Harry Potter*, dan *Kinect Adventures*.

4. Sports

Sports game merupakan bentuk simulasi dari olahraga yang bersifat tradisional, seperti sepak bola maupun bulutangkis. Beberapa *game* dari genre ini berfokus pada aturan olahraga sebenarnya atau strategi yang dilakukan saat permainan berlangsung. Contoh dari *sports game* adalah *Kinect Sports : Season 2*.

5. Simulation

Simulation game merupakan sebuah simulasi aktivitas berdasarkan kehidupan nyata atau fiksi. Genre dari *game* ini memungkinkan pemain untuk melakukan analisis terhadap objek-objek yang terlibat di dalam *game*. Beberapa contoh dari *simulation game* adalah *Sim City* dan *Conduct Flight Simulator 3*.

6. Role-playing game (RPG)

Role-playing game, atau biasa disebut RPG merupakan sebuah *game* dimana pemain berperan menjadi karakter tertentu pada jalan cerita yang sudah ditentukan sebelumnya. Tergantung dari karakter yang diperankan, setiap karakter memiliki keunikan tersendiri dilihat dari kemampuan bertarung maupun perannya dalam sebuah cerita. Pemain biasanya berpetualang dan menyelesaikan misi untuk mendapatkan jalan menuju lokasi lain seperti istana, atau *dungeon*, dan kepemilikan benda dengan jumlah yang cukup besar seperti senjata maupun uang. Beberapa contoh dari *Role-playing game* adalah *Lost Odyssey*, *Fable Heroes*, dan seri *game Etrian Odyssey*.

7. Fantasy

Fantasy game merupakan sebuah gambaran fiksi atau khayalan dari dunia virtual. Beberapa contoh dari *fantasy game* adalah seri *game Final Fantasy* dan seri *game Baldur's Gate*.

8. Card

Card game adalah sebuah permainan yang melibatkan permainan kartu secara virtual. Salah satu *card game* yang terkenal adalah *Solitaire*.

9. Board

Board game adalah sebuah permainan yang melibatkan perpindahan benda-benda dengan aturan yang berlaku pada sebuah papan permainan virtual. Beberapa diantaranya memerlukan strategi, kesempatan, atau keduanya untuk mencapai tujuan masing-masing. Beberapa contoh terkenal dari *board game* adalah *Backgammon* atau catur.

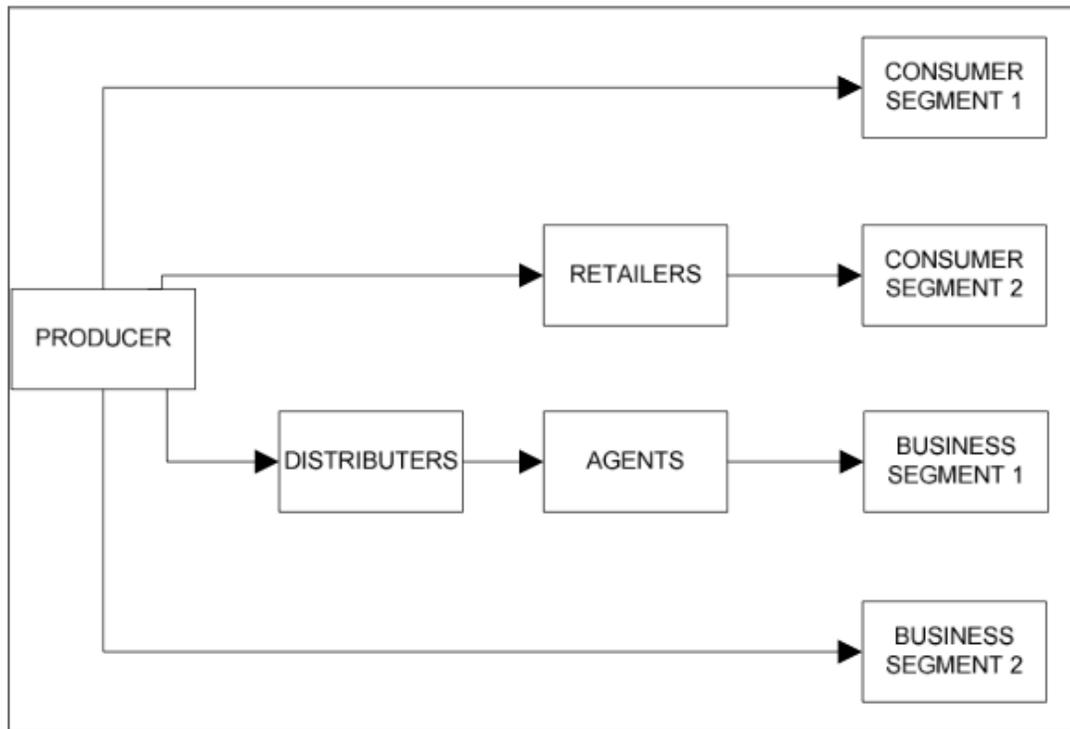
10. Strategy

Strategy game adalah sebuah permainan yang pemikiran dan perencanaan yang penuh perhitungan dan ketelitian untuk memenangkan permainan tersebut. Cakupan aktivitas dari permainan ini memiliki keanekaragaman di mulai dari penguasaan dunia hingga taktik yang melibatkan sejumlah orang. Beberapa contoh dari *strategy game* adalah *Warhammer*, *Plants vs. Zombies*, dan *Civilization*.

2.4 Distribusi Digital

Berdasarkan definisi yang diberikan daripada badan hukum Amerika Serikat (definitions.uslegal.com) distribusi digital (*digital distribution*), atau dikenal sebagai distribusi perangkat lunak elektronik (*electronic software distribution*) adalah metode distribusi yang mana konten yang disampaikan tidak menggunakan media fisik, melainkan menggunakan internet sebagai sarana pengunduhan yang akan langsung tertuju pada komputer pengguna, biasanya terletak di rumah masing-masing. Kemunculan distribusi digital menggantikan metode distribusi konvensional menggunakan media fisik, seperti kertas maupun DVD. Metode ini memungkinkan konsumen untuk masuk pada situs resmi dan melihat sampel produk yang dapat diunduh dan langsung tersimpan pada penyimpanan komputer daripada masing-masing konsumen. Distribusi digital memiliki tingkat keamanan lebih tinggi dalam hal transaksi secara *online*, pemberian lisensi pada konsumen, dan anti-pembajakan.

Phillipe Kern memberikan penjelasan pada rapat Think Tank yang berjudul “*The Impact of Digital Distribution – A Contribution*”, bahwa distribusi digital biasanya digunakan oleh para *artist*, *developer*, dan pihak pengembang lainnya untuk menyalurkan karya nya pada ruang lingkup publik tanpa harus mengeluarkan biaya lebih. Dengan metode ini, konsumen mendapatkan barang dengan harga lebih murah, dan pihak pengembang mendapatkan keuntungan dan kebebasan lebih dalam berkarya maupun mencari pasar.



Gambar 2.7
Sistem Distribusi
(Kotler, 2005:870)

Pada gambar 2.7 mengenai sistem distribusi tradisional, produsen mengeluarkan biaya lebih dalam proses percetakan, promosi, dan penyebaran produk pada beberapa toko fisik. Konsumen harus menggunakan jasa yang disediakan toko fisik untuk dapat membeli, menggunakan, dan/atau menikmati produk. Hal ini menyebabkan konsumen harus menggunakan waktu ekstra untuk dapat mencapai produk yang diinginkan. Contoh sistem distribusi tradisional meliputi toko buku, toko DVD, toko *video game*, siaran televisi, dan siaran radio. Distribusi digital telah mengubah model bisnis tradisional dalam industri musik maupun berita. Metode distribusi ini telah memberikan dampak yang sangat besar pada seluruh bentuk konsumsi media, termasuk penerbitan buku dan televisi. Distribusi digital merupakan sebuah bentuk tantangan pada tayangan dan siaran tradisional.

Dewasa ini, setiap acara televisi yang terdapat di dunia, khususnya negara maju seperti Amerika Serikat, dapat ditonton secara *online*. YouTube dan beberapa penyedia jasa tayangan *online* lainnya dapat diakses menggunakan TiVo dan perangkat elektronik lainnya.

Televisi saat ini semakin berkembang untuk dapat tersambung pada internet untuk mendapatkan akses konten yang baru setiap harinya. *Video on Demand* (VOD) yang ditonton maupun diunduh semakin bertambah dari hari ke hari. (Kern, 2009:1)

Beberapa regulator memiliki pandangan yang berbeda-beda mengenai VOD. Negara-negara dengan industri sinema yang kuat memiliki kemungkinan untuk menunjukkan ketertarikan terhadap industri tayangan dibandingkan dengan ketertarikan terhadap internet dan telekomunikasi. Adapun beberapa lingkungan bisnis yang mungkin terpengaruh distribusi digital (Kern, 2009:18):

1. Krisis ekonomi yang sedang berlangsung
2. Meningkatnya persaingan bisnis daripada media hiburan lainnya, seperti video game, internet, dan lain sebagainya.
3. Pembajakan
4. Kurangnya investasi terhadap sinema dari distributor dan perusahaan televisi.
5. Meningkatnya batasan-batasan aturan yang dikeluarkan dalam produksi film.
6. Munculnya bentuk baru dalam budaya konsumsi yang dipicu oleh internet.

Terdapat beberapa contoh perusahaan distribusi digital yang mengutamakan *video game* sebagai produk utamanya. Tergantung daripada media yang digunakan untuk memainkan *video game* tersebut, perusahaan berfokus pada *video game* yang dapat dimainkan pada konsol, PC, maupun mesin *arcade*. Beberapa contoh perusahaan yang berkecimpung di dunia distribusi digital pada *video game* adalah PlayStation Store, Steam, Uplay, dan Xbox Games Store.

2.5 Steam

Berdasarkan dokumentasi *online* mengenai yang didapat daripada situs resmi *Steam* (<https://partner.steamgames.com/doc/home>), *Steam* adalah *platform* digital yang dibangun untuk kebutuhan hiburan, khususnya untuk bermain *video game* secara *offline*, *multiplayer*, *local*, maupun *online*. Bagi *user*, *Steam* memiliki fungsi sebagai sarana untuk mencari *video game* yang sesuai dengan preferensi *user* dilihat dari segi genre, popularitas, tipe pemain, dan lain sebagainya. Harga dari *video game* yang disediakan oleh *Steam* berkisar dari gratis hingga membayar, dimana harga tersebut ditentukan oleh *developer* dan/atau *publisher* dari *video game* yang bersangkutan. Bagi *developer* dan/atau *publisher*, *Steam* memiliki fungsi sebagai sarana untuk mempromosikan dan menjual *video game* secara *online*. Meskipun uang yang diperoleh oleh *developer* dipotong oleh *Steam* per unit pembelian, *Steam* tetap menjaga hak cipta atas produk yang dibuat oleh *developer* dan tidak meng-klaim bahwa produk tersebut menjadi milik *Steam*.

2.5.1 Fitur-fitur pada Steam

Terdapat beberapa fitur yang dimiliki *Steam*, dimana fitur utama adalah layanan penyediaan *video game* secara *online* (<https://store.steampowered.com/about/>). Fitur-fitur yang dimiliki *Steam* antara lain :

1. Penawaran *video game*

Video game yang ditawarkan *Steam* memiliki keberagaman tinggi sehingga memungkinkan beragam *user* yang menggunakan layanan *Steam* untuk mengakses *video game* yang menjadi preferensi *user* masing-masing. *Game* yang ditawarkan tidak hanya memiliki keberagaman dalam hal *genre*, namun juga harga yang ditawarkan, popularitas, kesesuaian terhadap *system requirement* dari alat yang digunakan *user*, dan lain sebagainya. Pada tanggal-tanggal tertentu seperti *Black Friday*, *Valentine's Day*, hari raya imlek, dan lain sebagainya, *Steam* umumnya memberikan potongan harga dan/atau penawaran paket *game* kepada *user*.

2. Komunitas *online*

Sesama *user* memungkinkan untuk saling berkomunikasi dan membentuk grup seperti layaknya pada media sosial. *User* dapat berkomunikasi pada saat bermain *video game* dalam *Steam*. Fitur ini berguna untuk menyediakan komunikasi antar tim apabila *user* bermain *video game* yang bersifat *multiplayer*, *local*, atau *online* bersama dengan *user* yang lain.

3. Workshop

Workshop memungkinkan *user* untuk memberikan ide atau membuat konten baru pada *game* yang mereka suka. Ide-ide dan konten-konten tersebut nantinya ditampung untuk diberikan kepada *developer* untuk dipertimbangkan pada proses pengembangan *game*. Konten baru yang diusulkan atau dibuat oleh *user* akan muncul dalam *game* sebagai *mod*.

2.5.2 Informasi Video Game pada Steam

Pada situs resmi *Steam*, *user* dapat melihat informasi yang terdapat pada sebuah *video game* yang hendak dibeli atau diulas. Secara umum, terdapat beberapa *parameter* yang menjadi informasi yang terkandung dalam sebuah *video game*.

1. Judul

Judul dari *video game*. *Parameter* ini bertipe data String.

2. Pengecekan umur *user*

Pengecekan umur yang dilakukan apabila *video game* yang bersangkutan dikhususkan untuk batas umur tertentu, biasanya 18 tahun. *Parameter* ini bertipe data Boolean.

3. Harga

Harga yang ditentukan *developer* dan/atau *publisher* atas *video game* yang terdaftar dalam *Steam*. Tergantung dari posisi *user* secara regional, mata uang yang ditampilkan mungkin berbeda-beda. *Parameter* ini bertipe data Integer.

4. Genre

Setiap *video game* memiliki setidaknya satu *genre*. *Parameter* ini bertipe data String.

5. Developer

Nama grup atau tim yang menjadi pengembang dari *video game*. *Video game* mungkin memiliki nama *developer* dan *publisher* yang sama. *Parameter* ini bertipe data *String*.

6. Publisher

Nama grup atau tim yang menjadi penerbit dari *video game*. *Video game* mungkin memiliki nama *developer* dan *publisher* yang sama. *Parameter* ini bertipe data *String*.

7. Tanggal terbit

Tanggal diterbitkannya *video game* yang bersangkutan. *Parameter* ini bertipe data *Date*.

8. Tipe pemain

Tipe pemain mengacu pada berapa banyak pemain yang dapat berpartisipasi dalam *video game*. Tipe pemain yang tersedia adalah *single*, *multiplayer*, *local*, dan *online*. *Single* memungkinkan hanya satu pemain yang dapat bermain *game*. *Multiplayer* memungkinkan dua atau lebih pemain yang dapat bermain *game*. *Local* memungkinkan beberapa pemain dapat bermain menggunakan *Local Area Network*. *Online* memungkinkan sesama pemain dapat saling bermain menggunakan koneksi internet. *Parameter* ini bertipe data *String*.

9. Bahasa

Bahasa yang tersedia dalam *video game* umumnya bahasa Inggris, meskipun tidak tertutup kemungkinan adanya bahasa lain yang digunakan. *Parameter* ini bertipe data *String*.

10. Achievements

Achievements mengacu pada ada atau tidaknya pencapaian tertentu pada saat *game* dimainkan. Pencapaian tersebut bukan bagian utama dari *game*. *Parameter* ini bertipe data *Boolean*.

11. Full Controller Support

Full Controller Support mengacu pada penggunaan *joystick* atau kontroler lainnya selain *keyboard* dan *mouse* pada saat *game* dimainkan. *Parameter* ini bertipe data *Boolean*.

12. Workshop

Workshop mengacu pada ada atau tidaknya *workshop* pada sebuah *game* dimana pemain dapat berpartisipasi untuk memberikan usul mengenai konten dalam *game*. *Parameter* ini bertipe data Boolean.

13. Trading Cards

Trading Cards mengacu pada ada atau tidaknya *trading card* yang didapat pemain pada saat *game* dimainkan. Berbeda dengan *achievements*, *trading card* dapat ditukar dan dijual-belikan kepada pemain lain sebagai kartu koleksi. *Parameter* ini bertipe data Boolean.

14. Cloud

Steam Cloud adalah layanan yang disediakan *Steam* untuk menyimpan data *game* dan file lainnya secara otomatis pada server Valve. Pemain dapat mengakses data ini menggunakan perangkat keras apapun yang menjalankan aplikasi *Steam*. *Parameter* ini bertipe data Boolean.

15. Level Editor

Level Editor adalah fitur dalam *game* yang memungkinkan pemain untuk membuat desain level pada *game* yang dimainkan. Desain level tersebut nantinya dimungkinkan untuk dimainkan oleh pemain lain. *Parameter* ini bertipe data Boolean.

16. Captions

Captions memungkinkan *game* yang dimainkan untuk mengganti bahasa pada dialog dalam *game*. *Parameter* ini bertipe data Boolean.

17. Leaderboards

Leaderboards adalah sistem urutan berdasarkan pemain terbaik dan/atau pemain dengan pencapaian tertinggi dalam memainkan *game* tertentu. *Parameter* ini bertipe data Boolean.

18. Minimum System Requirements

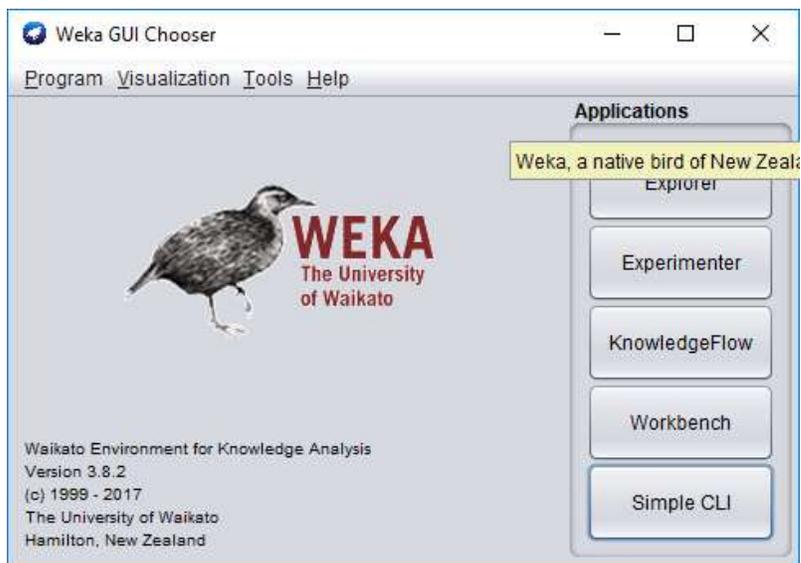
Minimum system requirements mengacu pada prasyarat perangkat keras yang dibutuhkan untuk memainkan *video game* dengan lancar. *Parameter* ini bertipe data String.

2.6 Weka

Berdasarkan situs resmi Universitas Waikato di Selandia Baru (<https://www.cs.waikato.ac.nz/ml/weka/>), Weka (Waikato Environment for Knowledge Analysis) merupakan kumpulan algoritma *machine learning* yang digunakan untuk mengerjakan tugas-tugas daripada *data mining*. Algoritma-algoritma tersebut dapat diaplikasikan secara langsung pada sebuah dataset atau dipanggil pada *source code* yang ditulis dengan bahasa pemrograman Java. Weka mengandung bermacam-macam alat untuk *data preprocessing*, proses klasifikasi, analisis regresi, asosiasi, dan visualisasi. Aplikasi ini cocok digunakan untuk mengembangkan skema baru daripada *machine learning*. Weka merupakan aplikasi *open source* yang dilisensikan oleh GNU General Public License.

Weka menyediakan implementasi mengenai algoritma pembelajaran yang dapat diaplikasi dengan mudah terhadap dataset yang dibutuhkan. Aplikasi ini meliputi berbagai macam peralatan yang dapat digunakan untuk mengubah bentuk dataset. *User* dimungkinkan untuk melakukan *data preprocessing*, membuat skema, dan melakukan analisis terhadap performa dan hasil, tanpa harus mengetik baris program. Fitur-fitur daripada aplikasi meliputi metode untuk menyelesaikan masalah mengenai permasalahan standar mengenai *data mining*: analisis regresi, klasifikasi, pengklasteran, asosiasi, dan visualisasi data. Seluruh algoritma menerima *input* dalam bentuk sebuah tabel relasi dalam format *.arff*, yang mana dapat dibaca pada sebuah *file* atau dibangun melalui *query*. (Witten, 2005:366)

Cara termudah dalam menggunakan WEKA adalah melalui antarmuka yang disebut *Explorer*. Ini memberikan akses menuju seluruh fasilitas, dalam bentuk pemilihan menu. WEKA dapat membaca beberapa format file selain *.arff*, seperti *.csv*, *.json*, *.dat*, *.xrff*, dan lain sebagainya.



Gambar 2.8
Antarmuka Awal Weka 3.8.2

Berdasarkan gambar 2.8, terdapat lima tombol yang dapat dipilih *user* untuk memulai aplikasi, meliputi *Workbench* yang pada Weka versi 3.5.8 belum ditambahkan (Bouckaert, 2013:31):

1. Explorer

Explorer adalah sebuah lingkungan yang memberikan panduan pada *user* dalam menjelajahi data menggunakan Weka.

2. Experimenter

Experimenter adalah sebuah lingkungan yang diperuntukkan untuk melakukan eksperimen dan pengujian statistik diantara beberapa skema pembelajaran.

3. KnowledgeFlow

KnowledgeFlow adalah sebuah lingkungan yang memiliki fungsionalitas yang serupa dengan *Explorer*, namun antarmuka yang disajikan berbentuk *drag-and-drop*. Salah satu kelebihan mengenai *KnowledgeFlow* adalah membantu mengenai pembelajaran inkremental.

4. Workbench

Workbench memiliki kesamaan dengan *Explorer*, dengan tampilan antarmuka yang lebih sederhana.

5. SimpleCLI

SimpleCLI adalah antarmuka *simple command-line* pada Weka yang serupa dengan *Command Prompt*, yang memungkinkan *user* untuk melakukan eksekusi langsung terhadap perintah Weka untuk sistem operasi yang tidak menyediakan antarmukanya masing-masing.

2.7 Penelitian Terdahulu

Perbedaan penelitian sebelumnya dengan penelitian yang akan penulis lakukan adalah:

1. Penelitian dilakukan terhadap data sekunder *video game* dari Steam. Metode yang digunakan adalah metode asosiasi menggunakan algoritma Apriori dengan fokus terhadap tren pengembangan *video game*. Sementara penelitian dengan objek yang sama menggunakan metode asosiasi dengan menggunakan algoritma Apriori dan FP-Growth dengan fokus pada perbandingan ke 2 algoritma tersebut.
2. Melakukan penelitian terhadap *association rule* guna menciptakan aturan kombinasi genre daripada *video game*.

Berikut ini merupakan kumpulan penelitian dan jurnal mengenai asosiasi dan aplikasinya pada bisnis:

Tabel 2.1
Tabel Penelitian dan Jurnal.

No.	Nama Peneliti	Tahun	Metode	Hasil
1	Haoyuan Li, et al.	2008	FP-Growth	Metode PFP dapat secara efektif digunakan dalam keperluan <i>web mining</i> , khususnya dalam menambang <i>tag</i> dan <i>URL</i> . Di masa depan, PFP dapat diaplikasikan pada <i>query log</i> daripada Google, atau <i>search engine</i> sejenis.

No.	Nama Peneliti	Tahun	Metode	Hasil
2	Cajetan Rodrigues, et al.	2012	Apriori	Desainer maupun orang lain yang mengumpulkan data harus memutuskan data yang benar, dan darimana sumber data tersebut. Desainer seharusnya merefleksikan aspek sebuah <i>video game</i> yang ditujukan untuk meningkatkan faktor kesenangan dan keuntungan bagi pada pengembang <i>video game</i> . Tergantung daripada faktor-faktor tersebut, pengetahuan yang didapat daripada sekumpulan orang akan mengijinkan desainer untuk mengkustomisasi <i>video game</i> , yang nantinya ditujukan untuk menghasilkan keuntungan, semenjak nantinya <i>video game</i> melibatkan sejumlah besar orang yang memainkannya.
3	Aparna S. Varde, et al.	2004	Apriori	<p>Algoritma Apriori dan proses Game-of-Life telah digunakan sebagai dasar untuk analisis prediktif untuk membangun alat yang disebut QuenchMiner untuk dukungan keputusan dalam Ilmu Material.</p> <p>Ini menyediakan estimasi parameter dan simulasi mikro untuk panas mengobati proses. Alat ini adalah yang pertama dari jenisnya mengintegrasikan penambangan data yang bergantung pada domain dan visualisasi data untuk mendukung keputusan proses teknik mesin. Karena ini adalah alat berbasis web, ia menghubungkan klien di seluruh dunia dan memungkinkan mereka untuk bertukar pengetahuan yang bermanfaat untuk mendukung keputusan bisnis. Ini mempromosikan E-Bisnis dalam Materi Ilmu.</p>

BAB III

OBJEK DAN METODOLOGI PENELITIAN

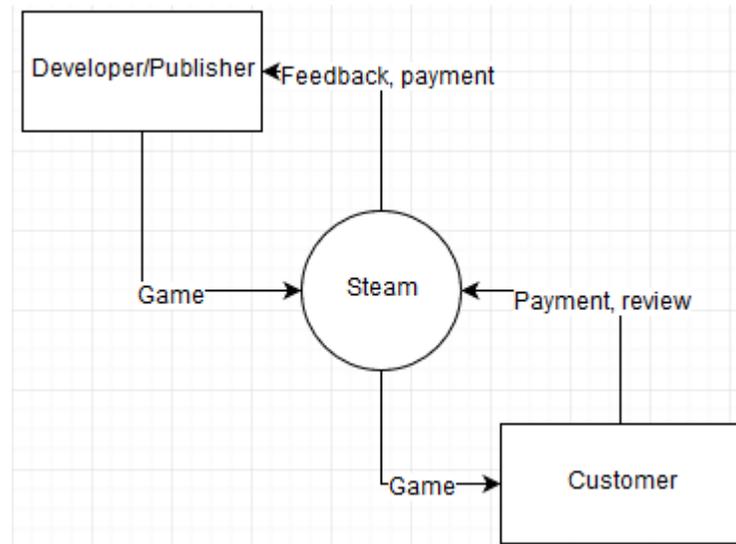
3.1 Profil Steam

Sejarah Steam dimulai pada saat Valve Corporation menghadapi masalah dalam memperbaharui *video game online* yang mereka publikasikan, seperti Counter-strike; dimana pengadaan *patch* akan mengakibatkan sebagian besar *user* untuk tidak bisa terhubung pada *server* dalam beberapa hari. Valve memutuskan untuk membuat sebuah *platform* yang memiliki kapabilitas untuk memperbaharui *game* secara otomatis, serta dilengkapi dengan sistem anti-pembajakan dan anti-kecurangan yang lebih kuat. Lewat pemberian suara daripada *user* mengenai peluncuran Steam pada tahun 2002, Valve juga menyadari bahwa setidaknya 75% daripada pengguna memiliki akses internet berkecepatan tinggi, yang mana akan terus berkembang dari segi kecepatan maupun luas jangkauan seiring bertambahnya tahun, dan menyadari bahwa Valve dapat memberikan konten *video game* lebih cepat kepada para pemain atau *user* daripada lewat saluran ritel lainnya. Valve mendekati beberapa perusahaan seperti: Microsoft, Yahoo!, dan RealNetworks untuk membangun platform tersebut, namun ditolak.

Pengembangan Steam dimulai pada tahun 2002, dan diberitahukan kepada publik pada tanggal 22 Maret, 2002, dan dirilis sebagai *beta* pada hari yang sama. *Client* Steam secara resmi keluar dari status *beta* pada tanggal 11 September 2003. Pada saat itu, fungsi utama Steam adalah melakukan optimisasi proses *patch* daripada *video game online*. Steam merupakan komponen opsional pada seluruh *game* lainnya. Diantara 80000 hingga 300000 *user* menguji sistem pada saat kondisi *beta*, namun hal tersebut menyebabkan kendala performa daripada sistem dan situs, dimana ribuan *user* mencoba untuk memainkan versi terbaru daripada *Counter-Strike*. Pada tahun 2004, World Opponent Network, jasa layanan *video game online*, tutup dan digantikan oleh Steam. Fitur *online* daripada *video game* yang membutuhkan World Opponent Network berhenti bekerja, dengan pengecualian dilakukan konversi menjadi Steam.

Sekitar tahun 2004, Valve mulai melakukan negosiasi kontrak dengan beberapa penerbit dan pengembang *video game* independen untuk mengeluarkan produk mereka, meliputi *Rag Doll Kung Fu* dan *Darwinia*, pada Steam. Strategy First, penerbit asal Kanada mengumumkan bahwa pada bulan Desember 2005, mereka akan mengadakan kerjasama dengan Valve untuk kepentingan distribusi digital daripada *video game* di masa yang akan datang. Pada tahun 2002, Gabe Newell, direktur daripada Valve, mengatakan bahwa dia menawarkan lisensi dan distribusi *game engine* pada *mod team* seharga 955 Dollar. *Half-Life 2* merupakan *video game* pertama yang membutuhkan Steam *client* untuk dapat dimainkan, bahkan dapat bentuk salinan ritelnya. Keputusan ini dibuat berdasarkan hasil pertimbangan mengenai kepemilikan perangkat lunak, kebutuhan perangkat lunak, dan isu-isu lainnya mengenai beban *server* yang terlalu besar. Pada waktu ini, *user* mengalami bermacam-macam masalah dalam berusaha untuk memainkan *game*.

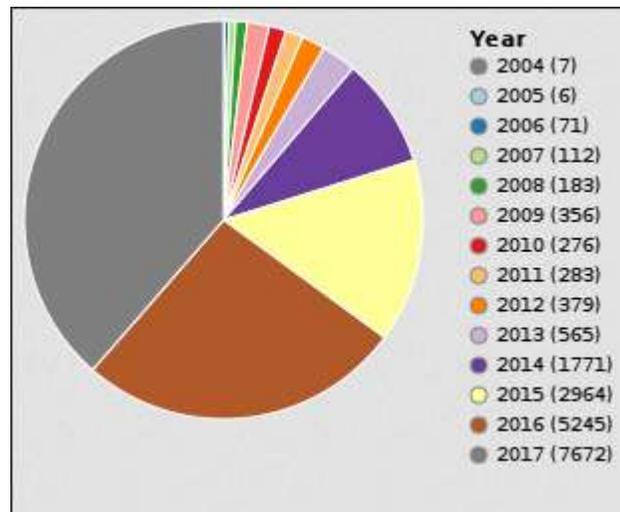
Dimulai dengan *Rag Doll Kung Fu* pada bulan Oktober 2005, *video game* pihak ketiga menjadi dapat dibeli dan diunduh pada Steam, dan Valve mengumumkan bahwa Steam menghasilkan untung yang disebabkan oleh beberapa *video game* buatan Valve yang sukses. Meskipun distribusi digital belum dapat menyaingi volume ritel, garis laba untuk Valve dan pengembang lain lebih besar pada Steam. Pengembang sekaligus penerbit besar, seperti id Software, Eidos Interactive, dan Capcom, memulai menyebarkan *video game* mereka lewat Steam pada tahun 2007. Pada bulan Mei di tahun yang sama, terdapat 13 juta akun yang dibuat, dan 150 judul *video game* dijual pada Steam. Hingga tahun 2014, total penjualan *video game* Steam diperkirakan mencapai 1,5 juta Dollar. Semenjak tahun 2007, Valve telah meneruskan dan meluaskan fungsionalitas dan layanan Steam untuk konsumen dan pengembang.



Gambar 3.1
DFD Steam.

Gambar 3.1 menunjukkan alur data diantara pengembang dan konsumen, dan sebuah sistem, yaitu Steam. Steam merupakan wadah yang menjembatani antara pihak pengembang yang disebut sebagai *developer* atau *publisher*, dan pihak yang memiliki permintaan atas *video game* yang dikembangkan.

Pihak pengembang membuat *video game* dengan spesifikasi tertentu, termasuk kebutuhan minimum mengenai perangkat keras yang dibutuhkan dan deskripsi dalam bentuk paragraf yang menjelaskan *video game* yang disajikan. *Video game* yang hendak dimasukkan pada Steam dapat bersifat *early access*, yaitu status dimana *video game* yang disajikan dapat dimainkan dengan baik, namun tidak memungkinkan untuk dapat menyelesaikan judul tersebut karena *video game* tersebut belum selesai dikembangkan. Pihak pengembang menerima pembayaran dan respon dan penilaian pemain terhadap *video game* yang ditawarkan lewat Steam. Transaksi pembayaran dapat dilakukan melalui kartu kredit, WebMoney, atau Bitcoin.



Gambar 3.2
Grafik *Video Game Steam*.

(https://www.gamasutra.com/view/news/312843/7672_games_hit_Steam_in_2017_alone_says_Steam_Spy.php)

Berdasarkan gambar 3.2, terdapat peningkatan jumlah *video game* yang dipublikasikan pada Steam dari tahun ke tahun, dimulai dari tahun 2004, tahun dimana Valve mulai melakukan negosiasi kontrak dengan beberapa penerbit dan pengembang *video game* independen untuk mengeluarkan produk mereka, meliputi *Rag Doll Kung Fu* dan *Darwinia*, pada Steam. Tahun 2017 merupakan tahun dimana Steam menerbitkan jumlah *video game* yang terbesar, yakni 7672 judul.

Karena banyaknya judul yang terbit tahun lalu, terdapat pula pola-pola tertentu yang dapat ditelusuri berdasarkan kombinasi nilai daripada parameter yang menjadi informasi daripada setiap judul pada *video game*. Pola-pola tersebut ditelusuri menggunakan metode asosiasi, atau *association rule* yang biasanya digunakan untuk analisis keranjang belanja. Pada penelitian ini, *association rule* digunakan untuk mencari tren tertentu yang ditawarkan oleh setiap judul *video game* yang diterbitkan pada tahun 2017. Pola-pola tersebut merupakan kombinasi daripada genre dari *video game*, tipe keterlibatan pemain pada *video game*, dan fitur yang ditawarkan pada setiap judul.

Dalam industri *video game*, distribusi digital adalah sebuah proses pengiriman konten *video game* sebagai informasi digital, tanpa pergantian atau pembelian media fisik yang baru. Proses ini telah dimulai dari tahun 1980, namun pada tahun 2000 distribusi digital menjadi lebih populer dalam metode penjualan *video game* karena kapabilitas *bandwidth* dan teknologi jaringan pada masa itu. Saat ini, proses ini didominasi oleh distribusi *online* menggunakan internet.

Distribusi digital daripada *video game* menjadi lebih umum dewasa ini, dimana penerbit dan pemilik ritel besar memperhatikan lebih terhadap penjualan secara digital, termasuk Steam, PlayStation Store, Amazon.com, GAME, GameStop, dan lain sebagainya. Berdasarkan studi yang dilakukan oleh SuperData Research, volume distribusi digital *video game* di seluruh dunia, mencapai 6,2 juta Dollar per bulan pada Februari 2016, dan mencapai 7,7 juta Dollar pada bulan April 2017.

Tabel 3.1
Tabel Parameter *Video Game* Steam.

No.	Field	Tipe Data
1	title	String
2	release date	Date
3	agecheck	Boolean
4	price	Integer
5	genre	String
6	developer	String
7	publisher	String
8	player type	String
9	feature	String
10	system requirements	String
11	all reviews	String

Dengan kondisi data sebelum melalui *data preprocessing*, terdapat 11 *field* pada tabel 3.1 yang menjadi bagian dalam *record* sebuah *video game* pada Steam. Sebagian *field* bersifat alfanumerik atau komposit, yang nantinya dapat diturunkan menjadi beberapa *field* yang lebih rinci.

1. Title

Title adalah judul daripada *video game* yang ditampilkan, bertipe data String.

2. Release Date

Release date adalah tanggal terbit daripada *video game* yang bersangkutan, bertipe data Date dengan format dd/mm/yyyy.

3. Agecheck

Agecheck adalah pemberitahuan daripada Steam kepada *user* apakah *user* sudah cukup umur untuk memainkan *video game* yang bersangkutan, bertipe data Boolean. Sebelum *user* dapat mengakses *video game*, Steam akan menanyakan tanggal lahir daripada *user* atau memberikan notifikasi bahwa konten yang disajikan *video game* ditujukan untuk pemain dewasa.

4. Price

Price adalah harga *video game* yang awalnya ditampilkan dalam kurs Dollar Amerika, namun secara otomatis dikonversi oleh Steam menjadi kurs masing-masing *region* daripada *user*, bertipe data Integer.

5. Genre

Genre adalah aliran atau gaya bermain yang disajikan sebuah *video game*, bertipe data String. Pada Steam, terdapat 10 genre yang dikategorikan: *Action*, *Adventure*, *Casual*, *Indie*, *Massively Multiplayer*, *Racing*, *RPG*, *Simulation*, *Sports*, dan *Strategy*. Sebuah judul *video game* dapat memiliki lebih dari satu jenis genre dengan kombinasi genre lainnya.

6. Developer

Developer adalah nama entitas pengembang daripada *video game*, bertipe data String. Nama *developer* dapat sama dengan *publisher*.

7. Publisher

Publisher adalah nama entitas penerbit daripada *video game*, bertipe data String. Nama *publisher* dapat sama dengan *developer*.

8. Player Type

Player type adalah tipe keterlibatan pemain atau *user* pada saat bermain sebuah *video game*, bertipe data String. Pada Steam, terdapat 4 tipe keterlibatan pemain: *Single*, *Multi*, *Local*, dan *Online*.

9. Feature

Feature adalah fitur yang dimiliki daripada *video game* yang bersangkutan, bertipe data String. Pada Steam, terdapat 8 fitur yang dikategorikan: *Achievements*, *Controller Support*, *Workshop*, *Trading Cards*, *Cloud*, *Level Editor*, *Captions*, dan *Leaderboards*.

10. System Requirements

System Requirements adalah kebutuhan dasar daripada perangkat keras yang dibutuhkan untuk dapat memainkan *video game*, bertipe data String. Kebutuhan dasar yang dispesifikasikan oleh Steam adalah sistem operasi, prosesor, memori, kartu grafik, versi DirectX, koneksi internet, ukuran *video game*, dan kartu suara yang dibutuhkan.

11. All Reviews

All Reviews adalah ulasan secara keseluruhan mengenai *video game* yang ditampilkan, bertipe data String. Pada Steam, ulasan dengan nilai terendah diberi label “*Overwhelmingly Negative*”, sedangkan ulasan dengan nilai tertinggi diberi label “*Overwhelmingly Positive*”.

Tabel 3.2
Tabel Parameter *Video Game* Steam, Setelah *Data Preprocessing*

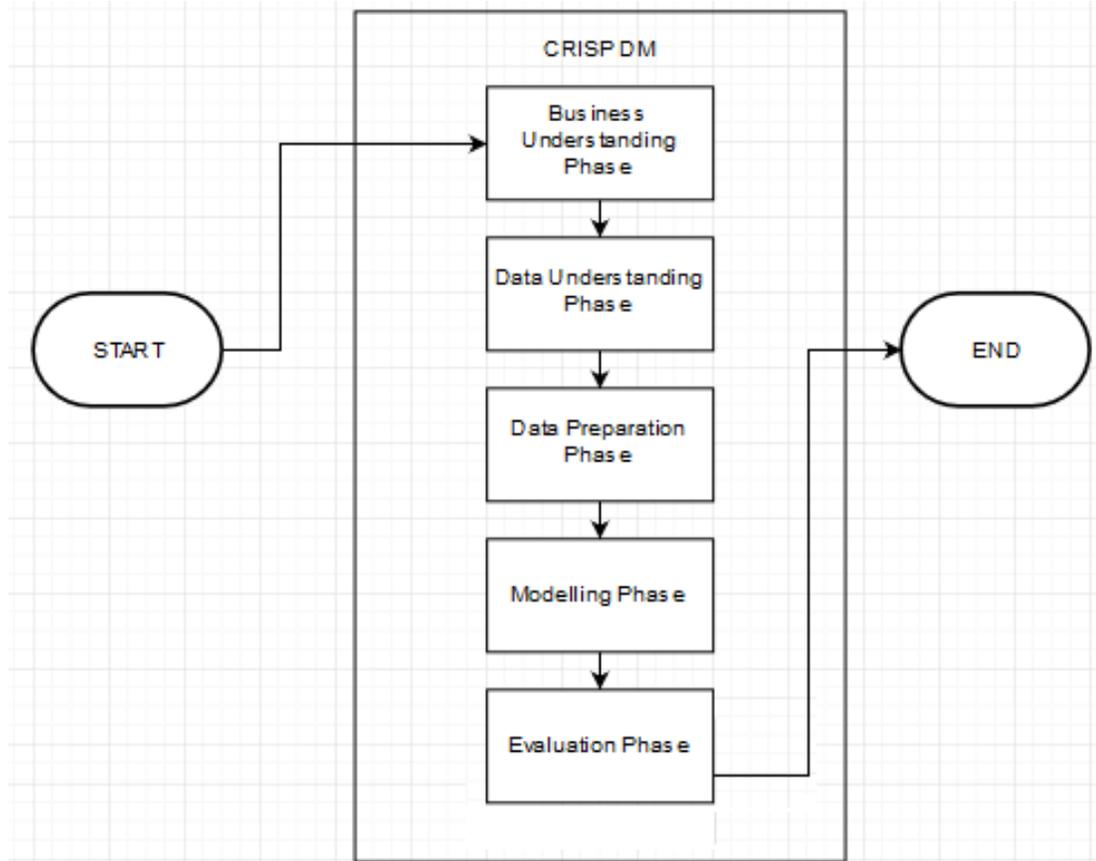
No.	Field	Tipe Data
1	title	String
2	release date	Date
3	price	Integer
4	developer	String
5	publisher	String
6	all reviews	String
7	agecheck	Boolean
8	action	Boolean
9	adventure	Boolean
10	casual	Boolean
11	indie	Boolean
12	massively multiplayer	Boolean
13	racing	Boolean

No.	Field	Tipe Data
14	rpg	Boolean
15	simulation	Boolean
16	sports	Boolean
17	strategy	Boolean
18	single	Boolean
19	multi	Boolean
20	local	Boolean
21	online	Boolean
22	achievements	Boolean
23	controller support	Boolean
24	workshop	Boolean
25	trading cards	Boolean
26	cloud	Boolean
27	level editor	Boolean
28	captions	Boolean
29	leaderboards	Boolean

Pada tabel 3.2, beberapa parameter yang memiliki tipe data String dilakukan konversi menjadi tipe data Boolean. Sebagai contoh, parameter “player type” dijabarkan menjadi “single”, “multi”, “local”, dan “online”. Konversi ini merupakan langkah penting dalam proses analisis metode asosiasi, karena tipe data Boolean merupakan bentuk yang tepat yang mendekati format data tabular, yang nantinya digunakan untuk merpresentasikan pola-pola yang terdapat pada setiap judul *video game*. Parameter “system requirement” tidak dimasukkan pada analisis ini karena isinya yang terdiri atas tipe-tipe data lain, tidak terbatas hanya String saja. Tipe-tipe data lain ada yang numerik dan tidak dapat dilakukan *data discretization* atau perubahan menjadi tipe data Boolean.

3.2 Metodologi Penelitian

Pada bagian ini akan dijelaskan mengenai metode penelitian yang digunakan untuk membantu penulis dalam melakukan penelitian. Metodologi penelitian, CRISP-DM, meliputi enam tahap yang harus dikerjakan, dimulai dari pemahaman bisnis (*Business Understanding Phase*) hingga penyebaran (*Deployment Phase*).



Gambar 3.3
Metodologi Penelitian

Merujuk pada subbab 2.1.3. mengenai siklus *data mining*, terdapat 6 fase dalam melakukan penelitian, dimana setiap fase memiliki fungsi tersendiri untuk memberikan kontribusi terhadap proses pengerjaan *data mining*, dimulai dari:

1. *Business Understanding Phase*

Tujuan bisnis yang akan dicari dalam penelitian masalah adalah ditemukannya pengetahuan berupa pola keterhubungan yang terjadi dalam setiap data *video game* pada Steam. Dengan mengetahui pola tersebut, selanjutnya akan diketahui genre-genre dan fitur-fitur yang dibeli secara bersamaan. Informasi tersebut berikutnya akan menghasilkan saran yang dibutuhkan kepada para *game developer* untuk menentukan elemen apa yang sebaiknya dimasukkan dalam pembuatan *video game* selanjutnya.

2. *Data Understanding Phase*

Pemahaman data atau *data understanding phase*, adalah tahapan pengumpulan data awal dan mempelajari data tersebut untuk dapat mengenal dan memahami apa saja yang dapat dilakukan terhadap data-data tersebut. Pemahaman data mengacu pada *record* pada *video game* yang didapatkan dari Steam. Pada tahap ini akan dilakukan analisis terhadap tabel yang terlibat dalam pengolahan data.

3. *Data Preparation Phase*

Persiapan data atau *data preparation phase*, berisi mengenai penetapan struktur basis data yang akan digunakan untuk mempermudah proses *data mining*. Persiapan data dibagi menjadi 3 bagian, yaitu pemilihan data, pemrosesan data, dan perubahan data. Pemilihan data dilakukan untuk memilih atribut yang digunakan dalam *data mining*. Data *video game* yang dipilih adalah sampel acak yang diperoleh pada tahun 2017, dimulai dari tanggal 1 Januari 2017 hingga 31 Desember 2017. Jumlah sampel acak yang dipilih adalah 2000.

4. *Modelling Phase*

Sebelum proses pemodelan berlangsung, proses *data mining* terhadap data yang terpilih dilakukan terlebih dahulu. Pemrosesan data dilakukan guna menentukan kualitas data yang akan digunakan dalam *data mining*. Pada tahapan ini akan dilakukan pemilihan data yang telah terisi dengan benar. Beberapa langkah yang diambil sebagai pemroses data adalah data yang tidak lengkap akan dihapus, kolom yang berisi atribut yang tidak

relevan dengan penelitian yang dilakukan, dan membuat keseragaman data sesuai dengan yang disepakati. Pengubahan data adalah langkah konversi atribut yang memiliki tipe data String dan memiliki nilai-nilai yang pasti, menjadi tipe data Boolean dalam format data tabular. Tujuan pengubahan, atau transformasi data, adalah untuk mengubah data yang sebelumnya tidak dapat dibaca, menjadi dapat dibaca oleh perangkat lunak. *Data mining* adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma yang digunakan yang tepat sangat bergantung pada tujuan dan proses pencarian pengetahuan secara keseluruhan. Dalam penelitian ini, algoritma daripada teknik *association rule* yang digunakan adalah Apriori dan FP-Growth. Tahapan yang dilakukan setelah data terpilih adalah menentukan nilai *support* dan *confidence*.

5. *Evaluation Phase*

Tahap akhir dalam metodologi penelitian ini menghasilkan sebuah pola informasi. Pola informasi yang dihasilkan daripada proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak berkepentingan, dalam hal ini, para pengembang *video game*. Proses evaluasi meliputi pemilihan aturan *association rule* yang memenuhi nilai *confidence* setidaknya 90% sebagai aturan asosiasi yang kuat (*recommended*), lalu melakukan evaluasi terhadap aturan asosiasi yang dihasilkan. Hasil tersebut adalah *rule* yang dihasilkan dari proses *data mining* selanjutnya, dan akan disampaikan kepada para *game developer* untuk menentukan genre dan fitur apa saja yang direkomendasikan untuk ditambahkan pada sebuah *game*.

3.3 Alat Penelitian

Adapun spesifikasi perangkat keras, dalam bentuk *laptop* yang digunakan untuk kebutuhan penelitian ini:

Tabel 3.3
Tabel Spesifikasi Perangkat Keras

Sistem operasi	Windows 10 Pro 64-bit (10.0, Build 16299)
Perusahaan	LENOVO
Model Sistem	80E1
BIOS	InsydeH2O Version CCB.03.74.03A2CN22WW(v1.04)
Prosesor	AMD A6-6310 APU with AMD Radeon R4 Graphics
Memory	10240MB RAM
DirectX	DirectX 12

Aplikasi yang digunakan dalam penelitian antara lain:

1. Microsoft Excel 2017

Microsoft Excel 2017 digunakan sebagai wadah penampung daripada data mentah dari Steam yang belum melalui *data preprocessing*, dan menjadi tempat terjadinya *data preprocessing* daripada data yang akan digunakan oleh aplikasi *data mining* lain seperti IBM SPSS Statistics 24, IBM SPSS Modeler 18.0, dan Weka.

2. IBM SPSS Statistics 24

IBM SPSS Statistics 24 digunakan sebagai tahap tambahan untuk menyesuaikan spesifikasi data yang akan digunakan. Setelah melalui *data preprocessing* menggunakan Microsoft Excel, tahap selanjutnya adalah membuat *dataset* dengan format *.sav* yang nantinya akan lebih mudah dibaca oleh IBM SPSS Modeler 18.0. Beberapa spesifikasi data yang harus disesuaikan adalah tipe data, metode pengukuran (*scale*, *ordinal*, atau *nominal*), dan *role* (*input*, *target*, atau *both*).

3. IBM SPSS Modeler 18.0

IBM SPSS Modeler 18.0 digunakan untuk melakukan *data mining* terhadap *dataset* yang sudah melalui tahap *data preprocessing*. Algoritma yang akan digunakan adalah Apriori.

4. Weka 3.8

Weka 3.8 digunakan untuk melakukan *data mining* terhadap *dataset* yang sudah melalui tahap *data preprocessing*. Algoritma yang akan digunakan adalah Apriori dan FP-Growth.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Business Understanding Phase

Pada tahun 2017, diketahui bahwa jumlah judul *video game* yang terdaftar pada situs, merupakan yang tertinggi dibandingkan dengan tahun-tahun sebelumnya, yakni mencapai 7672 judul. *Stakeholder* utama adalah store.steampowered.com, yang mana merupakan situs yang menyediakan *platform* untuk distribusi digital *video game*. *Video game* yang menjadi sampel acak diambil dimulai dari tanggal 1 Januari 2017 hingga 31 Desember 2017, yang mana memiliki cakupan *official* dan *indie* pada keseluruhan judul yang terdaftar. Tujuan bisnis daripada penelitian ini adalah untuk mengetahui tren genre *video game* yang mendominasi pada tahun 2017.

Tujuan analisis dari penelitian ini adalah untuk mencari kecenderungan genre yang saling berkaitan dengan genre lainnya pada setiap judul *video game*. Pengambilan data dilakukan secara manual dari store.steampowered.com ke Microsoft Excel sebagai *software* pendukung. Sesuai dengan lampiran A, proses pembersihan data dilakukan dan terdapat 2000 *record* yang siap diproses, berikut dengan 29 *field* yang terdapat pada masing-masing *record*, dimana 23 *field* yang terkandung merupakan variabel yang akan digunakan pada proses *data mining*.

Pemilihan data dilakukan dengan memilih sampel acak sebesar 2000 *record* yang nantinya akan digunakan untuk keperluan penelitian. Terlepas dari banyaknya *video game* yang masuk dan/atau dapat diambil, terdapat beberapa kriteria yang digunakan dalam pemilihan *video game*. *Video game* yang diambil merupakan yang bersifat *top seller*, dengan alasan *top seller* memiliki informasi yang lebih lengkap dibandingkan *video game* lainnya. Kriteria lainnya adalah data yang diambil secara eksklusif merupakan *video game* dan tidak termasuk pada *software* lainnya, serta sistem operasi yang berlaku hanya terbatas pada Windows.

Pemilihan dilakukan dengan sumber data diperoleh dari tabel *video game* dari situs resmi Steam (store.steampowered.com) periode Januari 2017 hingga Desember 2017. *Stakeholder* utama adalah store.steampowered.com, sebuah situs yang menyediakan *platform* untuk distribusi digital *video game*, atau dikenal secara umum dengan nama Steam. Secara umum, *video game* yang disediakan terbagi atas 2 jenis, yaitu yang bersifat *official*, dan yang bersifat *indie*. Perbedaan yang mendasar antara 2 jenis tersebut adalah pihak yang mengeluarkan dan membangun *video game* tersebut.

Dari sekian judul yang ditawarkan dari sejumlah besar nama *publisher* dan *developer*, terpilih sampel acak yang terdaftar pada situs dimulai dari tanggal 1 Januari 2017 hingga 31 Desember 2017. Tujuan bisnis daripada penelitian ini, adalah untuk mengetahui tren genre daripada *video game*, baik yang diproduksi secara *official* maupun *indie*. Tujuan lainnya adalah untuk menciptakan perbandingan performa yang ditunjukkan dua algoritma *association rule*, yaitu Apriori dan FP-Growth. Data yang diambil secara eksklusif diperoleh dari store.steampowered.com. Jumlah keseluruhan data yang terlampir berjumlah sekitar 2000 *record*. Setelah pengumpulan selesai, dilakukan *data preprocessing* untuk menyesuaikan format data untuk nantinya dapat dibaca dan/atau diproses oleh aplikasi.

Pengumpulan sampel acak setidaknya dilakukan selama 2 minggu, dimana per hari terhitung rata-rata pengumpulan informasi daripada 150 judul *video game*. Setelah pengumpulan informasi selesai dilakukan, tahap selanjutnya adalah pembersihan data dan *data preprocessing* yang dapat memakan waktu dari 1 hingga 2 minggu, tergantung daripada kompleksitas data yang digunakan dalam penelitian. Untuk penelitian ini, tahap persiapan data diestimasi mencapai waktu 1 bulan.

4.2 Data Understanding Phase

Association rule merupakan teknik *data mining* untuk menemukan pola hubungan antar *item*. Dalam hal ini, yang dianalisis adalah hubungan antar genre, keterlibatan jumlah pemain, dan fitur dalam tabel *video game* melalui 2 parameter: *minimum support* dan *minimum confidence*. Eksplorasi data mengacu pada penjelasan kumpulan *field* yang terkandung pada tabel *video game* yang menjadi objek penelitian. Terdapat 29 *field* secara keseluruhan, dan terdapat 23 *field* yang merupakan hasil perubahan dari format String menjadi Boolean. Kumpulan *field* tersebut dibagi menjadi 4 kategori.

Tabel 4.1
Tabel Kategori *Itemset* Berdasarkan Subjek Analisis

Subjek Analisis	Itemset
Informasi Dasar	Title, Release Date, Price, Developer, Publisher, All Reviews
Genre	Action, Adventure, Casual, Indie, Massively Multiplayer, Racing, RPG, Simulation, Sports, Strategy
Tipe Keterlibatan Pemain	Single, Multi, Local, Online
Fitur	Agecheck, Achievement, Controller Support, Workshop, Trading Card, Cloud, Level Editor, Captions, Leaderboards

4.2.1 Informasi Dasar

Informasi dasar adalah tulisan ringkas mengenai *video game* yang terdiri atas tulisan yang tidak dapat dikategorikan karena sifatnya yang unik.



Gambar 4.1
Informasi Dasar *Cuphead* pada Situs Resmi Steam
(<https://store.steampowered.com/app/268910/Cuphead/>)

Dalam penelitian ini, terdapat enam informasi dasar yang ada pada setiap judul *video game* yang dilampirkan.

1. Title

Title merupakan judul daripada *video game*. Dalam penulisannya, judul pada sebuah *video game* dapat ditulis dengan bahasa Inggris, bahasa Jepang, bahasa Mandarin, dan bahasa lainnya.

2. Release Date

Release date merupakan tanggal diterbitkannya sebuah *video game* pada Steam. Waktu yang dibutuhkan dalam memroses *video game* yang masuk pada Steam adalah 3 hingga 5 hari.

3. Price

Harga yang tercantum langsung disesuaikan dengan mata uang yang digunakan pada *region* tertentu. Secara *default*, harga ditulis dalam kurs USD.

4. Developer

Developer merupakan nama pihak atau kelompok yang bertanggung jawab atas pengembangan *video game* yang bersangkutan. *Developer* dan *publisher* dapat menjadi pihak yang sama atau berbeda.

5. Publisher

Publisher merupakan nama pihak atau kelompok yang bertanggung jawab atas penerbitan *video game* yang bersangkutan pada Steam. *Developer* dan *publisher* dapat menjadi pihak yang sama atau berbeda.

6. All Reviews

All reviews merupakan hasil keseluruhan penilaian orang-orang yang memainkan *video game* tertentu, dalam format data String. *Review* yang akan muncul adalah: *Overwhelmingly Negative*, *Very Negative*, *Negative*, *Mostly Negative*, *Mixed*, *Mostly Positive*, *Positive*, *Very Positive*, dan *Overwhelmingly Positive*.

4.2.2 Genre

Genre adalah jenis, atau ciri-ciri umum daripada sebuah *video game* yang dapat jatuh dalam kategori tertentu. Pada situs resmi Steam, genre *video game* ditulis dalam format tertentu, dan dipisahkan dengan tanda koma.



Gambar 4.2
Genre *Cuphead* pada Situs Resmi Steam
(<https://store.steampowered.com/app/268910/Cuphead/>)

Dengan menjumlahkan kemunculan genre per judul, dapat diketahui bahwa judul dari *record* yang valid, memiliki variasi genre antara 1 hingga 10, dimana sebagian besar judul memiliki kombinasi 3 genre. Terdapat 10 genre yang menjadi kategori pada penelitian ini:

1. Action

Video game bergenre action, memberikan instruksi pada pemain untuk melakukan gerakan cepat, biasanya untuk menyerang atau menghindari serangan lawan. Terdapat dua sub tipe dari genre ini, yaitu shooter game dan fighting game. Shooter game melibatkan perlawanan yang menggunakan senjata api atau senjata lainnya yang digunakan untuk menembak. Fighting game melibatkan serangan fisik pada lawan hingga lawan kalah. Beberapa contoh dari genre action pada Steam, adalah *Zombie Estate 2*, *I Am The Hero*, dan *Pinkman*.

2. Adventure

Video game bergenre *adventure*, berfokus pada pemecahan teka-teki dibandingkan dengan aksi maupun tantangan. Biasanya, pemecahan teka-teki tersebut dapat dilakukan dengan berinteraksi dengan karakter lainnya pada lingkungan *game*. Beberapa contoh

dari genre *adventure* pada Steam, adalah *Panty Party*, *Resident Evil 7: Biohazard*, dan *Mainlining*.

3. Casual

Video game bergenre *casual*, adalah tipe *game* yang memiliki aturan yang mudah dan tidak membutuhkan waktu yang panjang dan/atau kemampuan tertentu untuk dapat memainkan *game* tersebut. Dalam proses pengembangannya, *game* bergenre *casual* biasanya tidak membutuhkan biaya produksi dan distribusi yang rendah. Beberapa contoh dari genre *casual* pada Steam, adalah *Outrunner*, *NEKO-NIN exHeart*, dan *aMAZE 2*.

4. Indie

Video game yang dibuat oleh pengembang *indie* berbeda dengan yang dibuat oleh perusahaan resmi. Salah satu ciri yang sering ditemui pada proses pengembangan *indie game*, adalah tidak adanya saluran dana oleh pihak penerbit atas pengembangan tersebut. Beberapa contoh dari genre *indie* pada Steam, adalah *Outlast 2*, *Immortal Redneck*, dan *I Expect You to Die*.

5. Massively Multiplayer

Video game bergenre *massively multiplayer*, menandakan bahwa *game* dapat dimainkan lebih dari satu orang, dan biasanya memerlukan koneksi internet untuk dapat bermain, secara opsional maupun mandatori. Beberapa contoh dari genre *massively multiplayer* pada Steam, adalah *CarX Drift Racing Online*, *Grey Hack*, dan *OrbusVR*.

6. Racing

Video game bergenre *racing*, merupakan subgenre daripada genre *sports*. Genre ini berfokus pada perlombaan adu kecepatan, pacuan, atau balapan. Tergantung daripada tema visual dan *gameplay* yang ditawarkan, *game* bergenre *racing* menampilkan sejumlah peralatan atau kendaraan yang akan digunakan pada saat permainan dimulai, yang nantinya akan dipilih oleh pemain. Beberapa contoh dari genre *racing* pada Steam, adalah *TO THE TOP*, *Hover*, dan *Crashday Redline Edition*.

7. RPG

Video game bergenre RPG, merupakan sebuah *game* dimana pemain berperan menjadi karakter tertentu pada jalan cerita yang sudah ditentukan sebelumnya. Tergantung daripada karakter yang diperankan atau dimainkan, setiap karakter memiliki keunikan tersendiri dilihat daripada kemampuan bertarung maupun berinteraksi dalam sebuah cerita. Pemain biasanya berpetualang dan menyelesaikan misi untuk mendapatkan jalan menuju lokasi lain seperti istana, *dungeon*, dan kepemilikan benda dengan jumlah yang cukup besar seperti senjata maupun uang. Beberapa contoh dari genre RPG pada Steam, adalah *Has-Been Heroes*, *Weapon Shop Fantasy*, dan *Fallout Shelter*.

8. Simulation

Video game bergenre *simulation*, merupakan sebuah simulasi aktivitas berdasarkan kehidupan nyata maupun fiksi. Genre dari *game* ini memungkinkan pemain untuk melakukan analisis terhadap objek-objek yang terlibat di dalam *game*. Beberapa contoh dari genre *simulation* pada Steam, adalah *King Kaiju*, *Angels with Scaly Wings*, dan *We Need to Go Deeper*.

9. Sports

Video game bergenre *sports*, berfokus pada *gameplay* yang mensimulasikan olahraga yang bersifat tradisional, seperti sepakbola maupun bulutangkis. Beberapa *game* dari *genre* ini berfokus pada aturan olahraga sebenarnya, atau strategi yang dilakukan pada saat permainan berlangsung. Beberapa contoh dari genre *sports* pada Steam, adalah *Disc Jam*, *The Cat Games*, dan *Out of the Park Baseball 18*.

10. Strategy

Video game bergenre *strategy*, adalah sebuah permainan yang membutuhkan pemikiran dan perencanaan yang penuh dengan perhitungan dan ketelitian untuk memenangkan permainan tersebut. Cakupan aktivitas dari permainan ini memiliki keanekaragaman dimulai dari penguasaan dunia hingga taktik yang melibatkan sejumlah orang. Beberapa contoh dari genre *strategy* pada Steam, adalah *Weed Shop 2*, *Domina*, dan *Warstone TD*.

Adapun statistik daripada *genre* yang muncul pada setiap judul pada tahun 2017:

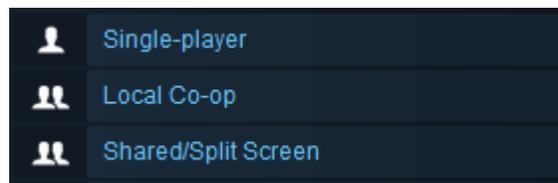
Tabel 4.2
Tabel Jumlah Kemunculan Genre

Genre	Jumlah Kemunculan
Action	925
Adventure	806
Casual	720
Indie	1429
Massively Multiplayer	92
Racing	72
RPG	352
Simulation	483
Sports	128
Strategy	427

Genre yang paling banyak muncul pada tahun 2017 adalah *indie* yakni pada 1426 judul, sedangkan yang paling sedikit adalah *racing* yakni pada 72 judul.

4.2.3 Tipe Keterlibatan Pemain

Tipe keterlibatan pemain adalah sebuah indikator mengenai seberapa banyak pemain yang dapat berpartisipasi dalam memainkan sebuah *video game*.



Gambar 4.3
Tipe Keterlibatan Pemain yang Ditawarkan *Cuphead*
(<https://store.steampowered.com/app/268910/Cuphead/>)

Dengan menjumlahkan tipe keterlibatan pemain per judul, dapat diketahui bahwa judul dari *record* yang valid, memiliki variasi tipe keterlibatan pemain 1 hingga 4, dimana kebanyakan judul memiliki setidaknya 1 keterlibatan pemain, atau *single-player*.

Terdapat 4 tipe keterlibatan pemain yang menjadi kategori pada penelitian ini:

1. Single

Single-player menandakan bahwa *game* dapat, atau hanya dapat dimainkan oleh satu orang saja. Beberapa contoh dari *single-player game* pada Steam, adalah *Sakura no Mori – Dreamers, Change*, dan *VODKA*.

2. Multi

Multiplayer menandakan bahwa *game* dapat dimainkan lebih dari satu orang, atau setidaknya melibatkan dua orang pemain. Beberapa contoh dari *multiplayer game* pada Steam, adalah *Fire Pro Wrestling World, Heart of Crown PC*, dan *Stationeers*.

3. Local

Local multiplayer menandakan bahwa *game* dapat dimainkan dengan koneksi LAN (*Local Area Network*), biasanya melibatkan belasan orang pemain. Beberapa contoh dari *local multiplayer game* pada Steam, adalah *Sky Force Reloaded, Dominions 5 – Warriors of the Faith*, dan *Door Kickers: Action Squad*.

4. Online

Online menandakan bahwa *game* dapat dimainkan dengan melibatkan seluruh pemain yang memiliki koneksi internet di seluruh belahan dunia, biasanya melibatkan ribuan hingga jutaan orang pemain. Beberapa contoh dari *online game* pada Steam, adalah *Sparc, Tannenberg*, dan *Racket Fury: Table Tennis VR*.

Adapun statistik daripada tipe keterlibatan pemain yang muncul pada setiap judul pada tahun 2017:

Tabel 4.3
Tabel Jumlah Kemunculan Tipe Keterlibatan Pemain

Tipe Keterlibatan Pemain	Jumlah Kemunculan
Single	1894
Multi	340
Local	192
Online	371

Game dengan tipe pemain *single player* merupakan yang paling banyak muncul, yakni 1894 judul, sedangkan *game* yang menyediakan keterlibatan pemain pada area lokal merupakan yang paling sedikit, yakni sebanyak 192 judul.

4.2.4 Fitur

Fitur adalah karakteristik atau properti tertentu yang dimiliki sebuah *video game* yang memiliki kesamaan dengan *video game* lainnya.



Gambar 4.4
Fitur yang Ditawarkan *Cuphead*
(<https://store.steampowered.com/app/268910/Cuphead/>)

Dengan menjumlahkan kemunculan fitur per judul, dapat diketahui bahwa judul dari *record* yang valid, memiliki variasi fitur antara 0 hingga 7, dimana kebanyakan judul memiliki 2 fitur. Terdapat 9 fitur yang menjadi kategori dalam penelitian ini:

1. Agecheck

Fitur *agecheck* menandakan bahwa *game* yang dimainkan ditujukan bagi pemain dewasa. Tergantung dari batas umur yang ditentukan masing-masing negara, nilai batasan berkisar antara 15 hingga 21 tahun. Di Indonesia, kategori dewasa berkisar antara umur 15 hingga 18 tahun.

2. Achievements

Fitur *achievements* mengacu pada ada atau tidaknya pencapaian tertentu pada saat *game* dimainkan. Pencapaian tersebut bukan bagian utama dari *game*.

3. Controller Support

Fitur *controller support* mengacu pada penggunaan *joystick* atau kontroler lainnya selain *keyboard* dan *mouse* pada saat *game* dimainkan.

4. Workshop

Fitur *workshop* mengacu pada ada atau tidaknya *workshop* pada sebuah *game* dimana pemain dapat berpartisipasi untuk memberikan usul mengenai konten dalam *game*.

5. Trading Cards

Fitur *trading cards* mengacu pada ada atau tidaknya *trading card* yang didapat pemain pada saat *game* dimainkan. Berbeda dengan *achievements*, *trading card* dapat ditukar dan dijual-belikan kepada pemain lain sebagai kartu koleksi.

6. Cloud

Steam Cloud adalah layanan yang disediakan *Steam* untuk menyimpan data *game* dan file lainnya secara otomatis pada server Valve. Pemain dapat mengakses data ini menggunakan perangkat keras apapun yang menjalankan aplikasi *Steam*.

7. Level Editor

Level Editor adalah fitur dalam *game* yang memungkinkan pemain untuk membuat desain level pada *game* yang dimainkan. Desain level tersebut nantinya dimungkinkan untuk dimainkan oleh pemain lain.

8. Captions

Captions memungkinkan *game* yang dimainkan untuk mengganti bahasa pada dialog dalam *game*.

9. Leaderboards

Leaderboards adalah sistem urutan berdasarkan pemain terbaik dan/atau pemain dengan pencapaian tertinggi dalam memainkan *game* tertentu.

Adapun statistik daripada fitur yang muncul pada setiap judul pada tahun 2017:

Tabel 4.4
Tabel Jumlah Kemunculan Fitur

Fitur	Jumlah Kemunculan
Agecheck	329
Achievements	1268
Controller Support	780
Workshop	90
Trading Cards	861
Cloud	720
Level Editor	89
Captions	73
Leaderboards	255

Fitur yang paling sering muncul pada tahun 2017 adalah *achivements*, yakni terdapat pada 1268 judul. Fitur yang paling sedikit muncul pada tahun 2017 adalah *captions*, yakni terdapat pada 73 judul.

4.3 Data Preparation Phase

Setelah data mentah Steam diambil dan tercatat pada Microsoft Excel, tahap selanjutnya adalah *data preprocessing*. Secara garis besar, normalisasi data yang dilakukan meliputi standarisasi data dan perubahan *field* bertipe data String menjadi biner.

Pemilihan genre didasarkan pada genre yang dicantumkan pada situs resmi Steam. Genre yang tercantum tidak sama dengan *tag* yang tercantum.



Gambar 4.5
Perbandingan antara Genre dengan *Tag*
(<https://store.steampowered.com/app/268910/Cuphead/>)

Berdasarkan gambar 4.5, kolom kiri menunjukkan genre yang ditetapkan oleh Steam sebagai standar yang berlaku untuk seluruh *video game* yang masuk. Hanya terdapat sepuluh genre yang menjadi tetapan, sedangkan kategori lainnya jatuh sebagai *tag*.

Tidak semua *tag* sama dengan genre, begitu pula sebaliknya. *Tag* tidak ditetapkan oleh Steam, melainkan dimasukkan oleh *user* untuk mengkategorikan *video game* tertentu untuk memberikan spesifikasi mengenai konten yang disajikan *video game*. Awalnya, informasi yang terkandung dalam sebuah *video game* ditulis dalam format String, dan dipisahkan menggunakan tanda koma. Informasi dasar tidak dapat dikategorikan karena nilainya yang unik, berbeda dengan genre, tipe keterlibatan pemain, dan fitur sebuah *video game*.

Judul	Genre					
Delicious - Emily's Hopes and Fears	Adventure	Casual	Indie	Strategy		
Camper Jumper Simulator	Indie	Racing				
Blaite	Action	Adventure	Casual	Indie		
Word Killer: Zorgilonian Chronicles	Action	Casual	Indie	Simulation		
Nuclear Contingency	Action	Adventure	Indie			
UNO	Casual					
LSD Driver	Action	Casual	Indie	Racing	Simulation	Sports
Football Blitz	Action	Casual	Indie	Sports		
Redneck Rampage	Action	Adventure	Casual	RPG		
Milkmaid of the Milky Way	Adventure	Indie				

Gambar 4.6
Data Sebelum Mengalami *Data Preprocessing*.

Genre, tipe keterlibatan pemain, maupun fitur yang tertulis dalam format String akan diubah menjadi data biner dengan memberikan tanda “1” apabila muncul, sedangkan “0” apabila tidak muncul dalam kategori tertentu.

Judul	Adventure	Casual	Indie	Strategy	Racing	Simulation	Action	Sports	RPG
Delicious - Emily's Hopes and Fears	1	1	1	1	0	0	0	0	0
Camper Jumper Simulator	0	0	1	0	1	0	0	0	0
Blaite	1	1	1	0	0	0	1	0	0
Word Killer: Zorgilonian Chronicles	0	1	1	0	0	1	1	0	0
Nuclear Contingency	1	0	1	0	0	0	1	0	0
UNO	0	1	0	0	0	0	0	0	0
LSD Driver	0	1	1	0	1	1	1	1	0
Football Blitz	0	1	1	0	0	0	1	1	0
Redneck Rampage	1	1	0	0	0	0	1	0	1
Milkmaid of the Milky Way	1	0	1	0	0	0	0	0	0

Gambar 4.7
Data Setelah Mengalami *Data Preprocessing*.

Setelah mengalami *data preprocessing*, maka data diubah dalam format .csv yang nantinya dapat dibaca menggunakan Weka. Berbeda dengan aplikasi *data mining* lainnya seperti IBM SPSS Modeler yang dapat secara langsung menggunakan data dengan format .xls sebagai sumber data, Weka hanya bisa membuka data dengan format .arff, .names, .data, .csv, .json, .libsvm, .m, .dat, .bsi, dan .xrf.

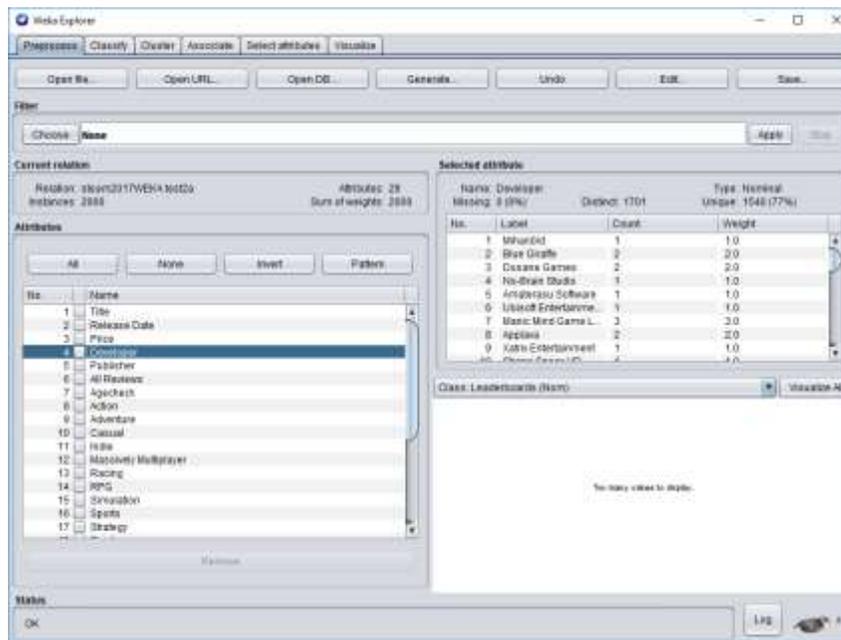
Apabila Weka membuka tipe data lain yang tidak dapat dibaca, maka *message box* akan muncul yang menginstruksikan bahwa tidak ada *file loader* yang dapat dibaca dan *user* diminta untuk mengatur CSVLoader untuk dapat membaca data.



Gambar 4.8
Frame dan Message Box yang muncul apabila terdapat kesalahan format dalam pembacaan data dalam Weka.

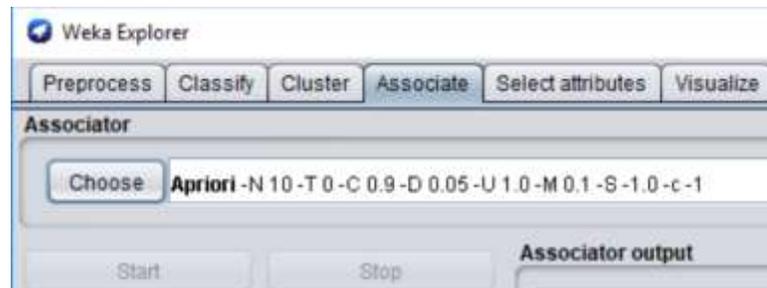
Kegagalan pembacaan data biasanya tidak hanya disebabkan karena perbedaan *delimiter*. Beberapa kasus pembacaan data pada Weka, meskipun *delimiter* yang dipakai sudah sesuai, data biasanya tidak dapat dibaca karena terdapat *special character* yang dibaca oleh Weka sebagai fungsi lain. Beberapa *special character*, selain spasi yang sebaiknya dihindari adalah tanda kutip, tanda garis, tanda seru, tanda tanya, dan karakter lainnya yang tidak sesuai dengan *Unicode* standar.

Weka memiliki batasan nilai unik daripada data yang dibaca. Apabila data unik yang dilampirkan melebihi batas yang ditentukan, maka Weka tidak dapat membaca data tersebut. Informasi dasar tidak dimasukkan sebagai atribut daripada *association rule*, hanya genre, tipe keterlibatan pemain, dan fitur.



Gambar 4.9
Weka Explorer, pada tab "Preprocess", Panel pada pojok kanan bawah memberikan peringatan "Too much values to display."

Apabila terdapat sebuah kolom yang memiliki nilai-nilai yang terlalu panjang atau terlalu banyak, maka Weka memberi tahu *user* bahwa terlalu banyak nilai yang ditampilkan. Dengan terlalu banyaknya nilai yang ditampilkan, maka algoritma *association rule* tidak dapat dieksekusi.



Gambar 4.10

Pada tab "Associate", tombol *Start* tidak dapat digunakan disebabkan karena terlalu banyaknya data yang ditampilkan.

4.4 Modelling Phase

Fase pemodelan data meliputi penentuan nilai minimum daripada *confidence* dan *support*, serta penggunaan nilai tersebut pada parameter dari algoritma Apriori dan FP-Growth. Pada fase ini pemodelan data daripada kedua algoritma dilakukan dan diteliti perbedaannya untuk mengetahui efektivitas dan efisiensi yang dimiliki masing-masing algoritma. Algoritma yang lebih efektif dan efisien akan nantinya digunakan sebagai rekomendasi untuk penelitian selanjutnya.

4.4.1 Penentuan Nilai Minimum *Support* dan *Confidence*

Untuk menentukan nilai minimum *support* dan *confidence* daripada penelitian ini, maka perlu diketahui nilai rata-rata dan nilai maksimum daripada jumlah kemunculan daripada *field* yang terdapat pada tabel 4.5.

Tabel 4.5
Tabel Jumlah Kemunculan Atribut yang Digunakan

Variabel	Jumlah Kemunculan
Action	925
Adventure	806
Casual	720
Indie	1429
Massively Multiplayer	92
Racing	72
RPG	352
Simulation	483
Sports	128
Strategy	427
Single	1894
Multi	340
Local	192

Variabel	Jumlah Kemunculan
Online	371
Agecheck	329
Achievements	1268
Controller Support	780
Workshop	90
Trading Cards	861
Cloud	720
Level Editor	89
Captions	73
Leaderboards	255
Maksimum nilai	1894
Rata-rata	552

Parameter yang diketahui untuk mencari nilai *support* minimum adalah nilai rata-rata kemunculan dan jumlah *record* keseluruhan. Setelah mengetahui nilai parameter masing-masing, maka *input* akan dimasukkan pada rumus yang sebelumnya ditulis. Nilai *support* minimum yang dihasilkan adalah 0,276 atau 27,6%.

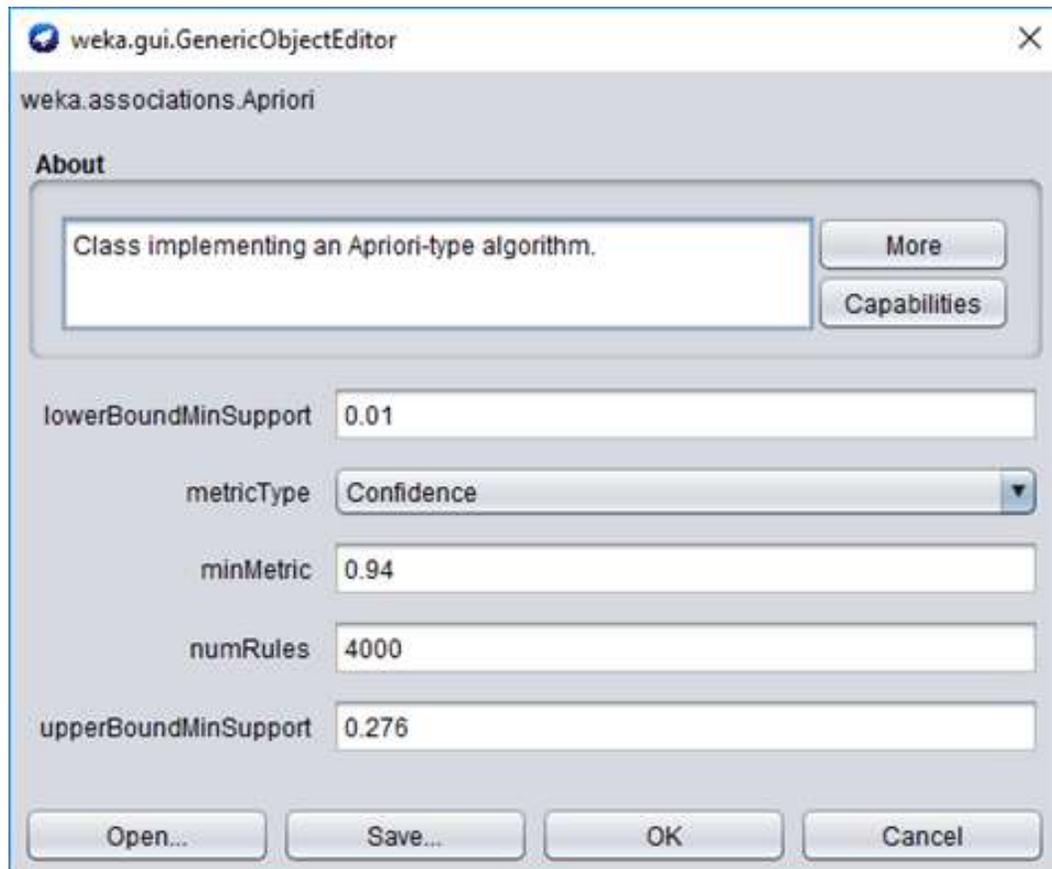
$$\text{Nilai support minimum} = \frac{552}{2000} = 0,276$$

Parameter yang diketahui untuk mencari nilai *confidence* minimum adalah nilai tertinggi kemunculan dan jumlah *record* keseluruhan. Setelah mengetahui nilai parameter masing-masing, maka *input* akan dimasukkan pada rumus yang sebelumnya ditulis. Nilai *confidence* minimum yang dihasilkan adalah 0,947 atau 94,7%.

$$\text{Nilai support minimum} = \frac{1894}{2000} = 0,947$$

4.4.2 Metode 1 – Algoritma Apriori

Dengan menggunakan Weka, 23 *field* dibaca sebagai *flag*. Kumpulan parameter yang terdapat pada `weka.associations.Apriori` adalah sebagai berikut:



Gambar 4.11

Pengaturan nilai minimum *support* dan *confidence* pada `weka.associations.Apriori`.

Pada gambar, Weka memungkinkan pengguna untuk dapat mengatur parameter yang mempengaruhi hasil yang dikeluarkan setelah algoritma *association rule* berjalan. Berikut merupakan keterangan parameter yang dilampirkan:

1. `lowerBoundMinSupport`

Nilai *support* minimum daripada *association rule*.

2. `metricType`

Parameter ini menentukan tipe metrik atau ukuran yang digunakan untuk menentukan urutan daripada *rules* yang dihasilkan. Pengguna dapat memilih tipe metrik berupa *confidence*, *lift*, *conviction*, dan *leverage*.

3. `minMetric`

Nilai minimum daripada tipe metrik yang ditentukan.

4. `numRules`

Jumlah *rule* yang harus ditemukan.

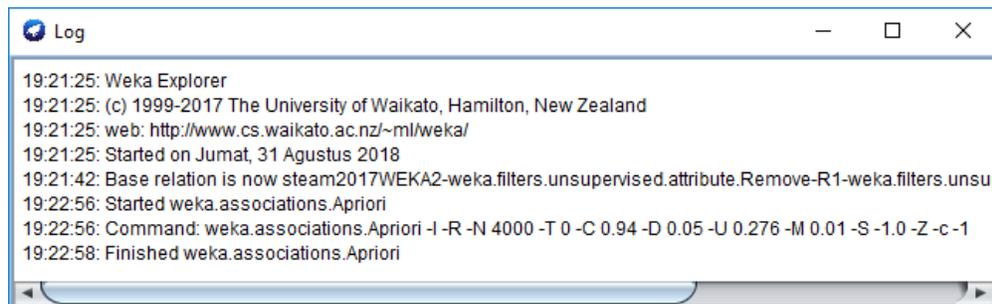
5. `upperBoundMinSupport`

Nilai *support* maksimum daripada *association rule*.

Tabel 4.6
Tabel Pengaturan Nilai Parameter Algoritma Apriori

Parameter	Nilai
<code>lowerBoundMinSupport</code>	0.01
<code>metricType</code>	<code>confidence</code>
<code>minMetric</code>	0.94
<code>numRules</code>	4000
<code>upperBoundMinSupport</code>	0.276

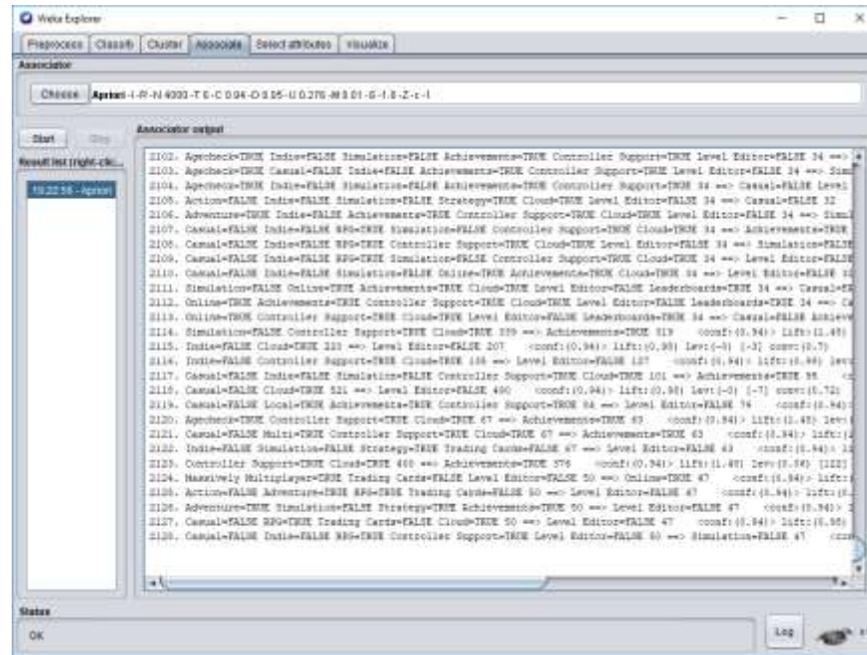
Setelah seluruh parameter dimasukkan, dan tombol *Start* ditekan, maka *Log* akan secara otomatis mencatat algoritma yang digunakan berikut dengan waktu berlangsungnya eksekusi.



Gambar 4.12
Log daripada eksekusi algoritma.

Setelah proses *association rule* selesai dieksekusi, maka akan muncul hasil yang keluar sebagai informasi yang dibutuhkan untuk proses evaluasi. Berdasarkan gambar, *Log*

menunjukkan bahwa waktu yang dibutuhkan untuk mengeksekusi algoritma Apriori hingga tuntas, adalah 2 detik.



Gambar 4.13
Rule yang Dihasilkan.

Terdapat 2128 *rule* yang dihasilkan dalam algoritma Apriori, namun *rule* ini perlu diolah lebih lanjut untuk mengetahui *rule* yang hanya mengandung nilai TRUE saja. Untuk mengolah *rule* menjadi bentuk yang dicari, maka *rule* yang memiliki nilai FALSE harus dihapuskan, dihilangkan, atau disembunyikan.

Tabel 4.7
Rule yang Dihasilkan

No.	Antisenden	Konsekuensi	Support	Confidence
1	Massively Multiplayer=TRUE RPG=TRUE	Online=TRUE	2,15%	100%
2	Online=TRUE Controller Support=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,95%	100%
3	Adventure=TRUE Massively Multiplayer=TRUE	Online=TRUE	1,60%	100%

No.	Antisenden	Konsekuen	Support	Confidence
4	Massively Multiplayer=TRUE RPG=TRUE Multi=TRUE	Online=TRUE	1,50%	100%
5	Adventure=TRUE Massively Multiplayer=TRUE Multi=TRUE	Online=TRUE	1,35%	100%
6	Multi=TRUE Online=TRUE Controller Support=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,35%	100%
7	Massively Multiplayer=TRUE Multi=TRUE Achievements=TRUE	Online=TRUE	1,30%	100%
8	Adventure=TRUE Massively Multiplayer=TRUE RPG=TRUE	Online=TRUE	1,15%	100%
9	Controller Support=TRUE Cloud=TRUE Captions=TRUE	Achievements=TRUE	1,15%	100%
10	Adventure=TRUE Massively Multiplayer=TRUE RPG=TRUE Multi=TRUE	Online=TRUE	1,00%	100%
11	Online=TRUE Controller Support=TRUE Leaderboards=TRUE	Achievements=TRUE	2,63%	98%
12	Online=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	2,43%	98%
13	Strategy=TRUE Leaderboards=TRUE	Achievements=TRUE	2,18%	98%
14	Multi=TRUE Online=TRUE Controller Support=TRUE Leaderboards=TRUE	Achievements=TRUE	1,88%	97%
15	Massively Multiplayer=TRUE Achievements=TRUE	Online=TRUE	1,78%	97%
16	Local=TRUE Cloud=TRUE Leaderboards	Controller Support=TRUE	1,78%	97%
17	Massively Multiplayer=TRUE Strategy=TRUE	Online=TRUE	1,73%	97%

No.	Antisenden	Konsekuen	Support	Confidence
18	Multi=TRUE Online=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,73%	97%
19	Multi=TRUE Controller Support=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,68%	97%
20	Local=TRUE Achievements=TRUE Cloud=TRUE Leaderboards=TRUE	Controller Support=TRUE	1,68%	97%
21	Controller Support=TRUE Cloud=TRUE Leaderboards=TRUE 98	Achievements=TRUE	4,88%	97%
22	Adventure=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,58%	97%
23	Agecheck=TRUE Adventure=TRUE Controller Support=TRUE Cloud=TRUE	Achievements=TRUE	1,53%	97%
24	Adventure=TRUE Controller Support=TRUE Cloud=TRUE	Achievements=TRUE	8,15%	96%
25	Massively Multiplayer=TRUE Strategy=TRUE Multi=TRUE	Online=TRUE	1,28%	96%
26	Strategy=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,23%	96%
27	Adventure=TRUE Controller Support=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,23%	96%
28	Multi=TRUE Controller Support=TRUE Leaderboards=TRUE	Achievements=TRUE	2,35%	96%
29	Adventure=TRUE Local=TRUE Cloud=TRUE	Controller Support=TRUE	1,18%	96%
30	Strategy=TRUE Local=TRUE Online=TRUE	Multi=TRUE	1,18%	96%
31	Local=TRUE Online=TRUE Controller Support=TRUE Leaderboards=TRUE	Achievements=TRUE	1,18%	96%

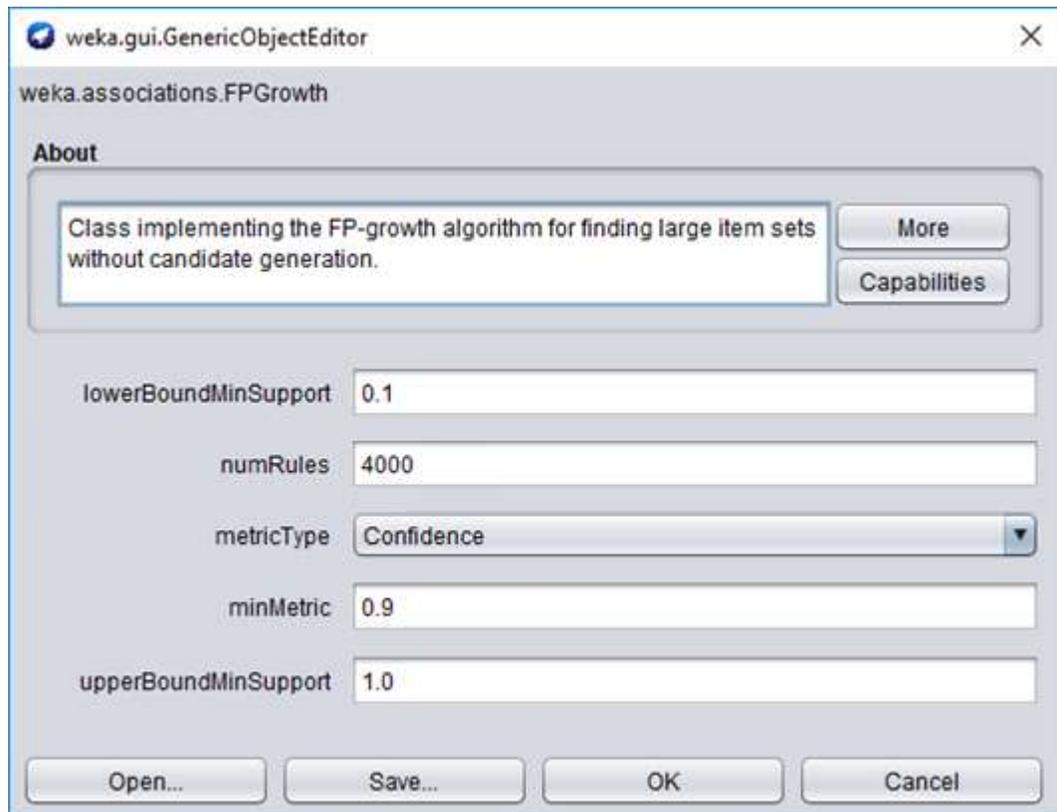
No.	Antisenden	Konsekuensi	Support	Confidence
32	Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	6,95%	96%
33	Massively Multiplayer=TRUE	Online=TRUE	4,50%	96%
34	Massively Multiplayer=TRUE Multi=TRUE	Online=TRUE	3,33%	96%
35	RPG=TRUE Cloud=TRUE Captions=TRUE	Adventure=TRUE	1,08%	95%
36	Multi=TRUE Online=TRUE Controller Support=TRUE Cloud=TRUE	Achievements=TRUE	3,18%	95%
37	Cloud=TRUE Captions=TRUE	Achievements=TRUE	2,10%	95%
38	Multi=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	2,10%	95%
39	Local=TRUE Online=TRUE Controller Support=TRUE Cloud=TRUE	Achievements=TRUE	2,05%	95%
40	Sports=TRUE Multi=TRUE Cloud=TRUE	Achievements=TRUE	1,03%	95%
41	Adventure=TRUE Online=TRUE Controller Support=TRUE Cloud=TRUE	Achievements=TRUE	1,03%	95%
42	Controller Support=TRUE Leaderboards=TRUE	Achievements=TRUE	7,60%	95%
43	Local=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,75%	94%
44	Multi=TRUE Local=TRUE Online=TRUE Controller Support=TRUE Cloud=TRUE	Achievements=TRUE	1,75%	94%
45	Local=TRUE Controller Support=TRUE Cloud=TRUE Leaderboards=TRUE	Achievements=TRUE	1,70%	94%
46	Agecheck=TRUE Controller Support=TRUE Cloud=TRUE 67	Achievements=TRUE	3,25%	94%

Tabel 4.7 memperlihatkan isi dari 46 *rule* dalam kurun waktu 1 tahun yang disusun berdasarkan nilai *confidence* tertinggi. 46 *rule* tersebut memiliki nilai *confidence* yang berkisar antara 94,7% hingga 100%. Berikut ini adalah analisis terhadap 3 *rule* teratas pada tabel 4.7 yang didasarkan dengan nilai *support* dan *confidence* yang tinggi.

1. Dalam rule 24 berisikan: IF (Adventure=TRUE Controller Support=TRUE Cloud=TRUE) THEN (Achievements=TRUE). Nilai support yang dimiliki adalah 8,15% sedangkan nilai confidence yang dimiliki adalah 96%.
2. Dalam rule 32 berisikan: IF (Cloud=TRUE Leaderboards=TRUE) THEN (Achievements=TRUE). Nilai support yang dimiliki adalah 6,95% sedangkan nilai confidence yang dimiliki adalah 96%.
3. Dalam rule 42 berisikan: IF (Controller Support=TRUE Leaderboards=TRUE) THEN (Achievements=TRUE). Nilai support yang dimiliki adalah 7,60% sedangkan nilai confidence yang dimiliki adalah 95%.

4.4.3 Metode 2 – Algoritma FP-Growth

Dengan menggunakan Weka, 23 *field* dibaca sebagai *flag*. Kumpulan parameter yang terdapat pada `weka.associations.FPGrowth` adalah sebagai berikut:



Gambar 4.14
Pengaturan nilai minimum *support* dan *confidence* pada `weka.associations.FPGrowth`.

Pada gambar 4.14, Weka memungkinkan pengguna untuk dapat mengatur parameter yang mempengaruhi hasil yang dikeluarkan setelah algoritma *association rule* berjalan. Berikut merupakan keterangan parameter yang dilampirkan:

1. `lowerBoundMinSupport`

Nilai batas bawah daripada nilai *support* minimum.

2. `maxNumberOfItems`

Jumlah maksimum *item* yang dapat dimasukkan dalam sebuah *itemset*. Nilai -1 menandakan tidak ada batas/jumlah maksimum.

3. *metricType*

Parameter ini menentukan tipe metrik atau ukuran yang digunakan untuk menentukan urutan daripada *rules* yang dihasilkan. Pengguna dapat memilih tipe metrik berupa *confidence*, *lift*, *conviction*, dan *leverage*.

4. *minMetric*

Nilai minimum daripada tipe metrik yang ditentukan.

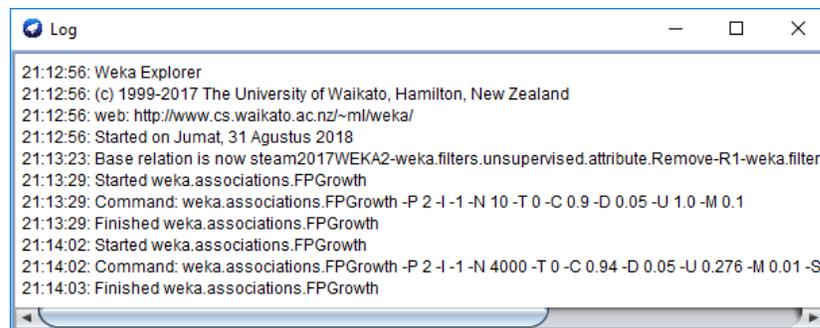
5. *numRulesToFind*

Jumlah *rule* yang dihasilkan.

Tabel 4.8
Tabel Pengaturan Nilai Parameter Algoritma FP-Growth

Parameter	Nilai
<i>lowerBoundMinSupport</i>	0.01
<i>metricType</i>	<i>confidence</i>
<i>minMetric</i>	0.94
<i>numRulesToFind</i>	4000
<i>upperBoundMinSupport</i>	0.276

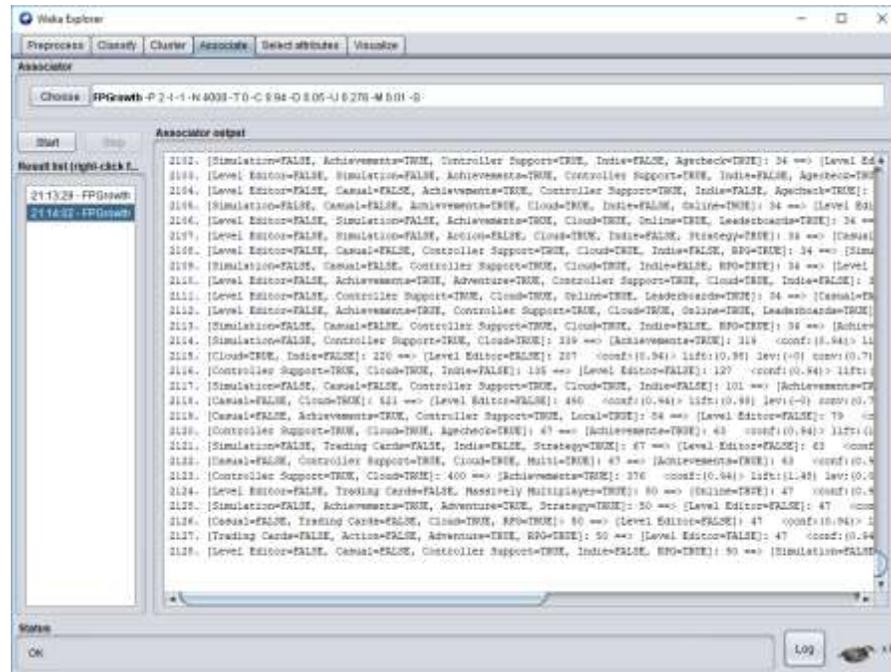
Setelah seluruh parameter dimasukkan, dan tombol *Start* ditekan, maka *Log* akan secara otomatis mencatat algoritma yang digunakan berikut dengan waktu berlangsungnya eksekusi.



Gambar 4.15
Log daripada eksekusi algoritma.

Setelah proses *association rule* selesai dieksekusi, maka akan muncul hasil yang keluar sebagai informasi yang dibutuhkan untuk proses evaluasi.

Berdasarkan gambar 4.15, Log menunjukkan bahwa waktu yang dibutuhkan untuk mengeksekusi algoritma FP-Growth hingga tuntas, adalah 1 detik.



Gambar 4.16
Rule yang Dihasilkan.

Terdapat 2128 *rule* yang dihasilkan dalam algoritma Apriori, namun *rule* ini perlu diolah lebih lanjut untuk mengetahui *rule* yang hanya mengandung nilai TRUE saja. Untuk mengolah *rule* menjadi bentuk yang dicari, maka *rule* yang memiliki nilai FALSE harus dihapuskan, dihilangkan, atau disembunyikan.

Tabel 4.9
Rule yang Dihasilkan

No.	Antisenden	Konsekuensi	Support	Confidence
1	[Adventure=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,60%	100%
2	[RPG=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	2,15%	100%

No.	Antisenden	Konsekuen	Support	Confidence
3	[Controller Support=TRUE, Cloud=TRUE Captions=TRUE]	[Achievements=TRUE]	1,15%	100%
4	[Achievements=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,30%	100%
5	[Adventure=TRUE, RPG=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,15%	100%
6	[Adventure=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,35%	100%
7	[RPG=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,50%	100%
8	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,95%	100%
9	[Adventure=TRUE, RPG=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,00%	100%
10	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,35%	100%
11	[Controller Support=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,63%	98%
12	[Cloud=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,43%	98%
13	[Strategy=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,18%	98%
14	[Controller Support=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,88%	97%
15	[Achievements=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,78%	97%
16	[Cloud=TRUE, Leaderboards=TRUE, Local=TRUE]	[Controller Support=TRUE]	1,78%	97%

No.	Antisenden	Konsekuen	Support	Confidence
17	[Strategy=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,73%	97%
18	[Cloud=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,73%	97%
19	[Controller Support=TRUE, Cloud=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,68%	97%
20	[Achievements=TRUE, Cloud=TRUE, Leaderboards=TRUE, Local=TRUE]	[Controller Support=TRUE]	1,68%	97%
21	[Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	4,83%	97%
22	[Adventure=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,58%	97%
23	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE, Agecheck=TRUE]	[Achievements=TRUE]	1,53%	97%
24	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE]	[Achievements=TRUE]	8,15%	96%
25	[Strategy=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,28%	96%
26	[Cloud=TRUE, Strategy=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,23%	96%
27	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,23%	96%
28	[Controller Support=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,35%	96%
29	[Adventure=TRUE, Cloud=TRUE, Local=TRUE]	[Controller Support=TRUE]	1,18%	96%

No.	Antisenden	Konsekuen	Support	Confidence
30	[Strategy=TRUE, Online=TRUE, Local=TRUE]	[Multi=TRUE]	1,18%	96%
31	[Controller Support=TRUE, Online=TRUE, Leaderboards=TRUE, Local=TRUE]	[Achievements=TRUE]	1,18%	96%
32	[Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	6,95%	96%
33	[Massively Multiplayer=TRUE]	[Online=TRUE]	4,50%	96%
34	[Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	3,33%	96%
35	[Cloud=TRUE, RPG=TRUE, Captions=TRUE]	[Adventure=TRUE]	1,08%	95%
36	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE]	[Achievements=TRUE]	3,18%	95%
37	[Cloud=TRUE, Captions=TRUE]	[Achievements=TRUE]	2,10%	95%
38	[Cloud=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,10%	95%
39	[Cloud=TRUE, Multi=TRUE, Sports=TRUE]	[Achievements=TRUE]	1,03%	95%
40	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE, Online=TRUE]	[Achievements=TRUE]	1,03%	95%
41	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Local=TRUE]	[Achievements=TRUE]	2,05%	95%
42	[Controller Support=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	7,60%	95%
43	[Cloud=TRUE, Leaderboards=TRUE, Local=TRUE]	[Achievements=TRUE]	1,75%	94%
44	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE, Local=TRUE]	[Achievements=TRUE]	1,75%	94%

No.	Antisenden	Konsekuensi	Support	Confidence
45	[Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE,Local=TRUE	[Achievements=TRUE]	1,70%	94%
46	[Controller Support=TRUE, Cloud=TRUE, Agecheck=TRUE]	[Achievements=TRUE]	3,25%	94%
47	[Controller Support=TRUE, Cloud=TRUE]	[Achievements=TRUE]	19,40%	94%

Tabel 4.9 memperlihatkan isi dari 47 *rule* dalam kurun waktu 1 tahun yang disusun berdasarkan nilai *confidence* tertinggi. 47 *rule* tersebut memiliki nilai *confidence* yang berkisar antara 94,7% hingga 100%. Berikut ini adalah analisis terhadap 3 *rule* teratas pada tabel 4.9 yang didasarkan dengan nilai *support* dan *confidence* yang tinggi.

1. Dalam rule 24 berisikan: IF (Adventure=TRUE Controller Support=TRUE Cloud=TRUE) THEN (Achievements=TRUE). Nilai support yang dimiliki adalah 8,15% sedangkan nilai confidence yang dimiliki adalah 96%.
2. Dalam rule 42 berisikan: IF (Controller Support=TRUE Leaderboards=TRUE) THEN (Achievements=TRUE). Nilai support yang dimiliki adalah 7,60% sedangkan nilai confidence yang dimiliki adalah 95%.
3. Dalam rule 47 berisikan: IF (Cloud=TRUE Leaderboards=TRUE) THEN (Achievements=TRUE). Nilai support yang dimiliki adalah 19,40% sedangkan nilai confidence yang dimiliki adalah 94%.

4.5 Evaluation Phase

Setelah melakukan *data mining* terhadap 2000 *record* dengan dua metode *association rule* yang berbeda, adapun beberapa pengetahuan yang diperoleh dari penelitian ini. Berdasarkan hasil *association rule* dalam periode transaksi Januari 2017 hingga Desember 2017, dilakukan analisis *video game* jenis apa yang muncul dalam *rule* yang dihasilkan. Hasil dari analisis tersebut digambarkan pada tabel 4.10 yang diurutkan berdasarkan nilai *support* dan *confidence*. Nilai *support* menentukan seberapa banyak sebuah *video game* terdapat pada

Dari tabel 4.10 diperoleh informasi *rule* yang memiliki nilai *confidence* 100% sebagai rujukan dalam menentukan strategi komposisi *video game* dalam proses pengembangan *video game* tersebut. Terdapat 10 *rule* dengan nilai *confidence* 100%.

Tabel 4.10
Rule dengan Nilai Confidence 100%

No.	Antisenden	Konsekuen	Support	Confidence
1	[Adventure=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,60%	100%
2	[RPG=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	2,15%	100%
3	[Controller Support=TRUE, Cloud=TRUE, Captions=TRUE]	[Achievements=TRUE]	1,15%	100%
4	[Achievements=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,30%	100%
5	[Adventure=TRUE, RPG=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,15%	100%
6	[Adventure=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,35%	100%
7	[RPG=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,50%	100%
8	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,95%	100%

No.	Antisenden	Konsekuen	Support	Confidence
9	[Adventure=TRUE, RPG=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,00%	100%
10	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,35%	100%

Berdasarkan tabel 4.10, terdapat genre, tipe keterlibatan pemain, dan fitur tertentu yang menjadi dominan atau tren pada tahun 2017. Genre yang menjadi tren adalah *Adventure*, *RPG*, dan *Massively Multiplayer* dengan melibatkan lebih dari satu pemain, atau area lokal daripada permainan. Hal tersebut dapat dilihat daripada antisenden dan konsekuen daripada tabel 4.10. Untuk menunjang pengembangan tren *video game* tersebut, para *developer* disarankan untuk menambahkan beberapa fitur yang relevan dengan tipe keterlibatan pemain dan genre.

Fitur yang menjadi umum dan/atau wajib adalah *Achievements* dimana para pemain berlomba untuk meraih sebanyak mungkin prestasi, rekor, dan peningkatan dalam *video game*, yang nantinya akan dipasang pada *Leaderboards* yang akan mempengaruhi ketenaran atau reputasi pemain yang bersangkutan. Fitur lainnya adalah penggunaan *controller* seperti *joystick* atau *Xbox controller* untuk menunjang efisiensi bermain, dan penggunaan *Steam Cloud* yang berfungsi sebagai manajemen *file* apabila sewaktu-waktu *user* harus menggunakan perangkat keras lain dalam memainkan game, seperti pada saat *game tournament*.

Dari tabel 4.11 diperoleh informasi *rule* yang memiliki nilai *support* melebihi 3% sebagai rujukan dalam mengetahui *video game* apa yang sedang menjamur pada Steam atau pasar *video game* saat ini. Terdapat 9 *rule* dengan nilai *support* melebihi 3%.

Tabel 4.11
Rule dengan Nilai *Support* Melebihi 3%

No.	Antisenden	Konsekuensi	Support	Confidence
1	[Controller Support=TRUE, Cloud=TRUE]	[Achievements=TRUE]	19,40%	94%
2	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE]	[Achievements=TRUE]	8,15%	96%
3	[Controller Support=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	7,60%	95%
4	[Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	6,95%	96%
5	[Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	4,83%	97%
6	[Massively Multiplayer=TRUE]	[Online=TRUE]	4,50%	96%
7	[Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	3,33%	96%
8	[Controller Support=TRUE, Cloud=TRUE, Agecheck=TRUE]	[Achievements=TRUE]	3,25%	94%
9	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE]	[Achievements=TRUE]	3,18%	95%

Berdasarkan tabel 4.11, *video game* yang menjamur tidak terlalu berfokus pada genre, melainkan berfokus pada keterlibatan pemain berikut dengan fitur yang berhubungan dengan pemain tersebut. Tidak ada genre yang terlalu spesifik seperti *Adventure*, *Strategy*, *RPG*, dan lain sebagainya yang benar-benar menonjol. Hal ini mengindikasikan bahwa pada tahun 2017, para *developer* berfokus dalam membuat *video game* yang melibatkan banyak orang dalam memainkannya, berikut dengan fitur yang bersangkutan.

Setelah mengetahui hasil yang diperoleh daripada *data mining*, maka terdapat informasi yang seharusnya diketahui oleh *developer* mengenai *video game* apa yang seharusnya dibuat dengan tujuan mengikuti tren pasar *video game* saat ini. Apabila *developer* berfokus daripada genre *video game*, maka disarankan untuk berfokus pada genre *Adventure* dan *RPG*. Apabila *developer* berfokus pada keterlibatan pemain dibandingkan dengan genre, maka disarankan untuk mengembangkan *video game* yang melibatkan banyak orang dalam memainkannya, setidaknya dalam jangkauan *multiplayer*, dan apabila memungkinkan dalam jangkauan *online*. *Video game* yang bersifat *online* lebih cenderung dipilih oleh *game developer* yang sudah berbentuk perusahaan dan sudah beroperasi setidaknya 3 tahun. Para *indie developer* memiliki kecenderungan untuk membuat *video game* yang bersifat *multiplayer*, apabila mengikuti jalur ini.

Fitur yang direkomendasikan untuk menjadi bahan pertimbangan pengembangan *video game* adalah *achievements*, *controller support*, *leaderboards*, dan *cloud*. *Achievements* dan *leaderboards* memiliki korelasi yang kuat terhadap tipe keterlibatan pemain karena salah satu faktor yang membuat para pemain tertarik dalam bermain *video game* yang bersifat *online* maupun *multiplayer*, adalah kompetisi. Kompetisi tersebut nantinya memberikan ketenaran dan reputasi tertentu pada pemain. Fitur lainnya yang menjadi bahan pertimbangan adalah penggunaan *controller* dalam bermain, khususnya *video game* yang membutuhkan kontrol yang gesit dalam *gameplay* yang ditawarkan, seperti *Adventure*, *RPG*, *Action*, dan lain sebagainya. Penggunaan *controller* juga memudahkan *video game* untuk dapat dimainkan. *Steam Cloud* berkorelasi dengan *video game* yang melibatkan banyak pemain, khususnya dalam *gaming competition* dan *gaming tournament*, yang memungkinkan pemain untuk menggunakan akun pribadi dengan perangkat keras yang berbeda tanpa harus memperlumahkan mengenai konsistensi data yang tersimpan daripada *video game* yang bersangkutan.

Apabila *developer* memilih untuk *video game* yang bersifat offline, maka disarankan untuk menambahkan fitur *achievements* dan *controller support* untuk membuat *gameplay* menjadi menarik dan mudah bagi para pemain. Apabila *video game* menggunakan sistem *file* untuk menyimpan atau membuka data, maka disarankan untuk diintegrasikan dengan menggunakan *Steam Cloud*.

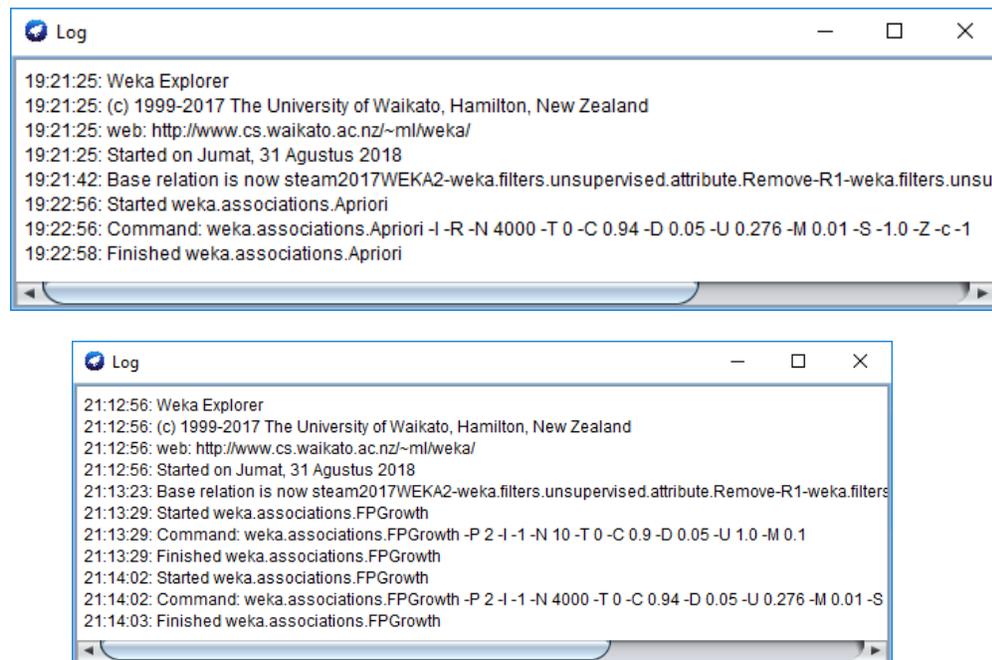
Tabel 4.12
Tabel Rekomendasi Komposisi *Video Game*

Field	Rekomendasi
Genre	Adventure, RPG, Sports, Strategy, Massively Multiplayer
Tipe Keterlibatan Pemain	Multi, Local, Online
Fitur	Achievements, Controller Support, Cloud, Leaderboards, Agecheck, Captions

Secara singkat, tren *video game* tahun 2017 pada Steam merupakan kombinasi daripada *itemset* diatas. Baik pihak *developer* maupun pemain dapat menggunakan tabel 4.12 sebagai acuan mengenai *video game* apa yang sebaiknya dibuat dan apa yang saat ini sedang mendominasi Steam pada tahun 2017.

4.6 Perbandingan Algoritma

Berdasarkan riset yang dilampirkan pada bab 2.2.3 mengenai perbandingan algoritma Apriori dan FP-Growth, maka terdapat pengujian yang harus dilakukan terhadap dua algoritma tersebut menggunakan aplikasi *data mining* yang bersangkutan. Pada saat dua algoritma dijalankan dengan tujuan mencari *rule*, *log* pada Weka menunjukkan perbedaan waktu antara dua algoritma tersebut.



Gambar 4.17
Perbandingan waktu eksekusi kedua algoritma

Berdasarkan gambar 4.17, algoritma Apriori membutuhkan waktu 2 detik untuk melakukan eksekusi terhadap 2000 *record*, sedangkan algoritma FP-Growth membutuhkan waktu 1 detik. Dengan menggunakan tabel dengan *record* dan *field* yang lebih banyak, akan terlihat perbedaan yang lebih besar mengenai performa daripada kedua algoritma.

Pengujian lebih lanjut melibatkan pengaturan nilai *lowerBoundMinSupport* untuk menguji performa kedua algoritma secara langsung. Parameter lainnya seperti *metricType*, *minMetric*, dan *numRules* memiliki nilai yang sama dengan pengaturan sebelumnya. Pengujian dilakukan selama 10 kali dengan nilai *lowerBoundMinSupport* yang semakin kecil, yang akan menambah beban kerja algoritma.

```

12:21:51: Weka Explorer
12:21:51: (c) 1999-2017 The University of Waikato, Hamilton, New Zealand
12:21:51: web: http://www.cs.waikato.ac.nz/~ml/weka/
12:21:51: Started on Kamis, 1 November 2018
12:22:22: Base relation is now steam2017WEKA2-weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsu
12:23:23: Started weka.associations.Apriori
12:23:23: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.01 -S -1.0 -Z -c -1
12:23:25: Finished weka.associations.Apriori
12:23:38: Started weka.associations.Apriori
12:23:38: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.009 -S -1.0 -Z -c -1
12:23:40: Finished weka.associations.Apriori
12:23:49: Started weka.associations.Apriori
12:23:49: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.008 -S -1.0 -Z -c -1
12:23:51: Finished weka.associations.Apriori
12:24:08: Started weka.associations.Apriori
12:24:08: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.007 -S -1.0 -Z -c -1
12:24:10: Finished weka.associations.Apriori
12:24:16: Started weka.associations.Apriori
12:24:16: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.006 -S -1.0 -Z -c -1
12:24:18: Finished weka.associations.Apriori
12:24:25: Started weka.associations.Apriori
12:24:25: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.005 -S -1.0 -Z -c -1
12:24:27: Finished weka.associations.Apriori
12:24:34: Started weka.associations.Apriori
12:24:34: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.004 -S -1.0 -Z -c -1
12:24:37: Finished weka.associations.Apriori
12:24:57: Started weka.associations.Apriori
12:24:57: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.003 -S -1.0 -Z -c -1
12:25:00: Finished weka.associations.Apriori
12:25:07: Started weka.associations.Apriori
12:25:07: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.002 -S -1.0 -Z -c -1
12:25:14: Finished weka.associations.Apriori
12:25:22: Started weka.associations.Apriori
12:25:22: Command: weka.associations.Apriori -I -R -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.001 -S -1.0 -Z -c -1
12:25:48: Finished weka.associations.Apriori

```

Gambar 4.18
Log Eksekusi Algoritma Apriori

```

12:26:29: Started weka.associations.FPGrowth
12:26:29: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.01 -S
12:26:31: Finished weka.associations.FPGrowth
12:26:39: Started weka.associations.FPGrowth
12:26:39: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.009 -S
12:26:40: Finished weka.associations.FPGrowth
12:26:46: Started weka.associations.FPGrowth
12:26:46: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.008 -S
12:26:47: Finished weka.associations.FPGrowth
12:26:52: Started weka.associations.FPGrowth
12:26:52: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.007 -S
12:26:53: Finished weka.associations.FPGrowth
12:27:00: Started weka.associations.FPGrowth
12:27:00: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.006 -S
12:27:01: Finished weka.associations.FPGrowth
12:27:08: Started weka.associations.FPGrowth
12:27:08: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.005 -S
12:27:09: Finished weka.associations.FPGrowth
12:27:14: Started weka.associations.FPGrowth
12:27:14: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.004 -S
12:27:16: Finished weka.associations.FPGrowth
12:27:21: Started weka.associations.FPGrowth
12:27:21: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.003 -S
12:27:23: Finished weka.associations.FPGrowth
12:27:28: Started weka.associations.FPGrowth
12:27:28: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.002 -S
12:27:35: Finished weka.associations.FPGrowth
12:27:40: Started weka.associations.FPGrowth
12:27:40: Command: weka.associations.FPGrowth -P 2 -I -1 -N 4000 -T 0 -C 0.94 -D 0.05 -U 0.276 -M 0.001 -S
12:28:09: Finished weka.associations.FPGrowth

```

Gambar 4.19
Log Eksekusi Algoritma FP-Growth

Setelah kedua algoritma diuji masing-masing selama sepuluh kali, maka tahap selanjutnya adalah mencari tahu waktu eksekusi berdasarkan waktu mulai dan waktu selesai dari masing-masing algoritma, lalu membandingkannya. Pada tabel 4.13, A adalah algoritma Apriori, sedangkan B adalah algoritma FP-Growth, dan hasil adalah *flag* yang menentukan bahwa waktu A tidak lebih dari waktu B.

Tabel 4.13
Tabel Perbandingan Waktu

Test	A-START	A-END	A-Time	B-START	B-END	B-Time	Hasil
0,010	12.23.23	12.23.25	00.00.02	12.26.29	12.26.31	00.00.02	FALSE
0,009	12.23.38	12.23.40	00.00.02	12.26.39	12.26.40	00.00.01	TRUE
0,008	12.23.49	12.23.51	00.00.02	12.26.46	12.26.47	00.00.01	TRUE
0,007	12.24.08	12.24.10	00.00.02	12.26.52	12.26.53	00.00.01	TRUE
0,006	12.24.16	12.24.18	00.00.02	12.27.00	12.27.01	00.00.01	TRUE
0,005	12.24.25	12.24.27	00.00.02	12.27.08	12.27.09	00.00.01	TRUE
0,004	12.24.34	12.24.37	00.00.03	12.27.14	12.27.16	00.00.02	TRUE
0,003	12.24.57	12.25.00	00.00.03	12.27.21	12.27.23	00.00.02	TRUE
0,002	12.25.07	12.25.14	00.00.07	12.27.28	12.27.35	00.00.07	TRUE
0,001	12.25.22	12.25.48	00.00.26	12.27.40	12.28.09	00.00.29	FALSE

Hasil menunjukkan dari antara 10 kasus, terdapat delapan *flag* yang menunjukkan bahwa algoritma FP-Growth memiliki efisiensi waktu yang lebih baik daripada algoritma Apriori. Berdasarkan perbandingan diatas, dapat disimpulkan bahwa algoritma FP-Growth lebih efisien dan efektif dibandingkan dengan algoritma Apriori. Algoritma FP-Growth memiliki penggunaan memori yang lebih kecil karena penyimpanan hanya melibatkan basis data yang dipadatkan, sedangkan algoritma Apriori menyimpan seluruh *itemset* dalam memori. Secara kecepatan *runtime*, algoritma FP-Growth dapat diukur lebih mudah karena sifatnya yang linear, berbeda dengan algoritma Apriori yang memiliki kecepatan *runtime* yang bersifat eksponensial. Tabel atau basis data yang memiliki *record* dan/atau *field* dengan jumlah besar, sebaiknya menggunakan algoritma FP-Growth.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan terhadap data *video game* yang masuk dalam Steam pada tahun 2017, terdapat beberapa kesimpulan yang dibagi menjadi dua bagian, yaitu pada sisi algoritma, dan hasil dari *data mining*. Adapun kesimpulan yang didapatkan pada penelitian ini adalah:

1. *Rule* yang dihasilkan daripada *data mining* menunjukkan bahwa *video game* yang muncul pada tahun 2017 membentuk kombinasi tertentu pada masing-masing fiturnya. Hal ini menunjukkan bahwa tahun 2017 merupakan tahun dimana *video game* bertipe MOBA (*Multiplayer Online Battle Arena*) sedang naik daun dan/atau pada puncak popularitas pada Steam. Terdapat beberapa fitur dalam *rule* yang menunjukkan karakteristik daripada MOBA. Genre yang muncul adalah RPG, *Massively Multiplayer*, dan Strategy. Tipe keterlibatan pemain melebihi satu orang. Fitur lainnya yang ditawarkan adalah munculnya fitur *achievements* dan *leaderboards*, serta penggunaan *cloud storage* dan perangkat keras dan/atau kontroler lainnya untuk dapat memainkan *video game*. Beberapa dari fitur tersebut memerlukan koneksi internet untuk dapat berfungsi.
2. Algoritma FP-Growth memiliki cara kerja yang berbeda daripada algoritma Apriori. Algoritma FP-Growth mengutamakan *pattern growth* dengan menggunakan *tree*, sedangkan algoritma Apriori mengutamakan penciptaan kandidat yang sifatnya *breadth-first*. Perbedaan tersebut mengarah pada performa efisiensi dan efektivitas algoritma yang ditunjukkan. Meskipun secara keseluruhan algoritma FP-Growth lebih baik dibandingkan dengan algoritma Apriori, algoritma FP-Growth tidak akan ditemukan tanpa adanya algoritma Apriori.

5.2 Saran

Apabila terdapat penelitian selanjutnya yang melibatkan penggunaan metode *association rule*, sebaiknya disarankan menggunakan algoritma FP-Growth sebagai langkah pencarian *rule*, karena sifatnya yang efektif dan efisien. Apabila terdapat penelitian selanjutnya yang melibatkan pencarian tren pada *video game*, sebaiknya menggunakan data keseluruhan yang muncul dalam satu tahun, atau menggunakan beberapa sampel acak yang diambil pada beberapa tahun ke belakang.

DAFTAR PUSTAKA

- Adams, Ernest, 2010, "*Fundamentals of Game Design – 2nd Edition*", Pearson.
- Anggraini, Rizka Ainul, 2017, "*Sistem Analisa Keranjang Belanja dengan Menggunakan Metode Algoritma Apriori pada Penjualan Suku Cadang (Studi kasus: AHASS PELITA Motor Kediri)*", Universitas Nusantara Kediri.
- Ariana, A.A. Gede Bagus, et al., 2013, "*Analisis Keranjang Belanja dengan Algoritma Apriori pada Perusahaan Retail*", STMIK STIKOM INDONESIA.
- Bouckaert, Remco R., et al., 2013, "*WEKA Manual for Version 3-7-8*", University of Waikato.
- Chapman, Pete, et al., 2000, "*CRISP-DM 1.0 - Step-by-step data mining guide by Pete Chapman (NCR)*", SPSS.
- Diwantara, Heru, et al., 2013, "*Perancangan Aplikasi Data Mining dengan Algoritma Apriori untuk Frekuensi Analisis Keranjang Belanja pada Data Transaksi Penjualan (Studi kasus di Swalayan KPRI Universitas Brawijaya)*", Jurusan Teknik Industri Fakultas Teknik Universitas Brawijaya.
- Griffith, Benjamin, 2013, "*College Fads*", St. James Encyclopedia of Popular Culture – via Gale Virtual Reference Library.
- Groeneveld, Liza, et al., 2016, "*2016 Global Games Market Report*", Newzoo Games.
- Han, Jiawei, et al., 2004, "*Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach*", Kluwer Academic Publishers.
- Han, Jiawei, et al., 2012, "*Data Mining: Concept and Techniques Third Edition*", Morgan Kaufmann Publishers.
- Hegland, Markus, 2005, "*The Apriori Algorithm – a Tutorial*", Australian National University.
- Iqbal, M., et al., 2017, "*Analisa Keranjang Belanja Konsumen Pada Data Penjualan Bulan Ramadhan Menggunakan Algoritma Apriori (Studi kasus: Distro Coffeepark Clothes Pekanbaru)*", Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau Pekanbaru.
- Jensen, Kenneth, 2016, "*IBM SPSS Modeler 18.0 User's Guide*", IBM Corporation.
- Kelly, Christian, et al., 2014, "*The Pulse of Gaming – Gaming Disruption*", Accenture.
- Kent, Steven L., 2000, "*The Ultimate History of Video Games*", Three Rivers Press.
- Kotler, Ph, et al., 2005, "*Principles of Marketing*", Pearson Education Limiter, Prentice-Hall.
- Larose, Daniel T., 2005, "*Discovering Knowledge in Data*", John Wiley & Sons Inc.
- Li, Haoyuan, et al., 2008, "*PFP: Parallel FP-Growth for Query Recommendation*", Google Beijing Research.

Rodrigues, Cajetan, et al., 2012, "*Better Game Design using Association Analysis*", Department of Information Technology, St. Francis Institute of Technology, Mumbai, India.

Tama, Bayu Adhi, 2010, "*Penetapan Strategi Penjualan Menggunakan Association Rules dalam Konteks CRM*", Fakultas Ilmu Komputer, Universitas Sriwijaya.

Varde, Aparna S., et al., 2004, "*Apriori Algorithm and Game-of-Life for Predictive Analysis in Materials Science*", Worcester Polytechnic Institute (WPI), Worcester, USA.

Witten, Ian H., Eibe Frank, 2005, "*Data Mining Practical Machine Learning Tools and Techniques Second Edition*", Morgan Kaufmann Publishers.

<https://www.slideshare.net/INSOFE/apriori-algorithm-36054672>, diakses tanggal 23 Mei 2018, jam 09:45 WIB.

https://www.gamasutra.com/view/news/312843/7672_games_hit_Steam_in_2017_alone_says_Steam_Spy.php, diakses tanggal 16 Mei 2018, jam 12:13 WIB.

<http://www.thegamesjournal.com/articles/WhatsaGame.shtml>, diakses tanggal 26 Januari 2018, jam 22:19 WIB.

<http://store.steampowered.com/about/>, diakses tanggal 23 Februari 2018, jam 11:40 WIB.

http://store.steampowered.com/app/291860/Pit_People/, diakses tanggal 23 Februari 2018, jam 11:40 WIB.

<https://www.pcmag.com/article2/0,2817,41291,00.asp>, diakses tanggal 2 Juni 2018, jam 13:45 WIB.

<https://www.gamesindustry.biz/articles/the-last-of-the-independents->, diakses tanggal 2 Juni 2018, jam 13:48 WIB.

<https://kotaku.com/steam-is-10-today-remember-when-it-sucked-1297594444>, diakses tanggal 2 Juni 2018, jam 13:53 WIB.

<http://www.gameguru.in/pc/2007/24/steam-registers-13-million-active-accounts/>, diakses tanggal 2 Juni 2018, jam 13:59 WIB.

<https://www.pcgamer.com/steam-versions/>, diakses tanggal 2 Juni 2018, jam 14:10 WIB.

<https://www.gamesradar.com/history-of-valve/>, diakses tanggal 2 Juni 2018, jam 15:01 WIB.

<http://www.ign.com/articles/2003/09/10/steam-client-delayed>, diakses tanggal 2 Juni 2018, jam 15:21 WIB.

<http://news.bbc.co.uk/2/hi/technology/4019095.stm>, diakses tanggal 2 Juni 2018, jam 15:26 WIB.

<https://www.gamespot.com/articles/half-life-2-now-preloading-via-steam/1100-6105848/>, diakses tanggal 2 Juni 2018, jam 15:27 WIB.

<https://www.pcgamer.com/market-data-firm-claims-valve-made-730-million-last-year/>, diakses tanggal 2 Juni 2018, jam 15:45 WIB.

<https://www.cs.waikato.ac.nz/ml/weka/>, diakses tanggal 7 Juni 2018, jam 16:25 WIB.

<https://definitions.uslegal.com/d/digital-distribution/>, diakses tanggal 22 November 2018, jam 14:45 WIB

LAMPIRAN

HASIL ASSOCIATION RULE BULAN JANUARI 2017 – DESEMBER 2017

No.	Antisenden	Konsekuen	Support	Confidence
1	[Adventure=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,60%	100%
2	[RPG=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	2,15%	100%
3	[Controller Support=TRUE, Cloud=TRUE Captions=TRUE]	[Achievements=TRUE]	1,15%	100%
4	[Achievements=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,30%	100%
5	[Adventure=TRUE, RPG=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,15%	100%
6	[Adventure=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,35%	100%
7	[RPG=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,50%	100%
8	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,95%	100%
9	[Adventure=TRUE, RPG=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,00%	100%
10	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,35%	100%
11	[Controller Support=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,63%	98%
12	[Cloud=TRUE, Online=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,43%	98%
13	[Strategy=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,18%	98%

No.	Antisenden	Konsekuen	Support	Confidence
14	[Controller Support=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,88%	97%
15	[Achievements=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,78%	97%
16	[Cloud=TRUE, Leaderboards=TRUE, Local=TRUE]	[Controller Support=TRUE]	1,78%	97%
17	[Strategy=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,73%	97%
18	[Cloud=TRUE, Online=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,73%	97%
19	[Controller Support=TRUE, Cloud=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,68%	97%
20	[Achievements=TRUE, Cloud=TRUE, Leaderboards=TRUE, Local=TRUE]	[Controller Support=TRUE]	1,68%	97%
21	[Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	4,83%	97%
22	[Adventure=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,58%	97%
23	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE, Agecheck=TRUE]	[Achievements=TRUE]	1,53%	97%
24	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE]	[Achievements=TRUE]	8,15%	96%
25	[Strategy=TRUE, Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	1,28%	96%
26	[Cloud=TRUE, Strategy=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,23%	96%

No.	Antisenden	Konsekuen	Support	Confidence
27	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	1,23%	96%
28	[Controller Support=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,35%	96%
29	[Adventure=TRUE, Cloud=TRUE, Local=TRUE]	[Controller Support=TRUE]	1,18%	96%
30	[Strategy=TRUE, Online=TRUE, Local=TRUE]	[Multi=TRUE]	1,18%	96%
31	[Controller Support=TRUE, Online=TRUE, Leaderboards=TRUE, Local=TRUE]	[Achievements=TRUE]	1,18%	96%
32	[Cloud=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	6,95%	96%
33	[Massively Multiplayer=TRUE]	[Online=TRUE]	4,50%	96%
34	[Multi=TRUE, Massively Multiplayer=TRUE]	[Online=TRUE]	3,33%	96%
35	[Cloud=TRUE, RPG=TRUE, Captions=TRUE]	[Adventure=TRUE]	1,08%	95%
36	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE]	[Achievements=TRUE]	3,18%	95%
37	[Cloud=TRUE, Captions=TRUE]	[Achievements=TRUE]	2,10%	95%
38	[Cloud=TRUE, Multi=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	2,10%	95%
39	[Cloud=TRUE, Multi=TRUE, Sports=TRUE]	[Achievements=TRUE]	1,03%	95%
40	[Adventure=TRUE, Controller Support=TRUE, Cloud=TRUE, Online=TRUE]	[Achievements=TRUE]	1,03%	95%

No.	Antisenden	Konsekuen	Support	Confidence
41	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Local=TRUE]	[Achievements=TRUE]	2,05%	95%
42	[Controller Support=TRUE, Leaderboards=TRUE]	[Achievements=TRUE]	7,60%	95%
43	[Cloud=TRUE, Leaderboards=TRUE, Local=TRUE]	[Achievements=TRUE]	1,75%	94%
44	[Controller Support=TRUE, Cloud=TRUE, Online=TRUE, Multi=TRUE, Local=TRUE]	[Achievements=TRUE]	1,75%	94%
45	[Controller Support=TRUE, Cloud=TRUE, Leaderboards=TRUE,Lo cal=TRUE	[Achievements=TRUE]	1,70%	94%
46	[Controller Support=TRUE, Cloud=TRUE, Agecheck=TRUE]	[Achievements=TRUE]	3,25%	94%
47	[Controller Support=TRUE, Cloud=TRUE]	[Achievements=TRUE]	19,40%	94%

Keterangan:

1. *Rule 1:*

IF [Adventure=TRUE, Massively Multiplayer=TRUE] THEN [Online=TRUE] memiliki nilai *support* atau jumlah kemunculan sebanyak 1,60% daripada keseluruhan sampel acak yang diambil pada tahun 2017. Nilai *confidence* 100% menandakan bahwa apabila sebuah *video game* muncul dengan genre *Adventure* dan *Massively Multiplayer*, maka tipe keterlibatan pemain *Online* akan selalu muncul pada setiap transaksi yang muncul.

2. *Rule 2:*

IF [RPG=TRUE, Massively Multiplayer=TRUE] THEN [Online=TRUE] memiliki nilai *support* atau jumlah kemunculan sebanyak 2,15% daripada keseluruhan sampel acak yang diambil pada tahun 2017. Nilai *confidence* 100% menandakan bahwa apabila sebuah

video game muncul dengan genre *RPG* dan *Massively Multiplayer*, maka tipe keterlibatan pemain *Online* akan selalu muncul pada setiap transaksi yang muncul.

3. *Rule 3:*

IF [Controller Support=TRUE, Cloud=TRUE, Captions=TRUE] THEN [Achievements=TRUE] memiliki nilai *support* atau jumlah kemunculan sebanyak 1,15% daripada keseluruhan sampel acak yang diambil pada tahun 2017. Nilai *confidence* 100% menandakan bahwa apabila sebuah *video game* muncul dengan fitur *Controller Support*, *Cloud*, dan *Captions*, maka fitur *Achievements* akan selalu muncul pada setiap transaksi yang muncul.

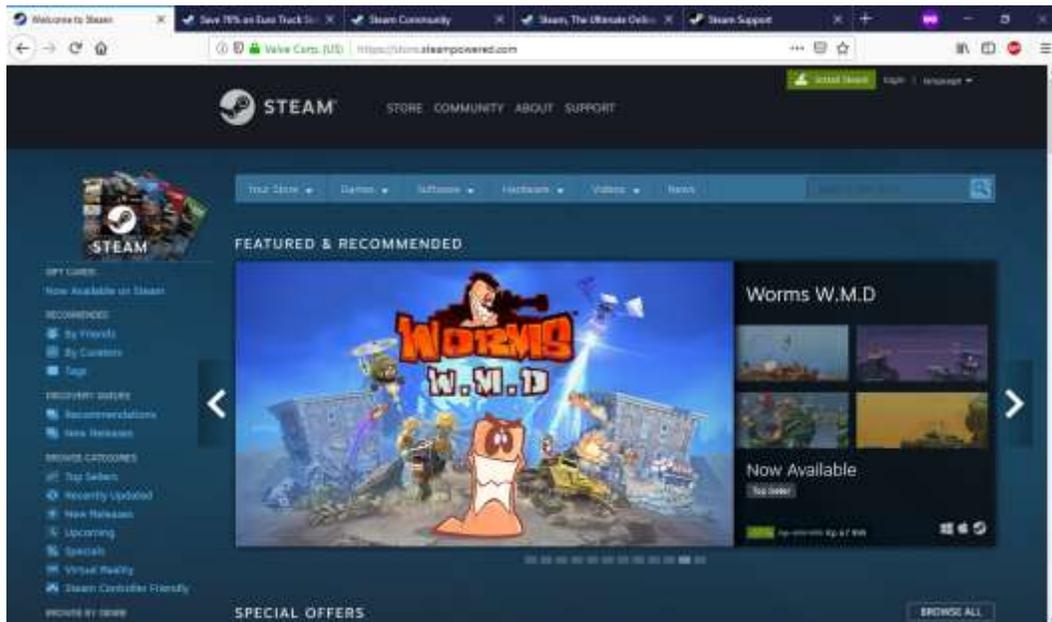
4. *Rule 4:*

IF [Achievements=TRUE, Multi=TRUE, Massively Multiplayer=TRUE] THEN [Online=TRUE] memiliki nilai *support* atau jumlah kemunculan sebanyak 1,30% daripada keseluruhan sampel acak yang diambil pada tahun 2017. Nilai *confidence* 100% menandakan bahwa apabila sebuah *video game* muncul dengan fitur *Achievements*, tipe keterlibatan pemain *Multiplayer*, dan genre *Massively Multiplayer*, maka tipe keterlibatan pemain *Online* akan selalu muncul pada setiap transaksi yang muncul.

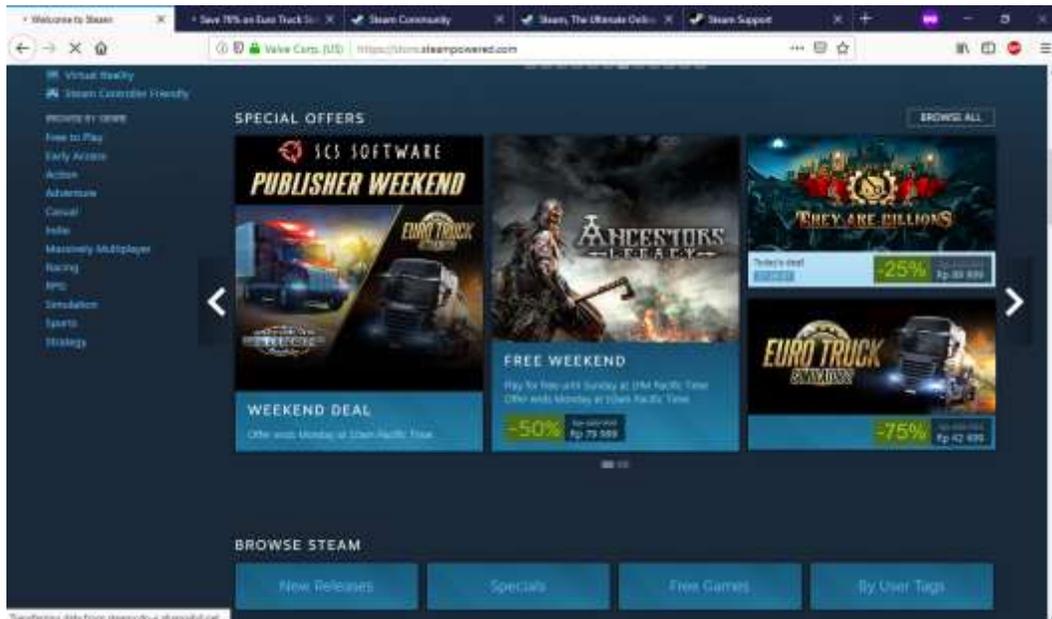
5. *Rule 5:*

IF [Adventure=TRUE, RPG=TRUE, Massively Multiplayer=TRUE] THEN [Online=TRUE] memiliki nilai *support* atau jumlah kemunculan sebanyak 1,15% daripada keseluruhan sampel acak yang diambil pada tahun 2017. Nilai *confidence* 100% menandakan bahwa apabila sebuah *video game* muncul dengan genre *Adventure*, *RPG*, dan *Massively Multiplayer*, maka tipe keterlibatan pemain *Online* akan selalu muncul pada setiap transaksi yang muncul.

FOTO STEAM



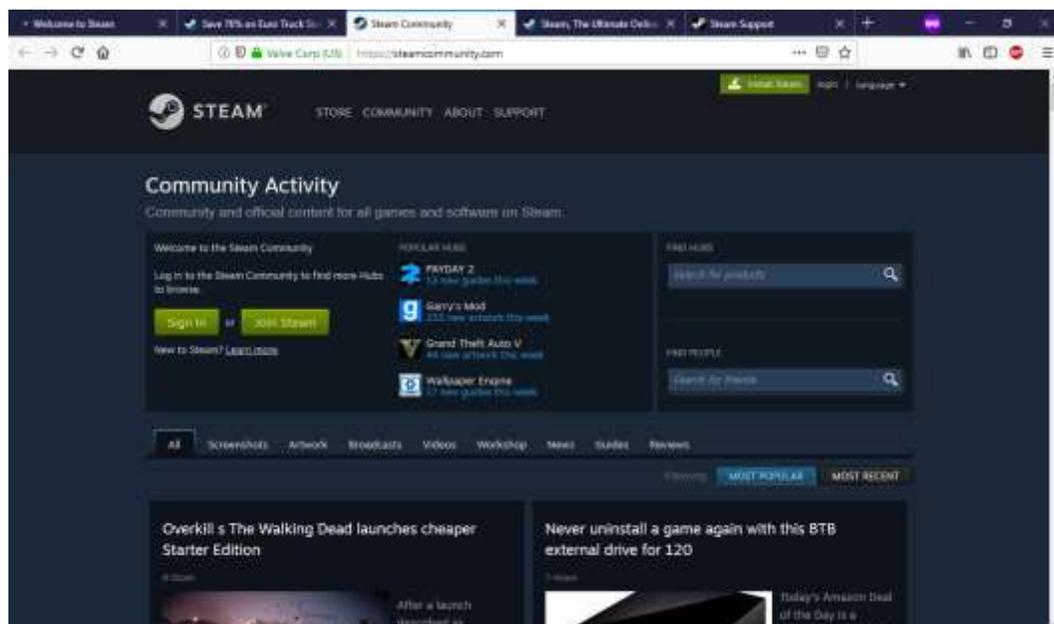
Halaman Utama Steam



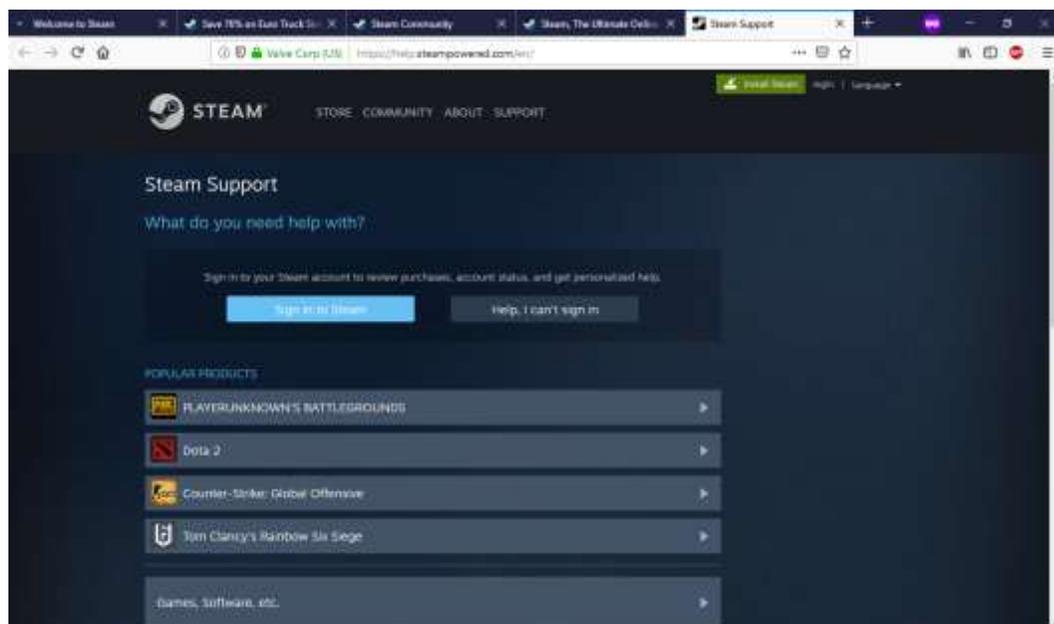
Halaman Penawaran Steam



Halaman Informasi *Video Game* pada *Steam*



Halaman Komunitas pada *Steam*

Halaman Tentang pada *Steam*Halaman Bantuan pada *Steam*