

Ekstraksi Pola Hubungan Penerimaan Mahasiswa Baru Dengan Sebaran Wilayah Asal Sekolah Menggunakan ARM Algoritma Apriori

Deden Prayitno

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Perbanas Institute,
Jl. Perbanas, Karet Kuningan, Setiabudi, Jakarta, Indonesia
deden@perbanas.id

Abstract — Competition between Private Universities (PTS) to get new students increases tight course, making the marketing department at any educational institution undertakes various promotional model that is exciting and innovative, plus competitors from new opening cooperation Asean Economic Community (AEC). One of the breakthrough strategies undertaken apart from the various models of the promotion is to open up the new student registration to be sufficiently large. The goal is to get a pattern of relationships between data attributes of the student database. Patterns can provide prefix information (knowledge base) as a base to get a map of the area students who apply to the Banks Association Institute. The pattern is extracted from database using data mining techniques with Association Rule Mining (ARM) method and apriori algorithms. This research uses secondary data that already exists in Marketing Bureau and Library and Information Technology Bureau, the entire database of students who enroll in PERBANAS Institute in period of 3 years (2013-2015). The results of data mining, can describe patterns of relationships between data attributes, which describe the distribution area prospective students who apply and are accepted on any existing courses at the Institute Banks Association. This pattern can be used to help define marketing strategies on the Bureau to create a campaign targeted at high school and equivalent territory of origin as prospective new student Perbanas Institute.

Keywords — students, data mining, patterns, ARM, algorithms apriori

I. PENDAHULUAN

Perguruan Tinggi terus mencari berbagai strategi yang bertujuan menjaring para lulusan SMA dan SMK menjadi mahasiswanya, baik strategi yang langsung terkait tim pemasaran (eksternal), maupun strategi yang dilakukan di dalam perguruan tinggi (internal) [1]. Jumlah lulusan SMU dan sederajat yang jauh lebih sedikit dibanding daya tampung perguruan tinggi, menjadi dasar berbagai perguruan tinggi untuk melakukan berbagai terobosan dalam memasarkan perguruan tingginya, seperti yang disampaikan oleh M. Nuh (mantan mendiknas) pada liputan6.com di bulan Mei 2014 menyatakan jumlah lulusan SMK sebanyak 1.171.907 siswa dan untuk lulusan SMA sebanyak 1.632.757 siswa, dan jumlah inilah yang

akan diperebutkan oleh berbagai perguruan tinggi di Indonesia yang jumlahnya sebanyak 4315 Perguruan Tinggi. Lain lagi yang disampaikan oleh mantan Mendikbud periode Presiden Jokowi jilid I, bapak Anis Baswedan pada bulan Mei 2015 di Kaltimpost.co.id, mnyampaikan bahwa yang lulus SMA dan SMK dan akan melanjutkan pendididkan ke jenjang lebih tinggi di berbagai perguruan tinggi adalah sebanyak sekitar 60 persen saja, sementara yang lainnya memilih bekerja dengan alasan ekonomi dan non ekonomi. Contoh strategi yang dilakukan seperti sistem pembayaran kuliah yang dinamis, pelayanan maksimal untuk kemudahan berinteraksi (satu atap), pemanfaatan Teknologi Informasi diberbagai bidang pelayanan, khususnya di bidang administrasi akademik (sistem pembelajaran), jadwal kuliah yang dinamis, harga yang menarik beserta kemudahan yang didapat, dan kelengkapan sarana prasarana sebagai pendukung kenyamanan, keamanan dan kepuasan mahasiswa selama menjalani proses pembelajaran [2].

Biro marketing berperan begitu penting dalam mencari dan mendapatkan mahasiswa baru, agar roda perkuliahan di Perguruan Tinggi tetap berjalan, kemampuan yang kreatif dan inovatif dalam merencanakan strategi pemasarannya, bukan tidak mungkin keberadaan biro ini dapat mendatangkan calon mahasiswa dalam jumlah yang besar sepanjang tahun.

Salah satu sumber informasi untuk strategi pemasaran yang jarang sekali dilakukan adalah dengan menggali database pendaftaran calon mahasiswa selama beberapa periode yang telah lalu, yang selama ini tersimpan saja, dan biasanya dibuka hanya jika dibutuhkan untuk membuat laporan. Peneliti akan menggunakan database ini untuk menjabarkan data sehingga mampu bernilai dan mampu memberikan informasi penting untuk menjaring calon mahasiswa baru.

Kemampuan sistem aplikasi untuk *data mining*, mampu melakukan menganalisis data, meringkas dan mengekstraksi pengetahuan dari data dalam jumlah yang besar, yang tidak akan mampu bila ditangani secara konvensional [5]. Disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data untuk

mendapatkan informasi penting yang sifatnya implisit dan sebelumnya tidak diketahui adalah *data mining* [6].

Data mining dengan dengan metoda asosiasi dan algoritma apriori dapat menggambarkan hubungan atau keterkaitan antar data pada database dengan nilai tingkat kepercayaan (*confidence level*) yang dimunculkan dari hubungan tersebut [8]. Data adalah fakta yang terekam dan tidak membawa arti sedangkan pengetahuan adalah pola, aturan atau model yang muncul dari data. Pola data adalah hubungan atribut yang ditunjukkan oleh nilai dari setiap atribut tersebut yang ada pada database. Data yang digunakan dalam penelitian ini adalah selama 3 tahun lalu (2013 - 2015), Karena pada tahun-tahun sebelumnya data tersimpan pada unit organisasi yang berbeda dengan struktur data yang berbeda pula, sehingga cukup sulit untuk mendapatkan data untuk diolah. Untuk melakukan pengolahan data yang menghasilkan asosiasi antar atribut menggunakan program aplikasi WEKA 3.6.10 [7].

II. KAJIAN PUSTAKA

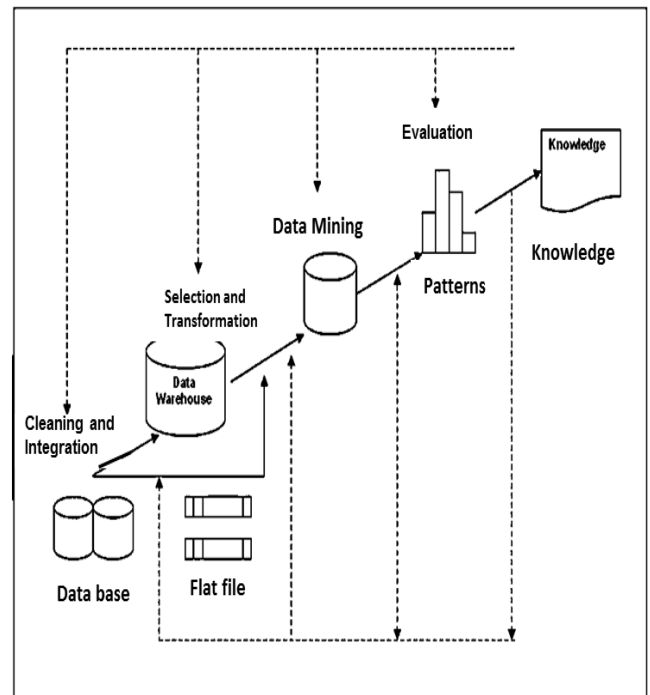
A. Data Mining

Data mining adalah proses menggali informasi dari database yang besar dan tersimpan dengan waktu yang relative lama dengan istilah untuk menemukan pengetahuan yang tersembunyi di dalam database. Proses yang dilakukan dengan menggunakan teknik statistik dan *machine learning* dengan mengekstraksi dan mengidentifikasi informasi yang potensial dan berguna untuk dimanfaatkan yang tersimpan di dalam database dengan ukuran besar [5]. *Data mining* adalah proses menemukan informasi yang bermakna dari korelasi antar atribut data, pola dan tren dengan memilah-milah sejumlah besar data yang disimpan dalam database, pengenalan pola menggunakan teknologi serta teknik statistik dan matematika [8].

Data mining memiliki beberapa tahapan yang dikenal sebagai ekstraksi pengetahuan, arkeologi data, eksplorasi dalam pemrosesan pola data dan mendapatkan informasi yang sesuai. *Data mining* terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery in Database* (KDD) [10]. Proses pencarian pengetahuan ini menggunakan berbagai teknik-teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksikannya. Proses pencarian bersifat iteratif dan interaktif untuk menemukan pola atau model yang sah, baru, bermanfaat, dan dimengerti. Untuk diterapkan dalam menggali atau memahami data saham, penggunaan *data mining* memerlukan berbagai perangkat lunak analisis data untuk menemukan pola dan relasi data [9].

Data mining mampu mengekstrak pengetahuan atau menemukan pola dari sekumpulan database, yang dapat memberikan informasi penting yang sifatnya implisit dari sekumpulan database yang jumlahnya banyak [6], dan hal ini akan sulit dilakukan oleh system yang bersifat

konvensional. *Data mining* disebut juga *knowledge discovery in data base* (menemukan pengetahuan dari database), kegiatannya meliputi pengumpulan, penggunaan data, mengetahui sejarah dari pola hubungan data, dan inter relasi set data yang berukuran besar [5]. Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap, dan tahap-tahap tersebut bersifat interaktif dimana pemakai terlibat langsung atau dengan perantara *knowledge base*. Tahapan tersebut diantaranya:



Gambar 1. Langkah-langkah untuk melakukan *Data Mining* (Kusrini & Luthfi, 2009)

Penjelasan gambar 1. adalah:

1) Data Warehouse

Transaksi yang dilakukan di Biro Marketing yang ada di kampus Perbanas Institute dalam setiap tahun penerimaan mahasiswa baru disimpan dalam beberapa database yang terpisah, pada tahun berjalan masih ada di Biro Pemasaran dan sebagian besar ada di Biro LPTI (Layanan Perpustakaan dan Teknologi Informasi). Sudah bertahun-tahun data ini disimpan dan hanya digunakan untuk keperluan tertentu tanpa melihat adanya nilai yang bermakna dari data tersebut yang dapat dianalisa dan mampu memberikan informasi yang dibutuhkan oleh biro marketing [2]. Timbul pertanyaan mau diapakan data yang begitu banyak tersebut tersimpan dalam sebuah computer, apakah dibiarkan saja menjadi data sampah, menghabskan media penyimpanan atau bahkan dapat menjadikan error bila data yang begitu besar tidak dipelihara dengan baik.

Oleh karena itu timbulah gagasan untuk mengambil manfaat dari database yang Selama ini disimpan di Biro

LPTI, agar mampu memberikan informasi yang menjadi pengetahuan dan digunakan untuk bahan pertimbangan pimpinan dalam mengambil keputusan untuk mencari mahasiswa baru yang tepat sasaran. Proses ETL (*Extract, Transform, Loading Data*) dilakukan untuk mendapatkan sekumpulan data yang banyak dan tersimpan dalam bentuk subjek-subjek, terintegrasi, memiliki jangka waktu tertentu atau dimensi waktu dan memiliki ketetapan data (*non-volatile*). Bila kondisi data yang banyak tersebut telah memiliki sifat-sifat tersebut, maka data tersebut sudah tersimpan dalam media yang disebut *data warehouse*.

Tahapan yang dilakukan untuk menjadikan *data warehouse* antara lain dilakukan pembersihan data (*data cleaning*) dan menggabungkan data (*integration*). Penjelasan dari rangkaian proses pembuatan data mining seperti pada gambar 1 di atas, beberapa tahap selanjutnya dengan penjelasan di bawah ini:

2) Pembersihan data (*Pre-processing / Cleaning*)

Pembersihan data maksudnya adalah database PMB (penerimaan mahasiswa baru) yang berisi banyak data, dimungkinkan memiliki isian yang tidak lengkap, pengulangan atau redudansi, terjadinya kesalahan cetak, biasanya ada saja data yang hilang, tidak valid, tidak sesuai kebutuhan atau ada juga terkait atribut / field yang tidak dibutuhkan atau tidak sesuai dengan kebutuhan pengolahan data mining. Ketidaksesuaian data ini harus dibersihkan (tidak diikutkan/dibuang) karena mempengaruhi kehandalan data. Memisahkan data yang tidak baik inilah yang disebut dengan pembersihan data (*cleaning*). Dari data yang diperoleh selama periode tahun 2013 sampai 2015 terdapat banyak data yang harus dibersihkan karena masalah yang dihadapi seperti disebutkan di atas.

3) Integrasi data (*Data integration*)

Pembentukan database untuk menjadi *data mining* yang berasal dari banyak file dilakukan dengan menggabungkan banyak data tersebut menjadi satu (integrasi). Integrasi dibuat dengan menggabungkan atribut-atribut yang unik dari setiap *file* yang akan digabungkan, seperti atribut nama siswa, asal sekolah, asal wilayah, jurusan sekolah, wilayah sekolah.

4) Transformasi data (*data diubah menjadi bentuk kategorikal*)

Data mining membutuhkan format data yang khusus sesuai dengan kebutuhan aplikasi weka. Data yang berasal dari format yang tidak sesuai dengan kebutuhan pembacaan *data mining*, dilakukan perubahan dahulu dengan melalui tahapan yang disebut *mining*. Transformasi yang dilakukan contohnya pada data alamat sekolah atau wilayah sekolah harus dirubah menjadi data kategorikal sesuai kebutuhan system aplikasi, supaya dapat dieksekusi. Tahapan ini menentukan kualitas data yang akan diolah, karena tahapan

ini menentukan karakteristik data dari *data mining* dan tahapan pengolahan selanjutnya bergantung pada kualitas pengolahan pada tahap transformasi data ini.

5) *Patterns yang ditemukan (untuk mendapatkan pola data yang bernilai)*

Hasil keluaran yang berupa pola asosiasi antar atribut data yang berelasi dari pengolahan data, akan dijelaskan pada bab hasil dan pembahasan. Penggunaan teknik *data mining* yang berupa model pola keterkaitan antar parameter yang diukur ataupun pola kesamaan yang berkelompok-kelompok atau yang lainnya, yang tentunya dapat memberikan gambaran dari hubungan antar atribut yang saling berasosiasi membentuk pola (*patterns*) dari seluruh database siswa yang menjadi mahasiswa berdasarkan pola yang sebelumnya telah ditetapkan. Hasil keluaran dari sebuah model pola dapat saja tidak sesuai, maka dapat dicari lagi dengan menggunakan nilai *support* dan *confidence* yang dirubah-rubah. Dengan menampilkan keluaran secara visualisasi dari pola data yang dihasilkan, memungkinkan untuk dapat membantu pengambil keputusan dalam memahami *data mining* dengan mudah dan jelas.

Sebagai contoh pola hubungan dari program studi dan asal sekolah $mnj=P$ $fti=T$ 486 \implies $akun=T$ SMA= P SMK= T 390, *confidence* 0,8. Dengan menggunakan nilai *minimum support sebesar 10 % dan nilai confidence 80%*, menyatakan 80% dari transaksi di database memuat pilihan manajemen dan bukan pilihan *fti* sebanyak 486 transaksi, juga akan memuat bukan pilihan akuntansi dan pilihan SMA serta bukan pilihan SMK sebanyak 390 transaksi. Sedangkan 10% dari seluruh transaksi yang ada di database memuat kelima item tersebut. Pola hubungan tersebut merupakan kemampuan *association rule mining* dan algoritma apriori dalam menemukan aturan asosiatif atau pola hubungan antar satu atau lebih item dalam suatu dataset, sehingga dengan ARM dan algoritma apriori dapat mengetahui kepeminatan program studi dari berbagai kriteria siswa yang akan dibahas pada bab hasil dan pembahasan.

6) *Presentasi pola yang ditemukan untuk menghasilkan aksi*

Pada tahap presentasi pola hasil dari pengolahan *data mining* ini, harus dipastikan dapat dibaca dan dimengerti oleh orang yang tidak memahami *data mining*. Bentuk keluaran yang berupa pengetahuan, dimana hasil keluaran yang membentuk pola dengan jelas dapat dibaca oleh siapa saja yang akan menggunakannya, dalam hal ini pimpinan di biro marketing institute Perbanas. Untuk kemudahan dalam pembacaan keluaran, maka dibuat dalam bentuk visualisasi yang mudah dimengerti dan dipahami, sehingga hasil dari *data mining* dapat dijadikan bahan pertimbangan untuk mengambil keputusan dalam melakukan kegiatan pencarian calon mahasiswa baru yang tepat sasaran.

Hasil dari pengolahan data dapat mengetahui hubungan asosiatif antara data program studi dengan wilayah asal sekolah dari mahasiswa, yang muncul dalam aturan asosiasi yaitu memunculkan nilai tingkat kepercayaan (*confidence level*) dari hubungan tersebut dan seringkali muncul dari data prodi dan wilayah, asal sekolah, jurusan siswa yang menjadi mahasiswa (*Minimum support*).

B. Analisa Pola Frekuensi Tinggi

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

$$support(A) = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total Transaksi}} \dots \dots \dots (1)$$

Sedangkan nilai *support* dari 2 *item* diperoleh dari rumus berikut:

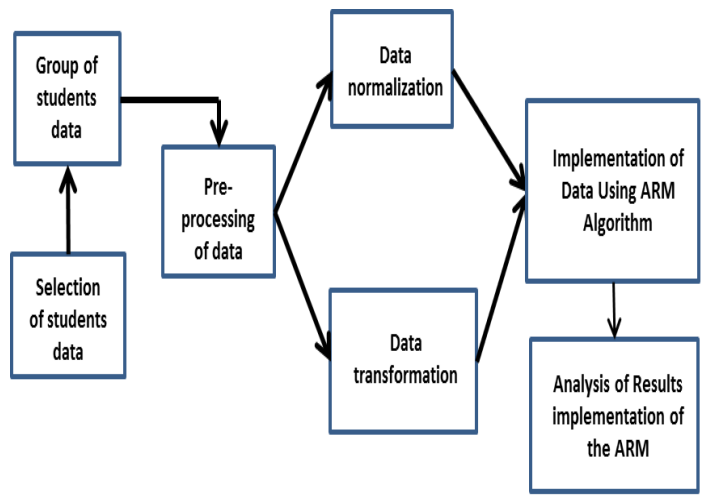
$$support(A \cap B) = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Total Transaksi}} \dots (2)$$

C. Pembentukan Aturan Asosiatif

Setelah semua pola frekuensi tinggi ditemukan, selanjutnya mencari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A→B. Nilai *confidence* dari aturan A→B diperoleh dari rumus berikut:

$$confidence = P(B|A) = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Jumlah transaksi mengandung A}} \dots (3)$$

Kerangka Berfikir



Gambar 2. Kerangka Berpikir penelitian

Tahapan kerangka berpikir mulai dari pemilihan data siswa sampai melakukan analisa dari pola yang dihasil yang terdapat pada gambar 2 di atas, penjelasannya adalah:

1) Pemilihan data siswa yang mendaft.

Dalam hal ini, yang dipilih adalah seluruh data siswa yang mendaftar ke Perbanas Institute, baik yang dari Jabodetabek maupun diluar Jabodetabek. Atribut yang dipilih adalah utamanya pada wilayah sekolah dan asal sekolah.

2) Pengelompokan data siswa

Kelompok dibagi menjadi 6 kelompok data berdasarkan wilayah pendaftar, yaitu Jakarta, Bogor, Depok, Tangerang, Bekasi, dimana kelima wilayah ini disebut

dengan Jabodetabek dan 1 lagi wilayah diluar Jabodetabek. Untuk perwilayah pendaftar nantinya dapat dibagi lagi menjadi per asal sekolah, sehingga dapat diketahui dari wilayah mana jumlah pendaftar calon mahasiswa baru dan dari wilayah tersebut dapat diketahui dari sekolah mana pendaftarnya terbanyak.

3) Pra-pemrosesan data.

Setelah tahap pemilihan data, data kemudian memasuki tahap pra-pemrosesan data dimana dalam tahap ini terdapat proses normalisasi dan transformasi data:

4) Normalisasi Data

Metode ini bertujuan untuk mendapatkan hasil dari setiap pendaftar dari wilayah dan sekolah mana siswa itu berasal, dimana data ini dihitung dari pendaftaran harian. Sebelum dilakukan normalisasi data, terlebih dahulu menghitung nilai rata-rata dari setiap data pendaftaran perhari.

5) Transformasi Data

Setelah dilakukan normalisasi, data kemudian ditransformasi dari numerik menjadi kategorikal. Hal ini dilakukan agar data bisa di analisis dan diolah dengan menggunakan program WEKA 3.6.10 [7], karena jika ada salah satu atribut yang masih berbentuk numerik, program tidak dapat berjalan dengan baik dan tidak bisa menghasilkan aturan-aturan yang menarik. Sebelum dilakukan transformasi, terlebih dahulu dilakukan penghitungan dari setiap data perhari pada setiap pendaftaran, guna mencari asal wilayah dan asal sekolah dari setiap pendaftar.

6) Melakukan pengujian data set pendaftar yang telah dipetakan.

Data digunakan sepenuhnya sebagai training data set. Metode ini bertujuan untuk menganalisa dan mendapatkan rules yang menarik dari data pendaftaran mahasiswa baru yang digunakan, dan menggunakannya untuk memahami perilaku data tersebut

7) Model rule terpilih

Pada pengujian data set dengan menggunakan aplikasi WEKA, didapat model terpilih yang dapat menggambarkan pergerakan penerima mahasiswa baru dari tiap program studi selama 3 tahun terakhir, sehingga dapat mengetahui kenaikan dan penurunan jumlah penerimaan dari tiap program studi, juga dapat menampilkan mahasiswa yang diterima berdasarkan jenis kelamin, asal SMA / SMK, asal daerah/wilayah tempat SMA / SMK itu berada, untuk selanjutnya dapat mengetahui prosentase dari tiap program studi yang diterima di kampus ini, serta inti dari penelitian ini adalah untuk mengetahui prosentase yang mendaftar dan yang diterima, terkait dengan asal smu, sehingga mengetahui berapa jumlah yang mengundurkan diri dan dari daerah mana asal smu yang mengundurkan diri. Hal ini terkait dengan strategi pencarian pendaftaran mahasiswa baru.

PENELITIAN YANG TELAH DILAKUKAN

No	Peneliti	Tahun	Obyek Penelitian	Hasil Penelitian
1	Rumaisa, F	2012	Perguruan Tinggi, Univ. Widyatama Bandung	ARM dengan Algoritma Apriori dapat mengetahui kepeminatan siswa atas program studi, dengan mengetahui asosiasi antar prodi untuk pilihan 1 dan pilihan 2, 3, dst
2	Kurniawati, A	2014	Perguruan Tinggi, UNNES	Algoritma Apriori dapat mengetahui pola pemetaan hubungan program studi dengan pilihan mahasiswa baru Unnes
3	Beta Noranita dan Nurdin Bahtiar	2010	Perguruan Tinggi, UNDIP	aplikasi <i>data mining</i> ini dapat digunakan untuk mengetahui factor yang mempengaruhi tingkat kelulusan yang diukur oleh nilai <i>support</i> dan <i>confidence</i> antar item.
4	Meidi Dian LP, Dian Eka R dan Marji	2014	Perguruan Tinggi	Dapat mengetahui pola penerimaan dan kelulusan Fakultas Ekonomi dan Bisnis Universitas Brawijaya, yang dapat diketahui dari nilai pola asosiasi yang dihasilkan.

TABEL I

D. Keterkaitan Penelitian

Beberapa penelitian yang berhubungan dengan perencanaan strategi untuk mendapatkan mahasiswa baru telah dilakukan oleh para peneliti sebelumnya seperti yang ada pada tabel I, menjelaskan dengan menggunakan *Association Rule Mining* (ARM) dan algoritma apriori, seperti yang telah dilakukan di perguruan tinggi Univ. Widyatama Bandung, membuktikan bahwa algoritma tersebut dapat menggambarkan hubungan antar program studi dari peminatan mahasiswa sebagai pilhan 1, pilihan 2 dan pilihan 3 [12]. Kurniawati Anis juga melakukan penelitian menggunakan ARM yang menghasilkan nilai *confidence* yang bervariasi dari asosiasi antar jurusan yang ada di UNNES [6]. Hasil penelitian lainnya yang dilakukan di kampus UNDIP, yang menunjukkan bahwa data mining dapat digunakan untuk mengetahui factor yang mempengaruhi tingkat kelulusan yang diukur oleh nilai *support* dan *confidence* antar item [10], dan penelitian yang dilakukan di UB menyimpulkan bahwa data mining dapat melihat pola hubungan penerimaan dan kelulusan mahasiswa di Fakultas Ekonomi dan Bisnis yang ditunjukkan oleh nilai *minimum support* dan *minimum confidence* dari pola asosiasi yang dihasilkan [1].

Penelitian yang peneliti lakukan adalah lebih dititik beratkan pada data calon mahasiswa yang memilih program studi, dimana data yang diekstrak jauh lebih rinci dari penelitian sebelumnya, yaitu data wilayah tempat tinggal, wilayah asal sekolah, jurusan saat sekolah dan hubungannya dengan kepeminatan terhadap program studi yang menjadi pilihannya, sehingga program studi dapat mengetahui persis pendaftar yang menjadi mahasiswanya. Kontribusi hasil penelitian ini sangat berguna untuk Biro pemasaran yang ada di kampus Perbanas Institute dalam melakukan perencanaan promosi ke SMU yang dikunjunginya. Dari asosiasi rule yang ditunjukkan dapat mengetahui wilayah dan sekolah yang memiliki potensi untuk menjadi mahasiswa pada program studi yang dipilihnya, untuk menjadi prioritas dikunjungi tim promosi.

III. METODE PENELITIAN

A. Jenis Penelitian

Pengujian menggunakan metode Deskriptif (*Description Methods*), yaitu menjelaskan dan menggambarkan suatu objek dengan kenyataan yang ada. Analisa yang dilakukan dengan fungsi Analisis Asosiasi (*Association Rule Mining*) dan algoritma apriori. Sedangkan Aplikasi untuk melakukan pengujiannya dengan WEKA 3.6 [11] yang menghasilkan *rule* asosiatif yang ditunjukkan dengan dua parameter, *support* (nilai penunjang) yaitu persentase kombinasi item tersebut dalam *database* dan *confidence* (nilai kepastian) yaitu kuatnya hubungan antar item yang ada dalam *database* dengan aturan asosiatif. Untuk Instrumen penelitian, karena penelitian ini tidak menggunakan kuesioner dalam memperoleh data, maka dalam penelitian ini tidak menggunakan instrument penelitian. Data yang diolah

berasal dari transaksi pendaftaran mahasiswa baru (PMB) secara harian selama 3 tahun yang telah terjadi, yaitu tahun 2013 sampai 2015.

B. Populasi Dan Sampel

Dalam penelitian yang dilakukan ini, data yang digunakan adalah data yang bersumber dari Biro Layanan Perpustakaan dan Teknologi Informasi (LPTI) dan biro Marketing. Setiap tahun penerimaan mahasiswa baru ada sekitar 1500 mahasiswa baru, tanda dari pendaftar menjadi mahasiswa adalah dengan diberikannya nomor induk mahasiswa (NIM), sementara yang mendaftar ada sekitar 2000 sampai 2500 siswa pertahun. Melihat persaingan begitu ketat, dengan perhitungan keketatan yang sudah disampaikan seperti pada pendahuluan di atas, dan ini terjadi di Institut Perbanas dimana sejak tahun 2013 sampai 2015, pendaftar mulai berkurang, yaitu hanya disekitar 2000 siswa, sehingga peneliti tertarik untuk melakukan penelitian dengan sampel dari tahun 2013 sampai 2015. Informasi dari hasil penelitian ini diharapkan dapat menjadi masukan yang baik untuk membantu tim *marketing* dalam melakukan strategi pemasaran untuk mendapatkan mahasiswa baru.

C. Metode Pengumpulan Data

Dalam mengumpulkan data, peneliti menggunakan data sekunder yang ada di *database* yang berada di server biro PTI dan biro Marketing, dengan cara mengunduh dari server, yang diambil selama 3 tahun (2013 – 2015), tidak menggunakan metoda kuesioner atau lainnya yang biasa digunakan untuk mendapatkan data primer.

D. Metode Analisis Data

Setelah melalui prapemrosesan data dengan melakukan normalisasi dan transformasi data, selanjutnya dengan menggunakan program aplikasi WEKA 3.6. [11] data dimasukan untuk diketahui nilai hubungannya melalui 2 jenis hubungan data, yaitu nilai pada *support* dan nilai *confidence*. Kedua nilai ini akan menggambarkan tingkat hubungan antar data siswa, dari yang mendaftar dengan yang diterima dan menjadi nomor induk mahasiswa (NIM).

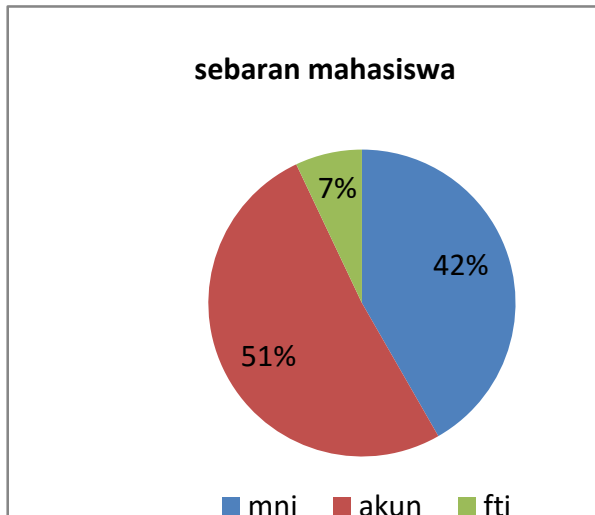
Hubungan antar data tersebut diekstrak berdasarkan wilayah lokasi siswa yang mendaftar dan berbagai atribut lainnya yang ada pada *entity* siswa, yang akan dibandingkan keterkaitannya dengan siswa yang diterima, sehingga penelusuran jejak siswa yang diterima dapat diekstrak dan selanjutnya dapat dijadikan bahan masukan untuk melakukan perencanaan strategi pemasaran ditahun berikutnya, artinya kegiatan pemasaran bisa lebih besar prosentasinya pada daerah yang menjadi langganan mengirim siswanya menjadi mahasiswa di Institut Perbanas.

IV. HASIL DAN PEMBAHASAN

A. Deskripsi Obyek Penelitian

1) Sebaran mahasiswa di setiap program studi

Untuk sebaran penerimaan mahasiswa baru selama perioda tiga (3) tahun, dari data yang terkumpul terlihat pada gambar 3 di bawah ini, dimana jumlah mahasiswa yang diterima untuk jurusan manajemen adalah 42%, untuk jurusan Akuntansi adalah 51%, dan untuk jurusan pada fakultas teknologi informasi (FTI) adalah 7%.



Gambar 3. Grafik Sebaran mhs di setiap prodi

2) Sebaran wilayah asal mahasiswa

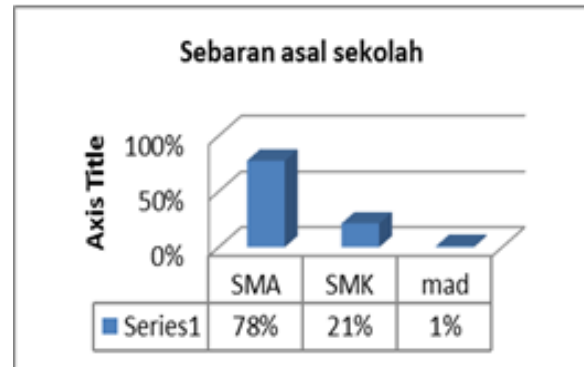


Gambar 4. Grafik Sebaran wilayah asal mahasiswa

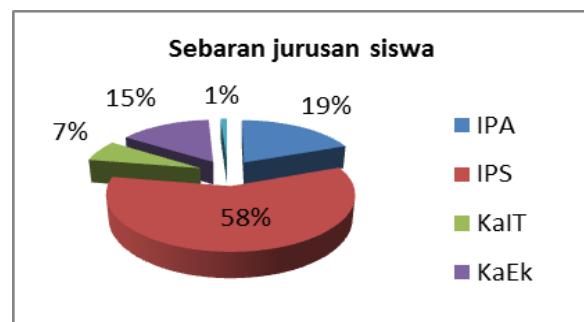
Seperti yang terlihat pada gambar 4 di atas, memperlihatkan sebaran wilayah asal mahasiswa berasal. Jakarta merupakan wilayah yang paling besar, yaitu 55,6%, selanjutnya Bekasi sebanyak 14,6% dan urutan berikutnya adalah wilayah ring 2 (jawa tengah dan jawa timur) sebesar

7,6%, Ring 1 (jawa barat, banten sebesar 5,1%), Bogor sebesar 6,1%, pada Ring3 (wilayah sebaran diluar pulau jawa) sebesar 4,5%, sementara Depok dan Tangerang sama pada jumlah 3,5%.

3) Sebaran jurusan asal sekolah dan jurusan saat siswa



Gambar 5. Grafik Sebaran asal sekolah



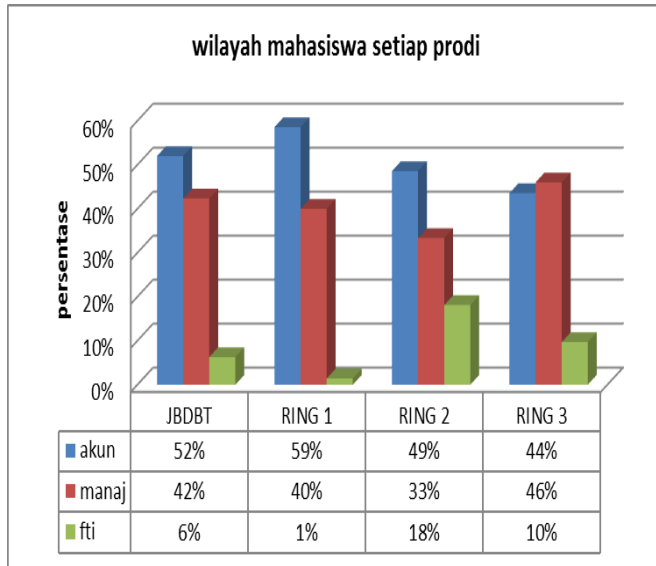
Gambar 6. Grafik jurusan siswa saat sekolah

Pada gambar 5 dan 6 di atas ini menggambarkan asal sekolah dan jurusan asal siswa yang yang menjadi mahasiswa Institut Perbanas, yaitu yang berasal dari SMU adalah sebesar 78%, SMK sebesar 21% mahasiswa dan Madrasah sebanyak 1% mahasiswa. Untuk asal jurusan sekolahnya dari mana, digambarkan pula di atas, yaitu yang berasal dari jurusan IPA sebesar 19%, yang berasal dari IPS sebesar 58%, yang berasal dari SMK yang terkait dengan IT adalah sebesar 7% orang dan yang berasal dari SMK yang jurusannya terkait dengan ekonomi sebesar 15%, sedangkan untuk jurusan yang umum (agama, bahasa, tataboga, dsb) adalah sebesar 1% dari mahasiswa yang diterima selama perioda 2013 - 2015.

4) Sebaran mahasiswa di setiap prodi berdasarkan asal wilayah saat siswa

Untuk mengetahui wilayah pemilihan calon mahasiswa yang mendaftar ke institut Perbanas, peneliti menggambarkan dalam bentuk grafik sebaran wilayah yang disempitkan menjadi 3 wilayah besar dengan membuat

kelompok kode wilayah, yaitu dengan kode JBDBT (wilayah meliputi area asal sekolah di Jakarta, Bogor, Depok, Bekasi dan Tangerang), berikutnya kode Ring 1 (wilayah meliputi area asal sekolah jawa barat dan banten), Ring 2 (wilayah meliputi area asal sekolah jawa tengah dan jawa timur), serta yang terakhir adalah kode Ring 3 (wilayah meliputi area asal sekolah di luar pulau jawa).



Gambar 7. Grafik mhs per prodi berdasar wilayah siswa

Pada gambar 7 di atas merupakan perbandingan sebaran jumlah mahasiswa dari wilayah mana saja untuk setiap prodi yang ada. Untuk wilayah JBDBT, jumlah mahasiswa yang diterima di program studi Akuntansi sebesar 52%, Manajemen sebesar 42% dan untuk prdi yang ada di FTI sebesar 6%. Untuk wilayah lainnya, yaitu Ring1, jumlah mahasiswa yang diterima di program studi Akuntansi sebesar 59%, Manajemen sebesar 40% dan untuk prdi yang ada di FTI sebesar 1%. Begitu juga untuk persentase komposisi Ring2 dan Ring3

B. Pengujian data set

Pengujian beberapa atribut pada file yang ada di database mahasiswa dipilih untuk menghasilkan nilai confidence yang optimal, pada system aplikasi yang digunakan dilakukan dengan mengubah rentan nilai dari dua variabel yang terdapat dalam algoritma apriori yaitu *Minimum Support* dan *Minimum Metric (Confidence)*. seperti yang dapat dilihat pada hasil pengujian tabel dibawah ini:

1) Data set 1 (pertama)

Pengujian atribut dilakukan untuk mengetahui sebaran mahasiswa di setiap prodi yang diterima di kampus Institut Perbanas berasal dari wilayah asal sekolah dengan kode wilayah yang telah disebutkan sebelumnya, menggunakan

nilai nilai *minimum support 0.1* dan *minimum metric (confidence)* yang digunakan 0,65 dan dengan jumlah rule yag dipilih adalah 100.

TABEL II
PENGUJIAN DATA TES 1 (PERTAMA)

1	JBDBK=P RING 1=T RING 2=T RING 3=T mnj=T 660	==> akun=P fti=T 588	conf:(0.89)
2	JBDBK=P RING 2=T RING 3=T mnj=T 660	==> RING 1=T akun=P fti=T 588	conf:(0.89)
3	JBDBK=P RING 1=T RING 3=T mnj=T 660	==> RING 2=T akun=P fti=T 588	conf:(0.89)
4	JBDBK=P RING 1=T RING 2=T mnj=T 660	==> RING 3=T akun=P fti=T 588	conf:(0.89)
5	JBDBK=P RING 3=T mnj=T 660	==> RING 1=T RING 2=T akun=P fti=T 588	conf:(0.89)
6	mnj=P fti=T 581	==> JBDBK=P RING 2=T akun=T 488	conf:(0.84)
7	mnj=P akun=T 581	==> JBDBK=P RING 2=T fti=T 488	conf:(0.84)
8	mnj=P 581	==> JBDBK=P RING 2=T akun=T fti=T 488	conf:(0.84)
9	mnj=P akun=T fti=T 581	==> JBDBK=P RING 3=T 488	conf:(0.84)
10	akun=T fti=T 581	==> JBDBK=P RING 3=T mnj=P 488	conf:(0.84)
11	mnj=P fti=T 581	==> JBDBK=P RING 3=T akun=T 488	conf:(0.84)
12	mnj=P akun=T 581	==> JBDBK=P RING 1=T RING 2=T RING 3=T 488	conf:(0.84)
13	mnj=P 581	==> JBDBK=P RING 1=T RING 2=T RING 3=T akun=T 488	conf:(0.84)
14	mnj=P fti=T 581	==> JBDBK=P RING 1=T RING 2=T RING 3=T 488	conf:(0.84)
15	akun=T 679	==> JBDBK=P RING 1=T RING 2=T mnj=P fti=T 488	conf:(0.72)
16	akun=T 679	==> JBDBK=P RING 1=T RING 3=T mnj=P fti=T 488	conf:(0.72)
17	akun=T 679	==> JBDBK=P RING 2=T RING	conf:(0.72)

		3=T mnj=P fti=T 488	
18	akun=T 679	==> RING 1=T RING 2=T RING 3=T mnj=P fti=T 488	conf:(0.72)
19	fti=T 1295	==> JBDBK=P RING 1=T RING 2=T 1076	conf:(0.83)
20	fti=T 1295	==> JBDBK=P RING 1=T RING 3=T 1076	conf:(0.83)

Pada tabel II di atas ini menggunakan *minimum support* dengan nilai 0,1, yang artinya persentase dari kemunculan *item set* dalam membentuk jumlah *rule* adalah kurang dari atau sama dengan 10%, dan nilai *minimum metric (confidence)* yang ada pada aplikasi diberikan nilai 0,65 yang artinya tingkat kepercayaan atau tingkat kebenaran dari *rule* yang terbentuk minimum pada nilai 65% dan maksimal 100%, dengan jumlah *rule* yang akan ditampilkan sebanyak 20 *rule* dari 100 *rule* yang dihasilkan dari system aplikasi, dengan data mahasiswa dari pendaftar dan yang diterima pada tahun ajaran 2013 – 2015. Sebagai gambaran untuk *rule* pada nomor 1, menunjukkan bahwa wilayah jabodetabek (JBDBK) menjadi pilihan dan wilayah Ring1, Ring2, Ring3 yang muncul sebanyak 660 kali, akan diikuti oleh akuntansi yang menjadi pilihan dan FTI tidak menjadi pilihan yang muncul sebanyak 588 kali, sehingga menghasilkan nilai *confidence* 0,89, yang artinya sebanyak 89% kemunculan JBDBK merupakan wilayah pemilih yang menjadi mahasiswa baru di Institut Perbanas yang diikuti oleh kemunculan prodi Akuntansi yang menjadi pilihan dan prodi FTI yang bukan menjadi pilihan dengan hubungan kemunculan transaksi sebanyak 660 dan 588 kali. Penjelasan yang lainnya pada nomor 12, menunjukkan prodi manajemen yang menjadi pilihan dan prodi akun yang bukan menjadi pilihan yang muncul sebanyak 581 kali, akan diikuti oleh wilayah JBDBK yang menjadi wilayah pilihan dan Ring1, Ring2, Ring3 yang bukan menjadi pilihan yang muncul sebanyak 488 kali, sehingga menghasilkan nilai *confidence* 0,84, yang artinya sebanyak 84% data menyatakan prodi manajemen menjadi pilihan dan bukan prodi akuntansi yang diikuti oleh JBDBK yang menjadi wilayah pemilih dan bukan wilayah Ring1, Ring2 dan Ring3 dengan hubungan kemunculan transaksi sebanyak 581 dan 488 kali. Pola asosiasi ini menggambarkan perbandingan antara prodi yang dipilih dengan wilayah yang memilih, sehingga setiap prodi dapat mengetahui wilayah mana yang menjadi pemilih terbanyak yang muncul secara bersamaan dengan prodi lain yang tidak dipilih dan wilayah lainnya yang bukan wilayah pemilih pada setiap *rule* yang ada pada aturan asosiasi dengan ARM.

2) Data set 2 (kedua)

Pengujian atribut dilakukan untuk mengetahui jumlah mahasiswa yang diterima di kampus institut Perbanas yang dipilih berdasarkan asal sekolah yang daftar (SMU, SMK atau Madrasah), dengan ketentuan nilai *minimum support* yang digunakan 0.1 dan *minimum metric (confidence)* yang digunakan 0,65 dan dengan jumlah *rule* 100.

TABEL III.
PENGUJIAN DATA TES 2 (KEDUA)

1	akun=T fti=T SMA=P SMK=T 396	==> mnj=P 390	conf:(0.98)
2	mnj=P fti=T SMK=T 396	==> akun=T SMA=P 390	conf:(0.98)
3	akun=T fti=T SMA=P SMK=T mad=T 396	==> mnj=P 390	conf:(0.98)
4	mnj=P akun=T fti=T SMK=T 396	==> SMA=P mad=T 390	conf:(0.98)
5	akun=P fti=T SMK=T mad=T 474	==> SMA=P 462	conf:(0.97)
6	mnj=T akun=P SMK=T mad=T 474	==> fti=T SMA=P 462	conf:(0.97)
7	akun=P SMK=T mad=T 474	==> mnj=T fti=T SMA=P 462	conf:(0.97)
8	akun=T fti=T SMK=T 402	==> mnj=P SMA=P mad=T 390	conf:(0.97)
9	mnj=T fti=T SMK=T mad=T 480	==> akun=P SMA=P 462	conf:(0.96)
10	akun=P fti=T SMK=T 480	==> mnj=T SMA=P mad=T 462	conf:(0.96)
11	mnj=T fti=T SMK=T 486	==> akun=P SMA=P mad=T 462	conf:(0.95)
12	mnj=T SMA=P SMK=T 528	==> akun=P fti=T 462	conf:(0.88)
13	akun=T SMA=P SMK=T 456	==> mnj=P fti=T 390	conf:(0.86)
14		==> akun=T fti=T SMA=P mad=T 462	conf:(0.85)
15	mnj=P mad=T 480	==> SMA=P 390	conf:(0.81)
16	mnj=P akun=T fti=T 486	==> SMA=P 390	conf:(0.8)

Pada tabel III di atas ini menggunakan *minimum support* dengan nilai 0,1 yang artinya persentase dari kemunculan *item set* dalam membentuk jumlah *rule* adalah kurang dari atau sama dengan 10%, dan nilai *minimum metric*

(confidence) yang ada pada aplikasi diberikan nilai 0,65, yang artinya tingkat kepercayaan atau tingkat kebenaran dari rule yang terbentuk minimum pada nilai 65% dan maksimal 100%, dengan jumlah rule yang dipilih sebanyak 100 untuk data mahasiswa dari pendaftar dan diterima pada tahun ajaran 2013 – 2015, sedangkan untuk menjelaskan asosiasi rule pada table III di atas yang ditampilkan hanya sebanyak 16 rule saja. Sebagai gambaran untuk rule pada nomor 4, menunjukkan bahwa ada transaksi yang muncul sebanyak 396 kali yang memilih prodi mnj dan tidak memilih prodi akuntansi dan FTI dan tidak berasal dari SMK dan diikuti oleh SMA yang memilih Perbanas dan Madrasah yang tidak memilih dan muncul sebanyak 390 kali, sehingga menghasilkan nilai confidence 0,98, yang artinya sebanyak 98% kemunculan data transaksi memilih mnj, tidak memilih fti, bukan dari SMK dan bukan dari mad dan berasal dari SMA. Contoh lainnya pada rule no. 9 menunjukkan bahwa kemunculan mnj, fti, SMK dan mad sebanyak 480 dan semuanya bukan menjadi pilihan, diikuti oleh akun dan SMA sebanyak 462 dan merupakan pilihan, sehingga menghasilkan nilai confidence 0,96 yang artinya sebanyak 96% kemunculan prodi mnj, fti, dan asal sekolah SMK serta mad yang tidak memilih muncul sebanyak 480 kali diikuti oleh prodi akun yang dipilih dan asal sekolah SMA yang dipilih yang muncul sebanyak 462 kali.

3) Data set 3 (ketiga)

Pengujian atribut dilakukan untuk mengetahui jumlah mahasiswa yang diterima di kampus institut Perbanas yang dipilih berdasarkan jurusan apa dari asal sekolahnya (IPA, IPS, KaIT, KaEk, UMUM), dengan ketentuan nilai minimum support yang digunakan 0.1 dan minimum metric (confidence) yang digunakan 0,65 dan dengan jumlah rule 100.

TABEL IV.
PENGUJIAN DATA TES 3 (KETIGA)

1	akun=P IPA =T KaEk=T 450	==> mnj=T fti=T IPS=P KaIT=T UMUM=T 428	conf:(0.95)
2	akun=T fti=T IPA =T KaIT=T KaEk=T 371	==> mnj=P IPS=P UMUM=T 351	conf:(0.95)
3	mnj=P fti=T IPA =T KaEk=T UMUM=T 380	==> akun=T IPS=P KaIT=T 351	conf:(0.92)
4	mnj=P akun=T IPA =T KaEk=T UMUM=T 380	==> fti=T IPS=P KaIT=T 351	conf:(0.92)
5	mnj=P IPA =T KaEk=T 387	==> akun=T fti=T IPS=P KaIT=T UMUM=T 351	conf:(0.91)
6	akun=T fti=T IPA =T KaEk=T UMUM=T 386	==> mnj=P IPS=P KaIT=T 351	conf:(0.91)

7	akun=P IPA =T KaEk=T UMUM=T 467	==> IPS=P KaIT=T 391	conf:(0.84)
8	mnj=P KaIT=T 547	==> akun=T fti=T KaEk=P UMUM=T 456	conf:(0.83)
9	mnj=P akun=T IPA =T KaIT=T UMUM=T 442	==> fti=T IPS=P KaEk=T 351	conf:(0.79)
10	mnj=P IPA =T KaIT=T UMUM=T 442	==> akun=T fti=T IPS=P KaEk=T 351	conf:(0.79)
11	akun=T fti=T IPA =T KaIT=T UMUM=T 448	==> mnj=P IPS=P 351	conf:(0.78)
12	akun=T fti=T IPA =T KaIT=T UMUM=T 448	==> mnj=P IPS=P KaEk=T 351	conf:(0.78)
13	akun=P IPA =T KaIT=T UMUM=T 554	==> IPS=P KaEk=T 428	conf:(0.77)
14	akun=P IPA =T KaIT=T 554	==> IPS=P KaEk=T UMUM=T 428	conf:(0.77)
15	mnj=P akun=T fti=T KaIT=T KaEk=T UMUM=T 456	==> IPA =T IPS=P 351	conf:(0.77)
16	mnj=P fti=T KaIT=T KaEk=T UMUM=T 456	==> akun=T IPA =T IPS=P 351	conf:(0.77)

Pada tabel IV di atas ini menggunakan minimum support dengan nilai 0,1 yang artinya persentase dari kemunculan item set dalam membentuk jumlah rule adalah kurang dari atau sama dengan 10%, dan nilai minimum metric (confidence) yang ada pada aplikasi diberikan nilai 0,65, yang artinya tingkat kepercayaan atau tingkat kebenaran dari rule yang terbentuk minimum pada nilai 65% dan maksimal 100%, dengan jumlah rule yang dipilih sebanyak 100 untuk data mahasiswa dari pendaftar dan diterima pada tahun ajaran 2013 – 2015, sedangkan untuk menjelaskan asosiasi rule pada tabel di atas yang ditampilkan hanya sebanyak 16 rule saja. Sebagai gambaran untuk rule pada nomor 1, menunjukkan bahwa ada transaksi yang muncul sebanyak 450 kali yang memilih prodi akun dan bukan berasal dari jurusan IPA dan bukan dari jurusan yang ada kaitannya dengan ekonomi (KaEk) dan diikuti oleh yang tidak memilih mnj, fti, KaIT dan UMUM serta jurusan IPS yang memilih prodi akun yang muncul sebanyak 428 kali, sehingga menghasilkan nilai confidence 0,95, yang artinya sebanyak 95% kemunculan data transaksi memilih prodi Akuntansi, bukan dari IPA, KaEk, yang diikuti oleh mnj, fti, IPS, KaIT, UMUM yang tidak memilih dan dipilih yang muncul sebanyak 428 kali transaksi dari IPS. Contoh lainnya pada rule no. 5 menunjukkan manajemen (mnj) menjadi pilihan dan bukan dari jurusan IPA dan KaEk yang muncul sebanyak 387 transaksi diikuti oleh bukan memilih

akun, fti, dan bukan jurusan KaIT dan UMUM tetapi berasal dari jurusan IPS merupakan pilihan, yang muncul sebanyak 351, sehingga menghasilkan nilai *confidence* 0,91 yang artinya sebanyak 91% prodi mnj menjadi pilihan bukan dari jurusan IPA, KaEk, yang diikuti oleh bukan memilih akun dan fti dan bukan berasal dari jurusan KaIT dan UMUM tetapi berasal dari jurusan IPS.

V. KESIMPULAN

Setelah melakukan penelitian pada lingkup sebaran mahasiswa yang ada di lingkungan kampus Institut Perbanas, maka dapat disimpulkan aplikasi *data mining* dengan *association rule mining* dan algoritma apriori dapat membangun *knowledge base* dari sekumpulan database mahasiswa yang diterima di Institut Perbanas dengan memperlihatkan hubungan antara identitas mahasiswa dengan program studi yang menjadi pilihan mahasiswa, baik data wilayah asal mahasiswa tinggal (Jabodetabek, pulau Jawa dan luar pulau Jawa), atau asal sekolah saat menjadi siswa (SMU, SMK atau Madrasah), dan juga jurusan yang ditempuhnya pada saat siswa (IPA, IPS atau yang lainnya).

Dengan mengetahui sebaran mahasiswa pada setiap program studi, tim dari biro pemasaran penerimaan mahasiswa baru dapat menggunakan informasi hasil penelitian ini untuk menjadi salah satu strategi dalam melakukan promosi ke sekolah-sekolah. Urutan prioritas untuk promosi dilakukan utamanya di wilayah Jabodetabek adalah Jakarta, selanjutnya ke Bekasi, Bogor, Tangerang dan Depok. Selanjutnya untuk diluar Jabodetabek adalah Jawa tengah, Jawa Timur, Jawa Barat dan Banten, dan pemasaran yang ada pada prioritas terakhir adalah diluar pulau Jawa.

Untuk asal sekolah yang menjadi mahasiswa di Institut Perbanas terbanyak ada pada SMU selanjutnya SMK dan Madrasah, sedangkan untuk jurusan siswa SMU dan sederajat yang terbanyak adalah dari IPS selanjutnya IPA dan jurusan lainnya yang ada di SMK (IT, Akuntansi, pemasaran, tata boga dan lainnya). Urutan informasi ini juga dapat dijadikan strategi promosi yang dilakukan biro Pemasaran dalam melakukan sosialisasi kampus Perbanas, dengan tujuan akhir dapat menerima jumlah mahasiswa dengan kuantitas dan kualitas sesuai yang diharapkan.

VI. REKOMENDASI

Untuk melaksanakan penelitian lebih lanjut, ada beberapa hal yang sebaiknya diperhatikan, diantaranya:

1. Database yang digunakan untuk dianalisa, sebaiknya pada bagian pendaftaran (PMB) dipastikan seluruh atribut yang ada di *form* terisi semua, sehingga kebutuhan penelitian yang menggunakan atribut yang ada menjadi mudah dalam melakukan pengolahannya.
2. Penelitian dapat dikembangkan dengan menggunakan algoritma yang berbeda, untuk melihat potensi calon mahasiswa dari berbagai sudut pandang yang berbeda, sesuai dengan peruntukan algoritma yang akan digunakan.

DAFTAR PUSTAKA

- [1] Dian M, et al. (2014). *Penemuan Pola Penerimaan Dan Kelulusan Mahasiswa S1 Fakultas Ekonomi Dan Bisnis Universitas Brawijaya Dengan Menggunakan Algoritma Fp-Growth*. Jurnal Mahasiswa PTIIK UB, Volume 3 nomor 12. Universitas Brawijaya, Malang.
- [2] Freitas, Alex A. *Data mining and knowledge discovery with evolutionary algorithms*. Springer Science & Business Media, 2013.
- [3] Halim, S. (2015). *Faktor-Faktor yang Mempengaruhi Siswa SMA dalam Memilih Perguruan Tinggi* (Doctoral dissertation, Petra Christian University).
- [4] Haryanto, R., & Rozza, S. (2012). *Pengembangan Strategi Pemasaran Dan Manajemen Hubungan Masyarakat Dalam Meningkatkan Peminat Layanan Pendidika. Ekonomi & Bisnis, 11*.
- [5] Han, Jiawei, Jian Pei, and Micheline Kamber. *Data mining: concepts and techniques*. Elsevier, 2011
- [6] Kurniawati, A. (2014). *Pemetaan Pola Hubungan Program Studi Dengan Algoritma Apriori-Studi Kasus Spmu Unnes*. Edu Komputika Journal, 1.
- [7] Kuswari, Hernawati, Insani Nur, and Sumarno Bambang. "Application Of Association Rules With Apriori Algorithm To Determine The Pattern Of The Relationship Between SBMPTN Database And Student's Grade Point Average." *International Seminar on Innovation in Mathematics and Mathematics Education*. Departement of Mathematics Education Faculty of Mathematics and Natural Science Yogyakarta State University, 2014.
- [8] Larose, Daniel T. (2006). *Data Mining Methodes and Models*, John Wiley & Sons, Inc.
- [9] Miranda, Eka, and Eli Suryani. "Implementation of datawarehouse, datamining and dashboard for higher education." *Journal of Theoretical & Applied Information Technology* 64.3 (2014).
- [10] Noratita B., Bactiar N. (2010). *Implementasi Data Mining Untuk Menemukan Pola Hubungan Tingkat Kelulusan Mahasiswa Dengan Data Induk Mahasiswa*. Seminar Nasional Vocational Education in IT Polytecnic; Competitive Advantage in ICT. JIT Telkomuniversiy, Bandung
- [11] Remco R. Bouckaert. (2011), "Weka Manual 3-6-5", Software manual, GNU General Public License version 2
- [12] Rumaisa, F. (2012, June). *Penentuan Association Rule Pada Pemilihan Program Studi Calon Mahasiswa Baru Menggunakan Algoritma Apriori Studi Kasus Pada Universitas Widyatama Bandung*. Seminar Nasional Aplikasi Teknologi Informasi 2012, Jurusan Teknik Informatika, Fakultas Teknologi Industri, Universitas Islam Indonesia, Yogyakarta.