

PENENTUAN ASSOCIATION RULE PADA PEMILIHAN PROGRAM STUDI CALON MAHASISWA BARU MENGGUNAKAN ALGORITMA APRIORI STUDI KASUS PADA UNIVERSITAS WIDYATAMA BANDUNG

Fitrah Rumaisa, S.T., M.Kom

Jurusan Teknik Informatika, Fakultas Teknik, Universitas Widyatama

Jl. Cikutra 204A Bandung 40125

Telp (022) 7278860 Fax (022) 7278860 ext 120

E-mail: fitrah.rumaisa@widyatama.ac.id

ABSTRAK

Diketahui mulai tahun 2001 Universitas Widyatama (UTama) memiliki 11 program studi. Pada Penerimaan Mahasiswa Baru, calon mahasiswa diperbolehkan memilih maksimal 3 (tiga) pilihan program studi. Selama ini, belum pernah diketahui rule pemilihan program studi yang dilakukan oleh calon mahasiswa. Sehingga strategi khusus yang dilakukan dalam kegiatan promosi. Salah satu alternatif solusi dari masalah tersebut menerapkan teknik data mining, sehingga dapat dilakukan penelusuran pada data historis untuk mengidentifikasi pola data yang dimiliki terhadap subyek yang diteliti yang didasarkan pada sifat-sifat data yang teridentifikasi sebelumnya. Dalam penelitian ini akan menggunakan algoritma apriori untuk mengetahui asosiasi rule pemilihan program studi calon mahasiswa.

Kata kunci: association rule, data mining, algoritma apriori

1. PENDAHULUAN

1.1 Latar Belakang

Diketahui mulai tahun 2001 UTama memiliki 13 program studi, yaitu Manajemen S1, Manajemen D3, Manajemen Retail D3, International Program S1, Akuntansi S1, Akuntansi D3, Teknik Informatika S1, Teknik Industri S1, Sistem Informasi S1, Bahasa Inggris S1, Bahasa Jepang D3, Desain Grafis D4 dan Desain D3. Pada Penerimaan Mahasiswa Baru, calon mahasiswa diperbolehkan memilih maksimal 3 (tiga) pilihan program studi.

Selama ini, belum pernah diketahui rule pemilihan program studi yang dilakukan oleh calon mahasiswa. Sehingga belum dapat diketahui strategi khusus yang akan dilakukan pada kegiatan promosi.

Algoritma Apriori sudah digunakan di beberapa penelitian sebelumnya, namun belum ada yang menggunakannya untuk penentuan rule pemilihan program studi calon mahasiswa. Penelitian sebelumnya mengangkat topik pengidentifikasian pola kecocokan siswa dan organisasi dalam penempatan magang [Almahdi,2009], atau hanya perancangan perangkat lunak yang menggunakan algoritma apriori sebagai metodenya [Santoso,2003]

1.2 Rumusan Masalah

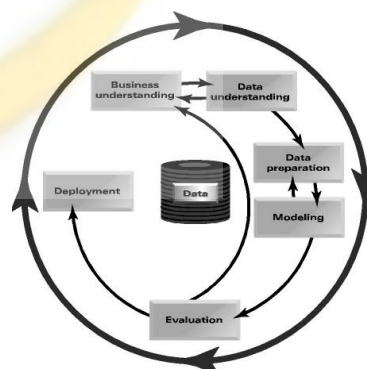
Permasalahan yang akan dibahas pada laporan ini adalah hal-hal yang berkaitan dengan aktifitas analisis rule pemilihan program studi calon mahasiswa. Dengan permasalahan yang telah diketahui maka penulis merumuskan masalah, sebagai berikut:

- Bagaimana rule-rule yang dihasilkan berkaitan dengan data data penerimaan mahasiswa baru?
- Bagaimana implikasi manajerial dari analisis association rule ?.

2. METODOLOGI PENELITIAN

2.1 Metodologi Penelitian

Metodologi penelitian data mining pada prinsipnya merupakan kegiatan pencarian pengetahuan atau lebih dikenal dengan Knowledge Discovery. Dalam penelitian ini menggunakan metodologi Cross-Industry Standard Process for Data Mining (CRISP-DM).



Gambar 1. Tahapan Data Mining Cross-Industry Standard Process for Data Mining

1. Business Understanding

Pada tahap awal ini, perlu dilakukan pemahaman mengenai data dan bisnis pada jasa pendidikan khususnya untuk Penerimaan Mahasiswa Baru meliputi informasi-informasi yang terkait

2. Data Understanding

Setelah melakukan beberapa kali diskusi, pada tahap ini dimulai dengan satu pengumpulan data awal dan dilanjutkan dengan

kegiatan-kegiatan untuk mempelajari data tersebut, untuk mengenal mutu data, untuk menemukan pemahaman terhadap data tersebut, atau mendeteksi subset yang menarik untuk membuat hipotesa dalam mencari informasi tersembunyi.

3. Data Preparation

Pemanfaatan data secara menyeluruh merupakan langkah yang ideal, akan tetapi kurang praktis dalam pengerjaan *data mining*. Pada fase ini persiapan data mencakup semua kegiatan untuk menyusun kumpulan data akhir (data yang akan dimasukkan dalam perangkat pembuatan model) dari data mentah.

3. Modelling

Dalam tahapan ini dilakukan pembuatan model. Hal yang terpenting dalam tahapan ini adalah proses Iterative dalam pemodelan. Dengan melakukan teknik dan algoritma *data mining* yang berbeda untuk mendapatkan penyelesaian permasalahan bisnis yang sesuai.

4. Evaluation

Kemudian setelah model dibuat, tahapan selanjutnya adalah melakukan evaluasi dari hasil model-model yang telah terbentuk berupa pola-pola pemilihan program studi calon mahasiswa. Model-model ini kemudian diterjemahkan sebagai informasi penting atau bernilai tinggi.

5. Deployment

Tahapan ini dapat dilakukan ketika model data mining yang sudah dibuat sudah divalidasi. Dari model-model yang dihasilkan, perlu diuji dengan data yang berbeda dan dilakukan evaluasi kembali mengenai keakuratan dari hasil model tersebut.

Dalam penelitian ini hanya dilakukan deploy secara sederhana, mengingat waktu yang diberikan dalam penelitian ini terbatas.

3. PEMBAHASAN DAN HASIL

3.1 Pembahasan

Sumber data merupakan data master yang tercatat setiap tahunnya mulai dari bulan Januari sampai dengan bulan Juli. Adapun sumber data yang digunakan untuk pembuatan data mining berasal dari data Penerimaan Mahasiswa Baru yang tercatat di Unit PR & Marketing. Untuk penelitian ini, data yang digunakan adalah kompilasi dari data mulai tahun 2007 sampai dengan 2011.

Daftar mengenai informasi mengenai calon mahasiswa baru didapat dengan cara melakukan eksplorasi terhadap database Penerimaan Mahasiswa Baru yang dimiliki oleh Universitas Widayatama. Adapun informasi yang didapat setelah melakukan eksplorasi adalah No USM, Nama Lengkap, Jenis Kelamin, Golongan Darah, Penghasilan Ortu, Asal sekolah, Jurusan Asal sekolah, Kabupaten asal sekolah, Provinsi asal sekolah, Pilihan prodi 1, ke 2 dan ke 3, Sumbangan sukarela

Tidak semua atribut dimasukkan dalam set data yang digunakan dalam proses data mining karena hanya berperan sebagai referensi identifikasi, seperti "No USM", "Nama", "Golongan Darah", "Kabupaten asal sekolah", "Provinsi Asal sekolah", dan "Sumbangan sukarela".

Pemilihan field/kolom data training untuk data mining berdasarkan data yang sudah dibuat. Serta informasi lainnya untuk mengakomodasi analisa selanjutnya agar relevan dengan permasalahan bisnis yang ada. Adapun field/kolom yang dijadikan sebagai data training adalah sebagai berikut:

Tabel 1. Field/kolom Data Training

No	Nama Field/Kolom	Deskripsi
1	JenisKelamin	Jenis Kelamin
2	Penghasilan Ortu	Penghasilan Orang Tua
3	Pil 1	Pilihan ke-1
4	Pil 2	Pilihan ke-2
5	Pil 3	Pilihan ke-3
6	Jurusan	urusan SMA

3.2 Pemodelan

Setelah penyiapan data training, maka selanjutnya adalah membuat model association rules. Model ini dibuat menggunakan tool dari clementine yang menyediakan berbagai macam algoritma. Di dalam penulisan ini, pembuatan model Association rules menggunakan algoritma apriori dan GRI. Seperti yang dijelaskan sebelumnya bahwa pencarian association rules dilakukan terhadap data training.

Dari objek validasi type data dilanjutkan ke proses data mining yaitu Algoritma Apriori

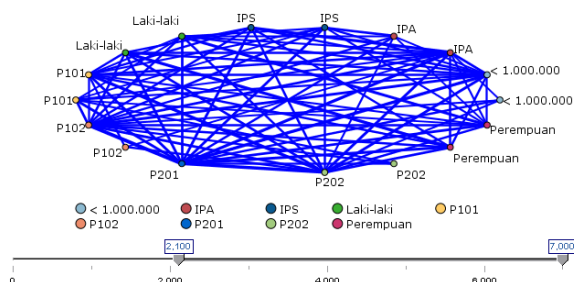
Data training yang disiapkan menggunakan Ms. Excel dikoneksikan ke tools clementine, dalam hal ini didefinisikan dalam objek data training. Kemudian objek ini diteruskan kepada objek type data, untuk memastikan type data yang diproses sudah sesuai dengan kebutuhan model

Setelah model algoritma apriori dijalankan menghasilkan model association rules seperti gambar 2 sebagai berikut:

Consequent	Antecedent	Support %	Confidenc
P101	P202 Perempuan < 1.000.000	10.153	94.255
P101	P202 Perempuan	14.315	93.854
P101	P202 IPS < 1.000.000	10.657	91.373
P101	P202 IPS	15.294	90.498
P101	P202 < 1.000.000	19.348	90.036
P101	P202 IPA	12.071	89.844
P101	P202	28.255	89.748
P102	P201	14.187	87.735
P101	P202 Laki-laki	13.94	85.532

Gambar 2. Association Rule dari algoritma Apriori

Kemudian secara visual hubungan antar field/variable dapat digambarkan dengan grafik web.



Gambar 3 Grafik Web dari Association Rule

Grafik web diatas dibuat dari field-field yang dominan saja, yaitu Jurusan sekolah, Jenis kelamin, pilihan 1 dan pilihan 2.

3.3 Hasil Evaluasi Model

Setelah model algoritma apriori dan GRI dijalankan menghasilkan association rules yang berkaitan dengan pola pemilihan program studi calon mahasiswa baru. Adapun pengetahuan-pengetahuan baru yang muncul dari:

(a). Kedua algoritma tersebut, terdapat beberapa pola yang sama dengan tingkat support dan confidence relative tinggi adalah sebagai berikut:

- 1) Association Rule: $P202 \rightarrow P101$ ($s=28,26\%$, $con=89,75\%$)
Hal ini berarti bahwa calon mahasiswa yang memilih pilihan pertama program studi Akuntansi kemungkinan memilih program studi Bisnis dan Manajemen S1 dengan tingkat confidence sebesar 89,75% dan support sebesar 28,26%.
- 2) Association Rule: $P201 \rightarrow P102$ ($s=14,19\%$, $con=87,74\%$)
Hal ini berarti bahwa calon mahasiswa yang memilih pilihan pertama program studi Bisnis dan Manajemen S1 kemungkinan memilih program studi Akuntansi dengan tingkat confidence sebesar 87,74% dan support sebesar 14,19%.
- 3) Association Rule: $P202, Perempuan, < 1.000.000 \rightarrow P101$ ($s=10,15\%$, $con=94,26\%$)
Hal ini berarti bahwa jika calon mahasiswa memilih pilihan kedua program studi Bisnis Manajemen S1, berjenis kelamin perempuan dan penghasilan orang tuanya dibawah 1 juta, maka kemungkinan pilihan satu pada program studi Akuntansi dengan tingkat confidence 94,26 dan support sebesar 10,15%.
- 4) Association Rule: $P202, IPS \rightarrow P101$ ($s=15,29\%$, $con=90,50\%$)
Hal ini berarti bahwa setiap calon mahasiswa yang memilih pilihan kedua program studi Bisnis dan Manajemen S1,

dan berasal dari jurusan sekolahnya IPS, kemungkinan memilih program studi Akuntansi S1 sebagai pilihan pertama dengan tingkat confidence 90,50% dan support 15,29%

3.4 Pemanfaatan Hasil Evaluasi

Dari hasil pengetahuan yang diperoleh dari association rules maka kemungkinan-kemungkinan pemanfaatan yang dapat diterapkan oleh Universitas Widyatama adalah:

1. Pemanfaatan pada Marketing
 - a. Rule $P202 \rightarrow P101$ ($s=28,26\%$, $con=89,75\%$) dan rule $P201 \rightarrow P102$ ($s=14,19\%$, $con=87,74\%$)
 - Dapat digunakan pada proses Roadshow atau promosi ke sekolah-sekolah, pihak marketing akan membawa serta program studi sesuai dengan rule diatas. Untuk target pasar yang memilih program studi Akuntansi, marketing membawa serta program studi Bisnis dan Manajemen dalam promosinya. Sehingga akan menjadi suatu paket promosi, jika akan mempromosikan program studi Akuntansi makan program studi Bisnis dan Manajemen sekaligus dipromosikan.
 - Visit Widyatama, calon mahasiswa dengan mudah diarahkan ke program studi yang diminatinya, tidak random seperti yang sudah dilakukan selama ini. Untuk calon mahasiswa yang ingin mengunjungi program studi Akuntansi, akan di ajak mengunjungi program studi Bisnis dan Manajemen juga.
 - Dalam pembuatan tools promosi, seperti brosur, leaflet dll, dapat dibuat lebih hemat, tidak perlu satu brosur untuk satu program studi. Untuk brosur program studi Akuntansi, disatukan dengan program studi Bisnis dan Manajemen, sehingga lebih tepat sasaran untuk calon mahasiswa yang mengetahui kedua program studi tersebut.
 - b. Rule $< 1.000.000, P202 \rightarrow P101$ ($s=19,35\%$, $con=90,04\%$) dapat digunakan untuk mempertimbangkan penentuan biaya kuliah setiap tahunnya atau cara pembayaran biaya kuliah. Hal ini dapat menekan angka mahasiswa yang mengundurkan diri dengan alasan biaya.
 - c. Rule $P202, IPS \rightarrow P101$ ($s=15,29\%$, $con=90,50\%$) dapat membantu program studi baik Akuntansi maupun Bisnis dan Manajemen dalam menentukan target pasarnya dalam melakukan promosi. Pihak program studi tidak perlu membuang-buang waktu untuk melakukan promosi pada

- siswa-siswa jurusan lain selain IPS. Sehingga promosi yang dilakukan menjadi lebih tepat pada sasaran.
2. Pemanfaatan untuk Kemahasiswaan
 - a. Rule $< 1.000.000, P202 \rightarrow P101$ (s=19,35%, con=90,04%) dan rule Perempuan, $< 1.000.000, P202 \rightarrow P101$ (s=10,15%, con=94,26%)
Rule tersebut menunjukkan bahwa program studi diminati oleh calon mahasiswa yang berpenghasilan orang tua dibawah satu juta dan sebagian besar berjenis kelamin perempuan. Informasi ini sangat dibutuhkan oleh bagian Kemahasiswaan yang bertugas mencari program beasiswa yang dapat dimanfaatkan oleh mahasiswa-mahasiswa tersebut. Bagian kemahasiswaan dapat menentukan porsi beasiswa berdasarkan program studi tersebut dan juga dapat ditentukan dari jenis kelamin.
 3. Pemanfaatan untuk Program Studi
 - a. Rule $P202 \rightarrow P101$ (s=28,26%, con=89,75%) dan Rule $P201 \rightarrow P102$ (s=14,19%, con=87,74%)
Rule menunjukkan bahwa sebenarnya ada sebagian dari mahasiswa yang diterima di pilihan kedua baik akuntansi maupun bisnis manajemen sebenarnya tidak menginginkan program studi tersebut. Maka rule ini dapat dimanfaatkan oleh program studi Akuntansi untuk mencari cara agar mahasiswa yang diterima di pilihan kedua baik program studi Akuntansi maupun Bisnis dan Manajemen tidak akan berpindah ke program studi Bisnis dan Manajemen pada semester selanjutnya atau kesulitan dalam mengikuti matakuliah, seperti yang selama ini terjadi. Misalkan dengan mengadakan *treatment* khusus terhadap mahasiswa-mahasiswa tersebut

4. KESIMPULAN

Dari hasil analisa Association rule data mining dengan data training dari pengembalian formulir calon mahasiswa baru Universitas Widyatama, maka dapat disimpulkan bahwa rule-rule yang dihasilkan dapat bermanfaat:

- (a). Untuk proses Roadshow atau promosi ke sekolah-sekolah, sehingga marketing akan membawa program studi sesuai dengan rule tersebut.
- (b). Untuk Proses Visit Widyatama, sehingga calon mahasiswa dengan mudah diarahkan ke program studi yang diminatinya
- (c). Untuk penghematan pada pembuatan tools promosi, seperti brosur, leaflet dll.
- (d). Untuk mempertimbangkan penentuan biaya kuliah setiap tahunnya atau cara pembayaran biaya kuliah

- (e). Untuk membantu pihak program studi dalam menentukan target pasarnya dalam melakukan promosi
- (f). Untuk mencari program beasiswa yang dapat dimanfaatkan oleh mahasiswa dan dalam teknis pemberiannya lebih tepat sasaran
- (g). Untuk mencari cara agar mahasiswa yang diterima dipilihan kedua tidak akan berpindah ke program studi lain.

PUSTAKA

- Agrawal, Rakesh, Ramakrishnan Srikant, *Fast algorithms for mining association rules*. Proceeding Of The 1994 International Conference Very Large Data Bases, Santiago, Chile, September 1994.
- Albion, Research Ltd. 2007. *Market Basket Analysis*. <http://www.albionresearch.com>. [23 Maret 2009]
- Almahdi, Mohammed Ahmed, Norita Md Norwawi, and Wan Hussain Wan Ishak; Identifying Student And Organization Matching Pattern Using Apriori Algorithm For Practicum Placement; 2009 International Conference On Electrical Engineering And Informatics 5-7 August 2009, Selangor, Malaysia
- HAN, Feng, ZHANG, Shu-Mao, and DU, Ying-Shuang;; The Analysis And Improvement Of Apriori Algorithm; Journal Of Communication And Computer, ISSN1548-7709, USA
- Han J,Kamber M.; Data Mining: Concepts and Techniques, Simon Fraser University, Morgan Kaufmann Publishers
- Larose D.T.. Discovering Knowledge in Data, An Introduction to Data Mining. A John Wiley & Sons, Inc. 2002
- Megaputer.. *Market Basket Analysis*. <http://www.megaputer.com>. 2007
- Meiwati, Leni, and Metty Mustikasari; Aplikasi Data Mining Menggunakan Aturan Asosiasi Dengan Metode Apriori Untuk Analisis Keranjang Pasar Pada Data Transaksi Penjualan Apotek; Proceeding Ikatan Profesi Komputer Dan Informatika Indonesia 2010 ISBN : 978-602-98154-0-5
- Olson, David and Y. Shi.. Pengantar Ilmu Penggalan Data Bisnis. Salemba Empat, Jakarta. 2008
- Santoso, Leo Willyanto; Pembuatan Perangkat Lunak Data Mining Untuk Penggalan Kaidah Asosiasi Menggunakan Metode Apriori; Jurnal Informatika, Vol 4, No 2, November 2003;