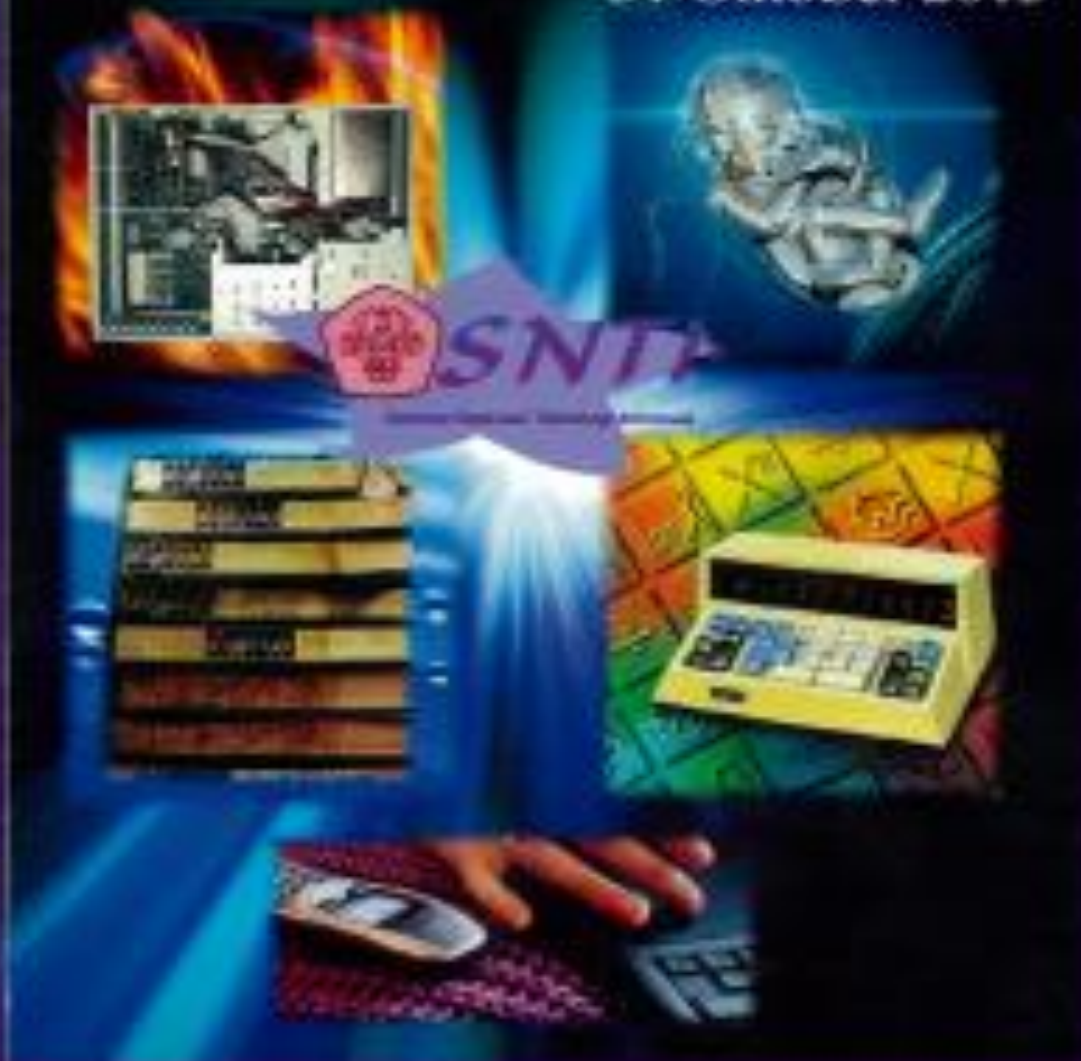


Prosiding

ISSN : 1829-8156
Vol. 12 No. 1 Tahun 2015

SNTI 2015

31 Oktober 2015



**Fakultas Teknologi Informasi
Universitas Tarumanagara
Jakarta**

A11	Implementasi <i>Global Extreme Programming</i> Dalam Analisis Dan Perancangan Sistem Penerimaan Mahasiswa Baru	Novel Hadinata	Universitas Bina Dharma	64
A12	Pengklasifikasian Minat Belajar Mahasiswa Dengan Model Data Mining Menggunakan Metode Clustering	Marliadawati Andri	Universitas Bina Dharma	73
A13	Aplikasi <i>Non-Cooperative game Theory</i> dalam Model Evolusi Virulensi	I Gusti Agung Surya Julawan	Universitas Indonesia	80
A14	Pembacaan Game Virtual Reality "Acha" Menggunakan Output RIL DK2 Dan Leap Motion Controller	Sri Mardiyati Andri Jeanny Pragantha Darius Andana Harris	Universitas Tarumanagara	85
A15	Game 3D No Way Out Dengan Fitur Virtual Reality	Erwin Sanders Jeanny Pragantha Darius Andana Harris	Universitas Tarumanagara	93
A16	Penerapan Algoritma Genetika Dalam Menemukan Rute Terpendek	Diana Fallo Alb. Joko Santoso Djoko Budiyanto	Universitas Atma Jaya Yogyakarta	99
A17	Penyelesaian Permasalahan Penjemputan Dan Pengantaran Travelling Salesman Sesuai Aturan FIFO Dengan Algoritma Iterated Local Search	Ajeng Dwi Andina Sri Mardiyati	Universitas Indonesia	104
A18	Penyelesaian Multiple Travelling Salesman Problem (M-TSP) Dengan Algoritma K-Means Clustering-Genetika	Iihan Sri Mardiyati	Universitas Indonesia	111
A19	Algoritma Genetic Ant Colony System Untuk Menyelesaikan Travelling Salesman Problem	Lutfiani Safitri Sri Mardiyati	Universitas Indonesia	118

PENGLASIFIKASIAN MINAT BELAJAR MAHASISWA DENGAN MODEL DATA MINING MENGGUNAKAN METODE CLUSTERING

Marlindawati¹⁾ Andri²⁾

¹⁾ Manajemen Informatika Universitas Bina Darma
Jl. Ahmad Yani No. 3, Palembang 30264 Indonesia
email : marlinda.wati@gmail.com

²⁾ Sistem Informasi Universitas Bina Darma
Jl. Ahmad Yani No. 3, Palembang 30264 Indonesia
email : andri@mail.binautama.ac.id

ABSTRACT

Dari waktu ke waktu dan dalam sebuah perguruan tinggi tentu data mahasiswa akan terus mengalami perkembangan. Hal ini menyebabkan data yang terdapat mahasiswa bertambah, terutama data akademik mahasiswa. Hal ini yang bertanggung tidak dimanfaatkan dengan maksimal. Sehingga, maka akan terjadi kemunduran terhadap data tersebut. Dengan pemanfaatan teknik data mining maka dari data yang banyak tadi akan didapat informasi yang berguna. Data Mining merupakan proses mencari data yang menggunakan perangkat lunak untuk menentukan pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat memetakan pengetahuan guna mendukung keputusan. Dalam penelitian ini akan dilakukan mining data yang berhubungan dengan skenario mahasiswa untuk dapat mengetahui minat belajar mahasiswa dengan metode clustering. Data penelitian diambil dari mahasiswa Sistem Informasi angkatan 2011-2012 yaitu 1.000 record data dari tb_rnk, 4.1367 record data dari tb_lsh dan 417 record data dari tb_rnk. Proses mining menggunakan tools SSAS (SQL Server 2008 Analysis Service). Setelah melakukan beberapa tahapan dari data mining, hasil analisis menunjukkan terdapat tiga cluster yang terbentuk berdasarkan konsentrasi matakuliahnya. Cluster 1 memiliki minat untuk konsentrasi A sebanyak 40 mahasiswa. Cluster 2 memiliki minat untuk konsentrasi B sebanyak 186 mahasiswa, dan cluster 3 memiliki minat untuk konsentrasi C sebanyak 188 mahasiswa.

Key words

mining informasi, data mining, clustering.

1. PENDAHULUAN

Sumber utama yang dapat dimanfaatkan untuk menghasilkan suatu informasi adalah data. Informasi yang dihasilkan sangatlah penting bagi sebuah perguruan tinggi. Pengolahan data yang maksimal akan menghasilkan informasi yang berkualitas dan bermanfaat bagi perguruan tinggi. Dari waktu ke waktu, data dalam sebuah perguruan tinggi akan terus mengalami peningkatan dan menjadi semakin data yang semakin banyak. Jika data tersebut tidak dikelola dengan baik maka data tidak ada manfaatnya. Dengan pemanfaatan teknik data mining dalam pengolahan data, maka dari data tersebut akan didapatkan informasi yang berguna dan bermanfaat bagi perguruan tinggi.

Data mining merupakan perambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar (Davis, 2004) [2]. Menurut Prasaditono (2007) [3] data mining dapat juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari kumpulan data. Data mining dapat dimanfaatkan untuk menggali informasi dari tampilan data akademik yang berhubungan dengan mahasiswa untuk mendapatkan informasi atau pola dari ketertarikan belajar mahasiswa. Informasi yang didapat dari teknik mining data ini mengkonstruksi sebuah keputusan yang dapat dimanfaatkan oleh pihak yang berwenang dalam sebuah perguruan tinggi dalam meningkatkan konsentrasi mahasiswa tersebut.

Menurut Davies (2004) [1], data mining memiliki beberapa karakteristik, yaitu:

- Data mining yang berhubungan dengan pemrosesan sesuatu yang (seseorang) dan pola data tertentu yang tidak diketahui sebelumnya.
- Data mining biasanya menggunakan data yang sangat besar yang digunakan untuk membuat hasil yang lebih dipercaya.
- Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

Data mining juga disebut dengan Knowledge Discovery in Database (KDD), yaitu kegiatan yang meliputi pengumpulan, pemilihan data untuk memverifikasi keterkaitan, pola dan hubungan dalam set data yang berukuran besar. Karena data mining dapat dipakai untuk pengambilalihan suatu keputusan.

Menurut Haa (2006) [2] terdapat enam tahap dalam data mining:

- Pembersihan data (cleaning)**
Merupakan proses untuk menghilangkan noise serta data yang tidak konsisten atau tidak relevan.
- Integrasi data (integration)**
Bertujuan untuk menggabungkan data dari berbagai sumber basisdata dalam satu himpunan yang baru.
- Seleksi data (selection)**
Bertujuan untuk memilih data-data apa saja yang akan diproses selanjutnya.
- Transformasi data (transformation)**
Bertujuan untuk mengubah data kedalam format tertentu yang sesuai untuk proses selanjutnya.
- Proses mining (data mining)**
Proses untuk menerapkan metode data mining yang bertujuan untuk memunculkan pengetahuan yang baru.
- Evaluasi pola (pattern evaluation)**
Bertujuan untuk mengidentifikasi pola-pola yang menarik kedalam knowledge base yang ditemukan.
- Presentasi pengetahuan (knowledge presentation)**
Merupakan penyajian pengetahuan untuk memperoleh pengetahuan yang diperoleh pengguna.

Tujuan dari penelitian ini membuat model data mining untuk membuat pengelompokan mahasiswa berdasarkan ketertarikan mereka terhadap mata kuliah yang diunggahnya. Dan diharapkan dapat menghasilkan sebuah model data mining yang dapat dimanfaatkan untuk memotivasi konsentrasi minat belajar mahasiswa yang ada pada program studi Sistem Informasi Fakultas Ilmu Komputer Universitas Bina Nusantara Palembang.

Metode penelitian yang digunakan adalah metode pengembangan sistem SDLC (*system development life cycle*). Metode pengembangan sistem ini memiliki beberapa tahap, yaitu:

- Tahap Perencanaan**
Tahap yang digunakan untuk melakukan identifikasi masalah, menentukan tujuan sistem, identifikasi kendala sistem dan membuat analisis kebutuhan.
- Tahap Analisis**
Tahapan untuk melakukan analisis terhadap permasalahan yang akan dihadapi dalam penelitian. Pada tahap ini akan dilakukan identifikasi penyebab masalah dan analisis kebutuhan sistem.
- Tahap Perancangan**
Dalam tahap ini akan dilakukan proses perancangan berdasarkan data yang didapat dari tahap analisis kebutuhan sistem.
- Tahap Penerapan**
Pada tahap ini dilakukan penerapan dari hasil rancangan dengan menggunakan tools yang ada.
- Tahap Pengujian**
Pada tahap ini dilakukan uji coba terhadap hasil rancangan yang telah dibuat.

Agar mendapatkan bahan-bahan yang relevan, akurat dan reliabel dalam penelitian ini, maka penulis melakukan pengumpulan data dengan cara sebagai berikut:

- Observasi**
Penulis melakukan pengamatan dari pertemuan data-data mahasiswa serta data-data yang berhubungan dengan nilai akademik fakultas ilmu komputer program studi Sistem Informasi angkatan 2011 s.d. 2013, Universitas Bina Nusantara Palembang.
- Studi pustaka**
Teknik pengumpulan data dilakukan dengan menggunakan buku-buku, referensi, dan dokumen yang berhubungan dengan penelitian ini.

2. PEMBAHASAN

Dalam pengelompokan minat mahasiswa dengan model data mining menggunakan metode Clustering, dibuatlah rancangan dari databasenya sebagai berikut:

2.1 Rancangan Database

- Ta_min

Tabel 1. Tabel Mahasiswa

Nama	Jenis	IPK
ibrahim	varchar	8
ibrahim	varchar	90
ibrahim_khalid	varchar	2
ibrahim	varchar	60
ibrahim_ahmad	date	60
ibrahim	varchar	3
ibrahim_panglima	varchar	270
ibrahim	varchar	20
ibrahim	varchar	20
ibrahim	varchar	20

k. **tb_khs**

Tabel 2. Tabel Kartu Hasil Studi (KHS)

Nama	Jenis	IPK
ibrahim	varchar	0
ibrahim	varchar	10
ibrahim	varchar	1
ibrahim_panglima	varchar	1
ibrahim_khalid	varchar	8
ibrahim	varchar	0
ibrahim	varchar	0
ibrahim	int	0
ibrahim	int	0
ibrahim	int	4
ibrahim	int	4
ibrahim_panglima	int	8
ibrahim_khalid	varchar	1

l. **tb_mk**

Tabel 3. Tabel Mata Kuliah

Nama	Jenis	IPK
ibrahim	varchar	10
ibrahim	varchar	10
ibrahim	varchar	1
ibrahim	int	1
ibrahim_panglima	varchar	0

2.2 ETL (Extract, Transform and Load)

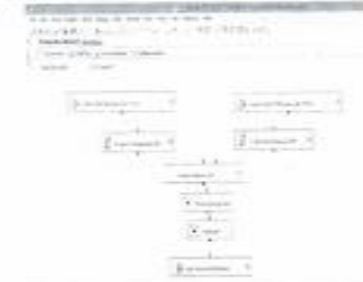
Proses ETL bertujuan untuk melakukan pembaruan data warehouse yang berasal dari data mentah dalam format database MySQL. Data yang di ekstrak akan dilakukan proses transformasi dan load ke dalam format baru yaitu format OLAP. Proses ETL dilakukan menggunakan tools SQL Server Integration Service (SSIS) yang ada dalam database *Microsoft Development Studio (BIDS)*. ETL yang dilakukan dalam penelitian dapat dilihat pada gambar berikut ini.

1. Pembuatan Control Object



Gambar 1. Pembuatan Control Object

2. Proses ETL



Gambar 2. Proses ETL

2.3 Seleksi Data (Data Selection)

Data yang digunakan dalam penelitian berasal dari data mahasiswa Sistem Informasi Universitas Bina Darma tahun angkatan 2011 sampai dengan 2013. Selain data mahasiswa, data IPS juga digunakan untuk penelitian ini, dikarenakan informasi yang terdapat di dalamnya sudah mewakili informasi yang dibutuhkan untuk dijadikan indikator prestasi dalam pengelolaan data kejuruan yang diinginkan. Data yang diperoleh adalah sebanyak 1.000 record data dari **tb_reks**, 4.1267 record data dari **tb_khs** dan 417 record dari **tb_mk**. Database mahasiswa terdiri dari 23 atribut yang menjelaskan identitas diri mahasiswa dan

informasi tentang keadaan mahasiswa yang bersangkutan. Atribut dan mahasiswa diambil dari penggabungan beberapa tabel yang terdiri dari tb_mhs, tb_khs, dan tb_rsk. Atribut tersebut diantaranya adalah rnk, nama, jenis_kelamin, temp_lahir, ugi_lahir, kd_progd, alamat, asal_sck, kota, gsk_smt, kd_rsk, srs, sma, pendak, tahun_akademik, kelas, kd_dowr, kd_progd, tagz, kwh, smd, ameskor, nilai_sjka, nilai_harf, dan ds. Sedangkan database IPK hanya terdiri dari 3 atribut yang memberikan informasi mengenai prestasi akademik dan beban studi yang diambil mahasiswa yang bersangkutan. Atribut tersebut diantaranya adalah srs, nama, ameskor, temp_sjka, dan IPK.

Tahap-tahap yang akan digunakan dalam penelitian ini:

1. Proses Data

Data yang digunakan dalam penelitian ini terdiri dari tabel mahasiswa, tabel ream kuliah, dan tabel khs.

a. Tb_mhs

Gambar 3. Record Tabel Mahasiswa

b. Tb_khs

Gambar 4. Record Tabel Kota Hasil Search

c. Tb_rsk

Gambar 5. Record Tabel Mata Kuliah

2. Pembersihan Data (Data Cleaning)

Tahap berikutnya pada proses data mining adalah cleaning data yaitu melakukan pembersihan data terhadap noise yang ditemukan berupa missing value, Ambiguous data, dan redundant data.

Gambar 6. Record Tabel Proses Data

Gambar data merupakan hasil proses data yang sudah di cleaning, sehingga sudah tidak ada lagi data yang missing value. Hal ini dilakukan karena data yang missing value tidak memberikan informasi apapun jika dipertanyakan keberadaannya. Dari tahap cleaning data yang telah dilakukan data diperoleh jumlah data dari tb_mhs sebanyak 995 record, 4.1367 record data dari tb_khs dan 417 record data dari tb_rsk.

3. Integrasi Data

Tahap ketiga adalah tahap integrasi data yang merupakan penggabungan data bertujuan meniadakan adanya data yang telah di-copying ke dalam satu tabel. Pada tahap ini, berbagai tabel *tbl_khs*, *tbl_rsk* dan *tbl_kon* diintegrasikan untuk mendapatkan data source mahasiswa dan data source IPK. Untuk mendapatkan data source mahasiswa proses penggabungan data dilakukan dengan menggunakan *tbl_khs* dan *tbl_rsk* dengan *join* antar table.

Proses selanjutnya menghasilkan *tbl_mhs* dan tabel hasil integrasi data join dari *tbl_khs* dan *tbl_rsk*. Dengan *join* antar tabel dengan *day* nya. Hasil dari integrasi data *join* antar tabel adalah data yang didapat dari source mahasiswa, seperti pada gambar berikut 7 berikut.

Gambar 7. Tabel Data Source Mahasiswa

4. Transformasi Data

Proses selanjutnya adalah tahap transformasi data yaitu data diubah menjadi bentuk yang sesuai untuk diproses dalam data *warehouse*. Dalam penelitian ini data yang akan diproses dari database mysql untuk digunakan di tool SSIS maka format tersebut diubah menjadi *sql*.

Hal ini melibatkan proses transformasi data, dilakukan dengan cara ke proses selanjutnya yaitu proses *mining* data.

Data mining merupakan proses mencari pola atau informasi tersembunyi dalam data yang terdapat dengan menggunakan teknik atau metode tertentu. Penelitian ini menggunakan algoritma yang tepat sangat bergantung pada proses *ETL* secara keseluruhan. Pada penelitian ini penelitian data mining menggunakan teknik *clustering* dan algoritma K-Means.

2.4 Analisis Data Mining

Analisa data mining yang digunakan dalam penelitian ini adalah teknik *clustering* atau pengelompokan yang bertujuan untuk mengelompokkan mahasiswa berdasarkan

kontribusi matematika yang telah dipilih dengan suatu nilai *ipk* yang didapat oleh mahasiswa tersebut.

Data yang digunakan untuk proses analisa data mining adalah data *warehouse* yang didapat dari proses awal dalam tahap data mining. Data *warehouse* berisikan data tentang nilai *ipk* dan kontribusi grafik matematika yang diambil oleh mahasiswa program studi sains inferensi.

Terdapat 700 data mahasiswa yang tersedia dalam data *warehouse* yang selanjutnya digunakan sebagai pemrosesan model data mining dalam penelitian ini. Informasi yang terdapat dalam data *warehouse* yaitu informasi tentang *nama*, *ipk* dan kontribusi yang diidentifikasi mahasiswa tersebut.

2.5 Implementasi Tools SSAS

Implementasi merupakan tahapan perwujudan hasil perancangan data *warehouse* yang telah dilakukan menggunakan SSAS (*SQL Server 2008 Analysis Services*). Hasil dari perancangan data *warehouse* menggunakan tool SSIS dapat dilihat pada gambar berikut. *ini*

Gambar 8. Data Warehouse Dll_Mhs

1. Koneksi Sumber Data

Langkah awal dalam analisa data mining menggunakan tools SSAS adalah dengan mendefinisikan sumber data yang digunakan sebagai data source untuk proses data mining, yang dirampas dalam database *SQL Server 2008* dengan nama *dbMahasiswa*. Dalam *db_Mahasiswa* terdapat sebuah tabel yang diberi nama *tbl_mhs*, *mhs* yang berisi field-field seperti: *Nim*, *Nama*, *Jumlah SKS*, *Jumlah Nilai*, *Jenis Kelamin*, *Angkatan*, *IPK* dan *Kontribusi Matematika* yang dipilih. Proses koneksi ke sumber data dalam penelitian ini dapat dilihat pada gambar 9 berikut.



Gambar 9. Koneksi Sumber Data

2. Pembentukan Data Source View

Langkah berikutnya dalam proses SSAS adalah menentukan *Data Source View* dari sumber data yang telah dikoneksikan. *Data source view* database DB_Mahasiswa dapat dilihat pada gambar 10.



Gambar 10. Data Source View Database DB_Mahasiswa

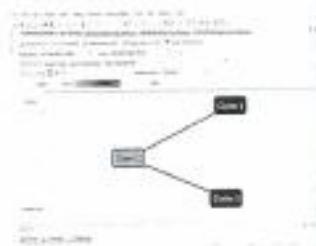
3. Data Mining Structure

Langkah terakhir dari proses SSAS dalam adalah proses analisis data mining, yang merupakan tahapan untuk melakukan analisis terhadap data akademik mahasiswa untuk menghasilkan sebuah model yang dapat digunakan untuk melakukan prediksi minat atau ketertarikan matakuliah pilihan terhadap mahasiswa Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Bina Darma.

Teknik data mining yang digunakan dalam penelitian menggunakan teknik clustering (pengelompokan). Dalam teknik clustering dilakukan

proses pengelompokan sejumlah objek yang memiliki kemiripan dalam suatu kelompok yang sama.

Dari hasil proses mining menggunakan tool SSAS didapatkan hasil pengelompokan data berdasarkan minat konsentrasi matakuliah pilihan yang ada pada Program Studi Sistem Informasi menjadi tiga kelompok (cluster). Hasil pengelompokan (cluster) dapat dilihat pada gambar 11 berikut.



Gambar 11. Cluster Hasil Analisis

Berdasarkan hasil analisis terdapat tiga cluster yang terbentuk berdasarkan konsentrasi matakuliah pilihan. Cluster 1 merupakan pengelompokan minat untuk konsentrasi A, Cluster 2 merupakan kelompok minat untuk matakuliah konsentrasi B, dan cluster 3 merupakan pengelompokan untuk minat matakuliah konsentrasi C.

	1	2	3
1	1	1	1
1	2	2	2
1	3	3	3

Gambar 12. Pengelompokan Data Mahasiswa Berdasarkan Konsentrasi

Dari data hasil analisis dapat dihitung jumlah mahasiswa yang mengambil konsentrasi A sebanyak 243 mahasiswa, konsentrasi B sebanyak 186 mahasiswa dan konsentrasi C sebanyak 188 mahasiswa.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan maka dapat disimpulkan:

1. Penelitian yang dilakukan telah menghasilkan sebuah model data mining pengalokasian minat mahasiswa konsentrasi rekrutasi.
2. Model data mining yang dihasilkan dapat digunakan untuk memprediksi minat mahasiswa konsentrasi rekrutasi program.

Dari selanjutnya pada penelitian sel diharapkan dapat meningkatkan proses analisis data mining menggunakan jumlah data mahasiswa yang cukup besar untuk menghasilkan kemampuan dalam prediksi, dan tidak ada dari program studi Sistem Informasi tapi juga dari semua program studi yang terdapat pada fakultas Ilmu Komputer.

REFERENSI

- [1] D. Inan, and Paul Beynon, 2004, "Database System Third Edition", Palgrave macmillan, New York.
- [2] Hsu, J., and Kamber, M., 2006, "Data Mining: Concepts and Techniques, Second Edition", Morgan Kaufman, San Francisco.
- [3] Rinaldiandri, I., 2007, Pengantar Data Mining : Meningkatkan Efisiensi Pengolahan di Gunung Data. <http://www.jurnal.stikom.ac.id> <http://www.stikom.ac.id> diakses pada tanggal 5 September 2014.

Biodata Penulis

Mahasiswa, sarjana di gelar Sarjana Komputer (S.Kom), Jurusan Manajemen Informatika Universitas Gadjadara, Jogyakarta tahun 1997. Menempuh gelar Magister Komputer (S.Kom) Program Pasca Sarjana Magister Teknik Informatika Universitas Bina Darma Palembang, tahun tahun 2011 Saat ini menjadi Dosen di Universitas Bina Darma Palembang.

Dokter, sarjana di gelar Sarjana Komputer (S.Kom), Fakultas Ilmu Informatika di Universitas Bina Darma Palembang, tahun tahun 2003. Menempuh gelar M. Cs, Program Pasca Sarjana, Ilmu Komputer pada Universitas Gajah Mada Yogyakarta, tahun tahun 2011. Saat ini menjadi Dosen di Universitas Bina Darma Palembang.