

ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI PENGUNDURAN DIRI MAHASISWA DENGAN APLIKASI DATA MINING ADD-INS – STUDI KASUS PADA STMIK MIKROSKIL

Afen Prana Utama Sembiring¹, Mbayak Ginting²

Program Studi Manajemen Informatika, STMIK Mikroskil

Jl. Thamrin No. 122, 124, 140 Medan 20212

afen@mikroskil.ac.id¹, mbayak@mikroskil.ac.id²

Abstrak

Pengambilan sebuah keputusan atau membuat sebuah kebijakan sangat membutuhkan informasi atau pengetahuan. Pengetahuan dapat diperoleh melalui eksplorasi data. Data itu sendiri sering disimpan dalam database atau data warehouse dalam jumlah banyak. Dalam kasus seperti ini umumnya menggunakan data mining. Penelitian ini berusaha mengeksplorasi pengetahuan dari basis data SIPT khususnya tentang faktor-faktor yang mempengaruhi pengunduran diri mahasiswa. Hasil penelitian ini menunjukkan bahwa nilai NEM yang sangat rendah dan nilai STTB yang sedang mempengaruhi pengunduran diri mahasiswa, disamping faktor lainnya.

Kata kunci: *data mining, add-ins, pengunduran diri mahasiswa, kebijakan.*

1. Pendahuluan

Pada umumnya perguruan tinggi baik negeri maupun swasta (PTN/PTS) telah memiliki sistem informasi perguruan tinggi (SIPT) yang mendukung sebagian besar dalam pengelolaan data perguruan tinggi. Jenis data yang terdapat pada SIPT dapat berupa data akademik, sumber daya manusia (dosen, laboran, teknisi, staf administrasi dan lain-lain), inventaris, keuangan dan sebagainya. Dalam fitur akademiknya biasanya berisi data mahasiswa, kurikulum, dosen, kartu rencana studi, jadwal kuliah, nilai, kartu hasil studi dan alumni. Apabila SIPT telah digunakan beberapa tahun tentunya pada basis data telah tersimpan data yang cukup banyak dan bersifat historis. Namun data yang melimpah ini belum banyak dimanfaatkan dalam menggali informasi atau pengetahuan yang dapat membantu pimpinan perguruan tinggi dalam membuat rencana yang bersifat strategis. Salah satu informasi yang *urgent* adalah informasi yang berhubungan dengan pengunduran diri mahasiswa dari sebuah perguruan tinggi (PT). Faktor penyebab pengunduran diri mahasiswa dapat dilihat dari dua sisi yaitu faktor internal PT atau faktor eksternal. Sebagian dari faktor eksternal akan sulit atau bahkan tidak mungkin dikendalikan oleh pimpinan PT, namun beberapa diantaranya masih dapat dihindari dengan mempertimbangkan faktor tersebut dalam seleksi calon mahasiswa, sedangkan faktor internal sangat memungkinkan dikendalikan dengan kebijakan-kebijakan yang tepat.

Dalam tulisan ini penulis melakukan studi kasus pada STMIK Mikroskil Medan yang tingkat pengunduran diri mahasiswanya rata-rata 88 orang per tahun berdasarkan data tahun 2007 – 2010 (sumber : SIPT Mikroskil). Perguruan tinggi swasta (PTS) ini telah menerapkan SIPT dan sistem informasi lainnya. STMIK Mikroskil telah mengembangkan SIPT sejak delapan tahun lalu, sehingga data sudah cukup banyak tersimpan di dalam basis data SIPT. Data yang tersimpan dalam SIPT ini dapat dijadikan sebagai sumber pengetahuan baru

melalui aplikasi data mining atau penambangan data, khususnya informasi tentang faktor-faktor yang dominan berpengaruh pada pengunduran diri mahasiswa. Faktor-faktor yang dimaksud berupa data pribadi mahasiswa sebagai faktor eksternal sedangkan faktor internal terdiri dari indeks prestasi mahasiswa, program studi, shift kuliah, dan grade ujian saringan masuk (USM), dan lain sebagainya.

Dari uraian diatas tergambar bahwa informasi atau pengetahuan yang dapat dieksplorasi dari data base SIPT dan hasilnya dapat dijadikan acuan dalam mendukung pembuatan kebijakan dalam merespon masalah yang muncul ke permukaan dan mengantisipasi munculnya masalah baru. Sehingga pertanyaan yang harus dijawab dalam penelitian ini adalah “Faktor apa saja yang mempengaruhi pengunduran diri mahasiswa dan faktor mana yang paling dominan?”

Penelitian ini dibatasi hanya pada analisis data mahasiswa, baik data akademik maupun biodatanya, dan aplikasi yang digunakan adalah *Data Mining Addins for MS Office 2007 (DMAddins)*. Adapun tujuan dalam penelitian ini adalah menambang pengetahuan dengan menggali (menganalisis) tumpukan data yang terdapat pada data base SIPT untuk mengetahui faktor-faktor yang dominan mempengaruhi pengunduran diri mahasiswa dengan menggunakan aplikasi *Data Mining Add-ins*.

Pengetahuan atau informasi yang diperoleh dari penelitian ini diharapkan dapat memberikan manfaat :

1. Dapat memperkirakan tingkat pengunduran diri mahasiswa pada tahun-tahun yang akan datang.
2. Dapat mengurangi tingkat pengunduran diri mahasiswa melalui pembuatan kebijakan yang tepat yang didasarkan pada hasil penelitian.
3. Dapat dijadikan sebagai dasar pembuatan kebijakan oleh pimpinan dalam hal :
 - Sistem seleksi mahasiswa baru.
 - Sistem pembelajaran dan proses belajar mengajar.

2. Kajian Pustaka

2.1. Pengunduran Diri Mahasiswa

Hampir tidak ada perguruan tinggi baik negeri maupun swasta (PTN/PTS) yang tidak pernah mahasiswanya mengundurkan diri sebagai mahasiswa di PTN/PTS tersebut. Hal ini dibuktikan dengan adanya peraturan pengunduran diri mahasiswa. Sebagai contoh, Universitas Indonesia (UI) sendiri memiliki peraturan tersebut dan khusus program studi *computer science* dapat diakses secara online[1], dan pada ketentuan pengunduran mahasiswa Institut Pertanian Bogor (IPB) dicantumkan ketentuan pengunduran diri mahasiswa baru dan reguler[2]. Pengunduran diri mahasiswa termasuk *dropout* di PTS bahkan cukup banyak, seperti pada STMIK Mikroskil mencapai 88 orang rata-rata per tahun dalam periode tahun 2007 hingga 2010 (sumber : SIPT Mikroskil).

Alasan pengunduran diri sangat beragam, seperti orangtua pindah tugas, mahasiswa pindah kerja, berhenti kuliah, tidak sanggup mengikuti pelajaran, dan lain sebagainya. Banyak diantara alasan tersebut ternyata hanya karangan belaka setelah dibuktikan melalui beberapa penelusuran. Informasi mengenai pengunduran diri mahasiswa pun masih sulit ditemui di jejaring sosial, sehingga dibutuhkan lebih banyak penelitian lagi tentang kasus ini.

Berbeda dengan pengunduran diri, *dropout* diakibatkan oleh ketidakmampuan mahasiswa dalam mencapai sejumlah SKS hingga semester tertentu. Faktor yang mempengaruhi *dropout* hampir sama dengan faktor pengunduran diri. Hasil penelitian Mega Khoirunnisak dan Nur Iriawan membuktikan bahwa mahasiswa yang *dropout* dari ITS Surabaya dipengaruhi oleh faktor perbedaan usia, perbedaan asal daerah mahasiswa, perbedaan penghasilan orang tua,

perbedaan fakultas mahasiswa, perbedaan jalur masuk, serta perbedaan nilai IPK dan nilai TPB[3].

2.2. Pengertian Data Mining

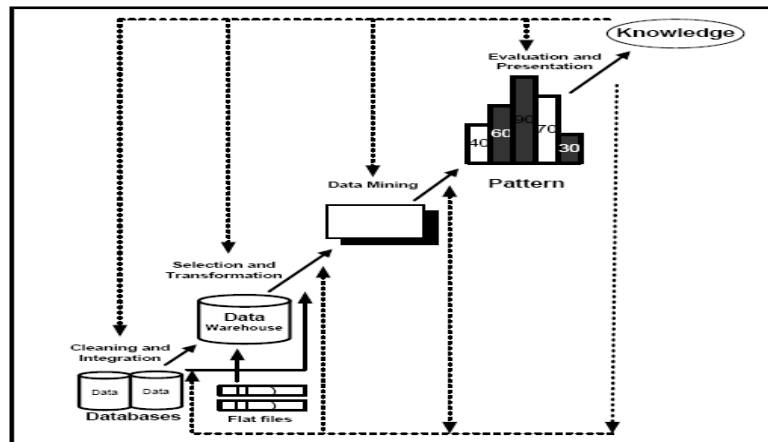
Data mining dapat didefinisikan sebagai penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar[4]. *Data mining*, sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD merupakan kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar[5]. Jadi, *data mining* dapat diartikan sebagai kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam *database*, *data warehouse*, atau penyimpanan informasi lainnya. Untuk menemukan pola yang menarik dalam jumlah data yang banyak dibutuhkan metode analisis sehingga *data mining* berkaitan erat dengan bidang ilmu – ilmu lain, seperti *database system*, *data warehousing*, statistik, *machine learning*, *information retrieval*, dan komputasi tingkat tinggi. Beberapa metode yang sering disebut-sebut dalam literatur *data mining* antara lain *clustering*, *classification*, *association rules mining*, *neural network*, *genetic algorithm* dan lain-lain[6].

Menurut Davies (2004)[4] ada tiga ciri atau karakteristik dari data mining, yaitu :

1. *Data mining* berkaitan dengan upaya menemukan sesuatu yang tersembunyi dan bentuk pola data tidak diketahui sebelumnya.
2. *Data mining* biasa menggunakan jumlah data yang sangat banyak untuk membuat hasil yang lebih dipercaya.
3. *Data mining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

2.3. Tahap-Tahap Data mining

Adapun tahapan atau rangkaian proses *data mining* dapat dibagi menjadi beberapa tahap (lihat Gambar 1). Tahap-tahap tersebut bersifat interaktif. Terdapat 6 tahapan pada *data mining* yaitu:[7]



Gambar 1. Tahapan Data Mining

1. Pembersihan data (*data cleaning*)

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari *database* suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data

yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa *data mining* yang dimiliki. Data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik *data mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data (*data integration*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk *data mining* tidak hanya berasal dari satu *database* tetapi juga berasal dari beberapa *database* atau file teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi Data (*Data Selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus *market basket analysis*, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data (*Data Transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*. Beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan *clustering* hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses *mining*,

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi pola (*pattern evaluation*),

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses *data mining*, mencoba metode *data mining* lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

7. Presentasi pengetahuan (*knowledge presentation*),

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses *data mining* adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami *data*

mining. Karenanya presentasi hasil *data mining* dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses *data mining*. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil *data mining*[7].

2.4. Data Mining dengan MS Excel (*Data Mining Add-ins*)

SQL Server 2005 Data Mining Add-ins for Office 2007 dapat digunakan untuk menemukan pola-pola tersembunyi dan hubungan dalam data dan dapat meningkatkan kualitas analisis data. Paket *Data Mining add-ins* ini terdiri dari tiga *add-ins* :

1. *Table analysis Tools for Excel*

Memungkinkan peneliti untuk menganalisis data *spreadsheet* dengan cara yang *powerful* dengan hanya beberapa klik *mouse*.

Tampilan *ribbon* menu *table tools* adalah seperti gambar-3 berikut ini.



Gambar 2. Toolbar dari Tabel Tools

2. *Data Mining Client for Excel*

Memungkinkan peneliti untuk menggunakan seluruh siklus hidup dari suatu pertambangan data, termasuk menyiapkan data, membangun, mengevaluasi dan mengelola *mining model*, dan memprediksi hasil menggunakan data *spreadsheet* atau data eksternal dapat diakses melalui *Analysis Services database* peneliti.

3. *Data mining template for Visio*

Memungkinkan peneliti untuk membuat, membubuhi keterangan dan berbagi pola *mining model* sebagai diagram visio antara peneliti dengan kolega dan para pengambil keputusan di suatu lembaga/institusi.

Berbagai analisis atau tugas yang sudah dapat dilakukan dalam *data mining add-ins* adalah seperti berikut :[8]

1. Menganalisa hubungan antara data pada satu kolom dengan data pada beberapa kolom yang lainnya
2. Mendeteksi kelompok dari baris-baris data yang ada dengan karakteristik yang sama
3. Mengumpulkan kolom dengan nilai yang sesuai dengan yang ditentukan
4. Melakukan peramalan untuk data dengan fungsi waktu (*time series*)
5. Mencari baris yang tampaknya tidak sejenis dengan baris yang lain (bisa saja baris ini adalah data yg menarik atau data dengan terjadinya anomali)
6. Melakukan sebuah skenario (*goal-seeking* atau *what-if*) *analysis*
7. Membuat sebuah kalkulator yang dapat memprediksi nilai yang diinginkan (dan tetap mempertahankan kemudahan untuk digunakan).
8. Melakukan analisis *shopping basket* dan mengidentifikasi peluang *cross-sales*.

3. Metode Penelitian

Langkah-langkah yang penulis lakukan dalam penelitian ini dapat diuraikan sebagai berikut :

1. Persiapan, termasuk di dalamnya adalah pembagian tugas, waktu pelaksanaan, dan diskusi untuk menyamakan pandangan.
2. Mengidentifikasi faktor penentu (determinan).
Diidentifikasi dari komponen data mahasiswa yang tersedia dalam basis data SIPT. Tahap ini dilakukan dengan observasi pada SIPT.
3. Menyeleksi faktor penentu yang dapat dijadikan variabel penelitian atau faktor penentu yang diduga dominan berpengaruh. Berdasarkan studi awal telah dipilih faktor-faktor yang diduga mempengaruhi pengunduran diri mahasiswa yakni :
 - a. asal sekolah,
 - b. tempat tinggal orang tua (Medan atau di luar Medan),
 - c. pekerjaan orangtua,
 - d. pendidikan orangtua,
 - e. usia mahasiswa,
 - f. jumlah bersaudara,
 - g. nilai ujian lokal dan ujian nasional SLTA,
 - h. jurusan SLTA,
 - i. nilai ujian saringan masuk,
 - j. indeks prestasi di bangku kuliah,
 - k. program studi,
 - l. shift kuliah dan
 - m. tahun masuk kuliah
 - n. hobbi mahasiswa.
4. Pengumpulan data dengan mengkopi database SIPT.
5. Seleksi data. Dilakukan untuk menyisihkan *record* data yang tidak lengkap.
6. Analisis data.
Analisis dilakukan dengan *Data Mining Add-ins (DMAdd-ins)*. Metode analisis yang digunakan sesuai dengan metode analisis yang tersedia pada *Data Mining Add-ins*, yaitu analisis hubungan antara data pada satu kolom dengan data pada beberapa kolom yang lainnya khususnya kolom status mengundurkan diri.
7. Melakukan interpretasi.
Interpretasi dari hasil analisis data merupakan tahap yang sangat penting. Pada tahap ini akan dijelaskan makna yang diperoleh dari penelitian dengan pemahaman umum.

4. Hasil dan Pembahasan

4.1. Deskripsi Data

Data dalam penelitian ini adalah data mahasiswa mulai tahun 1997 sampai dengan tahun 2012. Data tersebut diambil dari database SIPT Mikroskil dengan total records (baris) sebanyak 11.598, dan jumlah fields (kolom) sebanyak 74 atau sebanyak 858.252 sel atau komponen data. Oleh karena tidak semua field dianggap relevan serta adanya field yang tidak lengkap sehingga dilakukan penyortiran sehingga jumlah field yang dipergunakan sebanyak 14 field seperti yang tertulis pada metode penelitian, hal ini sesuai dengan tahapan ketiga pada tahapan data mining, yaitu seleksi data.

4.2. Analisis Data

Dengan menggunakan *Analyze Key Influencers* dan memilih faktor kunci mahasiswa mengundurkan diri (kolom *m_undur_diri*) serta menyenteng kolom-kolom yang diduga mempengaruhi, diperoleh hasil analisis sebagai berikut :

Column	Value	Favors	Relative Impact
11	THN_MASUK 2006	keluar	High
12	JNS_KLMN LAKI-LAKI	keluar	High
13	NIL_NEM < 23.8108165824	keluar	High
14	NIL_STTB 42.4374998208 - 48.266329696	keluar	High
15	NIL_NEM 23.8108165824 - 34.5018535296	keluar	High
16	THN_MASUK 2007	keluar	Low
17	THN_MASUK 2005	keluar	Low
18	THN_MASUK 2010	tidak	Low
19	JNS_KLMN PEREMPUAN	tidak	Low
20	NIL_NEM 47.3530101056 - 52.0941141568	tidak	Low
21	NIL_NEM >= 52.0941141568	tidak	Low
24	NIL_STTB >= 58.2657567744	tidak	Low

Gambar 3. *Key influencers report* untuk mahasiswa mengundurkan diri

Analyze influencers tool digunakan untuk menganalisis hubungan antar semua kolom pada tabel dengan kolom target yaitu kolom *M_undur_diri* (mahasiswa yang mengundurkan diri). Dari Gambar 3 dapat dilihat influencers atau yang mempengaruhi kolom target, yakni tahun masuk, jenis kelamin, nilai NEM, dan nilai STTB. Nilai NEM dalam istilah lain adalah nilai UN dan nilai STTB sama dengan nilai US.

4.3. Interpretasi

Faktor yang mempengaruhi mahasiswa mengundurkan diri lebih didominasi oleh angkatan 2006, atau bila ditinjau dari jenis kelamin maka laki-laki yang lebih mendominasi dari pada perempuan. Faktor berikutnya adalah nilai NEM yang sangat rendah (< 23,81). Kriteria sangat rendah ini dapat dilihat pada karakteristik kategori pada baris 97. Selanjutnya faktor nilai STTB juga mempengaruhi mundur mahasiswa namun nilai STTB-nya bukan pada level yang rendah tapi pada level medium.

Bila ditinjau kebalikannya, maka mahasiswa angkatan 2010 tergolong yang jarang mengundurkan diri, namun hal ini belum dapat dijadikan sebagai sebuah kepastian. Hal ini mengingat bahwa pada setiap semester kemungkinan mahasiswa mengundurkan diri tetap terbuka, sedangkan angkatan ini masih pada semester 4 (pada saat pengumpulan data penelitian ini). Selanjutnya menurut hasil analisis, mahasiswa perempuan lebih dapat bertahan dibanding mahasiswa laki-laki. Dan bila ditinjau dari nilai NEM maka mahasiswa dengan nilai NEM tinggi dan sangat tinggi tetap aktif kuliah. Sama halnya dengan mahasiswa yang mempunyai nilai STTB tinggi juga umumnya tidak mengundurkan diri.

Secara umum dapat disimpulkan bahwa nilai NEM dan nilai STTB cukup berpengaruh terhadap mengundurkan diri atau tidaknya mahasiswa. Mengacu pada grafik jumlah mahasiswa berdasarkan jurusan sekolah, jurusan IPA masih lebih banyak dibanding jurusan IPS dan SMK. Tetapi bila dibandingkan jumlah jurusan IPA dengan jurusan non-IPA maka menjadi seimbang. Karena program studi Teknik Informatika hanya menerima calon mahasiswa dari jurusan IPA, maka dapat dipastikan program studi Sistem Informasi, Manajemen Informatika dan Komputerisasi Akuntansi didominasi dari jurusan IPS dan SMK.

Jurusan sekolah tidak muncul sebagai *influencer* pada penelitian ini. Walaupun demikian kemungkinan faktor tersebut berpengaruh terhadap mahasiswa mengundurkan diri masih harus diperhitungkan. Kemungkinan tidak munculnya adalah ketidakseragaman data atau pada peng-inputan jurusan sekolah bervariasi.

Berdasarkan beberapa sampel yang penulis peroleh bahwa ada indikasi nilai IPK merupakan salah satu *influencer* bagi pengunduran diri mahasiswa, tetapi faktor ini juga tidak muncul dalam penelitian ini. Hal ini disebabkan tidak tercatatnya IPK mahasiswa saat mereka mengundurkan diri, bahkan status mengundurkan diri pada basis data belum terekam secara keseluruhan.

5. Penutup

Setelah dilakukan interpretasi terhadap hasil analisis data maka penulis dapat membuat beberapa kesimpulan serta saran-saran bagi perguruan tinggi khususnya pimpinan serta bagi para peneliti :

1. Pengunduran diri mahasiswa dari sebuah program studi lebih dipengaruhi oleh nilai yang mereka peroleh di bangku sekolah seperti nilai NEM dan nilai STTB. Nilai ini merupakan gambaran dari kemampuan mereka mengikuti pelajaran baik di bangku sekolah maupun di bangku kuliah. Penelitian ini membuktikan bahwa mahasiswa yang mengundurkan diri kebanyakan mendapat nilai rendah pada ujian nasional dan nilai ujian sekolah sedikit lebih baik dari nilai UN.
2. Belum seluruh faktor dapat jadi sebagai bagian dari penelitian sehubungan dengan kurangnya sempurnaan data base, baik dari kelengkapan field maupun nilai yang terdapat pada record.
3. Untuk dapat menetapkan suatu kebijakan yang lebih akurat masih diperlukan penelitian lebih lanjut dengan melengkapi data yang tidak diperoleh pada penelitian ini.
4. Agar peningkatan kreatifitas dosen tercapai, bagian yang berkaitan dengan pengelolaan data pada perguruan tinggi terutama data base SIPT supaya dikelola dengan baik.

Referensi

- [1] <http://panduan.cs.ui.ac.id/?cat=7>
- [2] <http://web.ipb.ac.id/~tpb/index.php?pilih=hal&id=20>.
- [3] Khoirunnisak, M., Iriawan, N, 2012, *Pemodelan Faktor-faktor yang Mempengaruhi Mahasiswa Berhenti Studi (Drop Out) di Institut Teknologi Sepuluh Nopember Menggunakan Analisis Bayesian Mixture Survival*, <http://digilib.its.ac.id/public/ITS-Undergraduate-13372-Paper.pdf>, tanggal akses : 15 Maret 2012.
- [4] Davies, Beynon, P, 2004, "*Database Systems Third Edition*", Palgrave Macmillan, New York.
- [5] Santosa, B, 2007, "*Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*", Graha Ilmu, Yogyakarta.
- [6] Pramudiono, I. 2003. *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*. <http://www.ilmukomputer.org/wp-content/uploads/2006/08/iko-datamining.zip>, Akses : tanggal 10 Mei 2011.
- [7] Han, J. and Kamber, M, 2006, "*Data Mining Concepts and Techniques Second Edition*". Morgan Kauffman, San Francisco.
- [8] MacLennan, J., Zhao Hui Tang, Bog, Crivat, 2009, "*Data Mining with Microsoft® SQL Server® 2008*", Wiley Publishing, Inc., Indianapolis, Indiana