

PEMANFAATAN APLIKASI DATA MINING DI PERGURUAN TINGGI

Haryanto Tanuwijaya¹⁾

1) Program Studi Sistem Informasi Sekolah Tinggi Manajemen Informatika & Teknik Komputer Surabaya (STIKOM),
email: haryanto@stikom.edu

Abstract: In facing tight competition, college capable of performing learning more efficient and effective and quality learning process, so that has excellent competition with quality services to the customers. Data mining application already used in many company in fact fit used in college. Algorithm and data mining model application help the institution more effective resource allocation and has excellent tool using model in predicting individual behaviour with high accuracy in large database. Using data mining application in college can increase education and learning process quality by using appropriate model analysis.

Keywords: Data mining, College, Algorithm, Modelling, Classification, Decision Tree

Perguruan Tinggi (PT) saat ini dihadapkan pada kondisi kompetisi ketat dan rumit dengan berbagai tuntutan kualitas penyelenggaraan oleh Direktorat Jenderal Pendidikan Tinggi (Dikti) dan kebutuhan/keinginan pelanggan (industri, pemerintahan, orang tua dan peserta didik) serta dunia kerja. Kondisi tersebut membuat perguruan tinggi harus memiliki keunggulan kompetitif dan pelayanan berkualitas untuk memuaskan para pelanggan agar dapat tetap eksis dan bersaing dengan para kompetitornya. Peningkatan kualitas tenaga pengajar sebagai aset perguruan tinggi harus senantiasa dilakukan melalui pemberdayaan secara efektif dan efisien. Untuk itu perguruan tinggi membutuhkan informasi cukup dan 'dalam' untuk melakukan proses penilaian, evaluasi, perencanaan, dan pengambilan keputusan yang lebih berkualitas.

Perguruan Tinggi mengalokasikan sumberdaya, tenaga pengajar dan karyawannya dengan lebih efektif, sehingga dapat menambah wawasan dan pengetahuan lebih ke dalam perilaku sivitas akademika. Perguruan Tinggi dapat memprediksi arah dari mahasiswa dan alumninya dan mengarahkan tantangan kepada para mahasiswa dan alumninya dengan lebih baik melalui analisa dan presentasi data.

Data mining memiliki keunggulan *tool* yang diinginkan dengan kemampuan pelaporan dengan jangkauan dan pola pemahaman dalam basisdata yang luas. Pola ini digunakan dalam pemodelan guna memprediksi perilaku individu dengan akurasi tinggi. Hasilnya *data mining* mampu membantu institusi mengalokasikan sumberdaya dan karyawan dengan lebih efektif, sehingga dapat menambah wawasan dan pengetahuan lebih ke dalam perilaku sumberdayanya.

Permasalahannya adalah bagaimana memanfaatkan aplikasi *data mining* termasuk algoritma dan model yang biasa digunakan di sektor bisnis ke dalam sistem perguruan tinggi. Kontribusi dari tulisan ini adalah mengemukakan kemampuan dan keunggulan aplikasi *data mining* pada sistem Perguruan Tinggi dengan harapan dapat meningkatkan kualitas proses pendidikan dan pengajaran menggunakan analisis model yang sesuai.

TINJAUAN PUSTAKA

1. Konsep Data Mining

Data Mining telah berkembang pesat seiring dengan perkembangan teknologi *Data Warehouse*. *Data mining* melakukan proses penggalian pengetahuan dan informasi yang berguna termasuk pola, asosiasi, perubahan, anomali dan struktur signifikan dari data transaksi tersimpan dalam basisdata, *data warehouse*, atau tempat penyimpanan informasi lainnya (Han and Kamber, 1999). *Data mining* merupakan sebuah *tool* yang digunakan pada berbagai bidang seperti keuangan, perbankan, telekomunikasi, kedokteran, industri retail, dan pendidikan (Beikzadeh and Delavari, 2004). *Data mining* menggali ketidakpahaman, validitas dan gagasan baru yang potensial dengan pola yang dapat dipahami didalam basisdata besar bidang pendidikan. *Data mining* dikenal sebagai langkah pencarian pengetahuan dimana terjadi interaksi dengan pengguna atau basisdata pengetahuan.

Menurut Rubenking (2001), *data mining* adalah “*the process of automatically extracting usefull information and relationships from immense quantities of data*”. Sedangkan menurut Gartner Group, *data mining* adalah “*the process of discovering meaningful new correlations, paterns, and trends mining by sifting through large amounts of data stored in repositories and by using pattern recognition technologies as well as statistical and mathematical techniques*”.

Kemunculan teknologi data mining membuka harapan bagi penyelenggara PT untuk mengadopsi data mining menyangkut 3 (tiga) tugas yang diemban PT yaitu: (1) penelitian berkaitan dengan pengembangan pengetahuan, (2) pengajaran berhubungan dengan transfer pengetahuan, dan (3) penelitian oleh institusi dalam rangka pemanfaatan pengetahuan untuk pengambilan keputusan. (Luan, 2001). Data mining memiliki keunggulan sebagai tool dengan kemampuan jangkauan pelaporan dan pola pemahaman basisdata yang besar. Pola ini digunakan dalam model guna memprediksi perilaku individu dengan akurasi tinggi. Dengan demikian data mining dapat membantu PT mengalokasi sumberdaya dan karyawan dengan lebih efektif, sehingga dapat menambah wawasan dan pengetahuan lebih ke dalam perilaku sivitas akademika.

2. Data Mining Pada Sistem Perguruan Tinggi

Tantangan terpenting yang dihadapi perguruan tinggi saat ini adalah bagaimana dapat menyelenggarakan PT dengan lebih efisien, efektif dan proses pengajaran yang lebih berkualitas. *Data mining* dikenal sebagai teknologi yang tepat untuk memberi wawasan atau pengetahuan lebih pada tenaga pengajar, mahasiswa, alumni, pejabat struktural, dan staf pendidikan lainnya dalam memberikan bantuan yang diperlukan guna pengambilan keputusan yang lebih baik.

Kesenjangan akan kedalaman dan kecukupan pengetahuan pada sistem PT telah menghalangi sistem manajemen untuk mencapai kualitas tujuan yang telah ditetapkan. Teknologi data mining dapat membantu menjembatani kesenjangan pengetahuan pada sistem PT melalui penemuan pola, asosiasi dan anomali yang tersembunyi, sehingga dapat meningkatkan efektifitas, efisiensi, dan kecepatan proses. Peningkatan ini akan memberi keunggulan pada sistem PT antara lain maksimalisasi sistem pendidikan yang efisien, penurunan angka drop out, peningkatan promosi, meningkatkan ingatan mahasiswa, peningkatan rasio pendidikan, peningkatan kesuksesan mahasiswa, peningkatan hasil yang dicapai mahasiswa dan pengurangan biaya proses dari sistem yang berjalan. Untuk mencapai kualitas dari peningkatan tersebut, maka aplikasi data mining sangat dibutuhkan untuk menyediakan kebutuhan pengetahuan dan wawasan bagi pengambil keputusan pada sistem PT. (Beikzadeh and Delavari, 2004).

3. Algoritma dan Tool

Data mining meliputi sejumlah algoritma yang dibedakan dari metode dan maksud tujuan yang ingin dicapai, yang terdiri atas eksplorasi data dan visualisasi hasil akhir yang sesuai kepada para pengguna. Terdapat beberapa algoritma dan *tools* yang dapat digunakan (Merceron and Yacef, ___) yaitu:

- *Tools*, terdapat beberapa tools yang dapat digunakan antara lain Excel dan Access untuk menampilkan hasil SQL *queries* dan visualisasi, Clementine dan Tada-Ed untuk clustering, klasifikasi dan asosiasi aturan, serta SODAS untuk menampilkan analisis data simbolik.
- Eksplorasi dan visualisasi data, untuk memvisualisasikan data mentah dan hasil algoritma dalam tabel dan grafik bentuk histogram atau graph melalui beberapa teknik spesifik seperti analisis data simbolik.
- Algoritma *clustering*, bertujuan menemukan kelompok data yang homogen. Dalam penggunaannya, *k-means clustering* dapat dikombinasikan dengan *hierarchic clustering*.
- Klasifikasi, digunakan untuk memprediksi nilai dari beberapa variabel, seperti hasil pekerjaan dari mahasiswa yang berguna untuk memprediksikan apakah mahasiswa dapat mengerjakan soal ujian dengan baik. Penggunaan *decision tree* dapat mewakili sekumpulan aturan dan merepresentasikan populasi yang digunakan untuk memprediksi nilai bagi individual baru.
- Aturan asosiasi, menemukan hubungan antar item yang dapat dilakukan menggunakan algoritma Apriori.

ANALISIS MODEL

Proses *tracking dan mining* data mahasiswa dalam meningkatkan pengajaran dan pembelajaran sangat relatif tetapi sudah terdapat sejumlah studi dan penelitian mulai menggabungkan gagasan mereka. Penggunaan algoritma yang akan terkait dengan prosedur dalam pemrograman bersifat umum untuk semua sektor baik bisnis maupun pendidikan tinggi. Hal ini dapat dilihat pada Tabel 1 yang memperlihatkan perbandingan pertanyaan *data mining* bidang bisnis dan perguruan tinggi.

Tabel 1. Perbandingan pertanyaan *data mining* bidang bisnis dan perguruan tinggi.

Pertanyaan di dunia bisnis	Pertanyaan di Perguruan Tinggi
Siapa pelanggan yang menguntungkan?	Siapa mahasiswa menempuh sks terbanyak?
Siapa pengunjung tetap website perusahaan?	Siapa yang suka mengulang di banyak kelas?
Siapa pelanggan loyal perusahaan?	Siapa bertahan lama di perguruan tinggi?
Siapa yang meningkatkan pembelian mereka?	Alumni mana yang suka menyumbang lebih?
Pelanggan apa yang berpindah ke pesaing?	Tipe pelajaran apa yang dapat ditawarkan untuk memperoleh lebih banyak mahasiswa?

(Sumber: Luan, 2001)

Salah satu pemodelan klasifikasi mahasiswa adalah menggunakan metodologi CRISP-DM (Cross Industry Process for Data Mining) mulai dari pemahaman tujuan Perguruan Tinggi (PT), pemahaman data akademik, persiapan data sampai dengan pemodelan dan evaluasi. Langkah dalam pemodelan klasifikasi ini adalah:

1. Memilih tujuan utama dari peningkatan kemampuan mahasiswa pada matakuliah tertentu.
2. Memahami data dengan inisialisasi koleksi data, kemudian menjelaskan, menggali dan mengecek kualitas data, mengidentifikasi permasalahan kualitas data.
3. Langkah ketiga adalah persiapan data, meliputi semua aktifitas dimana data set akhir dibangun untuk dimasukkan dalam teknik pemodelan.
4. Pemodelan, guna memperoleh hubungan antara variabel khusus (seperti sejarah pengambilan matakuliah mahasiswa), variabel target (status keberhasilan mahasiswa) dan variabel lain (indeks prestasi mahasiswa, jumlah mahasiswa yang gagal) dari data.

Model yang dianalisis yang terdapat pada Tabel 2 merupakan model parsial yang sudah diteliti dan diterapkan selama ini. Pengembangan model telah dilakukan untuk memperoleh hasil yang lebih optimal apabila diterapkan di perguruan tinggi. Tabel 3 menunjukkan proses utama dan sub-proses yang lebih detail dengan beberapa contoh proses *data mining* saja.

Tabel 2. Model Analisis Parsial

Proses Utama	Sub-Proses	Pengetahuan	Proses Baru melalui Data Mining	Fungsi Data Mining
Perencanaan	Perencanaan jadwal tenaga pengajar/ dosen	<ul style="list-style-type: none"> • Pola jadwal dosen yang lalu • Prediksi jadwal dosen yang akan datang 	<ul style="list-style-type: none"> • Prediksi jadwal dosen untuk tahun mendatang 	<ul style="list-style-type: none"> • Prediksi

	Perencanaan kegiatan alumni	<ul style="list-style-type: none"> • Pola kontribusi alumni masa lalu • Pola alumni yang melanjutkan studi • Pola alumni menemukan pekerjaan yang sesuai • Pola alumni yang tidak aktif dalam komunitas 	<ul style="list-style-type: none"> • Prediksi janji alumni • Prediksi kemungkinan alumni melanjutkan studi setelah lulus • Prediksi kemungkinan alumni yang menemukan pekerjaan sesuai • Prediksi alumni yang tidak aktif dalam komunitas 	<ul style="list-style-type: none"> • Prediksi • Prediksi • Prediksi • Prediksi
Evaluasi	Penilaian mahasiswa	<ul style="list-style-type: none"> • Pola mahasiswa dengan nilai rendah • Pola karakteristik mahasiswa berprestasi 	Prediksi kemungkinan sukses	<ul style="list-style-type: none"> • Prediksi
		<ul style="list-style-type: none"> • Pola nilai test mahasiswa wanita dan pria • Asosiasi perbedaan jenis kelamin pada nilai test 	Asosiasi perbedaan jenis kelamin dengan perbedaan nilai test mahasiswa	<ul style="list-style-type: none"> • Asosiasi
Konsultasi	Konsultasi perilaku mahasiswa	<ul style="list-style-type: none"> • Pola perilaku mahasiswa yang lalu • <i>Cluster</i> bermacam-macam karakteristik mahasiswa 	<ul style="list-style-type: none"> • Prediksi pola permasalahan perilaku mahasiswa • Clustering hasil analisis karakteristik mahasiswa 	<ul style="list-style-type: none"> • Prediksi • Clustering, Prediksi • Clustering
	Konsultasi pemilihan jurusan	<ul style="list-style-type: none"> • Pola karakteristik mahasiswa berbagai jurusan yang lalu • Asosiasi jurusan dengan tipe mahasiswa 	<ul style="list-style-type: none"> • Klasifikasi mahasiswa dengan jurusan yang sesuai • Asosiasikan mahasiswa dalam jurusan 	<ul style="list-style-type: none"> • Klasifikasi • Asosiasi
	Konsultasi pengambilan matakuliah	<ul style="list-style-type: none"> • Asosiasi mahasiswa dengan matakuliah • Pola karakteristik mahasiswa lalu pada beberapa matakuliah • Asosiasi matakuliah dengan tipe mahasiswa 	<ul style="list-style-type: none"> • Asosiasi tipe mahasiswa dengan matakuliah yang paling sesuai • Asosiasi tipe mahasiswa dengan ketepatan pemilihan matakuliah 	<ul style="list-style-type: none"> • Asosiasi • Asosiasi

(Sumber: Beikzadeh and Delavari, 2004)

Tabel 3. Pengembangan Model Untuk Diterapkan Di Perguruan Tinggi.

Proses Utama	Sub-Proses	Pengetahuan	Proses Baru melalui Data Mining	Fungsi Data Mining
--------------	------------	-------------	---------------------------------	--------------------

Evaluasi	Penilaian terhadap Mahasiswa	<ul style="list-style-type: none"> • Pola hasil belajar mahasiswa yang lalu • Prediksi hasil belajar 	<ul style="list-style-type: none"> • Prediksi hasil belajar mahasiswa • Mengartikan tipologi hasil belajar dengan lama studi 	<ul style="list-style-type: none"> • Prediksi • Clustering
		<ul style="list-style-type: none"> • Pola keberhasilan mahasiswa matakuliah tertentu • Pola keberhasilan yang dicapai mahasiswa pada matakuliah 	Pengelompokan mahasiswa dalam kelompok berhasil dan gagal	Klasifikasi
		<ul style="list-style-type: none"> • Pola mahasiswa dengan nilai test rendah • Pola karakteristik mahasiswa berprestasi 	Prediksi kemungkinan sukses	Prediksi
		<ul style="list-style-type: none"> • Pola nilai test mahasiswa pria dan wanita sebelumnya • Asosiasi informasi individu mahasiswa dengan nilai test 	<ul style="list-style-type: none"> • Asosiasi informasi individu mahasiswa (jenis kelamin, etnis, usia, status, tempat lahir) dengan nilai test 	<ul style="list-style-type: none"> • Asosiasi
	Penilaian terhadap Dosen	<ul style="list-style-type: none"> • Pola karakteristik dosen sebelumnya yang lebih efektif dari yang lain • Asosiasi pelatihan dosen dengan nilai test mahasiswa 	<ul style="list-style-type: none"> • Asosiasi pelatihan dosen dengan nilai test mahasiswa mereka 	<ul style="list-style-type: none"> • Prediksi, Klasifikasi
		<ul style="list-style-type: none"> • Pola dosen sebelumnya menurut tingkatan nilai test mahasiswa • Asosiasi pelatihan dosen dengan nilai test dosen 	<ul style="list-style-type: none"> • Asosiasi pelatihan dosen dengan nilai test mahasiswa 	<ul style="list-style-type: none"> • Asosiasi
	Penilaian matakuliah	<ul style="list-style-type: none"> • Pola karakteristik dosen yang lebih efektif dari yang lain 	<ul style="list-style-type: none"> • Prediksi dosen paling efektif dalam setahun terkait dengan hasil pembelajaran 	<ul style="list-style-type: none"> • Prediksi, Klasifikasi
		<ul style="list-style-type: none"> • Asosiasi pelatihan dosen dengan nilai mahasiswa mereka 	<ul style="list-style-type: none"> • Asosiasi pelatihan dosen dengan nilai mahasiswa mereka 	<ul style="list-style-type: none"> • Asosiasi
	Penilaian pelatihan industri	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> • 	<ul style="list-style-type: none"> •
Registrasi	Registrasi mahasiswa	<ul style="list-style-type: none"> • Pola pengambilan matakuliah mahasiswa sebelumnya • Asosiasi mahasiswa dengan matakuliah paling sesuai • Klasifikasi mahasiswa dengan matakuliah paling sesuai 	<ul style="list-style-type: none"> • Prediksi tipe mahasiswa sesuai dengan tipe matakuliah • Asosiasi mahasiswa dengan berbagai tipe matakuliah • Klasifikasi mahasiswa dengan matakuliah yang sesuai selama studi 	<ul style="list-style-type: none"> • Prediksi • Asosiasi • Klasifikasi
Kinerja	Kinerja mahasiswa	<ul style="list-style-type: none"> • Asosiasi kinerja dengan IPK • Asosiasi kinerja dengan 	<ul style="list-style-type: none"> • Asosiasi kinerja dengan IPK • Asosiasi kinerja dengan 	<ul style="list-style-type: none"> • Asosiasi • Asosiasi

		perilaku akademik <ul style="list-style-type: none"> • Asosiasi kinerja dengan perencanaan studi • Asosiasi kehadiran mahasiswa dengan situasi kelas • Klasifikasi waktu dan tempat kelas berbeda • Pola keberhasilan mahasiswa berperilaku buruk • Pola kinerja mahasiswa dengan kepuasan terhadap dosen 	perilaku akademik <ul style="list-style-type: none"> • Asosiasi kinerja dengan perencanaan studi • Asosiasi kehadiran mahasiswa dengan situasi kelas • Klasifikasi waktu dan tempat kelas berbeda • Prediksi kemungkinan keberhasilan mahasiswa berperilaku buruk • Prediksi kemungkinan kinerja mahasiswa yang tidak puas terhadap dosen 	<ul style="list-style-type: none"> • Asosiasi • Asosiasi • Klasifikasi • Prediksi • Prediksi
	Kinerja Dosen	<ul style="list-style-type: none"> • Asosiasi kemampuan mengajar dosen dengan nilai mahasiswa • Asosiasi dosen yang sering tidak masuk dengan nilai mahasiswa • Asosiasi kinerja dengan perilaku dosen • Asosiasi kinerja dosen • Asosiasi kompetensi dosen 	<ul style="list-style-type: none"> • mengajar dosen dengan nilai mahasiswa • Asosiasi dosen yang sering tidak masuk dengan nilai mahasiswa • Asosiasi kinerja dengan perilaku dosen • Asosiasi kinerja dosen • Asosiasi kompetensi dosen 	<ul style="list-style-type: none"> • Asosiasi • Asosiasi • Asosiasi • Asosiasi • Asosiasi
Ujian	Ujian mahasiswa	<ul style="list-style-type: none"> • Asosiasi tingkatan ujian dengan nilai mahasiswa • Asosiasi tingkatan ujian dengan kinerja kelas dosen 	<ul style="list-style-type: none"> • Asosiasi tingkatan ujian dengan nilai mahasiswa • Asosiasi tingkatan ujian dengan kinerja dosen 	<ul style="list-style-type: none"> • Asosiasi • Asosiasi
Konseling	Konsultasi perilaku mahasiswa	(Sama dengan Tabel 2)	(Sama dengan Tabel 2)	(Sama dengan Tabel 2)
	Konsultasi pemilihan program	(Sama dengan Tabel 2)	(Sama dengan Tabel 2)	(Sama dengan Tabel 2)

(Sumber: Delavari, 2005)

SIMPULAN

Dalam upaya agar tetap eksis dalam kompetisi ketat dan berbagai tuntutan kualitas proses penyelenggaraan, maka perguruan tinggi perlu memiliki keunggulan sistem penyelenggaraan perguruan tinggi. Perguruan tinggi perlu mengatasi kesenjangan pengetahuan pada sistem perguruan tinggi melalui penemuan pola, asosiasi dan anomali yang tersembunyi guna meningkatkan efektifitas, efisiensi dan kecepatan proses sehingga dapat memaksimalkan sistem pendidikan yang efisien. Untuk itu dibutuhkan *tool* yang dapat membantu meningkatkan kualitas penyelenggaraan dan pengambilan keputusan di perguruan tinggi.

Aplikasi *data mining* memiliki keunggulan *tool* menggunakan pemodelan untuk memprediksi perilaku individu dengan akurasi tinggi dalam basisdata besar. Aplikasi *data mining* membantu institusi mengalokasikan sumberdaya dengan lebih efektif. Algoritma dan pemodelan aplikasi *data mining* yang telah banyak digunakan dalam dunia bisnis, ternyata juga sesuai digunakan di perguruan tinggi. Aplikasi *data mining* dapat membantu manajemen dalam pengambilan keputusan berdasarkan individu, membangun strategi baru dan meningkatkan kemampuan dan menyempurnakan prosedur pengambilan keputusan

menjadi lebih baik. Penerapan aplikasi *data mining* di Perguruan Tinggi dapat meningkatkan kualitas proses pendidikan dan pengajaran menggunakan analisis model yang sesuai.

DAFTAR RUJUKAN

Beikzadeh, M.R., and Delavari, N. 2004. A New Analysis Model for Data Mining Processes in Higher Educational Systems. *M2USIC*.

Delavari, N. 2005. Application of Enhanced Analysis Model for Data Mining Processes in Higher Educational System. *IEEE*.

Han, J., and Kamber, M. 1999. Introduction to Data Mining and Knowledge Discovery. *Two Crows Corporation. Third Edition, USA*.

Luan, J. 2001. Data Mining as Driven by Knowledge Management in Higher Education- Persistence Clustering and Prediction. *Public Conference, UCSF*.

Merceron, A., and Yacef, K. _____. Educational Data Mining: a Case Study. ____ .

Rubenking, N. 2001. Hidden Messages. *PC Magazine, May 22 2001*.

STIKOM SURABAYA