

APLIKASI DATA MINING UNTUK MEMPREDIKSI MATAKULIAH PILIHAN PADA PROGRAM STUDI TEKNIK INFORMATIKA STMIK AKBA

Ashari

Program Studi Teknik Informatika, STMIK AKBA

[*ashari.akba36@gmail.com*](mailto:ashari.akba36@gmail.com)

Kemajuan teknologi informasi memungkinkan penyimpanan data dalam skala besar. Perkembangan ini merambah ke berbagai bidang, termasuk pendidikan, sehingga pemanfaatan teknologi informasi dalam bidang pendidikan sudah merupakan hal yang umum. Data mining merupakan kegiatan yang meliputi pengumpulan dan pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar. Dengan memanfaatkan data mahasiswa dan data matakuliah pada kurikulum program studi teknik informatika STMIK AKBA dapat dilakukan prediksi peminatan tertentu sebuah mata kuliah pilihan. Banyak teknik yang dapat digunakan untuk melakukan prediksi matakuliah. Penelitian ini bertujuan memprediksi mata kuliah pilihan apakah dapat dibuka atau sebaiknya ditutup saja berdasarkan data sebelumnya. Teknik yang digunakan dalam melakukan prediksi ini adalah dengan aturan Asosiasi (Association Rule). Aturan ini dipilih karena cukup sederhana, sehingga proses prediksi tidak memakan waktu terlalu lama. Hasil Pengujian menunjukkan bahwa matakuliah Aplikasi Internet sebaiknya ditutup saja karena supportnya hanya 10% dan confidencenya terhadap matakuliah pilihan lain hanya 33%, sedangkan mata kuliah Sistem teknologi Intranet/Extranet tetap dibuka karena supportnya 50% dan confidence diatas 75%.

Kata Kunci : *Data Mining, Prediksi, Association Rule, Mata Kuliah Pilihan*

1. PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi informasi sangat pesat dalam era globalisasi sekarang ini, sehingga manusia dituntut agar dapat mengikuti perkembangan dan selalu berusaha untuk mengadopsi teknologi informasi terbaru untuk membantu kelancaran organisasi. Teknologi informasi tidak lagi dipandang hanya sebagai pelengkap akan tetapi sudah menjadi salah satu penentu atas terlaksananya sasaran atau strategi organisasi yang handal. Dukungan teknologi informasi terhadap organisasi umumnya berbentuk sistem yang terdiri dari sistem informasi dan infrastruktur pendukungnya.

Banyaknya data yang dipunyai sebuah organisasi bisa menyebabkan kesulitan dalam proses menganalisis data, mengubah data menjadi informasi dan menemukan pola yang dapat menjadi pengetahuan untuk kepentingan organisasi. Proses yang dilakukan oleh manusia masih memiliki keterbatasan, terutama pada kemampuan manusia dalam menampung jumlah data yang ingin diolah. Selain itu bisa juga terjadi kesalahan dalam pengelolaan data yang dilakukan. Salah satu cara mengatasi masalah ini adalah dengan menggunakan *data mining*.

Sebagai cabang ilmu di bidang komputer, cukup banyak penerapan yang dilakukan *data mining*. *Data mining* berkembang dan diterapkan dalam dunia pendidikan karena ditunjang oleh ketersediaan data.

Program Studi Teknik Informatika merupakan salah satu program studi yang ada di STMIK AKBA menerapkan kurikulum berbasis kompetensi dengan kurikulum yang mencakup berbagai jenis matakuliah. Struktur kurikulum telah dirancang dan ditetapkan berdasarkan kebutuhan *stakeholders*. Untuk mendukung kebutuhan *stakeholders* matakuliah disusun sebagai matakuliah wajib yang terdiri dari matakuliah inti dan institusional serta matakuliah pilihan.

Matakuliah pilihan pada Program Studi Teknik Informatika STMIK AKBA diambil berdasarkan bidang peminatan rekayasa perangkat lunak atau peminatan sistem komputer jaringan. Matakuliah pilihan yang ditawarkan pada umumnya tidak semua diambil oleh mahasiswa. Matakuliah pilihan yang tersedia pada semester V ada tujuh, namun yang diambil hanya lima atau enam matakuliah saja. Banyak hal yang menjadi penyebab mahasiswa hanya memilih matakuliah pilihan tertentu saja. Faktor kepentingan dan ketertarikan mahasiswa terhadap sebuah mata kuliah pilihan, atau bahkan mungkin faktor pengajarnya. Penerapan aturan asosiasi (*Association Rule*) dalam area *data mining* dapat digunakan untuk memprediksi matakuliah pilihan yang kemungkinan diprogramkan atau tidak diprogramkan oleh mahasiswa teknik informatika pada STMIK AKBA.

Penerapan aturan asosiasi (*Association Rule*) sebagai *tools* untuk memprediksi mata kuliah pilihan yang dapat dibuka atau ditutup tidak menjamin untuk mendapatkan tingkat kebenaran 100%. Hasil yang diperoleh hanya dapat digunakan sebagai saran, sebab untuk membuka dan menutup mata kuliah pilihan perlu juga mempertimbangkan faktor-faktor lain. Seperti misalnya seorang mahasiswa pada tahun sebelumnya tidak lulus pada sebuah mata kuliah pilihan, maka sebaiknya mata

kuliah pilihan tersebut tetap dibuka meskipun mahasiswa yang mengambil jumlahnya di bawah jumlah minimal. Terlepas dari berbagai faktor non teknis tersebut *association rule* sebagai salah satu teknik data mining dapat membantu untuk memprediksi mata kuliah pilihan dapat dibuka atau ditutup berdasarkan data pada semester sebelumnya.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah tersebut, dapat dirumuskan permasalahan yang akan diselesaikan yaitu bagaimana membuat aplikasi data mining yang dapat memprediksi matakuliah pilihan pada Program Studi Teknik Informatika STMIK AKBA.

1.3. Tujuan Penelitian

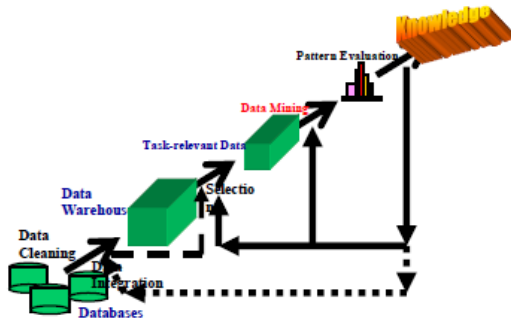
Tujuan penelitian ini adalah membuat aplikasi data mining yang dapat memprediksi matakuliah pilihan pada Program Studi Teknik Informatika STMIK AKBA.

2. LANDASAN TEORI

2.1 Data Mining

Data mining merupakan suatu metode menemukan suatu pengetahuan dalam suatu database yang cukup besar. Data mining adalah proses menggali dan menganalisa sejumlah data yang sangat besar untuk memperoleh sesuatu yang benar, baru, sangat bermanfaat dan akhirnya dapat dimengerti suatu corak atau pola dalam data tersebut (Han & Kamber, 2006).

Data mining adalah bagian integral dari penemuan pengetahuan dalam database (KDD), yang merupakan proses keseluruhan mengubah data mentah menjadi informasi yang bermanfaat, seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Proses penemuan pengetahuan dalam database / KDD (Han & Kamber, 2006)

Proses KDD ini terdiri dari serangkaian langkah-langkah transformasi, dari proses data *preprocessing* dan proses data *postprocessing* dari data yang merupakan hasil penggalian. Input data dapat disimpan dalam berbagai format (flat file, spreadsheet, atau relasional tabel) dan mungkin berada dalam penyimpanan data terpusat atau didistribusikan di beberapa alamat. Tujuan dari proses data *preprocessing* adalah untuk mengubah data input mentah menjadi format yang sesuai untuk analisis selanjutnya. Langkah-langkah yang dilakukan antara lain dengan memperbaiki data yang kotor atau ganda, dan memilih catatan dan fitur yang relevan dengan proses pengelolaan data selanjutnya. Karena banyak cara data dapat dikumpulkan dan disimpan, maka proses pengolahan data mungkin akan melelahkan dan memakan waktu yang lama dalam keseluruhan proses penemuan pengetahuan (Tan, dkk, 2006).

Terdapat empat tugas utama *data mining* seperti terlihat pada gambar 2, yaitu

1. Predictive Modelling

Predictive modelling digunakan untuk membangun sebuah model untuk target *variable* sebagai fungsi dari *explanatory variable*. *Explanatory variable* dalam hal ini merupakan semua atribut yang digunakan untuk melakukan prediksi,

sedangkan variabel target merupakan atribut yang akan diprediksi nilainya.

Predictive modelling dibagi menjadi dua tipe yaitu : *Classification* digunakan untuk memprediksi nilai dari target *variable* yang *discrete* (diskrit) dan *regression* digunakan untuk memprediksi nilai dari target *variable* yang *continue* (berkelanjutan).

2. Association Analysis

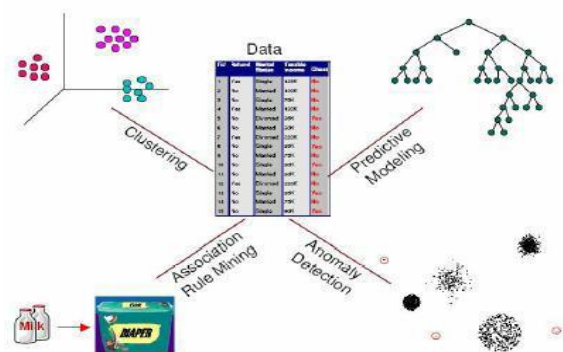
Association analysis digunakan untuk menemukan aturan asosiasi yang memperlihatkan kondisi-kondisi nilai atribut yang sering muncul secara bersamaan dalam sebuah himpunan data.

3. Cluster Analysis

Tidak seperti klasifikasi yang menganalisa kelas data obyek yang mengandung label. *Clustering* menganalisa objek data tanpa memeriksa kelas label yang diketahui. Label-label kelas dilibatkan di dalam data training. Karena belum diketahui sebelumnya. *Clustering* merupakan proses pengelompokkan sekumpulan objek yang sangat mirip.

4. Anomaly Detection

Anomaly detection merupakan metode pendeteksian suatu data dimana tujuannya adalah menemukan objek yang berbeda dari sebagian besar objek lain. *Anomaly* dapat di deteksi dengan menggunakan uji statistik yang menerapkan model distribusi atau probabilitas untuk data.



Gambar 2. Tugas utama data mining (Tan, dkk, 2006)

Penelitian ini akan difokuskan pada teknik *association* terutama dalam hubungannya untuk memprediksi mata kuliah pilihan yang mungkin dapat dibuka atau tidak dibuka.

2.2 Association Rule

Association Rule merupakan teknik data mining untuk mencari pola hubungan dalam data atau basis data. Yang paling populer adalah *market basket analysis* (MBA). Teknik MBA ini digunakan untuk mencari pelanggan pada sebuah toko yang biasanya membeli barang X, akan membeli barang Y dan seterusnya (Widodo, 2008).

Tabel 1. Contoh tabel transaksi untuk *association analysis*

Trans	Itemset
1	A,B,E
2	B,D
3	B,C
4	A,B,D
5	A,C
6	B,C
7	A,C
8	A,B,C,E
9	A,B,C

Dalam *association analysis* itemset merupakan kumpulan nol atau lebih item [...]. Pada contoh tabel 1 diumpamakan itemset adalah barang yang dibeli pada toko tersebut, trans adalah kode transaksi. Pada tabel tersebut dapat dilihat pada transaksi 1 yang dibeli adalah barang A,B, dan E. Transaksi 2 yang dibeli adalah B dan D, demikian seterusnya. *Association analysis* akan melakukan analisis hubungan dengan aturan tertentu sering disebut sebagai *association rule*.

Sebagai contoh pada tabel tersebut, pembeli yang biasanya membeli barang A dan B akan membeli barang C. Aturan ini dapat ditulis $\{A,B\} \rightarrow \{C\}$. Untuk memperoleh kuantifikasinya terminologi *support* dan *confidence* perlu diperhatikan.

Support dan *confidence* dapat dihitung sebagai berikut:

$$\text{Support, } s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N}$$

$$\text{Confidence, } c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)}$$

Dengan rumus tersebut maka *support* yang di dapat dari hubungan $s(A,B \rightarrow C)$ adalah: $2/9 = 0,22$. Sedangkan nilai *confidence* yang diperoleh adalah $c(A,B \rightarrow C) 2/4 = 0,5$. Nilai 2 pada *support* diperoleh dari itemset A,B,C berjumlah 2 yaitu pada trans 8 dan 9, jumlah keseluruhan trans adalah 9. Sedangkan 2 pada *confidence* sama seperti pada *support*, 4 diperoleh dari jumlah A,B yaitu pada trans 1,4,8, dan 9. Dari perolehan tersebut, kita dapat membuat aturan persentasi yang diinginkan untuk menentukan kepastian asosiasinya. Paling baik adalah apabila nilai *support* maupun *confidence* tinggi.

2.3 Prediksi Matakuliah Pilihan

Matakuliah pilihan dalam sebuah perguruan tinggi adalah pilihan bagi seorang mahasiswa, apakah akan diambil atau tidak. Hal ini karena matakuliah pilihan tersebut tidak wajib untuk diambil, sehingga dapat terjadi satu matakuliah pilihan diambil oleh mahasiswa yang jumlahnya di bawah jumlah minimum. Apabila hal tersebut terjadi, maka mata kuliah pilihan tersebut akan ditutup kembali. Aturan asosiasi (*association rule*) dapat melakukan prediksi untuk hal tersebut, tentu dengan data selama beberapa tahun yang dapat dicari kecenderungannya atau pola pada setiap tahunnya. Jika ini dilakukan mungkin akan dapat ditawarkan pada awal semester, mata kuliah mana yang dapat dibuka dan mata kuliah mana yang tidak perlu dibuka (Widodo, 2008).

Tabel 2. Matakuliah Pilihan Semester V Program Studi Teknik Informatika STMIK AKBA

Kode	Matakuliah Pilihan
A	Sistem Teknologi Intranet/Extranet
B	Manajemen Jaringan
C	Aplikasi Internet
D	Perancangan Sistem Informasi
E	Pemrograman Berorientasi Objek
F	Pemrograman Hardware dan Interface
G	Mikro Elektronika

Matakuliah pilihan pada semester V tersebut dalam tabel 2 dapat dijadikan sebagai itemset sebagaimana dalam contoh tabel 1, sedangkan yang menjadi transnya adalah mahasiswa semester V yang memprogramkan matakuliah pilihan tersebut.

3. Analisis dan Perancangan

Teknik analisis data dalam pembuatan perangkat lunak menggunakan pemodelan perangkat lunak dengan paradigma *waterfall* seperti terlihat pada Gambar 3 yang meliputi beberapa proses diantaranya:

a. System / Information Engineering

Merupakan bagian dari sistem yang terbesar dalam pengerjaan suatu proyek, dimulai dengan menetapkan berbagai kebutuhan dari semua elemen yang diperlukan sistem dan mengalokasikannya kedalam pembentukan perangkat lunak.

b. Analisis

Merupakan tahap menganalisis hal-hal yang diperlukan dalam pelaksanaan proyek pembuatan perangkat lunak.

c. Design

Tahap penerjemahan dari data yang dianalisis kedalam bentuk yang mudah dimengerti oleh *user*.

d. Coding

Tahap penerjemahan data atau pemecahan

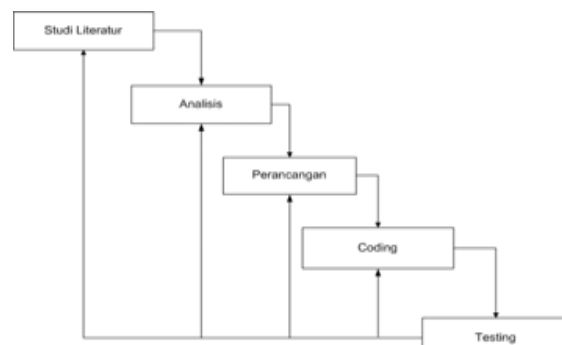
masalah yang telah dirancang kedalam bahasa pemrograman tertentu.

e. Pengujian

Merupakan tahap pengujian terhadap perangkat lunak yang dibangun.

f. Maintenance

Tahap akhir dimana suatu perangkat lunak yang sudah selesai dapat mengalami perubahan-perubahan atau penambahan sesuai dengan permintaan *user*.



Gambar 3. Paradigma Waterfall

3.1 Analisa Sistem

Analisis Sistem dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh ke dalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan yang terjadi dan kebutuhan yang diharapkan sehingga dapat diusulkan perbaikan-perbaikan. Atau secara lebih mudahnya, analisis sistem adalah penelitian atas sistem yang telah ada dengan tujuan untuk merancang sistem yang baru atau diperbarui. Tahap analisis sistem ini merupakan tahap yang sangat kritis dan sangat penting, karena kesalahan di dalam tahap ini akan menyebabkan juga kesalahan di tahap selanjutnya. Tugas utama analisis sistem dalam tahap ini adalah menemukan kelemahan dari sistem yang berjalan sehingga dapat diusulkan perbaikannya (Irdiansyah, 2009).

3.2 Analisa Basis Data

Struktur logika dari basis data dapat digambarkan dalam sebuah grafik dengan menggunakan *Entity Relationship Diagram* (ERD). ERD merupakan hubungan antara entitas yang digunakan dalam sistem untuk menggambarkan hubungan antara entitas atau struktur data dan relasi antar file.

Komponen utama pembentukan ERD yaitu *Entity* (entitas) dan *Relation* (relasi) sehingga dalam hal ini ERD merupakan komponen-komponen himpunan entitas dan himpunan relasi yang dideskripsikan lebih jauh melalui sejumlah atribut-atribut (*property*) yang menggambarkan seluruh fakta dari sistem yang ditinjau.

3.3 Perancangan Struktur Menu

Dalam perancangan sebuah aplikasi dibutuhkan struktur menu yang berisikan menu dan submenu yang berfungsi untuk memudahkan *user* dalam menggunakan aplikasi tersebut.

4. Implementasi

4.1 Lingkungan Pengembangan

Aplikasi ini dibuat dengan spesifikasi komputer sebagai berikut:

1. Sistem Operasi : Windows 7 Ultimate
2. *Processor* : Intel Core i3-2330M
3. RAM : 2 GB DDR3

4.2 Pemeliharaan Sistem

Pemeliharaan terhadap sistem yang telah dibangun secara khusus tidak harus dilakukan pemeliharaan, kecuali jika sistem ingin dikembangkan. Pemeliharaan yang dapat dilakukan adalah pada sisi database, dan program perhitungan prediksi peserta matakuliah pilihan.

4.2 Pengujian Program

Untuk menguji program ini disediakan data dari Kelas T.51 sebagaimana dalam tabel 3:

Tabel 3 Data Pengujian Program

NIM	Angkatan	MK Pilihan
20112205001	2011	A, D, E
20112205002	2011	B, D, E
20112205003	2011	A, D, E
20112205005	2011	B, F, G
20112205006	2011	C, D, E
20112205007	2011	A, D, E
20112205008	2011	B, F, G
20112205009	2011	A, D, E
20112205010	2011	B, F, G
20112205011	2011	A, D, E
20112205012	2011	A, F, G
20112205015	2011	A, D, E
20112205016	2011	A, D, E
20112205017	2011	C, F, G
20112205028	2011	A, D, E
20112205029	2011	A, D, E
20112205030	2011	B, D, E
20112205031	2011	C, F, G
20112205032	2011	C, D, E
20112205033	2011	A, D, E

Contoh Hasil Pengujian :

Pengujian 1:

Matakuliah $s(D,E \rightarrow A)$ adalah $10/20=0,5$

Matakuliah $c(D,E \rightarrow A)$ adalah $10/13=0,77$

Pengujian 2:

Matakuliah s(F,G→C) adalah $2/20=0,1$

Matakuliah c(F,G→C) adalah $2/6=0,33$

Kesimpulan Hasil Pengujian

Pengujian 1

Matakuliah Sistem Teknologi Intranet/Extranet (A) tetap dibuka karena *support* yang diperoleh terhadap matakuliah peminatan matakuliah perancangan sistem informasi dan matakuliah pemrograman berorientasi objek adalah 50% dan *confidence* di atas 75% dari persyaratan yang ditentukan oleh program studi teknik informatika untuk matakuliah yang berhubungan dengan matakuliah tersebut.

Pengujian 2

Matakuliah Aplikasi Internet (C) sebaiknya ditutup karena *support* yang diperoleh terhadap matakuliah peminatan matakuliah pemrograman hardware dan Interface dan matakuliah mikro elektronika adalah 10% dan *confidence* hanya 33% dari persyaratan *support* minimal atau sama dengan 50% dan *confidence* terhadap matakuliah lain minimal atau sama dengan 75% dari persyaratan yang ditentukan oleh program studi teknik informatika.

5. Penutup

Berdasarkan hasil analisis dan implementasi yang telah dilakukan guna pembuatan aplikasi data mining untuk memprediksi matakuliah pilihan pada program studi teknik informatika STM IK AKBA, maka dapat ditarik kesimpulan sebagai berikut:

- a. Teknik *association rule* (aturan asosiasi) dengan penghitungan *support* dan *confidencenya* dapat memprediksi mata kuliah pilihan yang ditawarkan untuk dibuka atau tidak dibuka.

- b. Dengan aplikasi yang berbasis teknologi informasi dihasilkan sebuah metode yang bisa memprediksi matakuliah pilihan dibuka atau tidak dibuka pada program studi teknik informatika STM IK AKBA.
- c. Nilai *support* dan *confidence* ditentukan oleh penentu kebijakan untuk memastikan prediksinya. Semakin besar nilai *support* dan *coinfidencenya*, maka semakin besar kemungkinan sebuah mata kuliah pilihan dapat dibuka.

Daftar Pustaka

- [1] Anonim, 2012, *Buku Pedoman STM IK AKBA 2012*, Makassar
- [2] Kusrini, dan Luthfi E,T. 2009. *Algoritma Data Mining*. Andi Offset: Yogyakarta
- [3] Han, J. & Kamber, M. (2006). *Data mining: Concepts and Techniques*. Morgan-Kaufman: New York
- [4] Helmy Mukhammad Yunan, Kushartantya, dan Nurdin bahtiar, 2012, *Implementasi data Mining untuk Memprediksi Kelayakan Permintaan Pinjaman Nasabahdi Lembaga Keuangan*, Prosiding Seminar Nasional Ilmu Komputer Universitas Diponegoro Tahun 2012, Graha Ilmu: Yogyakarta
- [5] Irdiansyah Enur, 2009, *Penerapan Data Mining pada Penjualan Produk Minuman di PT. PEPSI COLA INDOBEVERAGES Menggunakan metode clustering*, Jurnal TA/Skripsi Unversitas Gunadarma: Jakarta
- [6] Lindawati, 2008, *Data Mining dengan Teknik Clustering dalam Mengklasifikasian Data Mahasiswa Studi Kasus Prediksi Lama Studi*

- Mahasiswa Universitas Bina Nusantara, Seminar Nasional Informatika 2008 (semnasIF 2008) : Yogyakarta, ISSN: 1979-2328*
- [7] Meinanda Muhamad Hanif, dkk, 2009, *Prediksi Masa Studi Sarjana dengan Artificial Neural Network*, Internetworking Indonesia Journal, Vol. 1/No. 2 (2009).
- [8] Santosa Budi, 2007. *Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bisnis*. Graha Ilmu: Yogyakarta
- [9] Tan, P., Steinbach, M., & Kumar, V., (2006). *Introduction to Data Mining*, Pearson Education
- [10] Wahyudi Eko Nur, Arief Jananto dan Narwati, 2011, *Analisa Profil Data Mahasiswa Baru terhadap Program Studi yang dipilih di Perguruan Tinggi Swasta Jawa Tengah dengan Menggunakan Teknik Data Mining*, Jurnal Teknologi Informasi DINAMIK Volume 16, No.1, Januari 2011 : 29-43, ISSN : 0854-9524
- [11] Widodo, 2008, *Prediksi Mata Kuliah Pilihan dengan Aturan Asosiasi*, e-Indonesia Initiative 2008 (eII2008) : Jakarta.