

IMPLEMENTASI DATAMINING UNTUK MENDAPATKAN POLA PENGUNDURAN DIRI MAHASISWA STUDI KASUS MAHASISWA TEKNIK INFORMATIKA UNIVERSITAS DARMA PERSADA

Suzuki Syofian¹

¹ Dosen Teknik Informatika, Universitas Darma Persada

Abstrak

Program studi teknik informatika merupakan program studi yang baru berdiri di Fakultas Teknik Universitas Darma Persada pada tahun 2002. Animo masyarakat cukup besar terhadap program studi ini, terbukti dari laporan bagian panitia pendaftaran mahasiswa baru (PPMB) minat calon mahasiswa urutan kedua setelah program studi sastra jepang yang ada di Universitas Darma Persada.

Tahun akademik berjalan, ada kecenderungan mahasiswa tersebut mengundurkan diri tidak melakukan registrasi ulang. Hal ini ada dua kemungkinan pertama mahasiswa tersebut cuti dan yang kedua mahasiswa tidak melanjutkan kuliah.

Penelitian ini membahas mahasiswa yang tidak melanjutkan kuliah dengan kata lain mengundurkan diri. Pola undur diri ini hampir terjadi tiap semester. Namun sangat di sayangkan mahasiswa yang mungkin berpotensi tidak mampu dipertahankan untuk belajar di program studi teknik informatika.

Tujuan dari penelitian ini mengimplementasikan algoritma apriori untuk mendapatkan pola pengunduran diri mahasiswa. Sebagai parameter tambahan yang dijadikan isian data kuesioner yaitu dari variable-variabel salah mengambil jurusan, biaya, lokasi, dosen, hubungan antar mahasiswa dan sarana perkuliahan yang dijadikan sebagai tolok ukur dari penyebaran informasi yang diolah. Hasil yang diharapkan dalam penelitian ini adanya aplikasi datamining yang dapat menampilkan informasi pola pengunduran diri mahasiswa teknik informatika.

Keywords : Datamining, Algoritma Apriori, registrasi, undur diri

I. PENDAHULUAN

1.1. Latar Belakang Masalah

Awal berdiri program studi Teknik Informatika yang disingkat TIF di Universitas Darma Persada mendapat animo yang baik dari masyarakat. Tahun 2002 penerimaan mahasiswa baru sebanyak 49 orang. Sudah delapan tahun teknik informatika berdiri, tentunya berharap program studi ini lebih maju hendaknya. Sampai saat ini prodi ini masih di minati oleh masyarakat, terbukti tahun 2009 penerimaan mahasiswa baru sebanyak 110 orang.

Kalau dihitung sejak awal berdiri hingga sampai saat ini tahun 2010, jumlah mahasiswa teknik informatika sebanyak 500 orang belum dikurangi dengan mahasiswa yang sudah lulus. Namun dalam proses belajar mengajar mahasiswa yang aktif hanya mencapai 250 orang. Kurang lebih 50 % mahasiswa yang telah mendaftar bahkan sudah mengikuti perkuliahan di program studi TIF ini mengundurkan diri tanpa diketahui penyebabnya.

Melihat referensi bagaimana suatu prodi berhasil dalam mengolah mahasiswa tentunya rasio mahasiswa yang mendaftar dengan yang lulus minimal 5 berbanding 1, artinya mahasiswa yang masuk 5 orang dan yang lulus hendaknya 1 orang.

Kontradiksi dengan ketentuan yang di atas. Dalam program studi tif mahasiswa yang tidak registrasi kurang lebih 50%, bukan mahasiswa yang lulus. Untuk itu perlu dilakukan suatu penelitian, bagaimana mendapatkan informasi dari masalah yang dihadapi oleh program studi TIF, sehingga diharapkan dengan penemuan pola atau kecenderungan dari 50% data yang mengundurkan diri dapat didefinisikan dengan jelas bagaimana bentuk informasi dari data mahasiswa yang tidak melakukan registrasi ulang alias yang mengundurkan diri tersebut.

Salah satu cara untuk melakukan analisis kemungkinan pengunduran diri seorang mahasiswa adalah dengan melakukan analisis asosiasi yang dikenal sebagai salah satu metode data mining yang menjadi dasar dari berbagai metode data mining lainnya. Aturan asosiasi memberikan informasi dalam bentuk hubungan if-then atau jika-maka. Aturan ini dihitung dari data yang sifatnya probabilistik. [5]

1.2 Perumusan Masalah

- Bagaimana mendapatkan pola pengunduran diri mahasiswa?
- Parameter apakah yang digunakan untuk menentukan pola pengunduran diri?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini mengimplementasikan algoritma apriori untuk mendapatkan pola pengunduran diri mahasiswa. Sebagai parameter tambahan yang dijadikan isian data kuesioner yaitu dari variable-variabel salah mengambil jurusan, biaya, lokasi, dosen, biaya, hubungan antar mahasiswa dan sarana perkuliahan yang dijadikan sebagai tolok ukur dari penyebaran informasi yang diolah. Hasil dari penelitian ini berupa aplikasi datamining untuk mengetahui pola pengunduran diri mahasiswa teknik informatika unsada.

II. TINJAUAN PUSTAKA

Datamining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data.[2]

Hubungan atau pola yang dicari dalam datamining dapat berupa pola antara dua atau lebih dalam satu dimensi. Sebagai contoh dalam dimensi produk kita dapat melihat keterkaitan pembelian suatu produk dengan produk yang lain. Selain itu, hubungan dapat dilihat antara dua atau lebih atribut dan dua atau lebih objek [3].

Datamining disebut juga *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran datamining dapat digunakan untuk memperbaiki pengambilan keputusan di masa depan. [5]

Pengelompokan datamining berdasarkan tugas yang dapat dilakukan antara lain: deskripsi, estimasi, prediksi, klasifikasi, pengklusteran, dan asosiasi [2].

Algoritma untuk menemukan *frequent itemset* yang menjadi dasar pembacaan data dalam database adalah sebagai berikut: [4]

```

foreach item,
  ckcheck if it is afrequent itemset
  k=1
repeat
    foreach new frequent itemset  $I_k$  with  $k$  items
      Generate all itemset  $I_{k+1}$  with  $k+1$  items,  $I_k \subset I_{k+1}$ 
    Scan all transaction once and check if
    The generated  $k+1$  itemset are frequent
     $k = k+1$ 
until no new frequent itemsets are identified

```

Algoritma ini membagi beberapa tahap yang disebut iterasi. Tiap iterasi menghasilkan pola frekuensi tinggi dengan panjang yang sama dimulai dari pass pertama yang menghasilkan pola frekuensi tinggi dengan panjang satu. Di iterasi pertama ini, support dari setiap item dihitung dengan membaca data.

Algoritma apriori termasuk jenis aturan asosiasi pada datamining. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut affinity analysis atau market basket analysis. Analisis asosiasi adalah teknik datamining untuk menemukan aturan asosiatif antara satu kombinasi item. Dalam asosiasi terdapat istilah *antecedent* dan *consequent*, *antecedent* untuk mewakili bagian “jika” dan *consequent* untuk mewakili bagian “maka”. Dalam analisis ini, *antecedent* dan *consequent* adalah sekelompok item yang tidak punya hubungan secara bersama [5].

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap, yaitu tahap analisis pola frekuensi tinggi dan tahap pembentukan aturan asosiasi. [2]

Tahap analisis frekuensi tinggi dengan mencari item yang memenuhi syarat *minimum* dari nilai *support* dalam data. Nilai *support* dari sebuah item diperoleh dengan rumus:

$$Support(A) = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total transaksi}} \dots\dots\dots(1)$$

Rumus untuk *support* dari 2 item yaitu :

$$Support (A,B) = P (A \cap B)$$

$$Support (A,B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \dots\dots\dots(2)$$

Tahap pembentukan aturan asosiasi, frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat *minimum* untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus berikut:

$$Confidence = P(B| A) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi mengandung A}} \dots\dots\dots(3)$$

III. METODE PENELITIAN

Metode yang digunakan dalam pengumpulan data sebagai berikut:

1. Metode wawancara

Melakukan wawancara dengan pihak yang berkaitan langsung dengan permasalahan yang dibahas. Metode ini dilakukan dengan panduan daftar pertanyaan yang telah disiapkan.

IV. HASIL DAN PEMBAHASAN

Salah satu cara untuk melakukan analisis kemungkinan pengunduran diri seorang mahasiswa adalah dengan melakukan analisis asosiasi yang dikenal sebagai salah satu metode data mining yang menjadi dasar dari berbagai metode data mining lainnya.

Analisis asosiasi dikenal sebagai algoritma apriori. Untuk menemukan pola frekuensi tinggi digunakan algoritma apriori. Pola frekuensi tinggi adalah pola-pola item di dalam suatu data yang memiliki frekuensi atau support di atas ambang batas tertentu yang disebut dengan istilah *minimum support* atau *threshold*. *Threshold* adalah batas *minimum* transaksi. Jika jumlah transaksi kurang dari *threshold* maka item atau kombinasi item tidak akan diikutkan pada perhitungan selanjutnya. Penggunaan *threshold* juga dapat mempercepat perhitungan.

Tabel 1 pada lampiran menjelaskan analisis pembentukan kategori yang diambil dari dari database akademik yang ada. Penentuan kategori ini digunakan berdasarkan mahasiswa ada yang hanya aktif semester satu, semester dua, semester tiga, semester empat, semester lima, semester enam, lebih dari semester enam. Dari data pada tabel 3 (lampiran) didapat kandidat Pertama. Ditetapkan *threshold*= 3. Maka kandidat dengan nilai kurang dari 3 akan dieliminasi. Dari tabel 5 pada lampiran dapat dihitung dengan rumus *support* dan *confidence* [2] dan [3]

$$Support S1, Ya= 4 / 10$$

Support S2,Tidak = 3 / 10

Confidence S1, Ya = 4 / 5

Confidence S2, Tidak = 3 / 3

Oleh karena itu dapat disimpulkan bahwa alasan memutuskan kuliah pada semester ke 2 (S2) “bukan pada pola salah mengambil jurusan.”

Pada gambar 1 berikut ini merupakan tampilan awal aplikasi datamining. Ada beberapa fasilitas pada aplikasi ini antara lain: proses mining, sinkronisasi data, *cleaning* data, pengelolaan *user management*.



Gambar 1. Tampilan awal aplikasi

Gambar 2 berikut merupakan tampilan aplikasi untuk pengelolaan user. Bagi pengguna aplikasi ini wajib memasukkan user name dan password. Pada tampilan ini dapat juga untuk mengubah password yang sudah ada.

Gambar 2. Tampilan user managemen

Gambar 3. berikut merupakan form yang disediakan untuk memasukkan data atribut hasil kuesioner yang didapat.



NIM :

Jenis Kelamin :

Tempat / Tgl Lahir :

Alamat Sekarang :

Alamat Email :

No Telp / No HP : /

Asal Sekolah SMU / SMK :

Mengenal Unsada dari :

Masuk Kuliah Tahun :

Berhenti Kuliah Tahun : /

Gambar 3. Input Data Kuesioner

V. KESIMPULAN

Penelitian ini memberikan kesimpulan bahwa aplikasi datamining dapat digunakan untuk menampilkan informasi pola pengunduran diri mahasiswa teknik informatika unsada. Berdasarkan informasi pada pembahasan sebelumnya mahasiswa yang mengundurkan diri atau memutuskan kuliah pada semester ke 2 (S2) “bukan pada pola salah mengambil jurusan”. Aplikasi datamining ini dapat digunakan sebagai kajian untuk mempertimbangkan keputusan yang akan datang.

VI. SARAN

Untuk pengembangan lebih lanjut aplikasi datamining dapat diarahkan untuk mencari pola atribut-atribut lainnya yang ada dalam database akademik.

VII. DAFTAR PUSTAKA

- [1] Connolly Thomas. (2005) *“Database Systems: A Practical Approach to Design, Implementation, and Management”*, University of Paisley, Paisley, Addison Wesley.
- [2] Kusriani, dan Emha Taufik Luthfi, 2009, *“Algoritma Data Mining”*, Penerbit Andi, Yogyakarta.
- [3] Poniah, P, 2001, *“Datawarehouse Fundamentals: A Comprehensive Guide for IT Profesioanal”*, John Willey & Sons, Inc.
- [4] Ramakrisnan, Ragu, 2004, *“Sistem Manajemen Database”*, Edisi 3, Andi Offset, Yogyakarta.
- [5] Santosa, Budi, 2007, *“Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis”*, Graha Ilmu, Yogyakarta

VII. LAMPIRAN

Tabel 1. Kategori

Kategori	Keterangan
S1	Mahasiswa Aktif sampai semester satu
S2	Mahasiswa Aktif sampai semester dua
S3	Mahasiswa Aktif sampai semester tiga
S4	Mahasiswa Aktif sampai semester empat
S5	Mahasiswa Aktif sampai semester lima
S6	Mahasiswa Aktif sampai semester enam
S7	Mahasiswa Aktif sampai semester enam lebih

Tabel 2. Hubungan Kategori dengan Data Kuesioner

NIM	Kategori	Salah Jurusan
2010230018	S1	Ya
2010230054	S1	ya
2010230024	S1	Ya
2010230105	S1	Tidak
2010230084	S1	Ya
2005230021	S4	Tidak
2005230018	S5	Tidak
2010230027	S2	Tidak
2005230901	S2	Tidak
2010230001	S2	Tidak

Tabel 3. Hasil Eliminasi

Itemset	Jumlah
S1	5
S2	3
Ya	4
Tidak	6

Tabel 4. Kandidat kedua

Itemset	Jumlah
S1, Ya	4
S1, Tidak	1
S2, Ya	0
S2, Tidak	3

Tabel 5. Tabel Setelah Dieliminasi *Threshold* = 3

Itemset	Jumlah
S1, Ya	4
S2, Tidak	3