

ANALISIS DATA PENDIDIKAN TINGGI MENGUNAKAN PENDEKATAN DATA MINING

Gita Indah Marthasari¹⁾

¹⁾ProdiTeknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Malang

Jl. Raya Tlogomas 246, Malang

E-mail : gita@umm.ac.id

ABSTRAK

Data mining untuk pendidikan menekankan pada pemanfaatan metode untuk menemukan pengetahuan dari data di lingkungan pendidikan. Bidang ini menggunakan mekanisme transformasi atau inovasi dari pendekatan yang diturunkan dari ilmu statistik, pembelajaran mesin, psikometrik, dan komputasi ilmiah. Penelitian ini mengusulkan sebuah analisis terhadap status keaktifan siswa menggunakan salah satu teknik data mining yaitu *association rule mining* (ARM). Teknik ARM bertujuan menemukan pola-pola yang merepresentasikan informasi bernilai tinggi bagi para pengambil keputusan di perguruan tinggi. Salah satu algoritma yang umum digunakan dalam ARM adalah Apriori. Algoritma Apriori digunakan untuk mencari aturan-aturan asosiasi yang menarik dari basis data dalam rangka mengekstraksi pengetahuan dari data profil dan data akademik siswa. Aturan-aturan yang ditemukan selanjutnya dianalisis sebagai bahan rekomendasi bagi para pengelola akademik untuk meningkatkan kualitas proses pengambilan keputusan. Selain itu, hasil analisis dapat menjadi acuan bagi arah kurikulum yang mampu memperbaiki kualitas pembelajaran siswa. Uji coba dilakukan dengan melakukan pengaturan terhadap parameter nilai minimal *support* dan minimal *confidence*. Berdasarkan hasil analisis diperoleh pengetahuan antara lain adanya hubungan kuat antara asal sekolah mahasiswa dan pekerjaan orang tua terhadap tingkat keaktifan. Selain itu, juga diperoleh pengetahuan tentang nilai minimal mahasiswa tiap semesternya untuk tetap aktif di semester selanjutnya.

Kata kunci : Data mining untuk pendidikan, *Association Rule Mining*, Algoritma apriori, Status keaktifan siswa.

ABSTRACT

Educational data mining (EDM) concerns with developing methods for discovering knowledge from data that come from educational environments. EDM requires a transformation of existing or innovation of new approaches derived from statistics, machine learning, psychometrics, and scientific computing. In this paper, we propose an analysis of student activity status using one of data mining techniques that are association rule mining (ARM). ARM technique aims at discovering patterns that can provide valuable information for the decision maker in higher education. Apriori algorithm is one of many methods in ARM. We have used Apriori algorithm for finding interesting association rules from the transformed database which can be useful to extract knowledge of students' profile and academic evaluation. The identified rules are analyzed to offer a helpful and constructive recommendation to the academic planners in higher institutions to enhance their decision-making process. We analyze the data by setting minimal support and minimal confidence value. Based on the experiment, we conclude that the student high school location and parent's job have a strong correlation with student activity. Moreover, we also acquired a knowledge about student minimal grade point average in a semester to remain active in the next semester.

Keywords: Educational data mining, Association rule mining, Apriori algorithm, Students activity status

PENDAHULUAN

Perguruan tinggi merupakan salah satu elemen penting dalam komponen pendidikan. Sampai dengan tahun 2016, terdapat lebih dari 3500 perguruan tinggi aktif di Indonesia dengan jumlah mahasiswa lebih dari 4 juta orang (ristekdikti.go.id). Semakin besarnya jumlah universitas menunjukkan peningkatan dalam persaingan jasa pendidikan. Hal ini menuntut para pengelola perguruan tinggi untuk makin kompetitif dan mampu memberikan pelayanan dan pengelolaan yang profesional. Analitis akademis merupakan bidang yang sedang berkembang dalam dunia pendidikan tinggi. Bidang ini memanfaatkan penerapan *data mining* dan *business intelligence* baik untuk kepentingan operasional universitas maupun proses pembelajaran [1]. Proses analisis terhadap data akademik dimungkinkan dengan meningkatnya volume data seiring kemampuan universitas dalam menangkap dan menyimpan data. Pengetahuan yang diperoleh dari data akan menjadi bahan pertimbangan dalam pengambilan keputusan oleh pengelola universitas, memperbaiki kinerja siswa, menurunkan tingkat kegagalan studi, mempelajari perilaku siswa, dan perbaikan proses pembelajaran.

Analitik terhadap data akademik termasuk salah satu topik dalam data mining untuk pendidikan. Data mining dalam pendidikan adalah sebuah bidang yang memanfaatkan informasi statistik, algoritma pembelajaran mesin dan *data mining* untuk menemukan pola menarik dari data[2]. *Data mining* digunakan dalam proses analitik karena kemampuannya dalam mengekstraksi pengetahuan dari data. Kemampuan *data mining* dibedakan menjadi dua yaitu prediktif dan deskriptif [3]. Prediktif adalah kemampuan untuk memperkirakan nilai luaran berdasarkan nilai masukan tertentu dalam situasi acak, sedangkan deskriptif adalah kemampuan mengekstrak pola dan menemukan tren dari data. *Data mining* dianggap sebagai

teknologi yang paling sesuai untuk memberikan tambahan wawasan kepada para pengajar, siswa, alumni, pengelola, dan staf akademik lainnya dan mampu membantu dalam menjadi alat bantu otomatis pada proses pengambilan keputusan. Penggunaan *data mining* dalam pendidikan ditentukan oleh lingkup persoalan dan hasil analisis yang ingin diperoleh, antara lain: untuk mempelajari perilaku siswa [2], penyusunan kurikulum pembelajaran pada pendidikan tinggi [4], menemukan cara untuk meningkatkan retensi siswa[5], dan menganalisa hasil evaluasi siswa dalam sistem e-learning berbasis web [6]. Namun, belum banyak penelitian yang menerapkan *data mining* dalam rangka mempelajari faktor-faktor yang mempengaruhi status keaktifan siswa.

Teknik data mining telah banyak diterapkan pada bidang akademik. Dalam penelitiannya, Ahmed [8] menggunakan teknik penggalian aturan asosiasi untuk menganalisis variabel-variabel dari data akademik dan pribadi dari mahasiswa yang telah lulus dari jurusan Computer Science and Engineering. Adapun proses yang dilakukan adalah merubah data dari basis data relasional ke dalam format yang siap dianalisis. Beberapa aturan menarik berhasil diidentifikasi antara lain pengetahuan tentang pengaruh data-data pribadi mahasiswa seperti gender, tempat tinggal, dan lain-lain, pengetahuan yang berkaitan dengan hasil belajar mahasiswa, serta faktor-faktor yang mempengaruhi keaktifan mahasiswa.

El-Haales [9] menggunakan beberapa metode data mining untuk menganalisis data antara lain association rule mining, decision tree, algoritma clustering Expectation-Maximization (EM), dan analisis outlier. Data yang dianalisis adalah data akademik dan data dari sistem e-learning. Teknik association rule digunakan untuk menemukan faktor-faktor yang mempengaruhi nilai akhir mahasiswa. Decision tree digunakan untuk membentuk aturan logika yang dapat mengklasifikasikan nilai akhir mahasiswa. Sedangkan algoritma EM

digunakan dalam proses klasterisasi mahasiswa berdasarkan beberapa variabel.

Manvar [10] menerapkan teknik data mining untuk mempelajari profil siswa. Algoritma yang digunakan adalah k-Means dan Apriori. Algoritma Apriori digunakan untuk memperkirakan hasil ujian akhir. Algoritma k-means digunakan untuk membagi siswa ke dalam kelompok-kelompok berdasarkan hasil ujian akhirnya.

Teknik data mining juga diterapkan oleh Acharya [11] dalam proses penempatan siswa. Analisis penempatan siswa menggunakan metode *association rule mining* (ARM) terhadap data akademik siswa. Beberapa proses diterapkan sebelum analisis dengan ARM dilakukan yaitu pemfilteran, homogenisasi, generalisasi, dan integrasi. Algoritma Apriori digunakan dalam proses analisis untuk menemukan aturan-aturan yang menarik. Aturan menarik ditemukan dengan mengubah parameter nilai *support* dan *confidence*.

Metode ARM juga diterapkan untuk menganalisis kinerja siswa oleh Kumar [6]. Analisis diterapkan pada data hasil evaluasi siswa pada perkuliahan. Tujuannya adalah membandingkan kinerja siswa di level graduate dan post graduate untuk perkuliahan yang sama. Proses-proses yang dilakukan sebelum analisis dengan ARM (pre-processing) yaitu pembersihan dan transformasi data. Berdasarkan pengujian, akurasi hasil analisis meningkat setelah dilakukan pre-processing data.

Penelitian ini menggunakan data akademik mahasiswa Teknik Informatika Universitas Muhammadiyah Malang meliputi data pribadi dan data hasil evaluasi selama 4 semester. Data yang diambil adalah periode 2012 hingga 2015. Artikel ini menunjukkan bagaimana proses pengambilan data, pra-pemrosesan data, analisis dan pengetahuan yang dapat diambil dari hasil analisis. Metode *association rule mining* digunakan untuk menganalisis data. Aplikasi WEKA sebagai alat untuk pengolahan data[7].

METODE

Data diperoleh dari Biro Administrasi Akademik Universitas Muhammadiyah Malang meliputi data pribadi dan akademik mahasiswa Jurusan Teknik Informatika sebanyak 538 data. Adapun deskripsi data ditunjukkan pada Tabel 1, sedangkan penjelasan tiap variabel data dapat dilihat pada Tabel 2. Langkah selanjutnya adalah melakukan pra-proses (*pre-processing*) terhadap data yang diperoleh. Pra-proses bertujuan mempersiapkan data untuk dianalisis guna menghasilkan pengetahuan yang diharapkan. Pada penelitian ini, pra-proses yang dilakukan melalui langkah pembersihan data untuk menghapus tuple-tuple yang tidak relevan dan transformasi data untuk menyesuaikan format data yang diperoleh dengan format data untuk analisis.

Tabel 1 Deskripsi data

Jenis Data	Item Data
Data pribadi	jenis kelamin, usia saat mendaftar, kota asal, status domisili, agama, marital, asal sekolah, status kerja, asal biaya, pekerjaan orang tua
Data akademik	IP semester, jumlah SKS per semester, status keaktifan

Tabel 2 Deskripsi tiap variabel pada data

Atribut	Jenis	Keterangan
Jenis kelamin	Diskrit	Jenis kelamin mahasiswa
Usia mendaftar	Numerik	Usia mahasiswa saat masuk universitas
Kota asal	Diskrit	Kota tempat tinggal sebelum kuliah
Agama	Diskrit	Agama mahasiswa
Status marital	Diskrit	Status perkawinan mahasiswa saat mendaftar kuliah
Asal sekolah	Diskrit	Jenis SMU/ sederajat
Status pekerjaan	Diskrit	Status pekerjaan mahasiswa saat mendaftar kuliah
Asal biaya	Diskrit	Menunjukkan asal biaya kuliah mahasiswa

Pekerjaan orang tua	Diskrit	Pekerjaan orang tua
IP semester	Numerik	Nilai indeks prestasi semester 1 sampai dengan 4
SKS per semester	Numerik	Jumlah SKS yang diambil pada semester 1 sampai dengan 4
Status akademik	Diskrit	Status akademik mahasiswa pada semester genap 2015/2016

Transformasi data antara lain bertujuan untuk mengubah beberapa variabel yang nilainya numerik menjadi nilai diskrit. Teknik yang digunakan adalah teknik *equal width* yaitu mengelompokkan data dalam rentang dengan lebar seragam. Variabel-variabel yang diubah menjadi diskrit dan aturan perubahannya ditunjukkan pada Tabel 3.

Tabel 3 Penjelasan transformasi variabel

Variabel	Aturan Transformasi
Usia mendaftar	Sesuai : 17 sampai dengan 21 tahun Kurang sesuai : <17 dan >21 tahun
IP semester	Sangat Rendah : 0.0 sampai 1.0 Rendah : 1.1 sampai 2.0 Cukup : 2.1 sampai 3.0 Baik : 3.1 sampai 4.0
SKS per semester	Banyak : >18 Sedikit : <=18

Langkah berikutnya adalah mengeskrak aturan asosiasi menggunakan metode *Association Rule Mining* (ARM). ARM merupakan sebuah metode untuk menemukan aturan-aturan asosiasi yang memenuhi nilai minimal *support* dan minimal *confidence*[12]. Mekanisme ARM terbagi menjadi dua : (1) menemukan himpunan item yang kemunculannya memenuhi nilai support tertentu dimana himpunan item ini selanjutnya disebut sebagai frequent item set, dan (2) membangkitkan aturan-aturan asosiasi dari seluruh himpunan frequent item set yang memenuhi nilai confidence tertentu. Algoritma yang paling populer digunakan adalah Apriori.

Apriori adalah algoritma yang diusulkan oleh R. Agarwal dan R. Srikant pada tahun 1994 [6]. Penamaan ini berkaitan dengan cara kerja dari Apriori yang sesuai dengan namanya, 'prior' yang artinya 'awal', memanfaatkan pengetahuan awal dari himpunan frequent item. Berikut adalah langkah-langkah Apriori, yaitu : Jika C_k adalah himpunan kandidat item berukuran k dan L_k adalah himpunan frequent item berukuran k , maka proses iterasi yang berlangsung adalah :

- Cari frequent set L_{k-1}
- Langkah penggabungan : C_k dibangkitkan dengan menggabungkan L_{k-1} dengan dirinya sendiri ($L_{k-1} \times L_{k-1}$)
- Langkah pemangkasan (properti Apriori) : setiap himpunan item berukuran $(k-1)$ yang kemunculannya tidak sering tidak dapat menjadi himpunan bagian dari himpunan item berukuran k yang kemunculannya sering, maka harus dihapus
- Frequent set L_k telah dihasilkan.

HASIL DAN PEMBAHASAN

Analisis menggunakan Algoritma Apriori dilakukan melalui aplikasi WEKA. Sebelum diolah, data yang telah mengalami pra-proses diubah ke dalam format atribut-relasi. Proses uji coba dilakukan dengan mengatur nilai minimal support dan nilai minimal confidence.

Uji Coba dengan Pengaturan Nilai Minimal Support dan Nilai Minimal Confidence

Tujuan pengujian adalah menganalisis aturan-aturan yang dihasilkan untuk nilai support dan confidence tertentu. Variabel yang dianalisis adalah jenis kelamin, asal sekolah, kota asal, pekerjaan orang tua, IP1, IP2, IP3, IP4, SKS1, SKS2, SKS3, SKS4, dan status akademik. Nilai confidence yang digunakan = 0.9. Salah satu kriteria aturan disebut penting adalah jika memiliki nilai support dan confidence tinggi. Berdasarkan pengujian, nilai tertinggi yang dapat diberikan pada variabel support adalah 0.6 dan nilai confidence tertinggi adalah

0.9 Aturan yang dihasilkan menggunakan kedua nilai variabel ditunjukkan pada Tabel 4. Meski demikian, dari ke-8 aturan yang diperoleh, tidak ada pengetahuan yang dianggap menarik dan penting. Variabel SKS1 dan SKS2 dianggap tidak berpengaruh penting karena nilainya sama untuk sebagian besar mahasiswa.

Tabel 4 Uji Coba Nilai Minimal Support dan Minimal Confidence

No	Aturan
1	SKS2=BESAR 447 ==> SKS1=BESAR 447 conf:(1)
2	SKS2=BESAR Status Akademik=A 393 ==> SKS1=BESAR 393 conf:(1)
3	Jenis =Laki-laki SKS2=BESAR 359 ==> SKS1=BESAR 359 conf:(1)
4	SKS1=BESAR Status Akademik=A 400 ==> SKS2=BESAR 393 conf:(0.98)
5	Status Akademik=A 409 ==> SKS1=BESAR 400 conf:(0.98)
6	Jenis =Laki-laki 428 ==> SKS1=BESAR 416 conf:(0.97)
7	Status Akademik=A 409 ==> SKS2=BESAR 393 conf:(0.96)
8	Status Akademik=A 409 ==> SKS1=BESAR SKS2=BESAR 393 conf:(0.96)

Analisis variabel Status Akademik terhadap Jenis Kelamin

Di Jurusan Teknik Informatika, hampir 80% mahasiswa adalah laki-laki. Hal ini menunjukkan masih adanya paradigma bahwa jurusan teknik lebih sesuai untuk laki-laki. Analisis bertujuan untuk mengetahui hubungan antara jenis kelamin terhadap status akademik. Nilai minimum support yang digunakan adalah 0.1. Dari hasil uji coba, diperoleh aturan seperti pada Tabel 5. Aturan-aturan pada tabel tersebut menunjukkan bahwa mahasiswa laki-laki lebih berpotensi untuk non-aktif dibandingkan mahasiswa perempuan yang dinyatakan dengan nilai *confidence* yang tinggi. Sedangkan pada aturan yang lain tampak bahwa mahasiswa perempuan memiliki potensi yang tinggi untuk aktif.

Tabel 5 Hasil analisis terhadap variabel Jenis Kelamin

Aturan	Confidence
Status Akademik=N ==> Jenis =Laki-laki	0.83
Jenis =Perempuan	0.8
Status Akademik=A Status Akademik=A ==> Jenis =Laki-laki	0.79
Jenis =Laki-laki	0.75
Status Akademik=A Jenis =Laki-laki ==>	0.25
Status Akademik=N Status Akademik=A ==> Jenis =Perempuan	0.21

Analisis variabel Status Akademik terhadap variabel Usia Masuk

Tujuan analisis adalah mengetahui pengaruh usia masuk terhadap status akademik. Nilai minimal *support* yang digunakan adalah 0.1. Aturan yang dihasilkan dapat dilihat pada Tabel 6. Berdasarkan hasil tersebut dapat disimpulkan bahwa usia masuk tidak berpengaruh terhadap status keaktifan mahasiswa yang ditunjukkan dengan nilai *confidence* yang cukup tinggi pada aturan 1 dan 2.

Tabel 6 Hasil analisis terhadap Usia Masuk

Aturan	Confidence
Status Akademik=A ==> Usia mendaftar=SESUAI	0.96
Status Akademik=N ==> Usia mendaftar=SESUAI	0.94
==> Status Akademik=A Usia mendaftar=SESUAI	0.77
==> Status Akademik=N Usia mendaftar=SESUAI	0.23
Status Akademik=A ==> Jenis =Perempuan	0.21

Analisis variabel Status Akademik terhadap Asal Sekolah

Variabel Asal Sekolah menunjukkan jenis sekolah setara SMA. Penyebaran data untuk tiap jenis sekolah yaitu asal SMA sebanyak 315 orang, MA sebanyak 45 orang, SMK 173 orang, Pondok Pesantren sebanyak 1 orang, dan Kejar Paket sebanyak 1 orang. Hasil analisis ditunjukkan dalam Tabel 7. Berdasarkan pengujian, mahasiswa yang

berasal dari SMK cenderung untuk berstatus aktif di semester ke-5 yang ditunjukkan dengan *confidence* sebesar 0.82. Potensi mahasiswa non-aktif cenderung dimiliki oleh mahasiswa yang berasal dari SMA dengan nilai *confidence* sebesar 0.63.

Tabel 7 Hasil analisis terhadap variabel Asal Sekolah

Aturan	Confidence
Asal Sekolah=SMK Status Akademik=A	==> 0.82
Asal Sekolah=SMA Status Akademik=A	==> 0.75
Asal Sekolah=SMA Status Akademik=N	==> 0.63
Asal Sekolah=SMA Status Akademik=A	==> 0.57
Asal Sekolah=SMA Status Akademik=A	==> 0.34
Asal Sekolah=SMK Status Akademik=N	==> 0.25

Analisis variabel Status Akademik terhadap Pekerjaan Ortu

Berdasarkan data yang diperoleh, terdapat beberapa jenis pekerjaan orang tua mahasiswa yaitu PNS (161 orang), Pensiunan PNS (17 orang), Pensiunan TNI/POLRI (7 orang), wiraswasta (155 orang), pegawai swasta (97 orang), TNI/POLRI (14 orang), guru (45 orang), petani (39 orang), dan tidak bekerja (1 orang). Hasil analisis dapat dilihat pada Tabel 8. Dari hasil tersebut, mahasiswa yang berstatus aktif memiliki hubungan dengan orang tua yang bekerja sebagai pegawai swasta (dengan *confidence* 0.79), PNS (dengan *confidence* 0.76), dan wiraswasta (dengan *confidence* 0.75).

Tabel 8 Hasil analisis terhadap variabel Pekerjaan Ortu

Aturan	Confidence
Pekerjaan Ortu=Pegawai Swasta Status Akademik=A	==> 0.79
Pekerjaan Ortu=Pegawai Negeri Sipil Status Akademik=A	==> 0.76
Pekerjaan Ortu=Wiraswasta Status Akademik=A	==> 0.75

Analisis variabel Status Akademik terhadap variabel Indeks Prestasi

Penyebaran nilai Indeks Prestasi (IP) semester 1 sampai dengan 4 dapat dilihat pada Tabel 9. Adapun hasil analisis IP per semester terhadap variabel Status Akademik dapat dilihat pada Tabel 10 sampai dengan Tabel 12. Pada Tabel 10, terdapat aturan yang menghubungkan antara IP semester 1 dengan kategori Sangat Kurang terhadap status non-aktif dengan nilai *confidence* 0.95. Adapun pada hasil analisis terhadap nilai IP semester 2 sampai dengan 4 tidak berhasil menemukan aturan yang mengandung status non-aktif.

Tabel 9 Penyebaran nilai IP semester 1 sampai dengan 4

Kategori	IP1	IP2	IP3	IP4
Sangat Kurang (SK)	78	33	33	50
Kurang (K)	33	33	52	28
Cukup (C)	139	132	110	82
Baik (B)	280	261	245	262

Tabel 10 Hasil analisis terhadap variabel Indeks Prestasi 1

Aturan	Confidence
IP1=B Status Akademik=A	==> 0.95
IP1=SK Status Akademik=N	==> 0.95
IP1=C Status Akademik=A	==> 0.86

Tabel 11 Hasil analisis terhadap variabel Indeks Prestasi 2

Aturan	Confidence
IP2=B Status Akademik=A	==> 0.98
IP2=C Status Akademik=A	==> 0.86

Tabel 12 Hasil analisis terhadap variabel Indeks Prestasi 2

Aturan	Confidence
IP3=B Status Akademik=A	==> 0.98
IP3=C Status Akademik=A	==> 0.96

Tabel 13 Hasil analisis terhadap variabel Indeks Prestasi 3

Aturan		Confidence
IP4=C Akademik=A	==> Status	1
IP4=B Akademik=A	==> Status	0.99

SIMPULAN

Data mining telah digunakan banyak digunakan pada analisis data pendidikan tinggi. Penelitian ini telah menerapkan teknik Association Rule Mining (ARM) untuk menemukan faktor-faktor yang mempengaruhi status aktif mahasiswa. Analisis dilakukan terhadap data pribadi dan akademik mahasiswa Jurusan Teknik Informatika Universitas Muhammadiyah Malang.

Dari analisis diperoleh pengetahuan antara lain adanya hubungan kuat antara asal sekolah mahasiswa dan pekerjaan orang tua terhadap tingkat keaktifan. Selain itu, juga diperoleh pengetahuan tentang nilai minimal mahasiswa tiap semesternya untuk tetap aktif di semester selanjutnya. Hasil analisis diharapkan dapat memberikan dasar bagi pihak pengelola dalam merumuskan kebijakan dan arah pendidikan. Untuk selanjutnya, penelitian dapat dikembangkan dengan menganalisis data dari jurusan lain untuk mengetahui ada atau tidaknya kesamaan pola dan juga dapat dikembangkan untuk menganalisis faktor-faktor yang mempengaruhi rendahnya indeks prestasi mahasiswa.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada Universitas Muhammadiyah Malang yang telah mendanai penelitian ini.

DAFTAR PUSTAKA

[1] P. Baepler and C. J. Murdoch, "Academic Analytics and Data Mining in Higher Education," *Int. J. Scholarsh. Teach. Learn.*, vol. 4, no. 2, pp. 1–9, 2010.

[2] G. Elakia, N. J. Aarathi, Elakia, Gayathri, Aarathi, and Naren J, "Application of Data Mining in Educational Database for Predicting Behavioural Patterns of the Students," *Elakia al. (IJCSIT) Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 3, pp. 4649–4652, 2014.

[3] D. A. Alhammadi and M. S. Aksoy, "Data Mining in Education- An Experimental Study," *Int. J. Comput. Appl.*, vol. 62, no. 15, pp. 31–34, 2013.

[4] K. Chandra, E., Nandhini and E. Chandra, "Knowledge mining from student data," *Eur. J. Sci. Res.*, vol. 47, no. 1, pp. 156–163, 2010.

[5] E. J. M. Lauría, J. D. Baron, M. Devireddy, V. Sundararaju, and S. M. Jayaprakash, "Mining academic data to improve college student retention : An open source perspective," *Proc. 2nd Int. Conf. Learn. Anal. Knowl. - LAK '12*, no. May, pp. 139–142, 2012.

[6] V. Kumar and A. Chadha, "Mining association rules in student's assessment data," *Int. J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 211–216, 2012.

[7] S. S. Aksenova, "Machine Learning with WEKA," pp. 2–4, 2004.

[8] S. Ahmed, R. Paul, A. Sayed, and L. Hoque, "Knowledge Discovery from Academic Data using Association Rule Mining," no. December, pp. 22–23, 2014.

[9] A. El-Halees, "Mining Students Data To Analyze Learning Behavior : a Case Study Educational Systems," *Work*, no. February, 2008.

[10] M. Manvar and M. Rao, "Predicting students performance in higher education: A Data Mining Approach," *Ijser.Org*, vol. 5, no. 2, pp. 1024–1027, 2014.

[11] S. Acharya and N. Madhu, "Discovery of students ' academic patterns using data mining techniques," *Int. J. Comput. Sci.*

- Eng.*, vol. 4, no. 6, pp. 1054–1062, 2012.
- [12] S. Kotsiantis and D. Kanellopoulos, “Association Rules Mining: A Recent Overview,” *GESTS Int. Trans. Comput. Sci. Eng.*, vol. 32, no. 1, pp. 71–82, 2006.