

Penerapan Data Mining Untuk Menggali Informasi Tersembunyi Dalam *Big Data* Nilai Mata Kuliah

Elvira Asril¹, Fana Wiza², Taslim³

^{1,2}Prodi Sistem Informasi Fakultas Ilmu Komputer Universitas Lancang Kuning

³Prodi Teknik Informatika Fakultas Ilmu Komputer Universitas Lancang Kuning

(Jl. Yos Sudarso KM. 8 Rumbai, Pekanbaru, Riau, telp. 0811 753 2015

e-mail: ¹elvira@unilak.ac.id, ²fanawiza@unilak.ac.id, ³taslimmalano@unilak.ac.id

Abstrak

Jarang sekali perguruan tinggi melihat kompetensi lulusannya sebelum dilepas ke dunia nyata. Salah satu variabel yang bisa digunakan adalah nilai matakuliah yang telah diperoleh mahasiswa atau calon lulusan. Kemudian memetakan nilai matakuliah yang telah diperoleh tiap mahasiswa atau calon lulusan pada aspek kompetensi dasar lulusan Strata satu Informatika yang disusun oleh asosiasi perguruan tinggi komputer (APTIKOM) dengan menggunakan teknik data mining. Pemetaan dilakukan berdasarkan nilai matakuliah yang telah ditempuh oleh mahasiswa atau calon lulusan, dalam hal ini objek penelitian adalah mahasiswa angkatan 2012 s/d 2015 yang telah mencapai 120 sks. Daftar aspek kompetensi dasar yang digunakan adalah aspek kompetensi yang disusun oleh APTIKOM berdasarkan ACM/IEEE 2013. Kemudian dilakukan penentuan kelompok matakuliah pada tiap kompetensi tersebut. Topik-topik yang dikaji antara lain meliputi : database, data mining, association rule, apriori dan beberapa algoritma lain yang mungkin dapat digunakan, serta perangkat lunak yang digunakan untuk proses mining. Pengolahan data yang telah disiapkan menggunakan beberapa perangkat lunak bantu seperti Excel, dan Tanagra. Mining data yang telah dilakukan menghasilkan informasi mengenai kompetensi dari calon lulusan yang dapat digunakan sebagai bahan analisa untuk pengambilan keputusan.

Kata kunci : kompetensi, informasi, nilai mata kuliah

Abstract

Rarely college graduates look competence before being released into the real world. One of the variables that can be used is the value of the courses that have been acquired or prospective graduate students. Then mapping the value of the courses that have been taken by each student or graduate candidates on the basis of competence of graduates Strata aspects of the Information compiled by the association of colleges computer (APTIKOM) using data mining techniques. Mapping is done based on the value of the courses that have been taken by students or prospective graduates, in this case the object of study is the student of 2012 s / d in 2015 which has achieved 120 credits. List aspects of basic competencies that are used are compiled by the competence aspect APTIKOM based ACM / IEEE 2013. Then is the determination of subjects in each group that competency. Topics to be studied include: databases, data mining, association rule, a priori and some other algorithm that may be used, as well as the software used to process mining. Processing of the data which has been prepared using some assistive software such as Excel, and Tanagra. Data mining has been done to produce information concerning the competence of prospective graduates who can be used as material analysis for decision making.

Keywords: competence, information, mark

1. Pendahuluan

Pembentukan kurikulum berbasis KKNI sangat bergantung pada kemampuan tiap perguruan tinggi dalam memahami kompetensi dari mahasiswanya. Secara umum perguruan tinggi akan berpedoman pada visi dan misinya, dimana visi dan misi disusun berdasarkan kebutuhan dunia usaha dan industri atas lulusan jangka panjang dengan memasukkan ciri khas lembaganya. Di Universitas Lancang Kuning yang menjadi ciri khas adalah penambahan sisi

nilai-nilai budaya melayu. Untuk menelaah hal tersebut, di perguruan tinggi yang seharusnya hal ini menjadi perhatian. Jangan hanya mementingkan dapat menghasilkan lulusan dengan Indeks Prestasi Kumulatif (IPK) yang sesuai dengan kebutuhan dunia usaha dan industri tanpa melihat kompetensi yang dimiliki oleh lulusan tersebut.

Penelitian tentang pengolahan data nilai mahasiswa yang dihubungkan dengan pemetaan kompetensi mahasiswa sebelumnya sudah pernah dilakukan, tetapi tujuan dan algoritma yang di gunakan berbeda. Adapun penelitian sebelumnya yang digunakan untuk dapat dijadikan bahan pertimbangan dan di harapkan dapat membantu dalam penelitian ini dan penelitian berikutnya adalah penelitian Harsiti dari program studi sistem informasi Universitas Serang Raya Banten tahun 2013 yang menggunakan algoritma fuzzy c-means untuk pengelompokan konsentrasi jurusan berdasarkan data nilai matakuliah mahasiswa. Penelitian ini berhasil melakukan clustering data menurut bobot nilai mata kuliah dan nilai indeks prestasi mahasiswa untuk menentukan cluster konsentrasi jurusan pada program studi teknik informatika Universitas Serang Raya.

Penelitian lainnya yang juga menggunakan data nilai siswa sebagai objek adalah penelitian oleh Mardiani dari program studi Sistem Informasi STMIK GI MDP tahun 2014. Mardiani melakukan perbandingan algoritma k-means dan EM untuk cluterisasi data nilai mahasiswa berdasarkan asal sekolah dalam penelitiannya. Kemudian menyatakan bahwa k-means adalah algoritma terbaik dalam kasus yang ia teliti. Hal ini juga menjadi pertimbangan penulis untu menerapkan algoritma pada *data mining* dalam manambang knowledge yang terkandung di dalam data nilai mahasiswa. Penulis menggunakan algoritma yang berbeda yaitu algoritma apriori. Dari sisi pola kompetensi mahasiswa informatika yang ingin ditemukan penulis melalui penelitian ini, penulis juga merujuk pada beberapa penelitian yang juga mengarah kepada kompetensi mahasiswa informatika. Pertama adalah penelitian oleh Sofi Defiyanti dari Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang tahun 2015. Sofi Defiyanti menyatakan dalam kesimpulan penelitiannya bahwa nilai-nilai matakuliah yang dijadikan inputan dalam pengelompokkan peminatan tugas akhir mahasiswa teknik informatika adalah nilai- nilai yang telah didapatkan mahasiswa dari hasil belajarnya selama semester satu sampai dengan semester enam. Nilai-nilai mahasiswa dari semester 1 sampai semester 6 merupakan bekal yang akan dipakai dalam penulisan tugas akhir di semester 7 dan 8. Matakuliah yang akan dipergunakan adalah matakuliah wajib program studi teknik informatika karena matakuliah ini mendukung untuk pengelompokan peminatan tugas akhir mahasiswa program studi teknik informatika yaitu terdiri dari 35 nilai matakuliah.

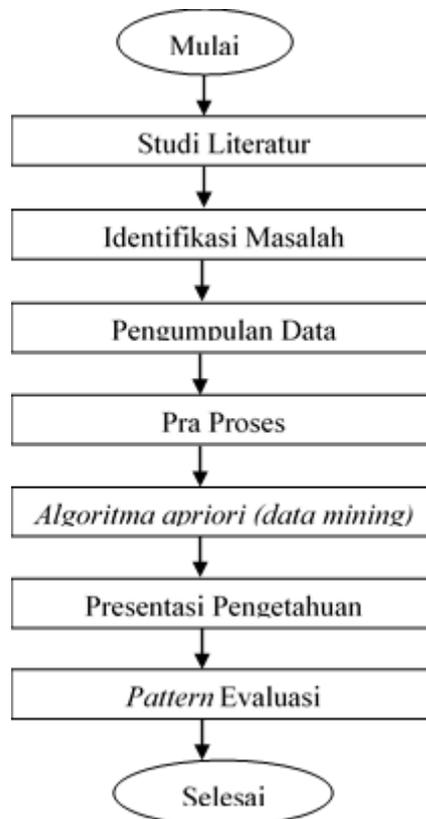
Selain itu, penelitian sejenis juga pernah dilakukan oleh Arif Jananto dari program studi sistem informasi Universitas Stikubank tahun 2012 yang membahas tema yang sama namun mengikuti standar kompetensi APTIKOM yang mengacu pada studi IEEE dan ACM yang disampaikan melalui dokumen publikasi computing curricula 2005. Dalam penelitiannya Arif Jananto menyatakan bahwa kompetensi calon lulusan dapat dibentuk dari beberapa matakuliah yang mempunyai kontent materi yang sejenis atau hampir sama. Kompetensi yang digunakan adalah 11 kompetensi dasar seorang lulusan informatika.

Umumnya penelaahan kesesuaian lulusan dilihat dari tempat kerja mereka saat ini atau dengan cara menelusuri dari angket-angket yang diisikan oleh alumni pada periode tertentu. Hal tersebut juga dapat dilakukan melalui temu alumni maupun melalui pusat informasi alumni pada tiap perguruan tinggi. Jarang sekali perguruan tinggi melihat kompetensi lulusannya sebelum dilepas ke dunia nyata. Salah satu variabel yang bisa digunakan adalah nilai matakuliah yang telah diperoleh mahasiswa atau calon lulusan.

2. Metode Penelitian

Pada penelitian ini, yang menjadi obyek penelitian adalah data-data yang didapat dari bagian akademik Program Studi Sistem Informasi Fasilkom Unilak berupa data nilai mahasiswa permatakuliah. Standar kompetensi dasar yang digunakan dalam penelitian adalah mangacu kepada ACM/IEEE 2013 yaitu terdiri dari : Pembentukan karakter, Matematika dan Statistika,

Algoritma dan Pemrograman, Sistem Cerdas, Rekayasa Perangkat Lunak, Komputer Arsitektur, Sistem Terdistribusi.



Gambar 1. Bagan Alir Kerangka Kerja Penelitian

Melakukan studi kepustakaan terhadap berbagai referensi yang berkaitan dengan penelitian yang dilakukan. Topik-topik yang dikaji antara lain meliputi : database, *data mining*, association rule, apriori dan beberapa algoritma lain yang mungkin dapat digunakan, serta perangkat lunak yang digunakan untuk proses mining.

Penjelasan:

1. Studi Literatur
Tahap ini melakukan kegiatan mempelajari algoritma *apriori* melalui buku dan beberapa tulisan ilmiah atau *paper*.
2. Identifikasi Masalah
Pada tahap ini, menggali permasalahan yang ditemukan pada obyek yang diteliti guna mencari alternatif solusi yang terkait dengan permasalahan.
3. Pengumpulan Data
Di tahap ini penulis mencari bahan-bahan dasar dalam penelitian ini yaitu data rekap nilai mahasiswa program studi sistem informasi yang sudah mencapai 138 sks dan didapat *dataset* yang berupa *record*, berupa data kompetensi dasar, mata kuliah dan nilai.
4. Pra Proses
Tahap pra proses ini dilakukan sebelum tahap proses yang meliputi pembersihan data (*data cleaning*), integrasi data (*data integration*), *task relevant data* yaitu melakukan seleksi data yang memiliki atribut yang relevan. Dengan melakukan seleksi data akan membantu tahapan proses *data mining* dengan algoritma apriori

dalam menemukan pola data yang berguna. Pada proses ini peneliti menggunakan software aplikasi sebagai alat bantu pengolahan data yaitu Ms. Excel.

5. Algoritma Apriori (Data Mining)
Tahap ini merupakan tahap *data mining*, yaitu proses dalam menerapkan algoritma *apriori* untuk menemukan pola hubungan atribut-atribut pada objek penelitian.
6. Presentasi pengetahuan (*knowledge presentation*),
Merupakan visualisasi dan penyajian pengetahuan mengenai algoritma 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*.
7. *Pattern* Evaluasi
Pada tahap ini dilakukan evaluasi terhadap informasi yang diperoleh atau data yang didapat dari objek penelitian. Tahap ini merupakan bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya. Dan pada proses ini juga pola atau informasi yang dihasilkan oleh penulis dibandingkan dengan pola informasi yang dihasilkan melalui *software data mining*.

3. Hasil dan Pembahasan

Pra Proses

Tahap pra proses ini meliputi pembersihan data (*data cleaning*), integrasi data (*data integration*), *task relevant data* yaitu melakukan seleksi data yang memiliki atribut yang relevan. Berikut ini atribut-atribut yang ada di dalam data rekap nilai.

Tabel 1 Data Rekap Nilai Mahasiswa

Atribut	Keterangan
Nama	Nama adalah nama mahasiswa.
NIM	NIM adalah nomor induk mahasiswa.
Mata Kuliah	Nama Mata Kuliah yang diambil.
M	Nilai mutu sesuai kurikulum yang berlaku
SKS	Sistem Kredit Semester
L	Lulus dengan keterangan nilai huruf

Task relevant data yaitu melakukan seleksi data yang memiliki atribut yang relevan. Atribut yang digunakan terdiri dari atribut pada berikut ini :

1. Atribut NIM digunakan sebagai penghubung dengan data rekap nilai.
2. Atribut Mata Kuliah digunakan untuk penghubung ke kompetensi dasar.
3. Nilai Huruf digunakan untuk penghubung ke kompetensi dasar.

Tahap selanjutnya adalah integrasi data. Integrasi data merupakan proses pengubahan atau penggabungan data ke dalam format yang sesuai untuk diproses dalam *data mining*. Dalam penelitian ini akan dicari nilai *support* dan *confidence* dari kombinasi atribut kompetensi dasar dengan atribut nilai dan matakuliah mahasiswa.

Parameter kompetensi dasar dapat dilihat dari keterangan kelulusan mahasiswa atas matakuliah yang diambil pada masing-masing kompetensi dasar berupa nilai yang diperoleh mahasiswa tersebut. *Format* data diubah menjadi *format* yang memudahkan untuk diproses.

Penggunaan Algoritma Apriori

Memulai aturan asosiasi dengan algoritma *apriori* terlebih dahulu menentukan minimum *support* dan minimum *confidence*. Peneliti menentukan minimum *support* adalah 35% dan minimum *confidence* adalah 95%. Setelah itu menemukan semua kombinasi dari *item*. Kemudian masuk ke titik persoalan *association rule*. Pada iterasi ke-*k*, akan ditemukan semua *itemsets* yang termasuk kategori *frequent itemsets*. Berikut ini adalah data *k-item* yang termasuk *frequent itemset*. Dari data awal pada dibentuk kombinasi *item* atau sering disebut *itemset* yang pada proses selanjutnya berfungsi sebagai kandidat *itemset*. Pada proses ini teknik aturan asosiasi dengan algoritma *apriori* mulai digunakan. Proses *mining* pertama selesai dan menghasilkan asosiasi yang *frequent itemset* yang pada pembahasan selanjutnya diberi nama *F1*, *F2*. Tahap kedua proses *mining* adalah pembentukan aturan asosiasi. Kombinasi *itemset* pada *F1* dan *F2* untuk selanjutnya disebut *rule*. Dari *rule* tersebut, dicari *rule* yang merupakan *association rule* atau merupakan aturan asosiatif karena memenuhi syarat minimum *confidence*, dimulai dari kombinasi *2-itemset*. Untuk menghitung *confidence* digunakan rumus berikut :

$$\text{Confidence} = \frac{\text{Support}(A \cup B)}{\text{Support} A}$$

Dari *F2* yang telah ditemukan, dapat dilihat besarnya nilai *confidence* dari *2-itemset*. Dari tabel calon aturan asosiasi *F2* diatas terlihat ada 6 aturan yang dihasilkan yang dipilih berdasarkan batas minimum *confidence*. Kemudian aturan atau *rules* tersebut disederhanakan menjadi 3 *rules*, karena beberapa diantaranya bermakna sama.

Rule yang dihasilkan merupakan pola yang ingin ditemukan. Pola kompetensi mahasiswa tersebut dapat memberikan knowledge tersembunyi di dalam data nilai mahasiswa. Beberapa knowledge tersebut adalah mengenai matakuliah apa yang dikuasai dengan baik oleh mahasiswa dan mata kuliah yang sulit untuk dikuasai. Matakuliah yang bisa dikuasai mahasiswa secara bersamaan dengan nilai baik sebelum mahasiswa tersebut lulus adalah matakuliah yang termasuk pada kompetensi dasar pembentukan karakter, algoritma dan pemrograman, rekayasa perangkat lunak, dan arsitektur komputer. Penjabarannya sebagai berikut :

Tabel 2 Kompetensi Dasar dengan Mata Kuliah

KOMPETENSI DASAR	MATA KULIAH
Pembentukan karakter	PENGANTAR MANAJEMEN
	BAHASA INDONESIA
	PENDIDIKAN PANCASILA
	BAHASA INGGRIS I
	PENDIDIKAN AGAMA
	KEWARGANEGARAAN
	BAHASA INGGRIS II
	ETIKA PROFESI
	KEWIRAUSAHAAN
	Algoritma dan Pemrograman
PEMROGRAMAN INTERNET I	
ALGORITMA PEMROGRAMAN II	
PEMROGRAMAN BERORIENTASI OBJEK I (JAVA)	
PEMROGRAMAN BERORIENTASI OBJEK II (JAVA)	
PEMROGRAMAN INTERNET II	

	PEMROGRAMAN JARINGAN NIRKABEL
	PEMROGRAMAN VISUAL I
	PEMROGRAMAN VISUAL II
Rekayasa Perangkat Lunak	BASIS DATA I
	DASAR AKUNTANSI
	ANALISIS PROSES BISNIS
	TEORI ORGANISASI
	AKUNTANSI LANJUTAN
	ANALISIS DAN PERANCANGAN SISTEM
	BASIS DATA II
	E-BISNIS
	MANAJEMEN PROYEK SISTEM INFORMASI
	METODOLOGI PENELITIAN
	REKAYASA PERANGKAT LUNAK
	SISTEM OPERASI
	SISTEM PENUNJANG KEPUTUSAN
	STRUKTUR DATA
	TESTING DAN IMPLEMENTASI SISTEM
Komputer Arsitektur	PENGANTAR SISTEM DAN TEKNOLOGI INFORMASI
	JARINGAN KOMPUTER

4. Kesimpulan

Dari hasil penelitian yang telah dilakukan dapat diambil kesimpulan bahwa kompetensi calon lulusan dapat dibentuk dari beberapa matakuliah yang mempunyai konten materi yang sejenis atau hampir sama dan perguruan tinggi jangan hanya mementingkan dapat menghasilkan lulusan dengan Indeks Prestasi Kumulatif (IPK) yang sesuai dengan kebutuhan dunia usaha dan industri tanpa melihat kompetensi yang dimiliki oleh lulusan tersebut.

Daftar pustaka

- [1] Defiyanti, Sofi (2015). Penentuan Peminatan Tugas Akhir Mahasiswa Teknik Informatika Unsika. *Jurnal ilmiah Solusi*, Vol. 2 No. 5 ISSN:2355-1119 Maret 2015 – Mei 2015: 9-16.
- [2] Jananto, Arif (2012). Penggunaan Market Basket Analysis untuk Menentukan Pola Kompetensi Mahasiswa. *Jurnal Teknologi Informasi DINAMIK*. Volume 17, No.2, ISSN : 0854-9524, Juli 2012 : 82-89.
- [3] Mardiani (2014). Perbandingan Algoritma K-Means dan EM untuk Clusterisasi Nilai Mahasiswa Berdasarkan Asal Sekolah. *Citec Journal*, Vol. 1 No. 4 ISSN: 2354-5771 Agustus 2014 – Oktober 2014 : 316-325.
- [4] Patil, Suraj (2012). The Novel Approach for Improving Apriori Algorithm for Mining Association Rule. *Proceeding of "National Conference on Emerging Trends in Computer Technology 2012*, ISSN: 2231-258.
- [5] Sarjon Defit (2012). "Data Mining". (Bahan Kuliah *Advance Database*). Padang.
- [6] Taufik, Emha (2009). *Algoritma Data Mining Edisi I*. Yogyakarta: CV ANDI OFFSET.
- [7] Warsito, Hadi (2009). Hubungan antar Self Efficacy dengan Penyesuaian Akademik dan Prestasi Akademik. *PEDAGOGI Jurnal Ilmiah Pendidikam*, Volume IX No.1 April 2009: 29-47.