

Analisis Data Lulusan dengan Data Mining untuk Mendukung Strategi Promosi Universitas Lancang Kuning

Elvira Asril¹, Fana Wiza², Yogi Yunefri³

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

³Program Studi 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, ²wiza_fana@yahoo.com, ³yogiyunefri@unilak.ac.id

Abstrak

Setiap perusahaan maupun organisasi yang ingin tetap bertahan perlu untuk menentukan strategi promosi yang tepat. Penentuan strategi promosi yang tepat akan dapat mengurangi biaya promosi dan mencapai sasaran promosi yang tepat. Salah satu cara yang dapat dilakukan untuk penentuan strategi promosi adalah dengan menggunakan teknik data mining. Teknik data mining yang digunakan dalam hal ini adalah dengan menggunakan algoritma Clustering K-Means. Clustering merupakan pengelompokan record, observasi, atau kasus ke dalam kelas-kelas objek yang mirip. K-Means adalah metode klaster data non-hirarkis yang mencoba untuk membagi data ke dalam satu atau lebih klaster. Penelitian dilakukan dengan mengamati beberapa variabel penelitian yang sering dipertimbangkan oleh perguruan tinggi dalam menentukan sasaran promosinya yaitu asal sekolah, daerah, dan jurusan. Hasil penelitian ini adalah berupa pola menarik hasil data mining yang merupakan informasi penting untuk mendukung strategi promosi yang tepat dalam mendapatkan calon mahasiswa baru.

Kata kunci: Data Mining, Clustering, K-Means

Abstract

Each company or organization that wants to survive needs to determine appropriate promotional strategies. Determination of appropriate promotional strategies will be able to reduce costs and achieve the goals the promotion of proper promotion. One way that can be done to determine campaign strategy is to use data mining techniques. Data mining techniques used in this case is to use a K-Means clustering algorithm. Clustering is the grouping of records, observation, or in the case of the object classes that are similar. K-Means is a method of non-hierarchical clustering of data that is trying to divide the data into one or more clusters. The study was conducted by observing some of the variables that are often considered by the college in determining the target of promotion that the school of origin, region, and department. Results of this study are interesting pattern of results in the form of data mining that is important information to support appropriate promotional strategies in getting new students.

Keywords: Data Mining, Clustering, K-Means

1. Pendahuluan

Universitas Lancang Kuning merupakan salah satu Universitas swasta tertua di Pekanbaru. Mahasiswa Universitas Lancang Kuning sangat banyak yang berasal dari berbagai daerah di Riau. Tidak hanya berasal dari Pekanbaru, mahasiswa Universitas Lancang Kuning pun banyak yang berasal dari luar Pekanbaru, seperti Dumai, Duri, Bangkinang, Siak, Sumatera Barat, Jambi dan lainnya. Seiring berjalannya waktu, maka mulai banyak bermunculan Universitas swasta dan Sekolah-sekolah Tinggi. Tentu saja untuk menghadapi persaingan ini

diperlukan strategi-strategi untuk memantapkan keberadaan Universitas Lancang Kuning. Banyaknya Universitas swasta dan Sekolah Tinggi yang berdiri mewajibkan Universitas Lancang Kuning melakukan serangkaian promosi-promosi aktif yang beragam, agar tidak kalah bersaing dengan Universitas lainnya.

Penelitian ini bertujuan untuk menganalisa dan mengolah data mahasiswa yang telah lulus dari Universitas Lancang Kuning untuk mendapatkan informasi yang penting dan bermanfaat. Atribut data yang akan digunakan adalah nama mahasiswa, jurusan yang diambil, sekolah asal, kota asal mahasiswa dan nilai Indeks Prestasi Kumulatif. Pada penelitian ini analisa dan pengolahan menggunakan teknik data mining dengan algoritma *K-Means clustering*. Manfaat yang diharapkan dari penelitian ini adalah membantu pihak pemasaran Universitas Lancang Kuning dalam melakukan promosi dan mencari calon mahasiswa baru dari berbagai kota di Indonesia. Melalui penelitian ini Penulis berusaha memaparkan informasi penting yang bermanfaat untuk mendukung strategi promosi Universitas Lancang Kuning.

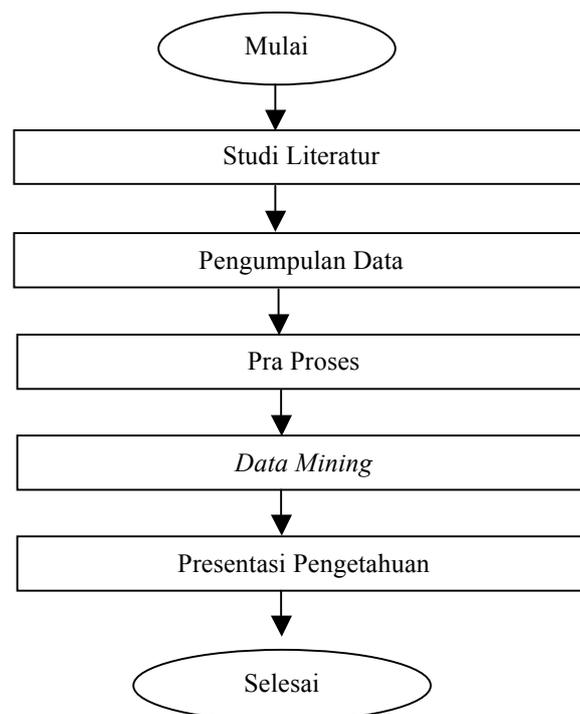
Penelitian dengan tema analisis data lulusan dengan data mining untuk mendukung strategi promosi perguruan ini sebelumnya sudah pernah dilakukan, tetapi tempat dan algoritma yang di gunakan berbeda-beda. Adapun penelitian sebelumnya yang digunakan untuk dapat dijadikan bahan pertimbangan dan di harapkan dapat membantu dalam penelitian ini dan penelitian berikutnya. Penelitian [1] menggunakan metode apriori untuk menganalisa penjualan barang yang pada akhirnya digunakan oleh pengambil keputusan untuk menentukan strategi penjualan. Peneliti [2] menggunakan studi kasus *President University* dalam penerapan algoritma K-Means untuk menentukan strategi promosi ke daerah di seluruh Indonesia agar lebih efektif dan efisien. Peneliti [3] melihat adakah hubungan antara nilai IPK dengan lamanya menyelesaikan sarjana atau gelas master seorang mahasiswa dengan membandingkan kedua algoritma *Affinity Propagation* dan K-Means untuk pengelompokannya. Penelitian Arlinda, Zefy and Arie ,Vatresia and Boko , Susilo tahun 2013 [4] yang berjudul Implementasi Algoritma K-Means Untuk Mengetahui Sebaran Mahasiswa Sebagai Informasi Penunjang Promosi Universitas Bengkulu. Penelitian ini bertujuan untuk mendapatkan informasi yang bermanfaat tentang sebaran mahasiswa di Fakultas Keguruan dan Ilmu Pendidikan Universitas Bengkulu, serta menerapkan algoritma k-means pada pengelompokan data. Pada penelitian ini sistem dibangun dengan menggunakan perangkat lunak Delphi 7, metode pengembangan sistem sekuensial linier, sedangkan analisis dan desain perangkat lunak menggunakan Unified Modeling Language (UML). Penelitian Rima Dias Ramadhani [5] yang berjudul Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro. Penelitian ini bertujuan untuk melakukan pengelompokan terhadap data mahasiswa Universitas Dian Nuswantoro dengan memanfaatkan proses data mining dengan menggunakan teknik Clustering. Metode yang digunakan adalah CRISP-DM dengan melalui proses business understanding, data understanding, data preparation, modeling, evaluation dan deployment. Algoritma yang digunakan untuk pembentukan cluster adalah algoritma K-Means. K-Means merupakan salah satu metode data non-hierarchical clustering yang dapat mengelompokkan data mahasiswa ke dalam beberapa cluster berdasarkan kemiripan dari data tersebut, sehingga data mahasiswa yang memiliki karakteri stik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain. Implementasi menggunakan RapidMiner 5.3 digunakan untuk membantu menemukan nilai yang akurat. Atribut yang digunakan adalah kota asal, program studi dan IPK mahasiswa. Cluster mahasiswa yang terbentuk adalah tiga cluster, dengan cluster pertama 804 mahasiswa, cluster kedua 2792 mahasiswa dan cluster ketiga sejumlah 223 mahasiswa. Hasil dari penelitian ini digunakan sebagai salah satu dasar pengambilan keputusan untuk menentukan strategi promosi berdasarkan cluster yang terbentuk oleh pihak admisi UDINUS.

2. Metode Penelitian

Metodologi penelitian merupakan langkah-langkah kerja yang perlu dilakukan agar penenelitian ini menjadi lebih mudah. Penelitian merupakan rangkaian kegiatan ilmiah. Sebagai langkah awal maka perlu adanya studi literatur untuk menentukan metode *data mining* untuk

pengolahan data dan penentuan alternatif solusi. Selanjutnya identifikasi masalah berkenaan dengan masalah yang di bahas, kemudian dilakukan pengumpulan data berkaitan dengan permasalahan yang akan diteliti, selanjutnya dilakukan pengumpulan data untuk menentukan parameter-parameter yang akan dijadikan landasan untuk melakukan penelitian ini. Secara umum proses-proses yang dilakukan dalam pelaksanaan penelitian ini sebagai berikut (gambar 1):

1. Studi literatur yang dilakukan untuk mendapatkan teori-teori tentang permasalahan yang telah dirumuskan. Teori-teori ini berfungsi sebagai panduan untuk mendapatkan solusi.
2. Mengumpulkan data-data yang diperlukan untuk penelitian ini. Data didapat dari bagian akademik yang berupa data mahasiswa angkatan 2010 dari 18 program studi yang telah dinyatakan lulus dari Universitas Lancang Kuning. Data-data ini berisi data diri dari mahasiswa.
3. Data-data yang telah didapatkan kemudian dilanjutkan ke pra proses data mining. Tahap pertama adalah *data cleaning*. Pada tahap ini data-data yang memiliki atribut yang tidak lengkap atau kosong akan dihapus. Selanjutnya tahap *data integration* dimana data yang berjenis data nominal diinisialisasikan ke dalam bentuk angka dapat diolah dengan menggunakan algoritma *K-means Clustering*. Tahap ini sejalan dengan proses *task relevan data*. Proses ini adalah seleksi atribut-atribut data yang dianggap relevan dengan rumusan masalah. Dalam penelitian ini hanya beberapa atribut data saja yang digunakan, seperti nama mahasiswa, nomor induk mahasiswa, program studi, sekolah asal, alamat, dan indeks prestasi kumulatif.
4. Semua data yang telah didapatkan tadi kemudian diolah dengan proses *data mining*. Pengolahan data ini menggunakan algoritma *K-Means Clustering*. Hasil proses ini adalah data yang dikelompokkan berdasar kemiripan karakteristik dari setiap data, sehingga dapat ditemukan pola atau informasi yang tersembunyi dari data-data tersebut.
5. Berdasarkan hasil tersebut, dilakukan *knowledge presentation*. *Knowledge presentation* yaitu menyajikan pola atau informasi yang ditemukan yang sesuai fakta atau hipotesa yang ada sebelumnya. Pola ini nantinya adalah pola yang berguna sebagai *output* penelitian ini. *Output* tersebut akan menjawab masalah yang telah dirumuskan di awal karena pada proses ini dilakukan formulasi keputusan atau aksi dari hasil yang didapat.



Gambar 1. Bagan Alir Kerangka Kerja Penelitian

3. Hasil dan Pembahasan

Bab ini adalah pembahasan mengenai proses-proses yang ada pada kerangka kerja penelitian.

3.1 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. Dengan melakukan seleksi data akan membantu tahapan proses *data mining* dalam menemukan pola data yang berguna, yang merupakan proses transformasi.

Pada pembersihan data, data-data yang tidak lengkap isinya tidak dilibatkan dalam penelitian. Dan data-data yang lengkap isinya, dipilih beberapa atribut agar data yang akan diolah benar-benar relevan dengan kebutuhan. Dengan demikian akan meningkatkan performa dalam proses *mining*.

Task relevant data yaitu melakukan seleksi data yang memiliki atribut yang relevan. Dengan melakukan seleksi data akan membantu tahapan proses *data mining* dalam menemukan pola data yang berguna. Oleh karena itu, tidak semua atribut data akan digunakan, hanya atribut yang dianggap peneliti berguna dan sebarannya tidak terlalu acak.

Tabel 1. Data Awal

No	NIM	Nama	Prodi	Alamat	TTL	Kelamin	Agama	Asal Sekolah	Nama Daerah	IPK
1	1061201251	JOHARI	MANAJEMEN	Jl. Raya Miras Km. 26 Perawang Mess IKPP 26K Blok N	Bagan Siapi-api,09-10-1990	L	Budha	SMA Methodist Bagan Siapi-api	Bagan siapi-api	3,37
2	1084205037	ERNAWATI	PENDIDIKAN BIOLOGI	Sibiruang Kec. Koto Kampar Hulu	Bagan Siapi-api,23-12-1963	P	Islam	SMA N 1 Bagan Siapi-api	Bagan siapi-api	3,15
3	1062201186	SAHIDIN	AKUNTANSI	JL. SUKAJAYA NO. 17 B	Selat Baru,09-08-1990	L	Islam	SMA N 1 BANTAN BENGKALIS	Bengkalis	3,04
4	1079201049	KAMASIAH	SASTRA INDONESIA	Jl. Teluk Leok Rumbai Pesisir	Sungai Selari,13-09-1990	P	Islam	SMAN 1 Bukit Batu	Bengkalis	2,75
5	1022201170	IDOLA MURSYID MAJA	TEKNIK SIPIL	Marpoyan GG.Amaliah 1	Bengkalis,28-10-1986	L	Islam	SMK Awang Mahmuda Bengkalis	Bengkalis	3,81
6	1063201213	NURAINI	ILMU ADMINISTRASI NEGARA	Perum Arengka Lestari Blok B26 Pekanbaru	Bengkalis,02-09-1991	P	Islam	SMKN 2 Bengkalis	Bengkalis	3,19
7	1079211095	MUSYAWIR	SASTRA DAERAH / MELAYU	Jl. Sembilang Rumbai	Bengkalis,24-04-1990	L	Islam	SMUN 2 Bengkalis	Bengkalis	3,91
8	1054211055	MUHAMMAD SOLICHIN	AGROTEKNOLOGI	Jl.Cemara Bengkalis	Bengkalis,28-09-1988	L	Islam	SMUN 3 Bengkalis	Bengkalis	3,83
9	1054201032	RATNA WULANDARI	AGRIBISNIS	Perum Cendana THP III No. 7 L Bukit Barisan	Bengkalis,16-01-1982	P	Islam	SMUN 3 Bengkalis	Bengkalis	3,17
10	1054251017	NINI METRIANA	KEHUTANAN	Jl. Umban Sari No. 21 Rumbai	Halalang,20-05-1988	P	Islam	SMAN 1 Tilatang Kamang Sumbar	Bukittinggi	3,15
11	1088203057	MERY SULASTRI	PENDIDIKAN BAHASA INGGRIS	Jl.Teluk Leok GG. Sengon Rumbai Pesisir	Bukit Tinggi,06-02-1978	P	Islam	SMK Pembina Bangsa Bukit Tinggi	Bukittinggi	3,42
...
300	1054201009	RUDIYANTO	AGRIBISNIS	Sorek Satu Pangkalan Kuras	Sorek,15-05-1986	L	Islam	SMK N 1 Pangkalan Kuras	Tembilahan	3,59

Karena data yang terlalu acak akan membuat proses *mining* memakan waktu lama dan tingkat hubungannya pun rendah. Atribut yang digunakan adalah pada tabel 2 yang terdiri dari :

Tabel 2. Atribut yang Digunakan

Atribut	Keterangan
NO	Nomor Urut
NIM	Nomor Induk Mahasiswa
NAMA	Nama Mahasiswa
ASAL SEKOLAH	Sekolah asal dari mahasiswa
NAMA DAERAH	Nama daerah berdasarkan sekolah asal mahasiswa
IPK	Indeks Prestasi Kumulatif terakhir

Tahap selanjutnya adalah integrasi data. Integrasi data merupakan proses perubahan atau penggabungan data ke dalam format yang sesuai untuk diproses dalam *data mining*. Seringkali data yang akan digunakan dalam proses *data mining* mempunyai format yang belum langsung bisa digunakan. Oleh karena itu, formatnya perlu dirubah. Dalam penelitian ini data yang berjenis data nominal diinisialisasikan ke dalam bentuk angka agar dapat diolah dengan menggunakan algoritma *K-means Clustering*. Tahap ini sejalan dengan proses *task relevan data*.

Data-data yang perlu diinisialisasi adalah data sekolah asal dan nama daerah. Pada data Sekolah asal, proses inisialisasi dimulai dengan mengelompokkan sekolah-sekolah tersebut ke jenis sekolahnya. Hasilnya menjadi 10 jenis sekolah pada tabel 3.

Tabel 3. Informasi Asal Sekolah

Asal Sekolah	Keterangan
MA Swasta	Madrasah Aliyah milik swasta
MA Negeri	Madrasah Aliyah milik negara
Paket C	Sekolah Kesetaraan dengan Sekolah Menengah Atas
SKMA	Sekolah Kehutanan Menengah Atas
SMA Swasta	Sekolah Menengah Atas milik Swasta
SMA Negeri	Sekolah Menengah Atas milik Negara
SMK Swasta	Sekolah Menengah Kejuruan milik swasta
SMK Negeri	Sekolah Menengah Kejuruan milik negara
SPG	Sekolah Pendidikan Guru

Berikut ini adalah data-data yang telah diinisialisasi ke dalam bentuk angka pada tabel 4 dan tabel 5. Tabel 6 adalah hasil akhir dari tahapan *preprocessing* data.

Tabel 4. Hasil Inisialisasi Asal Sekolah

Asal Sekolah	Keterangan	
MA Swasta	Madrasah Aliyah milik swasta	1
MA Negeri	Madrasah Aliyah milik negara	2
Paket C	Sekolah Kesetaraan dengan Sekolah Menengah Atas	3
SKMA	Sekolah Kehutanan Menengah Atas	4
SMA Swasta	Sekolah Menengah Atas milik Swasta	5
SMA Negeri	Sekolah Menengah Atas milik Negara	6
SMK Swasta	Sekolah Menengah Kejuruan milik swasta	7
SMK Negeri	Sekolah Menengah Kejuruan milik negara	8
SPG	Sekolah Pendidikan Guru	9
SPP	Sekolah Pendidikan Perawat	10

Tabel 5. Hasil Inisialisasi Nama Daerah

Nama Daerah	
Bagan siapi-api	1
Bengkalis	2
Bukittinggi	3
Dumai	4
Duri	5
Inhil	6
Inhu	7

Jambi	8
Jawa	9
Kab. Kampar	10
Kepri	11
kuansing	12
Minas	13
Palembang	14
Pekanbaru	15
Pelalawan	16
rohil	17
Rohul	18
Selat Panjang	19
Siak	20
Sumbar	21
Sumut	22
Tembilahan	23

Tabel 6. Data Test

NO	MAHASISWA	Xi	Yi	Zi
1	JOHARI	5	1	3,37
2	ERNAWATI	6	1	3,15
3	SAHIDIN	6	2	3,04
4	KAMASIAH	6	2	2,75
5	IDOLA MURSYID MAJA	7	2	3,81
6	NURAINI	8	2	3,19
7	MUSYAWIR	6	2	3,91
8	MUHAMMAD SOLICHIN	6	2	3,83
9	RATNA WULANDARI	6	2	3,17
10	NINI METRIANA	6	3	3,15
11	MERY SULASTRI (P1)	7	3	3,42
...
300	RUDIYANTO	8	23	3,59

Keterangan :

Xi = Asal sekolah

Yi = Nama daerah

Zi = IPK terakhir

3.1 Proses Data Mining dengan Algoritma K-means Clustering

Tahap ini merupakan tahap *data mining*, dimana pada tahap ini akan dibahas tahap proses *mining* dengan algoritma *K-Means Clustering*. Langkah-langkah melakukan clustering dengan metode *K-Means* adalah sebagai berikut:

- Memilih jumlah *cluster k*. Jumlah kluster yang diputuskan oleh peneliti adalah 4 kluster.
- Inisialisasi *k* pusat *cluster* dilakukan dengan cara yang paling sering dilakukan yaitu dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random. Pusat-pusat kluster adalah sebagai berikut pada tabel 7 :

Tabel 7. Pusat Cluster Awal

Cluster	No	Nama	Xi	Yi	Zi
P1	11	MERY SULASTRI (P1)	7	3	3,42
P2	178	FERNANDO DEDI MS.	7	15	2,95
P3	259	THERESIA TARI LESTARI S. (P2)	5	21	3,60
P4	295	SURIPAH	1	23	3,33

- c. Kemudian mengalokasikan semua data/objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster*. Jarak paling dekat antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster* mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:
- Dalam hal ini peneliti menggunakan aplikasi Microsoft Excel untuk menghitung dengan menggunakan rumus diatas. Hasil alokasi awal ini disebut iterasi 1 (terlampir).
- d. Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari semua data/objek dalam *cluster* tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Peneliti menggunakan rata-rata. Hasilnya sebagai berikut pada tabel 8 dan tabel 9 :

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \dots (1)$$

dimana:

$D(i, j)$ = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

Tabel 8. Perhitungan mencari Pusat Cluster ke-2

	P1-P2	P2-P3	P3-P4	P4-P1	Jumlah
BCV	12,01	6,36	4,48	20,88	43,73
WCV	Jumlah min kuadrat				2700,27
Rasio	BCV/WCV				0,02

Tabel 9. Pusat Cluster ke-3

Cluster	Xi	Yi	Zi
P1	6,25	4,925	3,298707
P2	5,479381	14,29897	3,282049
P3	5,758065	21,12903	3,314143
P4	1,5	22	3,260621

- e. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru. Jika rasio tidak berubah lagi maka proses *clustering* selesai. Pada penelitian ini proses berhenti pada iterasi ke-2.

3.2 Presentasi Pengetahuan (*Knowledge Presentation*)

Berdasarkan hasil yang diperoleh pada proses *data mining*, dilanjutkan ke tahap berikutnya *knowledge presentation*. *Knowledge presentation* yaitu menyajikan pola atau informasi yang ditemukan yang sesuai fakta atau hipotesa yang ada sebelumnya. Pola ini nantinya adalah pola yang berguna sebagai *output* penelitian ini. *Output* tersebut akan menjawab masalah yang telah dirumuskan di awal karena pada proses ini dilakukan formulasi keputusan atau aksi dari hasil yang didapat.

a. *Cluster 1 (P1)*

Klaster 1 terdiri dari mahasiswa yang berasal dari Bengkalis, Indragiri Hulu, Duri, dan Jawa sejumlah 47 orang. Sekolah asal mahasiswa-mahasiswa tersebut adalah dari SMA Swasta, SMK swasta, SMA Negeri dan SMK Negeri yang tersebar di daerah Bengkalis, Indragiri Hulu, Duri, dan Jawa. Jumlah terbesar berasal dari SMA Negeri. SMA Negeri tersebut terdiri dari SMAN 1, 2, 3 dan 4 Mandau; SMAN 1 Rengat, Pasir Penyu, Langgam, dan Ukui; SMA 1 Bantan, Bukit Batu, SMAN 2 Bengkalis, SMAN 3 Bengkalis; SMAN Conggeang Sumedang. Kemudian berisikan mahasiswa dengan perolehan rata-rata IPK 3,29.

b. *Cluster 2 (P2)*

Klaster 2 didominasi oleh mahasiswa yang berasal dari Pekanbaru dan Kabupaten Kampar sejumlah 178 orang. Sekolah asal mahasiswa-mahasiswa tersebut adalah dari SMA Swasta, SMK swasta, SMA Negeri dan SMK Negeri, MA Negeri, Paket C, dan MA Swasta. Jumlah terbesar berasal dari SMA Negeri dan SMA Swasta. SMA Negeri tersebut terdiri dari SMA Negeri 1, 2, 3, 4, 5, 6, 7, 8, 9, 11 dan 13 Pekanbaru; SMA Negeri 1 Tapung, Kampar Kiri, Kuala Kampar, Bangkinang dan SMA Negeri 2 Koto Kampar. SMA swasta terdiri dari SMA Annur, SMA Budhi Luhur, SMA Cendana, SMA Handayani, SMA Muhammadiyah, SMA Nurul Falah, SMA tri Bhakti, SMA Widya Graha Pekanbaru. Kemudian berisikan mahasiswa dengan perolehan rata-rata IPK 3,29.

c. *Cluster 3 (P3)*

Klaster 3 didominasi oleh mahasiswa yang berasal dari Sumatera utara, Siak, Sumatera Barat sejumlah 63 orang. Sekolah asal mahasiswa-mahasiswa tersebut adalah dari SMA Negeri dan SMA Swasta. Jumlah terbesar berasal dari SMA Negeri. SMA Negeri tersebut terdiri dari SMA 1, 2, 3, 9, dan 14 Siak; SMA Negeri 1 Suliki, SMA Negeri 1 Ranah Batahan, SMA Negeri 1 Sungai Geringging, SMA Negeri 1 Tarusan, SMA Negeri 1 Payakumbuh; SMA Negeri 1 Bagan Batu, Bagan Sinembah, Medan, Tarabintang, dan Tarutung. Kemudian berisikan mahasiswa dengan perolehan rata-rata IPK 3,30.

d. *Cluster 4 (P4)*

Klaster 3 didominasi oleh mahasiswa yang berasal dari Selat Panjang, Tembilahan, Sumatera utara, Siak, dan Sumatera Barat sejumlah 12 orang. Sekolah asal mahasiswa-mahasiswa tersebut adalah MA Negeri dan MA Swasta. Jumlah terbesar berasal dari MA Swasta. MA Swasta tersebut adalah MA Bustanul Ulum Siak, MA Darul Ikhlas Mandailing Natal, MA Budaya Cikampak Sumatera Utara, dan MA Al-Khairiyah. Kemudian berisikan mahasiswa dengan perolehan rata-rata IPK 3,33.

Dari data hasil clustering yang telah dilakukan di atas, maka dapat ditentukan beberapa strategi promosi, yaitu promosi pada persebaran wilayah berdasarkan tingkat akademik mahasiswa. Selanjutnya diharapkan dengan strategi yang diberikan pada masing-masing cluster dapat sesuai dengan tujuan yang telah dijelaskan.

4. Kesimpulan

Setelah dilakukan pengelompokan data mahasiswa melalui persebaran wilayah berdasarkan potensi akademik menggunakan K-Means clustering terbentuk empat cluster. Strategi promosi bagi calon mahasiswa baru yang tepat sasaran untuk setiap wilayah

berdasarkan cluster yang terbentuk adalah dengan mengirim tim promosi berdasarkan potensi akademik mahasiswa dengan melihat rata-rata IPK pada setiap *cluster*.

Dari hasil pengolahan data tersebut ada beberapa saran atau usul kepada bagian promosi Universitas Lancang Kuning dalam rangka meningkatkan kualitas mutu lulusan, yaitu:

1. Tim promosi yang dikirim sebaiknya memiliki pengetahuan dan potensi lebih dari satu jurusan untuk melakukan promosi pada kota yang memiliki minat lebih besar berdasarkan hasil klaster yang didapat agar yang dilakukan lebih efektif dan efisien. Pihak promosi dapat membagi tim-tim marketing untuk melakukan promosi di kota-kota tersebut sehingga sumber daya yang dibutuhkan disetiap kota tidak terlalu banyak.
2. Apabila pihak promosi Unilak ingin mendapatkan calon mahasiswa yang memiliki tingkat akademik di atas rata-rata, maka dapat dianalisa berdasarkan hasil rata-rata nilai IPK dari setiap mahasiswa yang telah lulus berdasarkan hasil klaster di penelitian ini.

Apabila Unilak memiliki rencana promosi hingga ke provinsi tetangga, hal ini juga akan berpeluang menjangkau mahasiswa baru yang banyak dan mampu lulus dengan IPK tinggi. Hal ini terlihat pada hasil klaster di penelitian ini, dimana peminat Unilak yang berprestasi juga banyak di provinsi tetangga.

Daftar Pustaka

- [1] Nurdin, Dewi Astika. Penerapan Data Mining untuk Menganalisa Penjualan Barang dengan Menggunakan Metode Apriori pada Supermarket Sejahtera Lhokseumawe. *Techsi*. 2015; vol 6(no 1): 133-155.
 - [2] Johan Oscar Ong. Implementasi Algoritma K-Means Clustering untuk Menentukan Strategi Marketing President University. *Jurnal Ilmiah Teknik Industri*. 2013; vol 12(no 1): 10-20.
 - [3] A. Juarna, et al . Analysis And Implementation Of Algorithm Clustering Affinity Propagation and K-Means At Data Student Based On GPA and Duration Of Bachelor-Thesis Completion. *JATIT & LLS*. 2012; vol.35 (no.1): 69-76.
 - [4] Arlinda, Zefry and Arie, Vatesia and Boko, Susilo. Implementasi Algoritma K-Means untuk Mengetahui Sebaran Mahasiswa Sebagai Informasi Penunjang Promosi Universitas (Studi Kasus : Mahasiswa Fakultas Keguruan dan Ilmu Pendidikan Universitas Bengkulu). Tesis. Fakultas Keguruan dan Ilmu Pendidikan Universitas Bengkulu. 2013.
 - [5] Rima, Ramadhani Dias. Data Mining Menggunakan Algoritma K-Means Clustering untuk Menentukan Strategi Universitas Dian Nuswantoro. Skripsi. Fakultas Ilmu Komputer Universitas Dian Nuswantoro. 2014.
-