

**PENERAPAN DATA MINING  
MENGUNAKAN ALGORITMA K-MEANS CLUSTERING  
UNTUK MENENTUKAN STRATEGI PROMOSI  
MAHASISWA BARU  
(Studi Kasus : Politeknik LP3I Jakarta)**

**Oleh :  
Rony Setiawan**

*Manajemen Informatika, Politeknik LP3I Jakarta  
Gedung Sentra Kramat Jl. Kramat Raya No. 7-9 Jakarta Pusat 10450  
Telp. 021-31904598 Fax. 021-31904599*

*Email : rony855185@yahoo.co.id*

**ABSTRAK**

Proses penerimaan calon mahasiswa baru di Politeknik LP3I Jakarta setiap tahun, menghasilkan data yang sangat banyak berupa profil calon mahasiswa. Kegiatan tersebut menimbulkan penumpukan data dalam jumlah besar, hal ini menjadi kesulitan untuk melakukan identifikasi terhadap calon mahasiswa. Penelitian ini membahas tentang penerapan *data mining*, menggunakan algoritma K-Means *Clustering* untuk menghasilkan profil yang memiliki kemiripan atribut yang sama. Metode yang digunakan adalah CRISP-DM dengan melalui proses *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* dan *deployment*. Implementasi proses K-Means *Clustering* menggunakan *Rapid Miner 7.0*. Atribut yang digunakan usia, agama, status ujian, kelulusan, registrasi, kelamin, gelombang pendaftaran, gelombang registrasi, Nilai Tes, Jurusan, program studi, sumber informasi, asal kota, lokasi kampus, pekerjaan orang tua, jenis sekolah, pembayaran uang registrasi. Menghasilkan jumlah kluster 4 ( $k=4$ ) dengan *cluster* pertama 17007 calon mahasiswa, *cluster* kedua sebanyak 83 calon mahasiswa, *cluster* ketiga sebanyak 12919 calon mahasiswa dan *cluster* keempat sebanyak 356 calon mahasiswa.

Kata Kunci : Data Mining, Algoritma, K-Means, *Clustering*, Strategi Promosi

**PENDAHULUAN**

Perkembangan teknologi informasi sekarang ini sangat pesat, hal ini menuntut setiap perusahaan untuk dapat saling bersaing. Dalam dunia bisnis yang dinamis dan penuh persaingan, perusahaan tidak lagi bisa unggul secara kompetitif hanya dengan memanfaatkan teknologi yang ada, karena teknologi tersebut digunakan hanya untuk mendukung kegiatan sehari-hari perusahaan.

Politeknik LP3I Jakarta setiap tahun melakukan penerimaan mahasiswa baru, hal ini akan berdampak terhadap pertumbuhan jumlah data yang sangat pesat dan menimbulkan tumpukan data yang berjumlah besar dalam basis data. Dengan jumlah data yang sangat besar Politeknik LP3I kesulitan untuk mendapatkan identifikasi terhadap calon-calon mahasiswa (calon mahasiswa) yang mendaftar ke Politeknik LP3I Jakarta. Penggunaan metode tradisional untuk

melakukan identifikasi konsumen dari data membutuhkan kemampuan manusia untuk menganalisa dan menginterpretasikan data. Dengan volume data yang berkembang sangat pesat, baik dari jumlah record dan jumlah field, analisa terhadap data tidak mungkin dilakukan secara manual oleh manusia. Penerapan data mining dapat membantu untuk menganalisa data yang diperoleh dari transaksi pada sistem informasi sehingga dapat menggali pola-pola yang dapat dijadikan pengetahuan baru untuk proses identifikasi calon mahasiswa di Politeknik LP3I Jakarta.

Untuk dapat melakukan promosi yang lebih efektif dan efisien, maka dalam penelitian ini dilakukan dengan cara mengolah data-data untuk mengetahui pola dari data-data tersebut sehingga kita dapat mengambil informasi-informasi yang tersembunyi dari data-data tersebut. Metode pengolahan data seperti ini sering disebut sebagai data mining. Pada penelitian ini analisa data mining dilakukan dengan menggunakan metode K-Means *Clustering*. Dengan menggunakan metode ini, data yang telah didapatkan dapat dikelompokkan kedalam beberapa cluster berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster dan yang memiliki karakteristik yang berbeda dikelompokkan dalam cluster yang lain yang memiliki karakteristik yang sama. Dengan adanya pengelompokan-pengelompokan data seperti ini, diharapkan bagian marketing dapat melakukan pemasaran dengan strategi yang tepat untuk mendapatkan calon mahasiswa baru.

## LANDASAN TEORI

### Data

Dalam Webster's New World's Dictionary tertulis bahwa datum: something known or assumed. Artinya, datum (bentuk tunggal data) merupakan

suatu yang diketahui/dianggap. Dengan demikian, data dapat memberi gambaran tentang suatu keadaan atau persoalan. Sedangkan, data menurut kamus Oxford Dictionary adalah The Facts. Jadi, dapat disimpulkan bahwa data adalah sesuatu yang nyata diketahui atau dianggap yang dipakai untuk keperluan suatu analisa, diskusi, presentasi ilmiah atau tes statistik (Supranto, 2000).

Data kualitatif secara sederhana dapat disebut data yang bukan berupa angka. Ciri utama data kualitatif didapat dengan cara menghitung, sehingga tidak memiliki nilai desimal. Selain itu data kualitatif memiliki ciri-ciri tidak bisa dilakukan operasi matematika, seperti penambahan, pengurangan, perkalian dan pembagian. Contoh data kualitatif adalah data gender, data golongan darah, data tempat tinggal atau data jenis pekerjaan. Agar dapat dilakukan proses pada data kualitatif atau non metric, data tersebut harus diubah ke dalam bentuk angka, proses ini dinamakan kategorisasi. Data kualitatif dibedakan menjadi dua jenis, yaitu data nominal dan data ordinal (Santoso, 2010).

Data nominal merupakan data yang paling rendah dalam level pengukuran data. Semua data memiliki posisi setara dalam arti tidak ada data yang memiliki tingkat yang lebih atau kurang dibandingkan dengan data yang lain. Jenis data nominal ini tidak memiliki jarak, urutan dan titik origin (Hidayat, 2011).

Data Ordinal adalah jenis data kualitatif namun memiliki level lebih tinggi dari data nominal. Data ordinal memiliki karakteristik nominal tapi terdapat perbedaan derajat, urutan, atau peringkat dalam objek tersebut (posisi data tidak setara) (Hidayat, 2011).

### Data Mining

Data mining memiliki pengertian lain yaitu *knowledge discovery* ataupun *pattern recognition* merupakan suatu istilah yang digunakan untuk mendapatkan pengetahuan yang tersembunyi dari

kumpulan data yang berukuran sangat besar. Tujuan utama data mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki (Susanto, 2010:2)

Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk pengenalan permasalahan pengambilan informasi dari database yang besar (Larose, 2005). Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database yang besar.

Menurut Larose, *data mining* dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Larose, 2005) :

1. Deskripsi  
Untuk mencari cara yang menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola dan kecenderungan yang sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan
2. Estimasi  
Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik daripada kearah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.
3. Prediksi  
Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.
4. Klasifikasi  
Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat

dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, sedang dan rendah.

5. Pengklusteran  
Pengklusteran merupakan pengelompokkan *record*, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record* dalam kluster lain. Algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal
6. Asosiasi  
Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja

*Clustering* merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*) yang memiliki arti bahwa karakteristik tiap kluster tidak ditentukan sebelumnya melainkan berdasarkan kemiripan atribut-atribut dari suatu kelompok atau *cluster*. *Cluster* membagi data menjadi kelompok-kelompok atau *cluster* berdasarkan suatu kemiripan atribut-atribut diantara sekumpulan data, atribut-atribut yang mirip tersebut dipresentasikan sebagai titik-titik dalam ruang multidimensi. Dalam data mining ada dua jenis metode *Clustering* yang digunakan dalam pengelompokan data, yaitu *hierarchical Clustering* dan *non-hierarchical Clustering* (Santosa, 2007).

### Algoritma K-Means Clustering

Salah satu lagoritma *Clustering* pada data mining adalah algoritma K-Means

*Clustering* untuk dapat menghasilkan kelompok yang memiliki kemiripan atribut yang sama. Algoritma pengelompokan K-Means yang akan menghasilkan kelompok catatan sebanyak  $k$  buah. Algoritma K-Means pertama kali digagas oleh J. MacQueen (Larose, 2005 dan Susanto, 2010).

Menurut Santosa (2007) dan Ong, langkah-langkah melakukan *Clustering* dengan metode K-Means adalah sebagai berikut:

1. Pilih jumlah *cluster*  $k$ .
2. Inisialisasi  $k$  pusat *cluster* ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat *cluster* diberinilai awal dengan angka-angka random,
3. Alokasikan 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 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:

$$D(i,j) = \sqrt{(X_{ki} - X_{kj})^2 + (X_{li} - X_{lj})^2 + \dots + (X_{ni} - X_{nj})^2} \quad (2.1)$$

dimana:

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

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

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

4. 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. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.

5. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses *Clustering* selesai. Atau, kembali ke langkah nomor 3 sampai pusat *cluster* tidak berubah lagi.

### Rapidminer

Rapiminer merupakan aplikasi *open source* berlisensi AGPL (GNU Affero General Public License) yang dapat digunakan untuk mengolah data mining yang dikembangkan oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer di *Artificial Intelligence Unit* dari University of Dortmund (Satria Wahono, 2015).

### CRISP-DM

CRISP-DM (*Cross-Industry Standard Process for Data Mining*) merupakan suatu metodologi data mining yang disusun oleh konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan telah ditetapkan sebagai proses standar dalam data mining. Menurut Larose, *data mining* memiliki enam fase CRISP-DM, seperti yang tertera pada gambar 2-1. (Larose, 2006:6).

1. Fase Pemahaman Bisnis (*Business Understanding Phase*)  
Fase pertama adalah memahami tujuan dan kebutuhan dari sudut pandang bisnis, kemudian menterjemakan pengetahuan ini ke dalam pendefinisian masalah dalam *data mining*. Selanjutnya akan ditentukan rencana dan strategi untuk mencapai tujuan tersebut
2. Fase Pemahaman Data (*Data Understanding Phase*)  
Fase ini dimulai dengan pengumpulan data yang kemudian akan dilanjutkan dengan proses untuk mendapatkan pemahaman yang mendalam tentang data, mengidentifikasi masalah kualitas data, atau untuk mendeteksi adanya bagian yang menarik dari data yang dapat digunakan untuk hipotesa untuk informasi yang tersembunyi.

### 3. Fase Pengolahan Data (*Data Preparation Phase*)

Fase ini meliputi semua kegiatan untuk membangun *dataset* akhir (data yang akan diproses pada tahap pemodelan/*modeling*) dari data mentah. Tahap ini dapat diulang beberapa kali. Pada tahap ini juga mencakup pemilihan tabel, *record*, dan atribut-atribut data, termasuk proses pembersihan dan transformasi data untuk kemudian dijadikan masukan dalam tahap pemodelan (*modeling*).

### 4. Fase Pemodelan (*Modeling Phase*)

Dalam fase ini akan dilakukan pemilihan dan penerapan berbagai teknik pemodelan dan beberapa parameternya akan disesuaikan untuk mendapatkan nilai yang optimal. Secara khusus, ada beberapa teknik berbeda yang dapat diterapkan untuk masalah *data mining* yang sama. Di pihak lain ada teknik pemodelan yang membutuhkan format data khusus. Sehingga pada tahap ini masih memungkinkan kembali ke tahap sebelumnya.

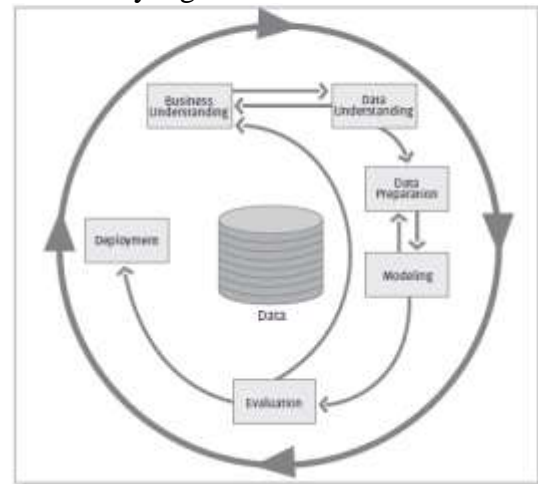
### 5. Fase Evaluasi (*Evaluation Phase*)

Pada fase ini, model sudah terbentuk dan diharapkan memiliki kualitas baik jika dilihat dari sudut pandang analisa data. Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas model sebelum digunakan dan menentukan apakah model dapat mencapai tujuan yang ditetapkan pada fase awal (*Business Understanding*).

### 6. Fase Penyebaran (*Deployment Phase*)

Pada tahap ini, pengetahuan atau informasi yang telah diperoleh akan diatur dan dipresentasikan dalam bentuk khusus sehingga dapat digunakan oleh pengguna. Tahap *deployment* dapat berupa pembuatan laporan sederhana atau mengimplementasikan proses *data mining* yang berulang dalam perusahaan. Dalam banyak kasus, tahap *deployment* melibatkan konsumen, di samping analis data,

karena sangat penting bagi konsumen untuk memahami tindakan apa yang harus dilakukan untuk menggunakan model yang telah dibuat.



Gambar 1 Fase CRISP-DM (Larose, 2006)

## Promosi

Menurut Kotler, “*promotion, the fourth marketing mix tools, stand for various activities, the company undertakes to communicate its products merits and to persuade target customers to buy them*”. Definisi tersebut mempunyai pengertian bahwa promosi meliputi semua alat yang terdapat dalam bauran promosi yang peranan utamanya adalah mengadakan komunikasi yang bersifat membujuk (Kotler, 1997). Strategi Promosi Menurut Kotler & Armstrong variabel variabel yang ada di dalam *promotional mix* ada lima, yaitu : (Kotler dan Armstrong, 2001)

1. Periklanan (*advertising*)
2. Penjualan Personal (*personal selling*)
3. Promosi penjualan (*sales promotion*)
4. Hubungan masyarakat (*public relation*)
5. Pemasaran langsung (*direct marketing*).

## METODOLOGI PENELITIAN

Penelitian ini dimulai dari melakukan pengamatan secara langsung pada Politeknik LP3I Jakarta untuk mendapatkan informasi tentang penerimaan mahasiswa baru dan promosi yang berjalan kepada bagian marketing dan melakukan seurvey kepada pengelola

data yaitu Unit Data dan Informasi (UDI) dan mendapatkan informasi tentang promosi yang digunakan dan sumber data yang dapat digunakan dalam penelitian sudah mendapat persetujuan dari bagian UDI.

### **Frame Work CRISP-DM**

Proses *data mining* yang dilakukan pada penelitian mengacu pada frame work CRISP-DM (Cross Industry Standard Process for Data mining)

1. Tahap Pemahaman Bisnis (Business Understanding)
  - a. Tujuan Bisnis
    - (1) Meningkatkan jumlah aplikasi yang mendaftar ke Politeknik LP3I Jakarta
    - (2) Meningkatkan jumlah mahasiswa yang melakukan registrasi penerimaan mahasiswa baru
  - b. Penilaian Situasi (*Situation Assesment*)
 

Politeknik LP3I Jakarta merupakan pendidikan tinggi yang memiliki 16 kampus yang tersebar di wilayah Jabetabek, dengan 4 (empat) program studi yaitu Program studi Administrasi Bisnis, Komputerisasi Akuntansi, Manajemen Informatika dan Hubungan Masyarakat
  - c. Tujuan *Data Mining*

tujuan data mining adalah untuk menentukan strategi promosi sehingga dapat menemukan profil yang bertujuan untuk :

    - (1) Mengetahui profil calon mahasiswa yang melakukan pendaftaran di Politeknik LP3I Jakarta.
    - (2) Mengetahui strategi promosi berdasarkan pada profil dari calon mahasiswa.
2. Pemahaman Data (Data Understanding)
  - a. Pengumpulan Data Awal
 

Kegiatan ini digunakan untuk mempelajari database dan struktur data yang ada pada Smart Campus Politeknik LP3I Jakarta, selanjutnya

menentukan tabel-tabel yang dipergunakan pada proses data mining.

- b. Deskripsikan Data
 

usia, agama, status ujian, status kelulusan, status registrasi, jenis kelamin, periode gelombang pendaftaran, periode gelombang registrasi, Nilai tes TPA, Jurusan di sekolah, program studi, sumber informasi, asal kota, lokasi kampus, pekerjaan orang tua, jenis sekolah dan jumlah pembayaran uang registrasi, dengan jumlah *record* sebanyak 30365.
- c. Eksplorasi Data
 

Untuk mendapatkan sumber data yang digunakan pada proses *data mining*, menggunakan database aplikasi dari calon mahasiswa 5 tahun terakhir yaitu dari tahun 2010 sampai tahun 2015.
- d. Verifikasi Kualitas Data.
 

Memastikan data yang digunakan sebagai atribut pada proses *data mining* tidak ada kesalahan, maka yang diperlukan adalah:

  - (1) Memastikan semua *record* terisi dengan data, dan semua kolom terisi dengan data yang benar jangan sampai ada data yang *anomaly*. Bila terjadi data konong maka diisi sesuai dengan ketentuan yang berlaku atau *record* tersebut dihapus.
  - (2) Tidak ada duplikasi data, jika terjadi maka hapus salah satu data.
3. Pengolahan Data (Data Preparation)
 

Pada tahap ini menetapkan struktur basis data yang digunakan untuk mempermudah dalam proses data mining, Data preparation dibagi menjadi 3 tahapan, yaitu :

  - a. Data Selection
 

Memilih data dari struktur basis data yang dimiliki untuk

dijadikan data yang akan digunakan pada data mining, data tersebut yang nantinya akan digunakan sebagai atribut data mining.

b. Data Processing

Memastikan kualitas data yang dipilih untuk proses data mining, tidak ada data yang noisy data dan missing value artinya data harus diyakinkan dulu data terisi dengan benar tidak ada data-data yang anomaly yang dapat memperngaruhi hasil *Clustering*. Untuk melakukan pembersihan dapat dilakukan dengan perintah query sederhana.

b. Data Transformation

Mengelompokan atribut-atribut yang telah terpilih dijadikan satu tabel, seperti pada tabel 1

Tabel 1.  
Atribut terpilih

No	Atribut	Keterangan	Alasan
1	Uts	Uts calon mahasiswa	Untuk mengetahui data calon mahasiswa yang dapat
2	Agama	1-Islam, 2-Kristen, 3-Hindu, 4-Budha	Program pendidikan agama
3	Ujian	1-Ya, 2-Tidak	Mengetahui apakah yang melakukan ujian berupa tes
4	Lulus	1-Ya, 2-Tidak	Benar atau tidak calon mahasiswa yang lulus
5	Registrasi	1-Ya, 2-Tidak	Mengetahui calon mahasiswa yang melakukan registrasi
6	Sex	1-Laki-laki, 2-Perempuan	Mengetahui calon mahasiswa yang dapat ikut seleksi LP 31 dan LP 32
7	Nilai_dasar	Diambil dari pendaftaran 1, 2 dan 3	Mengetahui calon mahasiswa melakukan pendaftaran di perguruan tinggi
8	Nilai_regis	Diambil dari pendaftaran 1, 2 dan 3	Mengetahui calon mahasiswa melakukan registrasi di perguruan tinggi
9	Nilai_uas	Hasil nilai tes potensi akademik nilai <60 Tidak lulus dan nilai >60 Lulus	Untuk mengetahui nilai potensi akademik dari calon mahasiswa
10	Ulang_regis	Revisi ulang dalam melakukan pendaftaran tidak terregistrasi	Mengetahui kemampuan ulang tes dalam melakukan tes registrasi
11	Jurusan_sekolah	Jurusan Calon mahasiswa pada waktu kuliah	Mengetahui asal jurusan calon mahasiswa untuk registrasi di LP 31
12	Nama_jurusan	Nama jurusan/konsentrasi dari program studi di LP 31	Untuk mengetahui jurusan apa yang paling banyak diterima oleh calon mahasiswa
13	Informasi	Informasi media (kita)	Untuk mengetahui media (kita) apa saja yang sering dilihat oleh calon mahasiswa dari media registrasi LP 31
14	Nama_kampus	Nama-nama Kampus LP 31	Untuk mengetahui kampus mana yang lebih banyak diterima oleh calon mahasiswa
15	Asal_kota	Wilayah tempat tinggal calon mahasiswa	Untuk mengetahui informasi asal wilayah tempat tinggal calon mahasiswa
16	Pekerjaan	Pekerjaan orang tua calon mahasiswa	Untuk mengetahui informasi jenis pekerjaan orang tua dari calon mahasiswa
17	Jenis_sekolah	Jenis Sekolah 1-SMA, 2-SMK, 3-SMA, 4-Diploma, 5-Lain-lain	Untuk mengetahui asal jenis sekolah dari calon mahasiswa

Data yang berjenis nominal seperti agama, ujian, lulus, registrasi, sex, jurusan\_sekolah, nama\_jurusan, informasi, nama\_kampus, asal\_kota, pekerjaan dan jenis\_sekolah harus dilakukan proses inisialisasi data terlebih dahulu ke dalam bentuk angka/numerical (Ong, 2013 dan

Ramadhani,2014). Untuk melakukan inisialisasi dapat dilakukan dengan :

1. Untuk atribut agama terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan agama selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 2.

Tabel 2  
Inisialisasi data agama

Agama	Frekuensi	Inisial
Islam	29099	1
Kristen	1227	2
Hindu	23	3
Budha	16	4

2. Untuk atribut ujian terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan kepersertaan ikut ujian selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 3.

Tabel 3  
Inisialisasi data ujian

Ujian	Frekuensi	Inisial
Ya	16510	1
Tidak	13855	2

3. Untuk atribut lulus terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan kelulusan hasil tes selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 4.

Tabel 4  
Inisialisasi data lulus

Lulus	Frekuensi	Inisial
Ya	16079	1
Tidak	14286	2

4. Untuk atribut registrasi terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan registrasi calon mahasiswa selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 5.

Tabel 5  
Inisialisasi data registrasi

Registrasi	Frekuensi	Inisial
Ya	16083	1



Tidak	14282	2
-------	-------	---

5. Untuk atribut sex terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan jenis kelamin (sex) calon mahasiswa selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 6.

Tabel 6  
Inisialisasi data jenis kelamin

Sex	Frekuensi	Inisial
Perempuan	16258	1
Laki-laki	14107	2

6. Untuk atribut Jurusan\_sekolah terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan jurusan yang ada di masing-masing sekolah calon mahasiswa selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 7.

Tabel 7  
Inisialisasi data Jurusan di sekolah

Jurusan_Sekolah	Frekuensi	Inisial
IPS	8413	1
IPA	4873	2
Administrasi Perkantoran	4314	3
Akuntansi	3718	4
Teknik	2115	5
Teknik Automotif	1128	6
Penjualan	729	7
Multimedia	718	8
Rekaya Perangkat Lunak	504	9
Sekertaris	502	10
Informatika Komputer	451	11
Administrasi Bisnis	366	12
Komputer Dan Jaringan	287	13
Perhotelan	282	14
Mesin	242	15

7. Untuk atribut nama\_jurusan terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan nama jurusan/konsentrasi yang diminati oleh calon mahasiswa selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 8.

Tabel 8

Inisialisasi data Prgram studi/konsentrasi LP3I

Jurusan/Konsentrasi LP3I	Frekuensi	Inisial
Administrasi Perkantoran	11413	1
Komputerisasi Akuntansi	7037	2
Informatika Komputer	4428	3
Administrasi Bisnis	2571	4
Teknik Informatika	2031	5
Multimedia	990	6
Sekretaris	795	7
Hubungan Masyarakat	368	8
Marketing Bisnis Komunikasi	267	9
Teknik Otomotif	231	10
Bisnis Internasional	209	11
Perbankan Syariah	25	12

8. Untuk atribut informasi terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan informasi yang diketahui oleh calon mahasiswa selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 9.

Tabel 9  
Inisialisasi data informasi

Informasi	Frekuensi	Inisial
Presentasi Sekolah	6536	1
Teman	5098	2
Brosur	3349	3
Spanduk Di Jalan	2535	4
Gedung	1886	5
Orang Tua/Saudara	1248	6
Sekolah Binaan	1248	7
Media Sosial	864	8
Koran	668	9
Hunting	623	10
Spanduk Kampus	578	11
Tele Selling	566	12
Website	397	13
Televisi	393	14
Datang Langsung	351	15

9. Untuk atribut nama\_kampus terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan nama kampus yang diminati oleh calon mahasiswa selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 10.



Tabel 10  
Inisialisasi data Kampus

Kampus	Frekuensi	Inisial
Kampus Kramat Raya	4028	1
Kampus Bekasi	3486	2
Kampus Depok	2670	3
Kampus Jakarta Utara	2634	4
Kampus Cimone-Tangerang	2390	5
Kampus Pondok Gede	2235	6
Kampus Pulogadung	1721	7
Kampus Cileungsi	1672	8
Kampus Ciputat	1628	9
Kampus Metropolitan	1496	10
Kampus Cikarang	1393	11
Kampus Pasar Minggu	1228	12
Kampus Ciledug	1204	13
Kampus Sumber Arta	1032	14
Kampus Blok M	961	15
Kampus Cibinong	587	16

10. Untuk atribut asal\_kota terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan asal kota alamat selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 11.

Tabel 11  
Inisialisasi data Asal Kota/Wilayah

Kota/Wilayah	Frekuensi	Inisial
Bekasi	6060	1
Tangerang	2726	2
Bogor	2650	3
Jakarta Timur	2376	4
Depok	2142	5
Jakarta Pusat	1827	6
Jakarta	1736	7
Jakarta Utara	1517	8
Jakarta Selatan	642	9
Tanjung Priuk	622	10
Jawa Barat	594	11
Tangerang Selatan	537	12

11. Untuk atribut pekerjaan terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan pekerjaan orang tua selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 12.

Tabel 12  
Inisialisasi Data Pekerjaan Orang Tua

Pekerjaan Orang Tua	Frekuensi	Inisial
Pegawai Swasta	16792	1
Wiraswasta	8129	2
Ibu Rumah Tangga	1612	3
Buruh	1572	4
Pegawai Negeri Sipil	925	5
Pensiunan	380	6
Pegawai Negeri	378	7
TNI/POLRI	303	8
Tani	139	9
BUMN	57	10
Dosen	20	11

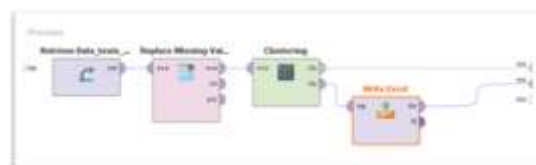
12. Untuk atribut jenissekolah terlebih dahulu dilakukan perhitungan frekuensi calon mahasiswa berdasarkan jenis sekolah (SMK, SMA, MA ) selanjutnya diurutkan dari terbesar ke terkecil seperti pada tabel 13.

Tabel 13  
Inisialisasi Data Jenis Sekolah

Jenis Sekolah	Frekuensi	Inisial
SMK	16118	1
SMA	13205	2
MA	992	3
Diploma	50	4

#### 4. Pemodelan (Modeling)

Metode *Clustering* yang digunakan adalah non-hierarchical *Clustering* dengan menentukan jumlah jumlah K cluster terlebih dahulu. Machine mining yang digunakan dalam proses data mining dengan Algoritma K-Means *Clustering* adalah Rapidminer 7.0, seperti pada gambar 2.



Gambar 2  
Proses Model *Clustering* pada Rapidminer 7.0

#### 5. Evaluasi (Evaluation)

Melakukan evaluasi hasil data mining terhadap algoritma K-Means *Clustering* apakah profil dari model yang diterapkan sudah dapat memenuhi

pencapai tujuan pada fase pertama (Business Understanding), Kunci dari tahap ini adalah menentukan apakah ada masalah bisnis yang belum dipertimbangkan.

#### 6. Penyebaran (Deployment)

Profil yang dihasilkan dari proses data mining, disampaikan kepada pihak Politeknik LP3I Jakarta sebagai rekomendasi untuk menentukan strategi promosi yang digunakan pada penerimaan calon mahasiswa.

## HASIL DAN PEMBAHASAN

### Proses Pengujian Jumlah Cluster

Proses pengujian untuk mencari jumlah kluster yang tepat untuk mendapatkan profil calon mahasiswa dengan algoritma K-Means Clustering, untuk menentukan jarak antar data/atribut ke pusat cluster (*Centroid*) diukur dengan menggunakan *Euclidean Distance*. Proses pengujian dilakukan mulai dari jumlah kluster 2 ( $k=2$ ) sampai jumlah kluster 5 ( $k=5$ ), pengujian pertama dilakukan dengan jumlah kluster 2 ( $k=2$ ), menghasilkan nilai pusat kluster (*centroid*), seperti pada tabel 14.

Tabel 14

Pusat kluster (*centroid*) jumlah kluster 2 ( $k=2$ )

Atribut	Cluster_0	Cluster_1
Usia	19	19
Agama	1	1
Ujian	1	1
Lulus	1	1
Registrasi	1	1
Sex	2	2
gel daftar	2	2
gel regis	2	2
nilai tes	42	80
uang regis	1751011	6317326
Jurusan Sekolah	5	5
Nama Jurusan	3	3
Informasi	7	8
Nama Kampus	6	5
Asal Kota	11	12
Pekerjaan	2	2
Jenis Sekolah	2	2

Cluster 0: 29353 items

Cluster 1: 1012 items

Total number of items: 30365

Berdasarkan hasil Clustering dengan jumlah kluster 2 ( $k=2$ ) profil yang ditampilkan terlalu global dengan banyak kemiripan antara kluster pertama dan kedua, karakteristik dari kedua kluster tersebut didominasi calon mahasiswa yang melakukan registrasi (registrasi=1), dengan rata-sata pembayaran uang registrasi pada kluster pertama sebesar 1.750.011 dan rata-rata pembayaran uang registrasi pada kluster kedua sebesar 6.317.326, kluster tersebut belum menunjukkan profil calon mahasiswa. Jumlah kluster ( $k=2$ ) menghasilkan profil yang telalu sederhana hanya membagi menjadi 2 kelompok yaitu kelompok calon mahasiswa yang tidak diterima dan calon mahasiswa yang mendaftar ke Politeknik LP3I Jakarta, maka jumlah kluster 2 ( $k=2$ ) belum dapat digunakan sebagai pengetahuan untuk strategi promosi. Pengujian dilanjutkan dengan menggunakan nilai kluster sebanyak 3 ( $k=3$ ) menghasilkan nilai pusat kluster (*centroid*), seperti pada tabel 15.

Tabel 15

Pusat kluster (*centroid*) jumlah kluster 3 (k=3)

Atribut	Cluster_0	Cluster_1	Cluster_2
Usia	20	19	19
Agama	1	1	1
Ujian	2	1	1
Lulus	2	1	1
Registrasi	2	1	1
Sex	1	2	2
gel daftar	2	2	2
gel regis	2	2	2
nilai tes	14	81	80
uang regis	183718	10028866	3991137
Jurusan Sekolah	5	5	5
Nama Jurusan	3	3	3
Informasi	7	6	7
Nama Kampus	6	5	6
Asal Kota	10	13	11
Pekerjaan	2	2	2
Jenis Sekolah	2	2	2

Cluster 0: 17007 items

Cluster 1: 224 items

Cluster 2: 13134 items

Total number of items: 30365

Hasil dari *Clustering* dengan jumlah kluster 3 (k=3), kluster pertama memiliki karakteristik didominasi calon mahasiswa yang tidak melakukan registrasi (registrasi=2) dengan rata-rata pembayaran uang registrasi sebesar 183.718, kluster kedua dengan karakteristik calon mahasiswa yang melakukan registrasi (registrasi=1) dengan rata-rata pembayaran uang registrasi sebesar 10.028.866, kluster ketiga memiliki karakteristik calon mahasiswa yang melakukan registrasi (registrasi=1) dengan rata-rata besar pembayaran uang registrasi sebesar 3.991.137, terlihat perbedaan karakteristik jumlah kluster 2 (k=2) dengan jumlah kluster 3 (k=3), jumlah kluster 3 (k=3) lebih baik dari jumlah kluster 2 (k=2). Berdasarkan atribut jumlah pembayaran uang registrasi (*uang\_regis*) dan calon mahasiswa yang melakukan registrasi (registrasi=1) menjadi dua golongan yaitu calon mahasiswa dengan rata-rata pembayaran uang registrasi sebesar 10.028.866 dapat diartikan dari kalangan atas dan rata-rata sebesar 3.991.137. Jumlah kluster 3 (k=3) belum dapat dijadikan sebagai pengetahuan disebabkan kelompok calon mahasiswa yang diterima kuliah dikelompokkan menjadi dua bagian

yaitu calon mahasiswa dengan ekonomi atas (10.028.866) dan rendah (3.991.137), selisih antara ekonomi atas dan rendah terlalu jauh jaraknya, maka pengujian dilanjutkan dengan jumlah kluster 4 (K=4), menghasilkan nilai pusat kluster (*centroid*), seperti pada tabel 16.

Tabel 16

Pusat kluster (*centroid*) jumlah kluster 4 (k=4)

Atribut	Cluster_0	Cluster_1	Cluster_2	Cluster_3
Usia	20	19	19	19
Agama	1	1	1	1
Ujian	2	1	1	1
Lulus	2	1	1	1
Registrasi	2	1	1	1
Sex	1	2	2	2
gel daftar	2	2	2	2
gel regis	2	2	2	2
nilai tes	14	82	80	80
uang regis	183718	12823831	3951274	7177444
Jurusan Sekolah	5	6	5	5
Nama Jurusan	3	3	3	3
Informasi	7	7	7	6
Nama Kampus	6	6	7	5
Asal Kota	10	14	11	12
Pekerjaan	2	2	2	2
Jenis Sekolah	2	2	2	2

Cluster 0: 17007 items

Cluster 1: 83 items

Cluster 2: 12919 items

Cluster 3: 336 items

Total number of items: 30365

Hasil *Clustering* dengan jumlah kluster 4 (k=4), pada kluster pertama memiliki persamaan dengan kluster pertama pada *Clustering* dengan jumlah kluster 3 (K=3) dapat dilihat pada tabel 4.2. pada kluster kedua menunjukkan karakteristik rata-rata jumlah pembayaran uang registrasi (*uang\_regis*) sebesar 12.823.831, kluster ketiga rata-rata pembayaran uang registrasi (*uang\_regis*) sebesar 3.951.274 dan kluster keempat memiliki rata-rata jumlah pembayaran (*uang\_regis*) sebesar 7.177.444. Berdasarkan atribut jumlah pembayaran uang registrasi (*uang\_regis*) dan calon mahasiswa yang melakukan registrasi (registrasi=1) *Clustering* dengan jumlah kluster 4 (k=4) memiliki tiga segmentasi calon mahasiswa berdasarkan rata-rata jumlah pembayaran uang registrasi (*uang\_regis*) yaitu calon mahasiswa dari masyarakat ekonomi atas (12.823.831), rendah (3.951.274) dan menengah (7.177.444) *Clustering* tersebut sudah

dapat dijadikan sebagai pengetahuan untuk menentukan strategi promosi.

Uji coba selanjutnya dengan menentukan jumlah kluster 5 ( $k=5$ ), dihasilkan pusat kluster (*centroid*), seperti pada tabel 17.

Tabel 17

Pusat kluster (*centroid*) jumlah kluster 5 ( $k=5$ )

Atribut	Cluster_0	Cluster_1	Cluster_2	Cluster_3	Cluster_4
Usia	20	19	19	19	19
Agama	1	1	1	1	1
Ujian	2	1	1	1	1
Lulus	2	1	1	1	1
Registrasi	2	1	1	1	1
Sex	1	2	2	2	2
gel daftar	2	2	2	2	2
gel regis	2	2	2	2	2
nilai tes	14	81	80	82	80
uang regis	183718	13645268	6645047	9999682	3948634
Jurusan Sekolah	5	5	5	5	5
Nama Jurusan	3	3	3	3	3
Informasi	7	7	6	6	7
Nama Kampus	6	6	5	6	7
Asal Kota	10	16	12	11	11
Pekerjaan	2	2	2	2	2
Jenis Sekolah	2	2	2	2	2

Cluster 0: 17007 items  
Cluster 1: 56 items  
Cluster 2: 320 items  
Cluster 3: 85 items  
Cluster 4: 12897 items  
Total number of items: 30365

*Clustering* dengan jumlah kluster 3 ( $k=3$ ), kluster 4 ( $k=4$ ) dan kluster 5 untuk kluster pertama memiliki karakteristik yang sama, sedangkan kluster kedua sampai ke kluster lima signifikan perbedaannya pada atribut jumlah pembayaran uang registrasi (*uang\_regis*), yaitu pada kluster kedua memiliki rata jumlah pembayaran uang registrasi (*uang\_regis*) sebesar 13.645.268, kluster ketiga sebesar 6.645.047, kluster keempat sebesar 9.999.682, kluster kelima sebesar 3.948.634, pada *Clustering* tersebut terlalu banyak katagori untuk menentukan karakteristik calon mahasiswa.

Berdasarkan pengamatan dan diinginkan segmentasi pasar berdasarkan jumlah uang pembayaran registrasi dari calon mahasiswa yang tidak lulus, dan calon mahasiswa yang memiliki kemampuan membayar biaya registrasi digolongkan menjadi tiga tingkatan yaitu tingkat ekonomi rendah, menengah dan tingkat atas. Maka kluster yang sesuai dengan kriteria berada pada kluster 4 ( $k=4$ ).

## Hasil Data Mining

Berdasarkan hasil uji coba jumlah kluster yang dapat mewakili profil atau karakteristik calon mahasiswa di Politeknik LP3I Jakarta, dipilih *Clustering* dengan jumlah kluster 4 ( $k=4$ ), memiliki tiga segmentasi calon mahasiswa berdasarkan rata-rata jumlah pembayaran uang registrasi (*uang\_regis*) yaitu calon mahasiswa dari masyarakat ekonomi atas (12.823.831), bawah (3.951.274) dan menengah (7.177.444) dapat dilihat pada tabel 4.3, yang dapat dijadikan untuk menentukan strategi promosi.

Karakteristik dan keanggotaan dari setiap kluster berdasarkan, jumlah kluster sebanyak 4 ( $K=4$ ), dapat dilihat pada tabel 18.

Tabel 18

Karakteristik Keanggotaan Kluster

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Calon Mahasiswa Tidak Melakukan Registrasi	Calon Mahasiswa melakukan Registrasi	Calon Mahasiswa melakukan Registrasi	Calon Mahasiswa melakukan Registrasi
Gelombang Pendaftaran, gelombang 1 (41.53%), gelombang 2 (24.52%)	Gelombang Pendaftaran, gelombang 1 (44.4%), gelombang 2 (34.53%)	Gelombang Pendaftaran, gelombang 1 (42.05%), gelombang 2 (31.16%)	Gelombang Pendaftaran, gelombang 1 (50%), gelombang 2 (26.97%)
Gelombang Registrasi, gelombang 1 (43.38%), gelombang 2 (34.94%)	Gelombang Registrasi, gelombang 1 (33.01%), gelombang 2 (34.94%)	Gelombang Registrasi, gelombang 1 (46.22%), gelombang 2 (33.58%)	Gelombang Registrasi, gelombang 1 (4.48%), gelombang 2 (30.34%)
Pembayaran Registrasi: 0 (34%), 1.900.000 (4.95%)	Pembayaran Registrasi: 13.030.000 (12.03%), 11.250.000 (10.84%), 13.100.000 (7.73%)	Pembayaran Registrasi: 4.000.000 (85.87%), 3.200.000 (4.18%), 5.000.000 (7.14%)	Pembayaran Registrasi: 8.000.000 (24.44%), 7.000.000 (18.26%), 7.500.000 (7.87%)
Program Studi Administrasi Perkantoran (35.54%), Keperawatan Akutansi (21.88%), Informatika Komputer (17.58%) dan Administrasi Biotek (9.38%), Teknik Informatika (7.25%)	Program Studi Administrasi Perkantoran (31.33%), Informatika Komputer (28.82%), Keperawatan Akutansi (15.66%) dan Administrasi (12.82%)	Program Studi Administrasi Perkantoran (46.57%), Keperawatan Akutansi (24.67%), Informatika Komputer (12.87%) dan Administrasi (15.89%)	Program Studi Administrasi Perkantoran (28.65%), Keperawatan Akutansi (28.37%), Informatika Komputer (22.19%) dan Administrasi (20.79%)
Biaya (9.64%)	Biaya (7.32%)	Biaya (18.18%), Teknik Informatika (5.34%)	

## Strategi Promosi

Berdasarkan pada pembahasan atribut yang berdampak pada strategi promosi, seperti segmentasi tingkat ekonomi orang tua/keluarga, peminatan terhadap program studi/ konsentrasi serta sumber informasi yang didapatkan oleh calon mahasiswa, maka Politeknik LP3I Jakarta dapat melakukan strategi promosi yang dapat diterapkan di semua kluster, sebagai berikut:

1. Strategi Promosi berdasarkan kemampuan ekonomi rendah dan menengah
  - a. Menerapkan biaya pendidikan yang dapat dijangkau oleh masyarakat

- dengan ekonomi menengah ke bawah.
- b. Memberikan banyak pilihan cara melakukan pembayaran biaya kuliah yang dapat dicicil sesuai dengan kemampuan calon mahasiswa.
  - c. Memberikan Dana Pinjaman Belajar, yang akan dikembalikan oleh calon mahasiswa setelah lulus dan ditempatkan kerja oleh Politeknik LP3I Jakarta, diberikan kepada calon mahasiswa yang yang tidak mampu secara ekonomi tetapi memiliki nilai akademis yang baik/sangat baik.
2. Strategi Peminatan Program Studi/Konsentrasi
- a. Melakukan promosi ke sekolah-sekolah yang linier dengan program studi/konsentrasi di Politeknik LP3I Jakarta.
  - b. Memberikan harga khusus untuk program studi/konsentrasi yang kurang diminati oleh calon mahasiswa, seperti program Hubungan Masyarakat, sekretaris, perbankan syariah, bisnis internasional.
3. Strategi Promosi Berorientasi Pada Sumber Informasi
- a. Menyusun *Tool* Multimedia Presentasi ke Sekolah yang dapat menarik perhatian dan tertarik untuk kuliah di Politeknik LP3I Jakarta
  - b. Menyusun tim presentasi sekolah yang memiliki kemampuan komunikasi
  - c. Promosi yang paling efektif bagi Politeknik LP3I Jakarta adalah melalui pemasaran langsung (direct marketing) seperti presentasi ke sekolah, sekolah binaan, kunjungan ke rumah calon mahasiswa (*Home Visit*).
  - d. Membuat program Member Get Member untuk keluarga/Teman dan diinformasikan kepada mahasiswa/keluarganya, apabila membawa temannya untuk masuk kuliah di Politeknik LP3I Jakarta
- akan mendapatkan insentif sebesar Rp. 500.000 (lima ratus ribu rupiah per mahasiswa registrasi) dan apabila jumlah yang registrasi mencapai 10 orang akan mendapatkan hadiah Smartphone.
- e. Melakukan penataan tampak gedung sebagai sarana promosi dengan memasang atribut promosi yang menarik perhatian.
  - f. Membentuk dan melatih tim promosi yang tangguh memahami *product knowledge*
  - g. Untuk meningkatkan kerjasama dengan sekolah binaan, menyelenggarakan Try Out gratis, menyelenggarakan doa bersama menjelang ujian nasional, donator kegiatan OSIS dan acara pelepasan kelas 12.
4. Strategi Promosi Berorientasi Pada Daerah asal calon mahasiswa hasil *Clustering*
- a. Melakukan promosi dengan memasang spanduk yang disesuaikan dengan tema disesuaikan dengan kalender kegiatan nasional pada daerah-daerah yang mendominasi pada setiap kluster.
  - b. Meningkatkan jumlah Sekolah SMK Binaan untuk daerah-daerah yang memiliki potensi sesuai dengan hasil kluster.
  - c. Membuka *counter* pendaftaran dan informasi di tempat-tempat pusat kegiatan masyarakat seperti Mall, station bis/kereta di masing-masing daerah potensi sesuai hasil *Clustering*.
- Dari data hasil *Clustering* dapat diterapkan bauran pemasaran/*marketing mix*, untuk menentukan strategi promosi yang sesuai dengan masing-masing *cluster* berdasarkan karakteristik keanggotaan sumber informasi yang mendekati ke pusat kluster/*centroid* dan besarnya kontribusi sumber informasi dari setiap kluster (Tabel 4.6).



1. Brosur menjadi bagian dari karakteristik anggota atribut sumber informasi yang besarnya kontribusi pada setiap kluster sebesar 10.87% pada kluster pertama, 6.02% pada kluster kedua, 11.34% pada kluster ketiga dan sebesar 8.71% pada kluster keempat.
2. Spanduk di jalan memiliki kontribusi sebesar 9.73% pada kluster pertama, 8.43% pada kluster kedua, 6.53% pada kluster ketiga dan 8,43% pada kluster keempat.
3. Media Cetak Koran, Majalah, Poster, memiliki kontribusi cukup kecil, tetapi dapat digunakan untuk menginformasikan kepada masyarakat pada sedang dilaksanakan penerimaan mahasiswa baru, pada kluster pertama sebesar 3.38% (Koran, Majalah), kluster kedua 3.61%, kluster ketiga 2.11% dan kluster keempat sebesar 1.12%
4. Televisi, memiliki kontribusi sedikit terhadap sumber informasi, tetapi masih dapat digunakan untuk pencitraan lembaga, besar kontribusi televisi sebesar 1.56% kluster pertama, 1.2% pada kluster kedua, 0.91% pada kluster ketiga, 2.81% pada kluster keempat.
5. Sosial Media, berada pada kluster pertama 2.6%, kluster kedua 1.2% dan kluster 4 sebesar 3.09%.
6. Website, diperlukan sebagai identitas pendidikan tinggi pendidikan wajib memiliki website *official*, pada kluster pertama 1.42%, kluster 2 1.2%, kluster ketiga 1.18% dan kluster keempat 0.56%.
7. Tampak Gedung Kampus, menjadi salah satu sumber informasi yang memberikan kontribusi sebesar 6.2% pada kluster pertama, 4.82% pada kluster kedua, 6.14% pada kluster 3 dan sebesar 9.83% pada kluster keempat.
8. Member Get Membet (MGM) terdiri dari informasi keluarga, orang tua, teman, alumni dapat dijadikan sebagai strategi promosi yang memberikan kontribusi pada setiap kluster, pada kluster pertama .76%, kluster kedua sebesar 27.71%, kluster ketiga sebesar 23.84%, 27.52%.
9. Strategi potongan biaya pendidikan diberikan kepada calon mahasiswa yang melakukan registrasi pada gelombang pertama untuk menarik calon mahasiswa agar melakukan pendaftaran dan registrasi lebih awal, potongan biaya pendidikan menjadi kebijakan Politeknik LP3I Jakarta dapat diberlakukan untuk semua kluster.
10. Strategi Dana Pinjaman dapat diberlakukan untuk calon mahasiswa pada kluster pertama dan ketiga yaitu untuk calon mahasiswa dengan pembayaran 0 (nol) atau dari kelompok ekonomi rendah.
11. Strategi harga khusus (Potongan Biaya Pendidikan) untuk Program Studi/ Konsentrasi yang kurang diminati dan dapat diberikan kepada semua kluster.
12. Untuk menjalin hubungan antara sekolah dengan LP3I, maka perlu adanya kerjasama antara Sekolah dengan Politeknik LP3I Jakarta untuk dapat memberikan keuntungan bagi sekolah dan Politeknik LP3I Jakarta dapat diberlakukan pada semua kluster.
13. Sebagai wujud membangun hubungan dengan pihak sekolah, maka Politeknik LP3I Jakarta menyelenggarakan Seminar, Training untuk para guru dan siswa dapat diterapkan pada semua kluster.
14. Presentasi di Sekolah, merupakan sumber informasi yang ada di setiap kluster dan memberikan kontribusi terhadap penerimaan calon mahasiswa, kluster pertama sebesar 22.28%, kluster kedua sebesar 24.1%, kluster ketiga sebesar 20.6% dan kluster keempat sebesar 18.54%
15. Tele selling, memiliki kontribusi kecil terhadap perolehan calon mahasiswa

baru sebesar 2.23% pada kluster pertama, 1.2% pada kluster kedua, 1.22% pada kluster 3 dan 1.69% pada kluster kedua, maka diperlukan untuk melakukan seleksi terlebih dahulu kelengkapan data calon mahasiswa seperti nomor telepon rumah *handphone*.

16. Strategi langsung berkunjung ke rumah orang tua calon mahasiswa (Hunting, home Visit) kontribusinya yang besar apda cluster pertama sebesar 1.69%, kluster kedua tidak ada, kluster ketiga sebesar 3.72%, sedangkan pada kluster keempat 1.12%.

Dari uraian sumber media informasi diatas, dapat disimpulkan strategi promosi yang dapat diterapkan pada setiap kluster berdasarkan pada bauran pemasaran/*Marketing mix*, seperti pada tabel 19.

Tabel 19  
Strategi Promosi berdasarkan bauran pemasaran/*Marketing mix*

No	Strategi Promosi	Cluster Satu	Cluster Dua	Cluster Tiga	Cluster Empat
1	Periklanan ( <i>advertising</i> ) a. Brosur b. Spanduk di jalan c. Media Cetak Koran, Majalah, Poster d. Televisi e. Sosial Media f. Website g. Tampak Gedung Kampus	✓	✓	✓	✓
2	Pengjualan Personal ( <i>personal selling</i> ) a. Member Get Member (MGM) b. Teman Sebanga	✓	✓	✓	✓
3	Promosi penjualan ( <i>sales promotion</i> ) a. Potongan harga pada Gelombang Pertama b. Dana pinjaman (dibayar setelah mahasiswa kerja) c. Harga Khusus (Potongan Biaya Pendidikan) untuk Program Studi Konsentrasi yang kurang diminati	✓	✓	✓	✓
4	Hubungan masyarakat ( <i>public relation</i> ) a. Sekolah Binaan b. Seminar, training untuk guru dan siswa	✓	✓	✓	✓
5	Pemasaran langsung ( <i>direct marketing</i> ) a. Presentasi di Sekolah b. Tele Selling c. Kunjungan ke rumah calon mahasiswa	✓	✓	✓	✓

Strategi bauran promosi periklanan dapat diterapkan disemua kluster seperti brosur, spanduk, media cetak, televisi, tampak gedung, sosial media dan Internet, berdasarkan jarak kluster mendekati ke pusat kluster, sedangkan iklan pada website dan social media menjadi mutlak harus diterapkan disemua kluster, dikalangan generasi muda sekarang internet sudah menjadi sebuah kebutuhan

dan identik dengan social media (facebook, tweeter, Instagram)

Bauran penjualan personal dengan melaksanakan member get member dapat diterapkan semua kluster kecuali kluster satu dikarenakan pada kluster satu persentasi MGM dibawah 1%, sedangkan penjualan personal dengan melibatkan teman/keluarga menjadi sangat efektif untuk semua kluster semua kluster memiliki persentase diatas 2%.

Bauran promosi penjualan dengan memberikan potongan biaya pada gelombang pertama dapat diterapkan pada semua kluster untuk memberikan daya tarik biaya lebih murah apabila melakukan pendaftaran pada gelombang pertama. Dana pinjaman diterapkan pada kluster satu dan tiga merupakan kelompok pendapat orang tua yang rendah, hal ini untuk memberikan kesempatan kepada calon mahasiswa yang berprestasi secara akademik tetapi kurang mampu dari faktor ekonomi. Harga khusus diterapkan disemua kluster untuk jurusan/konsentrasi yang kurang diminati oleh calon mahasiswa, sehingga jurusan/ konsentrasi yang kurang diminai oleh calon mahasiswa menjadi banyak peminatnya.

Hubungan masyarakat dapat diterapkan disemua kluster seperti kerjasama dengan sekolah binaan, memberikan seminar siswa/guru, training guru, *try out* untuk memberikan wawasan kepada guru dan mempersiapkan siswa untuk lulus ujian sekolah dan nasional, sebagai bentuk *win-win solution* antara Politeknik LP3I Jakarta dengan sekolah dan akhirnya dapat menyampaikan pesan promosi/presentasi produk Politeknik LP3I kepada siswa.

Pemasaran langsung dengan melakukan presentasi sekolah dan *Tele Selling* diterapkan pada semua kluster, rata-rata persentasi pada keempat kluster minimal 20%. Sedangkan kunjungan ke rumah calon mahasiswa dapat diterapkan disemua kluster kecuali kluster dua, karena kluster dua merupakan kelompok pendapatan orang tua dengan ekonomi



atas, dan kesibukan dari orang tua calon mahasiswa dan *sensitive* untuk dikunjungi langsung ke rumah.

## PENUTUP

### Kesimpulan

Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan sebagai berikut :

1. Profil calon mahasiswa yang dihasilkan dari *data mining* dengan algoritma K-Means *Clustering*, jumlah *cluster* terbentuk sebanyak 4 *cluster* ( $k=4$ ) yaitu *cluster* 1 dengan jumlah 17007 calon mahasiswa merupakan kelompok calon mahasiswa yang tidak lulus dengan rata-rata pembayaran uang registrasi sebesar Rp. 1.837.718, *cluster* 2 dengan jumlah calon mahasiswa sebanyak 83 merupakan kelompok calon mahasiswa yang lulus dan melakukan registrasi dengan rata-rata pembayaran uang registrasi sebesar Rp. 12.823.831 dikategorikan sebagai golongan masyarakat ekonomi atas, *cluster* 3 sebanyak 12919 calon mahasiswa baru dengan rata-rata pembayaran uang registrasi sebesar Rp. 3.951.274 dikategorikan sebagai golongan masyarakat ekonomi rendah, sedangkan *cluster* 4 sebanyak 356 calon mahasiswa baru dengan rata-rata pembayaran registrasi sebesar Rp. 7.177.444 digolongkan sebagai masyarakat ekonomi menengah.
2. Profil yang dihasilkan dari *data mining* dengan algoritma K-Means *Clustering* dapat memberikan pengetahuan untuk menentukan strategi promosi di Politeknik LP3I Jakarta. Pengetahuan yang didapat dari hasil *clustering* yaitu calon mahasiswa didominasi dari masyarakat ekonomi rendah dan menengah.  
Bauran penjualan personal dengan melaksanakan member get member dapat diterapkan semua kluster kecuali kluster satu dikarenakan pada kluster

satu persentasi MGM dibawah 1%, sedangkan penjualan personal dengan melibatkan teman/keluarga menjadi sangat efektif untuk semua kluster semua kluster memiliki persentase diatas 2%.

Bauran promosi penjualan dengan memberikan potongan biaya pada gelombang pertama dapat diterapkan pada semua kluster untuk memberikan daya tarik biaya lebih murah apabila melakukan pendaftaran pada gelombang pertama. Dana pinjaman diterapkan pada kluster satu dan tiga merupakan kelompok pendapat orang tua yang rendah. Harga khusus diterapkan disemua kluster untuk jurusan/konsentrasi yang kurang diminati oleh calon mahasiswa, sehingga jurusan/ konsentrasi yang kurang diminai oleh calon mahasiswa menjadi banyak peminatnya.

Bauran hubungan masyarakat dapat diterapkan disemua kluster seperti kerjasama dengan sekolah binaan, memberikan seminar siswa/guru, training guru, *try out* untuk memberikan wawasan kepada guru dan mempersiapkan siswa untuk lulus ujian sekolah dan nasional.

3. Bauran pemasaran langsung dengan melakukan presentasi sekolah dan *Tele Selling* diterapkan pada semua kluster, rata-rata persentasi pada keempat kluster minimal 20%. Sedangkan kunjungan ke rumah calon mahasiswa dapat diterapkan disemua kluster kecuali kluster dua.

## DAFTAR PUSTAKA

- Hidayat, T., dan N. Istiadah. 2011 . Panduan Lengkap Menguasai SPSS 19 untuk Mengolah Data Statistik Penelitian. Jakarta: Media Kita.
- Kotler, Philip. 1997. Manajemen Pemasaran: Analisis, Perencanaan,

Implementasi dan Kontrol. Jakarta: Prenhallindo.

Kotler, Philip, dan Armstrong Gary. 2001. Prinsip-Prinsip Pemasaran. Alih Bahasa Imam Nurmawan. Jakarta: Salemba Empat.

Larose, Daniel T. 2006. Data mining Methods and Models. Hoboken New Jersey: Jhon Wiley & Sons Inc.

Larose, Daniel T. 2005. Discovering Knowledge in Data: An Introduction to Data mining. John Willey & Sons. Inc.

Santosa, B. 2007. Data mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu

Santoso, S. 2010. Statistik Multivariat. Jakarta: Elex Media Komputindo

Supranto, J. M.A. 2000. Statistik: Teori dan Aplikasi Edisi Keenam. Jakarta: Erlangga

Susanto, Sani dan Dedy Suryadi. 2010. Pengantar Data mining menggali pengetahuan dari bongkahan data. Yogyakarta: Andi.