

# Penerapan Data Mining Untuk Menentukan Lokasi Promosi Penerimaan Mahasiswa Baru Pada Universitas Banten Jaya (Metode K-Means Clustering)

<sup>1</sup>Ramdani Budiman, <sup>2</sup>Rudianto

<sup>1,2</sup>Program Studi Teknik Informatika – Universitas Banten Jaya

ramdani.budiman@unbaja.ac.id <sup>1)</sup>, rudianto@unbaja.ac.id <sup>2)</sup>

**Abstrak**—Setiap perguruan tinggi menginginkan peningkatan dalam penerimaan mahasiswa baru pada setiap tahunnya. Sehingga setiap perguruan tinggi memiliki strategi masing-masing dalam menarik minat para calon mahasiswanya, serta menentukan daerah mana saja yang menjadi lokasi potensial dalam melakukan promosi. Penelitian ini bermaksud untuk menentukan lokasi atau daerah mana saja yang berpotensi untuk mendatangkan mahasiswa baru di perguruan tinggi tersebut. Dengan menerapkan data mining dengan metode clustering mengelompokkan item objek penelitian berdasarkan kemiripan sifatnya, sehingga akan diketahui daerah mana saja yang berpotensi untuk mendatangkan mahasiswa baru. Penentuan lokasi promosi penerimaan mahasiswa baru dengan metode data mining akan memberikan dampak yang bagus dan terarah dalam melakukan promosi, sehingga dapat meningkatkan jumlah mahasiswa baru setiap tahunnya.

**Kata Kunci**— Data Mining, Metode Clustering, Promosi

## I. PENDAHULUAN

Persaingan penerimaan mahasiswa baru disetiap perguruan tinggi negeri maupun swasta saat ini semakin pesat, ada yang menghabiskan banyak biaya untuk media promosi dan ada juga yang menawarkan berbagai program seperti murahnya biaya perkuliahan, setelah lulus cepat mendapatkan pekerjaan dan lain sebagainya.

Untuk mendapatkan calon mahasiswa sesuai dengan yang diharapkan, tentunya Universitas Banten jaya akan melakukan promosi ke berbagai tempat baik didalam maupun diluar provinsi. Namun dalam menentukan lokasi promosi secara cepat dan tepat bukanlah hal yang mudah,ada banyak hal yang harus diteliti dan dipertimbangkan sehingga memerlukan waktu yang tidak sedikit. Jumlah mahasiswa yang terdaftar di Universitas Banten Jaya terdiri dari berbagai lulusan SMA negeri maupun swasta serta

berbagai daerah, hal ini juga yang menjadi pertimbangan untuk melakukan promosi pada sekolah atau daerah tertentu agar hasilnya dapat lebih meningkat dari tahun sebelumnya.

Penerapan data mining pada penerimaan mahasiswa baru di Universitas Banten Jaya dapat digunakan untuk mengolah data-data mahasiswa sehingga memudahkan menentukan strategi poromosi, data sekolah mana saja yang telah dikunjungi pada tahun sebelumnya, dan daerah mana saja yang menjadi area potensial dalam melakukan promosi.Hal ini dapat membantu tim PMB untuk dalam pengambilan keputusan dalam hal promosi sertam embantu manajemen dalam mengambil keputusan dan evaluasi.

## II. TINJAUAN PUSTAKA

Ada beberapa kajian tentang definisi promosi menurut para ahli, untuk memahami hal tersebut disini penulis mengutip beberapa definisi

diantaranya adalah Charles W. Lamb (2001) promosi adalah Komunikasi oleh pemasar yang menginformasikan dan meningkatkan calon pembeli mengenai sebuah produk untuk mengetahui suatu pendapat atau memperoleh suatu respon[1].

Sedangkan menurut Irwan Dani promosi (1999) adalah usaha yang dilakukan agar calon pembeli memberi perhatian kepada usaha, barang atau jasa yang ditawarkan untuk kemudian mendorongnya untuk membeli[2]. Definisi lain menurut H. Frazier Moore (2005) mengatakan bahwa Promosi adalah aktivitas-aktivitas atau peristiwa-peristiwa yang direncanakan untuk menjamin dukungan atau pengakuan tentang diri seseorang, produk, lembaga atau gagasan.[3]

#### A. Literatur View

Penelitian Fina Nasari, Surya Darma 2015 “penerapan k-means clustering pada data penerimaan mahasiswa baru (studi kasus : universitas potensi utama)” penulis mengelompokkan data mahasiswa baru tahun ajaran 2014/2015 dengan teknik clustering. Pengelompokkan yang penulis terapkan menggunakan algoritma K-Means Clustering, algoritma K-Means Clustering mampu mengelompokkan data pada kelompok yang sama dan data yang berbeda pada kelompok yang berbeda. Sehingga akan terlihat kelompok data mahasiswa baru tahun ajaran 2014/2015 pada Universitas Potensi Utama yang tidak terstruktur menjadi terstruktur. Tujuan dari penelitian ini adalah menerapkan algoritma K-Means Clustering pada data penerimaan mahasiswa baru tahun ajaran 2014/2015 (studi kasus : Universitas Potensi Utama). Hasil K-Means Clustering yang diperoleh ada dua kelompok, pusat cluster dengan Cluster 1 = 1 ; 1.75; 1.5 dan Cluster 2 = 2.95; 1.65; 1.4, Cluster pertama jika asal sekolah adalah SMA atau Sekolah Menengah Pertama maka rata-rata jurusan yang diambil adalah Sistem Informasi dan kedua jika asal sekolahnya adalah SMK rata-rata jurusan yang diambil adalah Teknik Informatika.[4]

Penelitian Wiwit Agus Triyanto 2015 “algoritma k-medoids untuk penentuan strategi pemasaran produk” Pada penelitian ini akan menggunakan algoritma *K-Medoids* untuk

pengelompokan data penjualan, sehingga akan ditemukan informasi yang dapat digunakan untuk penentuan strategi pemasaran produk yang tepat. Hasil dari penelitian ini menghasilkan 5 *cluster* dengan *cluster* pertama terdiri dari 909 *record* transaksi, *cluster* kedua terdiri dari 166 *record* transaksi, *cluster* ketiga terdiri dari 66 *record* transaksi, *cluster* keempat terdiri dari 132 *record* transaksi, *cluster* kelima terdiri dari 87 *record* transaksi. Strategi pemasaran produk dapat dilakukan dengan melakukan promosi pada *cluster* kelima yang memiliki kombinasi jumlah barang dibeli yang paling tinggi.[5]

Penelitian O. J et al., (2010), Shovon & Haque, (2012) yang berjudul “*Application of k-Means Clustering algorithm for prediction of Students’ Academic Performance*” dan “*Prediction of Student Academic Performance by an Application of K-Means Clustering Algorithm*” menunjukkan bahwa *clustering* dapat digunakan untuk memonitor kinerja mahasiswa di suatu universitas. Metode ini juga dapat digunakan untuk memonitor kinerja persemester dalam meningkatkan prestasi akademik. Penelitian yang dilakukan O. J et al., (2010) menggunakan 79 data mahasiswa untuk uji coba clustering pada Universitas Nigeria, sedangkan penelitian Shovon & Haque, (2012) menggunakan 60 data mahasiswa untuk uji coba penelitiannya.[6]

Penelitian Arora & Badal, (2013) yang berjudul “*Evaluating Student’s Performance Using k-Means Clustering*”, menggunakan algoritma K-Means karena dinilai dapat dengan cepat dan efisien membantu memantau perkembangan kinerja mahasiswa di suatu instansi pendidikan. Jumlah data yang dianalisis adalah 118 data siswa untuk mendapatkan nilai rata-rata mahasiswa tiap semester. Metode ini dapat memainkan peran penting bagi analisis akademik untuk menentukan alasan penurunan kinerja mahasiswa selama semester tertentu sehingga dapat diambil tindakan untuk meningkatkan kinerja tersebut disemester berikutnya.[7]

### III. METODOLOGI

#### A. Jenis Penelitian

Metode pengumpulan data ialah teknik atau cara-cara yang dapat dilakukan oleh peneliti untuk mengumpulkan data (Sudaryono 2010: 125). Data

kuantitatif sendiri dibedakan menjadi dua macam, yaitu deskriptif dan kuantitatif. Data deskriptif diperoleh dari hasil menghitung, sehingga sering disebut juga data nominal. Sedangkan data kuantitatif diperoleh dari hasil eksplorasi atau survey. Data penelitian ini sendiri merupakan hasil eksplorasi terhadap hasil penjualan produk di setiap cabang, survey terhadap volume produksi yang kemudian berhubungan dengan volume pemakaian bahan baku.[8]

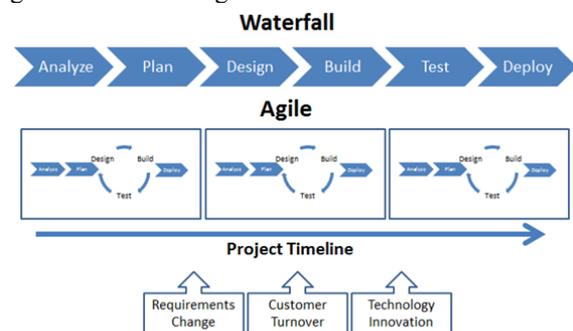
Dalam menentukan kebutuhan data di dalam lingkungan penelitian dapat dilakukan dengan pengamatan secara langsung data laporan pencapaian kinerja perusahaan secara umum. Data yang diperoleh dari data mart (data source) adalah output dari sebuah departemen yang merupakan bagian dari tim atau departemen dalam perusahaan. Setiap detail permasalahan yang dihadapi tim dan mempengaruhi data historis dapat diketahui secara langsung peneliti. Data historis sejak 3 tahun terakhir adalah penilaian terhadap kinerja perusahaan secara tim, dan oleh karenanya keputusan yang akan diambil merupakan keputusan level top management yang akan melibatkan manager tingkat menengah.

Penelitian ini pada dasarnya adalah kebutuhan penting perusahaan dalam rangka mencapai misi dan visinya. Hasil dari penelitian dapat langsung diterapkan sebagaimana kebijakan-kebijakan sebelumnya yang dilakukan secara manual yang lebih identik dengan intuisi direktur/ pengambil keputusan. Dengan adanya kegiatan penelitian ini maka dipandang sebagai tools untuk memberikan tingkat kepercayaan yang tinggi atas keputusan yang diambil dan segera dapat diterapkan oleh perusahaan.

### B. Desain Penelitian

Desain penelitian untuk makalah ini adalah studi kasus, di mana kasus yang ditemukan secara nyata pada lokasi penelitian pada waktu tertentu. Hasil dari penelitian ini nantinya akan menyelesaikan masalah yang sedang dihadapi. Pada siklus hidup pengembangan sistem informasi, tahap implementasi dapat dijalankan dengan strategi pilot project yaitu mengimplementasikan sistem baru pada salah satu bagian dalam organisasi untuk dapat digeneralisasi.

Adapun desain pengembangan aplikasinya diterapkan agile system development di mana tidak ada satu konsep yang disepakati dalam pengembangannya. Namun demikian pada kenyataannya agile system justru telah mampu melahirkan banyak teknik-teknik praktis dalam pelaksanaannya baik itu temuan baru maupun hasil penelitian lanjutan atas konsep-konsep yang telah ada (Abarahamsson, Salo, Ronkainen, & Warsta, 2002). Penerapan agile system memungkinkan pengembangan aplikasi dengan sejumlah kecil personil yang terlibat. Juga sangat fleksibel dalam hal pengembangan yang bersifat parsial dan insidental. Kaitannya dengan penelitian ini, pengembangan yang dilakukan menganut implementasi pilot project yang bersifat parsial, pelaksanaannya pun tidak menutup kemungkinan secara parallel, sehingga perbaikan di satu bagian akan segera diikuti pada bagian lainnya. Ini sangat sesuai dengan konsep agile system. Dalam Gbr.1 dapat dilihat perbedaan fleksibilitas pengembangan sistem dengan waterfall dan agile.



Gbr. 1 Waterfall vs Agile

### C. METODE ANALISIS

Metode analisis yang diterapkan dalam penelitian ini adalah sebagaimana dijelaskan Maimon & Rokach (2010) yaitu melalui tahapan sebagai berikut :

#### 1. Domain Understanding dan KDD Goals

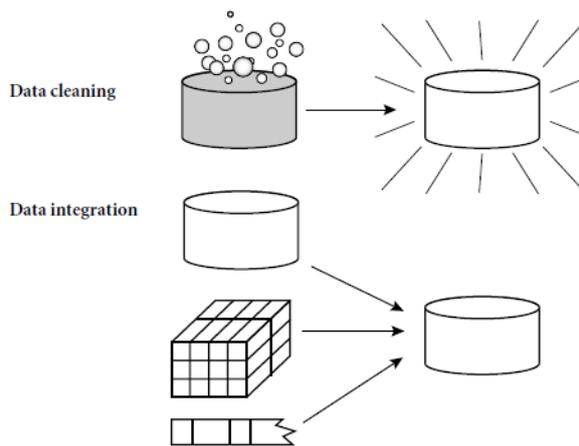
Pada tahap awal ini dilakukan usaha pemahaman terhadap tujuan yang ingin dicapai. Database yang tersaji sedemikian kompleks dalam data warehouse (DW) sangat berpotensi memberikan informasi yang berguna. Tujuan yang dimaksud tentunya dapat menyelesaikan akar masalah terhadap kasus yang sedang dihadapi.

#### 2. Selection dan addition

Sangat mungkin terjadi dalam lingkungan BI data yang diperoleh dari DW merupakan huge data (Han & Kamber, Data Mining : Concepts and Techniques, 2006) yaitu jumlah rekord yang banyak, atribut yang menyertai pun kompleks, dan type data yang heterogen. Dengan telah dipahaminya tujuan seperti pada tahap 1, maka yang segera harus dilakukan adalah memilih atribut data yang diperlukan. Atribut seperti tanggal order, tanggal kirim, nomor batch dan volume susut produksi tidaklah memberikan informasi apa-apa terkait dengan tujuan. Oleh karenanya atribut ini sebaiknya dibuang atau diabaikan. Namun demikian ada kalanya perlu ditambahkan sebuah atribut misalnya hasil perhitungan rata-rata, dan kegiatan ini disebut addition.

### 3. Preprocessing

Dapat dijelaskan mengenai tahapan preprocessing adalah upaya mendapatkan data siap pakai. Upaya pertama adalah data cleansing yaitu semacam pencucian data. Data yang tidak lengkap perlu dibuang, misalnya terdapat beberapa value atribut yang null. Kemudian hasil selection yang hanya menyisakan atribut yang berharga perlu diintegrasikan agar didapat data yang sah seperti dijelaskan pada Gbr. 2 berikut ini.

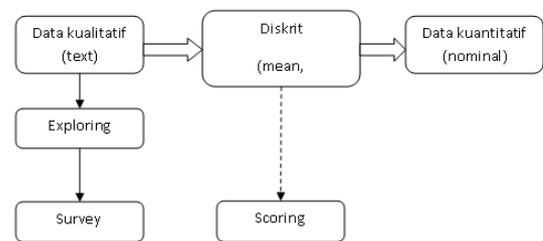


Gbr. 2 Proses *Cleansing* Dan *Integration*

### 4. Transformation

Transformasi atau perubahan bentuk dilakukan untuk mendapatkan data yang konsisten. Meskipun dalam DW sudah dilakukan ETL yang memungkinkan adanya penyeragaman format,

perlu diingat bahwa penyeragaman dalam DW tersebut umumnya berlaku untuk satu tabel saja. Sedangkan tabel yang diterima dari departemen lain bisa jadi telah mengalami editing format sebagai bagian dari kebutuhan pengolahan data mereka. Bukti transformasi dalam penelitian ini misalnya penggunaan tanda pemisah desimal. Oleh karena data yang diterima dapat berupa format xls dan ods, besar kemungkinan tanda pemisah ini menjadi masalah. Ketika disepakati pemisah desimal adalah “dot” (International style) maka harus dikenakan pada seluruh data yang ada. Tentu dapat dipahami selanjutnya bahwa “2.000” tidak sama dengan “2,000”. Selain itu transformasi juga dilakukan dalam hal penempatan jumlah digit di belakang koma untuk merepresentasikan bilangan desimal. Penjelasan ini dapat dilihat pada Gbr.3.



Gbr.3 Transformasi Data Kualitatif Ke Data Kuantitatif

### 5. Data mining : menentukan type data mining

Dengan memperhatikan keunggulan masing-masing type data mining yang telah dipelajari, maka memberikan keyakinan dalam menentukan type (metodologi) data mining. Dalam penelitian ini clustering dinilai paling efektif untuk menyelesaikan masalah yang ingin dipecahkan. Informasi penting yang diperoleh adalah segmentasi. Informasi ini memberikan gambaran bahwa produk yang dipertahankan bukan dalam satu segmen (cluster) yang sama.

### 6. Data mining : menentukan algoritma data mining

Algoritma adalah urutan kerja dari sebuah proses yang saling memiliki ketergantungan. Artinya proses secara berurutan dilaksanakan setelah menyelesaikan proses-proses sebelumnya. Dalam penelitian ini dipilih algoritma K-Means mengingat kemampuannya menganalisa sejumlah data dengan atribut yang relatif kompleks. Kunci

dari algoritma ini adalah perhitungan jarak Euclidean.

7. Data mining : implementasi

Tahapan ini adalah penerapan algoritma yang dipilih untuk menganalisa sejumlah data set (data history) dan dibuktikan juga pada data test.

8. Evaluasi dan interpretasi

Output dari analisa data set menggunakan metodologi yang ditentukan selanjutnya dilakukan evaluasi. Tahapan ini dianggap penting dalam rangka memberikan informasi kepada pengguna tentang tingkat kepercayaan, menggunakan metode Precision, Recall dan tingkat akurasi menggunakan metode F-1.

9. Visualisasi dan integrasi

Tahap terakhir dari serangkaian analisa penelitian ini adalah visualisasi dan integrasi. Dengan pengetahuan yang diterima dapat dilaporkan secara komprehensif kepada pengguna rekomendasi yang diusulkan berdasarkan penelitian. Dengan demikian pengetahuan dapat di-deliver secara efektif dari sistem kepada pengguna untuk menunjang pengambilan keputusan [9].

IV. ANALISIS DAN PERANCANGAN

A. Analisa Masalah

Penulis menyimpulkan permasalahan sebagai berikut :

1. Sebaran mahasiswa yang mendaftar di Universitas Banten Jaya tidak merata untuk kabupaten/kota yang ada di Provinsi Banten.
2. Penentuan lokasi promosi ke sekolah hanya menggunakan perkiraan dan pengalaman yang sudah ada dari tahun-tahun sebelumnya.
3. Sulitnya menentukan daerah potensial untuk melakukan promosi agar pencapaian jumlah mahasiswa baru yang daftar lebih meningkat setiap tahunnya.
4. Kurangnya informasi terhadap manajemen tentang pencapaian Tim PMB setiap tahunnya.

B. Analisa Sumber Data

Data yang digunakan mahasiswa tahun 2014 dan atribut yang akan digunakan berdasarkan data mahasiswa yang diperoleh.

Tabel 1  
Data pendaftar mahasiswa baru

No	NAMA SEKOLAH	PRODI											
		TIND	TL	TS	SI	TI	TID3	MID3	KAD3	PBI	PKN	PAK	JUMLAH
1	MAN SERANG 1	2	0	1	3	2	0	0	0	1	2	1	12
2	MAN SERANG 2	0	0	1	0	1	0	0	0	0	0	0	2
3	SMAN KOTA SERANG 1	0	0	0	0	1	0	0	0	0	1	0	2
4	SMAN KOTA SERANG 2	0	0	0	1	2	0	0	0	1	0	0	4
5	SMAN KOTA SERANG 3	0	0	0	0	0	0	0	0	1	0	0	1
6	SMAN KOTA SERANG 4	0	0	0	0	0	0	0	0	1	0	0	1
7	SMAN KOTA SERANG 5	0	0	1	2	11	0	0	0	0	1	1	16
8	SMAN KOTA SERANG 6	1	0	0	0	0	0	0	0	0	0	0	1
9	SMA AGRO TAKTAKAN	1	1	0	1	1	0	0	0	0	0	0	4
10	SMA FAHMI AL	0	2	2	2	1	0	2	0	0	0	0	9

C. Analisa Prosesing Data

Adapun tahapan dari analisis preprocessing data dalam penelitian ini adalah sebagai berikut :

1. Ekstraksi data (*data extraction*)  
Dalam tahapan ini data yang berasal dari flat file berformat microsoft excel di-ekstrak, kemudian disimpan kedalam sebuah database agar memudahkan dalam proses pengolahan data.
2. Pembersihan data (*data cleaning*)  
Pada tahap ini data yang memiliki nama lengkap, nomor telephone, alamat lengkap akan di hilangkan.

Tabel 2  
Pembersihan Data

No	NAMA SEKOLAH	PRODI										
		TIND	TL	TS	SI	TI	TID3	MID3	KAD3	PBI	PKN	PAK
1	MAN 1 SERANG	2	0	1	3	2	0	0	0	1	2	1
2	MAN 2 SERANG	0	0	1	0	1	0	0	0	0	0	0
3	SMAN 1 KOTA SERANG	0	0	0	0	1	0	0	0	0	1	0
4	SMAN 2 KOTA SERANG	0	0	0	1	2	0	0	0	1	0	0
5	SMAN 3 KOTA SERANG	0	0	0	0	0	0	0	0	1	0	0
6	SMAN 4 KOTA SERANG	0	0	0	0	0	0	0	0	1	0	0
7	SMAN 5 KOTA SERANG	0	0	1	2	11	0	0	0	0	1	1
8	SMAN 6 KOTA SERANG	1	0	0	0	0	0	0	0	0	0	0
9	SMA TAKTAKAN AGRO	1	1	0	1	1	0	0	0	0	0	0
10	SMA AL FAHMI	0	2	2	2	1	0	2	0	0	0	0

3. Pada tahap ini, data yang digunakan untuk penelitian adalah atribut asal sekolah, dan asal kecamatan, minat prodi. Ketiga atribut tersebut akan digunakan pada analisis data selanjutnya.

D. Analisa Metode Clustering

Langkah-langkah proses analisis menggunakan algoritma k-means adalah sebagai berikut :

- Tentukan banyaknya cluster adalah 3 cluster ( $k = 3$ ) yang akan dibuat.
- Tentukan centroid setiap cluster

Untuk penentuan *centroid* dilakukan dengan cara mengambil data dari tabel yang sudah di sediakan yaitu MAN 2 Kota Serang, SMK Informatika dan SMK 17 Kota Serang. Centroid itu di ambil sebanyak tiga dapat dilihat pada Table 3.

Tabel 3  
Penentuan Centroid Awal

Data Pusat Cluster										
c1 = SMAN 2 SERANG	0	0	1	0	1	0	0	0	0	0
c2 = SMA INFORMATIKA	3	2	4	2	12	0	0	0	4	2
c3 = SMK 17 KOTA SERANG	2	0	0	0	0	0	0	0	0	1

Langkah selanjutnya digunakan untuk menghitung jarak dari masing-masing hasil jarak data pada jumlah  $k$  di setiap *centroid*.

Jarak data dengan *centroid* pertama

$$d(x_i, c_1) = \sqrt{(a_1 - c_{11})^2 + (b_1 - c_{12})^2 + (c_1 - c_{13})^2 + (d_1 - c_{14})^2 + (e_1 - c_{15})^2 + (f_1 - c_{16})^2 + (g_1 - c_{17})^2 + (h_1 - c_{18})^2 + (i_1 - c_{19})^2 + (j_1 - c_{20})^2}$$

$$d(x_1, c_1) = \sqrt{(2-0)^2 + (0-0)^2 + (1-0)^2 + (3-0)^2 + (2-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (2-0)^2 + (1-0)^2}$$

$$d(x_1, c_1) = 4.795831523$$

Jarak data dengan *centroid* kedua

$$d(x_i, c_2) = \sqrt{(a_1 - c_{21})^2 + (b_1 - c_{22})^2 + (c_1 - c_{23})^2 + (d_1 - c_{24})^2 + (e_1 - c_{25})^2 + (f_1 - c_{26})^2 + (g_1 - c_{27})^2 + (h_1 - c_{28})^2 + (i_1 - c_{29})^2 + (j_1 - c_{30})^2}$$

$$d(x_1, c_2) = \sqrt{(2-12)^2 + (0-13)^2 + (1-10)^2 + (3-1)^2 + (2-13)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2 + (2-1)^2 + (1-0)^2}$$

$$d(x_1, c_2) = 17.83255$$

Jarak data dengan *centroid* ketiga

$$d(x_i, c_3) = \sqrt{(a_1 - c_{31})^2 + (b_1 - c_{32})^2 + (c_1 - c_{33})^2 + (d_1 - c_{34})^2 + (e_1 - c_{35})^2 + (f_1 - c_{36})^2 + (g_1 - c_{37})^2 + (h_1 - c_{38})^2 + (i_1 - c_{39})^2 + (j_1 - c_{40})^2}$$

$$d(x_1, c_3) = \sqrt{(2-0)^2 + (0-1)^2 + (1-3)^2 + (3-2)^2 + (2-13)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-1)^2 + (2-0)^2 + (1-0)^2}$$

$$d(x_1, c_3) = 4$$

Untuk seterusnya, hitung jarak pada setiap baris data, dengan rumus  $d(x_1, c_1) < d(x_1, c_2) < d(x_1, c_3)$

Hasil perhitungan dapat dilihat pada Tabel 4.

Tabel 4  
Hasil Perhitungan Iterasi Pertama

No	Nama Sekolah	Centroid	Centroid			Jarak Terpendek
			C1	C2	C3	
1	MAN SERANG	1	4,47213595	11,18034	4,1231106	4,123105626
2	MAN SERANG	2	0	13,0767	2,645751	0
3	SMAN KOTA SERANG	1	1,41421356	13,22876	2,236068	1,414213562
4	SMAN KOTA SERANG	2	2	12,12436	3,316625	2
5	SMAN	3	1,73205081	13,92839	2,44949	1,732050808

6	KOTA SERANG SMAN KOTA SERANG	4	1,73205081	13,92839	2,44949	1,732050808
7	SMAN KOTA SERANG	5	10,2956301	6,403124	11,44552	6,403124237
8	SMAN KOTA SERANG	6	1,73205081	14	1,414214	1,414213562
9	SMA AGRO TAKTAKAN	2	12,92285	2,236068	2	
10	SMA AL FAHMI	3	3,60555128	12,72792	4,690416	3,605551275

### E. Perancangan Tampilan Program

Perancangan tampilan program dari aplikasi data mining clustering yang akan dibangun adalah sebagai berikut :



Gbr 4 Halaman Menu Utama

Pada Gbr 4 Tampilan awal merupakan tampilan untuk melihat data yang sudah di inputkan sesuai kebutuhan untuk data clustering.



Gbr 5 Halaman Proses Iterasi

Tampilan untuk mengeksekusi data yang akan dibuat iterasi, sehingga ketika tombol proses iterasi di tekan, akan tampil hasil proses iterasinya dapat dilihat pada Gbr 5.



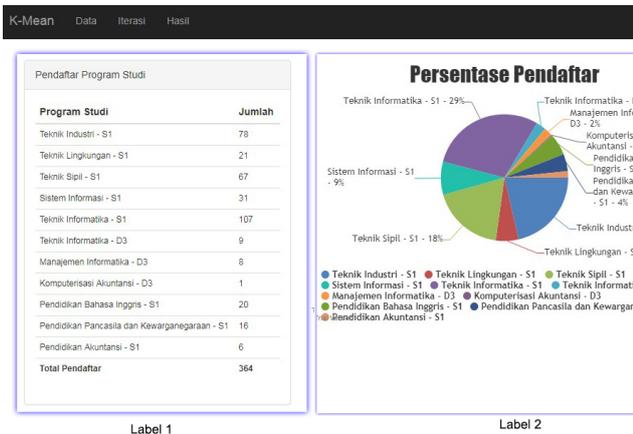
Gbr 6 Halaman Hasil Clustering

Tampilan hasil iterasi adalah tampilan untuk memetakan hasil iterasi dalam bentuk tabel dan bentuk peta dari setiap kecamatan yang terdaftar. Dan akan di simpulkan hasil dari cluster tersebut dalam bentuk keterangan hasil cluster.

## V. PEMBAHASAN

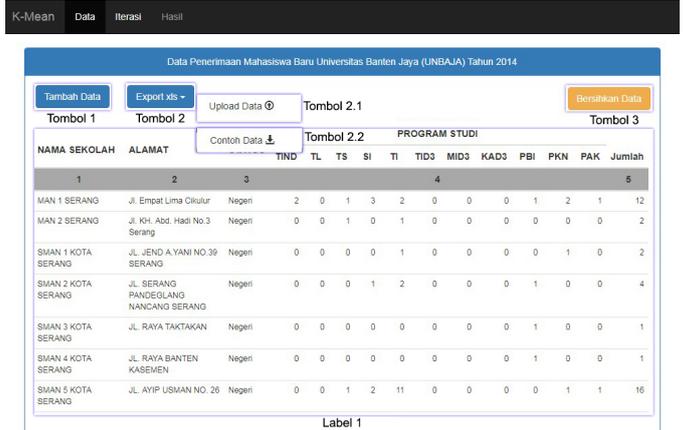
### A. Tampilan Utama

Halaman utama sistem memberikan informasi tentang banyaknya jumlah calon mahasiswa yang sudah mendaftar berdasarkan prodi yang di ambil. Pada Gbr 7 halaman utama label 1 menampilkan jumlah peminat prodi dan total pendaftar dari seluruh prodi-prodi, sedangkan Gbr 7 halaman utama label 2 merupakan informasi yang di tampilkan dalam bentuk presentasi peminat prodi berdasarkan peminat prodi.



Gbr 7 Halaman Utama

tambah data, tombol 2 yaitu export data, tombol 3 yaitu bersihkan data dan tabel 1 yaitu hasil dari input data.



Gbr 8 Menu Data

Fungsi dari tombol 1 yaitu tambah data pada Gbr 8 menu data digunakan untuk menambahkan data secara manual. Pada saat tombol tambah data di klik akan tampil halaman pada Gbr 9 tambah data, admin bisa menambahkan pada kolom nama sekolah, alamat sekolah, status sekolah dan jumlah peminat dalam prodi. Gbr 8 tombol 2.1 *upload* data digunakan mengambil data dan tampilkan kedalam halaman menu data dan tombol 2.2 digunakan mengambil contoh data yang akan ditampilkan pada Gbr 11 contoh data. Sedangkan tombol 3 pada Gbr 8 bersihkan data digunakan untuk menghilangkan semua data informasi di halaman data Gbr 8 label 1.

Gbr 9 Tambah Data

### B. Menu Data

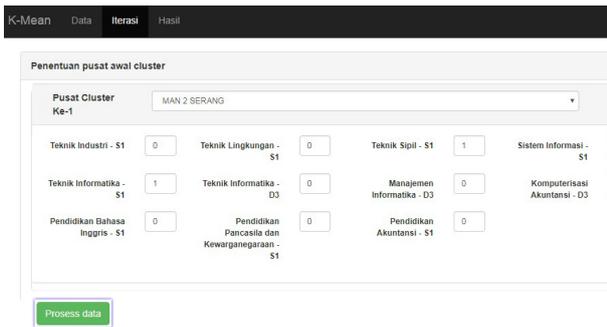
Pada Gbr 8 halaman menu data terdapat tiga tombol dan kolom data yaitu tombol 1 yaitu

1	NAMA SEKOLAH	ALAMAT	STATUS	TIND	TL	TS	SI	TI	TID3	MID3	KAD3	PBI	PKN	PAK	JUMLAH
2	MAN 1 SERANG	Jl. Empat Lima Cikukur	Negeri	2	0	1	3	2	0	0	0	3	2	1	12

Gbr 10 Contoh Data

C. Menu Iterasi

Pada Gbr 11 halaman iterasi ini admin memilih data pertama yang digunakan untuk cluster pertama, data kedua digunakan untuk cluster kedua, dan data ketiga digunakan untuk cluster ketiga. Kalau sudah ditentukan masing masing cluster yang diambil dari halaman tambah data, langkah selanjutnya adalah proses iterasi.



Gbr 11 Halaman Iterasi

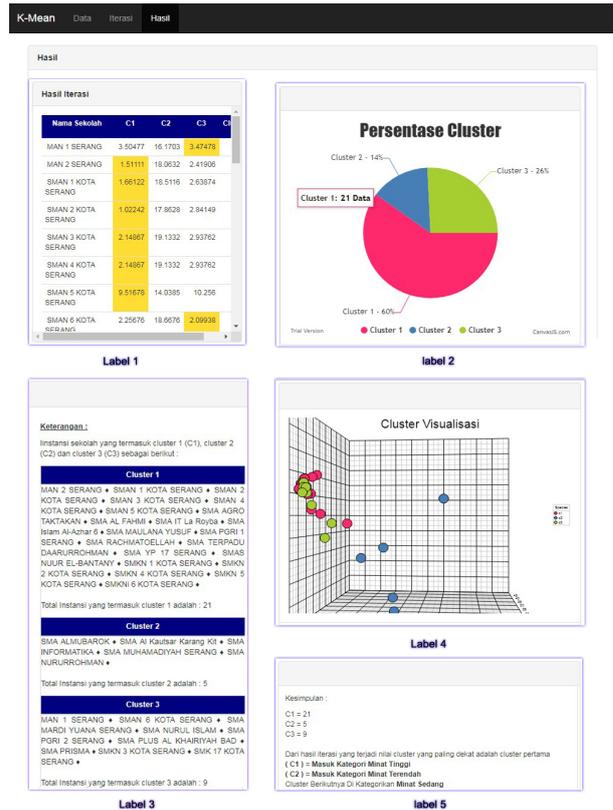
Setelah proses penentuan titik awal lokasi yang digunakan untuk iterasi, maka eksekusi dengan klik tombol proses data, tombol proses data inilah yang akan memproses semua data. Setelah selesai iterasi akan ada warna kuning dimana dari tiga cluster memiliki nilai titik terdekat.



Gbr 12 Proses Iterasi

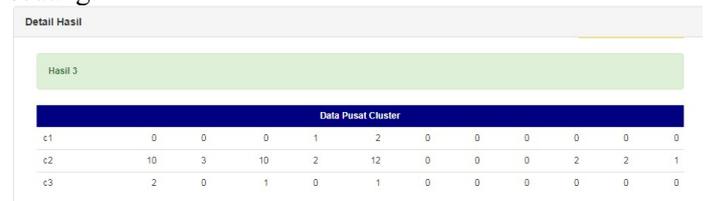
D. Menu Hasil

Gbr 13 menampilkan informasi-informasi dari hasil iterasi yang terjadi dari pusat cluster.



Gbr 13 Tabel dan Pai Informasi

Label 1 memfilter titik terdekat dari beberapa data yang sudah di hitung, dari setiap titik terdekat akan diberikan warna kuning agar menandakan bahwa angka tersebut adalah angka dari titik terdekat cluster. Label 2 memberikan informasi titik terdekat dan dikumpulkan dalam bentuk presentasi dan pembeda warna dari setiap cluster. Label 3 mengambil data dari label 1 yang diberikan warna kuning dan dikelompokkan berdasarkan instansi setiap cluster. Label 4 transformasi dari label 3 dan disajikan dalam bentuk 3D. Sedangkan label 5 mengumpulkan data dari label 4 untuk memberikan kesimpulan dalam bentuk informasi berdasarkan cluster yang tertinggi, terendah dan sedang.



Gbr 14 Cluster Terakhir

Gbr 14 menampilkan hasil akhir dari iterasi terjadi sebanyak 3 kali. Cluster 1 dapat diartikan sebagai kelompok peminat tinggi, Cluster 2 dapat diartikan sebagai kelompok terendah, Cluster 3 dapat diartikan sebagai kelompok peminat sedang.

## VI. KESIMPULAN

Berdasarkan hasil Analisa yang di dapat dengan metode *clustering k-means* dalam menentukan pengelompokan wilayah instansi pendaftar yang berada di daerah kota serang, dapat di simpulkan bahwa:

1. Setelah dilakukan pengelompokan data calon mahasiswa melalui persebaran berdasarkan wilayah menggunakan *K-Means clustering* terbentuk menjadi tiga *cluster*. *Cluster* pertama kategori tinggi, *cluster* kedua kategori rendah dan *cluster* tiga kategori sedang.
2. Strategi promosi bagi calon mahasiswa baru yang tepat sasaran untuk setiap wilayah berdasarkan *cluster* terbentuk
3. Setelah terbentuk *cluster* dengan metode *k-means* maka mengirim tim yang sesuai dengan program studi yang paling banyak diminati dan melakukan promosi berdasarkan peminatan calon mahasiswa dengan melakukan penyelarasan
4. Setelah didapatkan hasil dari pengelompokan berdasarkan *cluster* maka sistem akan menampilkan hasil pengelompokan wilayah calon mahasiswa dalam bentuk *dashboard*.

## DAFTAR PUSTAKA

- [1] Charles W. Lamb, "*Pemasaran*" (Jakarta: Salemba Empat, 2001), hlm. 146
- [2] Irwan Dani, "*Bagaimana Memperbaiki Pemasaran Anda*" (Jakarta: Freidrich Ebert Stiftung, 1999) hlm. 66
- [3] H. Frazier Moore, "*Humas : Membangun Citra Dengan Komunikasi*" (Bandung : Remaja Rosdakarya, 2005) Hlm. 5
- [4] Darma, Surya., Nasari, Fina., "*Penerapan k-means clustering pada data penerimaan mahasiswa baru (studi kasus : universitas potensi utama).*", Seminar Nasional Teknologi Informasi dan Multimedia 2015 STMIK AMIKOM Yogyakarta, 6-8 Februari 2015., ISSN : 2302-3805.
- [5] Triyanto, Agus, Wiwit., "*Algoritma K-Medoids Untuk Penentuan Strategi Pemasaran Produk*", Jurnal SIMETRIS, Vol 6 No 1 April 2015 ISSN: 2252-4983.
- [6] I. C.I., Obagbuwa., O.O. Oladipupo., O.J. Oyelade., "*Application of k-Means Clustering algorithm for prediction of Students' Academic Performance*", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, o. 1, 2010.
- [7] Arora, Kumar, Rakesh., Badal, Dharmendra, Dr., "*Evaluating Student's Performance Using k-Means Clustering.*", IJCST Vol. 4, ISSue 2, April - June 2013.
- [8] Sudaryono, "*Metodologi Riset di Bidang IT: Panduan Praktis, Teori dan Contoh Kasus*". Yogyakarta, Andi Offset, 2015
- [9] Maimon, O., & Rokach, L.. "*Data Mining And Knowledge Discovery Handbook (Second Edition)*". New York: Springer, 2010