

**TEKNIK CLUSTERING DENGAN ALGORITMA K-MEDOIDS UNTUK
MENANGANI STRATEGI PROMOSI DI POLITEKNIK TEDC BANDUNG****Novita Lestari Anggreini**

Prodi Pendidikan Teknik Komputer, Politeknik TEDC Bandung

E-mail : vienovie@poltektedc.ac.id**ABSTRACT**

This research was made based on promotional data that was conducted in 2015-2017. The unequal distribution of information about Politeknik TEDC Bandung can be seen in the data collected by two methods, first direct interviews to students from 2015-2017 and data taken at the time the student enrolled at Politeknik TEDC Bandung. In 2015-2017 there were several promotions carried out by Politeknik TEDC Bandung, including through Tribun Jabar newspaper, promotions to several educational institutions, installing billboards at several places, giving information about Politeknik TEDC Bandung on website, collaboration with other collages through alumni who work in industry or in schools, and last government scholarship programs. Unequal distribution of information to prospective students is one of the main problems of this research being made so that the promotions that have been carried out so far can be more efficient and effective. To be able to carry out promotions that are more effective and efficient, the researcher will process all data to produce a better promotional strategy. Data processing is carried out by grouping using the K-Medoids algorithm on the data of prospective students including Name, Place and Date of Birth, Address, Religion, Telephone Number, School Origin, Selection of Study Programs and who provide information. Which data is obtained from promotion activities carried out from 2015 until 2017. After cleaning and repairing the data, a detailed and more specific dataset can be obtained for grouping the data. If necessary, there is a possibility that some data will be discarded to be more effective in testing the clustering process. Clustering results are made in the form of a chart bar which shows that the most registrants at the TEDC Polytechnic in Bandung are from West Java Province, while the City of Sumedang is the least registrant.

Keywords: *k-medoids, promotion, clustering.*

INTISARI

Penelitian ini dibuat berdasarkan data promosi yang telah dilakukan pada tahun 2015-2017. Belum meratanya penyebaran informasi mengenai Politeknik TEDC Bandung dapat dilihat pada data yang telah dikumpulkan baik melalui wawancara langsung kepada Mahasiswa angkatan 2015-2017 maupun data yang diambil pada saat mahasiswa tersebut mendaftarkan diri ke Politeknik TEDC Bandung. Pada tahun 2015-2017 ada beberapa promosi yang dilakukan oleh Politeknik TEDC Bandung, antara lain melalui Tribun Jabar, promosi ke beberapa instansi pendidikan, memasang baliho di beberapa titik, memberikan informasi mengenai Politeknik TEDC Bandung pada laman web, kerjasama dengan perguruan tinggi lain, melalui alumni yang bekerja di industri maupun disekolah-sekolah, dan program beasiswa dari pemerintah. Tidak meratanya penyebaran informasi ke para calon mahasiswa menjadi salah satu penyebab utama penelitian ini dibuat agar promosi-promosi yang telah dilakukan selama ini bisa lebih efisien dan efektif. Untuk dapat melakukan promosi-promosi yang lebih efektif dan efisien maka peneliti akan mengolah semua data yang ada sehingga menghasilkan strategi promosi yang lebih baik lagi. Pengolahan data yang dilakukan adalah dengan cara pengelompokan menggunakan algoritma K-Medoids terhadap data para calon mahasiswa antara lain Nama, Tempat dan Tanggal Lahir, Alamat, Agama, Nomor Telepon, Asal Sekolah, Program Studi Pilihan dan Pembawa Informasi, yang mana data ini diperoleh dari kegiatan promosi yang dilakukan dari tahun 2015 hingga 2017. Setelah dilakukan

pembersihan dan perbaikan data, kemudian dapat diperoleh dataset yang lebih detail dan lebih spesifik dalam pengelompokan data. Bila diperlukan maka ada kemungkinan beberapa data yang akan dibuang agar lebih efektif dalam pengujian proses klasterisasi. Hasil klasterisasi dibuat dalam bentuk chart bar yang menunjukkan bahwa pendaftar terbanyak di Politeknik TEDC Bandung berasal dari Propinsi Jawa Barat, sedangkan Kota Sumedang menjadi pendaftar paling sedikit.

Kata Kunci : k-medoids, promosi, klasterisasi.

PENDAHULUAN

Politeknik TEDC Bandung merupakan salah satu Perguruan Tinggi Vokasi yang ada di Jawa Barat. Politeknik TEDC Bandung berkomitmen memberikan pendidikan yang terbaik kepada calon-calon mahasiswanya. Seiring dengan perkembangan zaman, serta teknologi dan informasi khususnya di lingkungan perguruan tinggi, maka daya saing antar perguruan tinggi pun semakin meningkat untuk mendapatkan mahasiswa baru. Begitu juga Politeknik TEDC Bandung yang selalu melakukan promosi-promosi ke seluruh Indonesia untuk meningkatkan jumlah penerimaan mahasiswa baru. Dilihat dari promosi-promosi yang dilakukan selama ini, dapat diketahui permasalahan yang terjadi bahwa kurang meratanya informasi mengenai Politeknik TEDC Bandung sehingga masih banyak sekali yang tidak atau belum mengetahui mengenai Politeknik TEDC Bandung.

Pada tahun 2015-2017 ada beberapa promosi yang dilakukan oleh Politeknik TEDC Bandung, antara lain melalui Tribun Jabar, Presentasi ke sekolah-sekolah, memasang baliho di beberapa titik, memasang informasi mengenai Politeknik TEDC Bandung di website, kerjasama dengan perguruan tinggi lain, melalui alumni yang bekerja di industri maupun di sekolah-sekolah, dan program beasiswa dari pemerintah. Kegiatan-kegiatan promosi tersebut mengakibatkan tidak meratanya penyebaran jumlah mahasiswa baru di tiap-tiap program studi yang ada di Politeknik TEDC Bandung.

Ada beberapa penelitian terdahulu yang menjadi referensi dalam penelitian ini antara lain :

1. Asril, dkk [1] pada penelitiannya memiliki tujuan untuk menentukan strategi promosi yang tepat agar dapat mengurangi biaya promosi dan mencapai sasaran promosi yang tepat pula, sehingga menghasilkan informasi penting yang dapat mendukung strategi promosi dalam mendapatkan calon mahasiswa baru. Penelitian Asril, dkk (2015) menjadi motivasi bagi peneliti untuk dapat menentukan strategi yang dapat menarik calon mahasiswa baru sehingga tepat sasaran agar dapat mengurangi biaya promosi yang terkadang diluar anggaran.

2. Pada penelitian Ong Johan Oscar [2] bertujuan untuk menunjukkan bahwa hasil dari pengolahan data mahasiswa membantu pihak Marketing President University dalam melakukan pemasaran dan mencari calon mahasiswa baru dari berbagai kota di Indonesia, dan hasilnya cukup efisien dan efektif. Metode pengolahan data yang digunakan dalam penelitian Ong (2015) ini adalah Data Mining, data yang telah didapatkan, pada tahap pengumpulan data kemudian diolah untuk mengetahui pola dari data-data tersebut sehingga dapat diambil informasi yang tersembunyi.

3. Adanya data dalam skala besar yang memungkinkan adanya Data Mining dengan teknik pengelompokan atau klasterisasi teori ini disebut dalam Penelitian Nugraha, dkk [3]. Pada Penelitian di Politeknik TEDC Bandung data yang digunakan sudah cukup banyak, sebanyak 1411 dataset. Sehingga dimungkinkan untuk dilakukan teknik clustering.

4. Digunakannya metoda pengelompokan dan peramalan pada penelitian tentang memprediksi nilai APBD dimasa depan dengan menggunakan algoritma *K-Means Clustering* dengan mengklasifikasikan dan mengidentifikasi data anggaran daerah [4].

Dari beberapa penelitian yang telah dilakukan, peneliti merasa tertarik untuk melakukan penelitian menggunakan Teknik Clustering untuk menangani Strategi Promosi di Politeknik TEDC Bandung.

Data Mining

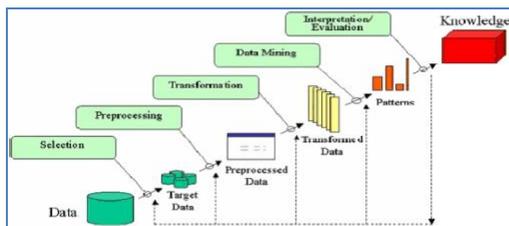
Istilah Data Mining digunakan untuk menjabarkan atau memaparkan penemuan ilmu pengetahuan di dalam database. Data Mining adalah proses yang menggunakan teknik statistik, matematika, *Artificial Intelligence* (AI), dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai database besar [5].

Data Mining adalah beberapa proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Di mana ketika hasil

yang didapat dari proses penggalian tersebut akan membentuk pola-pola berupa kumpulan data, yang sering disebut dengan pengenalan pola (*pattern recognition*). Ada beberapa bagian didalam data mining salah satunya adalah pengenalan pola.

Data Mining juga bisa disebut *knowledge discovery* adalah proses pengambilan pola pada data yang akan di proses lalu keluaran tersebut berupa informasi yang sangat penting dan berharga. Dimana data merupakan kumpulan fakta dan dapat memberikan gambaran, jadi setiap kita melakukan sesuatu dalam pengambilan data maka data tersebut tersimpan dan pola-pola data itu akan diteliti secara manual sehingga kita bisa mengetahui hal-hal yang akan terjadi. Data mining sering juga disebut *Knowledge Discovery in Database* (KDD), yaitu kegiatan yang meliputi pengumpulan data, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [6].

Logika proses *Knowledge Discovery in Database* ditunjukkan pada gambar di bawah ini.



Gambar 1. Aliran Informasi dalam Data Mining [9]

Dapat dijelaskan sebagai berikut :

1. Data Selection

Pemilihan data (seleksi data) dari beberapa data operasional perlu dilakukan sebelum memulai tahap penggalian informasi dalam *Knowledge Discovery in Database*. Data hasil seleksi yang akan digunakan untuk proses data mining dapat disimpan terpisah dari basis data operasional.

2. Pre-processing/Cleaning

Perlunya proses pembersihan data sebelum proses data mining dapat dilaksanakan yang menjadi fokus dalam *Knowledge Discovery in Database*.

3. Transformation

Perubahan atau transformasi pada data yang telah dipilih disebut juga dengan Coding, sehingga data tersebut sesuai untuk proses data mining. Proses coding didalam *Knowledge Discovery in Database* merupakan sebuah proses berpikir untuk menciptakan sesuatu

kreasi dan sangat tergantung pada jenis atau pola informasi yang akan dicari di basis data.

4. Data Mining

Proses mencari pola atau informasi menarik dari dalam data terpilih dengan menggunakan teknik atau metode tertentu sering disebut Data Mining. Teknik, metode, atau algoritma dalam data mining sangat beragam. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* secara penuh.

5. Interpretation/Evaluation

Tahap ini merupakan proses yang disebut *interpretation/Evaluation*. Pada Tahap ini mencakup pemeriksaan apakah pola atau informasi data yang ditemukan bertentangan dengan fakta atau hipotesis yang telah ada sebelumnya.

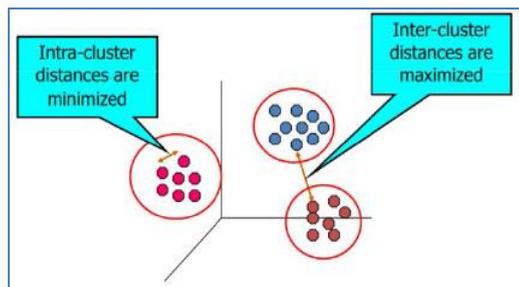
6. Knowledge Presentation (Persentasi Pengetahuan)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh oleh pengguna.

Clustering

Clustering merupakan suatu metode pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kesamaan (kemiripan). *Cluster* adalah kumpulan beberapa *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan beberapa *record* di *cluster* yang lain. *Clustering* berbeda dengan klasifikasi, yaitu tidak ada variabel target dalam *clustering*. *Clustering* tidak mencoba untuk melakukan klasifikasi, memperkirakan, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma *clustering* mencoba untuk melakukan pembagian terhadap semua data menjadi beberapa kelompok yang memiliki kesamaan/ kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai lebih besar, sedangkan kemiripan dengan *record* di kelompok lain akan bernilai lebih kecil [5].

Clustering melakukan pengelompokan data yang didasarkan pada banyaknya persamaan antar objek, oleh karena itu klusterisasi digolongkan sebagai metode *unsupervised learning* bukan *supervised learning* [6]. Logika dari konsep clustering ditunjukkan pada Gambar 2 di bawah ini.



Gambar 2. Konsep Clustering (Han, 2006)

K-Medoids

K-Medoids merupakan salah satu metode partitional *Clustering* yang meminimalkan jarak antara titik berlabel berada dalam *cluster* dan titik yang ditunjuk sebagai pusat klaster itu. Tidak sama dengan algoritma *K-Means*, *K-Medoids* memilih data points sebagai pusat (*medoids*).

K-Medoids adalah teknik partisi klasik *Clustering* yang mengelompokkan data set dari ni objek ke dalam kelompok *k* yang dikenal apriori. Dibandingkan dengan *K-Means*, *K-Medoids* lebih kuat untuk menangani kebisingan (*noise*) dan pencilan (*outlier*) karena meminimalkan beberapa dissimilarities berpasangan, bukan jumlah kuadrat jarak *Euclidean*. Sebuah medoid dapat diartikan sebagai objek cluster yang rata-rata perbedaan untuk semua objek dalam cluster minimal yaitu titik paling berlokasi di cluster.

Strategi Promosi

Terdapat beberapa pengertian mengenai strategi promosi, yaitu bahwa strategi promosi adalah kegiatan perusahaan untuk menggerakkan penjualan dengan mengarahkan komunikasi-komunikasi yang meyakinkan kepada para konsumen [7]. Strategi promosi adalah merencanakan penggunaan yang maksimal dari elemen-elemen promosi seperti periklanan, hubungan kepada masyarakat, penjualan pribadi dan promosi penjualan [8].

Dari beberapa pengertian di atas dapat dilihat bahwa strategi promosi merupakan kegiatan yang direncanakan dengan maksud merayu, mendorong pembeli agar mau atau berkeinginan membeli produk perusahaan, sehingga tujuan untuk meningkatkan penjualan diharapkan dapat dicapai.

Terdapat pula dua macam promosi, yaitu promosi pasif dan promosi aktif. Sesuai topik penelitian ini, promosi pasif dapat digunakan untuk meningkatkan kualitas internal seperti melalui alumni, kakak tingkat, ataupun karyawan yang bekerja di lingkungan Politeknik TEDC Bandung. Sedangkan promosi aktif adalah dengan cara melakukan promosi yang sudah dilakukan oleh tim selama ini, seperti memasang baliho, mencetak

brosur, melakukan kunjungan ke sekolah-sekolah dalam rangka mempromosikan Politeknik TEDC Bandung.

RapidMiner

RapidMiner adalah salah satu *tool* analisis yang berupa suatu *software* untuk pengolahan data mining. *RapidMiner* juga merupakan sebuah lingkungan (*environment*) untuk *machine learning*, *data mining*, *text mining*, dan *predictive analytics*. *Machine learning* merupakan suatu algoritma di mana perilaku komputer berevolusi berdasarkan data empiris, seperti sensor ataupun database. Data mining merupakan suatu proses merangkum pola-pola dari dataset yang besar dengan mengombinasikan metoda statistika, kecerdasan buatan dan database.

PENDEKATAN PEMECAHAN MASALAH

Metode

Metode penelitian yang digunakan pada penelitian ini adalah metode analisis deskriptif dengan pendekatan kuantitatif, artinya penelitian yang dilakukan adalah menekankan analisisnya pada data-data yang terdapat di bagian Promosi Penerimaan Mahasiswa Baru yang bertujuan untuk mendapatkan gambaran yang jelas mengenai strategi promosi yang paling efektif dan sesuai untuk meningkatkan Penerimaan Mahasiswa Baru di Politeknik TEDC Bandung.

Pengumpulan data dalam penelitian ini menggunakan beberapa cara antara lain :

1. Observasi

Observasi yang dilakukan peneliti adalah dengan cara mengamati secara langsung pada bagian Promosi Politeknik TEDC Bandung, serta secara cermat dan sistematis guna mengumpulkan data calon mahasiswa baru tahun 2015 sampai tahun 2017. Peneliti juga mengumpulkan informasi berapa persen ketercapaian promosi penerimaan mahasiswa baru dalam rentang waktu tiga tahun.

2. Wawancara

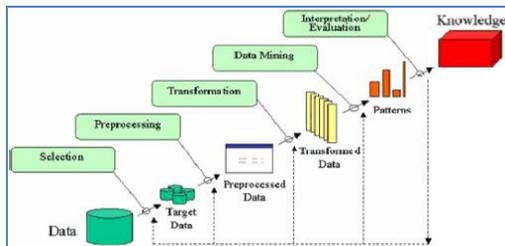
Peneliti juga melakukan wawancara langsung di bagian promosi Politeknik TEDC Bandung terkait pendataan Penerimaan Mahasiswa Baru. Wawancara juga dilakukan oleh peneliti terhadap para mahasiswa tahun 2015 sampai tahun 2017 mengenai seberapa mudah mereka memperoleh informasi mengenai penerimaan mahasiswa baru di Politeknik TEDC Bandung, maupun mengorek keterangan dari

mana mereka memperoleh informasi penerimaan mahasiswa baru tersebut.

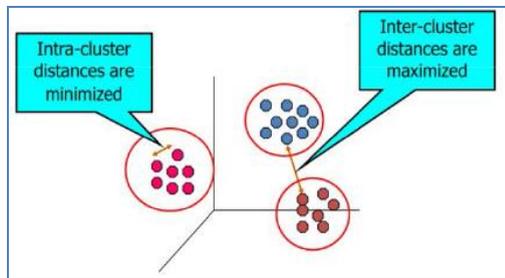
3. Studi Literatur

Peneliti melakukan studi literatur dengan membaca buku-buku, jurnal, dan paper yang terkait dan dapat mendukung penelitian ini.

Data yang akan dianalisis adalah data calon mahasiswa baru dari tahun 2015-2017 yang didapatkan dari bagian Promosi Penerimaan Mahasiswa Baru Politeknik TEDC Bandung yang selanjutnya data tersebut disesuaikan dengan kebutuhan dari penelitian sehingga akan didapat strategi promosi yang lebih efektif dan efisien. Adapun atribut yang dibutuhkan dalam penelitian ini adalah : 1) Kecamatan, 2) Kota, 3) Propinsi, 4) Asal Sekolah, 5) Jenis Sekolah, 6) Program Studi Pilihan, dan 7) Pembawa Informasi.



Gambar 3. Aliran Informasi dalam Data Mining [9]



Gambar 4. Konsep Clustering [9]

HASIL DAN PEMBAHASAN

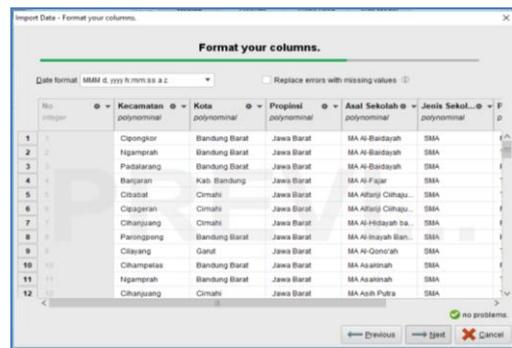
Pada penelitian ini Pengumpulan Data yang diperoleh dari bagian Promosi Politeknik TEDC Bandung adalah data mahasiswa baru dari tahun 2015 sampai tahun 2017 untuk 1.411 mahasiswa. Data yang didapatkan peneliti terdiri dari Nama Pendaftar, Tempat Tanggal Lahir, Alamat, Agama, Nomor HP/Telepon, Asal Sekolah, Program Studi Pilihan, dan Pembawa Informasi.

Setelah dilakukan pembersihan dan penyaringan maka diperoleh dataset yang harus dibuat dengan lebih detail, Data Alamat dapat dipecah lagi menjadi tiga buah atribut, yaitu Kecamatan, Kota, dan Propinsi. Hal ini agar peneliti dapat mengetahui lebih spesifik daerah

mana yang berpotensi besar dalam penyebaran informasi Promosi Penerimaan Mahasiswa Baru di Politeknik TEDC Bandung. Ada juga data yang dikira tidak efektif sehingga oleh peneliti dibuang seperti Nama Pendaftar, Tempat Tanggal Lahir, Agama dan Nomor HP/Telepon. Peneliti juga menambahkan sebuah atribut baru, yaitu Jenis Sekolah, yang dibedakan menjadi SMA dan SMK, dengan alasan untuk mengetahui jenis sekolah mana yang lebih banyak mendaftar di Politeknik TEDC Bandung.

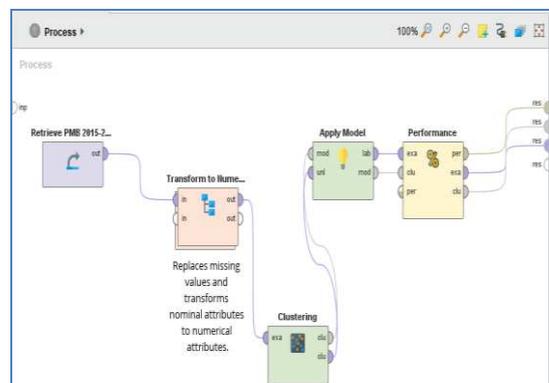
Dataset yang akan diimpor ke dalam Rapidminer adalah *dataset* yang dibuat dengan aplikasi Microsoft Excel, kemudian pilih *file* yang akan digunakan dalam penelitian ini. Setelah file itu diimpor, kemudian tentukan masing-masing indikator atribut yang terdapat di dalamnya. Karena semua atribut berisi lebih dari dua klasifikasi maka tipe dari atribut tersebut adalah *polynomial*.

Pada Gambar 4 di bawah ini dapat dilihat cara pengaturan *dataset* dalam melakukan impor ke dalam *tool RapidMiner*.



Gambar 5. Pengaturan dan Proses Impor Dataset ke dalam RapidMiner

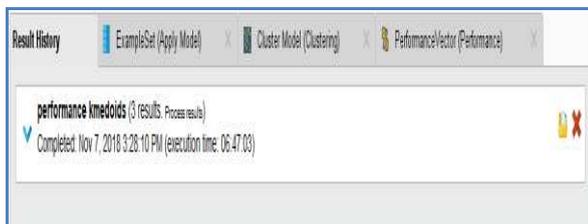
Peneliti membuat desain clustering menggunakan algoritma *K-Medoids* seperti yang ditunjukkan pada Gambar 4.



Gambar 6. Desain Clustering Menggunakan Algoritma *K-Medoids*

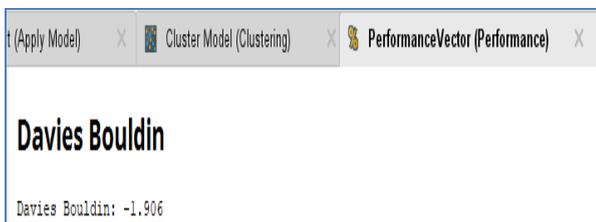
Pada Gambar 5 dapat dilihat bahwa proses pengujian dimulai dengan *retrive* data PMB 2015-2017, artinya dataset disimpan dengan nama PMB 2015-2017 Deal. Dataset itu kemudian disambungkan terlebih dahulu ke *Transform to Numerical* sebelum dilakukan pengujian menggunakan *Algoritma Clustering K-Medoids*. Setelah disambungkan ke *Algoritma Clustering K-Medoids* maka selanjutnya disambungkan lagi ke *Apply Model*, lalu ke *Performance* dan terakhir disambungkan ke *Result*.

Setelah pengujian berjalan maka didapatkan informasi bahwa lama waktu pemrosesan clustering dengan *Algoritma K-Medoids* adalah 6 Jam lebih, seperti yang tercatat pada Gambar 7.



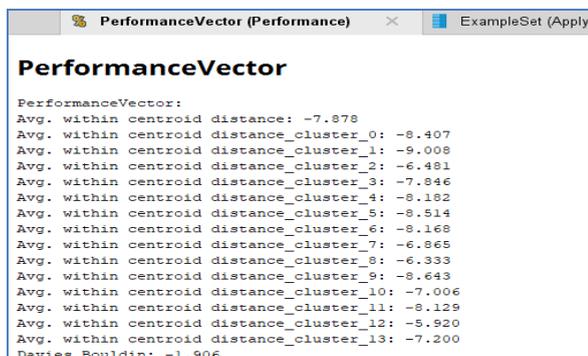
Gambar 7. Result History Clustering menggunakan Algoritma K-Medoids

Hasil dari nilai *Davies Bouldin Index* (DBI) yang terdapat di dalam *Performance Vector K-Medoids* ditunjukkan melalui Gambar 8.



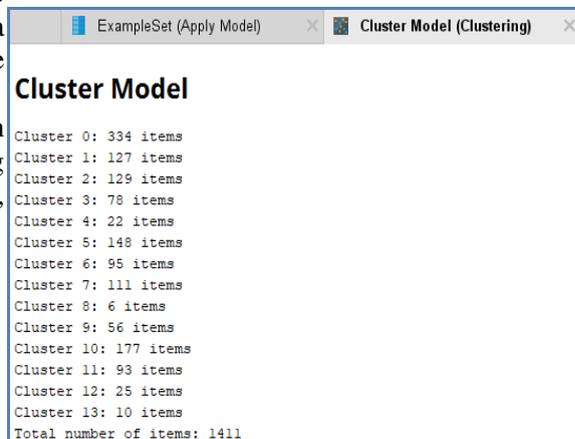
Gambar 8. Davies Bouldin Index untuk Performance Vector K-Medoids

Untuk menunjukkan deskripsi Performance yang menghasilkan *Centroid Distance* pada setiap cluster dapat dilihat pada Gambar 9.



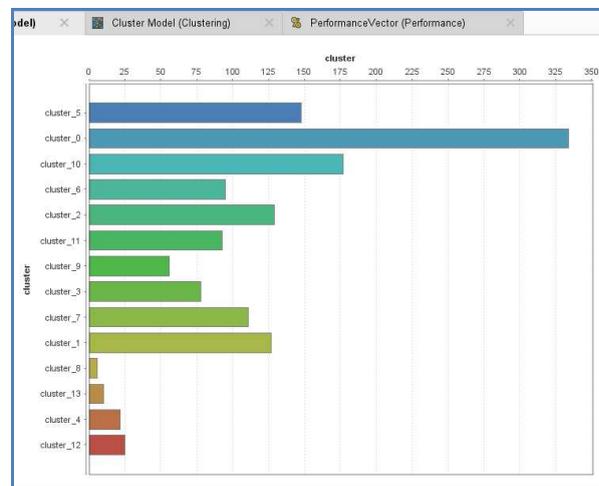
Gambar 9. Centroid Distance K-Medoids

Gambar 10. Menunjukkan bahwa pada penelitian ini *cluster* yang digunakan berjumlah 14, dimulai dari *cluster* 0 dan diakhiri dengan *cluster* 13, total *dataset* yang di-*clustering* sebanyak 1.411 dengan jumlah *items* pada tiap-*cluster*.



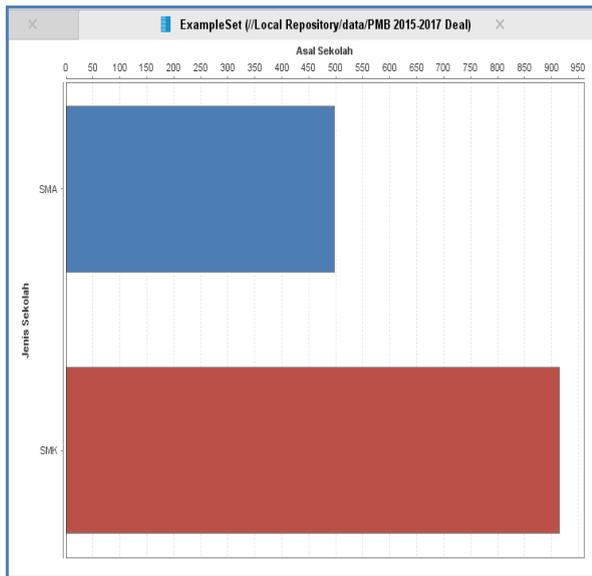
Gambar 10. Hasil Cluster Model K-Medoids

Pada penelitian ini hasil dari *cluster* model dapat digambarkan dalam bentuk *Charts Bars* seperti pada Gambar 11, yang menunjukkan bahwa *cluster* 0 (data Jenis Sekolah) memiliki nilai paling tinggi dan *cluster* 8 (kota Sumedang) memiliki nilai paling rendah.



Gambar 11. Charts Bars Hasil Cluster Model dengan Metode K-Medoids

Charts Bars pada Gambar 12 menunjukkan bahwa cluster SMK memiliki dominasi lebih besar dibandingkan cluster SMA.



Gambar 12. Hasil Cluster Model K-Medoids untuk Data Set Asal Sekolah

Dengan dilakukan analisis terhadap data para calon mahasiswa baru yang mendaftar di Politeknik TEDC Bandung dapat diketahui bahwa asal sekolah SMK yang berasal dari Jawa Barat merupakan pendaftar terbanyak dari tahun 2015 sampai dengan tahun 2017.

KESIMPULAN

Algoritma K-Medoids untuk metode clustering data dapat diterapkan untuk menghasilkan informasi-informasi yang dapat diusulkan kepada Direktur Politeknik TEDC Bandung untuk kedepannya agar dapat mendukung kegiatan promosi penerimaan mahasiswa baru dengan lebih efisien dan efektif. yaitu dengan cara bekerja sama dengan SMK-SMK yang ada di Provinsi Jawa Barat.

SARAN

Hasil clustering data penerimaan mahasiswa baru tahun 2015 sampai dengan tahun 2017 menggunakan 1.411 data. Pada penelitian ini dapat dikembangkan lagi dengan mengangkat topik tentang sistem pendukung keputusan dengan sasaran dapat menghasilkan pertimbangan yang dapat diusulkan kepada Direktur Politeknik TEDC Bandung untuk memilih media yang tepat dalam promosi penerimaan mahasiswa baru.

DAFTAR PUSTAKA

[1] Asroni, Ronald Adrian, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang", *Semesta Teknika*, pp: 76-82, 2015.

- [2] Ong, J. O, Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University, pp: 11-20, 2013.
- [3] Nugraha, D. D., Naimah, Z., Fahmi, M., & Setiani, N, Klasterisasi Judul Buku dengan Menggunakan Metode K-Means. *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, pp: G1-G4, 2016.
- [4] Dony Novalindry, Yeka Hendriyani, Cheng-Hong Yang, Hafilah Hamimi, "The Optimized K-Means Clustering Algorithms To Analyzed the Budget Revenue Expenditure in Padang", *Proceeding of the Electrical Engineering Computer Science and Informatics*, vol. 2, no. 1, pp. 35-40, 2015
- [5] Kusriani dan Lutfi, E.T, *Algoritma Data Mining*, Andi Offset, Yogyakarta, 2009.
- [6] Santosa, B, *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*, Graha Ilmu, Yogyakarta, 2007.
- [7] Moekijat, *Manajemen Pemasaran*, Mandar Maju, Bandung, 2000.
- [8] Lamb, Hair, dan McDaniel, *Pemasaran, Buku 1. Penerjemah David Octarevia*, Salemba Empat, Jakarta, 2001.
- [9] J. Han, M. Kamber, & J. Pei, *Data Mining: Concept and Techniques. Second Edition*, Morgan Kaufmann Publishers, Waltham, 2006.