



Penerapan K-Means Clustering Untuk Strategi Promosi Untuk Meningkatkan Jumlah Perolehan Mahasiswa Baru

Application of K-Means Clustering for Promotional Strategies to Increase the Number of New Students Accepted

*Herlina¹⁾, Siti Mundari²⁾, Elsa Amanda Putri³⁾, Amanda Dian Rahmawati⁴⁾
Program Studi Teknik Industri, Fakultas Teknik
Universitas 17 Agustus 1945 Surabaya, Indonesia

Diterima: Desember 2023; Disetujui: Januari 2024; Dipublikasi: Mei 2024

*Corresponding author: herlina@untag-sby.ac.id

Abstrak

Tiap tahun di sebuah program studi pasti ada mahasiswa baru yang masuk yang berasal dari berbagai wilayah. Penelitian ini bertujuan menganalisis data mahasiswa di prodi Teknik Industri Untag Surabaya berdasarkan sebaran asal SMA, program peminatan saat menempuh studi di SMA, jenis kelamin, jarak kedekatan rumah tempat tinggal dengan kampus, dan jenis pekerjaan orang tua. Data yang diperoleh akan diolah menggunakan algoritma K-Means Clustering. Data-data yang memiliki kesamaan karakteristik akan dikelompokkan menjadi satu kluster, sedangkan data dengan karakteristik yang berbeda akan dikelompokkan dalam kluster yang lainnya. Kluster yang terbentuk akan dibagi tiga yaitu tinggi, sedang dan rendah. Dari kluster yang tinggi akan dianalisis potensi terbesar yang dapat digunakan sebagai strategi marketing untuk meningkatkan jumlah mahasiswa baru di periode mendatang. Dari hasil pengolahan data menggunakan Rapidminer, diperoleh hasil cluster 0 terdiri dari 75 data, cluster 1 sebanyak 31 data, cluster 2 sebanyak 55 data. Dari hasil clustering, atribut yang membedakan cluster secara signifikan adalah jarak tempat tinggal ke kampus. Dari cluster yang tertinggi diperoleh data rata-rata jarak tempat tinggal mahasiswa ke kampus adalah 4,6 km dengan jangkauan antara 0,1 km sampai 10 km. Promosi dapat lebih ditingkatkan pada daerah-daerah dengan jangkauan jarak yang lebih jauh agar menambah perolehan jumlah mahasiswa baru.

Kata Kunci : K-Means, Klastering, Promosi, Mahasiswa

Abstract

Every year, new students enter a department from various regions. This research aims to analyze students' data in the Untag Surabaya Industrial Engineering department based on the distribution of high school, specialization program while studying at high school, gender, proximity distance of the house where they live to the campus, and type of work of their parents. The data obtained will be processed using the K-Means Clustering algorithm. Data that has similar characteristics will be grouped into one cluster, while data with different characteristics will be grouped into another cluster. The clusters formed will be divided into three, namely high, medium and low. From the high clusters, the greatest potential will be analyzed which can be used as a marketing strategy to increase the number of new students in the coming period. From the results of data processing using Rapidminer, the results obtained were that cluster 0 consisted of 75 data, cluster 1 consisted of 31 data, cluster 2 consisted of 55 data. From the clustering results, the attribute that significantly differentiates clusters is the distance from residence to campus. From the highest cluster, data was obtained that the average distance between students' residence and campus was 4.6 km with a range of between 0.1 km to 10 km. Promotion can be further increased in areas with a longer reach in order to increase the number of new students.

Keywords : K-Means, Clustering, Promotion, Students.

How to Cite: Herlina, dkk., (2024). Penerapan K-Means Clustering Untuk Strategi Promosi Untuk Meningkatkan Jumlah Perolehan Mahasiswa Baru. *JIME (Journal of Industrial and Manufacture Engineering)*. 8 (1) : 98-103

PENDAHULUAN

Teknik industri Untag Surabaya saat ini merupakan salah satu program studi yang banyak diminati oleh mahasiswa. Dalam beberapa tahun ini jumlah mahasiswa Teknik industri mengalami peningkatan. Penelitian ini bertujuan untuk menganalisis sampel data statistik dari mahasiswa Teknik Industri Untag Surabaya untuk penentuan strategi marketing bagi penerimaan mahasiswa baru di periode mendatang supaya dapat meningkatkan jumlah perolehan mahasiswa baru. Metode yang akan digunakan dalam penelitian ini menggunakan pendekatan klastering, dimana nantinya klaster yang terbentuk akan dibagi menjadi tiga klaster. Pembagian dalam tiga klaster tersebut untuk mendeteksi klaster yang paling tinggi yang menjadi potensi utama dalam penentuan strategi marketing, kemudian klaster sedang dan klaster rendah. Metode klastering yang akan digunakan dalam penelitian adalah K-Means Klastering. Teknik klastering dengan metode K-Means Clustering ini merupakan salah satu metode data mining. *Data mining* memiliki pengertian lain yaitu *knowledge discovery* ataupun *pattern recognition* di mana artinya *data mining* digunakan untuk mendapatkan pengetahuan yang tersembunyi dari kumpulan data yang berukuran sangat besar (Santosa, 2007). Tujuan utama *data mining* ialah untuk menggali, menemukan, atau menambang pengetahuan dari data atau informasi yang kita miliki. Atribut yang akan digunakan sebagai dasar penentuan klaster adalah jenis kelamin, asal SMA/SMK, program peminatan saat menempuh studi di SMA/SMK, jarak tempat tinggal ke kampus, asal kota, jenis pekerjaan orang tua. Data-data yang memiliki kemiripan sifat akan dikelompokkan menjadi satu klaster sedangkan data yang memiliki sifat berbeda akan dikelompokkan kedalam klaster yang lainnya (Santosa, 2007). Seluruh data yang diperoleh akan diolah menggunakan software Rapidminer. Berdasarkan data-data yang diperoleh dalam masing-masing klaster akan dianalisis berdasarkan atribut-atribut yang dominan untuk kemudian disimpulkan menjadi data input bagi strategi promosi atau marketing untuk memperoleh mahasiswa baru.

METODE PENELITIAN

Penelitian dimulai dengan pengumpulan data yang diperoleh melalui kuesioner yang disebarakan pada sejumlah sampel mahasiswa Teknik industri Untag Surabaya. Jumlah sampel yang akan dikumpulkan akan dihitung menggunakan rumus statistik uji kecukupan data. Kuesioner yang diperoleh akan diolah secara statistik deskriptif untuk memperoleh data komposisi prosentase dari masing-masing atribut yang digunakan. Setelah itu akan dilakukan transformasi data dari atribut-atribut untuk siap diolah menggunakan algoritma K-Means Clustering menggunakan *software Rapidminer*.

Algoritma K-Means Clustering sebagai berikut (Alhapizi, 2020):

- a) Menentukan jumlah klaster.
- b) Menentukan pusat klaster (centroid) awal dengan cara random.
- c) Menghitung jarak masing-masing data ke pusat klaster. Tempatkan data dengan jarak terdekat ke dalam satu klaster.
- d) Hitung kembali pusat klaster dengan menghitung rata-rata dari data anggota klaster yang terbentuk.
- e) Hitung kembali jarak tiap data dengan pusat klaster yang baru, jika pusat klaster baru yang terbentuk sudah tidak berubah maka proses clustering selesai.

Dari hasil klastering akan diperoleh jumlah data pada masing-masing klaster dan dapat dianalisa atribut yang menjadi faktor dominan.

HASIL DAN PEMBAHASAN

Jumlah mahasiswa aktif Program Studi Teknik Industri pada semester gasal 2023/2024 sebanyak 1.035 mahasiswa. Untuk memastikan bahwa sampel yang diambil sudah dianggap cukup mewakili populasi yang ada, maka peneliti melakukan uji kecukupan data menggunakan rumus Slovin (Rakhmawati, 2020). Berikut adalah perhitungan jumlah minimal sampel yang dibutuhkan dalam penelitian ini:

$$n = \frac{N}{1 + Nxe^2}$$

Dimana n adalah ukuran sampel, N adalah ukuran populasi, dan e adalah taraf kesalahan.

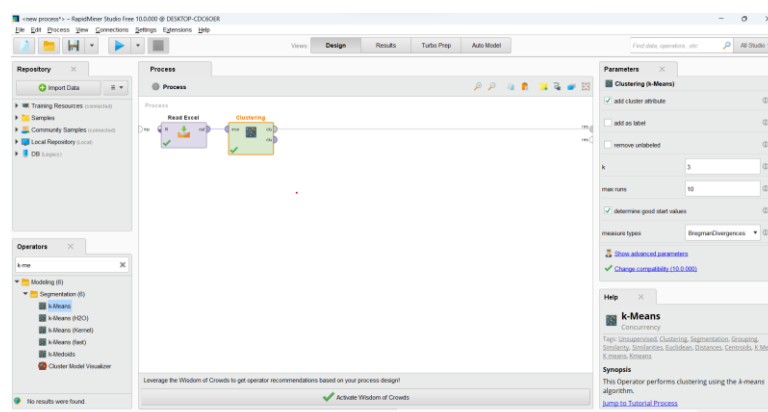
$$n = \frac{1035}{1 + 1035 \times 0,1^2} = 91$$

Berdasarkan perhitungan diatas, jumlah minimal sampel yang dibutuhkan adalah sebanyak 91 orang mahasiswa. Dari hasil pengumpulan data, didapatkan total 161 sampel data sehingga memenuhi uji kecukupan data karena sampel yang diperoleh lebih besar dari nilai minimal 91 sampel.

Atribut yang digunakan dalam penelitian ini antara lain jenis kelamin, jenjang pendidikan terakhir, program peminatan pada pendidikan terakhir, jarak dari tempat tinggal ke kampus Untag Surabaya, asal kota domisili, jenis pekerjaan orang tua.

Dari hasil pengolahan data secara statistik dekriptif, didapatkan hasil bahwa mahasiswa berjenis kelamin laki-laki sebesar 75% dan perempuan 25%, mayoritas mahasiswa adalah lulusan SMA dengan persentase 58%, sebagian lulusan SMK dengan persentase 40% dan MA dengan persentase 2%. Program peminatan saat SMA mayoritas adalah IPA. Jarak dari tempat tinggal ke kampus Untag Surabaya mayoritas berada di dalam jangkauan 0,1 km sampai 7,1 km. Jenis pekerjaan orang tua mayoritas sebagai pegawai swasta dengan persentase 60%.

Pembuatan model K-Means Clustering menggunakan software Rapid Miner ditampilkan pada Gambar 1.



Gambar 1. Model Rapid Miner untuk K-Means Clustering
(Sumber: Pengolahan Data)

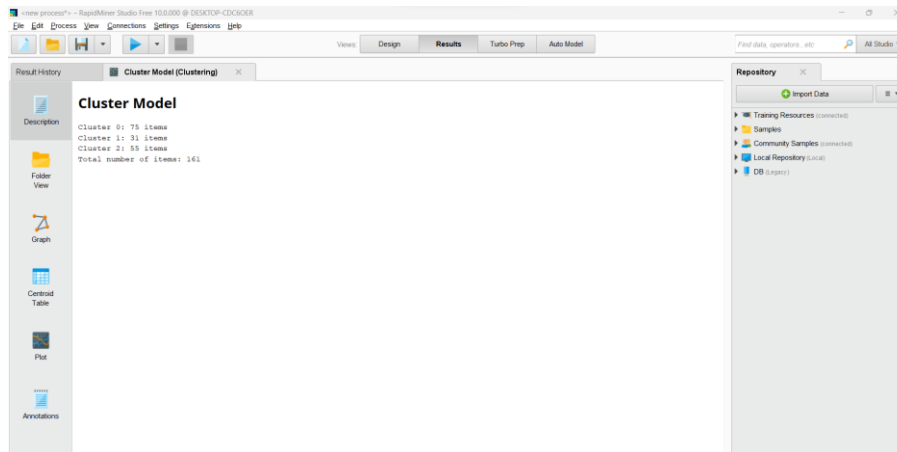
Pada Gambar 1 dibuat modul read excel untuk membaca data yang tersimpan pada file excel, kemudian dihubungkan dengan modul clustering menggunakan K-Means Clustering.

Data yang diolah ke dalam Rapidminer adalah data yang telah ditransformasi, dengan keterangan seperti ditampilkan pada Tabel 1.

Tabel 1. Pengkodean transformasi data

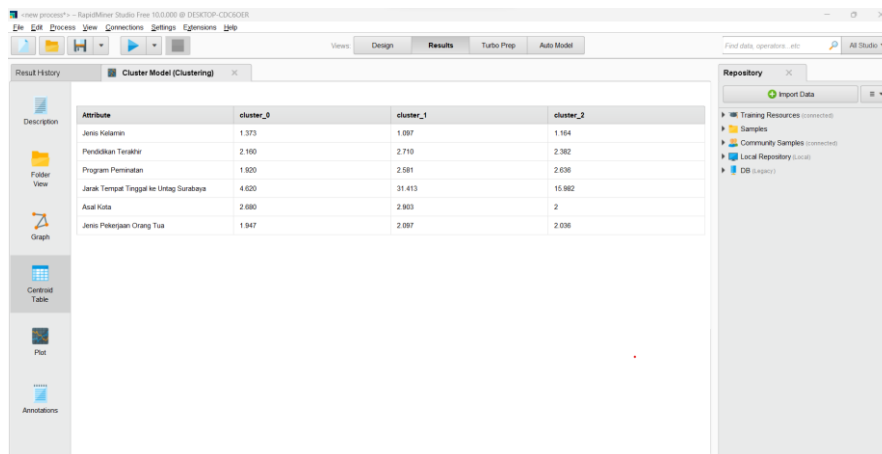
Jenis Kelamin	Pendidikan Terakhir	Program Peminatan	Asal Kota	Jenis Pekerjaan Orang Tua
Laki-laki = 1	SMA = 1	IPA=1	Surabaya=1	Pegawai Swasta=1
perempuan = 2	SMAN=2	IPS=2	Sidoarjo=2	Pegawai negeri=2
	SMK=3	Bahasa=3	Gresik=3	Pemilik Usaha=3
	SMKN=4	Lainnya=4	Lamongan,=4	Pensiunan=4
	MA = 5		Lainnya=5	Wiraswasta=5
	MAN = 6			Lainnya=6

K-Means clustering menggunakan tiga *cluster* (k=3). Gambar 2 menunjukkan jumlah data yang tersebar dalam tiap cluster.



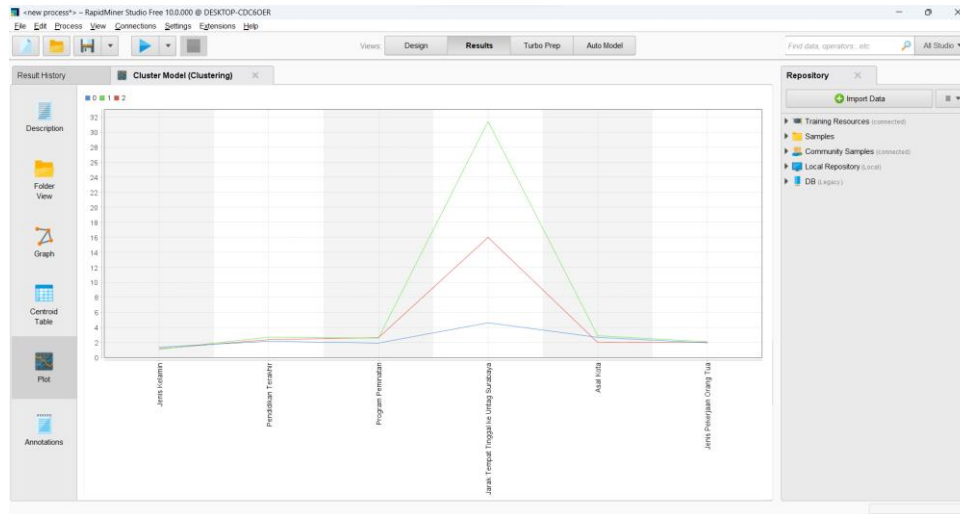
Gambar 2. Jumlah data tiap cluster (Sumber: Pengolahan Data)

Dari Gambar 2, dengan menggunakan Rapid Miner, diperoleh hasil cluster 0 sebanyak 75 data, cluster 1 sebanyak 31 data, dan cluster 2 sebanyak 55 data dengan total data sejumlah 161. Centroid dari masing-masing *cluster* ditampilkan pada Gambar 3.



Gambar 3. Centroid tiap cluster (Sumber: Pengolahan Data)

Pada gambar 3, ditampilkan centroid dari masing-masing cluster, yaitu cluster 0, 1, dan 2 pada masing-masing variable, yaitu jenis kelamin, pendidikan terakhir, program peminatan, jarak tempat tinggal ke kampus, asal kota, dan jenis pekerjaan orang tua.



Gambar 4. Plot Cluster
(Sumber: Pengolahan Data)

Berdasarkan gambar 4, atribut yang berperan penting dalam membedakan cluster adalah jarak dari tempat tinggal ke kampus. Cluster 0 dengan jumlah data terbesar berisi mahasiswa dengan range jarak tempat tinggal ke kampus antara 0,1 km sampai 10 km dan rata-rata jarak 4,6 km. Cluster 1 dengan jumlah data terkecil berisi mahasiswa dengan jangkauan jarak tempat tinggal ke kampus antara 24 km – 53 km dan rata-rata 31,4 km. Sedangkan pada cluster 2, jangkauan jarak tempat tinggal ke kampus anatara 11 km – 23 km dan rata-rata 16 km.

SIMPULAN

Dari hasil pengolahan data menggunakan Rapid Miner, hasil dari K-means Clustering, dengan menggunakan tiga cluster dan 6 variabel, antara lain jenis kelamin, pendidikan terakhir, program peminatan, jarak tempat tinggal ke kampus, asal kota, dan jenis pekerjaan orang tua terlihat bahwa jarak antara tempat tinggal ke kampus Untag Surabaya menjadi atribut yang dominan dalam membedakan data ke dalam cluster-cluster yang terbentuk. Jarak rata-rata mahasiswa dari tempat tinggal ke kampus pada cluster terbesar adalah 4,6 km.

Saran untuk strategi promosi adalah dengan melakukan promosi pada sekolah-sekolah di Surabaya dan wilayah sekitarnya yang letaknya dalam radius 4 sampai 5 km dan untuk memperluas pangsa pasar dapat melakukan promosi pada wilayah yang agak jauh dari kampus Untag Surabaya pada jarak diatas 10 km. Disamping itu disarankan melakukan promosi ke media sosial serta memberikan pelayanan yang terbaik kepada mahasiswa karena sebagian besar mahasiswa memilih berkuliah di Untag Surabaya karena rekomendasi dari teman dan keluarga.

DAFTAR PUSTAKA

- Alfina, T., & Santosa, B. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K-Means dan Gabungan Keduanya dalam Membentuk Cluster Data (Studi Kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS). *Analisa Perbandingan Metode Hierarchical Clustering, K-Means Dan Gabungan Keduanya Dalam Cluster Data*. Jurnal Teknik ITS.
- Alhapizi, M.R. et al. (2020). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang. *Journal of Software Engineering Ampera*.
- Hartati, T., Nurdiawan, O., Wiyandi, E. (2021). Analisis dan Penerapan Algoritma K-Means Dalam Strategi Promosi Kampus Akademi Maritim Suaka Bahari. *Jurnal Sains Teknologi Transportasi Maritim*. 3(1):1-7.
- Ong, J.O. (2013). Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University. *Jurnal Ilmiah Teknik Industri*. 12(1):10-20.
- Putri, K.L, et al. (2023). Penerapan Algoritma K-Means Untuk Penentuan Strategi Promosi. *JATI (Jurnal Mahasiswa Teknik Informatika)*. 7(3):2044-2050.
- Rakhmawati, D. Y., Herlina (2020). Pengaruh Pembelajaran Jarak Jauh Terhadap Mutu Belajar Mahasiswa Universitas 17 Agustus 1945 Surabaya. *Jurnal Penelitian dan Aplikasi Sistem & Teknik Industri (PASTI)*. 14(3):326-335.
- Santosa, B. (2007). *Data Mining. Teknik Pemanfaatan Data untuk Keperluan Bisnis*, First Edition ed. Yogyakarta.
- Suprawoto, T. (2016). Klasifikasi Data Mahasiswa Menggunakan Metode K-Means Untuk Menunjang Pemilihan Strategi Pemasaran. *JIKO (Jurnal Informatika Dan Komputer)*
- Susanto, S., & Suryadi, D. (2010). *Pengantar Data mining menggali pengetahuan dari bongkahan data*. Yogyakarta: Andi
- Triyansyah, D., Fitriannah, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *IncomTech, Jurnal Telekomunikasi dan Komputer*. 8(3):163-181.