

## ANALISIS DAN PENERAPAN ALGORITMA K-MEANS DALAM STRATEGI PROMOSI KAMPUS AKADEMI MARITIM SUAKA BAHARI

Tuti Hartati<sup>1\*</sup>, Odi Nurdiawan<sup>2</sup>, Eko Wiyandi<sup>3</sup>  
<sup>1,2,3</sup> STMIK IKMI Cirebon

\*Email: [toohart2013@gmail.com](mailto:toohart2013@gmail.com)

Email: [odynurdiawan@gmail.com](mailto:odynurdiawan@gmail.com)

Email: [eko.wiyandi@live.com](mailto:eko.wiyandi@live.com)

### ABSTRACT

The process of accepting new cadet candidates at the Akademi Maritim Suaka Bahari every year, produces a lot of data in the form of profiles of prospective cadets. The activity caused a large accumulation of data, it became difficult to identify prospective cadets. This research discusses the application of data mining to generate profiles that have similar attributes. One of the data mining techniques used to identify a group of objects that have the same characteristics is Cluster Analysis. The data clustering method is divided into one or more clusters that have the same characteristics called K-means. The method that the author uses is knowledge discovery in databases (KDD) consisting of Data, Data Cleaning, Data transformation, Data mining, Pattern evolution, knowledge. Implementation of K-means Clustering process using Rapid Miner. Attributes used by NIT, Level, Name, Student Status, Type of Registration, Gender, Place of Birth, Date of Birth, Religion, School Origin, School Origin Department, Religion, GPA, Subdistrict, District/ City, Province. Returns the number of clusters 30 ( $k=30$ ). From the research results based on davies bouldin test on K-means algorithm resulted in the closest value of 0 is  $k = 29$  with Davies bouldin: 0.070, with the most cluster member distribution in cluster 16 containing cluster members 115 items.

**Keywords:** Data mining, Algorithms, K-means, Clustering, Promotional Strategies

### ABSTRAK

Proses penerimaan calon mahasiswa baru di Akademi Maritim Suaka Bahari 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* untuk menghasilkan profil yang memiliki kemiripan atribut yang sama. Salah satu teknik *data mining* yang digunakan untuk mengidentifikasi suatu kelompok dari objek yang memiliki karakteristik yang sama adalah Analisis Cluster. Metode *clustering* data dibagi kedalam bentuk satu atau lebih *cluster*/kelompok yang memiliki karakteristik sama disebut *K-means*. Metode yang penulis gunakan adalah *knowledge discovery in databases* (KDD) yang terdiri dari Data, Data Cleaning, Data transformation, Data mining, Pattern evolution, knowledge. Implementasi proses *K-means Clustering* menggunakan Rapid Miner. Atribut yang digunakan NIT, Tingkat, Nama, Status Mahasiswa, Jenis Pendaftaran, Jenis Kelamin, Tempat Lahir, Tgl Lahir, Agama, Asal Sekolah, Jurusan Asal Sekolah, Agama, IPK, Kecamatan, Kab/Kota, Provinsi. Menghasilkan jumlah *cluster* 30 ( $k=30$ ). Dari hasil penelitian berdasarkan uji *davies bouldin* pada algoritma *K-means* menghasilkan nilai paling mendekati 0 adalah  $k = 29$  dengan *Davies bouldin*: 0.070, dengan persebaran anggota *cluster* terbanyak ada pada *cluster* 16 berisi anggota *cluster* 115 items.

**Kata Kunci:** Data mining, Algoritma, K-means, Clustering, Strategi Promosi

### PENDAHULUAN

Akademi Maritim Suaka Bahari sebagai lembaga penyedia jasa pendidikan dalam memberikan layanan pendidikan berupaya memberikan layanan terhadap mahasiswa dan untuk meningkatkan kepuasan terhadap layanan. Perguruan tinggi sebagai institusi pendidikan memiliki data akademik dan administrasi dalam jumlah yang besar, tetapi hanya sebagian kecil data tersebut dimanfaatkan dalam penyusunan evaluasi diri. Data akademik mahasiswa merupakan data yang dihimpun dari hasil kegiatan

proses belajar mengajar selama mengikuti studi di suatu perguruan tinggi. Di era sekarang ini, perkembangan dan persaingan antar lembaga pendidikan sangatlah ketat. Banyak lembaga pendidikan yang berlomba untuk menunjukkan kualitas maupun kuantitas dari lembaga itu sendiri. Minat dan bakat merupakan salah satu faktor yang dapat mempengaruhi keputusan siswa atau orang tua siswa dalam menentukan keputusan untuk memilih sekolah yang sesuai (Purwanto, et.al, 2017). Kenyataannya bahwa banyak sekolah yang mengabaikan faktor lain selain akademik dalam proses rekrutmen mahasiswa baru, atau dalam akademi maritim disebut calon mahasiswa. Oleh karena itu penentuan minat dan bakat calon mahasiswa dalam memilih sekolah sangatlah penting untuk menentukan strategi promosi.

Dalam menganalisis data pada penelitian ini menggunakan Teknik analisis *data mining* menggunakan proses tahapan *knowledge discovery in databases* (KDD) yang terdiri dari *Data, Data Cleaning, Data transformation, Data mining, Pattern evolution, knowledge*: Data Akademi Maritim Suaka Bahari dalam penelitian diambil dari data data mahasiswa tahun pelajaran 2019/2020. Data yang digunakan pada penelitian ini adalah sebanyak 675 *dataset*. Secara umum data yang didapatkan dari *database* akademi maritim suaka bahari berisi data yang tidak sempurna seperti data yang hilang, data tidak valid atau salah ketik. Data yang tidak digunakan sebaiknya dibuang karena keberadaannya dapat mengurangi mutu atau akurasi dari hasil *data mining*. Pembersihan data juga akan mempengaruhi performansi dari sistem *data mining* karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik Akademi Maritim Suaka Bahari. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil data yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya.

Penerimaan mahasiswa baru merupakan agenda yang harus dilakukan setiap tahun di berbagai universitas, seperti halnya Akademi Maritim Suaka Bahari. Sampai saat ini, strategi promosi yang dilakukan Akademi Maritim Suaka Bahari masih belum optimal dikarenakan setiap tahunnya dilakukan promosi ke setiap daerah secara *massive* baik secara langsung ke setiap sekolah dengan menggunakan media brosur dan baliho maupun media melalui penyedia media iklan seperti koran, radio dan iklan di *Google Ads* serta *Facebook Ads* tanpa melakukan pemetaan wilayah, baik yang potensial maupun sebaliknya terlebih dahulu. Hal tersebut mengakibatkan penggunaan anggaran yang besar dan kurang efisien. Dengan adanya penerimaan mahasiswa baru ini setiap tahunnya makan terjadi penumpukan data yang besar berupa profil calon mahasiswa dalam *database*, hal ini perlu dilakukan analisis data guna mendapatkan informasi-informasi yang bermanfaat dalam membantu menentukan strategi promosi penerimaan mahasiswa baru pada tahun berikutnya. Dalam melakukan promosi yang lebih tepat dan lebih efisien, maka perlu dilakukan pengolahan data mahasiswa terdahulu guna untuk mengetahui pola data mahasiswa tersebut, sehingga peneliti akan mendapatkan informasi-informasi yang tersembunyi dalam data mahasiswa tersebut misalnya informasi tentang pengelompokan data mahasiswa yang berpotensi berdasarkan asal kota atau wilayahnya.

Persaingan antar lembaga pendidikan sebenarnya memberikan keuntungan bagi konsumen, dalam hal ini calon mahasiswa, hal ini disebabkan konsumen memiliki banyak pilihan dalam memutuskan lembaga yang cocok untuk calon mahasiswa. Adanya persaingan antar lembaga yang semakin meningkat, maka upaya pemasaran untuk di sebuah lembaga pendidikan mutlak diperlukan. Jika pemasaran tidak ditentukan dengan baik, dalam arti tidak diupayakan mencari sasaran daerah yang potensial, maka hanya akan menghabiskan banyak waktu dan biaya yang seharusnya bisa diminimalisir melalui pemilihan daerah target promosi yang baik. Pemilihan Strategi Promosi dilakukan dengan menggunakan metode *K-means Clustering*. Dengan menggunakan metode ini, data-data yang telah didapatkan dapat dikelompokkan ke dalam beberapa *cluster* berdasarkan kemiripan dari data tersebut,

sehingga data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster* dan yang memiliki karakteristik berbeda dikelompokkan dalam *cluster* lain.

Pada penelitian yang telah dilakukan sebelumnya bahwa dalam rangka menarik minat para calon mahasiswa dan yang lebih banyak maka di perlukan strategi promosi yang tepat dengan menginventarisir *knowledge* asal sekolah, prestasi akademik, alamat tempat tinggal, tren keminatan calon mahasiswa, proses pembelajaran, serta prospek kelulusan para mahasiswa. (Suprawoto 2016)

Menurut Setiawan et.al (2016) menjelaskan bahwa dalam proses penerimaan mahasiswa baru tersebut 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. Dengan penerapan algoritma *k-means* diperoleh hasil profil mahasiswa yang memiliki kemiripan yang sama.

Sementara Menurut Dana et.al (2018) menjelaskan bahwa proses penerimaan mahasiswa baru, pihak pengelola sering menghadapi permasalahan dikarenakan belum diketahuinya pola atau strategi promosi yang tepat dan untuk menentukan strategi *marketing* diperlukan serangkaian Analisa untuk mengetahui karakteristik dari *customer*. Teknik *clustering* merupakan bagian dari disiplin ilmu *machine learning* yang sangat efisien dalam melakukan Analisa kelompok data untuk mengetahui karakteristik masing-masing kelompok data *customer*.

Menurut Wardhana (2018) menjelaskan bahwa dengan adanya persaingan antar sekolah semakin meningkat dewasa ini oleh karena itu diperlukan strategi promosi khususnya dibidang jasa pendidikan. Strategi dapat dipahami sebagai suatu rencana yang cermat, yang terkait dengan cara-cara pendayagunaan dan penggunaan sumber daya yang ada untuk meningkatkan efektifitas dan efisiensi dalam mencapai suatu sasaran tertentu. Strategi yang baik akan berdampak baik.

Minat dan bakat adalah salah satu faktor yang mempengaruhi keputusan siswa atau orang tua dalam memilih sekolah yang sesuai. Dengan demikian penentuan minat dan bakat siswa dalam pemilihan sekolah menjadi sangat penting untuk menentukan strategi promosi sekolah. Dari hasil analisa dapat diketahui bahwa minat siswa yang ingin melanjutkan Pendidikan ke perguruan tinggi memang mempunyai ragam yang berbeda. (Purwanto, et.al, 2017)

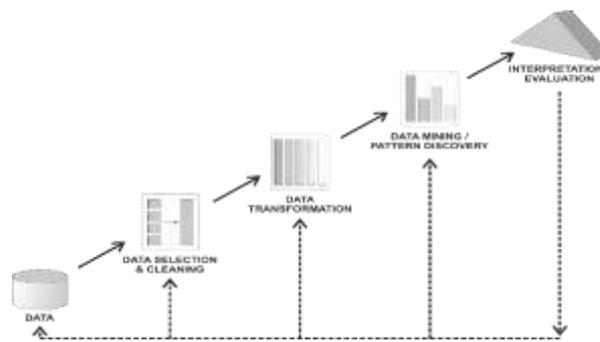
Analisis *cluster* merupakan teknik *data mining* yang bertujuan untuk mengidentifikasi suatu kelompok dari objek yang memiliki karakteristik yang sama. *K-means* adalah salah satu metode *clustering* data yang dibagi kedalam bentuk satu atau lebih *cluster*/kelompok yang memiliki karakteristik sama dihasilkan hasil pengelompokan berdasarkan daerah asal sekolah IPK rata-rata yang tertinggi berasal dari propinsi Daerah Istimewa Yogyakarta dan Jawa Tengah (Suprawoto 2016).

## METODE

Pada penelitian ini menggunakan pendekatan deskriptif kuantitatif. Dalam pengertiannya penelitian deskriptif dilakukan dengan mencari informasi terkait gejala yang ada, merencanakan cara melakukan pendekatannya, dijelaskan dengan gamblang tujuan yang akan diraih dan mengumpulkan berbagai macam data sebagai bahan untuk membuat laporan. Kemudian penelitian ini menggunakan pendekatan kuantitatif karena menggunakan angka, mulai dari pengumpulan data, penafsiran terhadap data tersebut, serta penampilan dari hasilnya. (Putra et al. 2019)

Sejalan dengan tujuan peneliti untuk mengelompokkan hasil proses pembelajaran maka membutuhkan data yang bersifat angka yang berasal dari gambaran atau deskripsi fakta yang ada di Akademi Maritim Suaka Bahari menggunakan teknik *machine learning* dengan metode *K-means*.

Adapun untuk menganalisis data dalam penerapan *data mining* ini menggunakan proses tahapan *knowledge discovery in databases* (KDD) yang terdiri dari Data, *Data Cleaning*, *Data transformation*, *Data mining*, *Pattern evolution*, *knowledge*:



Gambar 1. Tahapan KDD

Pada tahap penyajian data membahas proses *data mining* dan penerapannya, maka akan disajikan pembahasan mulai dari tempat penelitian, proses *data mining*, penggunaan *rapidminer*, hasil dan kesimpulan.

## HASIL DAN PEMBAHASAN

Dalam penelitian ini metode yang di gunakan adalah Metode *Davies bouldin index* yang mana berfungsi untuk menentukan jumlah *cluster* paling optimal yang tepat untuk digunakan dalam proses *clustering*. Dalam penentuannya hasil Metode *Devies Bouldin* adalah Semakin kecil nilai DBi yang diperoleh (non-negatif  $\geq 0$ ), maka semakin baik *cluster* yang diperoleh dari *clustering K-means* yang digunakan dimulai dari percobaan pertama yang diawali dengan jumlah  $k=2$  *cluster*, lalu  $k= 2$  *cluster* yang ditentukan sampai dengan jumlah  $k=30$ . Dari hasil percobaan dapat dilihat pada tabel 1 bahwa perhitungan nilai DBi yang paling mendekati dengan angka 0 (nol) adalah 0,070. maka dapat di tetapkan bahwa *cluster* yang paling baik/optimal adalah  $k=29$ . dengan nilai DBi = 0.070.

Tabel 1. Hasil *Cluster* Optimal

jumlah k	cluster yang dihasilkan	Nilai davies bouldin index	jumlah anggota cluster
29	29	0.070	Cluster 0: 16 items
			Cluster 1: 8 items
			Cluster 2: 36 items
			Cluster 3: 25 items
			Cluster 4: 20 items
			Cluster 5: 3 items
			Cluster 6: 10 items
			Cluster 7: 4 items
			Cluster 8: 26 items
			Cluster 9: 10 items
			Cluster 10: 1 items
			Cluster 11: 38 items
			Cluster 12: 1 items
			Cluster 13: 3 items
			Cluster 14: 8 items
			Cluster 15: 115 items
			Cluster 16: 10 items
			Cluster 17: 11 items

jumlah k	cluster yang dihasilkan	Nilai <i>davies bouldin index</i>	jumlah anggota cluster
			Cluster 18: 2 items
			Cluster 19: 5 items
			Cluster 20: 82 items
			Cluster 21: 21 items
			Cluster 22: 33 items
			Cluster 23: 21 items
			Cluster 24: 15 items
			Cluster 25: 34 items
			Cluster 26: 93 items
			Cluster 27: 20 items
			Cluster 28: 4 items

Dari uji algoritma *k-means* menggunakan aplikasi *rapidminer* dengan penggunaan parameter  $k=2$  sampai dengan  $k=30$  hasilnya seperti pada di atas. Maka, dilihat dari hasil tabel tersebut bahwa nilai *davies bouldin index* yang mendekati nol ada pada  $k=29$  dengan nilai DBi 0.070 dengan jumlah *Cluster* terbanyak ada pada *Cluster 16* yaitu sebanyak 115 *items*. Adapun profil mahasiswanya terdiri dari 88 mahasiswa berasal dari provinsi Jawa Barat, 12 mahasiswa dari provinsi Jawa Tengah, 6 mahasiswa dari provinsi Jawa Timur, 4 mahasiswa dari provinsi Provinsi DKI Jakarta, 2 mahasiswa dari provinsi Jambi, 2 mahasiswa dari provinsi Lampung, 1 mahasiswa dari provinsi D.I. Yogyakarta adapun program yang banyak diminati program studi Nautika.

Dilihat dari jumlah persebaran mahasiswa hasil penerapan algoritma *K-means*, maka strategi promosi yang baik digunakan yaitu turun langsung ke lapangan (*direct marketing*), jejaring sosial (*social network*) dan periklanan (*advertising*). Penentuan promosi data peminat dilakukan berdasarkan *cluster* yang terbentuk. Berikut strategi promosi yang digunakan pada masing-masing wilayah:

1. Strategi *Direct Marketing* atau turun langsung ke sekolah-sekolah terdekat yang masih berada di provinsi Jawa Barat dengan media promosi berupa brosur dan door prize.
2. Strategi *Social Network* seperti *Facebook*, *Twitter* dan *Instagram* digunakan sebagai media promosi untuk menjangkau seluruh wilayah, terutama untuk menjangkau wilayah-wilayah diluar provinsi Jawa Barat.
3. Strategi *Advertising* atau periklanan digunakan dengan media cetak (Koran dan Baliho), elektronik (TV dan Radio) dan digital (*Google Ads* dan *Facebook Ads*) untuk memberikan informasi terkait penerimaan mahasiswa baru ke provinsi-provinsi potensial yang berada di cluster tersebut.

Dari program studi yang paling diminati yakni Nautika, strategi yang tepat untuk diterapkan adalah Strategi Posisi Kompetitif, di mana materi pada media promosi harus lebih menonjolkan kelebihan program studi Teknik dan KPNK guna meningkatkan peminat untuk masuk ke program studi Teknik dan KPNK. Disamping itu juga guna menyeimbangkan jumlah mahasiswa di setiap program studi.

## KESIMPULAN

Dari hasil penelitian dan pembahasan yang telah diuraikan pada bab-bab sebelumnya, maka dapat ditarik Kesimpulan:

1. Berdasarkan karakteristik wilayah dan program studi hasil penerapan algoritma *K-means*, maka strategi yang paling tepat adalah menggunakan metode Posisi Kompetitif, di mana materi pada media promosi harus menonjolkan kelebihan program studi Teknika dan KPNK guna meningkatkan peminat untuk masuk ke program studi Teknika dan KPNK serta menyeimbangkan persebaran taruna di semua program studi. Untuk promosi di wilayah provinsi Jawa Barat menggunakan strategi *Direct Marketing* atau turun langsung ke sekolah-sekolah dengan langsung membagikan brosur dan presentasi. Untuk wilayah di luar Jawa Barat, bisa menggunakan strategi *Social Network* seperti *Facebook*, *Twitter* dan *Instagram*. Dan terakhir untuk memusatkan promosi ke wilayah potensial hasil *clusterisasi*, dapat dilakukan promosi menggunakan *Advertising* atau periklanan media cetak (Koran dan Baliho), elektronik (TV dan Radio) dan digital (*Google Ads* dan *Facebook Ads*).
2. Berdasarkan hasil dari *clusterisasi* mahasiswa dalam menentukan Strategi Promosi Kampus Akademi Maritim Suaka Bahari dapat disimpulkan bahwa profil mahasiswa terbaik yang dijadikan untuk promosi kampus terdapat pada *cluster 29* dengan *Davies bouldin index* 0,070
3. Berdasarkan hasil penelitian untuk mengukur *performance* algoritma *K-means* dengan parameter *range cluster* disimpulkan bahwa *Cluster 16* berjumlah 115 anggota. Adapun profil mahasiswanya terdiri dari 88 mahasiswa berasal dari provinsi Jawa Barat, 12 mahasiswa dari provinsi Jawa Tengah, 6 mahasiswa dari provinsi Jawa Timur, 4, mahasiswa dari provinsi Provinsi DKI Jakarta, 2 mahasiswa dari provinsi Jambi, 2 mahasiswa dari provinsi Lampung, 1 mahasiswa dari provinsi D.I. Yogyakarta adapun program yang banyak diminati program studi Nautika.

## SARAN

Berdasarkan kesimpulan sebagai hasil pembahasan dari penelitian ini, maka terdapat beberapa saran untuk pengembangan penelitian yaitu sebaran wilayah yang dihasilkan dari proses *clustering* terbaik baru terdapat 7 provinsi, sedangkan pada standar 3 BAN-PT tentang kualitas penerimaan mahasiswa baru yg baik itu terdiri dari 8 provinsi. Maka, hasil penelitian ini dapat dikembangkan lagi agar mendapat hasil yang lebih baik guna dijadikan sebagai dasar strategi untuk melakukan promosi penerimaan mahasiswa baru yang efektif dan efisien ke depan. Berikutnya, penelitian ini hanya melakukan klasifikasi menggunakan algoritma *K-means* dan untuk penelitian selanjutnya bisa di kembangkan lagi menggunakan algoritma yang lainnya.

## DAFTAR PUSTAKA

- Dana, Raditya Dinar, Cep Lukman Rohmat, and Ade Rizki Rinaldi. 2018. "Strategi Marketing Penerimaan Mahasiswa Baru Menggunakan Machine Learning Dengan Teknik Clustering." *Jurnal Pengembangan IT (JPIT)* 04(2):201–4. doi: 10.30591/jpit.v4i2-2.1879.
- Purwanto, Agus, Indrajit, Richardus Eko. 2017. "Menentukan Strategi Promosi Politeknik Tmkm." *Menentukan Strategi Promosi Politeknik Tmkm Berdasarkan Analisa Pengaruh Minat Siswa.* (November):1–2.
- Putra, Erik Ade, Iyus Jayusman, Oka Agus Kurniawan Shavab, Ki Hadjar Dewantara, Surakarta E-mail, Ananda Aulia, Ai Yulianti, and Wiwin Yuliani. 2019. "Deskripsi Kualitatif Sebagai Satu Metode Dalam Penelitian Pertunjukan." *Quanta* 3(3):9–19. doi: 10.25157/ja.v7i1.3180.

- Setiawan, Rony. 2016. "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru ( Studi Kasus : Politeknik Lp3i Jakarta )." *J. Lentera Ict* 3(1):76–92.
- Suprawoto, Totok. 2016. "Klasifikasi Data Mahasiswa Menggunakan Metode K-Means Untuk Menunjang Pemilihan Strategi Pemasaran." *JIKO (Jurnal Informatika Dan Komputer)* 1(1):12–18. doi: 10.26798/jiko.2016.v1i1.9.
- Wardhana, Ardiyanto. 2018. "Strategi Promosi Penerimaan Siswa Baru (Studi Kasus Smk Kesehatan Cipta Bhakti Husada Yogyakarta)." *CHANNEL: Jurnal Komunikasi* 6(1):96. doi: 10.12928/channel.v6i1.10215.