

ANALISIS DATA MINING MENGGUNAKAN ALGORITMA FUZZY C-MEANS UNTUK MENENTUKAN STRATEGI PROMOSI POLITEKNIK POS INDONESIA

Nisa Hanum Harani¹, Cahyo Prianto²,
¹Teknik Informatika, Politeknik Pos Indonesia
¹nisahanum@poltekpos.ac.id
²Teknik Informatika, Politeknik Pos Indonesia
²cahyoprianto@poltekpos.ac.id

Abstrak

Pengelolaan pendidikan menjadi sangat penting, dimana pertumbuhan dan perkembangan lembaga dipengaruhi oleh kemampuan administrator dalam melakukan scanning lingkungan eksternal, kompetitor lembaga lain dan memperhitungkan kompetensi internal. Seiring dengan persaingan yang ketat antar perguruan tinggi, strategi promosi adalah kunci utama dalam meningkatkan jumlah calon mahasiswa. Perguruan tinggi harus lebih proaktif dalam mencari calon mahasiswa istilahnya adalah “ jemput bola”. Kegiatan penerimaan mahasiswa baru politeknik pos indoneisa menghasilkan data mahasiswa yang sangat berlimpah berupa data profil mahasiwa dan data kegiatan perkuliahan. Pengolahan data mahasiwa dapat menghasilkan informasi yang bermanfaat bagi politeknik pos dalam membantu pihak PMB sebagai materi pertimbangan dalam menentukan strategi promosi penerimaan mahasiwa baru .

Metode cluster yang dipakai adalah FCM (Fuzz Logic Clustering), berbeda dengan metode clustering biasa FCM mengandalkan logika kabur dalam membagi data kedalam beberapa kelompok. Keluaran dari FCM berbeda dengan Fuzzy logic, namun berupa deretan cluster dan beberapa derajat keanggotaan untuk tiap kelompok. Metode penelitian yang di pakai adalah DSRM (Design Science Research Methodology). Hasil dari penelitian ini berupa pengelempokan data mahasiswa melalui pesebaran wilayah, pengelompokan tersebut dapat digunakan sebagai salah satu dasar pengambilan keputusan untuk menentuka strategi promosi berdasarkan cluster. Implementasi menggunakan Matlab untuk membantu memodelkan data yang akurat. Atribut yang digunakan adalah kota asal, prodi dan IPK mahasiswa.

Kata Kunci: data mining, FCM, akademik, promosi

I. PENDAHULUAN

Pengelolaan pendidikan menjadi sangat penting, dimana pertumbuhan dan perkembangan lembaga dipengaruhi oleh kemampuan administrator dalam melakukan *scanning* lingkungan eksternal, kompetitor lembaga lain dan memperhitungkan kompetensi internal [1] . Pemasaran adalah konsep generik yang berlaku untuk semua organisasi, banyak perguruan tinggi kini telah mengadopsi praktek perencanaan strategi

pemasaran [2]. Untuk perguruan tinggi, penerapan strategi pemasaran menciptakan nilai antara pemangku kepentingan didalamnya. Para pemangku kepentingan ini adalah calon siswa, mahasiswa, alumni, alumni yang telah bekerja dan yayasan. Perguruan tinggi perlu menyusun suatu strategi pemasaran dalam suatu program yang terkoordinasi [3]. Promosi merupakan salah satu elemen dari strategi pemasaran, promosi bisa melalui media cetak dan media elektronik [4].

Seiring dengan persaingan yang ketat antar perguruan tinggi, strategi promosi adalah kunci utama dalam meningkatkan jumlah calon mahasiswa. Perguruan tinggi harus lebih proaktif dalam mencari calon mahasiswa istilahnya adalah “jemput bola”. Untuk melakukan jemput bola yang efektif dan efisien perlu dilakukan proses pencarian informasi berharga dalam menentukan daerah promosi. Peningkatan dan penurunan jumlah calon mahasiswa baru dapat diukur dari proses PMB setiap semester. Data PMB dan data kegiatan belajar mengajar menjadi kunci dalam proses analisa karakteristik daerah.

Data mining adalah proses penemuan pola yang menarik dari sekumpulan data yang besar dengan bantuan algoritma untuk mengekstrak informasi tersembunyi [5]. Klustering memiliki peran yang penting dalam data mining, dimana teknik ini akan membagi data kedalam beberapa kelompok sesuai dengan kemiripannya [6]. Fuzzy klustering merupakan salah satu metode klustering yang paling dikenal karena memiliki keunggulan ketahanan terhadap ambiguitas dan mempertahankan lebih banyak informasi [7] daripada metode Clustering C-Means. FCM memungkinkan setiap sampel untuk dimiliki oleh setiap cluster dengan nilai fuzzy antara 0 dan 1. Jumlah semua keanggotaan untuk sampel harus [8]. Fungsi keanggotaan mempermudah identifikasi sampel yang tidak berkontribusi secara nyata terhadap kelompok mana pun (yaitu, tidak memiliki keanggotaan yang signifikan), dan perwakilan (yaitu, mereka yang memiliki keanggotaan 1 dalam sebuah cluster) [9]. Proses pembagian data kedalam beberapa kelompok dengan tingkat ambiguitas yang kecil dan didukung penemuan informasi yang berharga adalah acuan penulis dalam menjawab solusi permasalahan strategi promosi perguruan tinggi. Data yang akan dianalisis adalah data PMB

Politeknik Pos Indonesia dan data kegiatan belajar mengajar sedikitnya 3 periode. Berikut atribut yang diolah adalah kota asal, prodi dan IPK mahasiswa.

II. TINJAUAN PUSTAKA

2.1 Marketing Mix

Salah satu strategi yang berhubungan dengan kegiatan pemasaran perusahaan adalah marketing mix strategy yang didefinisikan oleh Kotler dan Armstrong [10] yang menyatakan bahwa “*marketing mix as the set of controllable marketing variables that the firm bleads to produce the response it wants in the target market*”. Dari definisi diatas dapat diartikan bahwa bauran pemasaran merupakan variable-variabel terkendali [11] yang digabungkan untuk menghasilkan tanggapan yang diharapkan dari pasar sasaran. Dan untuk usaha jasa terdapat 7 unsur marketing mix (Marketing Mix-7p) yaitu: Produk, Price, Promotion, Place, Partisipant, Proses, Dan Physical Evidence.

2.2 Promosi

Promosi adalah kegiatan mengkomunikasikan informasi dari penjual kepada konsumen atau pihak lain dalam saluran penjualan untuk mempengaruhi sikap dan perilaku. Melalui periklanan suatu perusahaan mengarahkan komunikasi persuasif pada pembeli sasaran dan masyarakat melalui media-media yang disebut dengan media massa seperti Koran, majalah, tabloid, radio, televisi dan *direct mail* [12] Media promosi yang dapat digunakan pada bisnis ini antara lain (1) Periklanan, (2) Promosi penjualan, (3) Publisitas dan hubungan masyarakat, dan (4) Pemasaran langsung. Penentuan media promosi yang akan digunakan didasarkan pada jenis dan bentuk produk itu sendiri.

2.3 Fuzzy Clustering

Clustering adalah proses mengelompokkan atau penggolongan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan

hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/cluster [5] (rong). Clustering dalam data mining berguna untuk menemukan pola distribusi di dalam sebuah data set yang berguna untuk proses analisa data. Kesamaan objek biasanya diperoleh dari kedekatan nilai-nilai atribut yang menjelaskan objek-objek data, sedangkan objek-objek data biasanya direpresentasikan sebagai sebuah titik dalam ruang multidimensi.

Berdasarkan himpunan bagian yang dihasilkan clustering dapat dibedakan menjadi crisp (hard) atau himpunan fuzzy. Metode hard clustering merupakan model yang berdasar pada teori himpunan klasik, yang mana suatu objek menjadi anggota atau tidak menjadi anggota secara penuh kedalam suatu kelompok. Hard clustering membagi data kedalam sejumlah himpunan bagian secara eksklusif. Sementara metode fuzzy clustering mengizinkan objek untuk menjadi bagian dari beberapa kelompok secara bersamaan dengan perbedaan level keanggotaan.

Terdapat dua metode pengclusteran dalam fuzzy clustering, yaitu Fuzzy Fuzzy C-means (FCM) yang merupakan algoritma pengclusteran yang terwarisi sebab pada FCM perlu diketahui terlebih dahulu jumlah cluster yang akan dibentuk. Sementara Subtractive Clustering merupakan algoritma pengclusteran yang dapat digunakan apabila jumlah cluster belum diketahui sebelumnya

2.3.1 Fuzzy C-Means Clustering

Algoritma fuzzy c-means diperkenalkan oleh Ruspini [16], kemudian dikembangkan oleh Dunn [17] dan Bezdek [18] dan telah banyak digunakan dalam analisis klaster, pengenalan pola dan pengelolaan citra dll. Algoritma

clustering fuzzy c-means (FCM) memperkenalkan logika kabur untuk keaslian setiap objek dan dapat menyimpan lebih banyak informasi tentang kumpulan data daripada algoritma clustering k-means (HCM) [6]. Meskipun algoritma FCM memiliki kelebihan yang cukup besar dibandingkan algoritma k-means clustering, ada juga beberapa kekurangan saat menggunakan algoritma FCM dalam praktiknya.

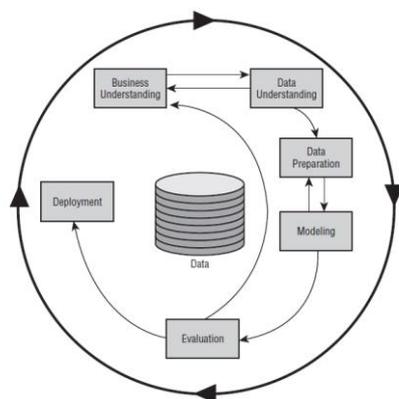
Bertentangan dengan metode analisis pengelompokan tradisional, yang mendistribusikan setiap objek ke kelompok yang unik, algoritma pengelompokan fuzzy mendapatkan nilai keanggotaan antara 0 dan 1 yang menunjukkan tingkat keanggotaan untuk setiap objek pada masing-masing kelompok. Jelas, jumlah nilai keanggotaan untuk setiap objek ke semua kelompok pasti sama dengan 1. Nilai keanggotaan yang berbeda menunjukkan probabilitas masing-masing objek terhadap kelompok yang berbeda.

III. METODOLOGI PENELITIAN

3.1 Alur Metodologi Penelitian

Metodologi penelitian yang akan ditempuh adalah kuantitatif, yakni suatu proses menemukan pengetahuan yang menggunakan data berupa angka sebagai alat menganalisis keterangan mengenai apa yang diketahui. Logika yang dipakai dalam penelitian yang bersifat kuantitatif adalah menggunakan logika *positivistic* sehingga menghindari hal-hal yang bersifat subjektif. Proses penelitian dilaksanakan dengan mengikuti prosedur standar dalam penggalian sebuah data yang biasa disebut dengan *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) [8]. *CRISP-DM* terdiri

dari enam tahap seperti yang terlihat dalam gambar 1:



Gambar 1 CRISP-DM

Metodologi CRISP-DM adalah standarisasi *data mining* yang disusun oleh tiga penggagas *data mining market* yaitu Daimler Chrysler (Daimler Benz), SPSS, NCR. Kemudian dikembangkan pada berbagai *workshop* (antara tahun 1997-1999). Lebih dari 300 organisasi yang berkontribusi dalam proses *modeling* ini dan akhirnya CRISP-DM 1.0 dipublikasikan pada tahun 1999 [9]

IV. HASIL PENELITIAN DAN PEMBAHASAN

4.1 Hasil Penelitian

Analisis yang dilakukan adalah berdasarkan tahapan-tahapan pada metodologi yang dilakukan menggunakan CRISP-DM, setelah melakukan pemahaman terhadap tujuan (*Business Understanding*) maka langkah selanjutnya beranjak pada pemahaman data (*Data Understanding*). Dalam penelitian ini, data yang diperoleh adalah berasal dari hasil pengumpulan data PMB di Politeknik Pos Indonesia. Atribut ada 3 yaitu, IPK, Program Studi dan Daerah asal. Pertama dilakukan adalah inisiasi terhadap nilai dari atribut Program Studi, terdapat 10 Program

studi. Kemudian dilakukan Inisiasi terhadap atribut Daerah Asal, inisiasi dilakukan dan diurutkan berdasarkan daerah yang paling terbanyak jumlah mahasiswanya..

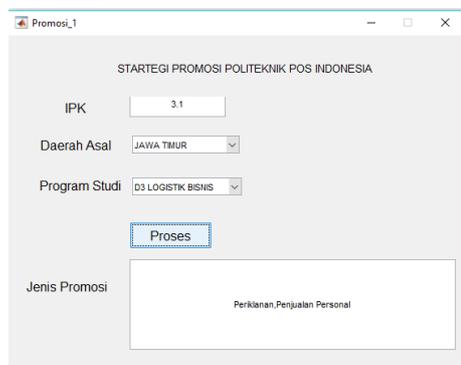
Jumlah sample dalam penelitian di Politeknik Pos Indonesia berjumlah 1717 data diambil dari data seluruh mahasiswa di tahun 2013-2015 seluruh program studi.

Dari 1717 data set yang telah diperoleh dalam penelitian ini, ditemukan bahwa terdapat sekitar 45 buah data yang terindikasi *missing value*. Kemudian strategi yang digunakan dalam menangani *missing value* pada penelitian ini adalah dengan cara mengisi data tersebut dengan mean/rata-rata. Cara ini diambil karena jumlah data set yang terindikasi *missing value* dapat mempengaruhi hasil cluster. Sehingga total data set yang dianalisis tetap berjumlah 1717.

4.2 Penerapan Algoritma Fuzzy C-Means

Setelah data yang diperlukan siap, maka selanjutnya proses penggalian data yang pertama dilakukan adalah dengan menyaring data menggunakan teknik *clustering* sehingga data set dapat terkelompok dengan jumlah *clustering* yang kita tentukan sebanyak 3. Teknik *clustering* yang dilakukan adalah dengan menggunakan algoritma Fuzzy C-Means. Proses penggalian data menggunakan algoritma Fuzzy C-means ini dilakukan dengan cara menggunakan bantuan perangkat lunak penggali data Matlab. Dalam proses *Clustering*, IPK dijadikan Xi1, Daerah asal Xi2, Program Studi dijadikan Xi3. Dari *clustering* yang dilakukan diperoleh hasil yaitu nilai fungsi obyektif dan iterasi, pusat cluster

atau center serta derajat keanggotaan lulusan untuk setiap cluster pada iterasi terakhir.



Gambar 2. Penerapan Fuzzy C-Means

Dari cluster yang dilakukan diperoleh hasil yaitu nilai fungsi obyektif selama iterasi, pusat cluster atau center serta derajat keanggotaan lulusan untuk setiap cluster pada iterasi terakhir. Dalam penelitian ini, proses iterasinya berhenti pada iterasi ke-50 karena nilai $|Pt - Pt-1| < \xi$. Nilai fungsi obyektif pada iterasi terakhir yang diperoleh adalah 10840.6242 seperti pada gambar 1.

4.3 Pembahasan Penerapan Algoritma Fuzzy C-Means

Data keluaran dari hasil perangkat lunak Matlab diperoleh hasil yang dapat dilihat di table 1:

Table 1 Hasil Akhir Algoritma K-Means

Jenis Cluster	Rata-rata IPK	Jumlah instance	Daerah Tertinggi
Cluster 0	2,755	716	Jawa Barat dan Sumatra Utara
Cluster 1	3,35	542	Jawa Barat, Jawa Timur
Cluster 2	2,85	459	Sumatra Barat, Jawa Tengah
Total		1717	

Jumlah cluster terbanyak ada di cluster 0 dimana mahasiswa banyak berasal dari Jawa barat dan Sumatra Utara dengan rata-rata IPK 2,755. Jumlah kedua terbanyak adalah cluste 1 dimana mahasiswa dari daerah Jawa Timur dan juga dari Daerah Jawa Barat dengan rata-rata IPK tertinggi yaitu 3,35. Kemudian disusul oleh cluster 2 dimana kebanyakan mahasiswa berasal dari Sumatra Barat dan Jawa Tengah dengan rata-rata IPK 2,85.

4.4 Strategi Promosi

Dari data hasil *clustering* yang telah dilakukan di atas, maka dapat ditentukan beberapa strategi promosi yang dapat dilakukan oleh pihak Politeknik Pos Indonesia. Pada klaster pertama didominasi oleh Prodi Diploma iii Teknik Informatika yang kebanyakan berasal dari daerah Jawa Barat. Sedangkan mahasiswa paling sedikit memilih prodi diploma iv Logistik Bisnis. Pada klaster kedua didominasi oleh Prodi Diploma iv Logistik Bisnis yang kebanyakan berasal dari daerah DKI Jakarta. Sedangkan mahasiswa paling sedikit memilih prodi Diploma iii Manajemen Informatika. Pada klaster ketiga didominasi oleh Prodi Diploma iii Manajemen Informatika yang kebanyakan berasal dari daerah Jawa Barat (bandung) . Sedangkan mahasiswa paling sedikit memilih prodi Diploma i Akuntansi . Secara keseluruhan kluster diperoleh informasi bahwa kebanyakan pilihan mahasiswa pada Prodi Diploma iv Logistik Bisnis yang berasal dari Jawa Barat dengan rata-rata IPK 2,755. Informasi yang diperoleh dai masing-masing kluster secara detail dapat digunakan untuk menentukan strategi promosi PMB Politeknik POs Indonesia dengan mempertimbangkan

nilai minimal yang perlu diambil tindakan dalam strategi promosi.

4.4.1 Penentuan Strategi Promosi

Penentuan strategi promosi PMB Politeknik Pos Indonesia dapat dilakukan dengan dua strategi, pertama mengacu pada hasil klaster yang masih nilai minimal. Dengan diketahuinya nilai minimal memungkinkan peningkatan usaha promosi yang tepat sasaran, dapat diambil :

- a. Promosi pada Prodi Diploma iii Manajemen Informatika dan Diploma I Akuntansi
- b. Promosi di daerah Jawa Barat dan Sumatra Utara lebih ditingkatkan lagi
- c. Promosi disarankan dengan mengirimkan tim Marketing yang sesuai dengan program studi yang diminati tiap daerah.

Kemudian akan dilakukan analisis strategi promosi berdasarkan promotion mix pada masing-masing cluster beserta detail rujukan realisasinya Untuk Cluster Pertama disarankan metode promosinya adalah Penyebaran brosur, website, baliho dan social media. Untuk Cluster dua Penjualan Personal ditingkatkan , lakukan workshop ke sma/smk dan juga berikan undangan untuk siswa berprestasi. Untuk Cluster ketiga fokus pada periklanan media cetak dan social media.

5. KESIMPULAN

Berdasarkan hasil dari analisis dan pembahasan dari bab sebelumnya, diperoleh kesimpulan sebagai berikut :

1. Dengan menggunakan algoritma Fuzzy C-Means clustering maka data persebaran wilayah yang berpotensi dijadikan sasaran promosi PMB dapat terlihat.

2. Dengan melihat persebaran wilayah potensi seluruh Indonesia, maka tim PMB dapat melihat jenis promosi yang sesuai dengan karakter masing-masing daerah.

DAFTAR PUSTAKA

- [1] E. Herawan and N. Hartini, *Manajemen Pendidikan, Pertama*. Bandung: Alfabeta, 2008.
- [2] D. Lewison and J. Hawes, "Student Target Marketing Strategies for Universities.," *J. Coll. Admiss.*, pp. 14–19, 2007.
- [3] A. Nurbasari, "Dinamika Manajemen Pemasaran Jasa.pdf," in *Dinamika Manajemen Pemasaran Jasa*, 2011, p. 77.
- [4] P. Kotler and G. Armstrong, *Marketing: An introduction*, 9th ed. Pearson Prentice Hall, 2009.
- [5] F. Chen, P. Deng, J. Wan, D. Zhang, A. V. Vasilakos, and X. Rong, "Data mining for the internet of things: Literature review and challenges," *Int. J. Distrib. Sens. Networks*, vol. 2015, no. i, 2015.
- [6] Y. Lu, T. Ma, C. Yin, X. Xie, W. Tian, and S. Zhong, "Implementation of the Fuzzy C-Means Clustering Algorithm in Meteorological Data," *Int. J. Database Theory Appl.*, vol. 6, no. 6, pp. 1–18, 2013.
- [7] Y. Zheng, B. Jeon, D. Xu, Q. M. J. Wu, and H. Zhang, "Image segmentation by generalized hierarchical fuzzy C-means algorithm," *J. Intell. Fuzzy Syst.*, vol. 28, no. 2, pp. 961–973, 2015.
- [8] E. H. Ruspini, "A THEORY OF FUZZY CLUSTERIN."
- [9] C. Rong and R. M. Esteves, "Using Mahout for clustering Wikipedia's latest articles: A comparison between k-means and fuzzy c-means in the cloud," *Proc. - 2011 3rd IEEE Int. Conf.*

- Cloud Comput. Technol. Sci. CloudCom 2011*, pp. 565–569, 2011.
- [10] S. R. A. Hendri Sukatjo, “Analisa Marketing Mix-7P(Produk, Price, Promotion, Place, partisipant, Process, dan Physical Evidence) terhadap Keputusan Pembelian Produk Klinik Kecantikan Teta di Surabaya,” *J. Mytra Ekon. dan manajemen Bisnis*, vol. 1, no. 2, pp. 216–228, 2010.
- [11] B. Gary, *Marketing Management*, 12th ed. 2000.
- [12] S. Natalius, “Matode Naive Bayes Classifier dan Penggunaanya Pada Klasifikasi Dokumen,” Institut Teknologi Bandung, 2010.
- [13] A. Karim, “Prestasi Akademik Mahasoswa Politeknik Poliprofesi Medan dengan Kernel K-Means CLustering,” pp. 151–160.
- [14] B. Everitt and S. Landau, *Cluster Analysis*, 5th ed. London: John Wiley & Son, 2011.
- [15] B. Muslimatin, “Perbandingan Metode K-Means dan Metode Fuzzy C-Means (FCM) Untuk CLUSTERING Data,” Universitas Islam Negeri Maulana Malik Ibrahim, 2011.
- [16] E. Ruspini, *Information Sciences*, 2nd ed. 1970.