

PENERAPAN FRAMEWORK CRISP-DM MENGGUNAKAN FUZZY C-MEANS UNTUK CLUSTERING DATA PERGURUAN TINGGI SWASTA DI LLDIKTI WILAYAH IV

Siti Nur Aulia^{1*}, Lena Magdalena², Muhammad Hatta³

^{1,2,3}Teknologi Informasi, Sistem Informasi, Universitas Catur Insan Cendekia, Cirebon, Indonesia
Email: ^{1*}siti.aulia.si.20@cic.ac.id, ²lena.magdalena@cic.ac.id, ³muhammad.hatta@cic.ac.id

Abstrak

Jumlah perguruan tinggi swasta di LLDIKTI Wilayah IV pada Desember 2023 yang terdata di PDDIKTI yaitu 437 perguruan tinggi dengan 2.782 program studi aktif, 816.512 mahasiswa dan 29.538 dosen. Dengan pertumbuhan jumlah perguruan tinggi yang terus meningkat menyebabkan daya saing yang ketat. Untuk menghadapi hal tersebut perguruan tinggi harus mempertahankan dan meningkatkan kualitas perguruan tinggi. Oleh karena itu, diperlukan pengelompokan perguruan tinggi swasta untuk mendapatkan informasi mengenai perkembangan dan kualitas perguruan tinggi. Penelitian ini bertujuan untuk membuat *clustering* perguruan tinggi swasta di LLDIKTI Wilayah IV ke dalam beberapa *cluster* sehingga dapat digunakan untuk menganalisis perkembangan dan kualitas perguruan tinggi swasta di LLDIKTI Wilayah IV. Penelitian ini menerapkan kerangka kerja CRISP-DM dengan enam tahapan yang terdiri dari *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation* dan *Deployment*. Untuk melakukan *clustering* data perguruan tinggi swasta di LLDIKTI Wilayah IV yaitu menggunakan metode Fuzzy C-Means berdasarkan akreditasi perguruan tinggi, jumlah program studi, jumlah mahasiswa dan jumlah dosen. Hasil *clustering* yang telah didapatkan yaitu terdapat dua *cluster* yaitu *Cluster 1* (Ideal) dengan 205 PTS dan *Cluster 2* (Memadai) dengan 110 PTS. Hasil *clustering* tersebut selanjutnya dilakukan evaluasi dengan menggunakan metode *Davies Bouldin Index* atau DBI didapatkan hasil evaluasi yaitu -0,854 yang menunjukkan bahwa hasil *clustering* yang didapatkan kurang baik karena bernilai negatif. Hasil akhir dari penelitian ini yaitu *dashboard* yang menampilkan hasil *clustering* dengan menampilkan informasi serta visualisasi yang dapat mudah dipahami. Hal ini diharapkan dapat membantu dalam analisis perkembangan dan kualitas perguruan tinggi swasta di LLDIKTI Wilayah IV.

Kata Kunci: CRISP-DM, fuzzy c-means, *clustering*, perguruan tinggi swasta, LLDIKTI Wilayah IV

Abstract

The number of private universities in LLDIKTI Region IV in December 2023 recorded in PDDIKTI is 437 universities with 2,782 active study programs, 816,512 students and 29,538 lecturers. With the growth of the number of universities that continue to increase, it leads to tight competitiveness. To deal with this, universities must maintain and improve the quality of universities. Therefore, it is necessary to group private universities to obtain information about the development and quality of universities. This study aims to create clusters of private universities in LLDIKTI Region IV into several clusters so that they can be used to analyze the development and quality of private universities in LLDIKTI Region IV. This study applies the CRISP-DM framework with six stages consisting of *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation* and *Deployment*. To cluster data on private universities in LLDIKTI Region IV, it uses the Fuzzy C-Means method based on university accreditation, the number of study programs, the number of students and the number of lecturers. The clustering results that have been obtained are that there are two clusters, namely *Cluster 1* (Ideal) with 205 PTS and *Cluster 2* (Adequate) with 110 PTS. The clustering results were then evaluated using the *Davies Bouldin Index* or DBI method, the evaluation result was -0.854 which showed that the clustering results obtained were not good because of negative values. The final result of this study is a *dashboard* that displays the results of clustering by displaying information and visualizations that can be easily understood. This is expected to help in analyzing the development and quality of private universities in LLDIKTI Region IV.

Keywords: CRISP-DM, fuzzy c-means, clustering, private universities, LLDIKTI Region IV

1. PENDAHULUAN

Dengan bertambahnya jumlah perguruan tinggi khususnya pada perguruan tinggi swasta (PTS) di wilayah Jawa Barat dan Banten yang memberikan dampak pada persaingan yang ketat dalam mendapatkan calon mahasiswa. Oleh karena itu, pada setiap perguruan tinggi swasta (PTS) mengharuskan untuk meningkatkan kualitas dan mutu perguruan tinggi secara berkala. Untuk mendapatkan informasi tersebut mengenai perkembangan dan kualitas perguruan tinggi swasta di LLDIKTI Wilayah IV cukup sulit. KEMENDIKBUDRISTEK telah melakukan klasifikasi dan pemeringkatan perguruan tinggi pada tahun 2015 untuk meningkatkan mutu pendidikan tinggi secara berkala yang berdasarkan lima faktor yaitu kualitas sumber daya manusia, kualitas kelembagaan, kualitas kegiatan kemahasiswaan, kualitas penelitian dan pengabdian masyarakat serta kualitas inovasi [1]. Untuk mengetahui informasi kualitas perguruan tinggi dengan cara melakukan *clustering* perguruan tinggi. Pada LLDIKTI Wilayah IV pada bulan Desember 2023 tercatat ada 437 perguruan tinggi, 2.782 program studi aktif, 816.512 mahasiswa dan 29.538 dosen. Sumber data tersebut berasal dari PDDIKTI. Berdasarkan hasil observasi pada LLDIKTI Wilayah IV yang telah dilakukan terdapat permasalahan yang ditemukan yaitu belum ada *clustering* perguruan tinggi swasta mengenai perkembangan kualitas perguruan tinggi swasta di LLDIKTI Wilayah IV. Hal ini diharapkan dapat membantu dalam memberikan informasi mengenai perkembangan dan kualitas perguruan tinggi secara berkala.

Beberapa penelitian yang telah dilakukan mengenai penerapan *framework* CRISP-DM dan menggunakan metode Fuzzy C-Means diantaranya yaitu penelitian yang dilakukan oleh (I. H. Zahro, U. A. Rosyidah, and L. Handayani, 2024) yang membahas mengenai pengelompokan 34 provinsi di Indonesia berdasarkan kualitas perguruan tinggi dengan menggunakan metode fuzzy c-means. Hasil dari penelitian ini yaitu terdapat 2 *cluster* yang terdiri dari *cluster* 1 dengan 20 provinsi dan *cluster* 2 dengan 14 provinsi [2]. Penelitian yang dilakukan oleh (N. Ulinnuha and D. C. R. Novitasari, 2023) berfokus pada pengelompokan kualitas pendidikan pada kota/kabupaten di Jawa Timur dengan menggunakan metode fuzzy c-means. Terdapat 2 *cluster* yaitu *cluster* 1 dengan tingkat kualitas pendidikan tinggi yang terdiri dari 14 anggota dan *cluster* 2 dengan tingkat kualitas pendidikan rendah yang terdiri dari 24 anggota [3]. Penelitian yang dilakukan oleh (D. Rizki, P. Noordi, I. A. Prastowo, N. A. Putri, Z. A. Ekovich, dan D. Hartanti, 2023)

bertujuan untuk menganalisis sekolah dasar di Karanganyar dengan mengelompokkan tingkat pendidikan sekolah dasar di Karanganyar menggunakan metodologi CRISP-DM dan algoritma Fuzzy C-Means. Tujuan penelitian ini yaitu membantu pemerintah dan sistem pendidikan dalam mengembangkan rencana dan kebijakan untuk meningkatkan pendidikan di Karanganyar [4]. Penelitian lain yang dilakukan oleh (Fatmawaty, V. S., Riadi, I., & Herman, H., 2024) membahas mengenai pengelompokan perguruan tinggi di LLDIKTI V berdasarkan indikator kinerja utama dan PDDIKTI menggunakan K-Means *clustering*, terdiri dari 4 *cluster* yaitu *cluster* 1 dengan 4 PTS, *cluster* 2 dengan 46 PTS, *cluster* 3 dengan 21 PTS dan *cluster* 4 dengan 33 PTS [5]. Penelitian lain yang dilakukan oleh Safaat, W. A. N., Kurniawan, R., & Wijaya, Y. A., 2024) berfokus pada *clustering* data pencari kerja di kabupaten kuningan menggunakan algoritma fuzzy c-means memperoleh *cluster* 2 sebanyak 31 item [6]. Penelitian yang dilakukan oleh (Sasmita, Y., Muhsi, M., & Walid, M. 2022) bertujuan untuk mengelompokkan perguruan tinggi swasta di Madura berdasarkan kinerja sumber daya manusia dan mahasiswa dengan menggunakan metode K-Means *clustering*. Hasil dari penelitian ini yaitu terdapat 3 *cluster* yaitu *cluster* 1 dengan 4 perguruan tinggi swasta, *cluster* 2 dengan 7 perguruan tinggi swasta dan *cluster* 3 dengan 19 perguruan tinggi swasta [7]. Penelitian yang dilakukan oleh (Dikarya, F., & Muharni, S., 2022) bertujuan untuk mengelompokkan universitas terbaik menjadi 3 *cluster* yaitu *cluster* tinggi, sedang dan rendah dengan menggunakan metode K-Means yang berdasarkan 4 atribut data yaitu institusi, negara, *world rank* dan *score*. Hasil dari penelitian ini yaitu pada *cluster* 0 *National Chung Cheng University* dengan 666 universitas, *cluster* 1 *University of Haiva* dengan 667 universitas dan *cluster* 2 *Harvard University* dengan 667 universitas [8]. Penelitian yang dilakukan oleh (Jannah, B., Utami, I. T., & Hakim, A. R., 2024) memberikan gambaran *cluster* perguruan tinggi di Semarang yang menjadi panduan mahasiswa dalam memilih PTS berdasarkan ciri khas PTS dan kualitas PTS dengan menggunakan metode *cluster ensemble* ROCK (*Robust Clustering Using Links*). Hasil pengelompokan data numerik menunjukkan bahwa jumlah kelompok optimal adalah 3, sedangkan untuk data kategorik, kelompok optimalnya adalah 4. Setelah melakukan pengelompokan untuk masing-masing jenis data, serta penggabungan dan pengelompokan ulang, diperoleh kelompok optimal sebanyak 3 dengan ambang batas (θ) sebesar 0,2. Hasil pengelompokan akhir menghasilkan 29 PTS dengan kualitas rendah, 7 PTS dengan kualitas sedang, dan 1 PTS dengan kualitas tinggi [9].

Berdasarkan uraian diatas, maka penulis menemukan peluang untuk menghasilkan suatu data yang mudah diolah agar dapat memberikan suatu informasi secara konkrit dan *real* mengenai informasi perkembangan dan kualitas perguruan tinggi di LLDIKTI Wilayah IV yang menerapkan framework CRISP-DM dan metode Fuzzy C-Means dengan penelitian yang berjudul “Penerapan Framework CRISP-DM Menggunakan Fuzzy C-Means Untuk *Clustering* Data Perguruan Tinggi di LLDIKTI Wilayah IV”.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining digunakan untuk mendefinisikan informasi dalam *database*. Data mining dapat digunakan untuk melakukan perhitungan dalam pengambilan keputusan dengan cara memperoleh informasi penting dari *database* [10]. Data mining yaitu proses ekstraksi informasi dari kumpulan data dengan menggunakan teknik dan algoritma khusus atau metode-metode seperti *clustering*, asosiasi dan klasifikasi [11].

2.2 CRISP-DM

CRISP-DM atau *Cross Industry Standard for Data mining* merupakan tolak ukur dalam pengolahan data mining secara terstruktur dan efisien. Tujuan CRISP-DM adalah untuk melakukan identifikasi pola-pola yang signifikan dalam data. Dalam CRISP-DM terdiri dari beberapa tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation* dan *Deployment*. Berikut ini merupakan enam tahapan dalam CRISP-DM [12] :

a. *Business Understanding*

Tahapan *business understanding* merupakan proses yang melibatkan penentuan tujuan bisnis, memahami situasi dan kondisi saat penelitian dan menetapkan tujuan penelitian untuk masalah yang dapat diselesaikan dengan data mining.

b. *Data Understanding*

Tahapan *data understanding* merupakan proses identifikasi data yang akan digunakan yaitu dengan mengumpulkan data awal dan menentukan kualitas data.

c. *Data Preparation*

Setelah melakukan pengumpulan data dilakukan tahap selanjutnya yaitu *data preparation*. Pada tahap ini terdapat tiga proses yaitu *select data*, *cleaning data*, dan *transformation data*.

d. *Modelling*

Tahapan *modelling* merupakan tahapan implementasi metode atau algoritma yang digunakan dalam mencari, mengidentifikasi dan

menghasilkan pola untuk digunakan dalam penelitian.

e. *Evaluation*

Evaluation adalah proses pengukuran hasil evaluasi dari metode atau algoritma yang digunakan pada tahapan *modelling*.

f. *Deployment*

Tahapan *deployment* ini merupakan tahapan terakhir dari CRISP-DM. Tahapan ini merupakan tahapan pembuatan laporan mengenai hasil penelitian.

2.3 Fuzzy C-Means

Fuzzy C-Means adalah metode data mining untuk mengelompokan data. Keberadaan seluruh data dalam suatu *cluster* ditentukan oleh nilai keanggotaan dan nilai derajat keanggotaan yang berkisar antara 0 hingga 1 [13]. Berikut ini adalah tahapan dalam Fuzzy C-Means:

1. *Input* data yang akan di *cluster* (X), berupa matriks yang berukuran n x m dengan n yaitu jumlah data, m yaitu atribut pada setiap data.
2. Menentukan parameter data yang akan digunakan yaitu :
 - a. Jumlah *cluster* (c)
 - b. Pangkat (w)
 - c. Maksimum iterasi (*MaxIter*)
 - d. *Error* terkecil yang diharapkan (ε)
 - e. Fungsi objektif awal ($P_0 = 0$)
 - f. Iterasi awal (t = 0)
3. Tentukan bilangan random (μ_{ik}) sebagai matriks partisi awal U dengan menggunakan rumus :

$$\sum_{j=1}^k \mu_{ij} = 1 \quad (1)$$

4. Hitung pusat *cluster* dengan menggunakan rumus :

$$V_{kj} = \frac{\sum_{i=1}^n ((U_{ik})^w * x_{ij})}{\sum_{i=1}^n (U_{ik})^w} \quad (2)$$

5. Menghitung fungsi objektif dengan menggunakan rumus :

$$P_t = \sum_{i=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (x_{ij} - v_{kj})^2] (u_{ik})^w) \quad (3)$$

6. Menghitung perubahan matriks partisi menggunakan rumus :

$$U_{ik} = \frac{[\sum_{j=1}^m (x_{ij} - v_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (x_{ij} - v_{kj})^2]^{-\frac{1}{w-1}}} \quad (4)$$

Memeriksa kondisi berhenti yaitu Jika $|P_t - P_{t-1}| < \epsilon$ atau $t < \text{MaxIter}$ maka iterasi berhenti. Jika tidak memenuhi syarat tersebut, maka iterasi dilanjutkan.

2.4 Davies Bouldin Index (DBI)

Davies Bouldin Index atau disingkat DBI adalah salah satu metode yang digunakan untuk melakukan evaluasi hasil *clustering*. Penggunaan metode DBI memiliki kelebihan yaitu memaksimalkan jarak antar *cluster* dengan *sum of square between cluster* dan meminimalkan jarak antar titik dalam satu *cluster* dengan *sum of square within cluster* sehingga akan menghasilkan *clustering* yang optimal. Dalam DBI terdapat tahapan-tahapan perhitungan yaitu sebagai berikut [14]:

- a. *Sum Of Square Within-Cluster (SSW)*

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(X_j, C_j) \quad (5)$$

m_i adalah untuk jumlah data dalam *cluster* ke- i . C_j adalah *centroid cluster* ke- i . Dan $d(X_j, C_j)$ adalah jarak setiap data ke *centroid* i yang dihitung menggunakan jarak.

- b. *Sum Of Square Between-Cluster (SSB)*

$$SSB_{ij} = d(X_i, X_j) \quad (6)$$

$d(X_j, C_j)$ adalah jarak antara data ke i dengan data ke i di *cluster* lain.

- c. *Ratio*

$$R_{ij, \dots, n} = \frac{SSW_i + SSW_j + \dots + SSW_n}{SSB_{i,j} + \dots + SSB_{n,i,j}} \quad (7)$$

SSW_i adalah *Sum Of Square Within-Cluster* pada *centroid* i . $SSB_{i,j}$ adalah *Sum Of Square Between-Cluster* data ke i dengan j pada *cluster* yang berbeda.

- d. *Davies Bouldin Index*

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_i, j, \dots, k) \quad (8)$$

R_i, j adalah rasio dari nilai SSW dan SSB . Maka diketahui k adalah jumlah *cluster*.

3. HASIL DAN PEMBAHASAN

3.1. Implementasi CRISP-DM

3.1.1 Business Understanding

Pada tahapan ini merupakan tahapan untuk menentukan tujuan bisnis dan tujuan dari dilakukannya data mining. Tujuannya yaitu untuk melakukan *clustering* data perguruan tinggi swasta di LLDIKTI Wilayah IV. Hal ini diharapkan dapat membantu dalam memberikan informasi mengenai

perguruan tinggi swasta dan melakukan peningkatan kualitas perguruan tinggi secara berkala.

3.1.2 Data Understanding

Pada tahapan ini yaitu melakukan pengumpulan data awal dan menentukan kualitas data. Penelitian ini menggunakan data perguruan tinggi swasta di LLDIKTI Wilayah IV pada Desember tahun 2023 yang bersumber dari PDDIKTI. Adapun data tersebut dapat dilihat pada Tabel 1.

Tabel 1. Data Perguruan Tinggi Swasta di LLDIKTI Wilayah IV

No.	Kode PTS	Akte PTS	Jmlh Prodi	Jmlh Mhs	Jmlh Dsn
1.	041001	B	36	37	8488
2.	041002	Unggul	33	38	14066
3.	041003	BS	26	31	4748
4.	041004	B	38	47	18594
5.	041005	B	19	19	3891

437	046005	-	2	2	0

3.1.3 Data Preparation

Dalam data preparation terdapat 3 proses untuk menyiapkan data yang akan digunakan yaitu proses *select data*, *cleaning data*, dan *transformation data*. Berikut ini merupakan proses *data preparation*:

- a. *Select data*

Menentukan atribut data yang akan digunakan untuk *clustering* data perguruan tinggi swasta di LLDIKTI Wilayah IV. Pada penelitian ini atribut yang akan digunakan yaitu nama perguruan tinggi, kode perguruan tinggi, akreditasi perguruan tinggi, jumlah program studi, jumlah mahasiswa dan jumlah dosen.

- b. *Cleaning data*

Cleaning data dilakukan dengan cara menghilangkan atau menghapus duplikat data atau data-data yang kosong. Data awal perguruan tinggi yaitu terdapat 437 perguruan tinggi, 2.782 program studi aktif, 816.512 mahasiswa dan 29.538 dosen. Setelah dilakukan *cleaning data* pada data perguruan tinggi di LLDIKTI Wilayah IV terdapat 315 perguruan tinggi, 1.827 program studi, 464.869 mahasiswa dan 18.393 dosen. Untuk data PTS yang sudah dilakukan proses *cleaning data* dapat dilihat pada Tabel 2.

Tabel 2. Cleaning Data

No.	Kode PTS	Akte PTS	Jmlh Prodi	Jmlh Mhs	Jmlh Dsn
1.	041002	Unggul	35	14066	616
2.	041003	BS	26	4748	261

3.	041006	Unggul	37	10839	455
.....
315	046003	-	1	58	5

c. Transformation data

Transformation data dilakukan pada atribut data yang digunakan yaitu akreditasi perguruan tinggi, jumlah program studi, jumlah mahasiswa dan jumlah dosen yang dapat dilihat pada Tabel 3.

Tabel 3. Transformation Data

No.	Kode PTS	Akte PTS	Jmlh Prodi	Jmlh Mhs	Jmlh Dsn
1.	041002	4	3	3	3
2.	041003	3	3	3	3
3.	041006	4	3	3	3
.....
315	046003	1	1	3	2

Transformasi data dilakukan pada atribut data akreditasi perguruan tinggi dengan mengubah format data dari data kategorikal menjadi data numeric. Selain itu, dilakukan pembobotan nilai pada atribut data jumlah program studi, jumlah mahasiswa dan jumlah dosen. Pembobotan nilai ini berdasarkan kriteria-kriteria tertentu yang telah ditentukan.

3.1.4 Modelling

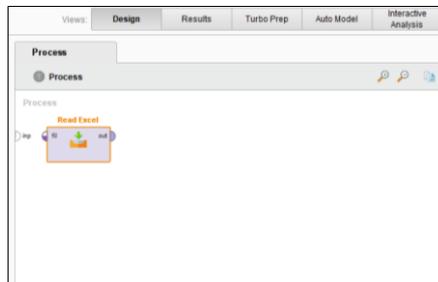
Model atau metode data mining clustering yang digunakan pada penelitian ini yaitu algoritma Fuzzy C-Means. Metode ini digunakan untuk menghasilkan cluster-cluster data perguruan tinggi swasta di LLDIKTI Wilayah IV yang terdiri dari dua cluster yaitu cluster 1 (Ideal) dan cluster 2 (memadai). Hal bertujuan untuk mempermudah dalam memberikan informasi mengenai perkembangan dan kualitas perguruan tinggi swasta di LLDIKTI Wilayah IV. Terdapat kriteria-kriteria clustering yang digunakan yaitu sebagai berikut :

1. Cluster 1- Ideal
 - a. Akreditasi perguruan tinggi unggul, baik sekali dan baik.
 - b. Dari 4 variabel data seperti akreditasi perguruan tinggi, jumlah program studi, jumlah mahasiswa dan jumlah dosen. Terdapat 2 variabel data yang memenuhi kriteria.
2. Cluster 2 – Memadai
 - a. Perguruan tinggi tidak terakreditasi
 - b. Jumlah program studi tidak memenuhi kriteria.

Berikut ini merupakan implementasi model Fuzzy C-Means dengan menggunakan Rapidminer :

- a. Import dataset

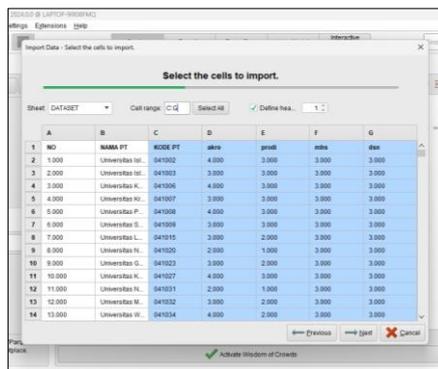
Import dataset menggunakan operator Read Excel. Operator ini digunakan untuk membaca file excel yang telah diinputkan. File tersebut berisi data yang akan digunakan dalam penelitian ini yaitu dataset PTS yang telah melalui tahap data preparation. Untuk import dataset dapat dilihat pada Gambar 1.



Gambar 1. Import Dataset

- b. Pemilihan atribut data dan role id

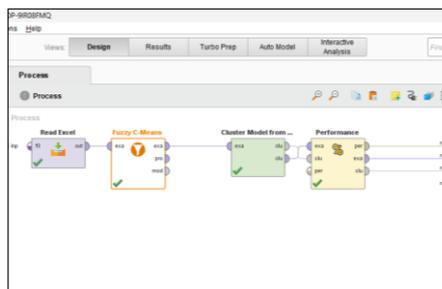
Pemilihan atribut data yang akan digunakan dalam penelitian ini yaitu kode PT, akreditasi PT, jumlah program studi, jumlah mahasiswa dan jumlah dosen. Data tersebut dapat dilihat pada Gambar 2.



Gambar 2. Pemilihan atribut data dan role id

- c. Proses Clustering Fuzzy C-Means

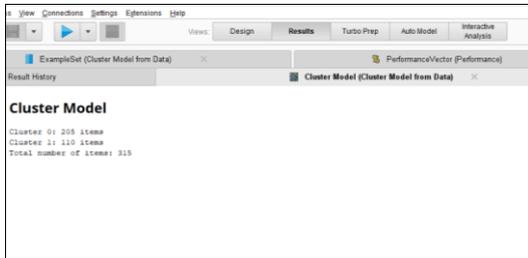
Pada Gambar 3 menunjukan bahwa Proses clustering Fuzzy C-Means menggunakan operator Fuzzy C-Means, Cluster Model From Data dan Cluster Distance Performance. Parameter yang digunakan yaitu jumlah cluster, iterasi dan fuzzyness.



Gambar 3. Proses Clustering Fuzzy C-Means

d. Hasil Clustering

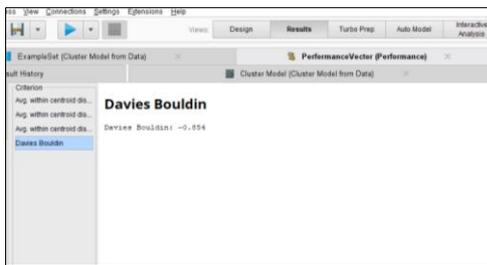
Pada Gambar 4 menunjukkan hasil clustering Fuzzy C-Means yang menunjukkan bahwa terdapat 2 cluster dari 315 PTS yaitu cluster 0 atau cluster 1 dengan kategori Ideal terdapat 205 PTS dan cluster 1 atau cluster 2 dengan kategori memadai terdapat 110 PTS.



Gambar 4. Hasil Clustering

3.1.5 Evaluation

Tahap evaluasi dilakukan menggunakan Rapidminer. Tahapan ini dilakukan untuk mengukur tingkat akurasi pada metode atau algoritma Fuzzy C-Means dalam melakukan clustering perguruan tinggi swasta di LLDIKTI Wilayah IV. Pengujian ini dilakukan dengan metode *Davies Bouldin Index* atau DBI. Berikut ini merupakan hasil evaluasi DBI menggunakan Rapidminer :



Gambar 5. Evaluation

Dari Gambar 5 menunjukkan bahwa hasil evaluasi menggunakan *Davies Bouldin Indeks* yaitu -0,854. Maka dapat disimpulkan bahwa hasil clustering kurang baik. Nilai DBI yang negatif menunjukkan adanya kesalahan dalam penggunaan metode Fuzzy C-Means, karena secara teori nilai DBI seharusnya bernilai positif semakin kecil nilainya semakin baik hasil clustering.

3.1.6 Deployment

Pada tahap deployment ini berupa dashboard yang menampilkan informasi mengenai hasil clustering perguruan tinggi swasta di LLDIKTI Wilayah IV dengan menyajikan visualisasi data yang dapat mudah dipahami.

3.2 Tampilan Website

1. Tampilan Login pada website



Gambar 6. Tampilan Login

Gambar 6 merupakan tampilan login pada website yaitu dengan melakukan input username dan password.

2. Tampilan Dashboard pada website



Gambar 7. Tampilan Dashboard

Gambar 7 merupakan tampilan untuk tampilan dashboard yang menampilkan jumlah PTS, jumlah program studi, jumlah mahasiswa dan jumlah dosen.

3. Tampilan Menu Data PTS pada website

No	Kode PT	Nama PT	Akreditasi	Jumlah Program Studi	Jumlah Mahasiswa	Jumlah Dosen
1	41002	Universitas Islam Bandung	Unggul	35	14056	616
2	41003	Universitas Islam Nusantara	Baik Sekali	26	4748	261
3	41006	Universitas Katolik Parahyangan	Unggul	37	10839	455
4	41007	Universitas Kristen Maranatha	Baik Sekali	32	6670	449
5	41008	Universitas Pasundan	Unggul	43	22748	514
6	41009	Universitas Swadaya Gunung Jati	Baik Sekali	26	13365	412
7	41015	Universitas Langlang Buana	Baik Sekali	21	5631	266
8	41020	Universitas Nusa Tenggara	Baik Sekali	9	1002	67
9	41023	Universitas Galuh	Baik Sekali	26	9029	326
10	41027	Universitas Komputer Indonesia	Unggul	29	10528	296

Gambar 8. Tampilan Menu Data PTS

Gambar 8 merupakan tampilan untuk menu data PTS yang menampilkan data PTS yang terdiri dari kode PT, nama PT, akreditasi PT, jumlah program studi, jumlah mahasiswa dan jumlah dosen.

a. Tampilan Menu Data Clustering



Gambar 9. Tampilan Menu Clustering

Gambar 9 merupakan tampilan untuk menu clustering yang menampilkan hasil clustering data

perguruan tinggi dan visualisasi hasil *clustering* yang terdiri dari 2 *cluster* yaitu *cluster* 1 (ideal) dengan total 205 PTS dan *cluster* 2 (memadai) dengan total 110 PTS. Selain itu, menampilkan kode PT, nama PT dan hasil *clustering*.

4. KESIMPULAN

Berdasarkan hasil *clustering* pada data perguruan tinggi di LLDIKTI Wilayah IV menggunakan metode Fuzzy C-Means berdasarkan akreditasi perguruan tinggi, jumlah program studi, jumlah mahasiswa dan jumlah dosen. Data dari PDDIKTI Desember 2023 mencatat 437 perguruan tinggi, 2.782 program studi, 816.512 mahasiswa, dan 29.538 dosen. Setelah dilakukan *cleaning* data, terdapat 315 perguruan tinggi, 1.827 program studi, 464.869 mahasiswa, dan 18.393 dosen yang digunakan untuk *clustering*. Hasil *clustering* menunjukkan 205 PTS terdapat pada *cluster* 1 dengan kategori ideal dan pada 110 PTS terdapat pada *cluster* 2 dengan kategori memadai. Hasil *clustering* tersebut dilakukan evaluasi menggunakan metode DBI yaitu didapatkan nilai -0,854 yang menunjukkan bahwa hasil *clustering* kurang baik, karena nilai DBI yang negatif menunjukkan adanya kesalahan dalam penggunaan metode Fuzzy C-Means, karena secara teori nilai DBI seharusnya bernilai positif semakin kecil nilainya semakin baik hasil *clustering*. Hasil akhir dari penelitian ini yaitu berupa *dashboard* yang menampilkan informasi mengenai perkembangan dan kualitas perguruan tinggi dan menampilkan visualisasi hasil *clustering* yang dapat mudah dipahami. Penelitian selanjutnya dapat mempertimbangkan untuk menggunakan metode *clustering* yang berbeda, seperti K-Means atau *clustering* lainnya untuk meningkatkan akurasi dan hasil yang lebih optimal. Selain itu, perbaikan dalam pembersihan data dan penambahan variabel lain, seperti rasio dosen-mahasiswa atau infrastruktur pendidikan, dapat membantu memberikan gambaran yang lebih komprehensif. Pengembangan *dashboard* yang lebih interaktif dan integrasi dengan data *real-time* juga menjadi peluang untuk meningkatkan pemantauan kualitas perguruan tinggi secara berkelanjutan.

5. REFERENCES

[1] M. Rozikin, M. Kurniawati, and K. Aliyyah, "Strategi Peningkatan Peringkat Perguruan Tinggi Versi Kemristekdikti (Studi pada Universitas Brawijaya)," *Kelola: Jurnal*

Manajemen Pendidikan, vol. 7, no. 1, pp. 86–97, 2020, doi: 10.24246/j.jk.2020.v7.i1.p86-97.

- [2] I. H. Zahro, U. A. Rosyidah, and L. Handayani, "Implementasi Algoritma Fuzzy C-Means untuk Pengelompokan Provinsi di Indonesia Berdasarkan Kualitas Perguruan Tinggi," *Jurnal Teknologi Informasi dan Rekayasa Komputer*, vol. 5, no. 1, pp. 80–86, 2024, doi: <https://doi.org/10.37148/bios.v5i1.102>.
- [3] A. R. Maulana, N. Ulinuha, and D. C. R. Novitasari, "PENERAPAN FUZZY C-MEANS UNTUK PENGELOMPOKKAN TINGKAT KUALITAS PENDIDIKAN DI JAWA TIMUR," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 14, no. 2, pp. 419–426, Nov. 2023, doi: 10.24176/simet.v14i2.9442.
- [4] D. Rizki, P. Noordi, I. A. Prastowo, N. A. Putri, Z. A. Ekovich, and D. Hartanti, "Penerapan Algoritma Fuzzy C-Means untuk Analisis Tingkat Pendidikan Sekolah Dasar di Karanganyar," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis*, pp. 12–18, 2023, doi: <https://doi.org/10.47701/senatib.v3i1>.
- [5] V. S. Fatmawaty, I. Riadi, and Herman, "Klasterisasi Perguruan Tinggi LLDIKTI V Berdasarkan Indikator Kinerja Utama dan PDDIKTI Menggunakan K-Means Clustering," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, pp. 878–888, 2024, doi: 10.30865/mib.v8i2.7497.
- [6] W. An-naziz Safaat, R. Kurniawan, and Y. Arie Wijaya, "Penerapan Data Mining Clustering Menggunakan Algoritma Fuzzy C-Means Pada Data Pencari Kerja Di Kabupaten Kuningan," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1507–1511, 2024, doi: 10.36040/jati.v8i2.8411.
- [7] Y. Sasmita, Muhsi, and M. Walid, "Klasterisasi Perguruan Tinggi Swasta di Madura Berdasarkan Kinerja Sumber Daya Manusia dan Mahasiswa Menggunakan Metode K-Means Clustering," *Jurnal Media Informatika Budidarma*, vol. 6, no. 4, p. 2157, 2022, doi: 10.30865/mib.v6i4.4431.
- [8] F. Dikarya and S. Muharni, "PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PENGELOMPOKKAN UNIVERSITAS TERBAIK DI DUNIA," *Jurnal Informatika*, vol. 22, no. 2, pp. 124–131, 2022, doi: 10.31294/coscience.v2i1.623.
- [9] B. Jannah, I. T. Utami, and A. R. Hakim, "Metode Ensemble Robust Clustering Using Links (ROCK) Untuk Pengelompokan Perguruan Tinggi Swasta (PTS) Di Kota Semarang," *Jurnal Gaussian*, vol. 12, no. 3, pp. 445–452, 2024, doi: 10.14710/j.gauss.12.3.445-452.

- [10] Oki Oktaviarna Tensao, I Nyoman Yudi Anggara Wijaya, and Ketut Queena Fredlina, "Analisa Data Mining dengan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Pada STMIK Primakara," *INFORMASI (Jurnal Informatika dan Sistem Informasi)*, vol. 14, no. 1, pp. 1–17, 2022, doi: 10.37424/informasi.v14i1.135.
- [11] M. Yamin Nurzaman and B. Nurina Sari, "Implementasi K-Means Clustering Dalam Pengelompokan Banyaknya Jumlah Petani Berdasarkan Kecamatan Di Provinsi Jawa Barat," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 10, no. September, pp. 131–144, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>
- [12] D. A. Dzulhijjah, M. B. Herlambang, and M. Haifan, "Implementasi Framework CRISP-DM untuk Proses Data Mining Aplikasi Credit Scoring PT.XYZ," *Prosiding Seminar Nasional Teknologi Informasi dan Bisnis*, vol. 1, no. 1, pp. 238–251, 2023.
- [13] V. Herlinda, D. Darwis, and Dartono, "Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means," *Jurnal Teknologi dan Sistem Informasi (JTSI)*, vol. 2, no. 2, pp. 94–99, 2021, doi: <https://doi.org/10.33365/jtsi.v2i2.890>.
- [14] F. Sandova, R. Kurniawan, and T. Supratati, "Penerapan Data Mining Menggunakan Metode K-Means Clustering Pada Penjualan Tas Di Asia Toserba Cirebon," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 245–251, 2024, doi: 10.36040/jati.v8i1.8330.