



Klasterisasi Perguruan Tinggi LLDIKTI V Berdasarkan Indikator Kinerja Utama dan PDDIKTI Menggunakan K-Means Clustering

Virdiana Sriviana Fatmawaty^{1,*}, Imam Riadi², Herman¹

¹Program Studi Magister Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

²Program Studi Sistem Informasi, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Email: ^{1,*}2307048011@webmail.uad.ac.id, ²imam.riadi@is.uad.ac.id, ³hermankaha@mti.uad.ac.id

Email Penulis Korespondensi: 2307048011@webmail.uad.ac.id

Abstrak—Pertumbuhan jumlah perguruan tinggi yang terus meningkat menjadi salah satu faktor krusial untuk memastikan mutu pendidikan tinggi agar berdaya saing. Komposisi perguruan tinggi di Provinsi Daerah Istimewa Yogyakarta terdiri atas 77% Perguruan Tinggi Swasta (PTS) dan sisanya adalah Perguruan Tinggi Negeri (PTN) menjadi tantangan tersendiri bagi PTS untuk mengoptimalkan pengelolaan, perencanaan strategis untuk peningkatan mutu pendidikan. Saat ini peran LLDIKTI Wilayah V sebagai Lembaga Layanan Pendidikan Tinggi di Yogyakarta adalah memfasilitasi peningkatan mutu penyelenggaraan pendidikan tinggi dengan melaksanakan pemetaan mutu perguruan tinggi khususnya PTS melalui Indikator Kinerja Utama (IKU) dan data Pangkalan Data Pendidikan Tinggi (PDDIKTI) yang wajib dilaporkan dan dipenuhi oleh PTS masih kesulitan dalam melakukan indentifikasi PTS berdasarkan mutu, sehingga perlu dilakukan pengelompokan PTS yang selama ini belum pernah dilakukan oleh peneliti lain. Salah satu model analisis yang dapat dilakukan adalah dengan metode klasterisasi. Oleh karena itu Metode ini biasa digunakan pada data mining untuk mengelompokkan data berdasarkan kesamaan karakteristik data. Penelitian ini melakukan klasterisasi PTS di LLDIKTI Wilayah V menggunakan algoritma K-Means Clustering. Hasil penelitian ini menunjukkan bahwa berdasarkan kesamaan karakteristik data IKU dan data PDDIKTI terbentuk empat klaster PTS, yaitu klaster 1 terdiri dari 4 PTS, klaster 2 terdiri dari 46 PTS, klaster 3 terdiri dari 21 PTS, dan klaster 4 terdiri dari 33 PTS. Hasil analisis ini akan sangat bermanfaat bagi LLDIKTI Wilayah V dalam melakukan fungsi pembinaan kepada PTS.

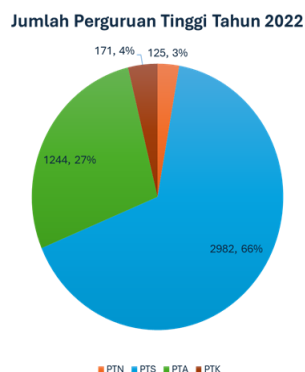
Kata Kunci: Klasterisasi; PTS; K-Means; Indikator Kinerja Utama; PDDIKTI

Abstract—The continuous growth in the number of higher education institutions is becoming a crucial factor in ensuring the competitiveness of higher education quality. The composition of higher education institutions in the Special Region of Yogyakarta Province consists of 77% Private Higher Education Institutions (PTS) and the rest are State Higher Education Institutions (PTN), posing a unique challenge for PTS to optimize management and strategic planning for educational quality improvement. Currently, the role of LLDIKTI Region V as a Higher Education Service Agency in Yogyakarta is to facilitate the improvement of the quality of higher education implementation by conducting quality mapping of higher education institutions, especially PTS, through Main Performance Indicators (IKU) and Higher Education Database (PDDIKTI) data, which must be reported and fulfilled by PTS. However, there are still difficulties in identifying PTS based on quality, thus the need for clustering of PTS, which has not been done by other researchers. One analytical model that can be used is the clustering method. Therefore, this method is commonly used in data mining to group data based on similar data characteristics. This study conducts clustering of PTS in LLDIKTI Region V using the K-Means Clustering algorithm. The results of this study show that based on the similarity of IKU and PDDIKTI data characteristics, four clusters of PTS are formed, namely cluster 1 consisting of 4 PTS, cluster 2 consisting of 46 PTS, cluster 3 consisting of 21 PTS, and cluster 4 consisting of 33 PTS. The results of this analysis will be very beneficial for LLDIKTI Region V in carrying out its mentoring functions for PTS.

Keywords: Clustering; PTS; K-Means; Key Performance Indicators; PDDIKTI

1. PENDAHULUAN

Pesatnya peningkatan jumlah perguruan tinggi di Indonesia telah meningkatkan juga tingkat persaingan dalam mendapatkan calon mahasiswa. Kualitas dan mutu perguruan tinggi menjadi isu krusial dalam persaingan tersebut [1], [2]. Berdasarkan [3], tercatat ada 4.522 jumlah perguruan tinggi di Indonesia dengan rincian 125 perguruan tinggi negeri (PTN), 2.982 PTS, 1.244 perguruan tinggi dibawah Kementerian Agama (PTA), dan 171 perguruan tinggi dibawah Kementerian/Lembaga lain (PTK) yang ditunjukkan pada Gambar 1.



Gambar 1. Jumlah Persebaran Perguruan Tinggi di Indonesia pada Tahun 2022



PTS mendominasi persebaran jumlah perguruan tinggi di Indonesia yaitu 66%. Provinsi Daerah Istimewa Yogyakarta (DIY) sebagai kota pelajar tercatat memiliki 135 perguruan tinggi dimana 104 adalah PTS [3], sehingga mencapai persentase 77% menjadi urgensi bagi PTS di Yogyakarta untuk meningkatkan mutu pendidikan secara menyeluruh dan tantangan bagi LLDIKTI Wilayah V yang saat ini kesulitan memetakan PTS berdasarkan mutu sebagai dasar untuk memfasilitasi PTS mencapai mutu unggul.

Kriteria Penilaian Indikator Kinerja Utama Perguruan Tinggi (IKU-PT) sebagai salah satu upaya Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) untuk memantau sistem penjaminan mutu internal di lembaga pendidikan tinggi [4]–[6]. Sumber data penilaian Indikator Kinerja Utama Perguruan Tinggi salah satunya adalah pelaporan Pangkalan Data Pendidikan Tinggi (PDDIKTI) sebagaimana [7] menyatakan PDDIKTI sebagai sumber informasi utama bagi sistem penjaminan mutu internal (SPMI) perguruan tinggi dan sistem penjaminan mutu eksternal (SPME) perguruan tinggi.

PDDIKTI sebagai lembaga pusat untuk pengelolaan data pendidikan tinggi, bertanggung jawab untuk mengumpulkan, menyusun, dan menyediakan data yang dapat digunakan untuk evaluasi kinerja perguruan tinggi [8]–[10]. Salah satu upaya mendorong perguruan tinggi terutama PTS untuk mencapai standar Indikator Kinerja Utama Perguruan Tinggi (IKU-PT), perlu dilakukan pemetaan PTS berdasarkan hasil penilaian IKU-PT dan pelaporan PDDIKTI.

Pemetaan PTS dapat dilakukan dengan menggunakan metode klusterisasi. Klusterisasi merupakan suatu teknik yang digunakan untuk proses pengelompokan data berdasarkan kemiripan [11]. Salah satu metode paling populer untuk klusterisasi yaitu K-Means Clustering [12] sebuah metode dalam bidang data mining, mempertahankan karakteristik non-hierarchy dalam proses analisisnya. Metode ini secara khusus dimanfaatkan untuk mengelompokkan data menjadi lebih dari satu kluster yang memiliki kemiripan karakteristik [13]. Kelebihan lain dari K-Means Clustering yaitu dinilai implementasi mudah dan hasilnya mudah dipahami [14].

Penelitian [15] membahas pemeringkatan perguruan tinggi di Indonesia dengan memanfaatkan metode Geometrical Analysis for Interactive Aid (GAIA) dan Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE). Penelitian ini menggunakan indikator/atribut aktivitas mahasiswa, standar institusi, kualitas pengajar, riset dan kontribusi sosial, serta kemajuan dalam inovasi. Penelitian membagi menjadi empat kluster. Data yang digunakan adalah data kinerja perguruan tinggi dengan menggunakan sampel lima perguruan tinggi vokasi.

Hasil penelitian perguruan tinggi alternatif 2 (ϕ 0;5125) memiliki nilai terbaik diikuti perguruan tinggi alternatif 3 dan perguruan tinggi alternatif 1. Penelitian lainnya [16] tentang pengelompokan perguruan tinggi negeri (PTN) dengan berdasarkan jumlah publikasi ilmiah. Proses pengelompokan data dilakukan dengan membaginya ke dalam tiga kluster yang masing-masing diberi label tinggi, sedang, dan rendah. Pengelompokan dilakukan dengan menerapkan dua metode, yaitu K-Medoids dan K-Means Clustering.

Tercatat bahwa menggunakan metode K-Means Clustering mendapatkan nilai yang lebih tinggi (0,8018) daripada menggunakan metode K-Medoids (0,7281) dengan hasil pengelompokan berdasarkan K-Means Clustering adalah PTN kelompok tinggi adalah ITB, ITS, UGM, dan UI, kelompok sedang ada 16 PTN, dan kelompok rendah ada 102 PTN. Penelitian menganalisa karakteristik perguruan tinggi juga dilakukan [17] dengan sampel tujuh universitas di Barranquilla, Kolombia dengan berdasarkan proses pembelajaran perguruan tinggi yang terdiri dari proses pembelajaran individu dan proses pembelajaran bersama. Metode K-Means Clustering dan Principal Component Analysis (PCA) digunakan untuk analisis pada penelitian ini. Hasil yang didapat menunjukkan kluster pertama memiliki ciri-ciri tingginya tingkat indikator pembelajaran kolektif, sumber eksternal, dan pentingnya kerja sama dalam melakukan penelitian bersama. kluster kedua memiliki tingkat yang tinggi dalam variabel pembelajaran individu dan sumber internal, menunjukkan pentingnya peningkatan kemampuan penelitian internal.

Penelitian [18] berfokus pada PTS di wilayah Madura dengan menggunakan indikator mahasiswa dan evaluasi kinerja pada Sumber Daya Manusia (SDM). Penelitian yang membagi PTS ke dalam tiga kluster menunjukkan hasil bahwa kluster 1 terdiri dari 4 PTS, kluster 2 memiliki 7 PTS, dan kluster 3 mencakup 19 PTS. Metode K-Means Clustering diterapkan pada penelitian ini. Penelitian lain oleh [19] yang meneliti tentang pengelompokan mahasiswa berdasarkan tingkat kesehatan mental dengan algoritma Fuzzy C-Means berbasis Heuristik (HFCA). Hasil penelitian mencatat bahwa pendekatan HFCA yang diusulkan berhasil mencapai tingkat kinerja mahasiswa sebesar 96,7%, tingkat perkembangan kognitif sebesar 97,2%, tingkat keterlibatan mahasiswa sebesar 97,5%, dan tingkat prediksi sebesar 95,1%. Penelitian [20] tentang prediksi kinerja akademis mahasiswa pada pendidikan tinggi di Kerala, India dengan memanfaatkan K-means Clustering sebagai metode klusterisasi dan SVM Classifier, Naïve Bayes Classifier, Decision Tree Classifier, dan Neural Network Classifier sebagai metode klasifikasi. Hasil penelitian mencatat untuk tingkat akurasi yaitu 0,7547 lebih baik daripada hanya menggunakan salah satu metode.

Permasalahan yang akan diselesaikan adalah saat ini LLDIKTI Wilayah V kesulitan untuk memetakan PTS berdasarkan mutu pendidikan yaitu data IKU-PT dan data pelaporan PDDIKTI. Penelitian ini menawarkan model klusterisasi dengan menggunakan K-Means Clustering. Model ini diharapkan dapat dikembangkan sebagai modul sistem monitoring penjaminan mutu LLDIKTI Wilayah V yang akan mengklusterkan PTS di LLDIKTI Wilayah V secara otomatis.

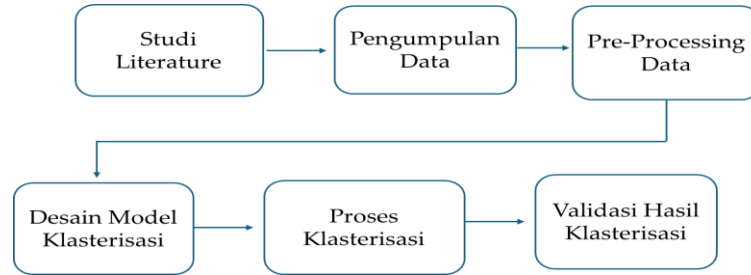


2. METODOLOGI PENELITIAN

Metodologi penelitian adalah serangkaian kegiatan, aturan, dan prosedur yang diterapkan oleh peneliti dalam suatu disiplin ilmu tertentu [21]. Penelitian ini mengadopsi pendekatan kuantitatif, suatu jenis penelitian yang berfokus pada pengumpulan dan analisis data numerik menggunakan prosedur statistik atau metode lainnya dari suatu pengukuran kuantitatif. Tujuannya adalah untuk mendapatkan temuan baru yang dapat diinterpretasikan secara objektif melalui penggunaan pendekatan kuantitatif dalam analisis data [22].

2.1 Tahapan Penelitian

Secara umum, ilustrasi proses penelitian dapat dilihat pada Gambar 2, yang mencakup langkah-langkah dan tahapan yang ditempuh dalam rangka menjalankan metodologi dari penelitian.



Gambar 2. Tahapan Penelitian Klasterisasi PTS

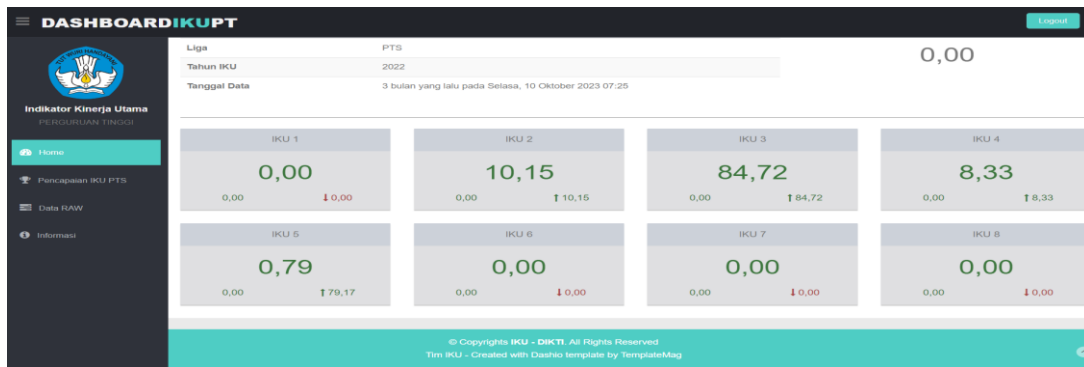
Langkah-langkah proses penelitian pada Gambar 2 terdapat tujuh tahapan yaitu :

1. Literature Study

Penelitian pendahuluan, sebagai langkah awal dalam proses penelitian, merupakan pendekatan yang digunakan untuk menggali data, sumber daya, dan tema yang relevan dengan fokus pada penelitian yang akan atau sedang dijalankan [23]. Penelitian ini memanfaatkan referensi dari jurnal ilmiah, buku, dan karya ilmiah lainnya dalam mendapatkan informasi penelitian tentang IKU-PT, PDDIKTI, klasterisasi, dan metode yang digunakan.

2. Pengumpulan Data

Tahap awal dari penelitian melibatkan pengumpulan data dengan mengakses laman <https://iku-pt.kemdikbud.go.id/> untuk mendapatkan data persentase IKU-PT tahun 2022. Selain itu, data persentase pelaporan PDDIKTI tahun 2022 diperoleh dengan mengakses laman <https://pddikti.kemdikbud.go.id/>. Informasi lebih rinci ditunjukkan pada Gambar 3 dan Gambar 4.



Gambar 3. <https://iku-pt.kemdikbud.go.id/>



Gambar 4. <https://pddikti.kemdikbud.go.id/>



Hasil penarikan data IKU-PT dan pelaporan PDDIKTI digabung dengan bantuan Microsoft Excel dengan mendapatkan sembilan indikator yaitu delapan indikator dari IKU-PT dan satu indikator dari pelaporan PDDIKTI. Penelitian ini akan menggunakan dataset yang mencakup 104 PTS dengan delapan indikator IKU-PT dan satu indikator pelaporan PDDIKTI dengan menggunakan standar atau aturan pada Tabel 1.

Tabel 1. Indikator IKU-PT dan Pelaporan Data PDDIKTI

Indikator	Keterangan	Satuan
I1 IKU 1	Persentase lulusan program Sarjana (S1) dan program Diploma Empat/Diploma Tiga/Diploma Dua/Diploma Satu (D4/D3/D2/D1) yang berhasil mendapatkan pekerjaan, studi lanjut, atau wirausaha, diperhitungkan terhadap jumlah lulusan	%
I2 IKU 2	Persentase mahasiswa yang mengikuti program Merdeka Belajar Kampus Merdeka (MBKM) dan mendapatkan prestasi yang memenuhi syarat, dihitung terhadap jumlah mahasiswa yang mengikuti MBKM dan prestasi mahasiswa	%
I3 IKU 3	Persentase dari jumlah dosen tetap yang terlibat dalam tridharma perguruan tinggi di luar institusi, berpraktik di industri, atau membimbing mahasiswa yang mengikuti kegiatan di luar program studi (prodi) dihitung terhadap jumlah dosen tetap (NIDN)	%
I4 IKU 4	Persentase dosen tetap (NIDN) atau dosen tidak tetap (NIDK) yang mempunyai sertifikat kompetensi (serkom) atau sertifikat profesi (serprof), serta jumlah pengajar dari praktisi profesional di industri atau dunia kerja yang dihitung terhadap total jumlah dosen tetap (NIDN), dosen tidak tetap (NIDK), dan yang berstatus non-dosen (NUP)	%
I5 IKU 5	Persentase rasio jumlah karya/penelitian dosen tetap yang ber NIDN atau tidak tetap yang ber NIDK yang mendapatkan pengakuan internasional atau digunakan oleh masyarakat/industri pemerintah dihitung terhadap total jumlah dosen berstatus dosen tetap dan dosen tidak tetap	%
I6 IKU 6	Persentase rasio jumlah kerjasama yang memenuhi kriteria pada prodi S1 atau D4/D3/D2/D1 dihitung terhadap total jumlah program studi S1 dan D4/D3/D2/D1.	%
I7 IKU 7	Persentase dari total jumlah mata kuliah yang kelasnya diselenggarakan pada tahun berjalan, yang menggunakan metode pembelajaran case method atau team-based project, serta diintegrasikan sebagai bagian dari bobot evaluasi, dihitung terhadap total jumlah mata kuliah tersebut.	%
I8 IKU 8	Persentase jumlah prodi S1 dan D4/D3 yang memegang akreditasi atau sertifikat internasional yang diakui pemerintah dihitung terhadap total jumlah prodi S1 dan D4/D3 dan minimal meluluskan satu kali lulusan.	%
P1 Pelaporan PDDIKTI	Persentase pelaporan PDDIKTI	%

3. Pre-processing Data

Hasil pengumpulan data IKU-PT dan pelaporan PDDIKTI menghasilkan data awal yang membutuhkan proses pre-processing data dengan dua tahap :

- 1) Integrasi data, yaitu proses penggabungan delapan indikator IKU-PT dan satu indikator data pelaporan PDDIKTI menjadi satu dataset.
- 2) Transformasi Data, yaitu melakukan proses mengatasi missing value dan normalisasi data. Hal ini diperlukan karena sembilan indikator memiliki nilai maksimum dan minimum yang tidak sama, sebagai contoh, pada indikator jumlah keluaran dosen yang berhasil mendapatkan rekognisi internasional atau diterapkan oleh masyarakat/industri/pemerintah per jumlah dosen, dapat mencapai lebih dari 100%. Sementara itu, pada indikator persentase lulusan S1 dan D4/D3/D2/D1 yang berhasil memperoleh pekerjaan, studi lanjut, atau menjadi wiraswasta, dapat mencapai kurang dari satu persen. Normalisasi menggunakan teknik min-max scaler dengan menghindari masalah numerik data yang tidak seragam dengan mengkonversi atribut pada range nilai [0,1] untuk optimalisasi klusterisasi [14]. Menghitung min-max scaler dengan rumus yang ditunjukkan pada persamaan 1.

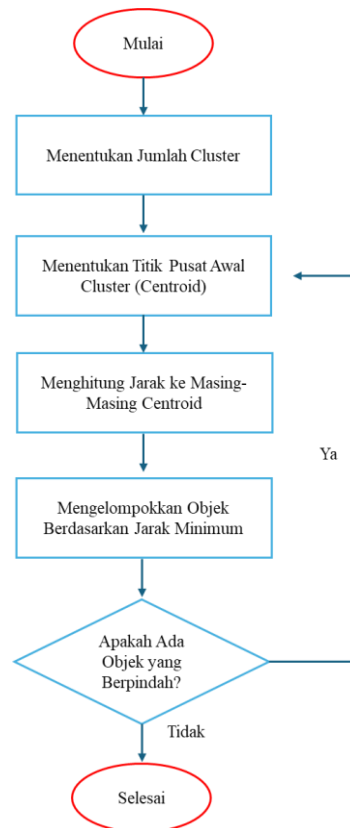
$$X_{\text{new}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \tag{1}$$

X_{new} mewakili nilai baru setelah melalui proses normalisasi, sedangkan X merupakan nilai awal dari atribut, X_{min} adalah nilai minimum dari atribut, dan X_{max} mewakili nilai maksimum dari atribut



4. Desain Model Klasterisasi Data

Analisis dalam penelitian ini menggunakan model klasterisasi. Model ini dibangun menggunakan algoritma K-Means Clustering. Algoritma ini dipilih karena dinilai sesuai dengan tujuan penelitian untuk mengelompokkan perguruan tinggi berdasarkan kesamaan data karakteristik PTS. Algoritma ini juga sesuai dengan dataset penelitian yang tidak terlalu besar, yaitu 104 data point. Flowchart algoritma K-Means Clustering yang digunakan adalah seperti yang ditunjukkan pada Gambar 5.



Gambar 5. Model Klasterisasi Data dengan algoritma K-Means Clustering

Setelah melewati tahap pre-processing data, dilanjutkan dengan proses klasterisasi data menggunakan algoritma K-Means Clustering. Tahapan umum dari proses tersebut melibatkan langkah-langkah sebagai berikut:

1. Pada iterasi pertama, tentukan jumlah klaster sesuai dengan kebutuhan.
2. Tentukan titik pusat awal klaster dengan acak.
3. Menghitung jarak dari setiap objek ke masing-masing centroid menggunakan persamaan rumus Euclidean Distance.
4. Mengelompokkan objek ke klaster berdasarkan jarak minimum. Jika ada perubahan dalam pengelompokkan, lakukan iterasi berikutnya dengan menentukan titik pusat klaster yang baru. Ulangi proses iterasi hingga posisi objek tidak berubah lagi, menunjukkan konvergensi atau stabilitas dalam pembentukan klaster [24].

3. HASIL DAN PEMBAHASAN

3.1 Dataset Penelitian

Penelitian ini menggunakan seluruh data indikator IKU-PT dan data persentase laporan PDDIKTI yang diambil pada laman <https://iku-pt.kemdikbud.go.id/> untuk data IKU-PT dan pada laman <https://pddikti.kemdikbud.go.id/> untuk data pelaporan PDDIKTI yaitu delapan indikator IKU-PT dan satu indikator pelaporan data PDDIKTI pada 104 PTS di LLDIKTI Wilayah V. Data 104 PTS dengan 9 Indikator melalui tahap pre-processing dengan melakukan penggabungan data (integrasi data) 8 indikator IKU-PT dan 1 indikator pelaporan PDDIKTI dengan bantuan Microsoft Excell yang ditunjukkan pada cuplikan data Tabel 2.

Tabel 2. Data Awal Penelitian

Nama PT	IKU 1	IKU 2	IKU 3	IKU 4	IKU 5	IKU 6	IKU 7	IKU 8	PDDIKTI
PT1	0.00	0.00	23.94	19.72	178.87	0.00	0.00	0.00	100.00
PT2	0.00	0.00	30.77	13.46	217.31	0.00	0.00	0.00	100.00
PT3	0.71	0.76	20.24	25.00	204.76	0.00	0.00	0.00	100.00



Nama PT	IKU 1	IKU 2	IKU 3	IKU 4	IKU 5	IKU 6	IKU 7	IKU 8	PDDIKTI
PT4	0.00	3.96	13.36	35.36	408.37	10.81	0.00	8.57	99.95
PT5	0.17	0.53	16.81	21.01	290.76	0.00	0.00	0.00	100.00

Hasil integrasi data menunjukkan meski satuan yang digunakan sama tetapi nilai maksimum dan minimum setiap indikator tidak sama, sebagai contoh, pada indikator jumlah keluaran dosen yang berhasil mendapatkan rekognisi internasional atau diterapkan oleh masyarakat/industri/pemerintah per jumlah dosen (IKU 5) dapat mencapai lebih dari 100%. Sementara itu, pada indikator persentase lulusan S1 dan D4/D3/D2/D1 yang berhasil memperoleh pekerjaan, studi lanjut, atau menjadi wiraswasta, dapat mencapai kurang dari satu persen (IKU 1). Selanjutnya pada proses missing value atau melihat rentang keseluruhan data, terdapat sebagian besar bernilai nol dan yang tidak bernilai nol memiliki rentang data yang terlalu jauh menyebabkan standar deviasi yang tidak normal. Sehingga perlu menggunakan metode Imputation atau menghitung nilai pengganti. Normalisasi data untuk mengubah rentang nilai menjadi seragam agar proses klusterisasi lebih mudah.

Tabel 3 menunjukkan dataset penelitian ini setelah melewati tahap data pre-processing yang meliputi integrasi data dan transformasi data. Data ini mencakup 104 PTS yang berlokasi di LLDIKTI Wilayah V Daerah Istimewa Yogyakarta.

Tabel 3. Dataset Penelitian

kode_pt	I1	I2	I3	I4	I5	I6	I7	I8	P1
1	0.20	0.02	0.18	0.27	0.44	0.56	0.01	0.10	1.00
2	0.22	0.02	0.23	0.18	0.53	0.45	0.09	0.10	1.00
.....
102	0.65	0.96	0.75	0.14	0.52	0.99	0.54	0.56	1.00
103	1.00	0.91	0.16	0.46	0.32	0.69	0.01	0.14	1.00
104	1.69	0.01	0.23	0.95	0.92	0.46	0.05	0.41	1.00

Pada kolom pertama Tabel 2, ditampilkan kode_pt yang merepresentasikan data point PTS mulai dari PTS 1 hingga PTS 104. Pada kolom 2 ditampilkan data indikator IKU 1(I1) dari PTS yaitu kesiapan kerja lulusan. Sebagai contoh PTS 1 memiliki IKU 1: 0.20 artinya hanya 20 persen dari lulusan yang sudah diterima bekerja. Berturut-turut kolom 3 sampai kolom 10 menunjukkan nilai 8 indikator lainnya: persentase partisipasi mahasiswa berkegiatan atau pencapaian prestasi di luar ruang lingkup program studi (I2), persentase dosen berkegiatan di luar kampus utama (I3), persentase kualifikasi dosen atau pengajar (I4), persentase penerapan karya dosen (I5), persentase kemitraan prodi (I6), persentase pembelajaran dalam kelas (I7), persentase akreditasi internasional (I8), dan persentase pelaporan PDDIKTI (P1).

3.2 Menentukan Jumlah Klaster

Salah satu tantangan utama dalam menerapkan algoritma K-Means Clustering adalah menetapkan jumlah klaster. Penelitian ini menetapkan menjadi empat klaster seperti rincian pada Tabel 4.

Tabel 4. Penentuan jumlah klaster

Cluster	Keterangan
C1	Klaster 1
C2	Klaster 2
C3	Klaster 3
C4	Klaster 4

Tabel 3 menggambarkan bahwa C1 merupakan representasi dari klaster pertama, C2 merupakan representasi dari klaster kedua, C3 merupakan representasi dari klaster ketiga, dan C4 merupakan representasi dari klaster keempat

3.3 Iterasi Pertama

3.3.1 Menentukan Centroid Awal

Setelah menentukan jumlah klaster, langkah selanjutnya adalah menentukan centroid untuk keempat klaster yang sudah ditentukan. Untuk permulaan atau iterasi pertama dari algoritma K-Means Clustering, centroid ini ditentukan secara acak seperti yang ditunjukkan pada Tabel 5.

Tabel 5. Titik pusat awal centroid

Titik Pusat Awal Centroid										
kode_pt	Centroid	I1	I2	I3	I4	I5	I6	I7	I8	P1
4	Center 1	0.17	0.12	0.10	0.49	1.00	0.11	0.14	0.10	1.00
14	Center 2	0.10	0.18	0.08	0.15	0.19	0.13	0.14	0.07	0.87
24	Center 3	0.11	0.84	0.25	0.09	0.20	0.56	0.88	0.30	1.00



Titik Pusat Awal Centroid										
kode_pt	Centroid	I1	I2	I3	I4	I5	I6	I7	I8	P1
51	Center 4	0.19	0.01	0.29	0.26	0.87	0.30	0.11	0.03	1.00

Pada iterasi awal atau pertama, penelitian ini memakai data dari kode_pt 4, 14, 24, dan 51 sebagai centroid awal. Data kode_pt 4 mewakili data awal centroid 1 (Center 1), kode_pt 14 mewakili data awal centroid 2 (Center 2), kode_pt 24 mewakili data awal centroid 3 (Center 3), dan kode_pt 51 mewakili data awal centroid 4 (Center 4).

3.3.2 Menentukan Euclidean Distance

Jarak Euclidean atau Metric Euclidean merujuk pada garis lurus yang umum dan sederhana antara dua elemen atau jarak minimum antara dua objek [11], [25], [26]. Menghitung jarak Euclidean Distance dengan rumus yang ditunjukkan pada persamaan 2.

$$D(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \tag{2}$$

D adalah jarak data ke centroid, data record diwakili dengan p, data centroid diwakili dengan q, dan n adalah data ke-n. Berikut cuplikan hasil perhitungan jarak dengan persamaan Euclidean yang disajikan pada Tabel 6.

Tabel 6. Euclidean Distance

kode_pt	dc1	dc2	dc3	dc4
1	0.77	0.58	1.25	0.53
2	0.68	0.56	1.22	0.39
3	0.56	0.47	1.29	0.41
4	0.00	0.89	1.46	0.40
5	0.94	1.05	1.10	0.93
6	1.08	0.62	1.00	0.89
.....
102	1.68	1.61	1.01	1.54
103	1.46	1.35	1.32	1.41
104	1.67	2.01	2.27	1.71

Tabel 5 menunjukkan hasil dari Euclidean distance pada 4 klaster untuk tahapan iterasi pertama. Pada kolom 2 ditampilkan hasil perhitungan jarak pada klaster 1 (dc1), pada kolom 3 ditampilkan hasil perhitungan jarak pada klaster 2 (dc2), pada kolom 4 ditampilkan hasil perhitungan jarak pada klaster 3 (dc3), dan pada kolom 5 ditampilkan hasil perhitungan jarak pada klaster 4 (dc4). Sebagai contoh kode_pt 2, hasil nilai jarak untuk klaster 1 (dc1): 0.68, klaster 2 (dc2): 0.56, klaster 3 (dc3): 1.22, dan klaster 4 (dc4): 0.39.

3.3.3 Pengelompokan Klaster

Setelah melaksanakan perhitungan Euclidean distance, langkah berikutnya adalah menentukan kelompok klaster untuk setiap elemen data berdasarkan prinsip jarak terdekat atau jarak minimum. Hal tersebut ditunjukkan dalam cuplikan data yang ditunjukkan pada Tabel 7.

Tabel 7. Cuplikan Pengelompokan Data untuk Iterasi Pertama

kode_pt	dc1	dc2	dc3	dc4	klaster 1
1	0.77	0.58	1.25	0.53	4
2	0.68	0.56	1.22	0.39	4
3	0.56	0.47	1.29	0.41	4
4	0.00	0.89	1.46	0.40	1
5	0.94	1.05	1.10	0.93	4
6	1.08	0.62	1.00	0.89	2
.....
102	1.68	1.61	1.01	1.54	3
103	1.46	1.35	1.32	1.41	3
104	1.67	2.01	2.27	1.71	1

Pada kolom 6 hasil jarak minimum diantara 4 klaster. Sebagai contoh kode_pt 1 memiliki hasil perhitungan jarak klaster 1 (dc1): 0.77, untuk klaster 2 (dc2): 0.58, untuk klaster 3 (dc3): 1.25, untuk klaster 4 (dc4): 0.53. Hasil perhitungan jarak 4 klaster tersebut yang memiliki jarak paling dekat atau jarak minimum adalah klaster 4 (dc4), artinya untuk kode_pt 1 masuk pada kelompok atau klaster 4. Contoh lain adalah kode_pt 104 memiliki hasil perhitungan jarak klaster 1 (dc1): 1.67, untuk klaster 2 (dc2): 2.01, untuk klaster 3 (dc3): 2.27, untuk klaster 4 (dc4): 1.71. Hasil perhitungan jarak 4 klaster tersebut yang memiliki jarak paling dekat atau jarak minimum adalah



klaster 1 (dc4), artinya untuk kode_pt 104 masuk pada kelompok atau klaster 1. Hasil iterasi 1 ditunjukkan pada Gambar 8.

Tabel 8. Cuplikan Hasil Iterasi Pertama

HASIL AKHIR ITERASI 1										
kode_pt	I1	I2	I3	I4	I5	I6	I7	I8	P1	Cluster iterasi 1
1	0.20	0.02	0.18	0.27	0.44	0.56	0.01	0.10	1.00	4
4	0.17	0.12	0.10	0.49	1.00	0.11	0.14	0.10	1.00	1
6	0.19	0.00	0.06	0.18	0.17	0.60	0.49	0.01	0.92	2
12	0.08	0.15	0.57	0.42	0.56	0.23	0.74	0.04	1.00	4
.....
103	1.00	0.91	0.16	0.46	0.32	0.69	0.01	0.14	1.00	3
104	1.69	0.01	0.23	0.95	0.92	0.46	0.05	0.41	1.00	1

3.4 Iterasi Kedua

Kemudian, pada tahap iterasi kedua, proses perhitungan dilakukan dengan langkah serupa seperti iterasi pertama, dimana centroid awal ditentukan. Namun, perbedaan mendasar dalam perhitungan terletak pada pendekatan untuk mendapatkan nilai centroid awal. Pada iterasi kedua, tidak lagi menggunakan pendekatan acak, melainkan dilakukan dengan menghitung rata-rata dari setiap hasil klaster yang ditunjukkan pada persamaan 3.

$$X_j = \frac{1}{N} \sum_{i=1}^N x_{ij} \tag{3}$$

X_j adalah rata-rata koordinat ke-j dari klaster, x_{ij} mewakili koordinat ke-j dari titik data ke-i, N adalah jumlah titik data pada klaster, dan \sum adalah operasi dari penjumlahan. Dengan menggunakan bantuan Microsoft Excel dan menerapkan rumus Averageif, langkah selanjutnya dilakukan untuk mendapatkan nilai centroid baru. Nilai-nilai ini mewakili pusat dari masing-masing klaster, yang diberi label sebagai Center 1 (centroid klaster 1), Center 2 (centroid klaster 2), Center 3 (centroid klaster 3), dan Center 4 (centroid klaster 4), sebagaimana ditunjukkan pada Tabel 9.

Tabel 9. Centroid Awal iterasi kedua

Centroid Awal									
Centroid	I1	I2	I3	I4	I5	I6	I7	I8	P1
Center C1	0.52	0.25	0.17	0.58	0.82	0.34	0.27	0.34	1.00
Center C2	0.22	0.16	0.23	0.23	0.23	0.42	0.12	0.20	0.93
Center C3	0.22	0.53	0.35	0.22	0.30	0.61	0.64	0.36	0.91
Center C4	0.27	0.09	0.24	0.24	0.57	0.51	0.18	0.24	1.00

Tahap iterasi kedua setelah menentukan centroid baru, proses selanjutnya sama dengan iterasi pertama yaitu menghitung jarak centroid menggunakan rumus Euclidean distance. Hasil dari perhitungan jarak kemudian dikelompokkan berdasarkan jarak terdekat. Cuplikan data hasil iterasi kedua ditunjukkan pada Tabel 10.

Tabel 10. Cuplikan Hasil Iterasi Kedua

HASIL AKHIR ITERASI 2										
kode_pt	I1	I2	I3	I4	I5	I6	I7	I8	P1	Cluster iterasi 2
1	0.20	0.02	0.18	0.27	0.44	0.56	0.01	0.10	1.00	4
4	0.17	0.12	0.10	0.49	1.00	0.11	0.14	0.10	1.00	1
6	0.19	0.00	0.06	0.18	0.17	0.60	0.49	0.01	0.92	2
12	0.08	0.15	0.57	0.42	0.56	0.23	0.74	0.04	1.00	3
.....
103	0.03	0.19	0.14	0.18	0.56	0.25	0.19	0.08	1.00	1
104	1.69	0.01	0.23	0.95	0.92	0.46	0.05	0.41	1.00	1

Pada iterasi kedua terjadi pergeseran klaster, seperti pada cuplikan data di hasil iterasi pertama untuk kode_pt 12 berada pada klaster 4 tetapi di iterasi kedua bergeser menjadi di klaster 3, begitu pula kode_pt 103 pada iterasi pertama berada di klaster 3 tetapi di iterasi kedua bergeser menjadi di klaster 1. Hasil iterasi pertama dan iterasi kedua masih ada pergeseran sehingga harus dilakukan iterasi ketiga sampai iterasi ke-n dimana sudah tidak ada lagi pergeseran hasil klasterisasi.

3.5 Hasil Klasterisasi

Proses klasterisasi PTS berdasarkan data IKU-PT dan data laporan PDDIKTI berakhir pada iterasi kelima, ditunjukkan bahwa hasil dari iterasi keempat dan kelima sudah tidak ada pergeseran klaster yang ditunjukkan pada cuplikan data Tabel 11 dan Tabel 12.



Tabel 11. Cuplikan Hasil Iterasi Keempat

HASIL AKHIR ITERASI 4										
kode_pt	I1	I2	I3	I4	I5	I6	I7	I8	P1	Cluster iterasi 4
1	0.20	0.02	0.18	0.27	0.44	0.56	0.01	0.10	1.00	4
2	0.22	0.02	0.23	0.18	0.53	0.45	0.09	0.10	1.00	4
3	0.25	0.02	0.15	0.34	0.50	0.24	0.06	0.07	1.00	4
4	0.17	0.12	0.10	0.49	1.00	0.11	0.14	0.10	1.00	4
5	0.20	0.02	0.13	0.29	0.71	0.23	1.00	0.03	1.00	3
6	0.19	0.00	0.06	0.18	0.17	0.60	0.49	0.01	0.92	2
7	0.24	0.07	0.06	0.28	0.65	0.67	0.03	0.06	1.00	4
8	0.08	1.00	0.06	0.09	0.41	0.79	0.01	0.09	1.00	2
9	0.19	0.94	0.46	0.12	0.27	0.46	0.06	0.27	0.91	2
10	0.12	0.73	0.12	0.40	0.04	0.51	0.09	0.15	1.00	2
11	0.28	0.31	0.64	0.11	0.19	0.24	1.00	0.05	1.00	3
12	0.08	0.15	0.57	0.42	0.56	0.23	0.74	0.04	1.00	3
.....
102	0.65	0.96	0.75	0.14	0.52	0.99	0.54	0.56	1.00	3
103	1.00	0.91	0.16	0.46	0.32	0.69	0.01	0.14	1.00	1
104	1.69	0.01	0.23	0.95	0.92	0.46	0.05	0.41	1.00	1

Tabel 12. Cuplikan Hasil Iterasi Kelima

HASIL AKHIR ITERASI 5										
kode_pt	I1	I2	I3	I4	I5	I6	I7	I8	P1	Cluster iterasi 5
1	0.20	0.02	0.18	0.27	0.44	0.56	0.01	0.10	1.00	4
2	0.22	0.02	0.23	0.18	0.53	0.45	0.09	0.10	1.00	4
3	0.25	0.02	0.15	0.34	0.50	0.24	0.06	0.07	1.00	4
4	0.17	0.12	0.10	0.49	1.00	0.11	0.14	0.10	1.00	4
5	0.20	0.02	0.13	0.29	0.71	0.23	1.00	0.03	1.00	3
6	0.19	0.00	0.06	0.18	0.17	0.60	0.49	0.01	0.92	2
7	0.24	0.07	0.06	0.28	0.65	0.67	0.03	0.06	1.00	4
8	0.08	1.00	0.06	0.09	0.41	0.79	0.01	0.09	1.00	2
9	0.19	0.94	0.46	0.12	0.27	0.46	0.06	0.27	0.91	2
10	0.12	0.73	0.12	0.40	0.04	0.51	0.09	0.15	1.00	2
11	0.28	0.31	0.64	0.11	0.19	0.24	1.00	0.05	1.00	3
12	0.08	0.15	0.57	0.42	0.56	0.23	0.74	0.04	1.00	3
.....
102	0.65	0.96	0.75	0.14	0.52	0.99	0.54	0.56	1.00	3
103	1.00	0.91	0.16	0.46	0.32	0.69	0.01	0.14	1.00	1
104	1.69	0.01	0.23	0.95	0.92	0.46	0.05	0.41	1.00	1

Hasil iterasi keempat dan iterasi kelima tidak ada perubahan klaster, sebagai contoh kode_pt 1 hingga kode_pt 104 pada iterasi keempat tidak mengalami pergeseran di iterasi kelima. Rangkuman hasil iterasi yang melibatkan iterasi pertama hingga iterasi kelima ditunjukkan pada Tabel 13.

Tabel 13. Rekap Hasil Klasterisasi

Iterasi	C1	C2	C3	C4
1	5	50	15	34
2	6	47	17	34
3	4	47	20	33
4	4	46	21	33
5	4	46	21	33

Pada Tabel 11, Ditunjukkan bahwa pada iterasi pertama, terdapat nilai-nilai berikut: C1=5 PTS, C2=50 PTS, C3=15 PTS, dan C4=34 PTS. Pada iterasi kedua, nilai-nilai berubah menjadi C1=6 PTS, C2=47 PTS, C3=17 PTS, dan C4=34 PTS. Iterasi ketiga menunjukkan nilai C1=4 PTS, C2=47 PTS, C3=20 PTS, dan C4=33 PTS. Sementara pada iterasi keempat, hasilnya adalah C1=4 PTS, C2=46 PTS, C3=21 PTS, dan C4=33 PTS. Iterasi kelima menunjukkan hasil yang sama dengan iterasi keempat, menandakan bahwa posisi klaster tidak mengalami perubahan lebih lanjut. Hal ini dapat disebut sebagai konvergensi, di mana pengelompokkan telah stabil dan tidak mengalami pergeseran antara iterasi keempat dan kelima.

Hasil akhir klaster menunjukkan bahwa C1 sebagai klaster 1 mewakili kategori responsif yaitu PTS yang selesai mendapatkan sosialisasi langsung menindaklanjuti dengan pendataan. C2 sebagai klaster 2 mewakili



kategori mengabaikan yaitu PTS yang sudah mendapatkan informasi terbaru melalui sosialisasi, sudah mendapatkan arahan tetapi tetap tidak melakukan pendataan. C3 sebagai klaster 3 mewakili kategori proaktif yaitu PTS mencari informasi terbaru sebelum ada sosialisasi untuk memastikan persiapan di perguruan tingginya. C4 sebagai klaster 4 mewakili kategori mengikuti aturan yaitu PTS yang menunggu arahan berikutnya setelah mendapat sosialisasi dan menindaklanjuti sesuai deadline pelaporan.

4. KESIMPULAN

Hasil penelitian menunjukkan bahwa dari 104 data dengan sembilan indikator, termasuk delapan indikator IKU-PT dan satu indikator pelaporan PDDIKTI, dapat dikelompokkan menjadi empat klaster. Klaster 1 (C1) terdiri dari 4 PTS, klaster 2 (C2) terdiri dari 46 PTS, klaster 3 (C3) terdiri dari 21 PTS, dan klaster 4 (C4) terdiri dari 33 PTS. Metode K-Means Clustering dapat diterapkan untuk klasterisasi PTS di LLDIKTI Wilayah V DIY. Hasil klasterisasi sudah memenuhi tujuan penelitian yaitu menawarkan solusi untuk mengatasi kesulitan LLDIKTI Wilayah V dalam memetakan PTS di wilayahnya berdasarkan mutu untuk dapat dijadikan salah satu dasar pengambilan keputusan dalam memfasilitasi PTS mencapai mutu unggul. Dalam mencapai tujuan penelitian secara optimal, penelitian lanjutan diperlukan dengan menambah indikator dari sisi pelaporan data PDDIKTI dan akreditasi sebagai penjaminan mutu eksternal perguruan tinggi. Selain itu, LLDIKTI Wilayah V harus mendorong PTS untuk melakukan update data pada aplikasi kementerian yang digunakan sebagai dasar penilaian oleh IKU-PT, seperti aplikasi PDDIKTI, SISTER, Laporankerma, dan Tracer study. Pengisian data tersebut juga sebagai penunjang untuk penilaian pada sistem penjaminan mutu eksternal yaitu akreditasi. Melalui penelitian ini, diharapkan LLDIKTI Wilayah V dapat melakukan tugas pembinaan dan layanan dengan lebih efektif dan memberikan pendampingan terutama kepada PTS yang masuk dalam klaster dengan capaian indikator mutu yang belum maksimal.

UCAPAN TERIMAKASIH

Terima kasih disampaikan kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) melalui beasiswa unggulan dengan nomor: 0379/J5.2.2/BP/PKS/X/2023 dan LLDIKTI Wilayah V, serta Magister Informatika, Universitas Ahmad Dahlan Yogyakarta.

REFERENCES

- [1] K. Kardoyo, L. K. Pitaloka, R. Rozman, and B. B. Hapsoro, "Analyzing Universities Service Quality to Student Satisfaction; Academic and Non-Academic Analyses," *Int. J. High. Educ.*, vol. 9, no. 1, p. 126, Dec. 2020, doi: 10.5430/ijhe.v9n1p126.
- [2] Yoga Budi Bhakti, Achmad Ridwan, and Riyadi, "Urgensi Sistem Penjaminan Mutu Internal Dan Eksternal Dalam Meningkatkan Mutu Perguruan Tinggi," *J. Penjaminan Mutu*, vol. 8, no. 02, pp. 251–260, Aug. 2022, doi: 10.25078/jpm.v8i02.1394.
- [3] D. J. P. T. Kemendikbudristek, Directorate General of Higher Education Ministry of Education and Culture Higher Education Statistics. 2022.
- [4] K. M. P. dan K. RI, Keputusan Menteri Pendidikan dan Kebudayaan RI No 3/ M/ 2021 tentang Indikator Kinerja Utama Perguruan Tinggi Negeri dan Lembaga Layanan Pendidikan Tinggi di Kementerian Pendidikan dan Kebudayaan. 2021, pp. 1–33.
- [5] K. Mendikbudristek, Keputusan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 210/M/2023 Tentang Indikator Kinerja Utama Perguruan Tinggi dan LL-DIKTI di Kemendikbudristek. 2023.
- [6] F. Rahmayani, "Penerapan Algoritma C4.5 Dengan Feature Forward Selection Untuk Analisis Capaian Indikator Kinerja Utama Berdasarkan Tracer Study (Studi Kasus: Fasilkom Unsika)," vol. 7, no. 4, pp. 2732–2738, 2023.
- [7] Kementrian Hukum dan HAM, "Undang-Undang Republik Indonesia No. 12 Tahun 2012 tentang Pendidikan Tinggi," Undang Undang, p. 18, 2012.
- [8] R. P. Wibowo et al., "Problem identification and intervention in the higher education data synchronization system in Indonesia," *Procedia Comput. Sci.*, vol. 197, no. 2021, pp. 484–494, 2021, doi: 10.1016/j.procs.2021.12.165.
- [9] N. Ngatmari, M. B. Musthafa, C. Rahmad, R. A. Asmara, and F. Rahutomo, "Pemanfaatan Data PDDIKTI sebagai Pendukung Keputusan Manajemen Perguruan Tinggi," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 555–564, 2020, doi: 10.25126/jtiik.2020722585.
- [10] S. Usman and U. Pancasakti, "Implementasi Sistem Informasi Akademik Dengan Feeder PDDIKTI Berbasis Android," *J. Syst. Comput. Eng. Vol. 1 No. 1 Juli 2020 ISSN*, vol. 1, no. 1, pp. 28–37, 2020.
- [11] T. M. Ghazal et al., "Performances of k-means clustering algorithm with different distance metrics," *Intell. Autom. Soft Comput.*, vol. 30, no. 2, pp. 735–742, 2021, doi: 10.32604/iasc.2021.019067.
- [12] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [13] A. A. Aldino, D. Darwis, A. T. Prastowo, and C. Sujana, "Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency," *J. Phys. Conf. Ser.*, vol. 1751, no. 1, 2021, doi: 10.1088/1742-6596/1751/1/012038.
- [14] A. F. Mohamed Nafuri, N. S. Sani, N. F. A. Zainudin, A. H. A. Rahman, and M. Aliff, "Clustering Analysis for Classifying Student Academic Performance in Higher Education," *Appl. Sci.*, vol. 12, no. 19, 2022, doi:



- 10.3390/app12199467.
- [15] R. Watrianthos, R. Handayani, W. Simatupang, D. Irfan, and M. Muskhir, "Penerapan Metode PROMETHEE-GAIA Dalam Pemingkatan Perguruan Tinggi di Indonesia," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 138, 2022, doi: 10.30865/mib.v6i1.3419.
- [16] E. Ermawati, I. Sriliana, and R. Sriningsih, "Clustering of State Universities in Indonesia Based on Productivity of Scientific Publications Using K-Means and K-Medoids," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 3, pp. 1617–1630, 2023, doi: 10.30598/barekengvol17iss3pp1617-1630.
- [17] J. H. E. Guzman, R. A. Zuluaga-Ortiz, L. E. G. Donado, E. J. Delahoz-Dominguez, A. Marquez-Castillo, and M. Suarez-Sánchez, "Cluster analysis in Higher Education Institutions' knowledge identification and production processes," *Procedia Comput. Sci.*, vol. 203, no. 2021, pp. 570–574, 2022, doi: 10.1016/j.procs.2022.07.081.
- [18] Y. Sasmita, M. Muhsni, and M. Walid, "Klasterisasi Perguruan Tinggi Swasta di Madura Berdasarkan Kinerja Sumber Daya Manusia dan Mahasiswa Menggunakan Metode K-Means Clustering," *J. MEDIA Inform. BUDIDARMA*, vol. 6, no. 4, p. 2157, Oct. 2022, doi: 10.30865/mib.v6i4.4431.
- [19] H. Han, "Fuzzy clustering algorithm for university students' psychological fitness and performance detection," *Heliyon*, vol. 9, no. 8, p. e18550, 2023, doi: 10.1016/j.heliyon.2023.e18550.
- [20] B. K. Francis and S. S. Babu, "Predicting Academic Performance of Students Using a Hybrid Data Mining Approach," *J. Med. Syst.*, vol. 43, no. 6, 2019, doi: 10.1007/s10916-019-1295-4.
- [21] F. R. Fiantika et al., *Metodologi Penelitian Kualitatif*. In *Metodologi Penelitian Kualitatif*, no. Maret. 2022. [Online]. Available: <https://scholar.google.com/citations?user=O-B3eJYAAAAAJ&hl=en>
- [22] M. M. Ali, T. Hariyati, M. Y. Pratiwi, and S. Affah, "Metodologi Penelitian Kuantitatif dan Penerapannya dalam Penelitian," *Educ. J.*, vol. 2, no. 2, pp. 1–6, 2022.
- [23] M. Ridwan, S. AM, B. Ulum, and F. Muhammad, "Pentingnya Penerapan Literature Review pada Penelitian Ilmiah," *J. Masohi*, vol. 2, no. 1, p. 42, 2021, doi: 10.36339/jmas.v2i1.427.
- [24] C. Shi, B. Wei, S. Wei, W. Wang, H. Liu, and J. Liu, "A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm," *Eurasip J. Wirel. Commun. Netw.*, vol. 2021, no. 1, 2021, doi: 10.1186/s13638-021-01910-w.
- [25] S. Saqib, A. Ditta, M. A. Khan, S. A. R. Kazmi, and H. Alquhayz, "Intelligent dynamic gesture recognition using CNN empowered by edit distance," *Comput. Mater. Contin.*, vol. 66, no. 2, pp. 2061–2076, 2020, doi: 10.32604/cmc.2020.013905.
- [26] M. Faisal, E. M. Zamzami, and Sutarman, "Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance," *J. Phys. Conf. Ser.*, vol. 1566, no. 1, 2020, doi: 10.1088/1742-6596/1566/1/012112.