

# Penerapan Algoritma K-Means Pada Pemetaan Kemampuan Penggunaan Teknologi Informasi Remaja dan Dewasa di Indonesia

Ronal Watrianthos<sup>1\*</sup>, Reti Handayani<sup>2</sup>, Ade Fitrah Putra Akhir<sup>3</sup>, Ambiyar<sup>4</sup>, Unung Verawardina<sup>4</sup>

<sup>1</sup>Fakultas Teknik, Teknik Informatika, Universitas Al Washliyah, Rantauprapat, Indonesia

<sup>2</sup>Fakultas Ekonomi, Manajemen Informatika, Universitas Mahaputra Muhammad Yamin, Solok, Indonesia

<sup>3</sup>Fakultas Teknik, Informatika, Universitas Prof. Dr. Hazairin SH, Bengkulu, Indonesia

<sup>4</sup>Fakultas Teknik, Pendidikan Teknologi Kejuruan, Universitas Negeri Padang, Padang, Indonesia

Email: <sup>1,\*</sup>ronal.watrianthos@gmail.com, <sup>2</sup>jeranikadun@gmail.com, <sup>3</sup>adefitrah.af@gmail.com, <sup>4</sup>ambiyar@ft.unp.ac.id, <sup>5</sup>unungverawardina@gmail.com,

Email Penulis Korespondensi: ronal.watrianthos@gmail.com

Submitted: 13/09/2022; Accepted: 21/11/2022; Published: 30/11/2022

**Abstrak**—Akses ke berbagai bentuk teknologi dianggap menjadi penting saat ini di Indonesia. Hal ini karena jika pemerintah tidak bergerak cepat untuk membuat teknologi lebih ramah pengguna, kesenjangan keterampilan akan teknologi informasi akan terus melebar. Menurut PISA, ada kemungkinan bahwa lembaga pendidikan dapat mengambil manfaat lebih banyak dari penggunaan teknologi informasi. Tujuan dari penelitian ini adalah untuk mengetahui ada tidaknya perbedaan tingkat keahlian teknologi informasi yang signifikan antar berbagai provinsi di Indonesia. Perhitungan klusterisasi menggunakan data resmi Badan Pusat Statistik tahun 2019-2021 dengan data jumlah remaja dan dewasa usia 15-59 tahun yang memiliki keterampilan di bidang Teknologi Informasi dan Komputer (TIK) di tingkat provinsi. Sebagai salah satu metode pada *unsupervised learning*, algoritma K-Means digunakan agar data set dapat diorganisasikan ke dalam kelompok dengan item lain yang sebanding dengannya. Hasil penelitian menunjukkan nilai DBI paling rendah yaitu -0,357 pada kluster ketujuh sehingga menjadi terbaik dari tiga kluster lainnya. Rata-rata persentase penduduk di kluster nol yang memiliki kemampuan di bidang teknologi informasi hanya 29,32 persen, sehingga menjadi kluster dengan persentase terendah secara keseluruhan. Kluster nol ini hanya dapat ditemukan di Provinsi Papua.

**Kata Kunci:** K-Means; Pemetaan; Teknologi Informasi; Data Mining; Rapid Miner

**Abstract**—In modern-day Indonesia, having access to various forms of technology is regarded as crucial. This is because the IT skills gap will only grow worse if the government does not act swiftly to make technology more accessible to the general public. PISA suggests that schools may gain much more from incorporating IT into their curriculum if they did so. The goal of this research is to establish whether or not there are notable variations in the level of knowledge in information technology across the different provinces in Indonesia. Official data from the Central Statistics Agency for 2019-2021, including information on the number of adolescents and adults aged 15-59 with abilities in the field of ICT at the provincial level, is used in the clustering calculation. The K-means algorithm is one unsupervised learning technique for clustering data into collections with other instances that share similar properties. The results showed that the seventh cluster had the lowest DBI value of the three examined, coming in at -0.357. Cluster 0 has the lowest average percentage of the population with IT skills, at 29.32%. Only in Papua Province will you find this particular zero cluster.

**Keywords:** K-Means; Mapping; Information Technology; Data Mining; Rapid Miner

## 1. PENDAHULUAN

Menghadapi era digital yang akan datang, yang ditandai dengan globalisasi dan teknologi baru, berbagai kelompok telah menekankan pentingnya mengubah pendidikan dengan apa yang disebut keterampilan abad ke-21. Dalam pendapat lain, kebutuhan akan keterampilan baru untuk menjadi bagian dari budaya digital ditekankan. Keterampilan digital adalah kunci dan esensial untuk hidup, bekerja, dan menjadi bagian dari *knowledge society*[1]. Agar siswa dapat bersaing di era pengetahuan, mereka perlu memiliki keterampilan dan kemampuan abad ke-21. Jadi, sekolah harus merubah paradigmanya untuk memenuhi kebutuhan tersebut[2].

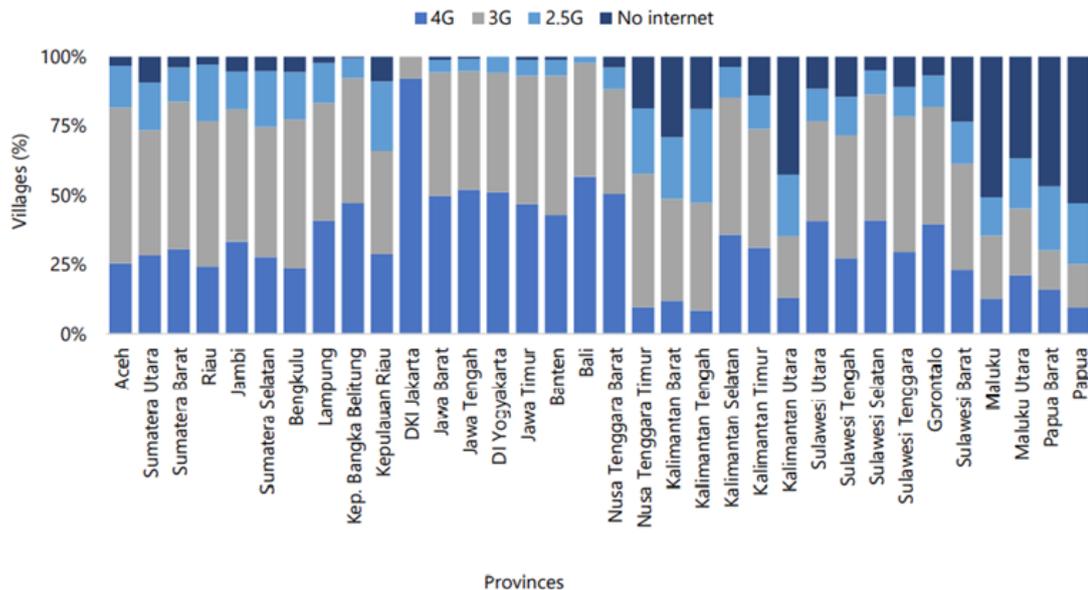
Peserta didik dan guru di lembaga pendidikan sekarang sudah memiliki kesempatan untuk mengadopsi teknik belajar-mengajar abad ke-21 untuk mendukung pengembangan kemampuan abad ke-21. Hal ini dimungkinkan oleh tersedianya infrastruktur TIK dan konektivitas internet di lembaga pendidikan. Konektivitas internet khususnya menyediakan platform untuk pergeseran dari penggunaan pedagogi yang berpusat pada guru (pembelajaran berbasis konten) dan menuju pedagogi yang berpusat pada peserta didik (inkuiri dan pembelajaran berbasis proyek), yang lebih interaktif dan berorientasi pada aktivitas. Pergeseran ini menjadi semakin lazim di lanskap pendidikan saat ini[3].

Peran lembaga pendidikan dalam proses pembelajaran adalah menyediakan infrastruktur pembelajaran, sehingga sekolah merupakan komponen eksternal yang penting dalam pembelajaran. Prasarana atau sarana pendidikan yang sangat pesat dimungkinkan oleh kemajuan Teknologi Informasi dan Komunikasi (TIK). Komputer, internet, website, dan proyektor adalah beberapa bentuk pembelajaran dan komunikasi berbasis teknologi informasi (TIK) yang tersedia di fasilitas tersebut untuk digunakan dalam proses belajar mengajar[4].

Akses terhadap teknologi merupakan isu yang sangat penting di Indonesia. Jika pemerintah tidak bertindak cepat untuk meningkatkan akses ke teknologi, kesenjangan keterampilan teknologi informasi akan semakin besar. Hasil PISA menunjukkan bahwa sekolah akan lebih maju dengan penggunaan komputer daripada tidak menggunakan komputer sama sekali. Namun, hambatan sistemis mempersulit sekolah dan perguruan tinggi untuk

menggunakan teknologi. Misalnya, sekitar 23,95% sekolah masih belum terhubung ke Internet, beberapa sekolah bahkan belum memiliki listrik. Peningkatan infrastruktur TIK diperlukan untuk memastikan bahwa *teaching factory* memiliki kualitas yang cukup baik untuk memenuhi kebutuhan industri, e-Learning dilakukan dengan baik, dan tidak ada masalah koneksi selama ujian nasional online. Hal ini membuat kebutuhan bagi banyak guru untuk meningkatkan keterampilan dasar TIK mereka[5].

Gambar 1 menunjukkan bahwa lebih dari 50% daerah perbatasan, terluar, dan miskin kekurangan akan akses internet. Desa Maluku, Maluku Utara, Papua, dan Papua Barat melaporkan sinyal yang buruk atau di bawah 3G. Hanya 34 Kota madya DKI Jakarta yang memiliki internet 4G. Bali dan NTB memiliki lebih banyak desa 4G daripada provinsi lain di Indonesia karena pariwisata. Untuk dua wilayah di mana sektor pariwisata mendukung hampir semua penduduk, kebutuhan koneksi internet berkualitas tinggi untuk menjangkau semua komunitas yang jauh menyebabkan akses yang lebih maju[6].



Gambar 1. Desa Memiliki Akses Internet

Penelitian ini bertujuan mengelompokkan sebaran kemampuan teknologi informasi pada remaja dan dewasa menurut provinsi di Indonesia. Data resmi Badan Pusat Statistik 2019-2021 terkait proporsi remaja dan dewasa usia 15-59 tahun dengan keterampilan Teknologi Informasi dan Komputer (TIK) menurut provinsi (persen) digunakan untuk menghitung pengelompokan[7]. Pendekatan *unsupervised learning* dengan algoritma K-Means digunakan agar objek di setiap provinsi dapat dikategorikan ke dalam kategori terkait. Teknik pembelajaran tanpa pengawasan menggunakan K-Means untuk pengelompokan adalah dasar dari pendekatan ini. Pendekatan-pendekatan ini telah digunakan secara luas. Teknik K-Means merupakan salah satu contoh dari *partitioning* klustering, yaitu suatu bentuk klustering yang memecah data menjadi bagian-bagian komponennya. K-Means adalah pendekatan yang terkenal karena penerapannya sederhana dan cepatnya mengelompokkan sejumlah besar data dan outlier[8][9].

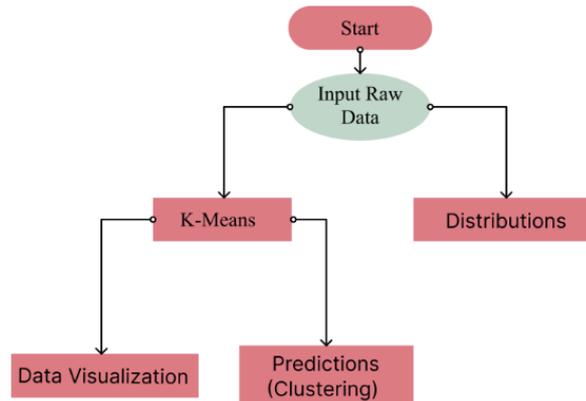
Riset menggunakan K-Means pernah dilakukan dalam rangka mengelompokkan provinsi-provinsi di Indonesia berdasarkan risiko pandemi COVID-19. Penelitian ini berdasarkan data dari virus Corona 2019 (COVID-19). Penelitian ini mengarah pada identifikasi tiga kluster kasus COVID-19 yang berbeda di provinsi-provinsi di Indonesia, yang masing-masing didasarkan pada proporsi kasus terkonfirmasi, kasus sembuh, dan kematian dalam satu provinsi tertentu[10]. Dalam studi lain dalam mengevaluasi data besar E-Commerce dengan teknik pengelompokan K-Means berbasis partisi menghasilkan temuan lokasi geografis dan nomor identifikasi pelanggan yang unik yang diperlakukan sebagai pembatasan pengelompokan dalam penelitian ini[11].

Proses klustering dilakukan dengan bantuan RapidMiner, pemetaan dengan Quantum GIS berdasarkan peta lengkap seluruh provinsi di Indonesia divisualisasikan menggunakan data yang dikumpulkan melalui proses klustering[12]. Hal ini dilakukan agar pemahaman visual tentang sebaran kluster kemampuan teknologi informasi di Indonesia dapat diselesaikan.

## 2. METODOLOGI PENELITIAN

Algoritma K-Means saat ini banyak digunakan walaupun ada banyak metode berbeda untuk memperluas K-Means, dan banyak literatur telah mengusulkan banyak pendekatan berbeda tersebut. Meskipun algoritme K-Means dan ekstensinya menggunakan pembelajaran tanpa pengawasan untuk melakukan pengelompokan dalam pengenalan pola dan pembelajaran mesin, algoritma ini selalu dipengaruhi oleh inisialisasi dengan jumlah kluster tertentu.

Terlepas dari apakah pembelajaran diawasi atau tidak, algoritma K-Means sebenarnya bukanlah sistem yang dapat mengelompokkan item tanpa bantuan manusia[13][14].



Gambar 2. Tahapan Penelitian

### 2.1 Algoritma K-Means

Algoritma K-Means adalah tentang metode pengelompokan iteratif dasar. Dengan menghitung jarak *mean* yang akan memasok centroid pertama dengan memanfaatkan jarak itu sendiri sebagai metrik dan kelas K yang terdapat dalam kumpulan data sebagai input. Ini akan memungkinkan untuk menentukan centroid pertama. Sebagai konsekuensi dari ini, setiap kelas akan dibedakan oleh centroidnya sendiri yang berbeda. Sebagai indeks kesamaan untuk kumpulan data X yang diberikan yang sudah memiliki n titik data multidimensi dan kategori K yang harus dipisahkan, jarak *Euclidean* telah diputuskan sebagai pilihan terbaik. Tujuan pengelompokan disesuaikan untuk mengurangi jumlah total kombinasi potensial sejauh mungkin, sekaligus meminimalkan jumlah kuadrat untuk masing-masing jenis yang berbeda[15]–[17].

$$D = \sum_{k=1}^k \sum_{i=1}^n ||(x_i - u_k)||^2 \tag{1}$$

Dimana k adalah jumlah pusat kluster,  $u_k$  adalah jumlah pusat ke-k, dan  $x_i$  adalah jumlah titik ke-i dalam kumpulan data. K merupakan bilangan positif dengan centroid  $u_k$  dapat ditemukan sebagai berikut:

$$\frac{\partial}{\partial u_k} = \frac{\partial}{\partial u_k} \sum_{k=1}^k \sum_{i=1}^n (x_i - u_k)^2 \tag{2}$$

$$= \sum_{k=1}^k \sum_{i=1}^n \frac{\partial}{\partial u_k} (x_i - u_k)^2 \tag{3}$$

$$= \sum_{i=1}^n 2(x_i - u_k) \tag{4}$$

Implementasi algoritma didasarkan pada prinsip pemilihan titik K secara acak dari kumpulan titik sampel untuk dijadikan sebagai pusat kluster pertama dengan memisahkan setiap titik sampel ke dalam kluster yang titik pusatnya paling dekat dengannya. Titik pusat dari setiap kluster ditentukan dengan mengambil rata-rata dari semua titik sampel yang ada di dalam kluster tersebut. Ulangi prosedur dari bagian sebelumnya sampai posisi pusat kluster tetap tidak berubah atau jumlah pengulangan maksimum telah tercapai. Temuan algoritma dapat berubah tergantung pada titik pusat yang dipilih. Lokasi titik pusat ditentukan oleh nilai K, yang merupakan fokus utama dari algoritma. Ini berdampak langsung pada hasil klustering, seperti menentukan apakah kluster tersebut lokal atau global[15], [18].

### 2.2 Sistem Informasi Geografis

Setiap sistem yang merekam, menyimpan, memproses, dan menyajikan data yang terhubung ke lokasi tertentu disebut sebagai Sistem Informasi Geografis, atau disingkat GIS. Istilah "GIS" mengacu pada berbagai kegiatan, termasuk pengelolaan, manipulasi, dan penggunaan data georeferensi digital[19]. GIS telah berkembang secara signifikan selama dekade terakhir, dengan jangkauan aplikasi yang lebih besar baik di industri dan pemerintahan, penggunaan konsumen yang luas, dan peran yang meningkat di sekolah, bisnis dan sistem informasi[20].

Desain GIS adalah genre kualitatif yang menggunakan metodologi deskriptif untuk menggambarkan sistematika penelitian. Dalam konteks penelitian ini, aplikasi Quantum GIS berfungsi sebagai instrumen untuk pengolahan dan pengelompokan distribusi kemampuan teknologi informasi[21].

### 2.3 Dataset

Penelitian ini menggunakan data dari Badan Pusat Statistik Republik Indonesia (BPS) periode tahun 2019-2021, dengan melihat persentase penduduk di setiap provinsi berusia 15-59 tahun yang mahir dalam kemampuan

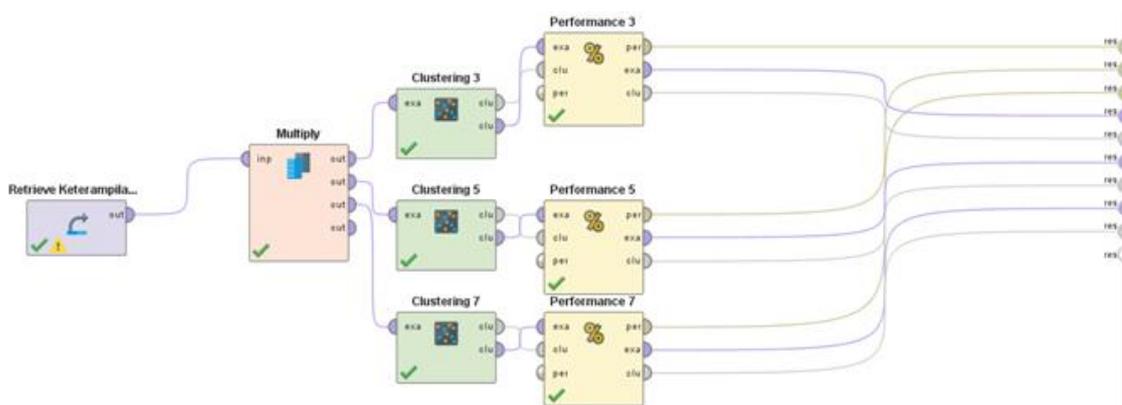
komputer dan teknologi[7]. Rapid Miner digunakan untuk implementasi algoritma pengelompokan untuk menampilkan data. Data dalam bentuk yang belum diproses dan dibersihkan diberikan dalam Tabel 1:

**Tabel 1.** Proporsi Remaja dan Dewasa Usia 15-59 Tahun dengan Ketrampilan Teknologi Informasi dan Komputer (TIK) Menurut Provinsi (Persen), 2019-2021[7]

Provinsi	2019	2020	2021
Aceh	46,77	54,25	60,21
Sumatera Utara	51,78	58,60	67,41
Sumatera Barat	52,85	58,67	68,00
Riau	55,37	62,67	70,69
Jambi	50,83	56,87	64,47
Sumatera Selatan	46,50	54,52	62,59
Bengkulu	48,70	53,42	62,10
Lampung	48,37	55,57	65,76
Bangka Belitung	54,93	60,37	66,33
Riau Kepulauan	77,18	81,73	89,06
DKI Jakarta	85,17	88,08	91,79
Jawa Barat	65,37	71,09	76,08
Jawa Tengah	58,75	65,78	71,15
Yogyakarta	75,04	81,36	84,72
Jawa Timur	57,23	63,91	68,07
Banten	66,96	69,35	75,69
Bali	65,48	72,56	77,09
Nusa Tenggara Barat	47,85	52,72	58,69
Nusa Tenggara Timur	36,33	42,89	53,16
Kalimantan Barat	47,04	54,10	62,04
Kalimantan Tengah	54,54	59,66	66,43
Kalimantan Selatan	57,82	62,88	70,39
Kalimantan Timur	69,44	75,33	81,17
Kalimantan Utara	65,36	71,99	76,94
Sulawesi Utara	57,48	63,03	69,77
Sulawesi Tengah	44,13	51,68	58,19
Sulawesi Selatan	54,85	60,50	67,29
Sulawesi Tenggara	53,36	60,35	65,75
Gorontalo	50,62	55,68	61,94
Sulawesi Barat	40,95	47,66	55,72
Maluku	44,02	49,96	59,26
Maluku Utara	38,11	45,22	51,53
Papua Barat	52,37	59,45	62,31
Papua	26,45	30,93	30,58

### 3. HASIL DAN PEMBAHASAN

Data Tabel 1 diolah menggunakan pendekatan K-Means dalam penelitian ini. Pada kluster ketiga, kelima, dan ketujuh, digunakan label pemetaan untuk pengelompokan menggunakan Rapid Miner dengan proses pemetaan digambarkan seperti pada Gambar 3.



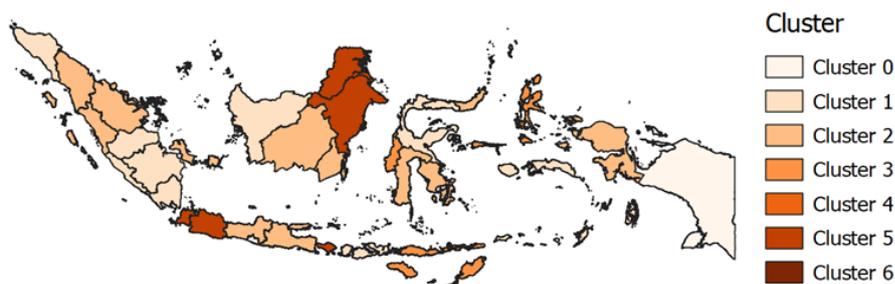
**Gambar 3.** Pemetaan dengan pemodelan Rapid Miner pada tiga kluster

Gambar 2 menunjukkan bagaimana teknik membaca Excel digunakan untuk memasukkan data berdasarkan data yang diberikan pada Tabel 1. Ini adalah implementasi model K-means untuk memetakan keterampilan TI di setiap provinsi. Sebagaimana ditunjukkan pada tabel 2, tiga kluster digunakan untuk meningkatkan ukuran kinerja berdasarkan Davies Bouldin Index (DBI) sehingga dapat ditentukan kluster mana yang paling akurat untuk pemetaan kemampuan teknologi informasi menurut provinsi.

Tabel 2. Kinerja Indeks Davied Bouldin

Cluster	DBI
Performance K=3	-0.408
Performance K=5	-0.439
Performance K=7	-0.357

Tabel 2 mengilustrasikan bagaimana kinerja setiap tabel berdasarkan Indeks DBI. Davies-Bouldin Index (DBI) adalah instrumen yang menggunakan pendekatan klustering untuk mengukur efektivitas evaluasi kluster. Semakin kecil nilai DBI (non-negatif  $\geq 0$ ) semakin kuat kluster (k) yang ditemukan oleh metode yang digunakan. Sehingga didapatkan cara paling efektif untuk mengatur 34 set data provinsi adalah dengan membagi tujuh kluster, dengan skor DBI terendah -0,357.



Gambar 4. Hasil Pemetaan Kemampuan Penggunaan Teknologi Informasi Provinsi Di Indonesia Tahun 2019-2021 Berdasarkan Tujuh Kluster

Provinsi Jawa Barat, Banten, Bali, Kalimantan Timur, dan Kalimantan Utara memiliki klaster terbesar dari total tujuh klaster, seperti terlihat pada Gambar 4. Hal ini menunjukkan bahwa provinsi yang termasuk dalam kategori ini memiliki persentase remaja terbesar (usia 15-24 tahun) dan orang dewasa (usia 15-59 tahun) yang memiliki kompetensi teknologi informasi dalam kurun waktu tiga tahun terakhir. Proporsi yang paling rendah terdapat pada klaster nol yang hanya terdapat di Provinsi Papua yang rata-rata memiliki persentase 29,32 persen untuk keterampilan teknologi informasinya selama tiga tahun terakhir. Hal ini menunjukkan bahwa kemampuan remaja dan orang dewasa di provinsi ini dalam menggunakan teknologi informasi berbeda secara signifikan dengan kemampuan orang dewasa dan remaja di provinsi lain.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian, pendekatan K-Means dapat digunakan untuk memetakan kemampuan teknologi informasi di Indonesia. Tiga kluster digunakan untuk meningkatkan bagaimana Davies Bouldin Index (DBI) digunakan untuk mengukur kinerja. Hal ini dilakukan agar dapat diketahui kluster mana yang memberikan pemetaan kompetensi TI paling akurat menurut provinsi. Kluster ketujuh memiliki skor DBI terendah sebesar -0,357 dari ketiga kluster, sehingga merupakan kluster terbaik. Ketujuh kluster ini ditampilkan menggunakan QGIS. Persentase terendah terdapat pada klaster nol yang hanya terdapat di Provinsi Papua dengan rata-rata 29,32 persen masyarakat memiliki keterampilan di bidang teknologi informasi.

#### REFERENCES

- [1] M. C. M. Bravo, C. S. Chalezquer, and J. Serrano-Puche, "Meta-framework of digital literacy: Comparative analysis of 21st century skills frameworks," *Revista Latina de Comunicacion Social*, vol. 2021, no. 79, 2021, doi: 10.4185/RLCS-2021-1508.
- [2] E. Y. Wijaya, D. A. Sudjimat, and A. Nyoto, "Transformasi Pendidikan Abad 21 sebagai Tuntutan Pengembangan Sumber Daya Manusia di Era Global," *Prosiding Seminar Nasional Pendidikan Matematika*, vol. 1, 2016.
- [3] S. A. Garba, Y. Byabazaire, and A. H. Busthami, "Toward the use of 21st century teaching-learning approaches: The trend of development in Malaysian schools within the context of Asia Pacific," *International Journal of Emerging Technologies in Learning*, vol. 10, no. 4, 2015, doi: 10.3991/ijet.v10i4.4717.
- [4] M. Megawati, Y. Andayani, and E. Junaidi, "Hubungan Kesiapan Teknologi Informasi Dan Komunikasi (TIK) Dengan Motivasi Belajar Kimia Siswa," *Chemistry Education Practice*, vol. 2, no. 2, 2019, doi: 10.29303/cep.v2i2.1414.
- [5] Petra Wiyakti Bodrogini, "Preparing ICT Skills for Digital Economy: Indonesia within the ASEAN context," Jakarta, Mar. 2018.

- [6] The SMERU Research Institute, *Digital Skills Landscape in Indonesia*. Jakarta: The SMERU Research Institute, 2022.
- [7] Badan Pusat Statistik, “Proporsi Remaja Dan Dewasa Usia 15-59 Tahun Dengan Keterampilan Teknologi Informasi Dan Komputer (TIK) Menurut Provinsi (Persen), 2019-2021,” 2022.
- [8] B. Setio and P. Prasetyaningrum, “PENERAPAN DATA MINING DALAM MENGELOMPOKKAN KUNJUNGAN WISATAWAN DI KOTA YOGYAKARTA MENGGUNAKAN METODE K-MEANS,” *Journal of Computer Science and Technology (JCS-TECH)*, vol. 1, no. 1, 2021, doi: 10.54840/jcstech.v1i1.9.
- [9] R. Watrionthos, W. A. Ritonga, A. Rengganis, A. Wanto, and M. Isa Indrawan, “Implementation of PROMETHEE-GAIA Method for Lecturer Performance Evaluation,” *J Phys Conf Ser*, vol. 1933, no. 1, p. 012067, Jun. 2021, doi: 10.1088/1742-6596/1933/1/012067.
- [10] D. Abdullah, S. Susilo, A. S. Ahmar, R. Rusli, and R. Hidayat, “The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data,” *Qual Quant*, vol. 56, no. 3, 2022, doi: 10.1007/s11135-021-01176-w.
- [11] I. Shaik, S. S. Nittela, T. Hiwarkar, and S. Nalla, “K-means clustering algorithm based on E-commerce big data,” *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, no. 11, 2019, doi: 10.35940/ijitee.K2121.0981119.
- [12] R. Watrionthos, M. Bobbi Kurniawan, Kusmanto, S. Budiman, and B. Ulya, “Mapping of Traffic Accidents in Labuhanbatu Regency using GIS Support,” *J Phys Conf Ser*, vol. 1566, no. 1, 2020, doi: 10.1088/1742-6596/1566/1/012104.
- [13] K. P. Sinaga and M. S. Yang, “Unsupervised K-means clustering algorithm,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [14] N. Azis *et al.*, “Mapping study using the unsupervised learning clustering approach,” *IOP Conf Ser Mater Sci Eng*, vol. 1088, no. 1, p. 012005, Feb. 2021, doi: 10.1088/1757-899X/1088/1/012005.
- [15] C. Yuan and H. Yang, “Research on K-Value Selection Method of K-Means Clustering Algorithm,” *J Multidisciplinary Scientific Journal*, vol. 2, no. 2, pp. 226–235, Jun. 2019, doi: 10.3390/j2020016.
- [16] Samsir *et al.*, “Naives Bayes Algorithm for Twitter Sentiment Analysis,” *J Phys Conf Ser*, vol. 1933, no. 1, p. 012019, Jun. 2021, doi: 10.1088/1742-6596/1933/1/012019.
- [17] M. Nasution, D. Irmayani, R. Watrionthos, S. Suryadi, and I. R. Munthe, “Comparative analysis of data mining using the rough set method with K-means method,” *International Journal of Scientific and Technology Research*, vol. 8, no. 5, 2019.
- [18] F. Rahman, I. I. Ridho, M. Muflih, S. Pratama, M. R. Raharjo, and A. P. Windarto, “Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country,” *IOP Conf Ser Mater Sci Eng*, vol. 835, no. 1, p. 012058, Apr. 2020, doi: 10.1088/1757-899X/835/1/012058.
- [19] R. Maliva and T. Missimer, “Geographic Information Systems,” in *Environmental Science and Engineering*, 2012. doi: 10.1007/978-3-642-29104-3\_19.
- [20] D. Farkas, B. Hilton, J. Pick, H. Ramakrishna, A. Sarkar, and N. Shin, “A tutorial on geographic information systems: A ten-year update,” *Communications of the Association for Information Systems*, vol. 38, no. 1, 2016, doi: 10.17705/1CAIS.03809.
- [21] Muttaqin, M. Zuhri, and A. Irfan, “Perancangan sistem pemetaan dan pendataan populasi penduduk miskin di Kota Banda Aceh menggunakan aplikasi Quantum GIS,” *Informatics and Computer Science*, vol. 5, no. 1, 2019.