



Signifikansi Pengaruh Akses Teknologi Informasi terhadap Indeks Pembangunan Manusia di Indonesia

Andriyan Rizki Jatmiko^{1*}, Nofrian Deny Hendrawan², Rizza Muhammad Arief³,
Firnanda Al-Islama Achyunda Putra⁴, Mochammad Daffa Putra Karyudi⁵

^{1,2,3,4,5} Universitas Merdeka Malang, Malang, Indonesia

^{1*}andriyan.jatmiko@unmer.ac.id, ²nofrian.hendrawan@unmer.ac.id, ³rizza@unmer.ac.id,

⁴firnanda.putra@unmer.ac.id, ⁵karyudi.mochammad@student.unmer.ac.id

^{*} andriyan.jatmiko@unmer.ac.id

Abstrak– Indeks Pembangunan manusia merupakan indikator majunya suatu negara, Teknologi Informasi merupakan pendukung penting untuk mengukur Indeks Pembangunan Manusia. Penelitian ini dapat memberikan gambaran umum untuk mengukur kemajuan suatu negara dilihat dari akses Teknologi Informasi. Penelitian ini mengolah data sekunder yang disediakan oleh Badan Pusat Statistik dari tahun 2017-2019. Menggunakan metode klustering secara *K-Means* maupun *K-Medoids*. *K-Means* adalah metode untuk mengelompokkan non-hierarkis populer yang mengelompokkan objek berdasarkan jarak ke titik pusat, yang bertujuan untuk memaksimalkan kemiripan di dalam kelompok. *K-Medoids* sebuah algoritma yang kuat menangani outlier menggunakan teknik seperti CLARA dan PAM. Pada tahun 2017 dengan rata-rata 0,294611827 Gorontalo termasuk kluster rendah sedangkan pada tahun 2018 dan 2019 Gorontalo memasuki kluster sedang dengan rata-rata 0,349570215 dan 0,394531648. Sama halnya dengan Sulawesi Tengah, Pada tahun 2017 dengan rata-rata 0,275848883 Sulawesi Tengah termasuk kluster rendah sedangkan pada tahun 2018 dan 2019 Sulawesi Tengah memasuki kluster sedang dengan rata-rata 0,291938731 dan 0,334276807 Dari hasil ini bisa dipastikan, dengan meningkatkan pengetahuan dalam Teknologi Informasi maka IPM pada suatu daerah dapat meningkat juga.

Kata Kunci: *K-Means*; *K-Medoids*; Klustering; Teknologi Informasi; Indeks Pembangunan Manusia

Abstract– Human Development Index is an indicator of the progress of a country, Information Technology is an important supporter to measure the Human Development Index. This research can provide an overview to measure the progress of a country in terms of access to Information Technology. This study processed secondary data provided by the Central Statistics Agency from 2017-2019. Using *K-Means* and *K-Medoids* clustering methods. *K-Means* is a popular non-hierarchical grouping method that groups objects by distance to a central point, aiming to maximize similarity within groups. *K-Medoids* is a powerful algorithm that handles outliers using techniques such as CLARA and PAM. In 2017 with an average of 0,294611827 Gorontalo was a low cluster while in 2018 and 2019 Gorontalo entered a medium cluster with an average of 0,349570215 and 0,394531648. Similar to Central Sulawesi, in 2017 with an average of 0,275848883 Central Sulawesi was included in the low cluster while in 2018 and 2019 Central Sulawesi entered the medium cluster with an average of 0,291938731 and 0,334276807 From this result, it can be ascertained, by increasing knowledge in Information Technology, the HDI in an area can increase as well.

Keywords: *K-Means*; *K-Medoids*; Clustering; Information Technology; Human Development Index

1. PENDAHULUAN

Teknologi informasi (TI) merupakan salah satu faktor penting yang dapat mendukung pembangunan di berbagai bidang, termasuk bidang sosial dan ekonomi. TI dapat memberikan akses informasi, pengetahuan, dan peluang bagi masyarakat untuk meningkatkan kualitas hidup dan kesejahteraan. Salah satu indikator yang dapat mengukur tingkat pembangunan suatu negara atau daerah adalah indeks pembangunan manusia (IPM). IPM merupakan ukuran komposit yang menggabungkan tiga dimensi dasar pembangunan manusia, yaitu harapan hidup, pendidikan, dan pendapatan. IPM dapat mencerminkan tingkat kesehatan, pengetahuan, dan kemampuan ekonomi masyarakat [1].



Indonesia merupakan negara yang memiliki keragaman geografis, budaya, dan sosial yang tinggi. Hal ini menyebabkan adanya perbedaan tingkat pembangunan antara provinsi- provinsi di Indonesia. Salah satu faktor yang dapat mempengaruhi perbedaan tersebut adalah akses TI yang dimiliki oleh masyarakat di setiap provinsi. Akses TI dapat diukur dengan beberapa indikator, seperti proporsi individu yang menggunakan internet dan proporsi individu yang menguasai atau memiliki telepon genggam. Akses TI dapat memudahkan masyarakat untuk mendapatkan informasi, berkomunikasi, belajar, bekerja, berbisnis, dan berpartisipasi dalam kegiatan sosial [2].

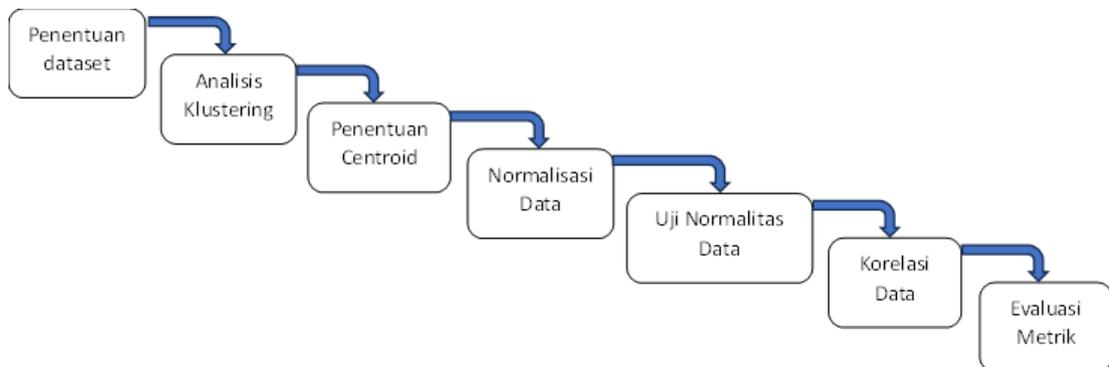
Tingkat signifikansi dari akses teknologi informasi terhadap Indeks Pembangunan Manusia Teknologi informasi memainkan peran penting dalam meningkatkan IPM karena memberikan aksesibilitas yang lebih luas terhadap informasi, pengetahuan, dan peluang bisnis ataupun lowongan kerja. Dari 3 variabel tersebut akan digunakan peneliti sebagai bahan acuan untuk proses penelitian signifikansi akses TI di suatu daerah [3].

Penelitian ini bertujuan untuk menganalisis pengaruh akses TI terhadap IPM di setiap provinsi di Indonesia. Penelitian ini menggunakan data IPM tahun 2017-2019 dari Badan Pusat Statistik (BPS) dan data proporsi individu yang menggunakan internet dan proporsi individu yang menguasai atau memiliki telepon genggam tahun 2017-2019 dari Survei Sosial Ekonomi Nasional (SUSENAS) BPS. Penelitian ini menggunakan metode data mining untuk menemukan pola dan hubungan antara variabel-variabel tersebut. Data mining merupakan proses mengekstrak informasi berguna dari data besar dengan menggunakan teknik-teknik analisis data, statistik, kecerdasan buatan, dan pembelajaran mesin [4].

2. METODE PENELITIAN

2.1 Tahapan penelitian

Tujuan dari dilakukannya penelitian ini untuk menganalisis pengaruh akses TI terhadap IPM di setiap provinsi di Indonesia dengan menggunakan metode penambangan data. Tahapan yang dilakukan penulis pada penelitian ini disajikan pada gambar 1.



Gambar 1. Tahapan penelitian

2.2 Penentuan Dataset

Data yang diolah dalam penelitian ini berupa data sekunder dengan pendekatan kuantitatif. Data ini didapatkan pada *website* Badan Pusat Statistik yang meliputi Indeks Pembangunan Manusia (IPM), Individu yang Menguasai/Memiliki Telepon Genggam, Individu Yang Menggunakan Internet, Lama Sekolah, Keterampilan Teknologi Informasi Dan Komputer (TIK), dan Umur Harapan Hidup (UHH). Analisis ini berdasarkan studi kasus 34 provinsi di Indonesia pada tahun 2017-2019.

Pada gambar 2 hingga gambar 4 merupakan tabel dataset dari tahun 2017 hingga 2019 setelah proses normalisasi data. Data yang telah dinormalisasi bertujuan mengubah nilai dalam dataset agar memiliki rentang yang seragam sehingga membantu meningkatkan performa. Berdasarkan gambar 2 ditunjukkan nilai rata-rata variabel. 4 urutan teratas yaitu Aceh dengan IPM bernilai 0.548879 hingga Bengkulu dengan nilai IPM 0.517883

dan 4 urutan terbawah adalah Sulawesi Utara dengan nilai IPM 0.599428 hingga Sumatera Utara dengan nilai IPM 0.547449.

KODE	IDProvinsi	IPM	Ponsel	Internet	lama_sekolah(tahun)	Melek_Teknologi	Umur_Harapan_Hidup	
0	11	ACEH	0.548879	0.439774	0.14386	0.570526	0.18503	0.498077
1	51	BALI	0.725322	0.735742	0.520616	0.48	0.539721	0.684615
2	36	BANTEN	0.587983	0.622419	0.474626	0.475789	0.483034	0.495192
3	17	BENGKULU	0.517883	0.54056	0.231083	0.463158	0.231737	0.408654
30	71	SULAWESI_UTARA	0.599428	0.630285	0.428863	0.604211	0.467265	0.644231
31	13	SUMATERA_BARAT	0.579399	0.568584	0.297236	0.515789	0.334132	0.426923
32	16	SUMATERA_SELATAN	0.465904	0.50885	0.207068	0.362105	0.214371	0.465385
33	12	SUMATERA_UTARA	0.547449	0.471731	0.251926	0.627368	0.275848	0.3875

Gambar 2. Tabel dataset tahun 2017 setelah normalisasi data

Pada gambar 3 menunjukkan provinsi dengan 4 urutan teratas yaitu Aceh dengan IPM bernilai 0.545321 hingga Bengkulu dengan nilai IPM 0.518373 dan 4 urutan terbawah adalah Sulawesi Utara dengan nilai IPM 0.594808 hingga Sumatera Utara dengan nilai IPM 0.544831.

KODE	IDProvinsi	IPM	Ponsel	Internet	lama_sekolah(tahun)	Melek_Teknologi	Umur_Harapan_Hidup	
0	11	ACEH	0.545321	0.517827	0.239741	0.567329	0.306936	0.494141
1	51	BALI	0.720725	0.766852	0.579482	0.470199	0.632773	0.693359
2	36	BANTEN	0.582558	0.662674	0.611447	0.463576	0.635608	0.494141
3	17	BENGKULU	0.518373	0.552925	0.282289	0.461369	0.305991	0.416016
30	71	SULAWESI_UTARA	0.594808	0.673538	0.453564	0.600442	0.510112	0.652344
31	13	SUMATERA_BARAT	0.571779	0.62312	0.365011	0.494481	0.439614	0.432617
32	16	SUMATERA_SELATAN	0.457129	0.575487	0.297192	0.326711	0.32319	0.47168
33	12	SUMATERA_UTARA	0.544831	0.528969	0.317063	0.622517	0.367038	0.393555

Gambar 3. Tabel dataset tahun 2018 setelah normalisasi data

Hingga tahun 2019, urutan nilai rata-rata dari seluruh variabel pada 4 provinsi teratas dan 4 provinsi terbawah tetap tidak berubah yang ditunjukkan pada gambar 4. Dimulai dari Aceh hingga Bengkulu untuk urutan teratas dan Sulawesi Utara hingga Sumatera Utara untuk urutan terbawah

KODE	IDProvinsi	IPM	Ponsel	Internet	lama_sekolah(tahun)	Melek_Teknologi	Umur_Harapan_Hidup	
0	11	ACEH	0.555221	0.482214	0.268547	0.573696	0.346049	0.5
1	51	BALI	0.72992	0.779309	0.62558	0.496599	0.66468	0.709901
2	36	BANTEN	0.582329	0.669088	0.667504	0.473923	0.689884	0.49703
3	17	BENGKULU	0.520582	0.568136	0.367465	0.471655	0.378917	0.434653
30	71	SULAWESI_UTARA	0.60994	0.736473	0.483578	0.630385	0.52844	0.669307
31	13	SUMATERA_BARAT	0.579819	0.604208	0.375773	0.514739	0.449591	0.444554
32	16	SUMATERA_SELATAN	0.460843	0.556112	0.31762	0.346939	0.341451	0.478218
33	12	SUMATERA_UTARA	0.547189	0.55511	0.380216	0.634921	0.431369	0.408911

Gambar 4. Tabel dataset tahun 2019 setelah normalisasi data

2.3 Analisis Klustering

Analisis Klustering adalah metode analisis yang digunakan untuk mengelompokkan karakteristik yang mirip sebagai objek penelitian ke dalam beberapa kelompok yang tidak sama. Metode klustering *K-Means* adalah metode klustering populer yang dengan unsur non-hierarki. Metode ini menggunakan algoritma yang berjalan dengan cara mengelompokkan objek demi objek berdasarkan jarak antar objek dan centroidnya. Jarak tersebut didapatkan dengan proses yang dijalankan berulang kali [5]. Tujuan dari penggunaan metode *K-Means* adalah guna memperoleh beberapa kluster objek dengan cara memaksimalkan kemiripan karakteristik antar objek dalam kluster dan memaksimalkan perbedaan antar kluster.

Metode klustering *K-Medoids* yaitu algoritma pengelompokan yang lebih kuat menghadapi nilai outlier dalam dataset. *K-Medoids* biasanya menggunakan *Clustering Large Application* (CLARA) dan *Partitioning Around Medoids* (PAM) [6]. Metode ini bekerja lebih baik bekerja dengan



menggunakan PAM. PAM ini bekerja lebih baik dalam memproses dataset dengan ukuran kecil. Sedangkan CLARA dirancang untuk memproses dataset dengan ukuran yang besar [7].

Metode *K-Means/K-Medoids* mengelompokkan objek dengan unsur non-hierarki sehingga dalam penggunaan metode ini harus menentukan sendiri berapa jumlah klusternya [8]. *Knee/Elbow* adalah sebuah teknik yang dapat digunakan untuk menentukan jumlah *centroid* (k). Dalam metode ini, untuk menentukan nilai k , akan dilakukan pengulangan secara terus menerus dari $k=1$ hingga $k=n$ (Di sini n adalah *hyperparameter* yang kita pilih sesuai dengan kebutuhan). Untuk setiap nilai k , kita menghitung nilai *Within-Cluster Sum of Squares* (WCSS). WCSS sendiri memiliki formula (1).

$$WCSS = \sum_i c_n \sum_j x_n (c_i - x_i)^2 \quad (1)$$

c_n = jumlah kluster

x_n = pusat kluster yang terdekat

c_i = pusat kluster ke- i

x_j = titik data ke- j dalam setiap kluster

Didefinisikan sebagai jumlah jarak kuadrat antara titik pusat dan setiap titik.

2.4 Penentuan Centroid

Centroid pada *K-Means* ditentukan secara acak dan bergantung pada keberuntungan yang hasilnya tidak konsisten pada setiap implementasi, sedangkan pada *K-Medoids* sendiri menggunakan *Partitioning Around Medoids* (PAM). PAM bertujuan untuk menemukan satu set awal dari *K-Medoids* dengan memilih titik-titik yang meminimalkan kerugian pengelompokan *K-Medoids* [9][10].

2.5 Normalisasi Data

Normalisasi data adalah serangkaian tahapan sistematis yang bertujuan mengubah data secara optimal. Melalui normalisasi mempermudah merancang model yang akan digunakan, sehingga data dapat digunakan dengan efisiensi maksimum dengan rumus (2) [8]:

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2)$$

x' = hasil normalisasi

x_i = nilai ke- i

x_{min} = nilai minimum dari dataset

x_{max} = nilai maximum dari dataset

2.6 Uji Normalitas Data

Dataset pada penelitian ini hanya terbatas sebanyak 34 sampel, maka yang dipilih adalah metode uji kenormalan *Shapiro-Wilk*. Karena metode *Shapiro-Wilk* lebih cocok digunakan pada dataset yang dibawah 50 sampel [11]. Uji ini dilakukan dengan membandingkan nilai statistik yang diamati dengan distribusi- t dengan menggunakan formula (3) [12]:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{(\sum_{i=1}^n x_i - \bar{x})^2} \quad (3)$$

$a_i x_{(i)}$ = titik data ke- i dalam sampel yang telah diurutkan



a_i = koefisien yang tergantung pada ukuran sampel dan ditentukan dari tabel distribusi normal standar.

$x_{(i)}$ = nilai ke-i terkecil dalam sampel.

x_i = nilai ke-i dari sampel data yang sedang diuji

\bar{x} = rata rata sampel

Kemudian dibandingkan dengan alpha (0.5) untuk menentukan apakah hipotesis nol dapat diterima atau ditolak yang akan menghasilkan nilai-p, jika nilai-p "kecil" ini dapat digunakan sebagai bukti yang menentang hipotesis nol dan mendukung hipotesis alternatif: bobot tidak diambil dari distribusi normal [12].

2.7 Korelasi Data

Uji korelasi dengan metode Pearson dan Spearman. Koefisien korelasi Pearson mengukur tingkat keterkaitan hubungan antara dua data. Nilai r (korelasi) mendekati -1 atau 1 mengindikasikan hubungan linier yang bagus, jika hampir 0 maka itu lemah [13].

2.8 Evaluasi Metrik

Evaluasi metrik menggunakan *Dunn Index* (DI) adalah salah satu ukuran evaluasi algoritma pengelompokan. Indeks ini paling sering digunakan untuk mengevaluasi kualitas pemisahan oleh algoritma pengelompokan untuk sejumlah kluster tertentu [14]. *Dunn Index* dihitung sebagai rasio dari jarak antar-kluster terkecil dan jarak intra-kluster terbesar. DI yang tinggi berarti pengelompokan yang lebih baik karena pengamatan di setiap kluster lebih dekat satu sama lain (lebih rapat).

Selain *Dunn Index* ada juga *Silhouette Coefficient* atau skor siluet adalah metrik yang digunakan untuk menghitung kualitas teknik pengelompokan. Nilainya berkisar dari -1 hingga 1. 1 berarti kluster terpisah satu sama lain dan dibedakan dengan jelas. 0 berarti cluster tidak peduli, atau dapat dikatakan bahwa jarak antar cluster tidak signifikan. -1 berarti cluster ditetapkan dengan cara yang salah. Ada pula *Davies-Bouldin Index* (DBI), DBI adalah ukuran seberapa baik pemisahan cluster dan berkisar dari 0 hingga tak terbatas. Sehingga nilai DBI yang lebih rendah menunjukkan pemisahan cluster yang lebih baik [15].

3. HASIL DAN PEMBAHASAN

3.1 Normalisasi Data

Langkah awal dalam melakukan analisis adalah dengan melakukan normalisasi data agar berada dalam skala yang sama. Pada penelitian ini dalam melakukan normalisasi data menggunakan metode *MinMax*. Setelah dilakukan normalisasi akan terlihat seperti ini dalam skala 0-1.

	KODE	IDProvinsi	IPM	Ponsel	Internet	lama_sekolah(tahun)	Melek_Teknologi	Umur_Harapan_Hidup
0	11	ACEH	0.548879	0.439774	0.14386	0.570526	0.18503	0.498077
1	51	BALI	0.725322	0.735742	0.520616	0.48	0.539721	0.684615
2	36	BANTEN	0.587983	0.622419	0.474626	0.475789	0.483034	0.495192
3	17	BENGKULU	0.517883	0.54056	0.231083	0.463158	0.231737	0.408654
30	71	SULAWESI_UTARA	0.599428	0.630285	0.428863	0.604211	0.467265	0.644231
31	13	SUMATERA_BARAT	0.579399	0.568584	0.297236	0.515789	0.334132	0.426923
32	16	SUMATERA_SELATAN	0.465904	0.50885	0.207068	0.362105	0.214371	0.465385
33	12	SUMATERA_UTARA	0.547449	0.471731	0.251926	0.627368	0.275848	0.3875

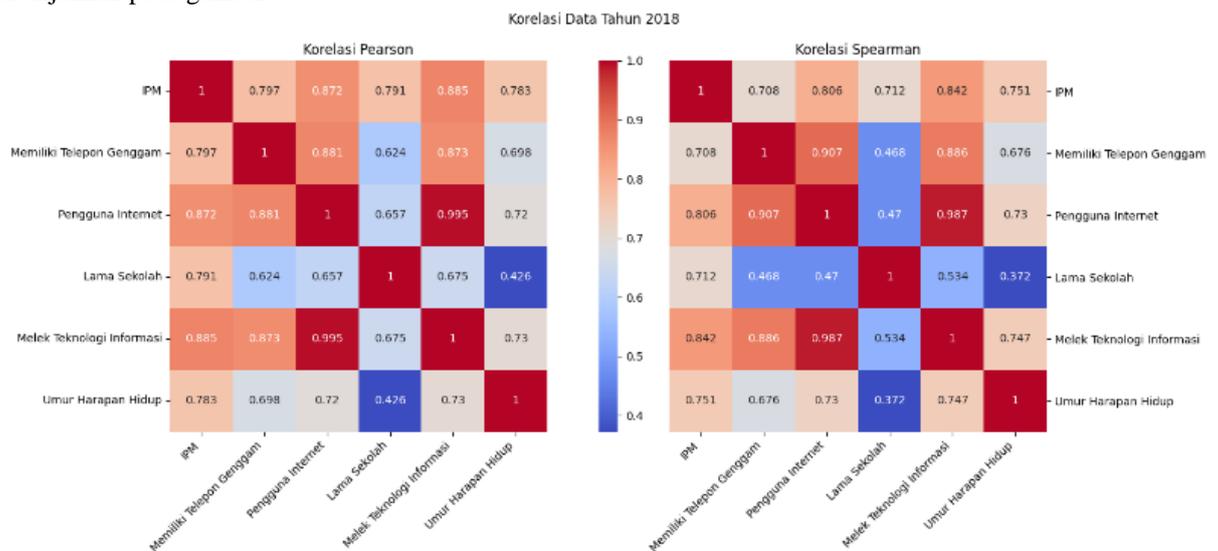
Gambar 5. Tabel dataset tahun 2017 setelah normalisasi data

3.2 Uji Normalitas

Hasil dari uji normalitas data 3 tahun ini bahwa dari 6 variabel dari masing-masing tahun, ada 2 buah variabel yang terdistribusi secara tidak normal. Variabel pengguna internet menunjukkan tahun 2017 terdistribusi secara tidak normal dengan nilai p-value 0,0113. Pada variabel melek teknologi informasi tahun 2017 juga menunjukkan ketidaknormalan dengan nilai p-value 0,0095.

3.3 Uji Korelasi

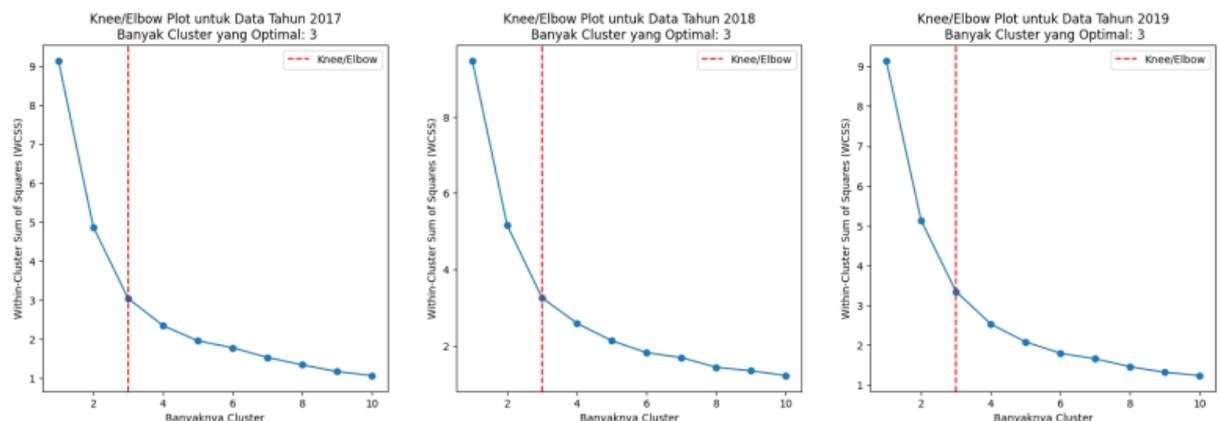
Saat dilakukan uji korelasi menggunakan 2 buah metode, dapat dihasilkan bahwa pasangan variabel yang memiliki korelasi yang kuat adalah IPM, Pengguna Internet, dan Melek Teknologi Informasi. Dapat dilihat pada gambar 5 bahwa 3 variabel tersebut mendekati nilai 1, dimana nilai 1 adalah nilai korelasi positif sempurna dan akan digunakan untuk acuan proses klustering. Grafik Heatmap korelasi data tahun 2018 ditunjukkan pada gambar 6.



Gambar 6. Grafik Heatmap korelasi data tahun 2018

3.4 Menentukan Jumlah Kluster Optimal

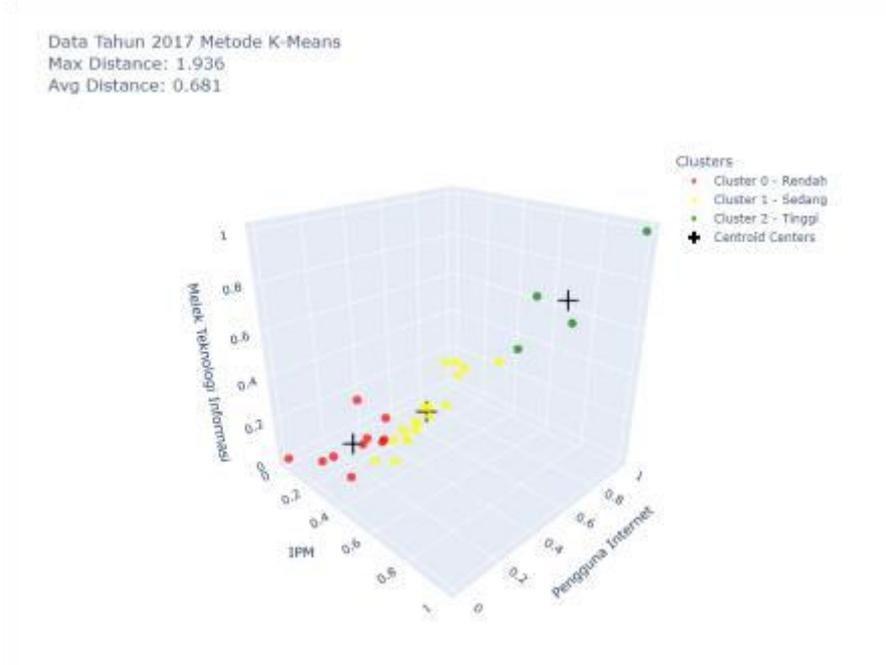
Langkah selanjutnya adalah menentukan jumlah kluster yang optimal dengan metode Knee/Elbow. Gambar 7 menunjukkan bahwa perbandingan dari 1-10 kluster, ditemukan bahwa yang paling optimal adalah 3 kluster menurut metode Knee/Elbow. Maka dari 34 provinsi di Indonesia nantinya akan dibagi menjadi 3 kluster menurut kemipiran karakteristiknya.



Gambar 7. Grafik Knee/Elbow data tahun 2017 hingga 2019

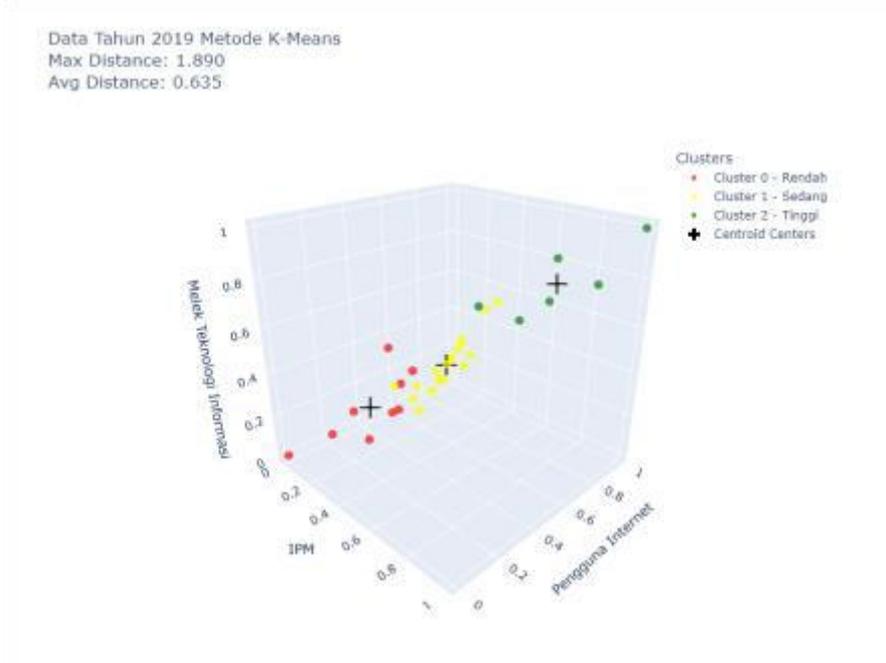
3.5 Analisis K-Means

Langkah selanjutnya adalah melakukan klustering dengan metode *K-Means*. Gambar 8 menunjukkan bahwa warna kuning lebih mendominasi. Kluster 0 sebanyak 9 unit yang mewakili 26% dari keseluruhan populasi. Kluster 1 sebanyak 20 unit yang mewakili 59%, dan Kluster 2 sebanyak 5 unit yang mewakili 15% dari keseluruhan populasi.



Gambar 8. Grafik 3D Scatter Plot hasil dari *K-Means* klusterling data tahun 2017

Tahun 2018 dan 2019 terepresentasikan pada gambar 9 ada sedikit perbedaan. Dimana kluster 2 bertambah 1 unit dan kluster 1 berkurang 1 unit.



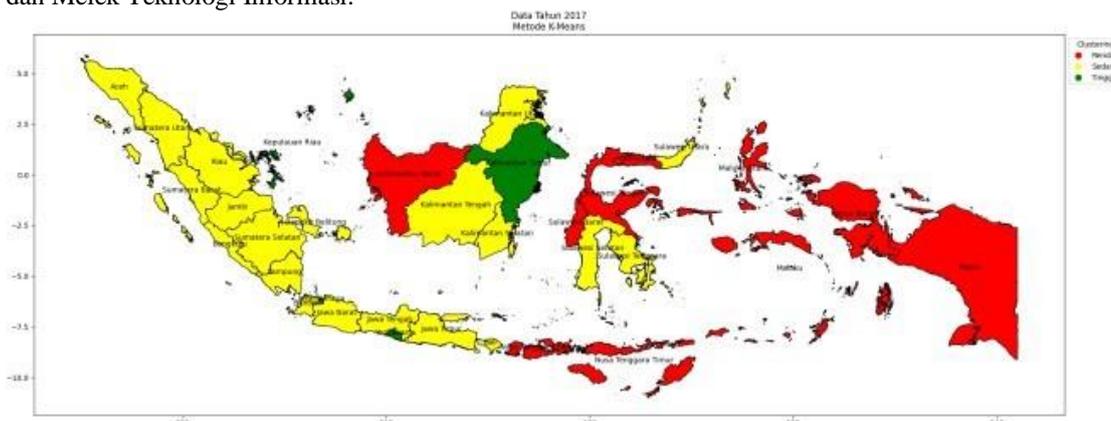
Gambar 9. Grafik 3D Scatter Plot hasil dari *K-Means* klusterling data tahun 2019

Setelah penggambaran pada Scatter Plot, dilakukan penggambaran geografis. Gambar 10 dan 11 disini ada sedikit perubahan pada provinsi Kalimantan Utara. Provinsi ini naik ke kluster 2 yang semulanya pada kluster 1.

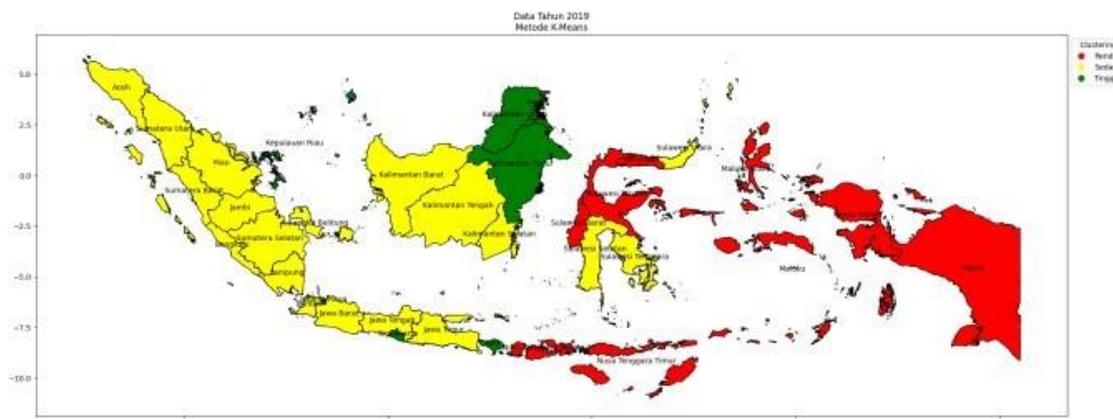
Tabel 1. Perubahan posisi kluster provinsi Kalimantan Utara

Tahun	Rata-rata	Kluster
2017	0,483358379	1
2018	0,58328207	2
2019	0,603345502	2

Hasil kalkulasi tabel 1 digunakan 3 variabel yang digunakan pada analisis klustering yaitu IPM, Pengguna Internet, dan Melek Teknologi Informasi.



Gambar 10. Grafik geografis hasil dari *K-Means* klustering data tahun 2017

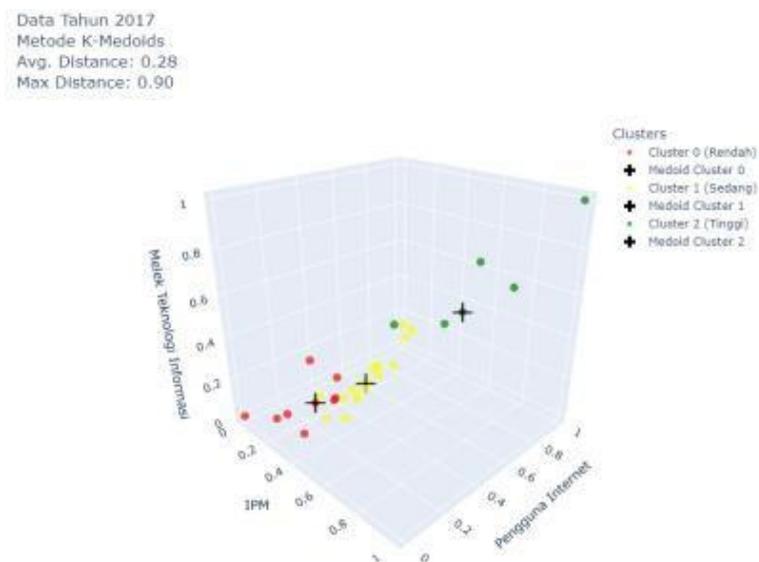


Gambar 11. Grafik geografis hasil dari *K-Means* klustering data tahun 2019

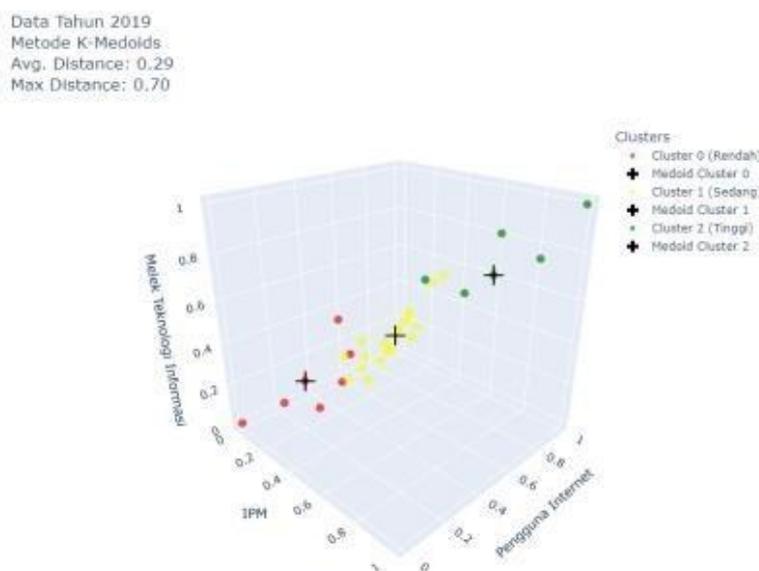
3.6 Analisis K-Medoids

Pada penggunaan *K-Medoids* masih sama pada jumlah kluster yang dipakai. Dari 3 kluster yang ada, warna kuning juga masih mendominasi. Gambar 12 menunjukkan perbedaan dengan metode *K-Means* pada tahun yang sama dengan letak perbedaan pada unit kluster 1 dan 2.

Pada gambar 8 metode *K-Means*, jumlah kluster 2 sebanyak 5 unit, sedangkan dapat dilihat perbedaan pada gambar 12 jumlah kluster 2 sebanyak 6 unit. Sama halnya dengan gambar 8 metode *K-Means*, jumlah kluster 1 sebanyak 20 unit, sedangkan pada gambar 13 jumlah kluster 1 sebanyak 19.



Gambar 12. Grafik 3D Scatter Plot hasil dari *K-Medoids* kluster ing tahun 2017



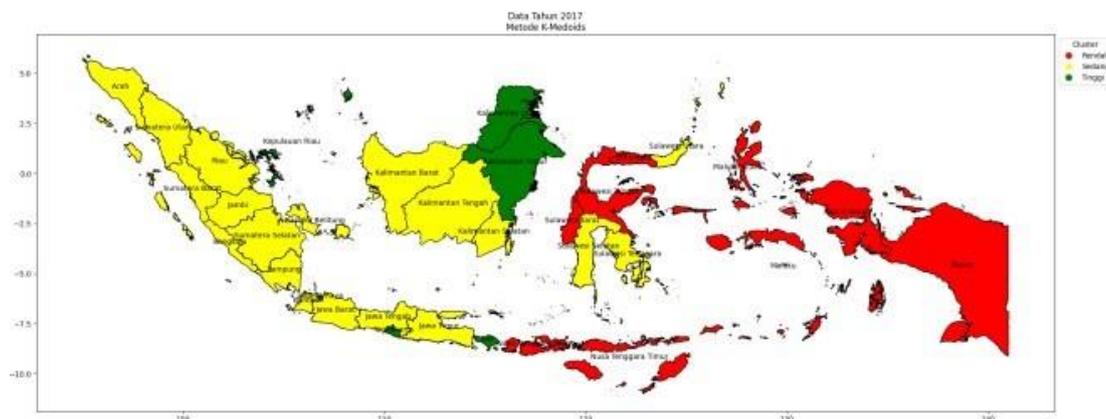
Gambar 13. Grafik 3D Scatter Plot hasil dari *K-Medoids* kluster ing tahun 2019

Sama seperti metode *K-Means*. Pada metode *K-Medoids* ini juga dilakukan penggambaran peta geografis agar lebih mudah dipahami secara visual. Seperti yang ditunjukkan gambar 14 dan 15. Pulau Sulawesi ada perbedaan warna pada 2 provinsi. Provinsi Gorontalo dan Sulawesi Tengah telah naik ke cluster 1 dari yang awalnya cluster 0.

Tabel 2. Perubahan posisi cluster provinsi Gorontalo dan Sulawesi Tengah

Tahun	Provinsi	Rata-rata	Kluster
2017	Gorontalo	0,294611827	0
	Sulawesi Tengah	0,275848883	0
2018	Gorontalo	0,349570215	1
	Sulawesi Tengah	0,291938731	1
2019	Gorontalo	0,394531648	1
	Sulawesi Tengah	0,334276807	1

Hasil kalkulasi pada tabel II digunakan 3 variabel yang digunakan pada analisis klustering yaitu IPM, Pengguna Internet, dan Melek Teknologi Informasi. Banyak provinsi yang masih tertinggal di negara Indonesia. Khususnya pada Indonesia bagian timur. Mayoritas berwarna merah pada gambar 10-11 dan 14-15.



Gambar 14. Grafik geografis hasil dari *K-Medoids* klustering data tahun 2017



Gambar 15. Grafik geografis hasil dari *K-Medoids* klustering data tahun 2019

3.7 Evaluasi yang Digunakan

Untuk melakukan perbandingan metrik, perlu dilakukan perhitungan *Dunn Index*, *Silhouette Score*, dan *DBI Score*. setelah dilakukan perhitungan hasil data sebagai berikut:

Tabel 3. Perbandingan *dunn index* pada tiap metode

Metode	Dunn Index	Tahun
<i>K-Means</i>	0.652445	2017
	0.620497	2018
	0.580649	2019
<i>K-Medoids</i>	0.594985	2017
	0.773590	2018
	0.717114	2019

Tabel 4. Perbandingan *silhouette score* pada tiap metode

Metode	Silhouette Score	Tahun
K-Means	0.369485	2017
	0.350184	2018
	0.332085	2019
K-Medoids	0.346985	2017
	0.361129	2018
	0.350570	2019

Tabel 5. Perbandingan *DBI score* pada tiap metode

Metode	DBI Score	Tahun
K-Means	0.863728	2017
	0.911409	2018
	0.952819	2019
K-Medoids	0.897080	2017
	0.892665	2018
	0.925938	2019

4. KESIMPULAN

Didapati perbedaan yang cukup signifikan pada tiap tahunnya, K-Means memiliki nilai Dunn Index yang tinggi pada tiap tahunnya dengan rata-rata 0,617863667, sedangkan K-Medoids 0,456191667, berbeda dengan Dunn Index, Silhouette Score pada K-Medoids memiliki rata-rata lebih tinggi yaitu 0,352894667, untuk K-Means sendiri memiliki rata-rata 0,350584667, begitu pula dengan DBI Score, dengan rendahnya rata-rata DBI Score K-Medoids 0,905227667, sedangkan K-Means 0,909318667 yang memiliki lebih tinggi pada bagian DBI Score. Walaupun K-Medoids memiliki nilai Dunn Index yang lebih rendah, namun Silhouette Coefficient yang tinggi serta diperkuat dengan DBI yang rendah sehingga dapat dibenarkan bahwa Provinsi Gorontalo dan Sulawesi Tengah telah terjadi peningkatan. Untuk data yang paling berdampak pada IPM adalah Melek Teknologi (Remaja dan Dewasa serta data Pengetahuan Teknologi) dan Pengguna Internet

REFERENSI

- [1] A. S. Kristiawan, "Analisis Pengaruh Produk Domestik Bruto, ICT Development Index Dan Good Governance Terhadap Indeks Pembangunan Manusia Anggota ASEAN Dan SAARC Kategori Lower-Upper Medium," *Jurnal Dinamika Ekonomi Pembangunan*, vol. 3, no. 2, pp. 140–156, 2020, doi: 10.14710/jdep.3.2.140-156.
- [2] S. Yoga, "Perubahan Sosial Budaya Masyarakat Indonesia Dan Perkembangan Teknologi Komunikasi," *Jurnal Al-Bayan*, vol. 24, no. 1, pp. 29–46, 2019, doi: 10.22373/albayan.v24i1.3175.
- [3] T. Oktavia, "Analisis Pengaruh Teknologi Informasi Dan Komunikasi (Tik) Serta Pendidikan Terhadap Pertumbuhan Ekonomi," *Prosiding National Simposium & Conference Ahlimedia*, vol. 1, no. 1, pp. 139–146, 2020, doi: 10.47387/nasca.v1i1.26.
- [4] Hermawan and H. Hasugian, "Penerapan Data Mining Untuk Clustering Indeks Pembangunan Manusia Berdasarkan Provinsi Di Indonesia," *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI) Jakarta-Indonesia*, no. September, pp. 525–532, 2022.
- [5] M. S. Yang and I. Hussain, "Unsupervised Multi-View K-Means Clustering Algorithm," *IEEE Access*, vol. 11, no. January, pp. 13574–13593, 2023, doi: 10.1109/ACCESS.2023.3243133.



- [6] H. L. Siregar, H. Lestari Siregar, M. Zarlis, and S. Efendi, "JURNAL MEDIA INFORMATIKA BUDIDARMA Cluster Analysis using K-Means and K-Medoids Methods for Data Clustering of Amil Zakat Institutions Donor," vol. 7, no. April, pp. 668–677, 2023, doi: 10.30865/mib.v7i2.5315.
- [7] I. H. Rifa, H. Pratiwi, and R. Respatiwan, "Clustering of Earthquake Risk in Indonesia Using K-Medoids and K-Means Algorithms," *Media Statistika*, vol. 13, no. 2, pp. 194–205, 2020, doi: 10.14710/medstat.13.2.194-205.
- [8] I. Fahmiyah and R. A. Ningrum, "Human Development Clustering in Indonesia: Using K-Means Method and Based on Human Development Index Categories," *Journal of Advanced Technology and Multidiscipline*, vol. 2, no. 1, pp. 27–33, 2023, doi: 10.20473/jatm.v2i1.45070.
- [9] K. E. Setiawan, A. Kurniawan, A. Chowanda, and D. Suhartono, "Clustering models for hospitals in Jakarta using fuzzy c-means and k-means," *Procedia Computer Science*, vol. 216, no. 2022, pp. 356–363, 2023, doi: 10.1016/j.procs.2022.12.146.
- [10] M. Tiwari, M. J. Zhang, J. Mayclin, S. Thrun, C. Piech, and I. Shomorony, "BanditPAM: Almost Linear Time k-Medoids Clustering via Multi-Armed Bandits," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., Curran Associates, Inc., 2020, pp. 10211–10222.
- [11] P. Mishra, C. M. Pandey, U. Singh, A. Gupta, C. Sahu, and A. Keshri, "Descriptive statistics and normality tests for statistical data," *Annals of Cardiac Anaesthesia*, vol. 22, no. 1, pp. 67–72, 2019, doi: 10.4103/aca.ACA_157_18.
- [12] G. D. Ahadi, N. Nur, and L. Ersela, "The Simulation Study of Normality Test Using Kolmogorov-Smirnov, Anderson-Darling, and Shapiro-Wilk," *Eigen Mathematics Journal*, vol. 6, no. 1, pp. 11–19, 2023.
- [13] A. Sergio, M. Zen, R. Wahyuni, and D. Nohe, "Hubungan Jumlah Penduduk Miskin Dengan Berat Badan Lahir Rendah Di Kalimantan Timur Menggunakan Kolerasi Person Dan Spearman," *Prosiding Seminar Nasional Matematika, Statistika dan Aplikasinya*, pp. 267–278, 2022.
- [14] H. Malikhatin, A. Rusgiyono, and D. A. I. Maruddani, "PENERAPAN k-MODES CLUSTERING DENGAN VALIDASI DUNN INDEX PADA PENGELOMPOKAN KARAKTERISTIK CALON TKI MENGGUNAKAN R-GUI," *Jurnal Gaussian*, vol. 10, no. 3, pp. 359–366, 2021, doi: 10.14710/j.gauss.v10i3.32790.
- [15] A. D, D. IC, and P. K, "Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," *Jurnal Matrix*, vol. 9, no. 3, p. 102, 2019.