Jurnal Sains Komputer & Informatika (J-SAKTI)
Volume 5 Nomor 2, September 2021, pp. 759-767
ISSN: 2548-9771/EISSN: 2549-7200
https://tunasbangsa.ac.id/ejurnal/index.php/jsakti

# Klasterisasi Angka Usia Muda Melek TIK Berdasarkan Algoritma K-Means Menurut jumlah Provinsi Indonesia

**Olivia Immanuela Massie[1], Tesa Nur Padilah[2]**
[1,2]Teknik Informatika, Fakultas Ilmu Komputer, Universitas SingaperbangsaaKarawang
Ds. Puseurjaya, Kec. Telukjambe Timur, Kab. Karawang, Prov. Jawa Barat
olivia.massie17169@student.unsika.ac.id[1]

***Abstract***

*Advances in technology that can't be separated from human life, making information easier to obtain. The results of a survey conducted by APJII on the penetration of internet users in 2018 based on age, stated that the use of the internet was dominated by young people. For this reason, it is hoped that in the future there will be improvements in numbers in the form of equitable use of ICT in all provinces in Indonesia. This research was conducted based on the clustering of young people who are ICT literate. The amount of data used is in accordance with the current number of Indonesian provinces, which are thirty-four provinces using the K-Means algorithm. The dataset in this study was obtained from the official government website https://.www.bps.go.id/ from 2017 to 2019. Clustering was carried out only to group provinces into two types of groups, which can later be used as evaluation material for the government in the framework of the equitable distribution of ICT in each province. The final result of this study is that there are twenty-five provinces that are included in cluster 1 and nine provinces are included in cluster group 2. Thus it is necessary to increase the ICT literacy rate for provincial clusters whose values are still lagging behind other provinces.*

***Keywords****: Clustering, K-Means, Data Mining, ICT, Youth, Indonesia*

***Abstrak***

*Kemajuan teknologi yang tak dapat dipisahkan dari kehidupan manusia, membuat informasi menjadi lebih mudah untuk didapatkan. Hasil survei yang dilakukan oleh APJII tentang penetrasi pengguna internet pada tahun 2018 berdasarkan umur, menyatakan bahwa penggunaan internet lebih banyak dikuasai oleh usia muda. Oleh karena alasan tersebut, maka diharapkan, kedepannya terdapat perbaikan angka berupa pemerataan penggunaan TIK di seluruh provinsi di Indonesia. Penelitian ini dilakukan berdasarkan klasterisasi angka usia muda yang melek TIK. Jumlah data yang digunakan sesuai dengan jumlah provinsi Indonesia saat ini, yaitu tiga puluh empat provinsi dengan menggunakan algoritma K-Means. Dataset pada penelitian ini diperoleh dari website resmi pemerintah https://.www.bps.go.id/ dari tahun 2017 sampai dengan tahun 2019. Clustering dilakukan hanya untuk mengelompokan provinsi menjadi dua jenis kelompok, yang nantinya dapat dijadikan bahan evaluasi terhadap pemerintah dalam rangka pemerataan TIK di setiap provinsinya. Hasil akhir dari penelitian ini adalah terdapat dua puluh lima provinsi yang termasuk ke dalam kelompok cluster 1 dan sembilan provinsi yang termasuk ke dalam kelompok cluster 2. Dengan demikian perlu adanya usaha peningkatan angka melek TIK untuk kluster provinsi yang nilainya masih tertinggal dengan berbagai provinsi lainnya.*

***Kata kunci****: Klasterisasi, K-Means, Data Mining, TIK, Usia Muda, Indonesia*

## 1. PENDAHULUAN

Kemajuan teknologi yang tak dapat dipisahkan dari kehidupan manusia, membuat informasi menjadi lebih mudah untuk didapatkan. Semua

informasi yang bernilai baik atau buruk perlahan-lahan telah mengubah pola pemikiran dan kehidupan manusia, sehingga perlu kebijakan dari pengguna teknologi itu sendiri untuk memanfaatkan penggunaa teknologi secara baik dan optimal (1).

Berdasarkan hasil *survey* APJII tentang penetrasi pengguna internet tahun 2018, terdapat kondisi yang cukup menarik, yaitu penggunaan internet umumnya, lebih banyak dikuasai oleh pengguna yang berusia 15-19 tahun untuk posisi pertama, kemudian disusul oleh pengguna yang berusia 20-24 tahun dan berbagai kategori usia untuk posisi lainnya. Perolehan posisi tersebut dapat menunjukkan keadaan, bahwa usia muda adalah usia yang lebih cenderung melek teknologi bila dibandingkan dengan usia-usia pengguna lainnya karena perolehan angka persentase yang lebih tinggi (2).

Konsep usia muda atau kaum muda itu sendiri dalam penelitian ini merujuk pada rekomendasi ILO *(International Labour Organization)* dan KILM 1999 *(Key Indicators of the Labour Market)* yang menyatakan, bahwa penduduk kelompok usia muda adalah penduduk yang berusia 15-24 tahun. Memiliki konsep yang serupa, GBHN (Garis Besar Haluan Negara) pun menyebutkan, bahwa kaum muda adalah penduduk yang berusia 15-29 tahun, namun sudah disesuaikan dengan pemahaman yang berlaku di dunia internasional (3).

Penelitian-penelitian terdahulu pun pernah melakukan berbagai penelitian serupa dalam berbagai aspek kehidupan manusia, misalnya pengelompokan persentase buta huruf pada rentang usia 15-44 tahun menurut Provinsi di Indonesia (4), *clustering* daerah miskin di Provinsi Riau (5), *clustering* Ujian Nasional untuk SMP di Indonesia periode ajaran tahun 2018/2019 (6) dan pengelompokan angka harapan hidup untuk kelahiran menurut Provinsi di Indonesia (7).

Dengan demikian, yang menjadi keterbaruan dalam penelitian ini adalah dengan melakukannya klasterisasi angka usia muda yang melek TIK berdasarkan jumlah provinsi di Indonesia yang menggunakan algoritma K-Means diharapkan, kedepannya terdapat perbaikan angka berupa terjadinya pemerataan penggunaan TIK pada usia muda yang melek TIK di seluruh provinsi yang tersebar di Indonesia. Penelitian ini menggunakan data primer yang diperoleh dari *website* resmi pemerintah di https://.www.bps.go.id/ dari tahun 2017 sampai dengan tahun 2019. Klasterisasi dilakukan menjadi 2 (dua) jenis kelompok yang berbeda.

## 2. METODOLOGI PENELITIAN

### 2.1. *Data Mining*

*Data mining* memanfaatkan beberapa ilmu seperti statistik, kecerdasan buatan dan *machine learning* untuk menemukan, menggali serta menambang pengetahuan yang memiliki nilai manfaat bagi penggunanya (4). Dalam penerapannya, terdapat beberapa teknik yang lumrah digunakan, seperti klasifikasi, *clustering*, asosiasi, predikisi dan estimasi (7).

Jurnal Sains Komputer & Informatika (J-SAKTI)
Volume 5 Nomor 2, September 2021, pp. 759-767
ISSN: 2548-9771/EISSN: 2549-7200
https://tunasbangsa.ac.id/ejurnal/index.php/jsakti

## 2.2. *Clustering*

*Clustering* merupakan salah satu teknik fungsionalitas dari *data mining*, bekerja dengan cara mengelompokkan data menjadi suatu kelompok data tertentu *(cluster)* (6). Pengelompokan dilakukan berdasarkan nilai kemiripan karakteristik antar data *(similarity)*. Memiliki sifat tanpa arahan *(unsupervised)*, *clustering* tidak memerlukan adanya data latih atau *data set*. Terdapat dua jenis metode *clustering* yang lumrah digunakan, yaitu *hierarchical clustering* dan *non-hierarchical clustering* (8).

Algoritma *cluster* bertujuan untuk membentuk klaster yang koheren secara internal, tetapi berbeda dengan klaster lainnya, atau dengan kata lain, data yang berada dalam satu *cluster* harus memiliki kemiripan dan harus berbeda dengan *cluster* lainnya (4). Jadi, *clustering* memiliki karakateristik untuk meminimalisasi variasi dan memaksimalisasi variasi (9).

## 2.3. *K-Means*

Pertama kali dipublikasikan pada tahun 1984 oleh Stuart Lloyd, *K-Means* menjadi salah satu algoritma yang populer (6). *K-Means* adalah contoh dari pengelompokkan data *non-hierarchical clustering*, cara kerja *K-Means* dalah dengan mengharuskan setiap data yang ada untuk masuk ke dalam suatu *cluster* dan memungkinkan bagi setiap data untuk berpindah ke *cluster* lain, jenis ini dinamakan *partitioning clustering*. *K-Means* terkenal karena kemudahannya yang mampu mengklaster data yang berjumlah besar dan mampu mengatasi *outlier* dengan cepat (Yunita, Efendi, & Rini, 2019).

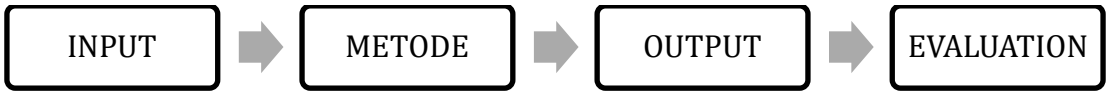
Secara umum, klasterisasi pada algoritma *K-Means* dapat dilakukan dengan menentukan jumlah nilai *cluster* (k) terlebih dahulu, lalu melakukan inisialisasi k pusat *cluster (centroid)* secara *random* dan menempatkan setiap objek ke *cluster* terdekat berdasarkan perhitungan jarak menggunakan persamaan *Euclidean Distance* (10). *Euclidean distance* adalah salah satu perhitungan untuk proses *K-Means* dengan cara mengukur jarak antar titik yang berbeda. Adapun rumus yang digunakan adalah :

$$d(i,j) = \sqrt{\left(X_{i1} - X_{j1}\right)^2 + \left(X_{i2} - X_{j2}\right)^2} \quad (1)$$

Keterangan :
$d(i,j) = jarak\ euclidean$
$X_i = nilai\ titik\ 1$
$X_j = nilai\ titik\ 2$

## 2.4. Metodologi Penelitian

Ada beberapa tahapan atau langkah yang akan diterapkan dalam penelitian ini, yaitu :




**Gambar 1.** Alur penelitian

a. *Input*
   Tahapan ini berguna untuk mempersiapkan data yang akan digunakan dalam penelitian. Tahapan ini memuat langkah *preparation* dan *preprocessing data*. Jenis data yang digunakan adalah *data public*. *Data public* berasal dari *website* resmi pemerintah yang berada di laman https://.www.bps.go.id/. Data disajikan dalam bentuk persen dengan periode tahun 2017 hingga tahun 2019.
b. Metode
   Tahapan ini berguna untuk menerapkan algoritma yang akan digunakan dalam penelitian, yaitu algoritma *K-Means*. Pemilihan algoritma ini berdasarkan alasan pada penerapannya yang mudah (11) dan penggunaannya yang populer (12).
c. *Output*
   Tahapan ini berguna untuk menampilkan proses dari *data mining* yang telah dilakukan dalam penelitian ini yang selanjutnya akan dimanfaatkan sebagai bahan evaluasi.
d. *Evaluation*
   Tahapan ini berguna untuk menilai hasil dari data mining yang telah didapatkan dari proses sebelumnya. Hasil evaluasi akan menggambarkan bagaimana kualitas dari proses *data mining* yang dilakukan.

## 3. HASIL DAN PEMBAHASAN

### 3.1. Input

Data primer yang digunakan berjumlah 34 baris dan setiap baris terdapat 4 buah atribut, yaitu Provinsi, 2017, 2018 dan 2019. Pada data tersebut, tidak ada *record* yang akan dihilangkan atau dihapus. Selain itu, tidak ditemukannya nilai *null* atau kosong, sehingga dapat dipastikan bila data sudah *clean* dan dapat digunakan untuk langkah selanjutnya. Perlu diketahui pula, bahwa setiap *record* yang ada pada data akan mewakili bagaimana nilai angka yang melek TIK untuk setiap provinsi, dan berikut ini adalah tabel data yang digunakan:

**Tabel 1.** Data BPS/Penelitian
(Persentase Keterampilan TIK Umur 15-24)

| Provinsi | 2019 | 2018 | 2017 |
|---|---|---|---|
| Yogyakarta | 97,91 | 95,48 | 92,19 |
| Jakarta | 95,41 | 92,15 | 89,93 |
| Riau | 93,05 | 86,37 | 84,48 |
| Bali | 91,40 | 87,56 | 81,45 |
| Jawa Tengah | 90,93 | 83,33 | 75,51 |
| Kalimantan Timur | 89,87 | 83,52 | 76,99 |
| Jawa Barat | 88,79 | 82,69 | 74,78 |
| Jawa Timur | 87,97 | 80,51 | 73,07 |
| Banten | 87,86 | 81,53 | 71,44 |
| Kalimantan Utara | 87,29 | 80,51 | 72,13 |
| Kalimantan Selatan | 87,16 | 78,81 | 69,00 |

| Provinsi | 2019 | 2018 | 2017 |
|---|---|---|---|
| Bangka Belitung | 84,49 | 77,06 | 64,40 |
| Sulawesi Selatan | 83,30 | 76,72 | 66,90 |
| Riau | 82,03 | 75,76 | 66,69 |
| Lampung | 81,61 | 73,25 | 58,71 |
| Jambi | 81,58 | 74,00 | 62,67 |
| Sulawesi Utara | 81,22 | 74,69 | 70,78 |
| Kalimantan Tengah | 80,96 | 68,48 | 61,19 |
| Sumatera Barat | 80,64 | 76,27 | 67,78 |
| Gorontalo | 79,60 | 74,22 | 61,68 |
| Sulawesi Tenggara | 79,51 | 70,66 | 59,71 |
| Nusa Tenggara Barat | 77,54 | 66,23 | 54,67 |
| Sumatera Utara | 77,28 | 71,24 | 63,90 |
| Sumatera Selatan | 77,26 | 72,36 | 61,55 |
| Bengkulu | 77,02 | 69,56 | 60,07 |
| Kalimantan Barat | 70,40 | 63,05 | 51,82 |
| Aceh | 69,01 | 62,96 | 51,78 |
| Sulawesi Tengah | 68,81 | 61,45 | 56,01 |
| Sulawesi Barat | 68,06 | 57,05 | 49,39 |
| Papua Barat | 65,40 | 61,29 | 50,04 |
| Maluku | 60,17 | 54,29 | 49,60 |
| Maluku Utara | 52,88 | 49,32 | 35,19 |
| Nusa Tenggara Timur | 51,03 | 43,32 | 38,68 |
| Papua | 32,88 | 33,48 | 30,42 |

### 3.2. Metode

Selanjutnya, untuk tahapan metode, langkah pertama yang dilakukan adalah mengimpor data. Kemudian melakukan pengecekan data dari keseluruhan atribut yang akan digunakan. Proses ini dilakukan dengan bantuan sebuah *tools* yang bernama *RStudio* dan berikut ini adalah rangkuman dari atribut 2017, 2018 dan 2019.

```
      2019            2018
 Min.   :32.88   Min.   :33.48
 1st Qu.:72.06   1st Qu.:63.84
 Median :81.09   Median :74.11
 Mean   :78.24   Mean   :71.74
 3rd Qu.:87.72   3rd Qu.:80.51
 Max.   :97.91   Max.   :95.48
      2017
 Min.   :30.42
 1st Qu.:55.01
 Median :63.28
 Mean   :63.37
 3rd Qu.:71.96
 Max.   :92.19
```

**Gambar 2.** Rangkuman nilai dari atribut 2017, 2019 dan 2019

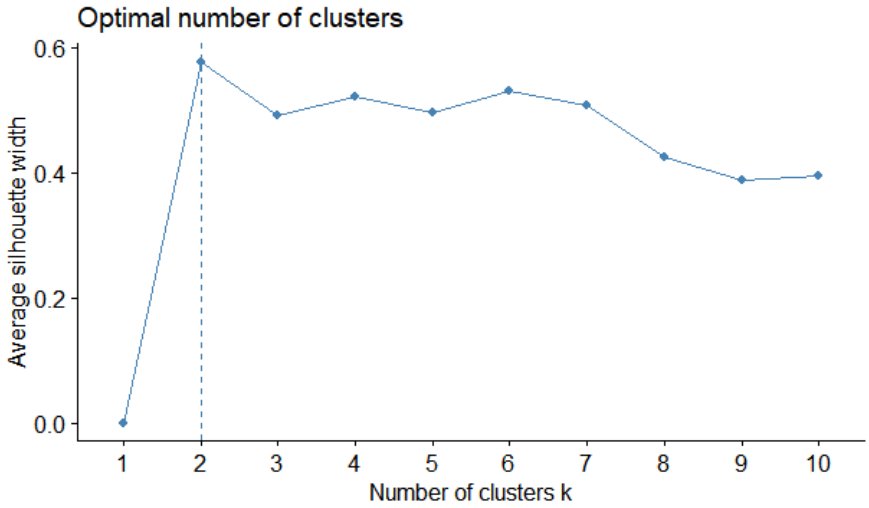
Berdasarkan rangkuman ketiga atribut tersebut, maka langkah yang dipilih selanjutnya adalah dengan melakukan transformasi data menjadi skala. Hal ini dilakukan untuk mencegah terjadinya penyebaran data menjadi tidak normal.

**Tabel 2.** Transformasi data yang digunakan

| Provinsi | 2019 | 2018 | 2017 |
|---|---|---|---|
| Yogyakarta | 1,4309 | 1,7667 | 2,0308 |
| Jakarta | 1,2490 | 1,5188 | 1,8716 |
| Riau | 1,0773 | 1,0887 | 1,4875 |
| Bali | 0,9572 | 1,1773 | 1,2740 |
| Jawa Tengah | 0,9230 | 0,8625 | 0,8554 |
| Kalimantan Timur | 0,8459 | 0,8766 | 0,9597 |
| Jawa Barat | 0,7673 | 0,8149 | 0,8040 |
| Jawa Timur | 0,7076 | 0,6526 | 0,6835 |
| Banten | 0,6996 | 0,7285 | 0,5686 |
| Kalimantan Utara | 0,6582 | 0,6526 | 0,6173 |
| Kalimantan Selatan | 0,6487 | 0,5261 | 0,3967 |
| Bangka Belitung | 0,4544 | 0,3959 | 0,0725 |
| Sulawesi Selatan | 0,3678 | 0,3706 | 0,2487 |
| Riau | 0,2754 | 0,2991 | 0,2339 |
| Lampung | 0,2449 | 0,1123 | -0,3284 |
| Jambi | 0,2427 | 0,1682 | -0,0494 |
| Sulawesi Utara | 0,2165 | 0,2195 | 0,5221 |
| Kalimantan Tengah | 0,1976 | -0,2426 | -0,1537 |
| Sumatera Barat | 0,1743 | 0,3371 | 0,3107 |
| Gorontalo | 0,0986 | 0,1845 | -0,1191 |
| Sulawesi Tenggara | 0,0921 | -0,0804 | -0,2580 |
| Nusa Tenggara Barat | -0,0513 | -0,4101 | -0,6131 |
| Sumatera Utara | -0,0702 | -0,0372 | 0,0373 |
| Sumatera Selatan | -0,0716 | 0,0461 | -0,1283 |
| Bengkulu | -0,0891 | -0,1623 | -0,2326 |
| Kalimantan Barat | -0,5708 | -0,6467 | -0,8139 |
| Aceh | -0,6719 | -0,6534 | -0,8168 |
| Sulawesi Tengah | -0.6865 | -0,7658 | -0,5187 |
| Sulawesi Barat | -0,7411 | -1,0932 | -0,9852 |
| Papua Barat | -0,9346 | -0,7777 | -0,9394 |
| Maluku | -1,3152 | -1,2986 | -0,9704 |
| Maluku Utara | -1,8456 | -1,6685 | -1,9858 |
| Nusa Tenggara Timur | -1,9802 | -2,1150 | -1,7399 |
| Papua | -3,3008 | -2,8472 | -2,3220 |

Lalu kemudian, menentukan nilai k=2. Proses pemilihan nilai k dilakukan dengan melihat dikoordinat berapa nilai k berada di puncak tertingginya berdasarkan perolehan nilai rata-rata *silhouette coefficient*. Setelah itu baru benar-benar mulai menerapkan algoritma *K-Means*.

Jurnal Sains Komputer & Informatika (J-SAKTI)
Volume 5 Nomor 2, September 2021, pp. 759-767
ISSN: 2548-9771/EISSN: 2549-7200
https://tunasbangsa.ac.id/ejurnal/index.php/jsakti




**Gambar 3.** Penentuan nilai k yang paling optimal

### 3.2. *Output*

Hasil yang didapatkan dari proses implementasi algoritma *K-Means* dengan jumlah k=2 adalah terdapat 9 *records* yang merupakan kluster 1 dan terdapat 25 *records* yang merupakan kluster 2.

```
K-means clustering with 2 clusters of sizes 9, 25
```

**Gambar 4.** Hasil implementasi algoritma *K-Means*

Berikut ini adalah rincian dari setiap klusternya, yaitu :

a) Kluster 1 terdiri dari provinsi : Yogyakarta, Jakarta, Riau, Bali, Jawa Tengah, Jawa Barat, Banten, Jawa Timur, Kalimantan Timur, Kalimantan Utara, Kalimantan Selatan, Bangka Belitung, Riau, Lampung, Jambi, Sumatera Utara, Sumatera Selatan, Sulawesi Utara, Kalimantan Tengah, Sumatera Barat, Gorontalo, Sulawesi Selatan, Sulawesi Tenggara dan Nusa Tenggara Barat.
b) Kluster 2 terdiri dari provinsi : Aceh, Bengkulu, Kalimantan Barat, Sulawesi Tengah, Sulawesi Barat, Nusa Tenggara Timur, Maluku, Maluku Utara, Papua Barat dan Papua.

```
Clustering vector:
 [1] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1
```

**Gambar 5.** Anggota setiap *cluster*

### 3.2. *Evaluation*

Berdasarkan klasterisasi yang dilakukan, diketahui nilai *within cluster sum of squares* yang didapatkan dari proses perhitungan, yaitu 14,22093 untuk *cluster* 1 dan 22,98090 untuk *cluster* 2 dengan nilai *between_SS* atau *between sum of squares* dibagi nilai *total_SS* sebesar 62,4%. Dengan

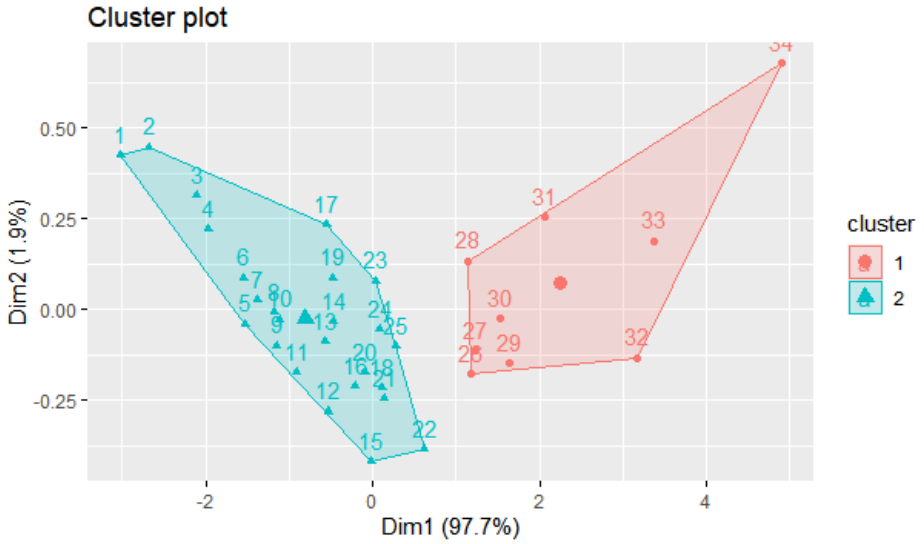
Jurnal Sains Komputer & Informatika (J-SAKTI)
Volume 5 Nomor 2, September 2021, pp. 759-767
ISSN: 2548-9771/EISSN: 2549-7200
https://tunasbangsa.ac.id/ejurnal/index.php/jsakti

demikian, maka dapat dinyatakan bahwa hasil klasterisasi k=2 memiliki hasil yang paling baik diantara nilai k yang lainnya.

```
Within cluster sum of squares by cluster:
[1] 14.22093 22.98090
 (between_SS / total_SS =  62.4 %)
```

**Gambar 6.** Hasil klasterisasi dengan k=2

Adapun untuk visualisasi penyebaran *cluster* itu sendiri bila ditampilkan dalam bentuk gambar akan seperti diagram yang ada pada gambar 6. Setiap *cluster* terdiri dari kumpulan titik-titik yang memiliki nilai kemiripan yang sama antar anggotanya dan berbeda dengan jenis klaster yang lainnya, dalam hal ini *cluster* 1 ditandai dengan area yang berwarna merah sedangkan *cluster* 2 ditandai dengan area yang berwarna biru.




**Gambar 7**. Visualisasi diagram penyebaran *cluster*

## 4. SIMPULAN

Dengan demikian, berdasarkan dari hasil klasterisasi yang dilakukan, terdapat dua jenis kelompok yang berbeda, yaitu :

a) Kluster 1 terdiri dari provinsi : Yogyakarta, Jakarta, Riau, Bali, Jawa Tengah, Jawa Barat, Banten, Jawa Timur, Kalimantan Timur, Kalimantan Utara, Kalimantan Selatan, Bangka Belitung, Riau, Lampung, Jambi, Sumatera Utara, Sumatera Selatan, Sulawesi Utara, Kalimantan Tengah, Sumatera Barat, Gorontalo, Sulawesi Selatan, Sulawesi Tenggara dan Nusa Tenggara Barat.
b) Kluster 2 terdiri dari provinsi : Aceh, Bengkulu, Kalimantan Barat, Sulawesi Tengah, Sulawesi Barat, Nusa Tenggara Timur, Maluku, Maluku Utara, Papua Barat dan Papua.

Oleh karena masih terdapat perbedaan hasil klasterisasi, yaitu sembilan dibanding dua puluh lima, maka diperlukanyalah pemerataan penggunaan TIK pada usia muda yang melek TIK di seluruh provinsi yang ada di

Indonesia untuk menjadi sama rata dan berangsur lebih baik. Dan untuk penelitian selanjutnya, diharapkan adanya penggunaan berbagai metode-metode yang lainnya yang dapat dijadikan sebagai bahan perbandingan dan perbaikan untuk penelitian ini atau pun yang serupa.

## DAFTAR PUSTAKA

[1] Aditya, A., Jovian, I., & Sari, B. N., "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019". Jurnal Media Informatika Budidarma, pp 51-58, 2020.

[2] APJII., "Laporan Survei : Penetrasi & Profil Perilaku Pengguna Internet Indonesia 2018". Jakarta: Polling Indonesia, 2019.

[3] Isyarah, F., Hasan, M., & Wiza, F., "Clustering Daerah Miskin di Provinsi Riau Menggunakan Metode K-Means". Prosiding-Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMASTER), pp 1-12, 2020.

[4] Oktavia, R., Hardinata, J. T., & Irawan., "Penerapan Metode Algoritma K-means dalam Pengelompokan Angka Harapan Hidup Saat Lahir Menurut Provinsi". KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen), pp 164-161, 2020.

[5] Priyatman, H., Sajid, F., & Haldivany, D., "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa". JEPIN (Jurnal Edukasi dan Penelitian Informatika), Vol. 5 No. 1, pp 62-66, 2019.

[6] Saifullah, & Hidayati, N., "Pengelompokan Persentase Buta Huruf Umur 15-44 Menurut Provinsi Menggunakan Algoritma K-Means". Kumpulan JurnaL Ilmu Komputer (KLIK), pp 230-240, 2020.

[7] Saragih, A. T., Sembiring, A. S., & Sayuthi, M., "Penerapan Metode Clustering K-Means untuk Proses Seleksi Calon Peserta Lomba MTQ". Jurnal Pelita Informatika, Volume 17, Nomor 2, pp 117-122, 2018.

[8] Sujanah, A., Sukroni, A., Satika, A., & Krismawati., "Statistik Ketenagakerjaan Usia Muda di Indonesia". Jakarta: Sub Direktorat Statistik Ketenagakerjaan, 2016.

[9] Triyansyah, D., & Fitrianah, D., "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing (Studi Kasus: Hoyweapstore)". IncomTech, Jurnal Telekomunikasi dan Komputer, Vol.8, No.3, pp 164-182, 2018.

[10] Vhallah, I., Sumijan, & Santony, J., "Pengelompokan Mahasiswa Potensial Drop Out menggunakan Metode Clustering K-Means". JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi), Vol. 2 No. 2, pp 572-577, 2018.

[11] Wahyudi, H. S., & Sukmasari, M. P., "Teknologi dan Kehidupan Masyarakat". Jurnal Analisa Sosiologi, pp 13-24, 2014.

[12] Yunita, Efendi, R., & Rini, D. P., "Metode K-Means & SAW dalam Seleksi Penerima Dana Zakat pada Badan Amil Zakat". TEKNOMATIKA, Vol.09, No.02, pp 163-174, 2019.