

Algoritma *K-Means* Dalam Pengelompokan Kabupaten/Kota Berdasarkan Indeks Pembangunan Manusia Di Sumatera Utara

Kristin Daya Rohani Sianipar^{1*}, Indra Gunawan²

^{1*2} *Teknik Informatika, STIKOM Tunas Bangsa Pematangsiantar*

^{1*}kristinsianipar7@gmail.com (penulis korespondensi)

²indra@amiktunasbangsa.ac.id

Abstrak— Indeks Pembangunan Manusia (IPM) merupakan salah satu parameter yang digunakan untuk memperkirakan tingkat quality secara fisik maupun non fisik masyarakat di suatu daerah. Dimensi dasar dalam membentuk Indeks Pembangunan Manusia (IPM) adalah umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak. Indeks Pembangunan Manusia (IPM) sangat mempengaruhi dalam pertumbuhan ekonomi, sehingga dengan meningkatnya Indeks Pembangunan Manusia (IPM) mampu mendorong perekonomian daerah, terutama dalam bidang industri seperti pada wilayah Sumatera Utara. Dalam melakukan penelitian ini, algoritma yang digunakan yaitu algoritma *k-means* dengan mengelompokkan menjadi tiga *cluster* yaitu tinggi, sedang, dan rendah. Adapun data yang digunakan adalah data kabupaten/kota berdasarkan Indeks Pembangunan Manusia (IPM) di Sumatera Utara tahun 2014-2019 sebanyak 33 data dan data penelitian ini diambil dari BPS (Badan Pusat Statistik). Hasil yang diperoleh dalam penelitian ini adalah *cluster* 1 terdapat 18 data, *cluster* 2 terdapat 11 data, *cluster* 3 terdapat 4 data. Tujuan dari penelitian ini adalah untuk mengetahui kabupaten/kota yang memiliki Indeks Pembangunan Manusia yang tergolong tinggi, sedang, dan rendah.

Kata kunci— IPM, *Data Mining*, *K-Means*, Pengelompokan, Masyarakat.

Abstract— The Human Development Index (HDI) is one of the parameters used to estimate the level of physical and non-physical quality of the community in an area. The basic dimensions in forming the Human Development Index (HDI) are a long and healthy life, knowledge, and a decent standard of living. The Human Development Index (HDI) greatly influences economic growth, so that the increase in the Human Development Index (HDI) is able to encourage the regional economy, especially in the industrial sector such as in the North Sumatra region. In conducting this research, the algorithm used is the *k-means* algorithm by grouping it into three clusters, namely high, medium, and low. The data used is district/city data based on the Human Development Index (HDI) in North Sumatra in 2014-2019 as many as 33 data and this research data was taken from BPS (Central Statistics Agency). The results obtained in this study are cluster 1 there are 18 data, cluster 2 there are 11 data, cluster 3 there are 4 data. The purpose of this study was to determine the regencies/cities that have a high, medium, and low Human Development Index.

Keywords— IPM, *Data Mining*, *K-Means*, *Cluster*, *Society*

I. PENDAHULUAN

Indeks pembangunan manusia adalah salah satu alat ukur pencapaian kualitas hidup satu wilayah bahkan negara. Indeks Pembangunan Manusia (IPM) merupakan salah satu indikator yang dijadikan untuk mengukur taraf kualitas fisik maupun non fisik penduduk di suatu wilayah.. Indeks Pembangunan Manusia dilakukan dengan tujuan manusia dijadikan sebagai tujuan akhir bukan menjadi alat Indeks Pembangunan Manusia. Adanya pembangunan manusia sendiri diharapkan masyarakat dapat menikmati umur yang panjang dan sehat, dan menjalankan kehidupan yang produktif.

Indeks Pembangunan Manusia (IPM) merupakan indikator komposit tunggal yang walaupun tidak dapat mengukur semua dimensi dari pembangunan manusia, tetapi mengukur tiga dimensi pokok pembangunan manusia yang dinilai mampu mencerminkan kemampuan dasar (basic capabilities) penduduk. Ketiga kemampuan dasar itu adalah umur panjang dan sehat, berpengetahuan dan berketerampilan, serta akses terhadap sumber daya yang dibutuhkan untuk mencapai standar hidup layak[1]. Indeks Pembangunan Manusia (IPM) sangat berpengaruh terhadap pertumbuhan ekonomi. Indeks Pembangunan manusia yang meningkat dapat menggerakkan perekonomian daerah terutama pada sektor industri yang ada di Indonesia terlebih di Sumatera Utara.

Data yaitu kumpulan fakta yang terekam atau sebuah entitas yang tidak memiliki arti dan selama ini terabaikan. *Mining* yaitu proses pembangunan. Sehingga *Data Mining*

dapat diartikan sebagai proses penambangan data yang menghasilkan sebuah output (keluaran) berupa pengetahuan. Menurut Pramudiono dalam [2] *Data Mining* adalah analisis otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya. Menurut Larose dalam [2] *Data Mining* merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. *Data Mining* merupakan proses pencarian pola dan relasi-relasi yang tersembunyi dalam sejumlah data yang besar dengan tujuan untuk melakukan klasifikasi, estimasi, prediksi, *association rule*, *clustering* dan visualisasi.

Penelitian ini dilakukan dengan menggunakan Algoritma *K-Means*. Algoritma *K-Means* merupakan salah satu algoritma yang terdapat dalam *data mining* yang cara pengerjaannya yaitu dengan pengelompokan data. Data pada penelitian ini diperoleh dari Badan Pusat Statistika (BPS). Dalam penelitian ini penulis menerapkan algoritma *K-Means* dengan tujuan untuk mengetahui kabupaten/kotaa yang memiliki Indeks Pembangunan Manusia yang tergolong tinggi, sedang dan rendah.

II. METODOLOGI PENELITIAN

Penelitian ini dilakukan menggunakan algoritma *K-means* metode dengan *clustering*. *K-Means* yaitu salah satu dari metode pengelompokan data nonhierarki (sekatan) yang dapat mempartisi data kedalam bentuk dua kelompok ataupun lebih[3]. *Clustering* merupakan aktivitas (*task*) yang bertujuan mengelompokkan data yang memiliki kemiripan antara satu data dengan data lainnya ke dalam klaster atau kelompok sehingga data dalam satu klaster memiliki tingkat kemiripan (*similarity*) yang maksimum dan data antar klaster memiliki kemiripan yang minimum[4]. *K-Means Clustering* merupakan salah satu metode data *clustering non* hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok[5].

Langkah-langkah dalam menyelesaikan perhitungan pada Algoritma *K-Means*, yaitu: [6]

1. Menentukan *k* sebagai jumlah *cluster* yang ingin dibuat. Menetapkan pusat *cluster*.
2. Menghitung jarak setiap data ke pusat *cluster* menggunakan persamaan Euclidean.

$$d_{ik} = \sqrt{\sum_j^m (Cij - Ckj)^2} \tag{1}$$

3. Kelompokkan data ke dalam *cluster* yang dengan jarak yang paling pendek menggunakan persamaan.

$$\text{Min} \sum_{k=1}^k d_{ik} = \sqrt{\sum_j^m (Cij - Ckj)^2} \tag{2}$$

4. Hitung pusat *cluster* yang baru menggunakan persamaan

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \tag{3}$$

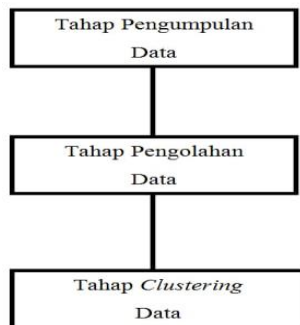
Dimana:

x_{ij} E kluster ke – k

P = banyaknya anggota *cluster* ke k

Ulangi langkah 2 sampai dengan 4 hingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

Dalam Algoritma *K-Means* memiliki tahapan-tahapan sebelum melakukan proses perhitungan yang di bentuk menjadi sebuah skema atau alur . Berikut adalah skema dari algoritma *K-Means* :



Gambar 1. Skema Algoritma *K-Means*

Pada penerapan algoritma *K-Means* dalam pengelompokan Kabupaten/Kota Berdasarkan Indeks Pembangunan Manusia di Sumatera Utara mendapatkan data dari Badan Pusat Statistika (BPS). Data akan diolah dengan melakukan *clustering* Kabupaten/Kota Berdasarkan Indeks

Pembangunan Manusia di Sumatera Utara dalam 3 kluster. Adapun data yang diperoleh yaitu sebagai berikut :

TABEL I
DATA KABUPATEN/KOTA BERDASARKAN IPM DI SUMATERA UTARA TAHUN 2014-2019

| NO | Kabupaten/ Kota | IPM | | | | | |
|----|----------------------|-------|-------|-------|-------|-------|-------|
| | | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
| 1 | Nias | 57,98 | 58,85 | 59,75 | 60,21 | 60,82 | 61,65 |
| 2 | Mandailing Natal | 63,42 | 63,99 | 64,55 | 65,13 | 65,83 | 66,52 |
| 3 | Tapanuli Selatan | 67,22 | 67,63 | 68,04 | 68,69 | 69,1 | 69,75 |
| 4 | Tapanuli Tengah | 66,16 | 67,06 | 67,27 | 67,96 | 68,27 | 68,86 |
| 5 | Tapanuli Utara | 70,7 | 71,32 | 71,96 | 72,38 | 72,91 | 73,33 |
| 6 | Toba | 72,79 | 73,4 | 73,61 | 73,87 | 74,48 | 74,92 |
| 7 | Labuhan batu | 70,06 | 70,23 | 70,5 | 71 | 71,39 | 71,94 |
| 8 | Asahan | 67,51 | 68,4 | 68,71 | 69,1 | 69,49 | 69,92 |
| 9 | Simalungun | 70,89 | 71,24 | 71,48 | 71,83 | 72,49 | 72,98 |
| 10 | Dairi | 67,91 | 69 | 69,61 | 70,36 | 70,89 | 71,42 |
| 11 | Karo | 71,84 | 72,69 | 73,29 | 73,53 | 73,91 | 74,25 |
| 12 | Deli Serdang | 71,98 | 72,79 | 73,51 | 73,94 | 74,92 | 75,43 |
| 13 | Langkat | 68 | 68,53 | 69,13 | 69,82 | 70,27 | 70,76 |
| 14 | Nias Selatan | 57,78 | 58,74 | 59,14 | 59,85 | 60,75 | 61,59 |
| 15 | Humbang Hasundutan | 65,59 | 66,03 | 66,56 | 67,3 | 67,96 | 68,83 |
| 16 | Pakpak Bharat | 65,06 | 65,53 | 65,81 | 66,25 | 66,63 | 67,47 |
| 17 | Samosir | 67,8 | 68,43 | 68,82 | 69,43 | 69,99 | 70,55 |
| 18 | Serdang Bedagai | 67,78 | 68,01 | 68,77 | 69,16 | 69,69 | 70,21 |
| 19 | Batu Bara | 65,5 | 66,02 | 66,69 | 67,2 | 67,67 | 68,35 |
| 20 | Padang Lawas Utara | 66,5 | 67,35 | 68,05 | 68,34 | 68,77 | 69,29 |
| 21 | Padang Lawas | 65,5 | 65,99 | 66,23 | 66,82 | 67,59 | 68,16 |
| 22 | Labuhan batu Selatan | 68,59 | 69,67 | 70,28 | 70,48 | 70,98 | 71,39 |
| 23 | Labuhan batu Utara | 69,15 | 69,69 | 70,26 | 70,79 | 71,08 | 71,43 |
| 24 | Nias Utara | 59,18 | 59,88 | 60,23 | 60,57 | 61,08 | 61,98 |
| 25 | Nias Barat | 57,54 | 58,25 | 59,03 | 59,56 | 60,42 | 61,14 |
| 26 | Sibolga | 71,01 | 71,64 | 72 | 72,28 | 72,65 | 73,41 |
| 27 | Tanjung balai | 66,05 | 66,74 | 67,09 | 67,41 | 68 | 68,51 |
| 28 | Pematang siantar | 75,83 | 76,34 | 76,9 | 77,54 | 77,88 | 78,57 |
| 29 | Tebing Tinggi | 72,13 | 72,81 | 73,58 | 73,9 | 74,5 | 75,08 |
| 30 | Medan | 78,26 | 78,87 | 79,34 | 79,98 | 80,65 | 80,97 |
| 31 | Binjai | 72,55 | 73,81 | 74,11 | 74,65 | 75,21 | 75,89 |
| 32 | Padang Sidempuan | 71,88 | 72,8 | 73,42 | 73,81 | 74,38 | 75,06 |
| 33 | Gunung Sitoli | 65,91 | 66,41 | 66,85 | 67,68 | 68,33 | 69,3 |

Data yang didapatkan akan diolah terlebih dahulu dan setelah diolah akan dilakukan *clustering*. Dalam tahap

pengumpulan data, data yang diperoleh akan dijumlahkan setiap kriterianya, maka pada tahap pengolahan data akan didapatkan kalkulasi nilai yang nantinya akan dioperasikan selanjutnya atau tahap *clustering*.

TABEL II
DATA AKUMULASI KABUPATEN/KOTA BERDASARKAN IPM DI SUMATERA UTARA TAHUN 2014-2019

| NO | Kabupaten/ Kota | IPM | | | | | | Rata-rata |
|----|---------------------|-------|-------|-------|-------|-------|-------|-------------|
| | | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | |
| 1 | Nias | 57,98 | 58,85 | 59,75 | 60,21 | 60,82 | 61,65 | 59,87666667 |
| 2 | Mandailing Natal | 63,42 | 63,99 | 64,55 | 65,13 | 65,83 | 66,52 | 64,90666667 |
| 3 | Tapanuli Selatan | 67,22 | 67,63 | 68,04 | 68,69 | 69,1 | 69,75 | 68,405 |
| 4 | Tapanuli Tengah | 66,16 | 67,06 | 67,27 | 67,96 | 68,27 | 68,86 | 67,59666667 |
| 5 | Tapanuli Utara | 70,7 | 71,32 | 71,96 | 72,38 | 72,91 | 73,33 | 72,1 |
| 6 | Toba | 72,79 | 73,4 | 73,61 | 73,87 | 74,48 | 74,92 | 73,845 |
| 7 | Labuhan batu | 70,06 | 70,23 | 70,5 | 71 | 71,39 | 71,94 | 70,85333333 |
| 8 | Asahan | 67,51 | 68,4 | 68,71 | 69,1 | 69,49 | 69,92 | 68,855 |
| 9 | Simalungun | 70,89 | 71,24 | 71,48 | 71,83 | 72,49 | 72,98 | 71,81833333 |
| 10 | Dairi | 67,91 | 69 | 69,61 | 70,36 | 70,89 | 71,42 | 69,865 |
| 11 | Karo | 71,84 | 72,69 | 73,29 | 73,53 | 73,91 | 74,25 | 73,25166667 |
| 12 | Deli Serdang | 71,98 | 72,79 | 73,51 | 73,94 | 74,92 | 75,43 | 73,76166667 |
| 13 | Langkat | 68 | 68,53 | 69,13 | 69,82 | 70,27 | 70,76 | 69,41833333 |
| 14 | Nias Selatan | 57,78 | 58,74 | 59,14 | 59,85 | 60,75 | 61,59 | 59,64166667 |
| 15 | Humbang Hasundutan | 65,59 | 66,03 | 66,56 | 67,3 | 67,96 | 68,83 | 67,045 |
| 16 | Pakpak Bharat | 65,06 | 65,53 | 65,81 | 66,25 | 66,63 | 67,47 | 66,125 |
| 17 | Samosir | 67,8 | 68,43 | 68,82 | 69,43 | 69,99 | 70,55 | 69,17 |
| 18 | Serdang Bedagai | 67,78 | 68,01 | 68,77 | 69,16 | 69,69 | 70,21 | 68,93666667 |
| 19 | Batu Bara | 65,5 | 66,02 | 66,69 | 67,2 | 67,67 | 68,35 | 66,905 |
| 20 | Padang Lawas Utara | 66,5 | 67,35 | 68,05 | 68,34 | 68,77 | 69,29 | 68,05 |
| 21 | Padang Lawas | 65,5 | 65,99 | 66,23 | 66,82 | 67,59 | 68,16 | 66,715 |
| 22 | Labuhanbatu Selatan | 68,59 | 69,67 | 70,28 | 70,48 | 70,98 | 71,39 | 70,23166667 |
| 23 | Labuhanbatu Utara | 69,15 | 69,69 | 70,26 | 70,79 | 71,08 | 71,43 | 70,4 |
| 24 | Nias Utara | 59,18 | 59,88 | 60,23 | 60,57 | 61,08 | 61,98 | 60,48666667 |
| 25 | Nias Barat | 57,54 | 58,25 | 59,03 | 59,56 | 60,42 | 61,14 | 59,32333333 |
| 26 | Sibolga | 71,01 | 71,64 | 72 | 72,28 | 72,65 | 73,41 | 72,165 |
| 27 | Tanjung balai | 66,05 | 66,74 | 67,09 | 67,41 | 68 | 68,51 | 67,3 |
| 28 | Pematang siantar | 75,83 | 76,34 | 76,9 | 77,54 | 77,88 | 78,57 | 77,17666667 |
| 29 | Tebing Tinggi | 72,13 | 72,81 | 73,58 | 73,9 | 74,5 | 75,08 | 73,66666667 |
| 30 | Medan | 78,26 | 78,87 | 79,34 | 79,98 | 80,65 | 80,97 | 79,67833333 |
| 31 | Binjai | 72,55 | 73,81 | 74,11 | 74,65 | 75,21 | 75,89 | 74,37 |
| 32 | Padang Sidempuan | 71,88 | 72,8 | 73,42 | 73,81 | 74,38 | 75,06 | 73,55833333 |
| 33 | Gunung Sitoli | 65,91 | 66,41 | 66,85 | 67,68 | 68,33 | 69,3 | 67,41333333 |

Selanjutnya, pada proses *clustering* untuk pengelompokan data Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia Di Sumatera Utara dengan membuat 3 *cluster* yaitu *cluster* tinggi, sedang, dan rendah. Proses pengelompokan data dilakukan menggunakan rumus *Euclidean Distance*.

berdasarkan Indeks Pembangunan Manusia. Dalam proses *clustering*, data yang diperoleh akan dihitung terlebih dahulu menggunakan jumlah Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia yang digunakan. Berikut adalah hasil perhitungan jumlah Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia :

III. HASIL DAN PEMBAHASAN

Proses perhitungan untuk mengelompokkan Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia di Sumatera Utara menggunakan algoritma *K-Means clustering* dengan menghitung jumlah Kabupaten/Kota

TABEL III
DATAJUMLAH RATA-RATA KABUPATEN/KOTA BERDASARKAN
IPM DI SUMATERA UTARA TAHUN 2014-2019

| NO | Kabupaten/ Kota | rata-rata |
|----|-----------------------|-------------|
| 1 | N i a s | 59,87666667 |
| 2 | Mandailing Natal | 64,90666667 |
| 3 | Tapanuli Selatan | 68,405 |
| 4 | Tapanuli Tengah | 67,59666667 |
| 5 | Tapanuli Utara | 72,1 |
| 6 | Toba | 73,845 |
| 7 | Labuhanbatu | 70,85333333 |
| 8 | A s a h a n | 68,855 |
| 9 | Simalungun | 71,81833333 |
| 10 | D a i r i | 69,865 |
| 11 | K a r o | 73,25166667 |
| 12 | Deli Serdang | 73,76166667 |
| 13 | L a n g k a t | 69,41833333 |
| 14 | Nias Selatan | 59,64166667 |
| 15 | Humbang Hasundutan | 67,045 |
| 16 | Pakpak Bharat | 66,125 |
| 17 | Samosir | 69,17 |
| 18 | Serdang Bedagai | 68,93666667 |
| 19 | Batu Bara | 66,905 |
| 20 | Padang Lawas Utara | 68,05 |

| | | |
|----|---------------------|-------------|
| 21 | Padang Lawas | 66,715 |
| 22 | Labuhanbatu Selatan | 70,23166667 |
| 23 | Labuhanbatu Utara | 70,4 |
| 24 | Nias Utara | 60,48666667 |
| 25 | Nias Barat | 59,32333333 |
| 26 | S i b o l g a | 72,165 |
| 27 | Tanjungbalai | 67,3 |
| 28 | Pematangsiantar | 77,17666667 |
| 29 | Tebing Tinggi | 73,66666667 |
| 30 | M e d a n | 79,67833333 |
| 31 | B i n j a i | 74,37 |
| 32 | Padangsidempuan | 73,55833333 |
| 33 | Gunungsitoli | 67,41333333 |

Selanjutnya, dilakukan perhitungan pusat *cluster* awal dan hasil perhitungan iterasi 1. Adapun centroid awal yang diterapkan dapat dilihat pada Tabel 4 dan hasil perhitungan iterasi-1 dapat dilihat pada Tabel 5 sebagai berikut :

TABEL IV
CENTEROID DATA AWAL

| <i>Centroid</i> | |
|-----------------|-------------|
| Tinggi | 70,85333333 |
| Sedang | 66,715 |
| Rendah | 59,32333333 |

TABEL V
HASIL PERHITUNGAN ITERASI 1

| NO | Kabupaten/ Kota | rata-rata | C1 | C2 | C3 | Cluster |
|----|--------------------|-------------|-------------|-------------|-------------|---------|
| 1 | N i a s | 59,87666667 | 10,97666667 | 6,838333333 | 0,553333333 | C3 |
| 2 | Mandailing Natal | 64,90666667 | 5,946666667 | 1,808333333 | 5,583333333 | C2 |
| 3 | Tapanuli Selatan | 68,405 | 2,448333333 | 1,69 | 9,081666667 | C2 |
| 4 | Tapanuli Tengah | 67,59666667 | 3,256666667 | 0,881666667 | 8,273333333 | C2 |
| 5 | Tapanuli Utara | 72,1 | 1,246666667 | 5,385 | 12,77666667 | C1 |
| 6 | Toba | 73,845 | 2,991666667 | 7,13 | 14,52166667 | C1 |
| 7 | Labuhanbatu | 70,85333333 | 0 | 4,138333333 | 11,53 | C1 |
| 8 | A s a h a n | 68,855 | 1,998333333 | 2,14 | 9,531666667 | C1 |
| 9 | Simalungun | 71,81833333 | 0,965 | 5,103333333 | 12,495 | C1 |
| 10 | D a i r i | 69,865 | 0,988333333 | 3,15 | 10,54166667 | C1 |
| 11 | K a r o | 73,25166667 | 2,398333333 | 6,536666667 | 13,92833333 | C1 |
| 12 | Deli Serdang | 73,76166667 | 2,908333333 | 7,046666667 | 14,43833333 | C1 |
| 13 | L a n g k a t | 69,41833333 | 1,435 | 2,703333333 | 10,095 | C1 |
| 14 | Nias Selatan | 59,64166667 | 11,21166667 | 7,073333333 | 0,318333333 | C3 |
| 15 | Humbang Hasundutan | 67,045 | 3,808333333 | 0,33 | 7,721666667 | C2 |
| 16 | Pakpak Bharat | 66,125 | 4,728333333 | 0,59 | 6,801666667 | C2 |
| 17 | Samosir | 69,17 | 1,683333333 | 2,455 | 9,846666667 | C1 |
| 18 | Serdang Bedagai | 68,93666667 | 1,916666667 | 2,221666667 | 9,613333333 | C1 |
| 19 | Batu Bara | 66,905 | 3,948333333 | 0,19 | 7,581666667 | C2 |

| NO | Kabupaten/ Kota | rata-rata | C1 | C2 | C3 | Cluster |
|----|---------------------|-------------|-------------|-------------|-------------|---------|
| 20 | Padang Lawas Utara | 68,05 | 2,803333333 | 1,335 | 8,726666667 | C2 |
| 21 | Padang Lawas | 66,715 | 4,138333333 | 0 | 7,391666667 | C2 |
| 22 | Labuhanbatu Selatan | 70,23166667 | 0,621666667 | 3,516666667 | 10,90833333 | C1 |
| 23 | Labuhanbatu Utara | 70,4 | 0,453333333 | 3,685 | 11,07666667 | C1 |
| 24 | Nias Utara | 60,48666667 | 10,36666667 | 6,228333333 | 1,163333333 | C3 |
| 25 | Nias Barat | 59,32333333 | 11,53 | 7,391666667 | 0 | C3 |
| 26 | S i b o l g a | 72,165 | 1,311666667 | 5,45 | 12,84166667 | C1 |
| 27 | Tanjungbalai | 67,3 | 3,553333333 | 0,585 | 7,976666667 | C2 |
| 28 | Pematangsiantar | 77,17666667 | 6,323333333 | 10,46166667 | 17,85333333 | C1 |
| 29 | Tebing Tinggi | 73,66666667 | 2,813333333 | 6,951666667 | 14,34333333 | C1 |
| 30 | M e d a n | 79,67833333 | 8,825 | 12,96333333 | 20,355 | C1 |
| 31 | B i n j a i | 74,37 | 3,516666667 | 7,655 | 15,04666667 | C1 |
| 32 | Padang Sidempuan | 73,55833333 | 2,705 | 6,843333333 | 14,235 | C1 |
| 33 | Gunung Sitoli | 67,41333333 | 3,44 | 0,698333333 | 8,09 | C2 |

Dari hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat. Setelah dilakukan perhitungan didapatkan jarak terdekat dengan pusat data yaitu 19 data pada *cluster* 1, 10 data pada *cluster* 2, dan 4 data pada *cluster* 3 sehingga terbentuklah *centroid* baru.

Proses *K-Means* akan berhenti jika nilai pengelompokan data sama dengan nilai pengelompokan data sebelumnya. Pada penelitian ini, proses *K-Means* berhenti di iterasi ke-6, karena pada iterasi ke-5 tidak memiliki perpindahan *cluster* di iterasi ke-6. Berikut adalah hasil perhitungan iterasi ke-6.

TABEL VI
HASIL PERHITUNGAN ITERASI 6

| NO | Kabupaten/ Kota | rata-rata | C1 | C2 | C3 | Cluster |
|----|--------------------|-------------|-------------|-------------|-------------|---------|
| 1 | N i a s | 59,87666667 | 14,24984848 | 8,356203703 | 0,044583337 | C3 |
| 2 | Mandailing Natal | 64,90666667 | 9,219848483 | 3,326203703 | 5,074583337 | C2 |
| 3 | Tapanuli Selatan | 68,405 | 5,72151515 | 0,17212963 | 8,57291667 | C2 |
| 4 | Tapanuli Tengah | 67,59666667 | 6,529848483 | 0,636203703 | 7,764583337 | C2 |
| 5 | Tapanuli Utara | 72,1 | 2,02651515 | 3,86712963 | 12,26791667 | C1 |
| 6 | Toba | 73,845 | 0,28151515 | 5,61212963 | 14,01291667 | C1 |
| 7 | Labuhanbatu | 70,85333333 | 3,273181817 | 2,620462963 | 11,02125 | C2 |
| 8 | A s a h a n | 68,855 | 5,27151515 | 0,62212963 | 9,02291667 | C2 |
| 9 | Simalungun | 71,81833333 | 2,308181817 | 3,585462963 | 11,98625 | C1 |
| 10 | D a i r i | 69,865 | 4,26151515 | 1,63212963 | 10,03291667 | C2 |
| 11 | K a r o | 73,25166667 | 0,874848483 | 5,018796297 | 13,41958334 | C1 |
| 12 | Deli Serdang | 73,76166667 | 0,364848483 | 5,528796297 | 13,92958334 | C1 |
| 13 | L a n g k a t | 69,41833333 | 4,708181817 | 1,185462963 | 9,586250003 | C2 |
| 14 | Nias Selatan | 59,64166667 | 14,48484848 | 8,591203703 | 0,190416663 | C3 |
| 15 | Humbang Hasundutan | 67,045 | 7,08151515 | 1,18787037 | 7,21291667 | C2 |
| 16 | Pakpak Bharat | 66,125 | 8,00151515 | 2,10787037 | 6,29291667 | C2 |
| 17 | Samosir | 69,17 | 4,95651515 | 0,93712963 | 9,33791667 | C2 |
| 18 | Serdang Bedagai | 68,93666667 | 5,189848483 | 0,703796297 | 9,104583337 | C2 |

| NO | Kabupaten/ Kota | rata-rata | C1 | C2 | C3 | Cluster |
|----|---------------------|-------------|-------------|-------------|-------------|---------|
| 19 | Batu Bara | 66,905 | 7,22151515 | 1,32787037 | 7,07291667 | C2 |
| 20 | Padang Lawas Utara | 68,05 | 6,07651515 | 0,18287037 | 8,21791667 | C2 |
| 21 | Padang Lawas | 66,715 | 7,41151515 | 1,51787037 | 6,88291667 | C2 |
| 22 | Labuhanbatu Selatan | 70,23166667 | 3,894848483 | 1,998796297 | 10,39958334 | C2 |
| 23 | Labuhanbatu Utara | 70,4 | 3,72651515 | 2,16712963 | 10,56791667 | C2 |
| 24 | Nias Utara | 60,48666667 | 13,63984848 | 7,746203703 | 0,654583337 | C3 |
| 25 | Nias Barat | 59,32333333 | 14,80318182 | 8,909537037 | 0,508749997 | C3 |
| 26 | Sibolga | 72,165 | 1,96151515 | 3,93212963 | 12,33291667 | C1 |
| 27 | Tanjungbalai | 67,3 | 6,82651515 | 0,93287037 | 7,46791667 | C2 |
| 28 | Pematangsiantar | 77,17666667 | 3,050151517 | 8,943796297 | 17,34458334 | C1 |
| 29 | Tebing Tinggi | 73,66666667 | 0,459848483 | 5,433796297 | 13,83458334 | C1 |
| 30 | Medan | 79,67833333 | 5,551818183 | 11,44546296 | 19,84625 | C1 |
| 31 | Binjai | 74,37 | 0,24348485 | 6,13712963 | 14,53791667 | C1 |
| 32 | Padangsidempuan | 73,55833333 | 0,568181817 | 5,325462963 | 13,72625 | C1 |
| 33 | Gunungsitoli | 67,41333333 | 6,713181817 | 0,819537037 | 7,581250003 | C2 |

Dari tabel di atas dapat dilihat hasil dari perhitungan yaitu, *cluster* 1 terdapat 18 data, *cluster* 2 terdapat 11 data, dan *cluster* 3 terdapat 4 data.

IV. KESIMPULAN

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa dengan adanya analisis dalam menentukan pengelompokan data pada Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia, dapat memepermudah untuk mengetahui Indeks Pembangunan Manusia yang tergolong tinggi, sedang dan rendah. *Data mining* dengan metode *K-Means Clustering* dapat dijadikan solusi dari permasalahan seperti pengelompokan data pada Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia.

Dari hasil analisa, maka dipeoleh hasil pengelompokan Kabupaten/Kota berdasarkan Indeks Pembangunan Manusia yaitu *cluster* 1 terdapat 18 data, *cluster* 2 terdapat 11 data, dan *cluster* 3 terdapat 4 data.

REFERENSI

- [1] M. B. Setiawan and A. Hakim, "Indeks Pembangunan Manusia Manusia," *J. Econ.* 9(1), 18-26, vol. 9(1), pp. 18–26, 2008.
- [2] D. Nofriansyah and G. W. Nurcahyo, "Algoritma Data Mining Dan Pengujian," *Algoritma Data Mining dan Pengujian*. pp. 1–3, 2019.
- [3] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [4] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi K-Means Clustering Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.
- [5] M. Anggara, H. Sujiani, and N. Helfi, "Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokan Member Di Alvaro Fitness," *J. Sist. dan Teknol. Inf.*, vol. 1, no. 1, pp. 1–6, 2016.
- [6] K. D. R. Sianipar, S. Wanti Siahaan, M. Siregar, and P. R. Fikrul Ilmi Zer, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19," *J. Teknol. Inf.*, vol. 4, no. 1, pp. 101–105, 2020.