

Klasterisasi Sumber Penyebaran Virus Covid-19 dengan Menggunakan Metode *K-MEANS* Di Daerah Kota Cimahi dan Kab. Bandung Barat

Arief Rachman¹, dan M. Reza Hidayat²

Teknik Elektro, Fakultas Teknik, Universitas Jenderal Achmad Yani, Kota Cimahi, Indonesia

arief3178@gmail.com, mreza@lecture.unjani.ac.id

Abstrak

Virus Covid-19 berasal dari Kota Wuhan di China. Penularannya sangat cepat dan pesat ke berbagai belahan dunia termasuk di Indonesia. Penelitian ini membahas tentang klasterisasi sumber penyebaran virus covid-19 dengan metoda *k-means* menggunakan *software* WEKA untuk Kota Cimahi dan Kab Bandung Barat. Tujuan dari penelitian ini yaitu memperoleh informasi tentang 5 klaster sumber penyebaran virus Covid-19, yaitu klaster 0 (kantor), klaster 1 (tempat makan), klaster 2 (rumah), klaster 3 (perbelanjaan), dan klaster 4 (transportasi umum) selama bulan Mei sampai Juli 2020 untuk Kota Cimahi dan Kab Bandung Barat. Hasil dapat disimpulkan pada klaster 0 tertinggi dari Kab. Bandung Barat dengan rata-rata 0,05%, untuk klaster 1 dampak covid-19 tertinggi dari Kab. Bandung Barat dengan rata-rata 0,55%, untuk klaster 2 dengan rata-rata 0,02%, untuk klaster 3 dampak covid-19 tertinggi dari Kab. Bandung Barat dengan rata-rata 0,04%, untuk klaster 4 dampak covid-19 tertinggi dari Kota Cimahi dengan rata-rata 0,79%.

Kata kunci: Covid-19, *Clustering*, *K-Means*, WEKA

Abstract

The Covid-19 virus originated from the city of Wuhan in China. Transmission is very fast and rapid to various parts of the world including Indonesia. This study discusses the clustering of sources of the spread of the covid-19 virus using the k-means method using software for Cimahi City and West Bandung Regency. The purpose of this study is to obtain information about 5 clusters of the source of the Covid-19 virus, namely cluster 0 (offices), cluster 1 (places to eat), cluster 2 (houses), cluster 3 (shopping), and cluster 4 (public transportation). during May to July 2020 for Cimahi City and West Bandung Regency. The results can be concluded that the highest cluster 0 from Kab. West Bandung with an average of 0,05%, for cluster 1 the highest impact of covid-19 was from Kab. West Bandung with an average of 0,55%, for cluster 2 with an average of 0,02%, for cluster 3 the highest impact of covid-19 was from Kab. West Bandung with an average of 0,04%, for cluster 4 the highest impact of covid-19 was from Cimahi City with an average of 0.79%.

Keywords: Covid-19, *Clustering*, *K-means*, WEKA

1. Pendahuluan

Indonesia telah mencatat 271.339 kasus positif infeksi virus corona sejak kasus pertama diumumkan Presiden Joko Widodo pada 26 September 2020 lalu. Dari jumlah itu, 199.403 dinyatakan sembuh dan 10.308 meninggal dunia. DKI Jakarta, Jawa Timur, dan Sulawesi Selatan menjadi penyumbang kasus terbesar Covid-19 (COVID-19, 2020). Namun, persebaran infeksi ini sebagian berasal dari klaster yang sama (Nurlaila Suci Rahayu Rais, 2018). Pertumbuhan klaster ini juga diakibatkan oleh faktor manusianya sendiri yang tidak patuh terhadap protokol kesehatan yang berlaku di Indonesia dengan memakai masker dan menjaga jarak 1-2 meter (Setiawan, Daryanto, 2018). Teknik yang dapat dilakukan untuk penanganan virus covid-19 salah satunya dengan mengetahui klaster sumber penularannya agar bisa mencegah penularan lebih banyak (Ahmad, 2019). Salah satu tahap dari data *mining* adalah bagian *descriptive* dengan tahapan klasterisasi (Zulfah, 2016). Salah satu keuntungan metoda *k-means* dari tahap klasterisasi yaitu mudah diimplementasikan dan dijalankan, waktu yang dibutuhkan untuk menjalankan pembelajaran ini cepat, mudah diadaptasi, dan umum digunakan (Setiawan, Daryanto, 2018).

Sesuai kebutuhannya, penggunaan algoritma *K-means Clustering* adalah suatu metode penganalisaan data atau metoda data *mining* yang melakukan proses pemodelan tanpa *supervised (unsupervised)* dan merupakan salah satu metoda yang melakukan pengelompokan data dengan sistem partisi. Adapun pada penelitian ini juga menggunakan metoda Algoritma *K-Means* dengan klasterisasi dampak covid-19 di Kota Cimahi dan Kab Bandung Barat pada bulan Mei sampai Juli 2020.

Info Makalah:

Dikirim : 08-27-20;
Revisi 1 : 09-17-20;
Revisi 2 : 09-28-20;
Revisi 3 : 11-11-20;
Diterima : 11-16-20.

Penulis Korespondensi:

Telp : +62-XXX-XXX
e-mail : mreza@lecture.unjani.ac.id

WEKA adalah sebuah paket *tools machine learning* praktis. "WEKA" merupakan singkatan dari *Waikato Environment for Knowledge Analysis*, yang dibuat di Universitas Waikato, New Zealand untuk penelitian, pendidikan dan berbagai aplikasi. WEKA mampu menyelesaikan masalah-masalah data *mining* di dunia nyata, khususnya klasifikasi yang mendasari pendekatan-pendekatan *machine learning*. Perangkat lunak ini ditulis dalam hirarki *class Java* dengan metode berorientasi objek dan dapat berjalan hampir di semua *platform*. WEKA mudah digunakan dan digunakan dan diterapkan pada

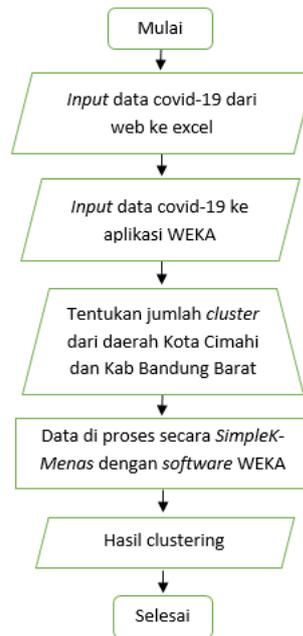
beberapa tingkatan yang berbeda. Tersedia implementasi algoritma-algoritma pembelajaran *state-of-the-art* yang dapat diterapkan pada data set dari *command line*. WEKA mengandung *tools* untuk *pre-processing* data, klasifikasi, regresi, *clustering*, aturan asosiasi, dan visualisasi. *User* dapat melakukan *preprocess* pada data, memasukkannya dalam sebuah skema pembelajaran, dan menganalisa *classifier* yang dihasilkan dan performasinya – semua itu tanpa menulis kode program sama sekali (WEKA, 2020). K-means merupakan salah satu metoda data *clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih klaster/kelompok. Metode ini mempartisi ke dalam klaster/kelompok sehingga data yang memiliki karakteristik yang sama (*High intra class similarity*) dikelompokkan ke dalam satu klaster yang sama dan yang memiliki karakteristik yang berbeda (*Low inter class similarity*) dikelompokkan pada kelompok yang lain (Narwati, 2016). Berikut adalah rumus matematis dasar untuk klasterisasi:

$$D(i, j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

Berdasarkan perkembangan riset, penelitian tentang hal ini yang telah dilakukan sebelumnya. Penelitian pertama dilakukan menggunakan data set *input*. Dalam mengimplementasikan algoritma ini data pengujian yang digunakan adalah data barang di supermarket Fadhilla Bengkulu. Hasil dari pemrosesan data menggunakan beberapa *software* data *mining* tersebut pada intinya sama. Menghasilkan kelompok data menjadi laris dan kurang laris (Benri Melpa Metisen, 2015). Penelitian kedua dilakukan untuk merealisasikan aplikasi yang dibangun akan memberikan rekomendasi pengevaluasian potensi pendapatan dan pemakaian air calon pelanggan baru PDAM disesuaikan dengan kriteria pemakaian air pelanggan baru yang telah ditentukan oleh PDAM. Pada penelitian ini menggunakan metode *K-Means* dengan beberapa tahapan yaitu tahap pertama menentukan jumlah klas (*k*) pada data set. Sistem menghasilkan *output* berupa kelompok potensi pendapatan dan pemakaian air berdasarkan tingkat kelayakannya (Gunawan Abdillah, 2016). Lalu penelitian ketiga dilakukan dengan Penerapan Data *mining* Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan *K-Means Clustering Method*. Sehingga diperoleh penilaian berdasarkan indeks ekspor buah-buahan dengan 2 negara klaster tingkat ekspor tinggi yakni India dan Pakistan, 3 negara klaster tingkat ekspor sedang yakni Singapura, Bangladesh dan Negara lainnya dan 6 negara klaster tingkat ekspor rendah yakni Hongkong, Tiongkok, Malaysia, Nepal, Vietnam dan Iran. Hasil yang dari penelitian dapat digunakan untuk mengetahui jumlah ekspor buah-buahan menurut negara tujuan (Windarto, 2017). Penelitian keempat dilakukan pada menerapkan algoritma *K-Means Clustering* untuk pengelompokan data Program SDP di PT. Bank Syariah. Berdasarkan hasil klaster dengan menerapkan beberapa kriteria dari daftar Program SDP menggunakan algoritma *K-Means* dapat diambil pengelompokan dengan rata-rata data program SDP yang dapat melakukan *Assessment center* lanjutan adalah yang lolos dan hasil klasifikasi program SDP yang hampir lolos harus memperbaiki administrasi seperti kedisiplinan dari bulan juni sampai bulan oktober agar dapat mengikuti *Assessment Center* lanjutan, sedangkan hasil klasifikasi daftar data program SDP yang tidak lolos harus memperbaiki data kedisiplinannya selama 1 Tahun. (In Parlina, 2018) Penelitian kelima yaitu penelitian ini membahas tentang Implementasi Rapidminer dengan Metode *K-Means* (Studi Kasus: Imunisasi Campak pada Balita berdasarkan Provinsi). Sehingga diperoleh penilaian berdasarkan pengelompokan Jumlah Persentase Balita yang Pernah Mendapat Imunisasi Campak berdasarkan Provinsi dengan klaster tinggi (C1) sebanyak 21 provinsi untuk klaster sedang (C2) sebanyak 12 provinsi dan untuk klaster rendah (C3) sebanyak 1 provinsi. Hasil dari penelitian ini dapat dijadikan masukan bagi pemerintah khususnya provinsi, agar provinsi yang masuk kedalam klaster tinggi mendapat perhatian lebih dan meningkatkan sosialisasi imunisasi campak terhadap balita. (Riyani Wulan Sari, 2018). Penelitian ke enam yaitu membahas tentang *clustering* data untuk rekomendasi penentuan jurusan perguruan tinggi menggunakan metode *k-means*. Dalam melakukan analisis penulis menggunakan alat bantu *tools* WEKA dan RapidMiner. Metode yang digunakan adalah metode *k-means clustering* dengan 24 atribut dan 5 klaster. Jumlah klaster pada perhitungan manual adalah, C1 terdapat 62 data, C2 terdapat 28 data, C3 terdapat 30 data, C4 terdapat 30 data, C5 terdapat 60 data. Jumlah klaster pada perhitungan RapidMiner adalah, C1 terdapat 35 data, C2 terdapat 55 data, C3 terdapat 58 data, C4 terdapat 35 data, C5 terdapat 27 data. Jumlah klaster pada perhitungan WEKA adalah, C1 terdapat 30 data, C2 terdapat 49 data, C3 terdapat 41 data, C4 terdapat 32 data, C5 terdapat 58 data (Pareza Alam Jusia, 2019). Penelitian ke tujuh yaitu membahas tentang *Bag of Word Clustering Using WEKA*. *SimpleKMeans* dan *XMeans* digunakan untuk mengolah dataset dan mengelompokkan berdasarkan jumlah klaster tetap yang digunakan, sedangkan *Farthest First* akan meletakkan semua pusat klaster pada titik terjauh dari pusat klaster yang sudah ada untuk mengelompokkan data. Dataset berasal dari UCI *machine learning* dengan menggunakan 3 koleksi data, yaitu *Enron Email*, *NIPS Proceedings*, dan *Daily Kos Blog entries*. Semua data merupakan teks tidak terstruktur yang direpresentasikan dalam *bags-of-word* (BoW). Performa dataset diuji dengan berbagai masukan parameter yang berbeda meliputi jumlah klaster hingga evaluasi *sum squared error* (SSE), serta iterasi selama proses pengolahan data. Pengujian dibagi dalam 3 yaitu membandingkan persentase klaster BoW yang terbentuk, kedua melakukan evaluasi dataset menggunakan *classes* to klaster *evaluation*, dan ketiga pengujian nilai SSE dibandingkan dengan parameter *k* dan jumlah iterasi yang terjadi (Mardiana, 2015). Penelitian ke delapan yaitu membahas Penerapan Metode *K-Means* Untuk *Clustering*

Mahasiswa Berdasarkan Nilai Akademik Dengan Weka *Interface* Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang. Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa algoritma *K-Means* bisa digunakan untuk mengelompokkan mahasiswa berdasarkan IPK dan beberapa atribut mata kuliah (Asroni, 2015).

Berdasarkan hasil dari perkembangan riset sebelumnya yang menggunakan metoda *K-Means* untuk klasterisasi, maka pada penelitian ini juga menggunakan metoda *K-Means Clustering* untuk mengetahui sumber penularan virus covid-19 paling banyak dari beberapa faktor yang ada di daerah Kota Cimahi dan Kab Bandung Barat.

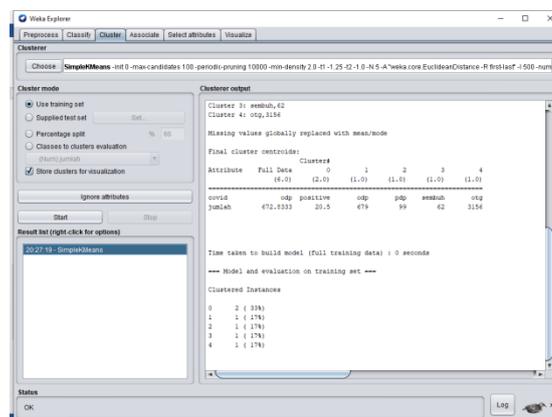


Gambar 1. Diagram Alir Penelitian

Pada Gambar 1 dijelaskan adalah pertama menyiapkan data harus dibuat dengan Microsoft Excel sesuai dengan persyaratan yang diharuskan oleh perangkat lunak WEKA dan dibuat perhari yang di *update* dari situs resmi pemerintah untuk Kota Cimahi dan Kab Bandung Barat, setelah disusun berdasarkan per tanggal masukan data tersebut ke aplikasi WEKA, sebelum itu harus mengubah data file tersebut menjadi “.arff” agar bisa terbaca di aplikasi.

Setelah itu menentukan kluster yang akan digunakan, di penelitian ini menentukan 5 kluster, sesudah di *setting* 5 kluster bisa memproses data, setelah itu akan dapatkan pembagian kluster dan harus mengubah format file tersebut untuk di input kembali di excel, jika ada data yang akan diubah harus mengubah kembali dari inputan data awal.

Hasil klasterisasi menggunakan WEKA



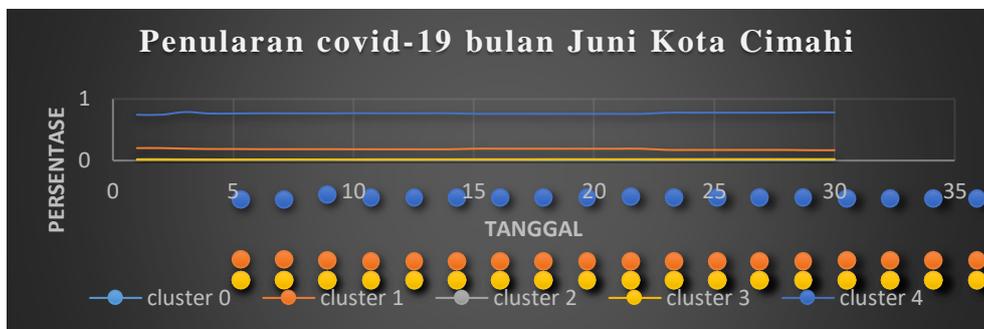
Gambar 2. Hasil Clustering di Software WEKA

Dari hasil kalsterisasi menggunakan WEKA seperti Gambar 2 dapat dijelaskan WEKA memakai kategori *SimpleKMeans* dengan konfigurasi 5 kluster dan diolah secara otomatis oleh WEKA memakai “*Use Training set*” yang artinya memakai metoda Algoritma *K-means* dasar digunakan untuk pemula dan hasilnya memakai satuan

persentase dari jumlah data yang di *input* kan sebelumnya. Yang diperlihatkan pada Gambar 2 di atas merupakan hasil akhir dari proses klasterisasi di aplikasi WEKA dan hasilnya menunjukkan persentase untuk tiap klaster bersama dengan diagram didalamnya.

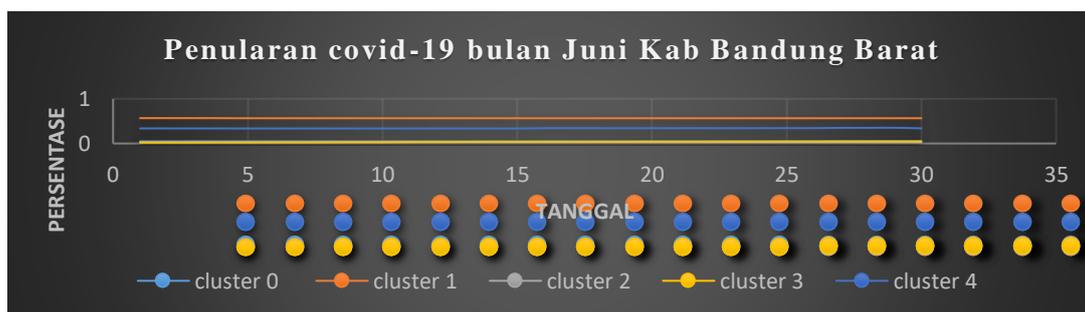
2. Hasil dan Pembahasan

A. Hasil Pengujian Perbulan



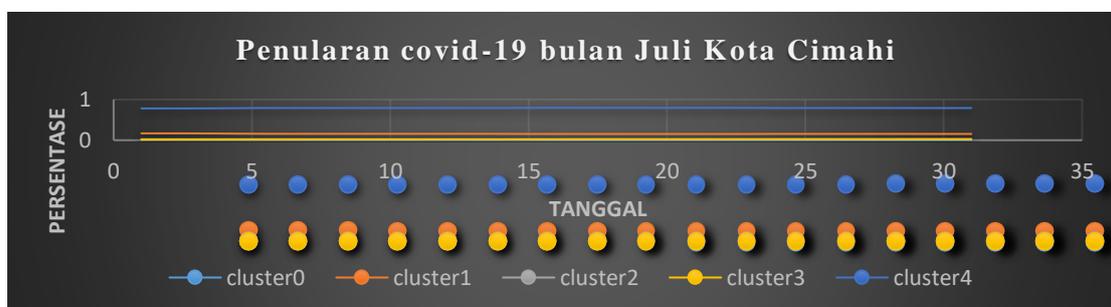
Gambar 3. Hasil 5 Klaster Bulan Juni 2020 Kota Cimahi

Pada gambar grafik diatas menunjukkan hasil 5 klaster yang terjadi pada bulan Juni 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada klaster 4 (lingkungan transportasi umum) dimana terdapat kenaikan yang signifikan pada tanggal 2–3 Juni 2020 sebesar $\pm 0,04\%$ dari 0,74%–0,78%, dan disusul posisi kedua pada klaster 1 (lingkungan tempat makan) terjadi penurunan yang signifikan pada tanggal 22–23 Juni 2020 sebesar $\pm 0,02\%$ dari 0,19%–0,17%, dan 3 klaster lainnya yaitu klaster 0 (lingkungan kantor), klaster 2 (lingkungan rumah), klaster 3 (lingkungan perbelanjaan) ada diantara 0% - 0,05%.



Gambar 4. Hasil 5 Klaster Bulan Juni 2020 Kab Bandung Barat

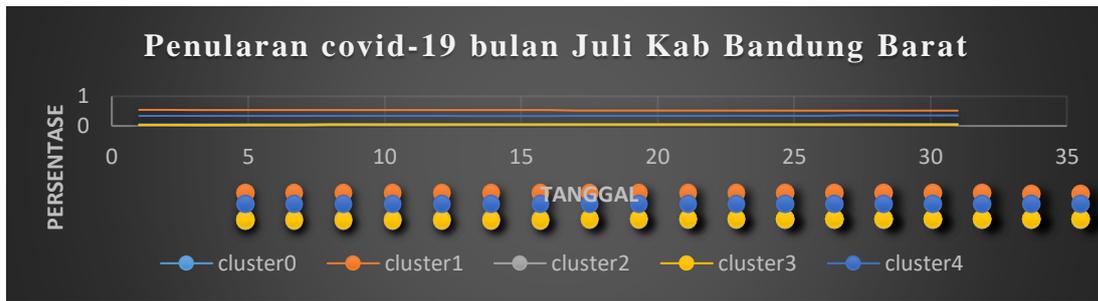
Pada gambar grafik diatas menunjukkan hasil 5 klaster yang terjadi pada bulan Juni 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada klaster 1 (lingkungan tempat makan) berada di presentase kisaran $\pm 0,55\%$, dan disusul posisi kedua pada klaster 4 (lingkungan transportasi umum) berada di presentase kisaran $\pm 0,33\%$, dan 3 klaster lainnya yaitu klaster 0 (lingkungan kantor), klaster 2 (lingkungan rumah), klaster 3 (lingkungan perbelanjaan) ada diantara 0%–0,05%.



Gambar 5. Hasil 5 Klaster Bulan Juli 2020 Kota Cimahi

Pada gambar grafik diatas menunjukkan hasil 5 klaster yang terjadi pada bulan Juli 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada klaster 4 (lingkungan transportasi umum) berada di presentase kisaran $\pm 0,79\%$, dan disusul posisi kedua pada klaster 1 (lingkungan tempat makan) berada di presentase

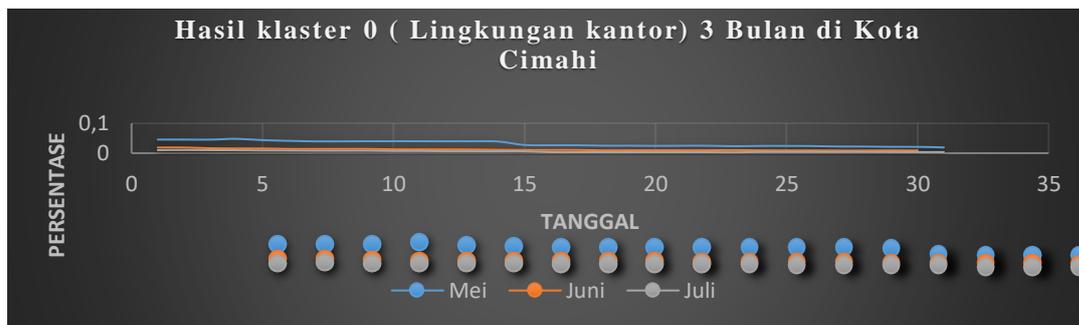
kisaran $\pm 0,15\%$, dan 3 kluster lainnya yaitu kluster 0 (lingkungan kantor), kluster 2 (lingkungan rumah), kluster 3 (lingkungan perbelanjaan) ada diantara $0\% - 0,05\%$.



Gambar 6. Hasil 5 Kluster Bulan Juli 2020 Kab Bandung Barat

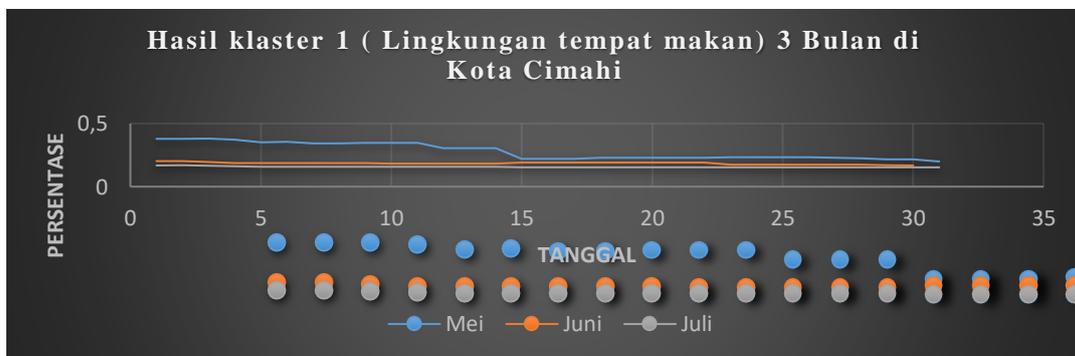
Pada gambar grafik diatas menunjukkan hasil 5 kluster yang terjadi pada bulan Juli 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada kluster 1 (lingkungan tempat makan) ada penurunan yang signifikan pada tanggal 16–17 Juli 2020 sebesar $\pm 0,01\%$ dari $0,53\% - 0,52\%$, dan disusul posisi kedua pada kluster 4 (lingkungan transportasi umum) terjadi kenaikan yang signifikan pada tanggal 16–17 Juli 2020 sebesar $\pm 0,01\%$ dari $0,33\% - 0,34\%$, dan 3 kluster lainnya yaitu kluster 0 (lingkungan kantor), kluster 2 (lingkungan rumah), kluster 3 (lingkungan perbelanjaan) ada diantara $0\% - 0,1\%$.

B. Hasil Perbandingan Per-kluster Tiap Bulan



Gambar 7. Hasil Perbandingan Kluster 0 dalam 3 Bulan di Kota Cimahi

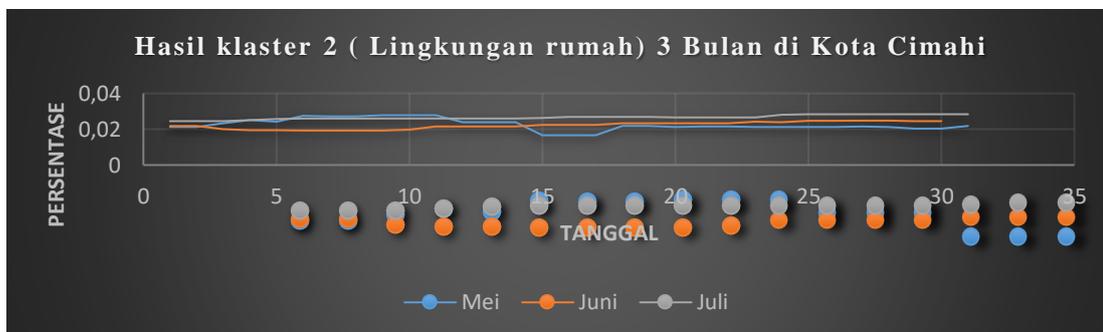
Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 0 (lingkungan kantor) yang terjadi pada bulan Mei-Juli 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Mei ada penurunan yang signifikan pada tanggal 14–15 Mei 2020 sebesar $\pm 0,01\%$ dari $0,03\% - 0,02\%$, disusul posisi kedua pada bulan Juni terjadi penurunan yang signifikan pada tanggal 2–3 Juni 2020 sebesar $\pm 0,002\%$ dari $0,018\% - 0,016\%$, dan terakhir posisi ketiga pada bulan Juli terjadi penurunan yang signifikan pada tanggal 15–16 Juli 2020 sebesar $\pm 0,003\%$ dari $0,006\% - 0,003\%$. Jadi dapat disimpulkan bahwa kluster 0 (lingkungan kantor) di Kota Cimahi dari bulan Mei–Juli 2020 mengalami penurunan yang relatif besar $\pm 0,04\%$.



Gambar 8. Hasil Perbandingan Kluster 1 dalam 3 Bulan di Kota Cimahi

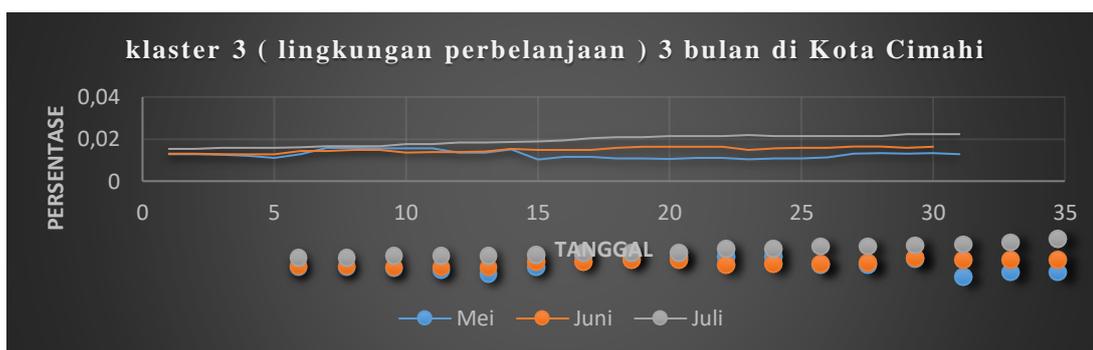
Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 1 (lingkungan tempat makan) yang terjadi pada bulan Mei-Juli 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Mei dimana terdapat penurunan yang signifikan pada tanggal 14–15 Mei 2020 sebesar $\pm 0,08\%$ dari $0,30\% - 0,22\%$,

disusul posisi kedua pada bulan Juni terjadi penurunan yang signifikan pada tanggal 22–23 Juni 2020 sebesar $\pm 0,02\%$ dari 0,19%-0,17%, dan terakhir posisi ketiga pada bulan Juli tidak terdapat kenaikan ataupun penurunan yang signifikan melainkan konsisten berada di kisaran $\pm 0,15\%$. Jadi dapat disimpulkan bahwa klaster 1 (lingkungan tempat makan) di Kota Cimahi dari bulan Mei–Juli 2020 mengalami penurunan yang relatif besar $\pm 0,22\%$.



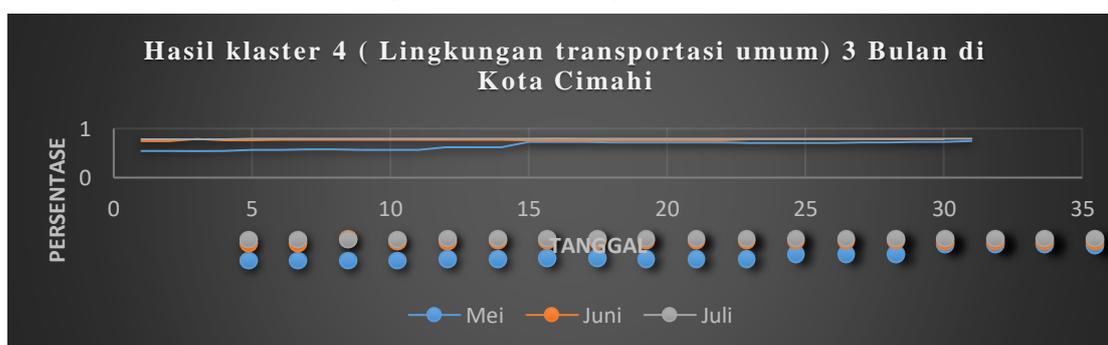
Gambar 9. Hasil Perbandingan Klaster 2 dalam 3 Bulan di Kota Cimahi

Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 2 (lingkungan rumah) yang terjadi pada bulan Mei-Juli 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Mei ada penurunan yang signifikan pada tanggal 14–15 Mei 2020 sebesar $\pm 0,01\%$ dari 0,02%–0,01%, disusul posisi kedua pada bulan Juni terjadi kenaikan yang signifikan pada tanggal 10–11 Juni 2020 sebesar $\pm 0,01\%$ dari 0,01%-0,02%, dan terakhir posisi ketiga pada bulan Juli terjadi kenaikan yang signifikan pada tanggal 23–24 Juli 2020 sebesar $\pm 0,001\%$ dari 0,026%-0,027%. Jadi dapat disimpulkan bahwa klaster 2 (lingkungan rumah) di Kota Cimahi dari bulan Mei–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,004\%$.



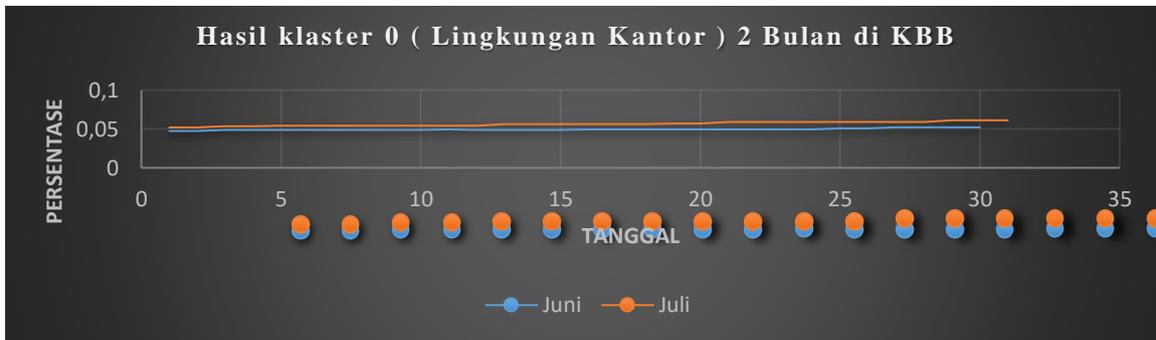
Gambar 10. Hasil Perbandingan Klaster 3 dalam 3 Bulan di Kota Cimahi

Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 3 (lingkungan perbelanjaan) yang terjadi pada bulan Mei-Juli 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Mei ada penurunan yang signifikan pada tanggal 14–15 Mei 2020 sebesar $\pm 0,004\%$ dari 0,014%–0,010%, disusul posisi kedua pada bulan Juni terjadi penurunan yang signifikan pada tanggal 22–23 Juni 2020 sebesar $\pm 0,002\%$ dari 0,016%-0,014%, dan terakhir posisi ketiga pada bulan Juli terjadi kenaikan yang signifikan pada tanggal 23–24 Juli 2020 sebesar $\pm 0,001\%$ dari 0,026%-0,027%. Jadi dapat disimpulkan bahwa klaster 3 (lingkungan perbelanjaan) di Kota Cimahi dari bulan Mei–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,004\%$.



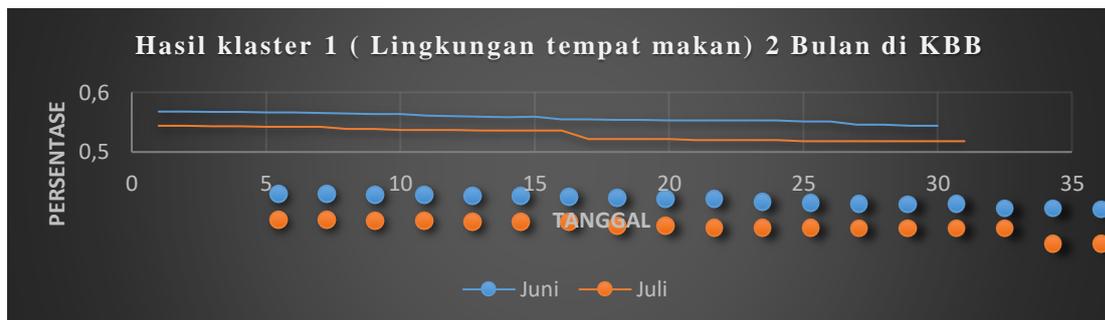
Gambar 11. Hasil Perbandingan Klaster 4 dalam 3 Bulan di Kota Cimahi

Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 4 (lingkungan transportasi umum) yang terjadi pada bulan Mei-Juli 2020 di Kota Cimahi dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Mei ada kenaikan yang signifikan pada tanggal 14–15 Mei 2020 sebesar $\pm 0,11\%$ dari $0,61\% - 0,72\%$, disusul posisi kedua pada bulan Juni terjadi penurunan yang signifikan pada tanggal 2–3 Juni 2020 sebesar $\pm 0,04\%$ dari $0,74\% - 0,78\%$, dan terakhir posisi ketiga pada bulan Juli tidak terdapat kenaikan ataupun penurunan yang signifikan melainkan konsisten berada di kisaran $\pm 0,79\%$. Jadi dapat disimpulkan bahwa klaster 4 (lingkungan transportasi umum) di Kota Cimahi dari bulan Mei–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,13\%$.



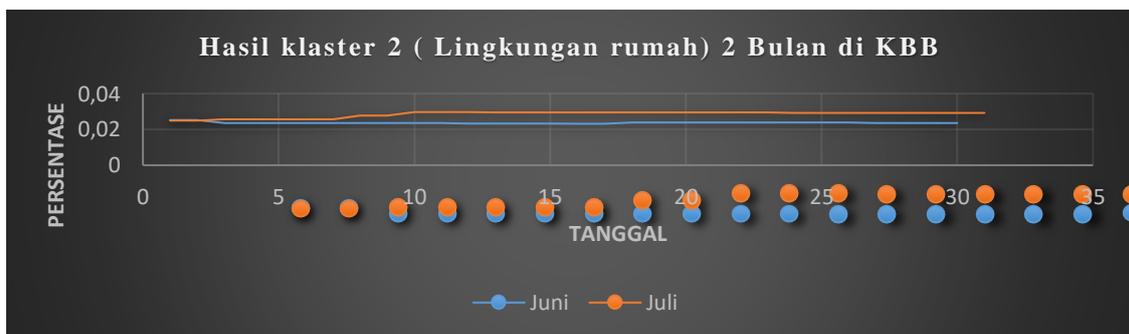
Gambar 12. Hasil Perbandingan Klaster 0 dalam 2 Bulan di Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 0 (lingkungan kantor) yang terjadi pada bulan Juni-Juli 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Juli ada kenaikan yang signifikan pada tanggal 12–13 Juli 2020 sebesar $\pm 0,002\%$ dari $0,054\% - 0,056\%$, dan terendah pada bulan Juni terjadi kenaikan yang signifikan pada tanggal 26–27 Juni 2020 sebesar $\pm 0,002\%$ dari $0,050\% - 0,052\%$. Jadi dapat disimpulkan bahwa klaster 0 (lingkungan kantor) di Kab Bandung Barat dari bulan Juni–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,007\%$.



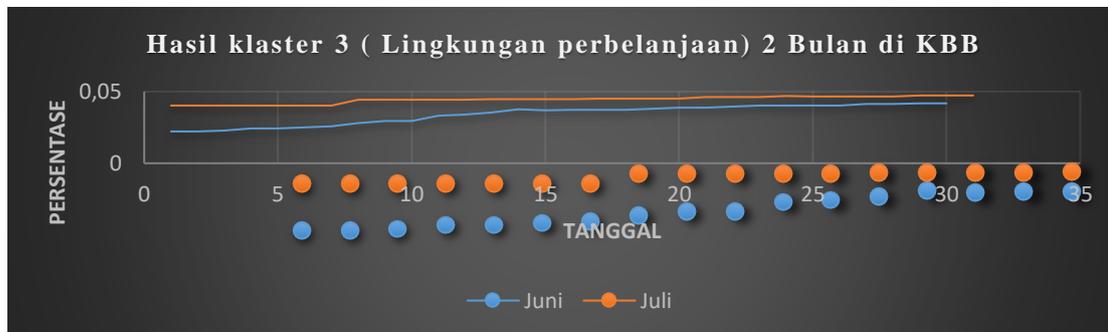
Gambar 13. Hasil Perbandingan Klaster 1 dalam 2 Bulan di Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 1 (lingkungan tempat makan) yang terjadi pada bulan Juni-Juli 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Juni ada penurunan yang signifikan pada tanggal 26–27 Juni 2020 sebesar $\pm 0,01\%$ dari $0,55\% - 0,54\%$, dan terendah pada bulan Juli terjadi penurunan yang signifikan pada tanggal 16–17 Juli 2020 sebesar $\pm 0,01\%$ dari $0,53\% - 0,52\%$. Jadi dapat disimpulkan bahwa klaster 1 (lingkungan tempat makan) di Kab Bandung Barat dari bulan Juni–Juli 2020 mengalami penurunan yang relatif besar $\pm 0,03\%$.



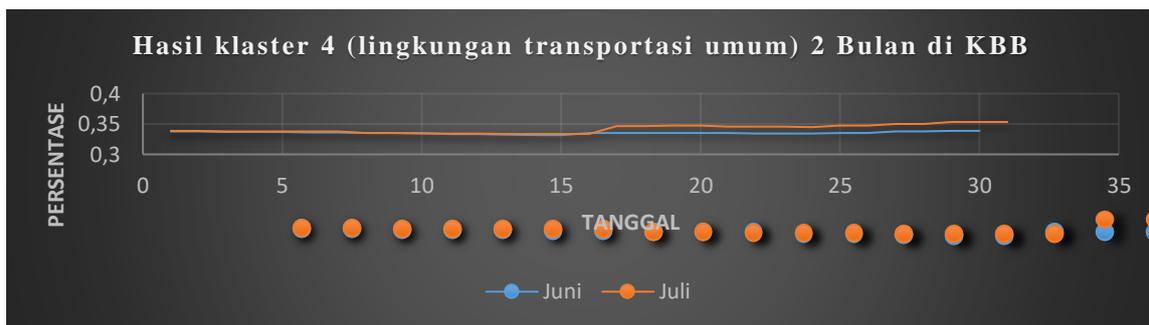
Gambar 14. Hasil Perbandingan Klaster 2 dalam 2 Bulan di Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 2 (lingkungan rumah) yang terjadi pada bulan Juni-Juli 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Juni ada penurunan yang signifikan pada tanggal 2–3 Juni 2020 sebesar $\pm 0,002\%$ dari $0,025\% - 0,023\%$, dan terendah pada bulan Juli terjadi kenaikan yang signifikan pada tanggal 9–10 Juli 2020 sebesar $\pm 0,002\%$ dari $0,027\% - 0,029\%$. Jadi dapat disimpulkan bahwa klaster 2 (lingkungan rumah) di Kab Bandung Barat dari bulan Juni–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,005\%$.



Gambar 15. Hasil Perbandingan Klaster 3 dalam 2 Bulan di Kab Bandung Barat

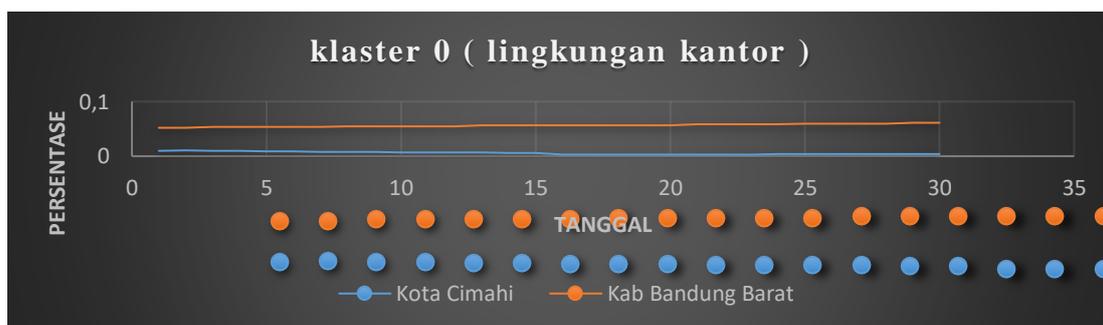
Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 3 (lingkungan perbelanjaan) yang terjadi pada bulan Juni-Juli 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Juni ada kenaikan yang signifikan pada tanggal 10–11 Juni 2020 sebesar $\pm 0,004\%$ dari $0,029\% - 0,033\%$, dan terendah pada bulan Juli terjadi kenaikan yang signifikan pada tanggal 7–8 Juli 2020 sebesar $\pm 0,004\%$ dari $0,040\% - 0,044\%$. Jadi dapat disimpulkan bahwa klaster 3 (lingkungan perbelanjaan) di Kab Bandung Barat dari bulan Juni–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,01\%$.



Gambar 16. Hasil Perbandingan Klaster 4 dalam 2 Bulan di Kab Bandung Barat

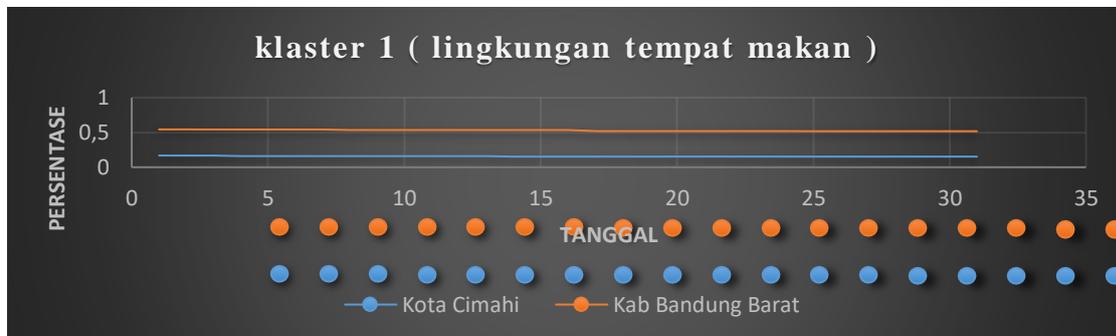
Pada gambar grafik diatas menunjukkan hasil perbandingan klaster 4 (lingkungan transportasi umum) yang terjadi pada bulan Juni-Juli 2020 di Kab Bandung Barat dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada bulan Juni ada kenaikan yang signifikan pada tanggal 15–16 Juni 2020 sebesar $\pm 0,003\%$ dari $0,332\% - 0,335\%$, dan terendah pada bulan Juli terjadi kenaikan yang signifikan pada tanggal 16–17 Juli 2020 sebesar $\pm 0,01\%$ dari $0,33\% - 0,34\%$. Jadi dapat disimpulkan bahwa klaster 4 (lingkungan transportasi umum) di Kab Bandung Barat dari bulan Juni–Juli 2020 mengalami kenaikan yang relatif besar $\pm 0,01\%$.

C. Hasil Perbandingan Klaster dan Bulan yang sama di Daerah yang Berbeda



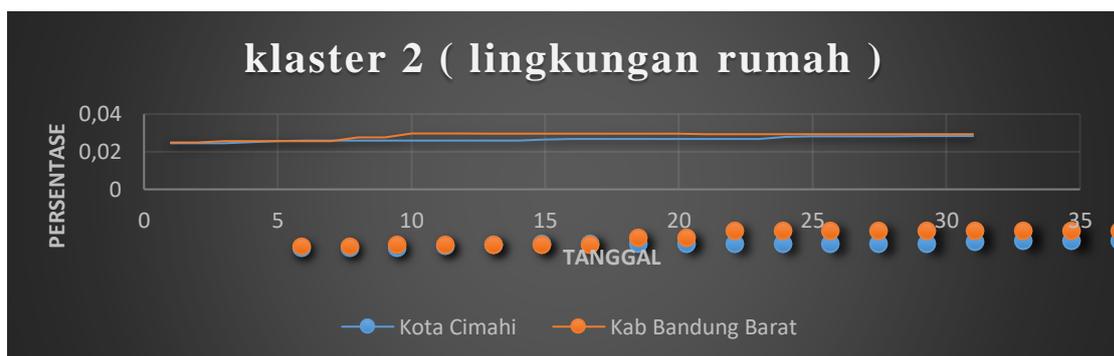
Gambar 17. Hasil Perbandingan Klaster 0 di Kota Cimahi dan Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 0 (lingkungan kantor) yang terjadi pada bulan Juni 2020 di 2 daerah yang berbeda dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada daerah Kab Bandung Barat ada kenaikan yang signifikan pada tanggal 24–25 Juni 2020 sebesar $\pm 0,01\%$ dari $0,04\%$ – $0,05\%$, dan terendah pada daerah Kota Cimahi terjadi penurunan yang signifikan pada tanggal 2–3 Juni 2020 sebesar $\pm 0,002\%$ dari $0,018\%$ – $0,016\%$. Jadi dapat disimpulkan bahwa kluster 0 (lingkungan kantor) di Bulan Juni 2020 antara Kota Cimahi dan Kab Bandung Barat mengalami selisih yang relatif besar $\pm 0,03\%$.



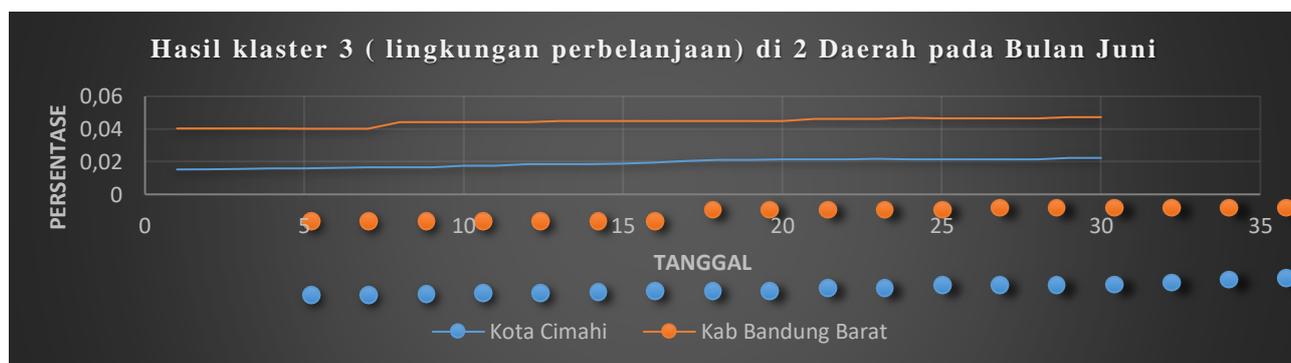
Gambar 18. Hasil Perbandingan Kluster 1 di Kota Cimahi dan Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 1 (lingkungan tempat makan) yang terjadi pada bulan Juli 2020 di 2 daerah yang berbeda dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada daerah Kab Bandung Barat ada penurunan yang signifikan pada tanggal 16–17 Juli 2020 sebesar $\pm 0,01\%$ dari $0,53\%$ – $0,52\%$, dan terendah pada daerah Kota Cimahi tidak ada penurunan atau kenaikan yang signifikan melainkan konsisten berada di kisaran $\pm 0,15\%$. Jadi dapat disimpulkan bahwa kluster 1 (lingkungan tempat makan) di Bulan Juli 2020 antara Kota Cimahi dan Kab Bandung Barat mengalami selisih dari rata–rata yang kisarannya $\pm 0,43\%$.



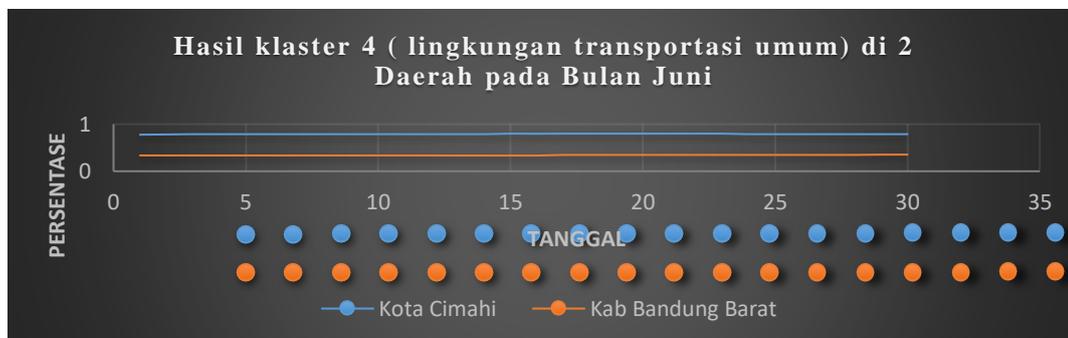
Gambar 19. Hasil Perbandingan Kluster 2 di Kota Cimahi dan Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 1 (lingkungan rumah) yang terjadi pada bulan Juli 2020 di 2 daerah yang berbeda dan dari hasil *clustering* menunjukkan penularan covid-19 paling banyak yang pertama pada daerah Kab Bandung Barat ada kenaikan yang signifikan pada tanggal 9–10 Juli 2020 sebesar $\pm 0,002\%$ dari $0,027\%$ – $0,029\%$, dan terendah pada daerah Kota Cimahi terjadi kenaikan yang signifikan pada tanggal 23–24 Juli 2020 sebesar $\pm 0,001\%$ dari $0,026\%$ – $0,027\%$. Jadi dapat disimpulkan bahwa kluster 2 (lingkungan rumah) di Bulan Juli 2020 antara Kota Cimahi dan Kab Bandung Barat mengalami selisih dari rata–rata yang kisarannya $\pm 0,002\%$.



Gambar 20. Hasil Perbandingan Kluster 3 di Kota Cimahi dan Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 3 (lingkungan perbelanjaan) yang terjadi pada bulan Juni 2020 di 2 daerah yang berbeda dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada daerah Kab Bandung Barat ada kenaikan yang signifikan pada tanggal 10–11 Juni 2020 sebesar $\pm 0,004\%$ dari $0,029\% - 0,033\%$, dan terendah pada daerah Kota Cimahi terjadi penurunan yang signifikan pada tanggal 5–6 Juni 2020 sebesar $\pm 0,002\%$ dari $0,012\% - 0,014\%$. Jadi dapat disimpulkan bahwa kluster 3 (lingkungan perbelanjaan) di Bulan Juni 2020 antara Kota Cimahi dan Kab Bandung Barat mengalami selisih dari rata-rata yang kisarannya $\pm 0,02\%$.



Gambar 21. Hasil Perbandingan Kluster 4 di Kota Cimahi dan Kab Bandung Barat

Pada gambar grafik diatas menunjukkan hasil perbandingan kluster 4 (lingkungan transportasi umum) yang terjadi pada bulan Juni 2020 di 2 daerah yang berbeda dan dari hasil *clustering* menunjukkan penularan covid-19 tertinggi pada daerah Kota Cimahi ada kenaikan yang signifikan pada tanggal 2–3 Juni 2020 sebesar $\pm 0,04\%$ dari $0,74\% - 0,78\%$, dan terendah pada daerah Kab Bandung Barat tidak ada penurunan atau kenaikan yang signifikan melainkan konsisten berada di kisaran $\pm 0,33\%$. Jadi dapat disimpulkan bahwa kluster 4 (lingkungan transportasi umum) di Bulan Juni 2020 antara Kota Cimahi dan Kab Bandung Barat mengalami selisih dari rata-rata yang kisarannya $\pm 0,4\%$.

Kesimpulan

Berdasarkan hasil penelitian algoritma *K-Means clustering analysis* pada virus covid-19 dengan aplikasi WEKA (studi kasus Kota Cimahi dan Kab Bandung Barat), dapat disimpulkan bahwa berdasarkan hasil simulasi *K-Means* dengan *software* WEKA pada Kota Cimahi dari tanggal 1 Mei sampai tanggal 31 Juli 2020 kluster tertinggi ditempati oleh kluster 4 (lingkungan transportasi umum) dengan rata-rata $0,79\%$ dan kluster yang terendah ditempati oleh kluster 0 (lingkungan kantor) dengan rata-rata $0,005\%$, sedangkan di Kab Bandung Barat dari tanggal 1 Juni sampai 31 Juli 2020 kluster tertinggi ditempati oleh kluster 1 (lingkungan tempat makan) dengan rata-rata $0,56\%$ dan kluster yang terendah ditempati oleh kluster 2 (lingkungan rumah) dengan rata-rata $0,024\%$. Kesimpulan ke 2 yaitu, berdasarkan hasil dari perbandingan 2 daerah yaitu Kota Cimahi dan Kab Bandung Barat di bulan Juni dan Juli 2020 dapat disimpulkan dari kluster 0 (lingkungan kantor) dampak covid-19 tertinggi dari Kab Bandung Barat dengan rata-rata $0,05\%$, untuk kluster 1 (lingkungan tempat makan) dampak covid-19 tertinggi dari Kab Bandung Barat dengan rata-rata $0,55\%$, untuk kluster 2 (lingkungan rumah) dengan rata-rata $0,028\%$, untuk kluster 3 (lingkungan perbelanjaan) dampak covid-19 tertinggi dari Kab Bandung Barat dengan rata-rata $0,04\%$, untuk kluster 4 (lingkungan transportasi umum) dampak covid-19 tertinggi dari Kota Cimahi dengan rata-rata $0,79\%$. Jadi dapat disimpulkan dari 5 kluster tersebut paling dominan penularan di Kab Bandung Barat dan yang tertinggi angka penularannya di Kota Cimahi. Kesimpulan ke 3 yaitu, berdasarkan hasil perbandingan yang telah dilakukan antara simulasi *K-Means* di *software* WEKA dan di Ms. Excel secara teori *K-Means* hasil dari kedua percobaan di sisi persentase tersebut tidak jauh berbeda namun yang membedakan di sisi kluster jika di hasil *software* WEKA kluster tertinggi yaitu kluster 4 (lingkungan transportasi umum) dan hasil teori di Ms. Excel yaitu kluster 2 (lingkungan rumah) dengan rata-rata persentase yang sama $0,54\%$ disebabkan karena untuk *software* WEKA pemilihan kluster secara otomatis dan teori algoritma sendiri pemilihan secara manual.

Saran

Adapun Saran untuk penelitian selanjutnya adalah pada penelitian *clustering data mining* ini bisa dikembangkan dengan menggunakan tahapan prediksi untuk kedepannya dan pada penelitian *clustering data mining* ini bisa dikembangkan lagi dengan cara menambahkan wilayah pengamatan kota-kota besar lain yang ada di Jawa Barat.

Daftar Notasi

- $D(i,j)$ = Jarak data ke i ke pusat kluster j
- X_{ki} = Data ke i pada atribut data ke k
- X_{kj} = Titik pusat ke j pada atribut ke k

Daftar Pustaka

- Ahmad, A. (2019). *Perkembangan Teknologi Komunikasi dan Informasi: Akar Revolusi dan Berbagai Standarnya*. Jakarta: Dakwah Tabligh, Vol. 13, No. 1.
- Asroni, R. A. (2015). *Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang*. Magelang: Ilmiah Semesta Teknika.
- Benri Melpa Metisen, H. L. (2015). *Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila*. Bengkulu: Media Infotama Vol. 11 No. 2.
- COVID-19. (2020, September 26). Retrieved from covid-19: <https://covid19.go.id/>
- Gunawan Abdillah, F. A. (2016). *Penerapan Data Mining Pemakaian Air Pelanggan untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru di PDAM Tirta Raharja Menggunakan Algoritma K-Means*. Cimahi: Seminar Nasional Teknologi Informasi dan Komunikasi.
- Iin Parlina, A. P. (2018). *Memfaatkan Algoritma K-Means Dalam Menentukan Pegawai Yang Layak Mengikuti Assessment Center Untuk Clustering Program SDP*. Pematangsiantar: CESS (Journal of Computer Engineering System and Science) Vol. 3 No. 1.
- Mardiana, T. (2015). *Bag Of Word Clustering Using WEKA*. Yogyakarta: Universitas Gadjah Mada.
- Narwati. (2016). *Pengelompokan Mahasiswa Menggunakan Algoritma K-Means*. Jakarta: Fakultas Teknologi Informasi.
- Nurlaila Suci Rahayu Rais, M. M. (2018). *Kemajuan Teknologi Informasi Berdampak Pada Generalisasi Unsur Sosial Budaya Bagi Generasi Milenial*. Jakarta: Mozaik Vol. X Edisi 2.
- Pareza Alam Jusia, F. M. (2019). *Clustering Data Untuk Rekomendasi Penentuan Jurusan Perguruan Tinggi Menggunakan Metode K-Means*. Jambi: IKRA-ITH Informatika.
- Riyani Wulan Sari, A. W. (2018). *Implementasi Rapidminer Dengan Metode K-Means (Study Kasus: Imunisasi Campak Pada Balita Berdasarkan Provinsi)*. Pematangsiantar: KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer) Volume 2, Nomor 1.
- Setiawan, Daryanto. (2018). *Dampak Perkembangan Teknologi Informasi dan Komunikasi Terhadap Budaya*. Deli Serdang: SIMBOLIKA, Vol. 4.
- Setiawan, Daryanto. (2018). *Impact of Information Technology Development and Communication on Culture*. Deli Serdang: SIMBOLIKA.
- WEKA. (2020). Retrieved from WEKA: <https://dokumen.tips/education/bab-1-pendahuluan-weka.html>
- Windarto, A. P. (2017). *Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering*. Pematangsiantar: Techno.COM, Vol. 16, No. 4.
- Zulfah, S. (2016). *Pengaruh Perkembangan Teknologi Informasi Lingkungan (Studi Kasus Kelurahan Siti Rejo I Medan)*. Medan: Universitas Islam Sumatera Utara.