

PENDETEKSIAN DAERAH (PROVINSI) RAWAN COVID19 DENGAN METODE UNSUPERVISED LEARNING & ALGORITMA K-MEDOIDS

Fitri Marisa^[1] Bagas Ariefia Pribady^[2] Ayuni Desi^[3] Anastasia L Maukar^[4]

Jurusan Teknik Informatika Fakultas Teknik Universitas Widyagama Malang^{[1][2][3]}

Jl. Borobudur No.35 Malang, 65142

Department Teknik Industri - President University^[4]

Jababeka Education Park, Jl. Ki Hajar Dewantara, Cikarang Bekasi

email : fitrimarisa@gmail.com^[1], gbagasarief@gmail.com^[2], ayunidesi997@gmail.com^[3]

almaukar@president.ac.id^[4]

Abstrak

Hingga kini dari 65 jumlah negara telah terpapar oleh Covid_19. Sesuai dari data Satuan Penanganan Covid_19 terkonfirmasi mulai dari tanggal 26 Januari 2021 Pasien yang terpapar Virus Covid_19 sudah mencapai angka 98.794.942 orang diseluruh dunia. Khusus indonesia sendiri pasien yang terpapar covid-19 masih terus bertambah diangka 999.256 orang terkonfirmasi covid-19. Selain itu ada 809.488 orang sembuh dan 28.132 orang meninggal dunia akibat pandemi ini. Pasien pertama di indonesia yang dinyatakan positif covid_19 bermula saat sedang menghadiri sebuah acara yang ada di Jakarta yang mana pasien positif covid_19 tersebut melakukan kontak dengan tamu warga negara asing yang berasal dari jepang yang mana migrasi ke Malaysia. Selesai dari acara itu pasien yang terdeteksi positif Covid_19 ini merasakan demam yang disertai batuk dan dada yang sesak napas setelah diperiksa ternyata positif virus covid_19.

Penulis menggunakan dalam melakukan penelitian akan menggunakan 3 buah metode adalah pertama adalah metode untuk pengembangan perangkat lunak, lalu metode untuk pengumpulan data dan terakhir pengelompokan data nya. Dalam hal ini Metode Pemodelan Perangkat Lunak menggunakan Waterfall, Pengumpulan dan Pengelompokan data menggunakan Algoritma K-Medoids dan Unsupervised Learning.

Setelah beberapa percobaan data set dilakukan, penulis memutuskan untuk menggunakan model atau metode tersebut agar data yang ada dapat dikelompokkan kemudian untuk hasil dan pembahasan akan disesuaikan dengan tujuan dari penelitian yang penulis akan penulis teliti. Konsep yang jauh berbeda dengan Supervised learning, Unsupervised Learning dengan tujuan pengelompokan data akan membuat penggunaanya dapat mengelompokkan berbagai objek yang ada dengan cluster yang berbeda beda.

Kata Kunci-- Covid_19;Peta ;Unsupervised Learning; Algoritma K-Medoids

I. PENDAHULUAN

Seperti yang kita ketahui bersama dimana pada masa sekarang, masa pandemic terjadi di berbagai negara diseluruh penjuru dunia, termasuk di Indonesia. Tercatat semakin banyak korban yang teridentifikasi positif covid_19.

Dari data yang ada hingga kini banyaknya 65 jumlah negara terpapar oleh Covid_19. Satgas Penanganan Covid_19 dari tanggal 26 Januari 2021 pasien positif covid_19 mencapai angka 98.794.942

orang dinyatakan positif diseluruh dunia. Khusus indonesia sendiri hingga data terakhir pasien positif covid_19 semakin bertambah dan mencapai angka 999.256 orang terkonfirmasi covid 19. Selain itu ada 809.488 orang sembuh dan 28.132 orang meninggal dunia akibat pandemik ini. Padahal saat penelitian sebelumnya sesuai data oleh Satgas Penanganan Covid_19 dari tanggal 02 Maret 2020 terkonfirmasi pasien yang menderita covid_19, 90.308 orang atau jiwa. Pada saat itu terinfeksi 2 orang. Diduga karena korban sedang mengadakan acara dan mengadakan

kontak dengan WNA, pada saat itu Angka kematian semakin menambah drastis hingga angka 3.087 atau sama dengan 2.3% dan untuk jumlah keseluruhan pasien yang dinyatakan sembuh atau negatif Covid_19 mencapai 45.726 pasien. [1]

Awal mula kemunculannya kasus Covid-19 pertama kali di provinsi Wuhan, China muncul dan menyerang manusia. Diiduga merupakan penyakit pneumonia, dengan varian ciri-ciri gejala yang serupa sakit flu seperti pada umumnya kemudian gejala demam yang disertai batuk, dada sesak, nafsu makan berkurang. Sangat berbeda dengan influenza, virus ini dapat berkembang dengan cepat melalui udara dengan berbagai varian gejala yang baru ditemukan hingga sangat fatal akibatnya infeksi lebih parah dan gagal organ. Kondisi darurat ini terutama terjadi pada pasien dengan masalah kesehatan sebelumnya. [2]

Berdasarkan jurnal yang ditulis oleh peneliti sebelumnya tentang Pengelompokan data (Clustering) Penyebaran covid_19 di Indonesia dengan algoritma K-Medoids [3] pada jurnal tersebut membahas tentang cara pengelompokan data covid 19 yang tersebar di seluruh Republik Indonesia menggunakan metode algoritma K-Medoids Clustering, namun belum dibuat menjadi sebuah system yang dapat membantu mempublikasikan data tersebut.

Dalam hal ini menurut Keputusan Presiden Republik Indonesia No 7 Tahun 2020 Tentang Gugus Tugas Percepatan Penanggulangan Corona Virus [4] akhirnya penulis berinisiatif untuk membuat sistem berbasis web, system yang pada masa sekarang sudah menjadi komoditas public sehingga bisa untuk membantu dalam melihat perkembangan pandemic yang dialami di Indonesia selama berjalannya waktu, selain itu juga dapat membantu juga dalam pembuatan keputusan pemerintah dalam menerapkan kebijakan baik secara nasional maupun lingkup daerah, karena data yang sudah dikelompokkan dan tersusun rapih secara sistematis.

Media penyimpanan pun akan lebih praktis karena semua akan disimpan dalam bentuk database. Kegiatan analisa data kini akan menjadi lebih cepat bila menggunakan sistem berbasis web tersebut.

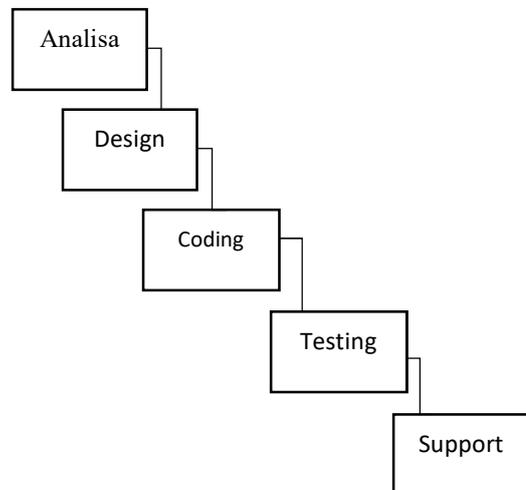
Perlunya pengelompokan data tersebut karena beberapa pokok masalah yang harus tersolusikan dan menjadi kebutuhan konsumsi baik untuk public maupun pemerintahan itu sendiri, diantaranya dalam proses pembuatan kebijakan, keterbukaan data dengan public,

penggambaran grafik sukses tidaknya kebijakan yang sudah dilakukan.

Sehingga kali ini penulis akan mencoba membuat dan mendesain sebuah system yang akan digunakan untuk menggambarkan perkembangan covid19 yang ada di Indonesia dan mempublikasikannya secara online melalui portal web yang nantinya dapat diakses seluruh warga Indonesia, sehingga keterbukaan data tersebut dapat tercapai dan membantu dalam proses kesadaran setiap warna negara pada daerah daerah tertentu yang ada di Indonesia terkait covid19 dengan status yang bisa diartikan dalam status bahaya, berhati-hati dan aman.

II. METODE PENELITIAN

Diagram Pemodelan Perangkat Lunak Model air terjun, atau orang mengenalnya dengan nama Waterfall ini adalah metode yang menyediakan pendekatan perangkat lunak secara berurutan yang berawal dari analisa, design, coding, testing, dan support.



A. Metode Mendapatkan Data

Penulis dalam mendapatkan data untuk pembuatan Makalah menggunakan teknik sebagai berikut:

1. Pencarian Data

Penulis melakukan pengumpulan data bersumber dari data covid19 yang dipublish oleh satgas covid19 guna pengamatan secara langsung terhadap perkembangan data yang

ada di kementerian. Sehingga bisa didapat data yang sesuai untuk ditampilkan pada sistem berbasis web ini.

2. Studi Pustaka

Untuk pengumpulan data sekunder penulis memperoleh data melalui studi kepustakaan, yaitu metode pencarian data atau informasi melalui buku-buku referensi, majalah, koran, dsb.

B. Teknik Pengolahan Data

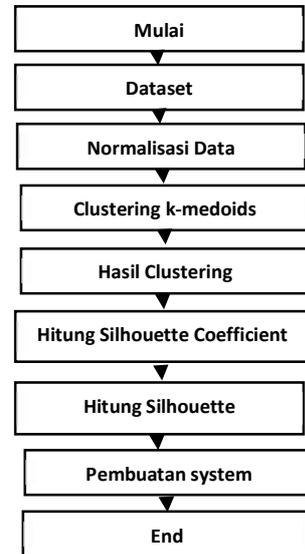
Teknik pengolahan data menggunakan algoritma K-Medoids yang merupakan turunan dari metode Unsupervised Learning yang tujuannya adalah untuk mengelompokan data. Kami memilih K-Medoids karena jika membandingkan antara Algoritma K-Medoids dan K-Means, K-Medoids bisa dibilang lebih baik, karena mendapatkan k yang disebut object representatif dan meminimalisir total object data yang tidak sama. [3] atau secara rincin untuk K-means merupakan algoritma yang digunakan sebagai pelatihan unsupervise dan dipublikasikan untuk pertama kalinya oleh Stuart Lloyd pada tahun 1984 dan merupakan algoritma clustering yang banyak digunakan.[5]

Langkah dalam algoritma K-medoids adalah sebagai berikut :

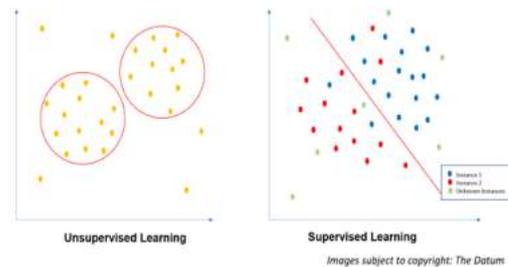
- a. Input data Covid19.
- b. Memproses perubahan data yang sudah di input untuk melakukan perhitungan lanjutan.
- c. Memproses clusterisasi menggunakan algoritma K-Medoids.
- d. Memvalidasi kualitas cluster dengan menghitung nilai Silhouette Coefficient
- e. Mendapatkan data yang sudah melalui proses clustering dengan kualitas cluster yang bisa dilihat dari nilai Silhouette Coefficient yang sudah didapat. Secara singkat langkah dari metode K-Medoids ditunjukkan melalui diagram dan rumus seperti yang ada dibawah ini [6]

$$\sum_{i=1}^n (x_i - y_i)^2; 1, 2, 3, \dots, n, d$$

$$(x, y) = \sqrt{x^2 + y^2} = \sqrt{z}$$



C. Unsupervised Learning

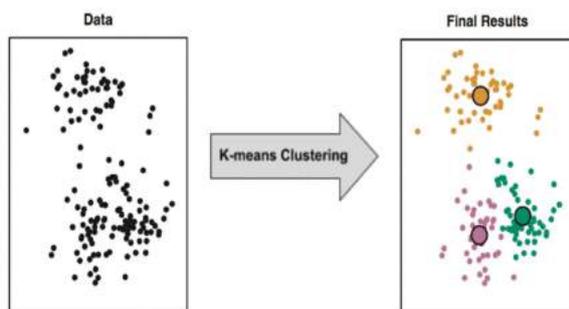


Sudah banyak percobaan dengan menggunakan model data set yang berbeda-beda, pada umumnya Machine Learning mempunyai 3 jenis pertama supervised learning, kedua unsupervised learning, yang terakhir adalah reinforcement learning, dari tipe machine learning memiliki tersebut memiliki tipe data set nya sendiri [7], dan disini penulis memilih menggunakan model Unsupervised learning tersebut.

Goals dari Unsupervised Learning adalah untuk membuat para penggunanya bisa mengelompokkan object-object yang memiliki value sama dalam scope tertentu. Unsupervised Learning dinilai sangat cocok untuk mengeloka atau mengklasifikasi suatu pola dari banyak object sejenis yang tidak sepenuhnya sama.

Jadi secara singkat dapat disimpulkan bahwa clustering merupakan kegiatan untuk memecah data ke dalam sejumlah kelompok (cluster) tertentu [8].

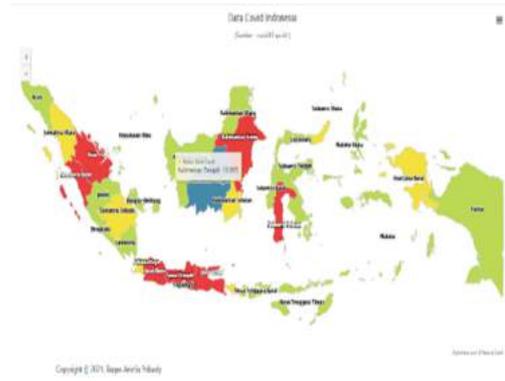
Contoh Kasus dalam penyelesaian masalah menggunakan metode Unsupervised Learning ini adalah misal lembaga negara (BPS) ingin melakukan pengelompokan data warna negara, yang dimana negara Indonesia ini adalah negara yang penuh kemajemukan, agar dapat mengelompokan segala kemajemukan tersebut bisa menggunakan unsupervised learning untuk mengelompokan tiap tiap kelompok data dari warganya tersebut, seperti pengelompokan berdasarkan jenis kelamin, agama, ras, suku dan budaya. Sehingga terbentuknya data untuk dibuat menjadi statistic yang digambarkan secara berkelompok yang ditentukan tersebut.



III. HASIL IMPLEMENTASI

No	Kode Provinsi	Nama Provinsi	Jumlah Data
1	id-ac	Aceh	9,053
2	id-su	Sumatera Utara	19,647
3	id-sb	Sumatera Barat	25,6
4	id-ri	Riau	27,371
5	id-kr	Kepulauan Riau	7,554
6	id-ja	Jambi	3,946
7	id-be	Bengkulu	4,226
8	id-sl	Sumatera Selatan	13,068
9	id-bb	Bangka Belitung	3,481
10	id-1024	Lampung	8,143
11	id-bt	Banten	22,11
12	id-jk	Jakarta	227,365
13	id-jr	Jawa Barat	112,587
14	id-jt	Jawa Tengah	102,904
15	id-ji	Jawa Timur	99,377

Dari hasil yang penulis paparkan, dapat dilihat bahwa daerah yang berwarna merah adalah daerah dengan potensi covid yang sedang kerawanan tinggi,



kuning kerawanan sedang dan hijau kerawanan rendah secara persentase.

Dengan rumus sebagai berikut :

JCI = Jumlah keseluruhan Data Covid Indonesia

JCP = Jumlah Data Covid Provinsi

$JCP / JCI \times 100\% =$ Persentase Covid Daerah tersebut.

Dari hasil presentase yang didapat penulis melakukan clustering dimana data yang mendapatkan hasil persentase lebih dari 2,5% penulis set menjadi cluster daerah yang berwarna merah atau kerawanan tinggi, kurang dari 2,5% & lebih dari 1% penulis set menjadi cluster daerah berwarna kuning atau kerawanan sedang dan terakhir persentase data kurang dari 1% penulis set menjadi cluster daerah warna hijau dengan kerawanan rendah.

V. DAFTAR PUSTAKA

- [1] M. M. C. Otálora, “Yuliana,” *Parq. los afectos. Jóvenes que cuentan*, vol. 2, no. February, pp. 124–137, 2020, doi: 10.2307/j.ctvzxxb18.12.
- [2] V. No and N. Mona, “Konsep Isolasi Dalam Jaringan Sosial Untuk Meminimalisasi Efek Contagious (Kasus Penyebaran Virus Corona Di Indonesia),” *J. Sos. Hum. Terap.*, vol. 2, no. 2, pp. 117–125, 2020, doi: 10.7454/jsht.v2i2.86.
- [3] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. Ilmi R.H.Zer, and D. Hartama, “Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia,” *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 166–173, 2020.

- [4] K. Karyono, R. Rohadin, and D. Indriyani, "Penanganan Dan Pencegahan Pandemi Wabah Virus Corona (Covid-19) Kabupaten Indramayu," *J. Kolaborasi Resolusi Konflik*, vol. 2, no. 2, p. 164, 2020, doi: 10.24198/jkrk.v2i2.29127.
- [5] D. Marlina, N. Lina, A. Fernando, and A. Ramadhan, "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 4, no. 2, p. 64, 2018, doi: 10.24014/coreit.v4i2.4498.
- [6] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan Malang," *J. Terap. Sains Teknol.*, vol. 1, no. 3, pp. 54–69, 2019.
- [7] A. Rauhan, "Pengolahan Data Menggunakan Machine Learning," *Student Pap. Pertamina Univ.*, no. 021, pp. 1–4, 2019, [Online]. Available: <https://library.universitaspertamina.ac.id/xmlui/bitstream/handle/123456789/162/JurnalIlmiah.pdf?sequence=1&isAllowed=y>.
- [8] F. Farahdinna, I. Nurdiansyah, A. Suryani, and A. Wibowo, "Perbandingan Algoritma K-Means Dan K-Medoids Dalam Klasterisasi Produk Asuransi Perusahaan Nasional," *J. Ilm. FIFO*, vol. 11, no. 2, p. 208, 2019, doi: 10.22441/fifo.2019.v11i2.010.