



Volume 20 No.2 September 2023

Jurnal Ilmiah Fakultas Teknik LIMIT'S

Rancang Bangun Aplikasi Kriptografi Dengan Rivest Code 4 (RC4) Untuk Pengamanan File Dokumen Berbasis Web
Muhammad Malay, R, Faizal Zuli

Aplikasi Data Mining Untuk Clustering Penyebaran Covid-19 Di DKI Jakarta Menggunakan Algoritma K-Means
Habibi K. Al Hanif, Turkamun Adi Kurniawan, T.W. Wisjhnuadji

Perancangan Alat Pendekripsi Kekeruhan dan Pengurasan Air
Riamma Sibarani, Ferry

Rancang Bangun Sistem Keamanan Rumah Berbasis *Internet Of Things* (IoT)
Mico Ardana, Bosar Panjaitan, Teguh Budi Santoso

Aplikasi Pembelajaran Matematika Berbasis android Studi kasus: siswa kelas ii SD Negeri 01 Pagi Kembangan Selatan Jakarta Barat
Kiki Kusumawati, Prionggo Hendradi, Muhammad Alif Fauzi

Sistem Informasi Penilaian Kinerja Pegawai Berbasis Web Pada Pt.PermataIndonesia (Studi Kasus : Kantor Cabang Mayestik Jakarta Selatan)
Wawan Kurniawan , Nurul Chafid, Indah Kurniati

Perancangan Sistem Informasi Pengarsipan Dokumen Di Universitas Satya Negara Indonesia
Prionggo Hendradi, Khey Khey Rakhmawati Dewi

Sistem Informasi Geografis Berbasis Location Based Service Untuk Pencarian Wilayah Krisis Pangan
Istiqomah Sumadikarta, Odi Kurniadi

Evaluasi Kinerja Tata Kelola Teknologi Informasi Dengan Framework Cobit. Studi kasus di sistem informasi Perikanan (SIP) Pada Direktorat Sumber Daya Ikan, Direktorat Jenderal Perikanan Tangkap Departemen Kelautan Dan Perikanan
Wahyu Fajar Arinto, Agung Priambodo

JURNAL ILMIAH FAKULTAS TEKNIK
LIMIT'S



ISSN 0216-1184



ISSN 0216-1184

Jurnal Ilmiah Fakultas Teknik

LIMIT'S

Volume 20

September

Nomor 2

Rancang Bangun Aplikasi Kriptografi Dengan Rivest Code 4 (RC4) Untuk Pengamanan File Dokumen Berbasis Web Muhammad Malay, R. Faizal Zuli	1-9
Aplikasi Data Mining Untuk Clustering Penyebaran Covid-19 Di DKI Jakarta Menggunakan Algoritma K-Means Habibi K. Al Hanif, Turkamun Adi Kurniawan, T.W. Wisjhnuadji	10-22
Perancangan Alat Pendekripsi Kekeruhan dan Pengurasan Air Riama Sibarani, Ferry	23-30
Rancang Bangun Sistem Keamanan Rumah Berbasis <i>Internet Of Things</i> (Iot) Mico Ardana, Bosar Panjaitan, Teguh Budi Santoso	31-36
Aplikasi Pembelajaran Matematika Berbasis android Studi kasus: siswa kelas ii SD Negeri 01 Pagi Kembangan Selatan Jakarta Barat Kiki Kusumawati, Prionggo Hendradi, Muhammad Alif Fauzi	37-44
Sistem Informasi Penilaian Kinerja Pegawai Berbasis Web Pada Pt.PermataIndonesia (Studi Kasus : Kantor Cabang Mayestik Jakarta Selatan) Wawan Kurniawan , Nurul Chafid, Indah Kurniati	45-49
Perancangan Sistem Informasi Pengarsipan Dokumen Di Universitas Satya Negara Indonesia Prionggo Hendradi, Khey Khey Rakhmawati Dewi	50-59
Sistem Informasi Geografis Berbasis Location Based Service Untuk Pencarian Wilayah Krisis Pangan Istiqomah Sumadikarta, Odi Kurniadi	60-72
Evaluasi Kinerja Tata Kelola Teknologi Informasi Dengan Framework Cobit. Studi kasus di sistem informasi Perikanan (SIP) Pada Direktorat Sumber Daya Ikan, Direktorat Jenderal Perikanan Tangkap Departemen Kelautan Dan Perikanan Wahyu Fajar Arinto, Agung Priambodo	73 -83

RANCANG BANGUN APLIKASI KRIPTOGRAFI DENGAN METODE *RIVEST CODE 4 (RC4)* UNTUK PENGAMANAN FILE DOKUMEN BERBASIS WEB
APLIKASI DATA MINING UNTUK CLUSTERING DAERAH PENYEBARAN COVID-19 DI DKI JAKARTA MENGGUNAKAN ALGORITMA K-MEANS

Habibi Kholil Al Hanif, Turkhamun Adi Kurniawan, T.W. Wisjhnuadji

Fakultas Teknik Informatika Universitas Satya Negara Indonesia, Fakultas Teknik Informatika Universitas Satya Negara Indonesia , Fakultas Teknik Informasi Universitas Budi Luhur
habibikalhanif@gmail.com, t.adikurniawan@usni.ac.id, 1wisjhnuadji@budiluhur.ac.id

ABSTRAK

Penyakit Koronavirus 2019 (*coronavirus disease2019*, disingkat *COVID-19*) adalah penyakit menular yang disebabkan oleh *SARS-Cov-2*, salah satu jenis korona virus. Penderita Covid-19 dapat mengalami demam, batuk kering, dan kesulitan bernapas. Infeksi menyebar dari satu orang ke orang lain melalui percikan (*droplet*) dari saluran pernapasan yang dihasilkan saat batuk atau bersin. Banyaknya jumlah penduduknya hingga tahun 2019 mencapai 11.063.324 jiwa yang tersebar di 6 Kota terdiri dari 44 Kecamatan serta 267 Kelurahan membuat Covid-19 mudah menyebar. Untuk dapat melihat daerah penyebaran Covid-19 maka perlu dibuatkan pengelompokan berdasarkan atribut yang dipakai yang terdiri dari Kasus Suspek, Kasus *Probable*, Kontak Erat, Kasus Konfirmasi dan Meninggal. Pada penelitian ini, untuk melakukan klasterisasi data digunakan metode K-Means dan metode pengukuran jarak Euclidean. Penelitian ini menghasilkan prototipe aplikasi pengelompokan data persebaran pasien Covid-19. Hasil dari implementasi Algoritma K-Means yaitu cluster penyebaran Covid-19 di DKI Jakarta dibagi dalam 3(tiga) cluster yaitu cluster 1, cluster 2 dan cluster 3. Cluster 1 merupakan zona kasus sedang, Cluster 2 merupakan zona kasus tinggi dan Cluster 3 merupakan zona kasus rendah.

Kata kunci: *Covid-19, Algoritma K-MEANS*

ABSTRACT

Coronavirus Disease 2019 (coronavirus disease2019, abbreviated as COVID-19) is an infectious disease caused by SARS-Cov-2, a type of coronavirus. Covid-19 patients can experience fever, dry batik, and difficulty breathing. The infection spreads from one person to another through a splash (droplet) from the respiratory tract produced when coughing or sneezing. The number of residents until 2019 reached 11,063,324 people spread across 6 cities consisting of 44 districts and 267 urban villages, making Covid-19 easy to spread. To be able to see the area of spread of Covid-19, it is necessary to group based on the attributes used consisting of Suspect Cases, Probable, Cases, Close Contacts, Confirmed Cases and Deaths. In this study, to cluster the data, the K-Means method and the Euclidean distance measurement method were used. This study produced a prototype application for grouping data on the distribution of Covid-19 patients. The result of the implementation of the K-Means Algorithm is that the Covid-19 spread cluster in DKI Jakarta is divided into 3 (three) clusters, namely cluster 1, cluster 2 and cluster 3. Cluster 1 is a medium case zone, Cluster 2 is a high case zone and Cluster 3 is a low case zone.

Keywords: *Covid-19, K-MEANS Algorithm*

PENDAHULUAN

Latar Belakang

Penyakit Koronavirus 2019 (*corona virus disease2019*, disingkat COVID-19) merupakan penyakit menular yang disebabkan oleh *SARS-Cov-2*, salah satu jenis korona virus [3]. Penderita Covid-19 bisa mengalami demam, batuk kering, dan kesulitan bernapas. Infeksi menyebar dari satu orang ke orang lainnya melalui percikan (*droplet*) dari saluran pernapasan yang dikeluarkan saat batuk atau bersin. Beberapa penelitian juga membuktikan bahwa virus corona dapat bertahan beberapa jam di benda-benda yang disentuh oleh penderita. Jika seseorang menyentuh benda yang sudah terkontaminasi dengan droplet yang mengandung virus dan menyentuh mata, hidung atau mulut, maka orang itu dapat tertular COVID-19. Oleh karena itu, salah satu pencegahan penularan virus ini adalah dengan menjaga jarak fisik, menggunakan masker dan senantiasa menjaga kesehatan fisik dan non-fisik. Menurut data dari Badan Pusat Statistik (BPS) banyaknya jumlah penduduk di DKI Jakarta hingga tahun 2021 mencapai 10,609,681 jiwa yang tersebar di 6 Kota terdiri dari 44 Kecamatan serta 267 Kelurahan. Banyaknya penduduk dan padatnya aktivitas di DKI Jakarta membuat virus covid-19 dapat menular dengan cepat dan tidak terkendali. Provinsi DKI Jakarta merupakan penyumbang terbesar kasus COVID19 di Indonesia. Berdasarkan data dari Gugus Tugas Percepatan Penanganan COVID-19. Beberapa penelitian telah melakukan klasterisasi kasus Covid-19 di tingkat dunia dan beberapa negara, namun

diperlukan analisis lebih mendalam untuk kasus penyebaran Covid-19 di Indonesia, khususnya di DKI Jakarta. Sejak diumumkan kasus pertama Covid-19 di DKI Jakarta, data-data yang berkaitan dengan jumlah kasus Covid-19 sudah mulai dikumpulkan (rekap) seperti jumlah kasus positif, PDP, ODP dan meninggal dunia termasuk penyebarannya di seluruh daerah di DKI Jakarta. Data jumlah kasus Covid-19 di DKI Jakarta dapat dimanfaatkan untuk memprediksi jumlah kasus dan membentuk cluster zona terbaru berdasarkan jumlah Kasus Suspek, Probable, Meninggal, Kontak Erat dan Kasus Konfirmasi atau positif dengan menggunakan algoritma clustering K-means. Salah satu algoritma yang termasuk dalam clustering adalah K-means. K-means adalah algoritma yang mengelompokkan data dengan memaksimalkan tingkat kemiripan antar cluster. Dengan dibuatnya data persebaran virus Corona yang sudah dikelompokan atau dicluster menggunakan Algoritma ini diharapkan dapat membantu pengambilan keputusan yang tepat dalam mengurangi persebaran virus Covid.

Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, peneliti merumuskan masalah yaitu Bagaimana cara melakukan Clustering terhadap penyebaran penyakit Covid-19 (*coronavirus disease2019*) di DKI Jakarta dengan menggunakan Algoritma K-means?

Tujuan Penelitian

Berdasarkan rumusan masalah yang penulis uraikan sebelumnya, tujuan dari penelitian ini yaitu membuat Aplikasi Data Mining Untuk Clustering Daerah Penyebaran *COVID-19* Di DKI Jakarta Menggunakan Algoritma K-means.

Manfaat Penelitian

Adapun manfaat penelitian ini diantaranya yaitu:

Manfaat bagi peneliti

1. Menambah ilmu untuk peneliti di bidang IT dalam merancang serta menghasilkan sebuah sistem perangkat lunak.
2. Dapat mempraktikan langsung ilmu yang dipelajari selama diperkuliahan
3. Sebagai salah satu syarat kelulusan mahasiswa Teknik Informatika di Universitas Satya Negara Indonesia

Manfaat bagi instansi

Sebagai pertimbangan atau pengambil keputusan dalam menangani penyebaran covid 19 berdasarkan hasil pengelompokan atau clustering dengan aplikasi yang sudah dibuat.

Manfaat bagi universitas

Penelitian ini diharapkan dapat digunakan untuk menambah referensi sebagai bahan penelitian lanjutan yang lebih mendalam pada masa yang akan datang

Penelitian Terdahulu

Pada penelitian (Teguh Budi Santoso ,Ahmad Sofian Hadi, 2019, dalam JURNAL ILMIAH FIFODOI dengan judul “Aplikasi Data Mining untuk Clustering Daerah Penyebaran Penyakit Diare di DKI Jakarta Menggunakan Algoritma K-MEANS”.

Pada penelitian (Achmad Solichin,Khansa Khairunnisa, 2020, dalam Fountain of Informatics Journal Volume 5, No. 2, November 2020). dengan judul “Klasterisasi Persebaran Virus Corona (Covid-19) Di DKI Jakarta Menggunakan Metode K-Means”,

Pada penelitian (Panjaitan, B., Kom, S., Kom, M., & Turkhamun Adi Kurniawan, 2016) dengan judul “CLASTERING DATA AKADEMIK MAHASISWA FAKULTAS TEKNIK USNI DENGAN ALGORITMA K-MEANS”).

Data Mining

Penggalian data atau dalam bahasa Inggris disebut data mining merupakan ekstraksi pola yang menarik dari data dengan jumlah besar. Suatu pola dikatakan menarik apabila pola tersebut tidak mudah, tidak diketahui sebelumnya, dan bermanfaat. Pola yang diberikan harus mudah dipahami. Data mining mempunyai beberapa nama lain, meskipun definisi eksaknya tidak sama, seperti KDD (Knowledge Discovery in Database), analisis pola, arkeologi data, pemanenan info, serta intelegensi bisnis. Data mining ialah proses analitik yang didesain untuk memeriksa sejumlah data yang besar dalam mencari suatu pengetahuan tersembunyi yang berharga serta konsisten [4]. Tujuan dari data mining ialah mencari trend atau pola yang diinginkan pada database besar untuk membantu pada pengambilan keputusan di waktu kedepannya.

Clustering

Clustering adalah suatu proses pengelompokan rekam, observasi, atau mengelompokkan kelas yang mempunyai kemiripan objek [8]. Yang membedakan clustering dengan klasifikasi yaitu tidak adanya variable sasaran dalam melakukan pengelompokan pada proses clustering. Clustering seringkali dilakukan untuk Langkah pertama pada proses data mining.

Algoritma K-means Clustering

Algoritma K-means adalah salah satu algoritma analisis kluster (*cluster analysis*) non hirarki yang bertujuan mendapatkan data dengan memaksimalkan kemiripan karakteristik dalam cluster. Prinsip utama K-Means adalah menyusun k prototype atau pusat massa (centroid) dari sekumpulan data berdimensi. Sebelum diterapkan proses algoritma K-means, dokumen akan di preprocessing terlebih dahulu. Kemudian dokumen direpresentasikan sebagai vektor yang memiliki term dengan nilai tertentu.

Covid-19

Korona virus 2019 (coronavirus disease 2019, disingkat COVID-19) ialah salah satu jenis korona virus yang merupakan penyakit menular dan disebabkan oleh SARS-CoV-2 [3]. Korona virus adalah kumpulan virus yang menginfeksi sistem pernapasan. Pada banyak kasus, virus ini hanya menyebabkan infeksi pernapasan ringan, Seperti flu. Namun, virus ini juga bisa menyebabkan infeksi pernapasan berat, seperti infeksi paru-paru (pneumonia). Penderita Covid-19 bisa mengalami batuk kering, demam serta sulit bernapas. Infeksi menyebar dari satu orang ke orang lain lewat percikan (droplet) melalui saluran pernapasan yang dikeluarkan ketika batuk atau bersin.

Metode Elbow

Metode Elbow adalah cara yang dipergunakan untuk menghasilkan informasi dalam menentukan jumlah cluster terbaik menggunakan cara melihat persentase hasil dari perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik [10]. Pada metode Elbow nilai cluster terbaik yang akan dipilih dari nilai Sum of Square Error (SSE) yang mendapati penurunan signifikan serta berbentuk siku.

Indikator Level Asesmen Situasi Tingkat Penularan COVID-19

Menurut Kemenkes RI, Panji Fortuna, dr, MPH, Tenaga Ahli Kementerian Kesehatan menyampaikan, "Asesmen Level Situasi Pandemi" merupakan dasar pengetatan atau pelonggaran upaya kesehatan masyarakat & pembatasan sosial. Tingkatan situasi pandemi menunjukkan kapasitas respon terhadap tingkat transmisi. Kapasitas respon yang memadai di tingkat transmisi rendah dapat menjadi tidak memadai jika transmisi meningkat. Dari perspektif ini, tujuan pengendalian pandemi adalah untuk mencapai tingkat situasi yang serendah mungkin (tingkat situasi nol).

METODE PENELITIAN

Waktu dan Tempat Penelitian

Waktu penelitian dilakukan selama 3 bulan dimulai dari bulan Februari 2022 sampai dengan bulan April 2022. Tempat penelitian dilaksanakan di Dinas Cipta Karya Tata Ruang dan Pertanahan Pemerintahan Provinsi DKI Jakarta Jl. Taman Jati Baru No.17, RT.17/RW.1, Cideng, Kecamatan Gambir, Kota Jakarta Pusat.

Metode Pengumpulan Data

Dataset

Dataset adalah sebuah kumpulan data yang berasal dari informasi-informasi pada masa lalu dan siap untuk dikelola menjadi sebuah informasi baru. Dataset merujuk pada kumpulan data. Secara umum, dataset berisi lebih dari satu variabel dan menyangkut suatu topik tertentu. Penelitian clustering Persebaran Virus Corona di DKI Jakarta dilakukan memakai dataset Kasus Suspek, Probable, Meninggal, Kontak Erat dan Kasus Konfirmasi atau positif Covid-19 DKI Jakarta Per-Kelurahan Tahun 2021. Data yang diperoleh dari situs Jakarta Open Data ini di olah oleh Pemerintah Provinsi DKI Jakarta.

Tabel 1 Data field Covid-19

NAMA FIELD	TIPE	KETERANGAN
ID_KECAMATAN	INT	Kode unik dari setiap kecamatan
NAMA_KOTA	VACHAR	Nama dari setiap kota di DKI
NAMA_KECAMATAN	VACHAR	Nama dari setiap kecamatan di DKI
SUSPEK	INT	Jumlah suspek
MENINGGAL	INT	Jumlah meninggal
KONTAK ERAT	INT	Jumlah kontak erat
PROBABLE	INT	Jumlah Probable
POSITIF	INT	Jumlah positif

Studi Pustaka

Metode yang dilakukan ini mempelajari jurnal ilmiah serta laporan penelitian yang berkaitan dengan clustering penyebaran covid-19, data mining dan Algoritma clustering K-means. Pada tahap inipeneliti mempelajari dan memahami berkaitan dengan sistem yang akan diterapkan pada penelitian

PEMBAHASAN

Penerapan Metode

Penerapan metode algoritma K-means dalam penelitian Clustering Persebaran Virus Corona di DKI Jakarta dilakukan beberapa tahap, yaitu:

1. Berikan nilai awal pusat cluster sejumlah k (banyaknya cluster). pusat cluster diambil dari objek yang dipilih secara acak yang menjadi perwakilan centroid.
2. Tentukan banyaknya setiap titik data (objek) dengan medoid yang paling mirip dengan menggunakan ukuran jarak serta menghitung cost. Berikut persamaan jarak Euclidian Distance:

$$D_{(i,j)} = \sqrt{\sum_{k=1}^n (X_{ik} - C_{jk})^2}$$

yang mana

D = jarak klaster

X_{ik} = nilai data (i,k)

C_{jk} = nilai centroid (j,k)

n = jumlah klaster

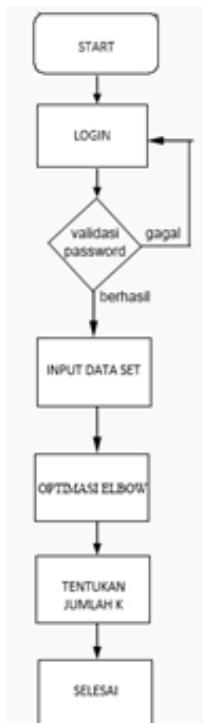
Gambar 1 Persamaan Euclidian Distance

3. Kelompokkan data ke dalam cluster yang mempunyai jarak yang paling pendek.
4. Hitung pusat cluster.
5. Ulangi Langkah ke 2 - ke 4 hingga sudah tidak ada lagi data yang berpindah ke cluster yang lain.

Prinsip Kerja Sistem

Prinsip kerja dari sistem yang dirancang ini ialah untuk melakukan clustering berdasarkan data set menggunakan algoritma K-means Sistem ini mempermudah perhitungan clustering covid-19 di DKI Jakarta dalam jumlah besar.

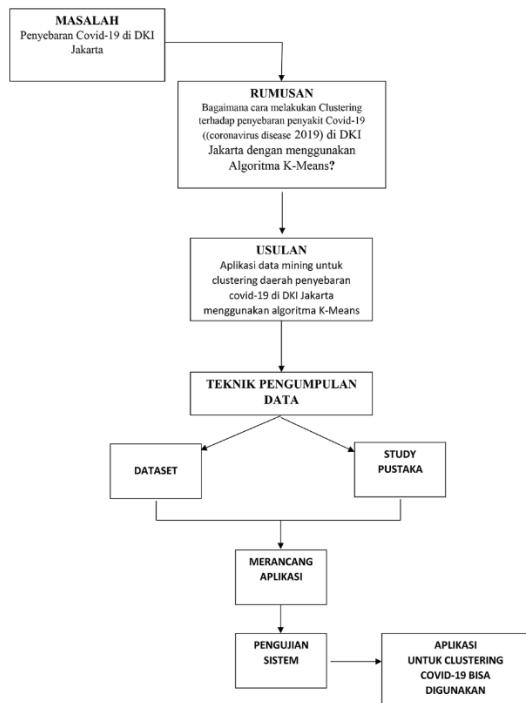
Flowchart



Gambar 2 Flowchart

Penjelasan diagram Flowchart diatas adalah user melakukan login untuk masuk kedalam sistem. Jika user memasukan username dan password dengan benar, maka login berhasil. Jika user salah memasukan username dan password maka user akan dibawa lagi kehalaman login. Setelah user berhasil login, masukan data set atau jumlah data persebaran yang akan diclustering. Selanjutnya di optimasi elbow, masukan *range* cluster dan jumlah maksimal pengulangan untuk menentukan cluster yang optimal. Tentukan jumlah K atau jumlah cluster yang optimal berdasarkan perhitungan elbow dan clusterisasi selesai.

Kerangka Berpikir



Gambar 3 Kerangka Berpikir

Berdasarkan kerangka berpikir diatas dijelaskan bahwa permasalahan yang menjadi dasar perencanaan ini adalah penyebaran covid-19 di DKI Jakarta. Maka dari itu dibuatkan rumusan masalah yaitu bagaimana cara melakukan clustering terhadap penyebaran virus covid-19 di DKI Jakarta dengan menggunakan algoritma K-Means. Dari rumusan masalah yang ada, diberikan usulan yaitu membuat aplikasi data mining untuk clustering daerah penyebaran covid-19 di DKI Jakarta menggunakan algoritma K-Means. Setelah usulan diberikan, peneliti melakukan teknik pengumpulan data, yaitu pengumpulan data set dan studi pustaka yang bertujuan untuk mendapatkan data data yang dibutuhkan untuk clustering covid-19 di DKI Jakarta. Tahap selanjutnya yaitu merancang aplikasi dan pengujian sistem agar aplikasi berjalan dengan baik. peneliti tertarik membuat aplikasi untuk clustering covid-19 untuk bisa melihat penyebaran covid 19 di DKI Jakarta

Rancangan Antar Muka

Dalam penelitian ini dilakukan pengembangan sebuah rancangan antar muka aplikasi clustering data Covid-19 yang menerapkan metode algoritma K-means sebagai gambaran peneliti. Rancangan antar muka ini dikembangkan menggunakan bahasa pemrograman PHP dan MySQL.

Tampilan rancangan login
Gambar 4 Tampilan Login

Rancangan awal halaman tampilan login dimana diberikan kolom username dan password untuk bisa masuk atau log in kedalam website. Terdapat 2 pilihan tombol. Pertama tombol login untuk masuk kedalam aplikasi. Kedua tombol register untuk mendaftar.

Tampilan rancangan beranda
Gambar 5 Tampilan rancangan beranda

Halaman rancangan beranda adalah halaman pertama saat pengguna melakukan login. Dihalaman beranda terdapat info singkat mengenai covid-19. Pada menu beranda pengguna bisa memulai proses clustering dengan memilih menu data kasus covid-19.

Tampilan rancangan data kasus

Beranda	Data Kasus COVID-19					
Data kasus COVID-19	tambah data					
Proses Clustering						
Hasil Clustering						
	KECAMATAN	PROBABLE	SUSPEK	KONTAK ERAT	POSITIF	MENINGGAL

Gambar 6 Tampilan Data Kasus

Pada rancangan tampilan data kasus diatas, ditampilkan data kasus mulai dari tanggal, kecamatan sampai meninggal. Pengguna bisa memasukan data dengan memilih tombol tambah data. Untuk lanjut ke proses berikutnya, pengguna bisa memilih menu proses clustering.

Tampilan rancangan pemilihan dan proses dataset

Beranda	Pilih data dan cluster					
Data kasus COVID-19	cari data					
Proses Clustering						
Hasil Clustering						
	KECAMATAN	PROBABLE	SUSPEK	KONTAK ERAT	POSITIF	MENINGGAL
	simpan					

Gambar 7 Tampilan Pemilihan Dataset

Pada rancangan tampilan proses clustering, data yang sudah dimasukan ke aplikasi akan ditampilkan di table data yang akan di proses. Pengguna bisa mencari data yang sudah dimasukan dengan memilih tombol cari data. Jika data sudah sesuai, pengguna bisa memulai proses dengan memilih tombol simpan.

Tampilan Hasil Clusterisasi

Beranda	Hasil Clustering	
Data kasus COVID-19	cari data	
Proses Clustering		
Hasil Clustering		
	KECAMATAN	CLUSTER
	JUMLAH CLUSTERING	
	CLUSTER	JUMLAH

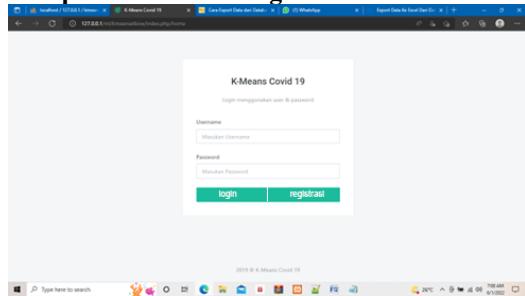
Gambar 8 Tampilan Hasil Cluster

Tampilan hasil clusterisasi berfungsi sebagai hasil laporan akhir dari clustering. Pada hasil clustering di tabel kecamatan akan menampilkan semua kecamatan dan pada tabel cluster akan menampilkan cluster ke berapa kecamatan itu berada. Pada jumlah clustering di tabel cluster akan menampilkan jumlah cluster dan pada tabel jumlah akan menampilkan jumlah dari cluster.

Implementasi Program

Implementasi program adalah tahap mewujudkan hasil rancangan menjadi sebuah program aplikasi. Berdasarkan rancangan interface yang telah dibuat, maka akan dijelaskan mengenai hasil program dari implementasi data mining dengan menggunakan metode clustering untuk menentukan daerah penyebaran Covid-19 di DKI Jakarta.

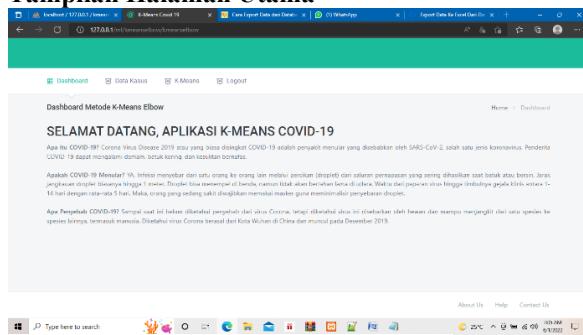
Tampilan Halaman Login



Gambar 9 Tampilan Halaman Login

Saat membuka aplikasi, pengguna langsung diarahkan pada halaman Login. Pada halaman login ini pengguna harus memasukan *username* dan *password* sebelum memulai aplikasi..

Tampilan Halaman Utama

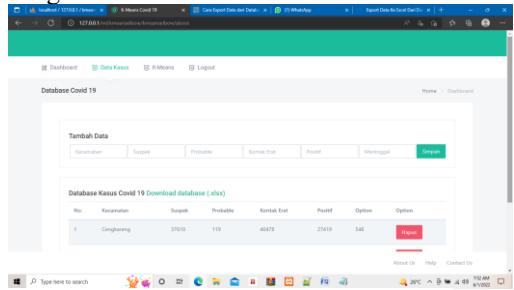


Gambar 10 Tampilan Hasil Utama

Saat login berhasil, pengguna diarahkan ke halaman Beranda seperti Tampilan Halaman Utama ini. Pada halaman ini ditampilkan penjelasan tentang kasus covid-19. Pengguna bisa melanjutkan proses memasukan data dengan memilih menu data kasus untuk lanjut ke proses selanjutnya

Tampilan Data Kasus

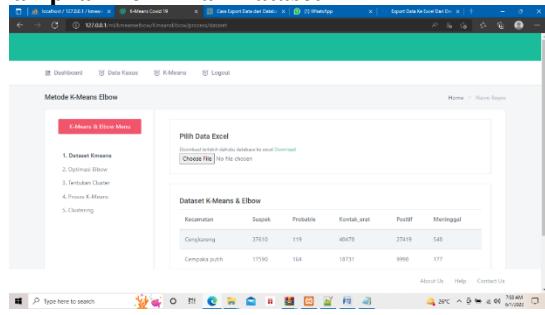
Di halaman data kasus Covid-19 seperti pada gambar 4.3 ditampilkan data kasus yang sudah dimasukan sebelumnya. Data kasus yang ditampilkan dalam bentuk tabel. Pengguna juga bisa menambah data kasus baru dengan memilih tombol “Tambah Data”.



Gambar 11 Tampilan Data Kasus

Setelah pengguna memasukan data dengan mengisi kolom kecamatan, suspek, probable, kontak erat, positif dan meninggal data akan tersimpan di database kasus covid-19 yang datanya bisa didownload berupa file xlsx atau excel. Data yang sudah dimasukan juga bisa dihapus dengan memilih tombol hapus maka data akan terhapus dan bisa melakukan tambah data ulang.

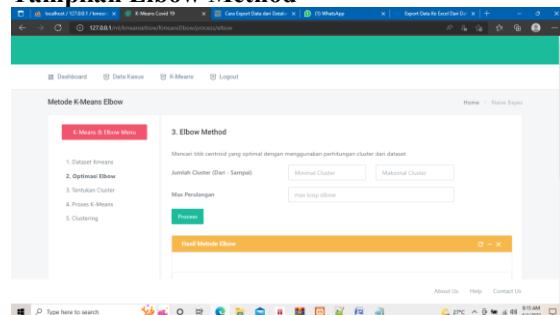
Tampilan Pemilihan Dataset



Gambar 12 Tampilan Pemilihan Dataset

Sebelum proses clustering, pengguna harus memasukkan file database kasus yang telah di download terlebih dahulu. File database kasus bisa didapatkan dari halaman sebelumnya yaitu data kasus atau bisa memasukan file yang didapat dari eksternal dengan ketentuan yang sama seperti file database yaitu nama kecamatan, suspek, probable, kontak erat, positif dan meninggal. Jika data berhasil dimasukan, maka data akan terlihat di tabel dataset K-means.

Tampilan Elbow Method



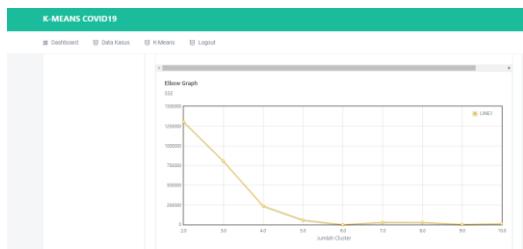
Gambar 13 Tampilan Elbow Method

tampilan untuk mencari titik centroid yang optimal menggunakan perhitungan cluster dari dataset dengan mengisi kolom minimal dan maximal cluster kemudian menekan tombol proses maka data akan ditampilkan dalam bentuk grafik Pada hasil algoritma elbow didapatkan

bawa cluster 2 mempunyai nilai sebesar 1304720.6989474 dengan selisih 1304720. Cluster 3 mempunyai nilai sebesar 498485.91346154 dengan selisih 806234.78548583 dan seterusnya sampai cluster 9 seperti yang di jelaskan pada tabel hasil algoritma elbow dibawah ini

Tabel 2 Hasil Algoritma Elbow

Cluster	Nilai	Selisih
2	1304720.6989474	1304720.6989474
3	498485.91346154	806234.78548583
4	261365.15294118	237120.76052036
5	203058.89090909	58306.262032086
6	203371.58863636	312.69772727287
7	172985.5	30386.088636364
8	143411.91111111	29573.588888889
9	146895.06111111	3483.15
10	135594.66666667	11300.394444444

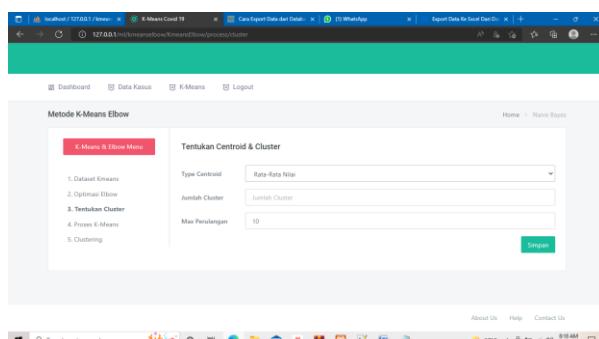


Gambar 14 Tampilan Elbow Graph

Pada tampilan diagram elbow menunjukkan sisi vertikal menunjukkan nilai SSE (square sum of error) yang bernilai dari 250.000 sampai 1.500.000 dan sisi horizontal menunjukkan jumlah cluster dari 2 sampai 9.

Tampilan Menu Cluster

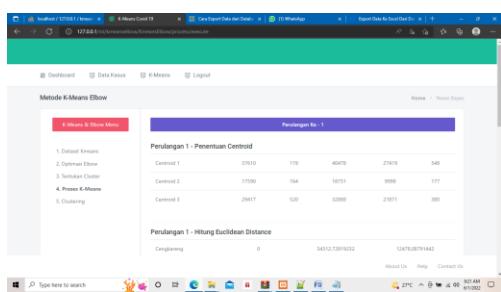
Pada gambar dibawah merupakan tampilan halaman untuk menentukan centroid dan cluster kemudian mengisi kolom jumlah cluster dan memilih maximal pengulangan kemudian tekan tombol simpan. Pada kasus ini dimasukan jumlah cluster adalah 3 seperti contoh gambar dibawah ini.



Gambar 15 Tampilan Menu Cluster

Menampilkan Proses K-Means

Halaman ini menampilkan hasil penentuan centroid sebagai berikut. Hasil dari perulangan ke-1 yaitu centroid 1 himpunan 1 berjumlah 3710, himpunan 2 berjumlah 119, himpunan 3 berjumlah 40476, himpunan 4 berjumlah 27419, himpunan 5 berjumlah 548 dan seterusnya sampai centroid 3.



Gambar 16 Tampilan Proses K-Means

Penentuan Centroid

Sebelum menentukan centroid terlebih dahulu memberikan jumlah cluster atau nilai k untuk mendapatkan titik pusat pusat dari setiap himpunan yang ada seperti tabel dibawah. Centroid atau titik pusat 1 memiliki nilai suspek 26.33, probable 229.46, kontak erat 29.46, positif sebesar 17.73, meninggal sebesar 314.8 dan seterusnya sampai centroid 3.

Tabel 3 penentuan centroid

Centroid	SUSPEK	PROBABLE	KONTAK ERAT	POSITIF	MENINGGAL
Centroid 1	26.33	229.46	29.46	17.73	314.8
Centroid 2	22.33	129.93	22.2	96.46	244.13
Centroid 3	24.5	148.07	23.57	14.42	258.07

Hitung Euclidean Distance

Melakukan perhitungan jarak antar cluster dengan menggunakan rumus Euclidean distance. Dalam penelitian ini menentukan titik terdekat dapat dilihat berdasarkan cluster 1, 2 dan cluster 3 dimana nilai terkecil akan dijadikan titik terdekat cluster yang disajikan pada tabel dibawah

Tabel 4 jarak Euclidean Distance

KECAMATAN	CLUSTER 1	CLUSTER 2	CLUSTER 3
CAKUNG	298.9345376128	421.7002779753	393.16179745697
CEMPAKA PUTIH	153.52391055178	115.60244134293	83.325002984992
CENGKARENG	258.6420778691	312.74919308893	292.38316173355
CILANDAK	81.049724106735	95.73535281064	32.981364558842
CILINCING	31.283506481496	143.92842356919	97.034892735362
DST..			

Kecamatan cakung dengan cluster 1 mempunyai nilai 298.9345376128, dengan cluster 2 mempunyai nilai 421.7002779753, dan cluster 3 mempunyai nilai 393.16179745697 dan seterusnya sampai 44 kecamatan.

Perulangan 3 - Hasil Cluster

Setelah melakukan perhitungan sampai iterasi ke tiga, diperoleh sudah tidak ada perubahan dari setiap kecamatan yang di cluster dan perhitungan bisa diberhentikan.

Menampilkan Hasil Cluster

halaman yang menampilkan hasil dari cluster K-Means seperti pada gambar dibawah

The screenshot shows a web application for K-Means clustering. The sidebar on the left lists navigation options: 1. Dataset Kmeans, 2. Optimalis Elbow, 3. Tentukan Cluster, 4. Proses K-Means, and 5. Clustering. The main content area displays two tables. The first table, 'Hasil Cluster KMeans', shows the assignment of three objects (Cengkareng, Cempaka putih, Cakung) to three clusters (1, 2, 3). The second table, 'Jumlah Cluster', shows the count of objects assigned to each cluster (Cluster 1: 1, Cluster 2: 1, Cluster 3: 1). The application is running on a local host (127.0.0.1) and is titled 'K-Means Covid 19'.

Gambar 17 Tampilan Hasil Cluster

Berdasarkan gambar hasil cluster diperoleh pada gambar hasil cluster kmeans memiliki hasil cengkareng berada di cluster 1, cempaka putih di cluster 2 dan cakung 1. Dan pada tabel jumlah cluster diperoleh cluster 1, 2, 3 serta jumlah dari seluruh cluster. Untuk lebih lengkap dapat di lihat pada tabel hasil cluster kecamatan.

Jumlah Cluster

Berdasarkan perhitungan yang sebelumnya dilakukan didapatkan bahwa cluster 1 berjumlah 16 kecamatan, cluster 2 berjumlah 26 kecamatan dan cluster 3 berjumlah 2 kecamatan. Untuk lebih jelasnya bisa dilihat pada tabel 4.9 dibawah ini

CLUSTER		
CLUSTER 1	CLUSTER 2	CLUSTER 3
CAKUNG	CEMPAKA PUTIH	KEP. SERIBU SELATAN
CENGKARENG	CILANDAK	KEP. SERIBU UTARA
CILINGCING	GAMBIR	
CIPAYUNG	GROGOL PETAMBURAN	
CIRACAS	JOHAR BARU	
DUREN SAWIT	KALI DERES	
JAGAKARSA	KEBAYORAN BARU	
JATINEGARA	KELAPA GADING	
KEBAYORAN LAMA	KEMBANGAN	
KEBON JERUK	KOJA	
KEMAYORAN	MAKASAR	
KRAMAT JATI	MAMPANG PRAPATAN	
PASAR MINGGU	MATRAMAN	
PASAR REBO	MENTENG	
PULO GADUNG	PADEMANGAN	
TANJUNG PRIOK	PALMERAH	
	PANCORAN	
	PENJARINGAN	
	PESANGGRAHAN	
	SAWAH BESAR	
	SENEN	
	SETIA BUDI	
	TAMAN SARI	
	TAMBORA	
	TANAH ABANG	
	TEBET	
C1 16 KECAMATAN	C2 26 KECAMATAN	C3 2 KECAMATAN

Berdasarkan hasil cluster pada tabel diatas, diperoleh c1 sebanyak 16 kecamatan yaitu cakung, cengkareng, cilingcing dan Seterus sampai tanjung priok. Untuk c2 diperoleh sebanyak 26 kecamatan yaitu cempaka putih, cilandak, Gambari dan seterusnya sampai tebet. Untuk c3 sebanyak 2 kecamatan yaitu Kep. Seribu Utara dan Kep. Seribu Selatan

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil pengujian dan uji coba pada penerapan metode k-means untuk melakukan pengelompokan atau clusterisasi kasus Covid-19 di DKI Jakarta, dapat disimpulkan bahwa aplikasi clustering metode k-means dapat digunakan dan berjalan dengan baik. jumlah cluster yang direkomendasikan berdasarkan perhitungan nilai SSE adalah 9 cluster. Tetapi pada praktiknya, jumlah cluster tersebut bisa ditentukan sendiri oleh pengguna program aplikasi berdasarkan hasil pertimbangan-pertimbangan tertentu. Jika jumlah 9 cluster sangat banyak, penelitian ini merekomendasikan menggunakan 3 (tiga) cluster dikarenakan jumlah tersebut memiliki kualitas yang cukup baik berdasarkan hasil perhitungan metode Elbow dan menempati urutan kedua pada SSE setelah 9 cluster.

Pada clusterisasi dengan 3 (tiga) cluster, diperoleh pengelompokan C1 sebanyak 16 kecamatan, C2 sebanyak 26 kecamatan, dan C3 sebanyak 2 kecamatan. Berdasarkan Indikator Level Asesmen Situasi Tingkat Penularan COVID-19 dan Jumlah kasus COVID-19 yang sudah dicluster dari jumlah penduduk DKI Jakarta per 2021 sebanyak 10.609.681 jiwa, C2 merupakan zona kasus yang tinggi, C1 merupakan zona kasus sedang dan C3 merupakan zona kasus rendah.

Saran

Untuk penelitian selanjutnya, clusterisasi bisa dikembangkan dengan menggunakan berbagai metode clusterisasi yang lainnya seperti Fuzzy K-Means, Fuzzy C-Means, K-Medoids, dan sebagainya dengan mempertimbangkan data yang cocok dan sesuai dan untuk hasil dari clustering K-Means ini bisa dikembangkan lagi agar bisa lebih bermanfaat untuk di bidang lainnya

DAFTAR PUSTAKA

Abdulloh, R. (2015). Web programming is easy. Jakarta: PT. Elex Media Komputindo, 7.

- Bu'ulolo, E., & Purba, B. (2021). Algoritma Clustering Untuk Membentuk Cluster Zona Penyebaran Covid-19. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*
- Calisher, C. H., Carroll, D., Colwell, R., Corley, R. B., Daszak, P., Drosten, C., & Turner, M. (2021). Science, not speculation, is essential to determine how SARS-CoV-2 reached humans. *The Lancet*, 398(10296), 209-211.
- Florin, G. (2011). *Data mining: concepts, models and techniques*. Springer-Verlag.
- Haviluddin, H. (2016). *Memahami Penggunaan UML (Unified Modelling Language)*. Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer. Jakarta Satu. (2022). Open Data Covid-19 Provinsi DKI Jakarta. <https://riwayat-file-covid-19-dki-jakarta-jakartagis.hub.arcgis.com/>. 1 Januari 2022.
- Keputusan Menteri Kesehatan Republik Indonesia Nomor Hk.01.07/Menkes/413/2020. (2020). Tentang Pedoman Pencegahan Dan Pengendalian Coronavirus Disease 2019 (Covid-19).
- Larose, D. T. (2015). *Data mining and predictive analytics*
- Madcoms, T. (2016). *Pemrograman PHP dan MySQL untuk pemula*. Yogyakarta: CV Andi Offset.
- Rahman, A. T. (2017). Coal trade data clustering using K-means (case study Pt. Global Bangkit Utama). *ITSMART: Jurnal Teknologi dan Informasi*, 6(1), 24-31.
- Santoso, T. B. (2019). Aplikasi Data Mining untuk Clustering Daerah Penyebaran Penyakit Diare di DKI Jakarta Menggunakan Algoritma K-MEANS. *J. Ilm. FIFO*, 11(2), 131.
- Saputro, R. A., & Widodo, P. P. (2014). Komparasi Algoritma Klasifikasi Data Mining Untuk Memprediksi Penyakit Tuberculosis (Tb): Studi Kasus Puskesmas Karawang Sukabumi.
- Noviyanto, N. (2020). Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan Negara di Benua Asia. *Paradigma Jurnal Komputer dan Informatika*, 183-188.
- Solichin, A., & Khairunnisa, K. (2020). Klasterisasi persebaran virus Corona (Covid-19) di DKI Jakarta menggunakan metode K-Means. *Fountain of Informatics Journal*, 5(2), 52-59.
- Panjaitan, B., Kom, S., Kom, M., & Turkhamun Adi Kurniawan (2016). CLASTERING DATA AKADEMIK MAHASISWA FAKULTAS TEKNIK USNI DENGAN ALGORITMA K-MEANS.
- Purwasandina, Y., Gunadhi, E., & Fatimah, D. D. S. (2015). Rancang Bangun Sistem Informasi Pengelolaan Data Rujukan Pasien Untuk Puskesmas Pembangunan Garut. *Jurnal Algoritma*, 12(2), 393-397.
- Suendri, S. (2019). Implementasi Diagram UML (Unified Modelling Language) Pada Perancangan Sistem Informasi Remunerasi Dosen Dengan Database Oracle (Studi Kasus: UIN Sumatera Utara Medan). *Algoritma: Jurnal Ilmu Komputer Dan Informatika*, 2(2), 1.
- Widodo, P. P., & Prabowo, H. (2011). Menggunakan uml. *Bandung: Informatika*, 19, 393-403.
- X. Wu et al., "Top 10 algorithms in data mining," in *Knowledge and Information Systems*, 2008, vol. 14, no. 1, hal-3