

APLIKASI DATA MINING UNTUK CLUSTERING DAERAH PENYEBARAN
COVID-19 DI DKI JAKARTA MENGGUNAKAN ALGORITMA K-MEANS

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar

SARJANA KOMPUTER

Program Studi Teknik Informatika



DISUSUN OLEH:

NAMA :HABIBI KHOLIL AL HANIF

NIM : 011701503125063

FAKULTAS TEKNIK

UNIVERSITAS SATYA NEGARA INDONESIA

JAKARTA

2022

DATA MINING APPLICATION FOR CLUSTERING *COVID-19* SPREAD
AREAS IN DKI JAKARTA USING THE K-MEANS ALGORITHM

Proposed As One Of The Requirements To Obtain A BACHELOR'S DEGREE IN COMPUTER

Informatics Engineering Study Program



JAKARTA

2022

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 *Probalbe*, Kontak Erat, Kasus Konfirmasi dan Meninggal. Pada penelitian ini, untuk melakukan klusterisasi 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 batuk, 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*