

**ANALISIS DAN PERBANDINGAN KELOMPOK PASIEN COVID-19  
BERDASARKAN KOMORBIDITAS MENGGUNAKAN *K-MEANS*  
*CLUSTERING***

**Hilma Fitri Solehah**

**ABSTRAK**

Pandemi COVID-19 yang terjadi sejak awal tahun 2020 merupakan sebuah permasalahan dan tantangan baru yang dirasakan oleh dunia. Penularan infeksi virus SARS-CoV-2 sangat mudah dan terjadi dengan cepat. Salah satu faktor risiko yang rentan terinfeksi COVID-19 adalah faktor penyakit penyerta/ komorbiditas yang ada pada masyarakat. Untuk dapat mengidentifikasi kelompok komorbid apa saja yang banyak ditemukan pada pasien COVID-19, penelitian ini melakukan kegiatan analisis dan pemodelan menggunakan algoritma data mining *K-Means Clustering* yang akan menghasilkan cluster pasien COVID-19. Adapun atribut data yang digunakan yakni nilai tengah usia, jenis komorbiditas, serta persentase penyakit melihat ukuran populasinya. Data yang akan diolah didapatkan dari sumber valid Covid Analytics yang berisikan rangkuman data riset medis pasien COVID-19 dari berbagai Negara. Cluster yang dihasilkan berjumlah 3 cluster dengan perbedaan derajat dari ringan hingga berat. Terdapat 321 data (39%) kelompok pasien COVID-19 yang diolah termasuk pada cluster pertama dengan rata-rata usia 48 tahun dan komorbid terbanyak ditemukan yaitu penyakit diabetes. Cluster kedua terdiri dari 370 data (44%) dengan rata-rata usia 64 tahun dan penyakit diabetes merupakan komorbid terbanyak. Sedangkan untuk cluster ketiga terdiri dari 140 data (17%) dengan rata-rata usia 68 tahun dan komorbid terbanyak yaitu penyakit Hipertensi. Analisis menggunakan *K-Means Clustering* ini dapat memberikan gambaran kelompok pasien dengan derajat yang berbeda serta bentuk penanganan yang akan berbeda pula.

**Kata Kunci:** Covid-19, Komorbiditas, *K-Means Clustering*

# **ANALYSIS AND COMPARISON OF COVID-19 PATIENTS BASED ON COMORBIDITY USING K-MEANS CLUSTERING**

**Hilma Fitri Solehah**

## **ABSTRACT**

The COVID-19 pandemic that has occurred since the beginning of 2020 is a new problem and challenge that is felt by all around the world. The transmission of the SARS-CoV-2 virus infection is very easy and fast to spread among us. Comorbidity is one of the risk factors that is susceptible to the virus infection. To identify any comorbid groups that can be found in COVID-19 patients, an analysis study using data mining method with K-Means Clustering algorithm will help to do grouping by the clustering results from the patients based on the proportion of comorbidities and age ranges that found in several populations. The data to be processed is obtained from a valid source Covid Analytics which contains a summary of medical research data for COVID-19 patients from various countries. The cluster results are 3 clusters with different clinical degrees from mild to severe. There are 321 data (39%) of the COVID-19 patients group who were treated in the first cluster with an average age of 48 years old and the most found comorbidity was diabetes. The second cluster consisted of 370 data (44%) with an average age of 64 years old with diabetes as the most found comorbidity. Meanwhile, the third cluster consisted of 140 data (17%) with an average age of 68 years and the highest comorbidity was hypertension. This analysis using K-Means Clustering can provide an overview of different groups of patients also with different clinical degrees and different forms of treatment.

**Keywords:** Covid-19, Comorbidities, K-Means Clustering