

PENGELOMPOKKAN DATA KASUS *COVID-19* DI JAWA BARAT DENGAN ALGORITMA *K-MEANS*

Laila Nur Fitriana

(170411100101)

Dosen Pembimbing I : Eka Mala Sari Rochman, S.Kom., M.Kom.

Dosen Pembimbing II : Dr. Bain Khusnul Khotimah, ST., M.Kom.

ABSTRAK

Covid-19 merupakan sebuah penamaan baru pada virus corona oleh WHO (Organisasi Kesehatan Dunia). Penyakit yang disebabkan oleh virus ini dapat menyebar secara cepat antar manusia. Dalam waktu singkat penyakit ini sudah banyak menyebar di *Indonesia*, salah satunya pada Provinsi Jawa Barat di wilayah Kabupaten dan sekitarnya. Banyaknya permasalahan yang belum diketahui dan dilaporkan, seperti masalah kesehatan, sosial, maupun perekonomian, membuat situasi di wilayah Kabupaten Jawa Barat menjadi buruk. Penambahan jumlah kasus *Covid-19* yang tidak dapat dikelompokkan mengakibatkan kurangnya persediaan alat kesehatan, seperti masker, alat pelindung medis, ruang isolasi ataupun sarana dan prasarana lainnya. Hal tersebut mengakibatkan penanganan yang kurang maksimal dan berdampak pada sulitnya menahan penyebaran penyakit *Covid-19* di Jawa Barat. Oleh karena itu pemerintah Jawa Barat membutuhkan sebuah sistem yang mampu mengelompokkan jumlah kasus *Covid-19* tersebut. Untuk mengatasi masalah mengenai pengelompokkan jumlah penderita *Covid-19* dimasa depan, peneliti membuat sistem yang mampu membagi jumlah wilayah kasus *Covid-19*. Metode pengelompokkan yang digunakan adalah metode *K-Means Clustering*. Hasil pengujian terhadap metode *K-Means Clustering* menggunakan metode *Silhouette Coefficient* dengan *SSE* menunjukkan bahwa metode *K-Means Clustering* dengan perhitungan jarak terdekat menggunakan *Minkowski Distance* adalah menjadi pilihan yang baik pada penelitian dengan didapat hasil pada penilaian *Silhouette Coefficient* dengan *SSE* adalah berada pada nilai 1 berada nilai $k=13$.

Kata Kunci – Covid-19, Jawa Barat, Cluster, K-Means, Silhouette Coefficient

DAFTAR ISI

HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN ORISINALITAS	iv
ABSTRAK.....	v
RIWAYAT HIDUP	vi
UCAPAN TERIMA KASIH	vii
KATA PENGANTAR	ix
DAFTAR ISI.	x
DAFTAR GAMBAR.....	Error! Bookmark not defined.
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	14
1.1 Latar Belakang	14
1.2 Perumusan Masalah.....	17
1.2.1 Permasalahan.....	18
1.2.3 Metode Usulan	18
1.2.3 Pertanyaan Penelitian	18
1.3 Tujuan dan Manfaat.....	18
1.3.1 Tujuan Penelitian.....	18
1.3.2 Manfaat Penelitian.....	18
1.4 Batasan Masalah.....	18
1.5 Sistematika Penulisan.....	19
BAB II KAJIAN PUSTAKA	20
2.1 Covid-19	20
2.2 Data Mining	21
2.3 Clustering	24
2.4 Algoritma K-Means	25
2.5 Analisis Cluster	26
2.6 Sum Of Square	27
2.7 Silhouette Coeficient.....	27
2.8 Penelitian Sebelumnya	28
BAB III METODOLOGI PENELITIAN	30
3.1.Tahapan Penelitian	30
3.1.1 Studi Pustaka.....	30
3.1.2 Rancangan Arsitektur Sistem.....	30

3.1.3	Implementasi Sistem	31
3.1.4	Uji Coba Sistem	31
3.1.5	Evaluasi Model.	31
3.2	Dataset	31
3.2.1	Korelasi Data.....	36
3.2.2	Pengolahan Dataset.....	39
3.3	<i>K-Fold Cross Validation</i>	40
3.4	Data Pra-Prosesing	41
3.4.1	<i>Data Cleaning</i>	42
3.5	<i>Data Transformasi</i>	42
3.6	<i>Flowchart</i> Sistem	42
3.7	Skenario Uji Coba	46
3.7.1	Gambarisasi Hasil Klasterisasi..	47
BAB IV HASIL DAN PEMBAHASAN		48
4.1	Lingkungan Uji Coba	48
4.2	Implementasi Sistem	48
4.2.1	Halaman Utama Sistem.....	48
4.2.2	Halaman Praproses.....	49
4.2.3	Halaman K-Means.....	50
4.2.4	Halaman Silhouette	51
4.2.5	Halaman Hasil dalam bentuk Grafik.....	51
4.3	Uji Coba Sistem.....	52
4.3.1	Pengujian 1.....	52
4.3.2	Pengujian 2.....	64
4.3.3	Pengujian 3.....	74
4.3.4	Pengujian 4.....	80
4.3.5	Pengujian 5.....	87
4.4	Hasil Pengujian.....	55
BAB V PENUTUP		93
5.1	Kesimpulan.....	93
5.2	Saran.....	93
REFERENSI		95
LAMPIRAN 1		99
LAMPIRAN 2		100