

**PENERAPAN DATA MINING UNTUK PENGELOMPOKAN
PENYEBARAN COVID-19 DI DKI JAKARTA
MENGUNAKAN ALGORITMA K-MEANS**



SKRISPI

Diajukan untuk memenuhi salah satu syarat kelulusan Program Sarjana

HETI SUHAETI

11211959

**Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Nusa Mandiri
Jakarta
2022**

PERSEMBAHAN

*Ketika aku tahu, aku semakin tahu kalau aku
tidak tahu apa-apa (Imam Safe'i)*

Dengan mengucap puji syukur kepada Allah S.W.T, Skripsi ini kupersembahkan

untuk:

1. Abah Acep dan Emak Marsiti (Almh) tercinta yang telah membesarkan aku dan selalu membimbing, mendukung, memotivasi, memberi apa yang terbaik bagiku serta selalu mendoakan aku untuk meraih cita-cita.
2. Suamiku (Agus Nuriman) yang telah menjadi support system, curahan hatiku, yang telah memberiku semangat.
3. Alina Sekar Arumingratri dan Hanum Aysaluna Ruby, anak-anakku tercinta yang selalu setia dan memberikan semangat.

*Tanpa mereka,
aku dan karya ini tak akan pernah ada*

UNIVERSITAS
NUSA MANDIRI

SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini, saya:

Nama : Heti Suhaeti :
NIM : 11211959
Program Studi : Sistem Informasi
Fakultas : Teknologi Informasi
Perguruan Tinggi : Universitas Nusa Mandiri

Dengan ini menyatakan bahwa Skripsi yang telah saya buat dengan judul: **“Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di DKI Jakarta Menggunakan Algoritma K-Means”**, adalah asli (orisinal) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Skripsi yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Universitas Nusa Mandiri dicabut/dibatalkan.

Dibuat di : Jakarta
Pada Tanggal : 15 Juli 2022
Yang menyatakan



Heti Suhaeti

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Heti Suhaeti :
NIM : 11211959
Program Studi : Sistem Informasi
Fakultas : Teknologi Informasi
Perguruan Tinggi : Universitas Nusa Mandiri

Dengan ini menyetujui untuk memberikan izin kepada pihak Universitas Nusa Mandiri, Hak Bebas Royalti Non-Eksklusif (Non-exclusive Royalti-Free Right) atas karya ilmiah kami yang berjudul: **“Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di DKI Jakarta Menggunakan Algoritma K-Means”**, beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak Universitas Nusa Mandiri berhak menyimpan, mengalih-media atau format-kan, mengelolanya dalam pangkalan data (database), mendistribusikannya dan menampilkan atau mempublikasikannya di internet atau media lain untuk kepentingan akademis tanpa perlu meminta izin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada Tanggal : 15 Juli 2022
Yang menyatakan



Heti Suhaeti

PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Heti Suhaeti
NIM : 11211959
Jenjang : Strata Satu (S1)
Program Studi : Sistem Informasi
Fakultas : Teknologi Informasi
Perguruan Tinggi : Universitas Nusa Mandiri
Judul Skripsi : Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di DKI Jakarta Menggunakan Algoritma K-Means

Telah dipertahankan pada periode 2022-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer (S.Kom) pada Program Strata Satu (S1) Program Studi Sistem Informasi Fakultas Teknologi Informasi di Universitas Nusa Mandiri.

Jakarta, 20 Juli 2022

PEMBIMBING SKRIPSI

Dosen Pembimbing : Yuni Eka Achyani, S.Kom,M.Kom



DEWAN PENGUJI

Penguji I : Rizki Aulianita, S.Kom,M.Kom



Penguji II : Tri Santoso, S.Kom,M.Kom



PEDOMAN PENGGUNAAN HAK CIPTA

Skripsi yang berjudul **“Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di DKI Jakarta Menggunakan Algoritma K-Means”** adalah hasil karya tulis asli HETI SUHAETI dan bukan hasil terbitan sehingga peredaran karya tulis hanya berlaku di lingkungan akademik saja, serta memiliki hak cipta. Oleh karena itu, dilarang keras untuk menggandakan baik sebagian maupun seluruhnya karya tulis ini, tanpa seizin penulis.

Referensi kepustakaan diperkenankan untuk dicatat tetapi pengutipan atau peringkasan isi tulisan hanya dapat dilakukan dengan seizin penulis dan disertai ketentuan pengutipan secara ilmiah dengan menyebutkan sumbernya.

Untuk keperluan perizinan pada pemilik dapat menghubungi informasi yang tertera di bawah ini:

Nama : Heti Suhaeti
Alamat : Jl. Gorontalo IV No. 16 RT/RW 015/001
Kel. Sungai Bambu Kec. Tanjung Priok, Jakarta Utara
No. Tlp : 08111391101
Email : hetisuhaeti17@gmail.com

UNIVERSITAS
NUSA MANDIRI

KATA PENGANTAR

Alhamdulillah, dengan mengucapkan puji syukur kehadiran Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tugas ini dengan baik. Dimana Skripsi ini penulis sajikan dalam bentuk buku yang sederhana. Adapun judul Skripsi, yang penulis ambil sebagai berikut, **“PENERAPAN DATA MINING UNTUK PENGELOMPOKAN PENYEBARAN COVID-19 DI DKI JAKARTA MENGGUNAKAN ALGORITMA K-MEANS”**.

Tujuan penulisan Skripsi ini dibuat sebagai salah satu syarat kelulusan Program Sarjana Universitas Nusa Mandiri. Sebagai bahan penulisan diambil berdasarkan Data data sekunder yang diperoleh tidak secara langsung dari objek penelitian, namun berasal dari data yang telah dikumpulkan sebelumnya oleh pihak lain dan beberapa sumber literatur yang mendukung penulisan ini yang mendukung penulisan ini. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan Skripsi ini tidak akan lancar. Oleh karena itu pada kesempatan ini, izinkanlah penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Nusa Mandiri
2. Wakil Rektor I Bidang Akademik Universitas Nusa Mandiri
3. Dekan Fakultas Teknologi Informasi Universitas Nusa Mandiri
4. Ketua Program Studi Informatika Universitas Nusa Mandiri.
5. Ibu Yuni Eka Achyani, M.Kom, selaku Dosen Pembimbing Skripsi.
6. Bapak/ibu dosen Program Studi Informatika Universitas Nusa Mandiri yang telah memberikan penulis dengan semua bahan yang diperlukan.
7. Staff / karyawan / dosen di lingkungan Fakultas Teknologi Informasi Universitas Nusa Mandiri.
8. Orang tua, Suami dan anak-anakku tercinta yang telah memberikan dukungan moral maupun spritual.
9. Rekan-rekan mahasiswa kelas 11.8i.01

Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi ini masih jauh

sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan dimasa yang akan datang.
Akhir kata semoga Skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 15 Juli 2022

Penulis



Heti Suhaeti

DAFTAR ISI

LEMBAR JUDUL SKRIPSI.....	ii
PERSEMBAHAN.....	ii
SURAT PERNYATAAN KEASLIAN SKRIPSI	iii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	iv
PERSETUJUAN DAN PENGESAHAN SKRIPSI.....	v
PEDOMAN PENGGUNAAN HAK CIPTA	vi
 KATA PENGANTAR	 vii
ABSTRAK	ix
DAFTAR ISI	xi
DAFTAR SIMBOL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL	xv
DAFTAR LAMPIRAN.....	xvi
 BAB I.....	 1
PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Identifikasi Masalah.....	2
1.3. Rumusan Masalah	3
1.4. Tujuan Penelitian	3
 BAB II.....	 5
LANDASAN TEORI.....	5
2.1. Tinjauan Pustaka	5
2.1.1 Covid-19.....	5
2.1.2 Data Mining.....	6
2.1.3 Clustering	10
2.1.4 K-Means	12
2.1.5 RapidMiner.....	13
2.2. Penelitian Terkait	16
2.3. Objek Penelitian	16
 BAB III.....	 18
METODOLOGI PENELITIAN	18
3.1. Tahapan Penelitian.....	18
3.2. Metode Penelitian.....	19
3.3. Analisa Data	20
 BAB IV	 Error! Bookmark not defined.
HASIL DAN PEMBAHASAN	22
4.1. Proses Algoritma K-Means Clustering	22
4.2. K-Means Pada Rapid Miner	35

4.3. Pembahasan	43
BAB V.....	44
PENUTUP	44
5.1. Kesimpulan.....	44
5.2. Saran	45
DAFTAR PUSTAKA	46
DAFTAR RIWAYAT HIDUP	48
LEMBAR BIMBINGAN SKRIPSI.....	49
LAMPIRAN	50
Lampiran A Bukti Hasil Pengecekan Plagiarisme	
Lampiran B Bukti Submit/Publish Artikel Ilmiah	

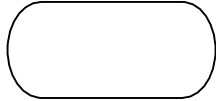


4.3. Pembahasan	43
BAB V.....	44
PENUTUP	44
5.1. Kesimpulan.....	44
5.2. Saran	45
DAFTAR PUSTAKA	46
DAFTAR RIWAYAT HIDUP	48
LEMBAR BIMBINGAN SKRIPSI.....	49
LAMPIRAN	50
Lampiran A Bukti Hasil Pengecekan Plagiarisme	
Lampiran B Bukti Submit/Publish Artikel Ilmiah	



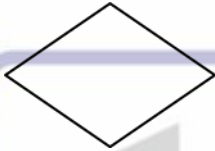
DAFTAR SIMBOL

a. Simbol Flowchart



TERMINAL

Digunakan untuk menggambarkan awal dan akhir dari suatu kegiatan



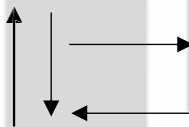
DECISION

Digunakan untuk menggambarkan proses pengujian suatu kondisi yang ada



PREPARATION

Digunakan untuk menggambarkan persiapan harga awal, dari proses yang akan dilakukan.



FLOW LINE

Digunakan untuk menggambarkan hubungan proses dari satu proses ke proses lainnya.



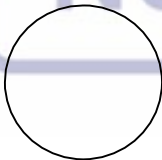
INPUT/OUTPUT

Digunakan untuk menggambarkan proses memasukan data yang berupa pembacaan data dan sekaligus proses keluaran yang berupa pencetakan data



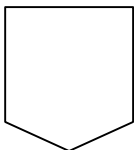
SUBROUTINE

Digunakan untuk menggambarkan proses pemanggilan sub program dari main program (recursivitas).



PAGE CONNECTOR

Digunakan untuk menghubungkan alur proses ke dalam satu halaman atau halaman yang sama.



CONNECTOR

Digunakan untuk menghubungkan alur proses dalam halamanyang berbeda atau ke halaman berikutnya

DAFTAR GAMBAR

	Halaman
Gambar II.1 Tampilan Rapid Miner versi 5.3.....	14
Gambar III.1 Kerangka Kerja Penelitian	17
Gambar III.2 Flowchart K-Means	20
Gambar IV.1 Tampilan Awal RapidMiner	35
Gambar IV.2 Tampilan New Process.....	36
Gambar IV.3 Data Import Tahap 1.....	36
Gambar IV.4 Tampilan Import Data Excel.....	37
Gambar IV.5 Tampilan Import Data Cluster Tahap 2	37
Gambar IV.6 Tampilan Import Data Cluster Tahap 3	38
Gambar IV.7 Tampilan Import Data Cluster Tahap 4	38
Gambar IV.8 Pemrosesan Tahap 1.....	39
Gambar IV.9 Pemrosesan Tahap 2.....	40
Gambar IV.10 Pemrosesan Tahap 3.....	40
Gambar IV.11 Nilai Cluster Model RapidMiner.....	41
Gambar IV.12 Hasil Pengelompokan RapidMiner.....	42
Gambar IV.13 Tampilan Data Perhitungan RapidMiner	43



DAFTAR TABEL

	Halaman
Tabel III.1 Data Kasus Covid-19 Di 44 Kecamatan Di DKI Jakarta	19
Tabel IV.1 Penyebaran Covid-19 di DKI Jakarta	21
Tabel IV.2 Centroid Data Awal	23
Tabel IV.3 Hasil Perhitungan Jarak Pusat Cluster Iterasi 1	23
Tabel IV.4 Hasil Cluster Iterasi 1.....	24
Tabel IV.5 Cluster Iterasi 1.....	25
Tabel IV.6 Centroid Baru Iterasi 2.....	26
Tabel IV.7 Hasil Perhitungan Jarak Pusat Cluster Iterasi 2	27
Tabel IV.8 Hasil Cluster Iterasi 2.....	27
Tabel IV.9 Cluster Iterasi 2.....	28
Tabel IV.10 Centroid Baru Iterasi 3.....	29
Tabel IV.11 Hasil Perhitungan Jarak Pusat Cluster Iterasi 3.....	29
Tabel IV.12 Hasil Cluster Iterasi 3	30
Tabel IV.13 Cluster Iterasi 3	31
Tabel IV.14 Centroid Baru Iterasi 4.....	32
Tabel IV.15 Hasil Perhitungan Jarak Pusat Cluster Iterasi 4.....	32
Tabel IV.16 Hasil Cluster Iterasi 4	33
Tabel IV.17 Cluster Iterasi 4	33
Tabel V.1 Kesimpulan hasil Pengelompokan Penyebaran Covid-19.....	44



DAFTAR LAMPIRAN

	Halaman
Lampiran A Bukti Hasil Pengecekan Plagiarisme	50
Lampiran B Bukti Submit/Publish Artikel Ilmiah	52



Abstrak

Heti Suhaeti (11211959), Penerapan Data Mining untuk Pengelompokan Penyebaran Covid-19 di DKI Jakarta menggunakan Algoritma K-Means

Covid-19 adalah penyakit yang disebabkan oleh virus SARS-CoV-2 yang menyerang pernafasan. Salah satu daerah yang terdampak Covid-19 adalah DKI Jakarta. Untuk mengatasi penyebaran virus yang sangat cepat di DKI Jakarta perlu pengelompokan daerah. Data mining sangat cocok diterapkan untuk mengetahui penyebaran virus Covid-19. Algoritma K-means merupakan salah satu teknik data mining untuk melakukan pengelompokan daerah penyebaran virus Covid-19. Dengan menggunakan metode ini data-data yang telah diperoleh dapat dikelompokkan ke dalam beberapa cluster, dimana penerapan proses K-Means Clustering menggunakan tools RapidMiner. Penelitian ini menggunakan sumber data terekam di situs pemerintah DKI Jakarta dengan alamat url <https://jakartasatu.jakarta.go.id/>. Data yang digunakan adalah bulan Januari 2022 yang terdiri dari 44 Kecamatan. Kriteria yang digunakan, yaitu 1) pasien positif 2) dirawat 3) sembuh 4) meninggal. Data diolah menggunakan K-means yang dibagi dalam 3 cluster yaitu tingkat penyebaran tinggi (C1), tingkat penyebaran sedang (C2), dan tingkat penyebaran rendah (C3). Proses iterasi berlangsung 4 kali sehingga diperoleh penilaian dalam mengelompokkan daerah penyebaran virus Covid-19 berdasarkan kecamatan. Hasil perhitungan menggunakan K-Means didapatkan daerah yang terdampak virus Covid-19 dengan tingkat tinggi (C3) sebanyak 17 kecamatan, tingkat sedang (C2) sebanyak 24 kecamatan, dan tingkat rendah sebanyak (C1) 3 Kecamatan. Hasil dari penelitian ini menjadi masukan untuk pemerintah DKI Jakarta agar daerah-daerah yang memerlukan prioritas penanganan virus Covid-19 dapat dilakukan secara efektif mungkin agar dapat memutus mata rantai penyebaran.

Kata kunci: Covid-19, Data Mining, K-Means, Clustering, RapidMiner

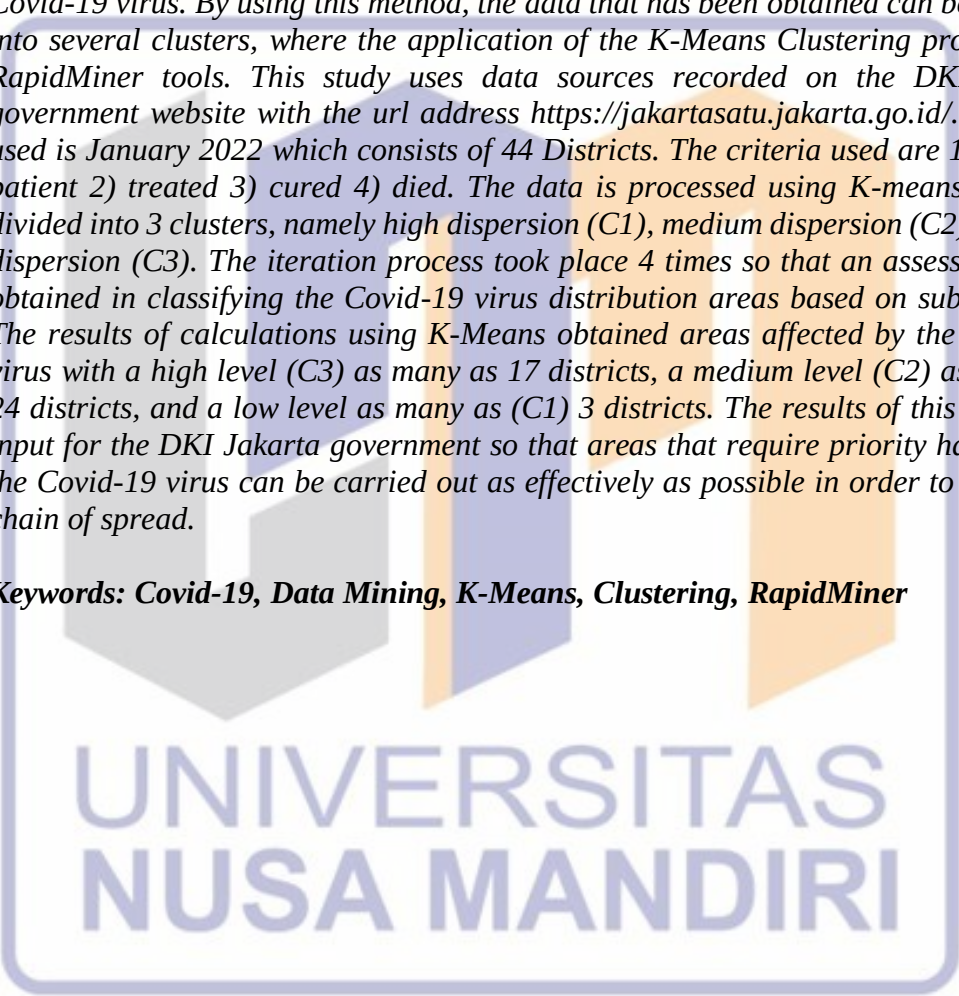
UNIVERSITAS
NUSA MANDIRI

Abstract

Heti Suhaeti (11211959), Application of Data Mining for Grouping the Spread of Covid-19 in DKI Jakarta using the K-Means Algorithm

Covid-19 is a disease caused by the SARS-CoV-2 virus that attacks the respiratory tract. One of the areas affected by Covid-19 is DKI Jakarta. To overcome the very fast spread of the virus in DKI Jakarta, it is necessary to group regions. Data mining is very suitable to be applied to determine the spread of the Covid-19 virus. The K-means algorithm is one of the data mining techniques for grouping the spread areas of the Covid-19 virus. By using this method, the data that has been obtained can be grouped into several clusters, where the application of the K-Means Clustering process uses RapidMiner tools. This study uses data sources recorded on the DKI Jakarta government website with the url address <https://jakartasatu.jakarta.go.id/>. The data used is January 2022 which consists of 44 Districts. The criteria used are 1) positive patient 2) treated 3) cured 4) died. The data is processed using K-means which is divided into 3 clusters, namely high dispersion (C1), medium dispersion (C2), and low dispersion (C3). The iteration process took place 4 times so that an assessment was obtained in classifying the Covid-19 virus distribution areas based on sub-districts. The results of calculations using K-Means obtained areas affected by the Covid-19 virus with a high level (C3) as many as 17 districts, a medium level (C2) as many as 24 districts, and a low level as many as (C1) 3 districts. The results of this study are input for the DKI Jakarta government so that areas that require priority handling of the Covid-19 virus can be carried out as effectively as possible in order to break the chain of spread.

Keywords: Covid-19, Data Mining, K-Means, Clustering, RapidMiner



UNIVERSITAS
NUSA MANDIRI

DAFTAR PUSTAKA

- [1] A. Susilo *et al.*, “Coronavirus Disease 2019 : Tinjauan Literatur Terkini Coronavirus Disease 2019 : Review of Current Literatures,” *J. Penyakit Dalam Indones.*, vol. 7, no. 1, pp. 45–67, 2020.
- [2] R. Widiyani, “Latar Belakang Virus Corona, Perkembangan hingga Isu Terkini, 2020,” 2019. <https://news.detik.com/berita/d4943950/latar-belakang-virus-corona-perkembangan-hingga-isu-terkini>
- [3] Y. F. S. Y. Damanik, S. Sumarno, I. Gunawan, D. Hartama, and I. O. Kirana, “Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means,” *J. Ilmu Komput. dan Inform.*, vol. 1, no. 2, pp. 109–132, 2021, doi: 10.54082/jiki.13.
- [4] N. Jannah and T. Yulianto, “Mengelompokkan Siswa Berprestasi Akademik dengan Menggunakan Metode K Means Kelas VII MTe,” 2016.
- [5] M. Rifaldi, *pandemic virus corona*. Bengkulu, 2021. [Online]. Available: https://www.google.co.id/books/edition/Pandemi_Virus_Corona/3skwEAAQBAJ?hl=en&gbpv=1&pg=PP1&printsec=frontcover
- [6] Anjar Wanto, “Data Mining : Algoritma dan Implementasi,” 2020, [Online]. Available: https://www.google.co.id/books/edition/Data_Mining_Algoritma_dan_Implementasi/gAnfDwAAQBAJ?hl=en&gbpv=1&pg=PP1&printsec=frontcover
- [7] L. Maulida, “Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.
- [8] I. Pramudiono, *Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data I*. 2007. [Online]. Available: http://www.ilmukomputer.org/wpcontent/uploads/2006/08/iko_datamining
- [9] R. A. Indraputra and R. Fitriana, “K-Means Clustering Data COVID-19,” *J. Tek. Ind.*, vol. 10, no. 3, pp. 275–282, 2020, doi: 10.25105/jti.v10i3.8428.
- [10] A. N. Manthovani, S. Serdawati, C. F. Sari, I. F. Widiawati, Z. Putri, and E. Widodo, “Pengelompokkan Jumlah Kasus Tuberculosis Paru di Indonesia Menggunakan Cluster K-Means Tahun 2016,” *Knpmp Iii*, vol. 3, pp. 731–740, 2018.
- [11] A. Agrawal and H. Gupta, “Global K-Means (GKM) Clustering Algorithm: A Survey,” *Int. J. Comput. Appl.*, vol. 79, no. 2, pp. 20–24, 2013, doi: 10.5120/13713-1472.
- [12] K. Fatmawati and A. P. Windarto, “Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi,” *Comput. Eng. Sci. Syst. J.*, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.

- [13] Z. Zahrotunnimah, “Langkah Taktis Pemerintah Daerah Dalam Pencegahan Penyebaran Virus Corona Covid-19 di Indonesia,” *SALAM J. Sos. dan Budaya Syar-i*, vol. 7, no. 3, pp. 247–260, 2020, doi: 10.15408/sjsbs.v7i3.15103.
- [14] Jakartasatu, “Jelajahi Peta Informasi Covid-19 Melalui Jakartasatu.” <https://jakartasatu.jakarta.go.id/portal/apps/sites/#/peta-persebaran-kasus-covid-19>
- [15] W. Widodo and D. Wahyuni, “Implementasi Algoritma K-Means Clustering Untuk Mengetahui Bidang Skripsi Mahasiswa Multimedia Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta,” *PINTER J. Pendidik. Tek. Inform. dan Komput.*, vol. 1, no. 2, pp. 157–166, 2017, doi: 10.21009/pinter.1.2.10.

