



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



# IMPLEMENTASI CLUSTERING K-MEANS UNTUK MENENTUKAN PENYEBARAN COVID-19 DI INDONESIA

LAPORAN SKRIPSI

MUHAMMAD RAMADHAN 4817040242

POLITEKNIK  
NEGERI  
JAKARTA

PROGRAM STUDI TEKNIK INFORMATIKA  
JURUSAN TEKNIK INFORMATIKA DAN  
KOMPUTER

POLITEKNIK NEGERI JAKARTA

2019/2020



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



# IMPLEMENTASI METODE K-MEANS UNTUK MENENTUKAN PENYEBARAN COVID-19 DI INDONESIA

LAPORAN SKRIPSI

POLITEKNIK  
NEGERI  
JAKARTA

Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan untuk Memperoleh  
Diploma Empat Politeknik

MUHAMMAD RAMADHAN

4817040242

PROGRAM STUDI TEKNIK INFORMATIKA  
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER  
POLITEKNIK NEGERI JAKARTA  
2019/2020

## HALAMAN PERNYATAAN ORISINALITAS

Skripsi/Tesis/Disertasi ini adalah hasil karya saya sendiri, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : **Muhammad Ramadhan**

NIM : **4817040242**

Tanggal : **10 Juni 2021**

Tanda Tangan : 

**© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta**

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta





**Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta**

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**LEMBAR PENGESAHAN**

Skripsi diajukan oleh

Nama : Muhammad Ramadhan

NIM : 4817040242

Program Studi : TI

Judul Skripsi : IMPLEMENTASI METODE K-MEANS  
UNTUK MENENTUKAN PENYEBARAN COVID-19 DI INDONESIA

Telah diuji oleh tim penguji dalam Sidang Skripsi pada hari Senin, Tanggal  
14, Bulan Juni, Tahun 2021 dan dinyatakan LULUS.

Disahkan oleh

Pembimbing I : Dr. Dewi Yanti Liliana, S.Kom., M.Kom. (.....)

Penguji I : Risna Sari, S.Kom., M.Ti. (.....)

Penguji II : Asep Taufik Muharram, S.Kom., M.Kom. (.....)

Penguji III : Anggi Mardiyono, S.Kom., M.Kom. (.....)

Mengetahui :

Jurusan Teknik Informatika dan Komputer  
Ketua

Mauldy Laya, S.Kom.,M.Kom  
NIP. 197802112009121003



## KATA PENGANTAR

Alhamdulillah syukur penulis panjatkan atas kehadiran Allah SWT karena atas rahmat dan karunia Nya-lah laporan skripsi ini dapat diselesaikan. Skripsi ini berjudul "IMPLEMENTASI METODE K-MEANS UNTUK MENENTUKAN PENYEBARAN COVID-19 DI INDONESIA". Pada kesempatan ini penulis mengucapkan banyak terima kasih kepada berbagai pihak atas bantuan, bimbingan dan dukungannya sehingga penulisan ini berjalan lancar, yaitu:

- a. Allah SWT tuhan yang maha esa, yang telah memberikan penulis kesehatan dan akal sehat sehingga laporan ini dapat terselesaikan dengan baik.
- b. Ibu Dr. Dewi Yanti Liliana, S.Kom., M.Kom. selaku dosen pembimbing skripsi yang telah meluangkan waktu, tenaga dan pikiran untuk membimbing dan membantu penulis dalam menyusun laporan skripsi.
- c. Orang tua dan keluarga penulis yang setiap saat mendoakan penulis serta memberikan dukungan dan bantuan moral maupun material kepada penulis.
- d. Sahabat dan teman-teman yang telah banyak membantu penulis dalam menyelesaikan penyusunan skripsi ini, terkhusus untuk Sarah, Maulida dan teman-teman TI angkatan 2017.

Akhir kata, penulis berharap Allah SWT berkenan membalas segala kebaikan Bapak/Ibu semua dan semoga laporan skripsi ini dapat bermanfaat bagi orang banyak dan pengembangan ilmu.

Jakarta, 10 Juni 2021

Muhammad Ramadhan

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI

### SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Muhammad Ramadhan

NIM 4817040242

Program Studi : Teknik Informatika

Jurusan : Teknik Informatika dan Komputer

Jenis Karya : Skripsi

demikian demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta **Hak Bebas Royalti Noneksklusif** (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

*IMPLEMENTASI METODE K-MEANS UNTUK MENENTUKAN PENYEBARAN COVID-19 DI INDONESIA*

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Politeknik Negeri Jakarta berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database) , merawat, dan memublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya. Dibuat di: Depok

Pada tanggal: 10 Juni 2021

Yang menyatakan

Muhammad Ramadhan



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**IMPLEMENTASI METODE K-MEANS UNTUK MENENTUKAN PENYEBARAN  
COVID-19 DI INDONESIA**

*Abstrak*

*Pada tahun 2020 dunia dikejutkan dengan penemuan virus baru yaitu coronavirus yang berasal dari Wuhan, Provinsi Hubei Tiongkok yang nantinya disebut sebagai Coronavirus disease 2019(Covid-19). Berdasarkan data yang dihimpun sampai 2 Februari 2020 dipastikan sudah ada 219 negara yang terjangkit Covid -19 dengan jumlah kasus mencapai 103.931.497 dan di Indonesia sendiri sudah mencapai total kasus 1,089,300. Banyaknya penduduk Indonesia menyulitkan pemerintah untuk dapat mencegah penularan Covid -19 dan memberikan pelayanan kesehatan yang tepat bagi masyarakat sehingga diperlukan suatu cara agar penularan Covid -19 dapat dicegah. Upaya yang bisa dilakukan salah satunya dengan mempelajari karakteristik Covid -19 dan mengelompokkannya berdasarkan kesamaan karakteristik di setiap provinsi di Indonesia sehingga ciri khas dari tiap masing-masing kelompok dapat diketahui. Pemerintah sendiri telah mengembangkan website untuk memberikan informasi mengenai perkembangan covid-19 di Indonesia namun dalam website tersebut kemampuan untuk mendapatkan data secara real time belum diaplikasikan sehingga data terbaru terkadang belum ditampilkan yang dimana masyarakat sangat membutuhkan data yang real time sehingga dapat mengetahui perkembangan covid 19 secara jelas dan tepat. Dari permasalahan tersebut, dibuatlah visualisasi data kasus covid19 berbasis website yang menggunakan teknik data mining untuk menganalisa data dengan tujuan mendapatkan pengetahuan. Data mining berfokus pada pengelompokan data covid 19 berdasarkan kasus harian dan kasus meninggal. Dalam pengelompokan data covid 19 pada aplikasi ini menggunakan metode clustering dengan algoritma k-means untuk pengelompokan data ke dalam bentuk dua atau lebih kelompok dengan tujuan untuk meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antarkelompok.*

**Kata Kunci:** Covid-19, Clustering, Data Mining, K-Means

## DAFTAR ISI

HALAMAN PERNYATAAN ORISINALITAS.....	iii
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI.....	vi
KRIPSI UNTUK KEPENTINGAN AKADEMIS.....	vi
Abstrak.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR .....	x
DAFTAR TABEL.....	xi
BAB I.....	1
PENDAHULUAN.....	1
1. Latar Belakang.....	1
2. Rumusan Masalah.....	2
3. Batasan Masalah.....	2
4. Tujuan.....	3
5. Manfaat.....	3
6. Metode Pelaksanaan Skripsi.....	3
BAB II.....	5
TINJAUAN PUSTAKA.....	5
2.1 Data Mining.....	5
2.2 Clustering.....	6
2.3 Algoritma K-Means.....	7
2.4 Covid-19.....	9
2.5 Website.....	9
2.6 Python.....	11
2.7 Flask.....	11
2.8 NOSQL.....	12
2.9 Firebase.....	13
2.10 Robust Scalar.....	14
2.11 Metode Elbow.....	15

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta





## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritis atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

12 Silhouette Coefficient.....	15
13 DB Index.....	17
14 Unified Modeling Language (UML).....	18
15 Penelitian Terdahulu.....	20
<b>BAB III.....</b>	<b>21</b>
<b>PERENCANAAN DAN REALISASI.....</b>	<b>21</b>
1 Perancangan sistem.....	21
1.1 Deskripsi program aplikasi.....	21
1.2 Analisis kebutuhan.....	21
1.3 Tahapan Data Mining.....	22
1.4 Cara Kerja Program Aplikasi.....	27
1.5 Rancangan Program Aplikasi.....	31
1.6 Perancangan Protoyping.....	36
2 Realisasi Sistem.....	37
2.1 Implementasi tampilan sistem.....	38
3.2.2 Implementasi metode K-Means.....	39
<b>BAB IV.....</b>	<b>50</b>
<b>PEMBAHASAN.....</b>	<b>50</b>
4.1 Pengujian.....	50
4.2 Deskripsi Pengujian.....	50
4.3 Prosedur Pengujian.....	50
4.3.1 Prosedur Pengujian Sistem.....	50
4.3.2 Prosedur Pengujian Algoritma K-Means.....	51
4.4 Data Hasil Pengujian.....	51
4.4.1 Hasil Pengujian Algoritma K-Means.....	51
4.4.2 Hasil Pengujian Sistem.....	57
4.5 Analisis Data / Evaluasi.....	58
4.5.1 Evaluasi Algoritma K-Means.....	59
4.5.2 Evaluasi Sistem.....	59
<b>BAB V.....</b>	<b>61</b>
<b>PENUTUP.....</b>	<b>61</b>

## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritis atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



1 Kesimpulan.....	61
2 Saran.....	61
DAFTAR PUSTAKA.....	62
LAMPIRAN.....	65





## DAFTAR GAMBAR

Gambar 2. 1 Flowchart Process Sumber (Muslehatin, et al., 2017).....	6
Gambar 2. 2 Fitur firebase (Maulana, 2019).....	13
Gambar 3. 1 Histogram kasus harian sebelum ditransformasi.....	25
Gambar 3. 2 Histogram kasus harian setelah ditransformasi.....	25
Gambar 3. 3 Histogram kasus meninggal sebelum ditransformasi.....	26
Gambar 3.4 Histogram kasus meninggal setelah ditransformasi.....	26
Gambar 3. 5 Flowchart Process.....	28
Gambar 3. 6 Flowchart Process.....	29
Gambar 3. 7 Flowchart K-Means.....	30
Gambar 3. 8 Use Case.....	31
Gambar 3. 9 Activity Diagram System.....	32
Gambar 3. 10 Activity Diagram Process.....	33
Gambar 3. 11 ER Diagram.....	34
Gambar 3. 12 Sequence Diagram System.....	35
Gambar 3. 13 Sequence Diagram Process.....	36
Gambar 3. 14 Prototype dashboard.....	37
Gambar 3. 15 Dashboard Aplikasi.....	38
Gambar 3. 16 Source Code pengambilan data.....	39
Gambar 3. 17 Source Code Map.....	40
Gambar 3. 18 Source Code clustering K-Means.....	42
Gambar 3. 19 Centeroid akhir.....	48
Gambar 4.1 Proses Evaluasi Jumlah Cluster Menggunakan Elbow Method.....	52
Gambar 4. 2 Hasil uji metode elbow.....	52
Gambar 4. 3 Scatterplot jumlah cluster sama dengan 3.....	53
Gambar 4.4 Proses Evaluasi Jumlah Cluster Menggunakan DB Index.....	54
Gambar 4.5 Hasil Uji DB Index.....	54
Gambar 4.6 Scatterplot jumlah cluster sama dengan 2.....	55
Gambar 4.7 Proses Evaluasi Jumlah Cluster Menggunakan Silhouette.....	56

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Gambar 4.8 Hasil Uji Silhouette.....	56
--------------------------------------	----



## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritis atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



## DAFTAR TABEL

Tabel 2. 1 Kriteria pengukuran silhouette coefficient Sumber : (Dewi & Pramita, 2019)...	16
Tabel 2. 2 Penelitian terdahulu.....	20
Tabel 3. 1 Data awal.....	23
Tabel 3. 2 Perhitungan rata-rata dan standar deviasi.....	24
Tabel 3. 3 Perhitungan threshold.....	24
Tabel 3. 4 Tabel Data clustering 27 maret 2021.....	40
Tabel 3. 5 Tabel Data setelah menggunakan RobustScaler().....	42
Tabel 3. 6 Tabel Centeroid awal.....	44
Tabel 3. 7 Tabel pengklusteran kmeans.....	45
Tabel 3. 8 Tabel centeroid baru.....	46
Tabel 3. 9 Tabel clustering baru.....	46
Tabel 3. 10 Hasil clustering K-means.....	48
Tabel 4. 1 Data Uji.....	51
Tabel 4. 2 Hasil DB Index.....	55
Tabel 4. 3 Hasil Silhouette score.....	57
Tabel 4. 4 Hasil Pengujian.....	58

© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



POLITEKNIK  
NEGERI  
JAKARTA



## BAB I

### PENDAHULUAN

#### 1. Latar Belakang

Pada tahun 2020 dunia dikejutkan dengan penemuan virus baru yaitu *coronavirus* yang berasal dari Wuhan, Provinsi Hubei Tiongkok yang nantinya disebut sebagai *coronavirus disease 2019* (Covid-19) (Yuliana, 2020). Covid -19 merupakan virus yang bersifat zoonosis yang berarti virus dapat ditularkan dari hewan ke manusia, namun beberapa kasus menemukan bahwa virus dapat ditularkan dari manusia ke manusia melalui *droplet*, kontak dengan *droplet* dan melalui penularan *fekal-oral* dengan Gejala awal yang tidak spesifik seperti demam, batuk yang kemudian dapat sembuh secara spontan atau berkembang menjadi sesak napas, dispnea, dan pneumonia yang menyebabkan Acute Respiratory Distress Syndrome (ARDS), gagal ginjal, disfungsi koagulasi, multipel kegagalan organ dan kematian (Harahap, 2020). Berdasarkan data yang dihimpun sampai 2 Februari 2020 dipastikan sudah ada 219 negara yang terjangkit Covid -19 dengan jumlah kasus mencapai 103.931.497 dan di Indonesia sendiri sudah mencapai total kasus 1,089,300 (worldometers, 2021). Indonesia menduduki posisi keempat sebagai negara dengan jumlah penduduk terbesar di dunia setelah Republik rakyat Cina, India, Amerika Serikat pada tahun 2019 yang diproyeksikan mencapai 267 juta jiwa (Satiti, 2019). Banyaknya penduduk Indonesia menyulitkan pemerintah untuk dapat mencegah penularan Covid -19 dan memberikan pelayanan kesehatan yang tepat bagi masyarakat sehingga diperlukan suatu cara agar penularan Covid -19 dapat dicegah. Upaya yang bisa dilakukan salah satunya dengan mempelajari karakteristik Covid -19 dan mengelompokkannya berdasarkan kesamaan karakteristik di setiap provinsi di Indonesia sehingga ciri khas dari tiap masing-masing kelompok dapat diketahui. Pemerintah sendiri telah mengembangkan website untuk memberikan informasi mengenai perkembangan covid-19 di Indonesia namun dalam website tersebut kemampuan untuk mendapatkan data secara real time belum diaplikasikan sehingga data terbaru terkadang belum ditampilkan yang dimana masyarakat sangat

#### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritikan atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



© Hak Cipta milik Jurusan Teknik Politeknik Negeri Jakarta

#### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan Teknik Politeknik Negeri Jakarta

memerlukan data yang real time sehingga dapat mengetahui perkembangan covid 19 secara jelas dan tepat.

K-Means sendiri adalah algoritma *Clustering* yang termasuk dalam kelompok *Unsupervised learning* yang digunakan untuk mengelompokkan data kedalam beberapa kelompok dengan sistem partisi. Algoritma ini menerima input berupa data tanpa label kelas. Pada algoritma K-Means, komputer menerima data-data yang tidak diketahui kelasnya terlebih dahulu lalu mengelompokkannya. Input yang diterima ialah data dan jumlah kelompok (cluster) yang diinginkan (Wanto, et al., 2020) dan telah dibuktikan bahwa K-Means dapat mengelompokkan penyebaran penyakit Covid -19 melalui penelitian yang dilakukan oleh Dwitri, et al., 2020.

Berdasarkan penjelasan di atas dapat disimpulkan bahwa Indonesia memerlukan suatu sistem untuk dapat mengelompokkan penyebaran penyakit Covid -19 sehingga masyarakat dan pemerintah dapat mengetahui kondisi penyebaran Covid 19 di Indonesia dan pemilihan metode K-means sendiri dikarenakan sudah terbukti dapat mengelompokkan penyebaran Covid -19.

### 1.2 Rumusan Masalah

Mengacu pada latar belakang yang telah diuraikan, maka rumusan masalahnya adalah:

1. Bagaimana mengumpulkan data penyebaran Covid -19 secara berkala
2. Bagaimana menerapkan *k-means* untuk menentukan penyebaran Covid -19 secara berkala
3. Bagaimana kinerja dari *k-means* untuk menentukan penyebaran Covid -19 secara berkala
4. Bagaimana memvisualisasikan hasil penyebaran Covid -19

### 1.3 Batasan Masalah

Adapun batasan masalah dari aplikasi ini adalah :

- Dijalankan melalui aplikasi berbasis website
- Bahasa yang digunakan di web adalah *python* dengan framework *flask*



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Data yang digunakan hanya data Covid -19 di Indonesia

Menggunakan API dari kawalcorona yang dimana jika tidak ada update data terbaru maka aplikasi tidak akan mengupdate data

#### 4 Tujuan

Berdasarkan rumusan masalah yang diuraikan maka sebagai berikut:

1. Mengumpulkan data penyebaran Covid -19 secara berkala
2. Menerapkan *k-means* untuk menentukan penyebaran Covid -19 secara berkala
3. Hasil clustering metode *k-means* diharapkan dapat mengelompokkan penyebaran Covid-19
4. Mengvisualisasikan hasil penyebaran Covid -19

#### 5 Manfaat

Mengelompokkan data penyebaran Covid -19 sehingga pemerintah bisa mengetahui zona rawan Covid -19 secara berkala

- Dapat dijadikan untuk pengambilan keputusan baik bagi pemerintah maupun masyarakat Indonesia

#### 1.6 Metode Pelaksanaan Skripsi

Penelitian ini dilakukan dalam beberapa tahap, antara lain:

##### 1.6.1 Identifikasi Masalah

Pada tahap ini akan mempelajari masalah yang terjadi dengan cara melakukan observasi melalui berita maupun data yang mendukung. Sehingga mendapatkan permasalahan berdasarkan penelitian terbaru maupun pengamatan keadaan sekarang.

##### 1.6.2 Metode Pengumpulan Data

Pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

###### 2.1 Studi Literatur

Studi literatur dilakukan dengan mencari sumber





**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

pustaka yang relevan dengan masalah yang dijadikan topik baik melalui buku, jurnal, maupun teori lainnya yaitu penelitian mengenai clustering, Covid -19 dan pemrograman python.

## 2.2 Analisis Permasalahan

Setiap studi literatur yang sebelumnya dikumpulkan akan dianalisis untuk mendapatkan pemahaman tentang bagaimana menerapkan metode *k-means* dalam mengelompokkan data dalam penyebaran virus Covid -19.

### 1.6.3 Analisis dan Perancangan

Setelah studi literatur didapatkan maka kemudian dianalisis kebutuhan fungsional dari sistem dengan cara mendefinisikan kebutuhan dan menentukan Batasan yang ada. Setelah itu perancangan arsitektur untuk menggambarkan alur sistem secara utuh.

### 1.6.4 Implementasi

Pada tahap ini akan diimplementasikan metode *k-means* untuk menentukan penyebaran virus Covid -19. Perancangan sistem merupakan web dengan bahasa pemrograman python dan framework flask.

### 1.6.5 Pengujian

Pengujian pada sistem dilakukan untuk mengevaluasi apakah sistem yang dibuat berjalan sebagaimana mestinya dan memastikan sistem sesuai dengan kebutuhan yang telah didefinisikan sebelumnya.

### 1.6.6 Kesimpulan

Kesimpulan didasarkan pada hasil analisis dan pengolahan data yang bertujuan untuk menjawab tujuan penelitian maupun memberikan gambaran penelitian yang ringkas dan mudah dipahami. Pada tahap ini juga terdapat masukan untuk penelitian selanjutnya.



## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Kesimpulan yang didapat berdasarkan hasil sistem implementasi metode k-means untuk menentukan penyebaran covid-19 di indonesia yaitu:

1. Sistem implementasi metode k-means untuk menentukan penyebaran covid-19 di indonesia telah berhasil memberikan informasi hasil clustering dari semua provinsi di Indonesia.
2. Pengujian black box testing terhadap sistem implementasi metode k-means untuk menentukan penyebaran covid-19 di indonesia dinyatakan berhasil.
3. Pengujian menggunakan *Elbow method* memberikan hasil cluster optimal adalah 3 namun pengujian menggunakan *DB Index* dan *Silhouette* memberikan hasil cluster optimal adalah 2 sehingga penggunaan cluster sama dengan 2 lebih tepat

#### 5.2 Saran

Berdasarkan pada permasalahan yang diangkat yaitu pengclusteran zona covid-19 yang menggunakan metode k-means, masih diperlukan perbaikan kedepannya. Saran untuk pengembang sistem selanjutnya yaitu:

1. Menggunakan database yang lebih kuat dan besar yang memang diperuntukan untuk data mining
2. Memberikan informasi mengenai vaksinasi covid-19 di Indonesia

#### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritis atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



## DAFTAR PUSTAKA

- Aggarwal, C. C., 2015. *Data Mining*. 1 ed. s.l.:Springer International Publishing.
- D. S. , B. J., 2018. PENINGKATAN HASIL EVALUASI CLUSTERING DAVIES-BOULDIN INDEX DENGAN PENENTUAN TITIK PUSAT CLUSTER AWAL ALGORITMA K-MEANS.
- Haroini, A. A. S. & Yustanti, W., 2016. PERBANDINGAN PENGGUNAAN NOSQL MONGODB DAN MYSQL PADA BASIS DATA FORUM KOMUNIKASI. *Manajemen Informatika*, VI(1), pp. 134-142.
- Hewi, D. A. I. C. & Pramita, D. A. K., 2019. Analisis Perbandingan Metode Elbow dan Silhouette pada Analisis Perbandingan Metode Elbow dan Silhouette pada Produksi Perajinan Bali. *MATRIX*, 9(3), pp. 102-109.
- Juhita , W. M. P., 2015. CLUSTERING MENGGUNAKAN METODE K-MEANS UNTUK MENENTUKAN STATUS GIZI BALITA. *Informatika*, 15(2), pp. 160-174.
- Radli, A. et al., 2020. Analisis Perbandingan Unjuk Kerja Database SQL dan Database NoSQL Untuk Mendukung Era Big Data. *Nasional Teknik Elektro*, IX(3), pp. 154-158.
- Sunawan, R., 2018. PENGUKURAN QUERY RESPON TIME PENGUKURAN QUERY RESPON TIME. *Siliwangi*, IV(2), pp. 100-103.
- Harahap, R. J. T., 2020. KARAKTERISTIK KLINIS PENYAKIT CORONAVIRUS 2019. *Jurnal Penelitian Perawat Profesional*, 2(3), pp. 317-324.
- Harismawan, A. F., Kharisma, A. P. & Afirianto, T., 2018. Analisis Perbandingan Performa Web Service Menggunakan Bahasa Pemrograman Python, PHP, dan Perl pada Client Berbasis Android. *Pengembangan Teknologi Informasi dan Ilmu Komputer*, II(1), pp. 234-245.
- Harison & Syarif, A., 2016. Sistem Informasi Geografis Sarana Pada Kabupaten Pasaman Barat. *TEKNOIF*, IV(2), pp. 41-43.
- Irsyad, R., 2018. Penggunaan Python Web Framework Flask Untuk Pemula.
- Mardi, Y., 2017. Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika Penelitian Bidang Komputer Sains dan Pendidikan Informatika*, 2(2), pp. 213-219.
- Maulana, I. F., 2019. Penerapan Firebase Realtime Database pada Aplikasi E-Tilang Smartphone berbasis Mobile Android. *RESTI*, IV(5), pp. 854-863.
- Mona, N., 2020. KONSEP ISOLASI DALAM JARINGAN SOSIAL UNTUK MEMINIMALISASI EFEK CONTAGIOUS (KASUS PENYEBARAN VIRUS CORONA DI INDONESIA). *Jurnal Sosial Humaniora Terapan*, 2(2), pp. 117-125.
- Muhardian, A., 2018. *Belajar Pemrograman Python: Pengenalan Dasar Python dan Persiapan Awal*. [Online]

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta

Available at: <https://www.petanikode.com/python-linux/>  
[Accessed 03 02 2020].

Muslehatin, W., Ibnu, M. & M., 2017. Penerapan Naive Bayes Classification untuk Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Klasifikasi Tingkat Kemungkinan Obesitas Mahasiswa Sistem Informasi UIN Suska Riau. *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 9*, Volume 9, pp. 250-256.

Ahlevi, O., Mulyani, A. & Khoir, M., 2018. SISTEM INFORMASI INVENTORI BARANG MENGGUNAKAN METODE OBJECT ORIENTED DI PT. LIVAZA TEKNOLOGI INDONESIA JAKARTA. *PROSISKO*, V(1), pp. 27-35.

Prehandayana, G., Yahya, W. & Nurwarsito, H., 2018. Implementasi Struktur Data Dictionary Untuk Sistem Monitoring Perangkat Internet Of Things. *Pengembangan Teknologi Informasi dan Ilmu Komputer*, II(10), pp. 3466-3473.

Uttra, R. E., 2021. *Wawancara* [Interview] (13 Juli 2021).

Ahman, A. T., W. & Anggrainingsih, R., 2017. Coal Trade Data Clustering Using K-Means (Case Study PT. Global Bangkit Utama). *ITSMART: Jurnal Ilmiah Teknologi dan Informasi*, 6(1), pp. 24-31.

Chohma, F. F., Ismail, I. E. & Waluyo, Y. S., 2018. Implementation of K-means Clustering on SIPP-KLING Dashboard Applications. *multinetics*, 4(2), pp. 38-42.

Sanadi, E. A. W., Achmad, A. & Dewiani, 2018. Pemanfaatan Realtime Database di Platform Firebase Pada Aplikasi E-Tourism Kabupaten Nabire. *JPE*, XII(1), pp. 20-26.

Sanusi, W., Zaky, A. & Afni, B. N., 2018. Analisis Fuzzy C-Means dan Penerapannya Dalam Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Faktor-faktor Penyebab Gizi Buruk.

Sari, R. P. & I., 2018. Analisis dan Perancangan Sistem Informasi Rapat Online FMIPA UNTAN menggunakan UML. *SISFOTEK*, pp. 154-165.

Satiti, S., 2019. GERAKAN AYO SEKOLAH DI KABUPATEN BOJONEGORO: PENINGKATAN SUMBER DAYA MANUSIA MELALUI PENDIDIKAN UNTUK MENYONGSONG BONUS DEMOGRAFI. *JURNAL KEPENDUDUKAN INDONESIA*, 14(1), pp. 77-92.

scikit-learn developers, B. L., 2020. *sklearn.preprocessing.RobustScaler — scikit-learn 0.24.2 documentation*. [Online]

Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html>  
[Accessed 12 07 2021].

Sinduningrum, E., Suhendra, R. F. & Kamayani, M., 2018. Implementasi CMS pada Media Pembelajaran Mengenal Alat Musik Tradisional Indonesia. *MULTINETICS*, pp. 28-37.

Sonata, F. & Sari, V. W., 2019. Pemanfaatan UML (Unified Modeling Language) Dalam Perancangan Sistem Informasi E-Commerce Jenis Customer-To-Customer. *Jurnal Komunika*, VIII(1), pp. 22-31.



## © Hak Cipta Milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Tomoto, M. N. Y., 2020. Pengembangan Model Migrasi Database Relational ke NoSQL Memanfaatkan Metadata SQL. *TEKNOLOGI ELEKTERIKA* , XVII(2), pp. 1-6.

Vahidah, I. et al., 2020. Pandemi Covid-19: Analisis Perencanaan Pemerintah dan Masyarakat dalam Berbagai Upaya Pencegahan. *Jurnal Manajemen dan Organisasi (JMO)*, 11(3), pp. 179-188.

Wanto, A., 2020. *Data Mining : Algoritma dan Implementasi*. Medan: Yayasan Kita Menulis.

Wanto, A. et al., 2020. *Data Mining : Algoritma dan Implementasi*. Medan: Yayasan Kita Menulis.

Worldometers, 2021. *Coronavirus Update (Live): 103,931,497 Cases and 2,247,715 Deaths from COVID-19 Virus Pandemic - Worldometer*. [Online] available at: <https://www.worldometers.info/coronavirus/> [Accessed 02 02 2021].

Widada, S., 2020. Corona virus diseases(COVID-19); Sebuah tinjauan literatur. *WELLNESS AND HEALTHY MAGAZINE*, 2(1), pp. 187-192.





## LAMPIRAN

Muhammad Ramadhan



Lahir di Jakarta 28 Januari 1997, putra pertama dari dua bersaudara dari pasangan Bapak Sawardi dan Ibu Sri Rahayu. Lulus dari SDN 05 Pagi Cipulir pada tahun 2009, SMPN 48 Jakarta pada tahun 2012 dan SMAN 32 Jakarta pada tahun 2015. Saat ini sedang menempuh Pendidikan Diploma IV Program Studi Teknik Informatika Jurusan Teknik Informatika dan Komputer di Politeknik Negeri Jakarta.



### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta