

TESIS

**CLUSTERING DAMPAK DAN PENANGANAN COVID 19 SE-ASIA
MENGUNAKAN METODE K-MEANS DENGAN VARIABEL-
VARIABEL PADA EPIDEMIOLOGI**



REYNALDI THIMOTIUS ABINENO

NPM: 205303224

PROGRAM STUDI MAGISTER INFORMATIKA

PROGRAM PASCASARJANA

UNIVERSITAS ATMA JAYA YOGYAKARTA

2022

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

CLUSTERING DAMPAK DAN PENANGANAN COVID 19 SE-ASIA MENGGUNAKAN METODE K-MEANS
DENGAN VARIABEL-VARIABEL PADA EPIDEMIOLOGI

yang disusun oleh

Reynaldi Thimotius Abineno

205303224

dinyatakan telah memenuhi syarat pada tanggal 27 Januari 2022

		Keterangan
Dosen Pembimbing 1	: Dr. Ir. Alb. Joko Santoso, M.T.	Telah Menyetujui
Dosen Pembimbing 2	: Paulus Mudjihartono, S.T., M.T., Ph.D	Telah Menyetujui
Tim Penguji		
Penguji 1	: Dr. Ir. Alb. Joko Santoso, M.T.	Telah Menyetujui
Penguji 2	: Dr. Andi Wahyu Rahardjo Emanuel, BSEE., MSSE	Telah Menyetujui

Yogyakarta, 27 Januari 2022

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. A. Teguh Siswanto, M.Sc.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

CLUSTERING DAMPAK DAN PENANGANAN COVID 19 SE-ASIA MENGGUNAKAN METODE K-MEANS
DENGAN VARIABEL-VARIABEL PADA EPIDEMIOLOGI

yang disusun oleh

Reynaldi Thimotius Abineno

205303224

dinyatakan telah memenuhi syarat pada tanggal 27 Januari 2022

Dosen Pembimbing 1
Dosen Pembimbing 2

: Dr. Ir. Alb. Joko Santoso, M.T.
: Paulus Mudjihartono, S.T., M.T., Ph.D

Keterangan

Telah Menyetujui

Telah Menyetujui

Tim Penguji
Penguji 1
Penguji 2

: Dr. Ir. Alb. Joko Santoso, M.T.
: Dr. Andi Wahyu Rahardjo Emanuel, BSEE., MSSE

Telah Menyetujui

Telah Menyetujui

Yogyakarta, 27 Januari 2022

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. A. Teguh Siswanto, M.Sc.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

CLUSTERING DAMPAK DAN PENANGANAN COVID 19 SE-ASIA MENGGUNAKAN METODE K-MEANS
DENGAN VARIABEL-VARIABEL PADA EPIDEMIOLOGI

yang disusun oleh

Reynaldi Thimotius Abineno

205303224

dinyatakan telah memenuhi syarat pada tanggal 27 Januari 2022

Dosen Pembimbing 1
Dosen Pembimbing 2

: Dr. Ir. Alb. Joko Santoso, M.T.
: Paulus Mudjihartono, S.T., M.T., Ph.D

Keterangan

Telah Menyetujui

Telah Menyetujui

Tim Penguji
Penguji 1
Penguji 2

: Dr. Ir. Alb. Joko Santoso, M.T.
: Dr. Andi Wahyu Rahardjo Emanuel, BSEE., MSSE

Telah Menyetujui

Telah Menyetujui

Yogyakarta, 27 Januari 2022

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. A. Teguh Siswanto, M.Sc.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini



UNIVERSITAS ATMA JAYA YOGYAKARTA
PROGRAM PASCASARJANA
PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA

PERNYATAAN

Bersama dengan penelitian ini, maka saya yang bertandatangan dibawah

ini:

Nama : Reynaldi Thimotius Abineno

Nomor Mahasiswa : 205303224

Konsentrasi : *Innovation of Computational Science*

Judul Tesis : CLUSTERING DAMPAK DAN PENANGANAN
COVID 19 SE-ASIA MENGGUNAKAN METODE K-
MEANS DENGAN VARIABEL-VARIABEL PADA
EPIDEMIOLOGI

Dengan lembar pernyataan ini, penulis menyatakan bahwa penelitian tesis ini merupakan hasil dari pemikiran sendiri serta bukan duplikasi pada karya tulis yang ada sebelumnya. Sepengetahuan penulis juga belum ada karya tulis tentang ini yang diterbitkan oleh orang lain. Karya tulis yang jadi rujukan atau acuan dan yang sudah ada digunakan penulis untuk melengkapi penelitian tesis ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 27 Januari 2022

Penulis

Reynaldi Thimotius Abineno

INTISARI

Sampai saat ini pandemi *COVID-19* masih melanda seluruh dunia. Disaat yang bersamaan seluruh negara di dunia tetap berupaya agar bisa cepat keluar dari pandemi ini. Kualitas penanganan kesehatan yang baik tentu dapat membantu suatu negara untuk keluar dari situasi ini. Namun, meski banyaknya penelitian yang dilakukan, hingga saat ini banyak negara masih mencari solusi untuk keluar dari pandemi *COVID-19*.

Oleh karena permasalahan di atas pada penelitian ini akan dilakukan *Clustering* dampak dan penanganan terhadap *COVID-19* pada negara-negara se-asia dengan metode *K-Means* menggunakan variabel atau atribut pada bidang ilmu *Epidemiologi* yaitu seperti *Case Fatality Rate*, *Mortality Rate*, *Recovery Rate* ditambah dengan kepadatan penduduk (*Density*) sebagai variabel bebas yang diolah dan didapatkan dari Raw data kasus *COVID-19*. Dengan tujuan menentukan dampak dan penanganan *COVID-19* dan pengaruh dari kepadatan penduduk (*Density*) terhadap dampak, dan penanganan pandemi *COVID-19* pada suatu negara.

Hasil *clustering* dibagi menjadi 3 cluster dan dianalisis dengan *aggregate value* yaitu nilai Min-Max dan nilai Mean dari masing cluster maupun variabel. Hasil didapatkan bahwa Tingkat Kepadatan Penduduk cukup berpengaruh terhadap tingkat fatalitas kasus dan kematian sebagai dampak dan tingkat kesembuhan sebagai gambaran penanganan. Selain itu ditemukan juga bahwa cluster 2 merupakan cluster yang paling baik baik dari dampak maupun penanganan terhadap pandemi *COVID-19* dengan nilai mean untuk *Case Fatality Rate* sebesar 1.27%, kemudian *Mortality Rate* sebesar 0.009%, *Recovery Rate* 93.56% walaupun nilai Mean *Density* atau Kepadatan penduduknya tergolong sangat tinggi yaitu 1295 populasi/km². Dari segi akurasi dengan karakteristik dataset yang digunakan dalam penelitian ini menggunakan metode *K-Means* lebih unggul dibandingkan *DBSCAN* yang dievaluasi dengan *Davies-Bouldin Index* dimana akurasi terbaik dari metode *K-Means* mencapai 97% sedangkan *DBSCAN* 80%.

Kata Kunci : *K-Means, DBSCAN, Covid-19, Clustering, Epidemiologi.*

ABSTRACT

Until now, the COVID-19 pandemic is still sweeping the world. At the same time, all countries in the world are still trying to get out of this pandemic quickly. The quality of good health care can certainly help a country to get out of this situation. However, despite the large number of studies being carried out, until now many countries are still looking for solutions to get out of the COVID-19 pandemic.

Because of the problems above, in this study, clustering of the impact and handling of COVID-19 in Asian countries with the K-Means method will be carried out using variables or attributes in the field of Epidemiology, namely Case Fatality Rate, Mortality Rate, Recovery Rate plus with population density (Density) as an independent variable which is processed and obtained from Raw data on COVID-19 cases. To determine the impact and handling of COVID-19 and the influence of population density (Density) on the impact and handling of the COVID-19 pandemic in a country.

The results of clustering are divided into 3 clusters and analyzed by aggregate value, namely the Min-Max value and the Mean value of each cluster and variable. The results show that the level of population density is quite influential on the level of case fatality and death as an impact and the rate of recovery as an illustration of treatment. In addition, it was also found that cluster 2 is the best cluster both in terms of impact and handling of the COVID-19 pandemic with a mean value for Case Fatality Rate of 1.27%, then Mortality Rate of 0.009%, Recovery Rate of 93.56% even though the Mean Density value is the population is classified as very high, namely 1295 population/km². In terms of accuracy, the characteristics of the dataset used in this study using the K-Means method are superior to DBSCAN evaluated by the Davies-Bouldin Index where the best accuracy of the K-Means method is 97% while DBSCAN is 80%.

Keywords: K-Means, DBSCAN, Covid-19, Clustering, Epidemiologi.

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas segala hikmat dan anugerah-Nya sehingga penulis dapat menyelesaikan tesis ini dengan baik. Adapun tesis ini berjudul **“Clustering Dampak dan Penanganan COVID 19 Se-Asia Menggunakan Metode K-Means Dengan Variabel-Variabel Pada Epidemiologi”**. Penyusunan tesis ini sebagai salah satu syarat akademik dalam meraih kelulusan dan mendapatkan gelar Strata Dua (S2) pada **Program Pascasarjana, Magister Informatika, Universitas Atma Jaya Yogyakarta**.

Penulis menyadari bahwa telah mendapat bantuan dari berbagai pihak baik itu berupa ide, pikiran, tenaga, waktu, dukungan maupun doa kepada penulis baik secara langsung maupun tidak langsung. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Tuhan Yang Yesus Kristus yang telah memberikan hikmat dan anugerah-Nya kepada penulis.
2. Kedua orang tua, Adik-adik, dan Keluarga yang telah memberikan dukungan dalam kelancaran penelitian ini.
3. Bapak Yonathan Dri Handarkho ST.,M.Eng.,Ph.D. selaku Ketua Program Studi Magister Informatika, Program Pascasarjana, Universitas Atma Jaya Yogyakarta.
4. Bapak Dr. Ir. Alb. Joko Santoso, M.T. selaku Dosen Pembimbing I yang telah bersedia memberi bimbingan, gagasan ide, koreksi dan pengarahan kepada penulis sehingga Tesis ini dapat diselesaikan dengan baik.
5. Bapak Paulus Mudjihartono, S.T., M.T., Ph.D. selaku Dosen Pembimbing II yang telah bersedia memberi bimbingan, gagasan ide, koreksi dan pengarahan kepada penulis sehingga Tesis ini dapat diselesaikan dengan baik.

6. Bapak/Ibu, selaku dosen penguji, terimakasih atas kritik dan saran yang diberikan kepada penulis.
7. Seluruh Dosen, Staf Pengajar, dan Karyawan Program Pascasarjana, Universitas Atma Jaya Yogyakarta yang telah membantu penulis selama masa kuliah di Program Studi Magister Informatika, Program Pascasarjana, Universitas Atma Jaya Yogyakarta.
8. Semua teman dan sahabat tahun angkatan 2020 Magister Informatika, Universitas Atma Jaya Yogyakarta yang telah memberikan semangat.
9. Semua orang yang secara tidak langsung memberikan dukungan kepada penulis.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna karena keterbatasan waktu dan pengetahuan yang dimiliki penulis. Oleh karena itu segala bentuk kritik maupun saran yang sifatnya membangun akan sangat diharapkan kedepannya. Akhir kata, semoga tesis ini dapat bermanfaat dan berguna bagi semua pihak.

Yogyakarta, 27 Januari 2022

Reynaldi Thimotius Abineno

MOTTO

Matius 6:33

Tetapi Carilah Dahulu Kerajaan Allah dan Kebenarannya, maka semuanya itu akan ditambahkan kepadamu.



DAFTAR ISI

HALAMAN PENGESAHAN TESIS	i
HALAMAN PENGESAHAN TESIS	ii
HALAMAN PENGESAHAN TESIS	iii
HALAMAN PERNYATAAN	iv
HALAMAN INTISARI	v
HALAMAN ABSTRAK	vi
HALAMAN KATA PENGANTAR	vii
HALAMAN MOTTO	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Keaslian Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Tujuan Penelitian	4
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
BAB III LANDASAN TEORI	17
3.1 COVID-19	17
3.2 Machine Learning	17
3.3 Clustering	18
3.4 K-Means	19
3.5 DBSCAN	20
3.6 Biostatistika dan Epidemiologi	21
3.7 Davies-Bouldin Index	22
BAB IV METODOLOGI PENELITIAN	24
4.1 Alat dan Bahan	24

4.1.1 Pengumpulan Data.....	24
4.1.2 Dataset	26
4.1.3 Perangkat Lunak	29
4.1.4 Perangkat Keras	30
4.2 Identifikasi Data	30
4.2.1 Normalize	30
4.3 Alur Implementasi Algoritma.....	31
4.4 Alur Penelitian.....	33
4.5 Implementasi Pada RapidMiner	34
4.6 Model Analisis Hasil Clustering.....	36
BAB V HASIL DAN PEMBAHASAN.....	38
5.1 Hasil Clustering Dengan K-Means.....	38
5.1.1 Hasil Clustering Model 1 (K-Means)	38
5.1.2 Hasil Clustering Model 2 (K-Means)	39
5.1.3 Hasil Clustering Model 3 (K-Means)	40
5.2 Hasil Clustering Dengan DBSCAN	41
5.2.1 Hasil Clustering Model 1 (DBSCAN).....	42
5.2.2 Hasil Clustering Model 2 (DBSCAN).....	43
5.2.3 Hasil Clustering Model 3 (DBSCAN).....	44
5.3 Analisis Hasil Clustering K-Means dan DBSCAN	45
5.3.1 Pengaruh Kepadatan Penduduk Pada Tingkat Fatalitas Kasus	45
5.3.2 Pengaruh Kepadatan Penduduk Pada Tingkat Kematian	49
5.3.3 Pengaruh Kepadatan Penduduk Pada Tingkat Kesembuhan.....	53
5.4 Model Clustering Dengan Semua Variabel	57
5.4.1 Model Clustering Dengan Semua Variabel (K-Means)	58
5.4.2 Model Clustering Dengan Semua Variabel (DBSCAN)	61
5.5 Evaluasi Perbandingan Metode K-Means dan DBSCAN	66
BAB VI KESIMPULAN DAN SARAN	68
6.1 Kesimpulan	68
6.2 Saran.....	69
Daftar Pustaka	70

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka	9
Tabel 4.1 Raw Data COVID-19 Benua Asia	24
Tabel 4.2 Raw Data Populasi Benua Asia	25
Tabel 4.3 Dataset.....	26
Tabel 4.4 Model Analisis Hasil Clustering.....	36
Tabel 5.1 Kriteria Cluster Pada Model 1	46
Tabel 5.2 Cluster Pada Model.....	47
Tabel 5.3 Kriteria Cluster Pada Model 2	50
Tabel 5.4 Cluster Pada Model 2.....	51
Tabel 5.5 Kriteria Cluster Pada Model 3	54
Tabel 5.6 Cluster Pada Model 3.....	55
Tabel 5.7 Clustering Dengan Semua Variabel Pada K-Means	58
Tabel 5.8 <i>Aggregate Value</i> K-Means Clustering	60
Tabel 5.9 Clustering Dengan Semua Variabel Pada DBSCAN.....	62
Tabel 5.9 <i>Aggregate Value</i> K-Means Clustering	65
Tabel 5.10 Perbandingan Akurasi K-Means dan DBSCAN	65

DAFTAR GAMBAR

Gambar 4.1 Alur Implementasi Algoritma K-Means	31
Gambar 4.2 Alur Implementasi Algoritma DBSCAN	32
Gambar 4.3 Alur Penelitian.....	33
Gambar 4.4 Model Implementasi K-Means Pada RapidMiner.....	34
Gambar 4.5 Model Implementasi DBSCAN Pada RapidMiner	35
Gambar 5.1 Scatter Plot Model 1 (K-Means)	39
Gambar 5.2 Scatter Plot Model 2 (K-Means)	40
Gambar 5.3 Scatter Plot Model 3 (K-Means)	41
Gambar 5.4 Scatter Plot Model 1 (DBSCAN).....	42
Gambar 5.5 Scatter Plot Model 2 (DBSCAN).....	43
Gambar 5.6 Scatter Plot Model 3 (DBSCAN).....	44