



**ANALISIS PERBANDINGAN ALGORITMA CLUSTERING DBSCAN
DENGAN SNN UNTUK PEMETAAN DAERAH PENYEBARAN COVID-
19 DI DKI JAKARTA**

TUGAS AKHIR

Adam Maulana Khosasih
41517010099

**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS ILMU
KOMPUTER UNIVERSITAS MERCU BUANA JAKARTA**

2021

MERCU BUANA



**ANALISIS PERBANDINGAN ALGORITMA CLUSTERING DBSCAN
DENGAN SNN UNTUK PEMETAAN DAERAH PENYEBARAN COVID-
19 DI DKI JAKARTA**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Adam Maulana Khosasih 41517010099

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA

2021
MERCU BUANA

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41517010099

Nama : Adam Maulana Khosasih

Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering DBSCAN Dengan SNN Untuk Pemetaan Daerah Penyebaran COVID-19 Di DKI Jakarta.

Menyatakan bahwa laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, Januari 2021



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Adam Maulana Khosasih
NIM : 41517010099
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering DBSCAN Dengan SNN Untuk Pemetaan Daerah Penyebaran COVID-19 Di DKI Jakarta.

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, Januari 2021

UNIVERSITAS
MERCU BUANA



Adam Maulana Khosasih

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Adam Maulana Khosasih
 NIM : 41517010099
 Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering
 DBSCAN Dengan SNN Untuk Pemetaan Daerah
 Penyebaran COVID-19 Di DKI Jakarta.

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan
		Jurnal Nasional Terakreditasi	
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
Disubmit/dipublikasikan di :	Nama Jurnal	: Jurnal Informatika	
	ISSN	:	
	Link Jurnal	: http://jurnalnasional.ump.ac.id/index.php/JUITA/author	
	Link File Jurnal Jika Sudah di Publish	:	

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui
 Dosen Pembimbing TA

Jakarta, 08 Januari 2021


 Dr. Devi Fitriana, S.Kom., M.TI


 Adam Maulana Khosasih



LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010099
Nama : ADAM MAULANA KHOSASIH
Judul Tugas Akhir : ANALISIS PERBANDINGAN ALGORITMA
CLUSTERING DBSCAN DENGAN SNN UNTUK
PEMETAAN DAERAH PENYEBARAN
COVID-19 DI DKI JAKARTA

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 08 Februari 2021



(Dr. Mujiono Sadikin, M.T.)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010099
Nama : ADAM MAULANA KHOSASIH
Judul Tugas Akhir : ANALISIS PERBANDINGAN ALGORITMA
CLUSTERING DBSCAN DENGAN SNN UNTUK
PEMETAAN DAERAH PENYEBARAN
COVID-19 DI DKI JAKARTA

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 08 Febuari 2021



(Sabar Rudiarto, M.Kom)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010099
Nama : ADAM MAULANA KHOSASIH
Judul Tugas Akhir : ANALISIS PERBANDINGAN ALGORITMA
CLUSTERING DBSCAN DENGAN SNN UNTUK
PEMETAAN DAERAH PENYEBARAN
COVID-19 DI DKI JAKARTA

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 08 Febuari 2021



(Hery Derajat Wijaya, S.Kom, MM)

UNIVERSITAS
MERCU BUANA

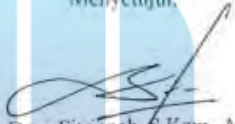
LEMBAR PENGESAHAN

NIM : Adam Maulana Khosasih
Nama : 41517010099
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering DBSCAN
Dengan SNN Untuk Pemetaan Daerah Penyebaran
COVID-19 Di DKI Jakarta.


Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 08 Februari 2021

Menyetujui,


(Dr. Devi Fitriana, S.Kom., M.TI)
Dosen Pembimbing

Mengetahui,


(Diky Firdaus, S.Kom., MM)
Koord. Tugas Akhir Teknik Informatika


(Desi Ramayanti, S.Kom., MT)
Ka. Prodi Teknik Informatika

ABSTRAK

Nama : Adam Maulana Khosasih
NIM : 41517010099
Pembimbing TA : Dr. Devi Fitriana, S.Kom, M.TI
Judul : Analisa Perbandingan Algoritma Clustering
DBSCAN Dengan SNN Untuk Pemetaan Daerah
Penyebaran COVID-19 Di DKI Jakarta.

Covid-19 merupakan penyakit yang menyerang saluran pernafasan sehingga menyebabkan demam tinggi, batuk, flu, sesak nafas serta nyeri tenggorokan. Dengan semakin meluasnya penyebaran Covid-19 di berbagai wilayah, khususnya di wilayah Provinsi DKI Jakarta, menyebabkan keluarnya kebijakan Pembatasan Sosial Berskala Besar (PSBB) di setiap wilayahnya. Untuk mencegah penyebaran COVID-19, khususnya di wilayah Provinsi DKI Jakarta, dibutuhkan pemetaan penyebaran COVID-19. Tujuan dari penelitian ini adalah untuk membandingkan antara Algoritma clustering DBSCAN dengan Shared Nearest Neighbor dalam mengidentifikasi area penyebaran COVID-19 di Provinsi DKI Jakarta. Dari kedua algoritma tersebut dapat disimpulkan bahwa Algoritma DBSCAN dapat lebih baik digunakan untuk memetakan daerah penyebaran COVID-19 di Provinsi DKI Jakarta, dengan hasil validasi menggunakan Dunn Index sebesar 0,99, sedangkan SNN sebesar 0,95. Hasil profiling penyebaran COVID-19 di wilayah Provinsi DKI Jakarta yang didapat dari Algoritma DBSCAN dengan jumlah cluster sebanyak 2 yaitu, cluster rawan dan waspada.

Kata kunci:

Covid-19, Clustering, DBSCAN, SNN

UNIVERSITAS
MERCU BUANA

ABSTRACT

Name : Adam Maulana Khosasih
Student Number : 41517010099
Counsellor : Dr. Devi Fitriana, S.Kom, M.TI
Title : Analisa Perbandingan Algoritma Clustering
DBSCAN Dengan SNN Untuk Pemetaan Daerah
Penyebaran COVID-19 Di DKI Jakarta.

Covid-19 is a disease that attacks the respiratory tract, causing high fever, cough, flu, shortness of breath and sore throat. With the increasingly widespread spread of Covid-19 in various regions, especially in the DKI Jakarta Province, this has led to the issuance of a Large-Scale Social Restriction (PSBB) policy in each region. To prevent the spread of COVID-19, especially in the DKI Jakarta Province, mapping the causes of COVID-19 is needed. The purpose of this study is to compare the DBSCAN clustering algorithm with Shared Nearest Neighbor in identifying the area of the spread of COVID-19 in DKI Jakarta Province. From the two algorithms, it can be concluded that the DBSCAN Algorithm can be better used to map the area of the spread of COVID-19 in DKI Jakarta Province, with validation results using the Dunn Index of 0.99, while the SNN is 0.95. The results of profiling the spread of COVID-19 in the DKI Jakarta Province area were obtained from the DBSCAN Algorithm with 2 clusters, namely vulnerable and alert clusters.

Key words:

Covid-19, Clustering, DBSCAN, SNN



UNIVERSITAS
MERCU BUANA

KATA PENGANTAR

Puji syukur atas kehadiran Tuhan Yang Maha Esa yang telah melimpahkan nikmat, serta hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “Analisa Perbandingan Algoritma Clustering DBSCAN Dengan SNN Untuk Pemetaan Daerah Penyebaran COVID-19 Di DKI Jakarta” dengan baik. Laporan Tugas Akhir ini merupakan salah satu persyaratan wajib untuk menyelesaikan Program Studi Strata Satu (S1) pada jurusan Informatika, Universitas Mercu Buana.

Dalam penyelesaian laporan Tugas Akhir ini, penulis mendapatkan banyak dukungan, bantuan, serta sumbangan ide maupun pikiran dari berbagai pihak. Untuk itu, penulis menyampaikan ucapan terimakasih kepada :

1. Ibu Dr. Devi Fitriana, S.Kom., MTI, selaku dosen pembimbing Tugas Akhir yang telah meluangkan waktunya untuk memberikan bimbingan serta arahan dalam penyusunan Tugas Akhir ini.
2. Ibu Desi Ramayanti, S.Kom., MT, selaku Ka. Prodi Teknik Informatika dan Dosen Akademik, Fakultas Ilmu Komputer Universitas Mercu Buana.
3. Bapak Dicky Firdaus, S.Kom., M.M selaku Koordinator Tugas Akhir Jurusan Informatika, Universitas Mercu Buana.
4. Orang tua yang senantiasa memberikan doa dan dukungan.
5. Teman-teman yang selalu memberi semangat dan memberi motivasi kepada penulis selama pelaksanaan tugas akhir.
6. Semua pihak yang telah membantu dalam penyusunan Tugas Akhir ini yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa penyusunan laporan Tugas Akhir ini masih jauh dari kesempurnaan dan masih terdapat banyak kekurangan. Walaupun demikian, penulis telah berusaha semaksimal mungkin untuk mendapatkan hasil yang terbaik. Penulis berharap tugas akhir ini dapat bermanfaat bagi pembaca guna menambah pengetahuan dan wawasan serta pembaca dapat memberikan kritik maupun saran.

Jakarta, Febuari 2021
Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR ...	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR	iv
LEMBAR PERSETUJUAN PENGUJI	v
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT	x
KATA PENGANTAR	xi
DAFTAR ISI	xii
NASKAH JURNAL	1
KERTAS KERJA	9
BAB 1. LITERATUR REVIEW	10
BAB 2. SOURCE CODE	17
BAB 3. DATASET	24
BAB 4. TAHAPAN EKSPERIMEN	25
BAB 5. HASIL SEMUA EKSPERIMEN	28
DAFTAR PUSTAKA.....	33
LAMPIRAN DOKUMEN HAKI	35
LAMPIRAN KORESPONDENSI	36

NASKAH JURNAL

COMPARISON ANALYSIS OF DBSCAN CLUSTERING ALGORITHM WITH SNN FOR MAPPING COVID-19 SPREAD AREA IN DKI JAKARTA

Adam Maulana Khosasih¹, Devi Fitriyah²

¹*Informatika/Ilmu Komputer-Universitas Mercubuana*, ²*Informatika/Ilmu Komputer-Universitas Mercubuana*
adammaulana1@gmail.com, devi.fitriyah@mercubuana.ac.id

Abstract - Covid-19 is a disease that attacks the respiratory tract, causing high fever, cough, flu, shortness of breath and sore throat. With the increasingly widespread spread of Covid-19 in various regions, especially in the DKI Jakarta Province, this has led to the issuance of a Large-Scale Social Restriction (PSBB) policy in each region. To prevent the spread of COVID-19, especially in the DKI Jakarta Province, mapping the causes of COVID-19 is needed. The purpose of this study is to compare the DBSCAN clustering algorithm with Shared Nearest Neighbor in identifying the area of the spread of COVID-19 in DKI Jakarta Province. From the two algorithms, it can be concluded that the DBSCAN Algorithm can be better used to map the area of the spread of COVID-19 in DKI Jakarta Province, with validation results using the Dunn Index of 0.99, while the SNN is 0.95. The results of profiling the spread of COVID-19 in the DKI Jakarta Province area were obtained from the DBSCAN Algorithm with 2 clusters, namely vulnerable and alert clusters.

Keywords: Covid-19, Clustering, DBSCAN, SNN

I. INTRODUCTION

Coronavirus Disease (Covid-19). In 2020, there's a spread of new type of corona virus (SARS-CoV-2) that caused a diseases called coronavirus disease 2019 (COVID-19). The virus was first discovered in Wuhan, China[1]. The virus speeded and became a pandemic and pushed government to put a regional lockdown for each city in Indonesia, also called Pembatasan Sosial Berskala Besar (PSBB) in Bahasa Indonesia. Fever, fatigue, dry cough, and chronic asphyxiate are some of the symptoms that occurred if someone caught the virus[2].

World Health Organization (WHO) declare Covid-19 as global pandemic on March 11th, 2020[3]. At high economic cost, countries around the world take several unprecedented measure to try to contain the spread of the virus, such as large scale quarantine and isolation, closing the border, limitation on public activities, and even a national lockdown[4]. By April 4th, 2020, 1,051,635 people from 208 countries reported has infected by the SARS-CoV-2, the virus that caused COVID-19, thus caused more than 75,000 death around the world. Although the number of new cases has been decreasing in China, new cases number drastically increasing in Europe and United States, total of death number in Italia, Spain, France and US exceeds the number in China. Based on laboratory reports of confirmed case and the number of death per country or region, on April 20th, 2020, Indonesia are on the second place (with most reported case) in South East Asia[5]. So, it's is important to be able to identified the risks of the COVID-19 virus contagion on any area. This will also help the government to keep the people save, and to provide better care and helps to areas with reported infected people.

DBSCAN algorithm is the method that typically used for special clustering, subsequently to be able to profile and identify which area with most number of COVID-19 patients, it requires a clustering method that comparing

between DBSCAN algorithm versus SNN algorithm. DBSCAN is an algorithm that used to data clustering or grouping. Every object inside the cluster must contain at least several data. Any object outside the cluster considered as noise[6]. On the other hand, Shared Nearest Neighbour (SNN) algorithm produced a neighbour graphs, a graphs that use a similarity between each data based on number of shared nearest neighbour[7].

The research conducted by using python programme and Dunn's validity index (DN) to calculate the minimum value of dissimilarities function between two clusters as separation and the maximum value of cluster's diameter as compactness[8]. Dunn Index is a cluster validation method that based on geometric calculation of compactness of each cluster and separation between clusters[9]. The higher the Dunn Index value, the better the cluster formed[10].

II. METHODE

The primary object of this research is to do a comparison between DBSCAN algorithm versus SNN algorithm on area mapping of the spreads of COVID-19 in Special Capital Region of Jakarta (DKI Jakarta). Steps of the research as seen on picture below:

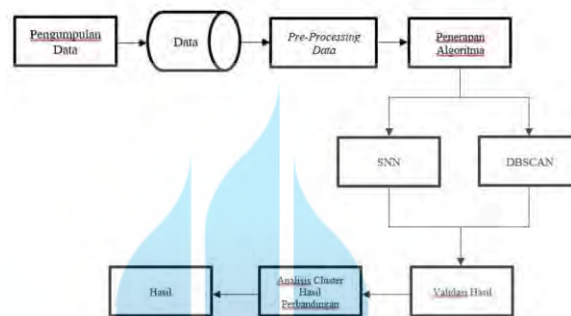


Fig 1. Steps of Research

A. Data Collection.

The research use data of sub-district in DKI Jakarta that affected by COVID-19 in corona.jakarta.co.id website. Data collection done from March 26th, 2020 to August 26th, 2020. There are total of 267 data collected, as shown on Picture 2.

	nama_kecamatan	nama_kelurahan	Isolasi di Rumah	POSITIF	Dirawat	Sembuh	Meninggal	KONTAK ERAT
0	PADEMANGAN	ANCOL	34	67	6	55	2	405
1	TAMBORA	ANGKE	25	77	2	64	2	380
2	KRAMAT JATI	BALE KAMBANG	76	55	1	49	1	306
3	JATINEGARA	BALI MESTER	9	11	0	10	1	49
4	CIPAYUNG	BAMBU APUS	36	58	0	45	0	215

Fig 2. Preliminary Data

DBSCAN algorithm versus SNN algorithm use spatial data that requires latitude and longitude parameters. So latitude and longitude parameters acquired from *Latlong.net* website added to the preliminary data. To make sure the accuracy of the latitude and longitude parameters writes did a double check through Google Maps. Final Data are shown on Picture 3.

	nama_kecamatan	nama_kelurahan	Latitude	Longitude	Isolasi di Rumah	POSITIF	Dirawat	Sembuh	Meninggal	KONTAK ERAT
0	PADEMANGAN	ANCOL	-6.12873	106.832314	34	67	6	55	2	405
1	TAMBORA	ANGKE	-6.14519	106.799150	25	77	2	64	2	380
2	KRAMAT JATI	BALE KAMBANG	-6.28200	106.799150	76	55	1	49	1	306
3	JATINEGARA	BALI MESTER	-6.22077	106.867210	9	11	0	10	1	49
4	CIPAYUNG	BAMBU APUS	-6.32372	106.895840	36	58	0	45	0	215

Fig 3. Final Data

B. Data Pre-Processing.

On this step, pre-processing done by using Encoder Label that turns float data into integer. It necessary because DBSCAN and SNN algorithm required integer data as the input. The result of Label Encoder shown on Picture 4.

	nama_kecamatan	nama_kelurahan	Latitude	Longitude	Isolasi di Rumah	POSITIF	Dirawat	Sembuh	Meninggal	KONTAK ERAT
0	28	0	-6.12873	106.832314	34	67	6	55	2	405
1	40	1	-6.14519	106.799150	25	77	2	64	2	380
2	23	2	-6.28200	106.799150	76	55	1	49	1	306
3	11	3	-6.22077	106.867210	9	11	0	10	1	49
4	5	4	-6.32372	106.895840	36	58	0	45	0	215
...
262	26	262	-6.20562	106.869400	71	103	4	88	4	449
263	26	263	-6.19483	106.870480	45	62	5	44	3	344
264	18	264	-6.16364	106.855420	32	69	2	59	6	197
265	42	265	-6.12080	106.878030	85	104	10	86	3	479
266	9	266	-6.15527	106.774971	21	85	1	81	3	524

Fig 4. Result of Label Encoder

After altering the type of data using Label Encoder, next pre-processing step is to use Minmax Scaler to transform or normalised original data with different range into range between 1 to 0. Result from Minmax Scaler shown on Picture 5.

UNIVERSITAS
MERCU BUANA

	nama_kecamatan	nama_kelurahan	Latitude	Longitude	Isolasi di Rumah	POSITIF	Dinawat	Sembuh	Meninggal	KONTAK ERAT
0	0.180278	0.000000	0.303194	0.413306	0.608158	0.290576	0.629653	0.565400	0.300535	0.623055
1	0.169632	0.004468	0.339374	0.367601	0.810968	0.287612	0.405844	0.385598	0.302201	0.509502
2	0.135384	0.014589	0.325071	0.620298	0.618050	0.164366	0.388358	0.370085	0.216936	0.129840
3	0.010629	0.029412	0.289351	0.434537	0.255523	0.254844	0.400219	0.531711	0.206766	0.652092
4	0.092395	0.023288	0.351218	0.484502	0.174339	0.297412	0.361268	0.255343	0.115072	0.486970
...
262	0.207455	0.980785	0.401629	0.520137	0.548090	0.244195	0.429365	0.488529	0.168567	0.434799
263	0.153545	0.991111	0.312733	0.467423	0.515607	0.394947	0.504746	0.591032	0.162577	0.516871
264	0.087142	1.000000	0.400091	0.440990	0.323113	0.278657	0.322090	0.729501	0.202444	0.542801
265	0.221620	0.992225	0.392349	0.564518	0.891901	0.280282	0.746689	0.586364	0.235772	0.516993
266	0.238526	0.990019	0.321217	0.309138	0.103609	0.189255	0.376660	0.417249	0.229414	0.417269

Fig 5. Hasil Minmax Scaler

Pre-processing is a step to eliminate unnecessary part of the data or text cleaning that done to get qualified data to be executed into the algorithms[11]. Pre-processing required to get a structured data. The object of pre-processing on data mining is to alter data format into a format that suit users requirements[12].

C. Implementation of Clustering Method.

Next step is to study which algorithm to use to design a better method. The program need to be tested to see in the algorithm works normally and to check the quality of clusters formed[13].

(1) DBSCAN Algorithm:

On this steps, clustering done to pre-processed data using DBSCAN Algorithm[14]. DBSCAN Algorithm is a partition clustering, that area with high density considered as cluster and area with low density or not included in any cluster considered as noise[15]. DBSCAN Algorithm is an algorithm used for data classification or data grouping. Any object in the cluster must contain a minimum amount of data. Any object outside the cluster considered as noise[6]. DBSCAN can differentiate noise in a large spatial database[16]. DBSCAN will do clustering bases on the input parameters, that are *epsilon* (ϵ) and MinPts. The number of cluster formed by DBSCAN will be depend on those two parameters[14]. Algorithm gather area with high density into cluster and found noise cluster in the scatter spatial database[17].

(2) SNN Algorithm:

Shared Nearest Neighbour (SNN) algorithm produced a neighbour graphs, a graphs that use a similarity between each data based on number of shared nearest neighbour[7]. Because of SNN algorithm similarity to clustering process, it introduced as way to solve distance measurement problem[18]. SNN is not depends on density, meaning it will keep the chain in consistent area and cut the chain in transition area. It's an important feature, because diversity of cluster can caused more problem in clustering process[19]. SNN will form clusters based on the input parameter, neighbour number dan min shared neighbour proportion. The number of cluster formed will be depend on those two parameters

D. Comparison Result Analysis

Results from each algorithm verified by using tuning scores. The verified results are compared to get the highest verified results to analyse the research result. Then, compare the graphics and clusters formed by each algorithm.

III. RESULT AND DISCUSSION

Results from two proposed algorithm had been compared. Trial and simulation implemented using python programming. Using Dunn Index as validation method on the research. The result on DBSCAN Algorithm using *epsilon* 0.5, 0.25, 0.35, 0.55 and 0.60 with *Minimum Points* 10 are shown on TABLE I.

TABLE I. DBSCAN VALIDATION RESULT

Eps	MinPts	Cluster	DunnIndex
0.5	10	2	0.89
0.25	10	4	0.89
0.35	10	2	0.90
0.55	10	2	0.91
0.60	10	2	0.99

After doing impression analysis to the clustering and do the experiment to find the necessity of tuning score for clustering. There are 5 experiments using different tuning scores. Clustering result from the experiment shown in Picture 7.

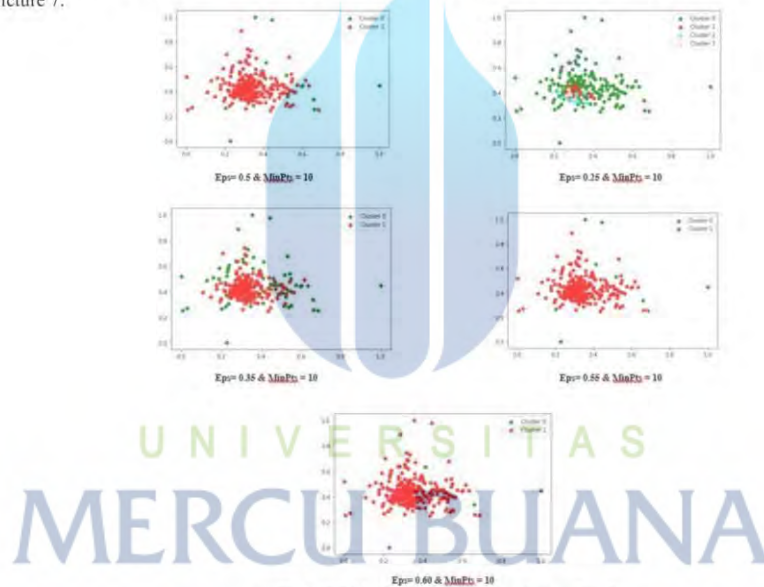


Fig 7. DBSCAN Method's Experiment Result

Picture 7 shown that changes on tuning scores affect the number of clusters formed. Higher tuning scores result a higher density. The experiment result shown that Eps 0.60 and MinPts 10 parameters results the higher density. Analysing the experiment of DBSCAN algorithm only by changing the eps value, because it's the eps value that affect the clustering the most. Eps value is used to form cluster with determined radius[20].

Two clusters formed by DBSCAN algorithm, cluster -1 and cluster 0. Cluster -1 can be interpreted as vulnerable area and cluster 0 can be interpreted as alert area.

In cluster -1 as vulnerable cluster, there are 4 sub-districts from 4 district and in cluster 0, there are 263 sub-districts from 44 district.

TABLE II. SNN VALIDATION RESULT

Neighbor_num	Min_SN_Proportion	Cluster	DunnIndex
15	0.15	8	0.89
15	0.25	4	0.89
15	0.35	5	0.90
15	0.60	5	0.91
30	0.70	3	0.95

Table II shown the Result on experiment of SNN method, also using phyton programming and Dunn Index for validation, parameters use are *Neighbor_num* 15,30 and *Min_SN_Proportion* 0.15, 0.25, 0.35, 0.60, 0.70.

There are 5 experiment executed using *Neighbor num* and *Min Shared Num Proportion*, best result shown on fifth experiment with tuning score *Neighbor num: 30* and *Min Shared Num Proportion: 0.70* resulting in 3 clusters with validation index of 0.95.

On the other hand, using SNN method to experiment with tuning scores and 5 different densities. Tunning scores with good density are *Neighbor Num 15* dan *Min Shared Num Proportion 0.70*. the clustering result are shown on Picture 8.

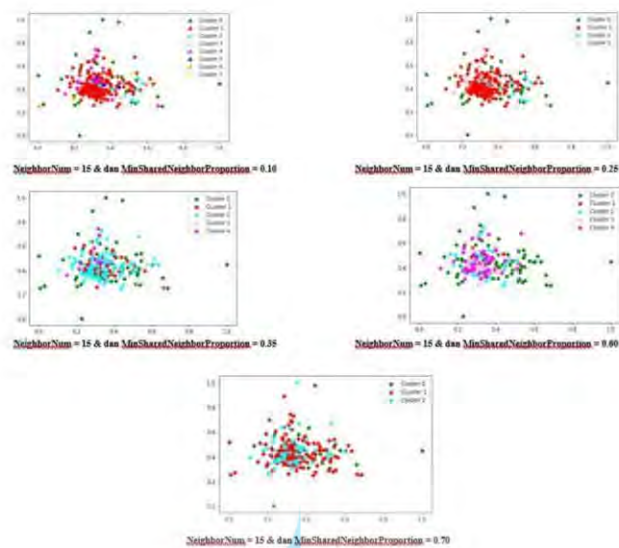


Fig 8. SNN Method's Experiment Result

There are three clusters form by experiment using SNN algorithm, cluster -1, cluster 0 and cluster 1. Cluster -1 can be interpreted as vulnerable area and cluster 0 can be interpreted as alert area, cluster 1 can be interpreted as save area.

Included in cluster -1 as vulnerable area are 20 sub-districts from 16 district, 190 sub-districts from 41 district included in cluster 0 as alert area, and in cluster 1 as save area are 57 sub-districts from 23 district.

After this, results of the experiment will be analyse to see if the object of the research have been met or not[14]. As shown on Table I, DBSCAN algorithm have higher validation index compare to SNN algorithm. The experiments of each methods executed by doing five trial process resulting in number of clusters with highest validation index.

IV. CONCLUSION

Based on the experiment using DBSCAN algorithm and SNN algorithm to do a profiling area affected by COVID-19 in DKI Jakarta, also using Dunn Index for validation, the result shown DBSCAN algorithm get the higher validation index of 0.99 compared to SNN algorithm with 0.95 validation index.

As the conclusion, compared to SNN algorithm, DBSCAN algorithm are the better method to be used to help public and government to do profile areas affected by COVID-19 within DKI Jakarta. As the result, there are 2 clusters: vulnerable area and alert area. Included in vulnerable area are 4 sub-districts from 4 district, while other 263 sub-districts from 44 district are on the alert area.

REFERENCES

- [1] Yuliana, "Corona virus diseases (Covid -19); Sebuah tinjauan literatur," *Wellness Heal. Mag.*, vol. 2, no. 1, pp. 187–192, 2020, [Online]. Available: <https://wellness.journalpress.id/wellness/article/view/v1i218wh>.
- [2] K. C. Liu *et al.*, "CT manifestations of coronavirus disease-2019: A retrospective analysis of 73 cases by disease severity," *Eur. J. Radiol.*, vol. 126, no. February, p. 108941, 2020, doi: 10.1016/j.ejrad.2020.108941.
- [3] R. Djalante *et al.*, "Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020," *Prog. Disaster Sci.*, vol. 6, p. 100091, 2020, doi: 10.1016/j.pdisas.2020.100091.
- [4] D. F. Gudbjartsson, D. Ph, A. Helgason, D. Ph, H. Jonsson, and D. Ph, "Early Spread of SARS-Cov-2 in the Icelandic Population," *medRxiv*, pp. 1–34, 2020, doi: 10.1056/NEJMoa2006100.
- [5] B. B. Practice, "Coronavirus disease 2019," *World Heal. Organ.*, vol. 2019, no. March, p. 2633, 2020, doi: 10.1001/jama.2020.2633.
- [6] A. S. Devi, I. K. G. D. Putra, and I. M. Sukarsa, "Implementasi Metode Clustering DBSCAN pada Proses Pengambilan Keputusan," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 6, no. 3, p. 185, 2015, doi: 10.24843/lkjiti.2015.v06.i03.p05.
- [7] R. F. Zainal and A. Djunaedy, "Algoritma Shared Nearest Neighbor Berbasis Data Shrinking," *JUTI J. Ilm. Teknol. Inf.*, vol. 7, no. 1, p. 3, 2008, doi: 10.12962/j24068535.v7i1.a56.
- [8] A. F. Khairati, A. A. Adlina, G. F. Hertono, and B. D. Handari, "Kajian Indeks Validitas pada Algoritma K-Means Enhanced dan K-Means MMCA," *Pros. Semin. Nas. Mat.*, vol. 2, pp. 161–170, 2019.
- [9] S. Monalisa, "Klusterisasi Customer Lifetime Value dengan Model LRFM menggunakan Algoritma K-Means," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 247, 2018, doi: 10.25126/jtiik.201852690.
- [10] A. D. Savitri, F. A. Bachtiar, and N. Y. Setiawan, "Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan (Studi Kasus: Belle Crown Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 9, pp. 2957–2966, 2018.
- [11] S. Sanjaya, S. Sanjaya, and E. A. Absar, "Pengelompokan Dokumen Menggunakan Winnowing Fingerprint dengan Metode K-Nearest Neighbour," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 1, no. 2, pp. 50–56, 2015, [Online]. Available: <http://ejournal.uin-suska.ac.id/index.php/coreit/article/view/1229>.
- [12] P. Meilina, "Penerapan Data Mining dengan Metode Klasifikasi Menggunakan Decision Tree dan Regresi," *J. Teknol. Univ. Muhammadiyah Jakarta*, vol. 7, no. 1, pp. 11–20, 2015, [Online]. Available: jurnal.ftumj.ac.id/index.php/jurtek.
- [13] I. Id, "Modifikasi DBSCAN (Density-Based Spatial Clustering with Noise) Pada Objek 3 Dimensi," *J. Komput. Terap.*, vol. 3, no. 1, pp. 41–52, 2017.
- [14] Aurelius, "Implementasi Clustering Pada Algoritma DBSCAN Untuk Deteksi Kejadian Bencana Alam Pada Microblogging Twitter," *Duke Law J.*, vol. 1, no. 1, pp. 1–13, 2019, doi: 10.1017/CBO9781107415324.004.
- [15] N. Arsih, N. Hajarisman, and S. Darwis, "Metode Pengclusteran Berbasis Densitas Menggunakan Algoritma DBSCAN," pp. 153–163, 2015, [Online]. Available: <http://karyailmiah.unisba.ac.id/index.php/statistika/article/view/3796/pdf>.
- [16] A. Kristianto, E. Sedyono, and K. D. Hartomo, "Implementation dbscan algorithm to clustering satellite surface temperature data in indonesia," *Regist. J. Ilm. Teknol. Sist. Inf.*, vol. 6, no. 2, p. 109, 2020, doi: 10.26594/register.v6i2.1913.
- [17] P. Desiana and W. Ayu, "Perbandingan Kinerja Fuzzy C-Means dan DBSCAN Dalam Segmentasi Citra USG Kepala Janin."
- [18] A. M. Rukmi, M. I. Irawan, and N. Hidayat, "Pengklasteran Data Kategoris Dengan Algoritma Shared Nearest Neighbor," *Limits J. Math. Its Appl.*, vol. 6, no. 1, p. 63, 2009, doi: 10.12962/j1829605x.v6i1.1432.
- [19] L. Ertoz, M. Steinbach, and V. Kumar, "A new shared nearest neighbor clustering algorithm and its applications," ... *Data its Appl.*, no. May, pp. 1–15, 2002, [Online]. Available: http://www-users.cs.umn.edu/~kumar/papers/sjam_hd_snn_cluster.pdf.
- [20] M. Ramadhani and D. Fitriah, "Implementation of data mining analysis to determine the tuna fishing zone using DBSCAN algorithm," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 5, pp. 706–711, 2019, doi: 10.18178/ijmlc.2019.9.5.862.

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “*Analisa Perbandingan Algoritma Clustering DBSCAN Dengan SNN Untuk Pemetaan Daerah Penyebaran COVID-19 Di DKI Jakarta*”. Kertas kerja ini berisi semua material hasil penelitian Tugas Akhir. Di dalam kertas kerja ini disajikan beberapa bagian yang terdiri dari literature review, dataset yang digunakan, tahapan eksperimen, dan hasil eksperimen secara keseluruhan.

Bagian I membahas mengenai literatur review yang berisi artikel jurnal yang menjadi dasar atau landasan dalam penelitian ini. Bagian II menjelaskan mengenai source code yang digunakan pada penelitian ini. Bagian III menjelaskan mengenai dataset yang digunakan dalam penelitian ini, meliputi penjelasan, cara perolehan data, atribut data, dan penyesuaian data akhir yang siap untuk diolah. Bagian IV membahas mengenai tahapan eksperimen yang disajikan dalam bentuk gambar dengan penjelasan dari setiap tahapan tersebut. Bagian V merupakan bagian terakhir dari kertas kerja ini yang menjelaskan hasil keseluruhan dari eksperimen yang telah dilakukan, meliputi penjelasannya.



UNIVERSITAS
MERCU BUANA