



**ANALISA PERBANDINGAN ALGORITMA CLUSTERING UNTUK
PEMETAAN STATUS PEREKONOMIAN MASYARAKAT DI
KELURAHAN TOMANG**

TUGAS AKHIR

Noviana Hapsari
41517010034

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

2021



**ANALISA PERBANDINGAN ALGORITMA CLUSTERING UNTUK
PEMETAAN STATUS PEREKONOMIAN MASYARAKAT DI
KELURAHAN TOMANG**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Noviana Hapsari
41517010034

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2021

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41517010034

Nama : Noviana Hapsari

Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang.

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



Noviana Hapsari



UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

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

Nama Mahasiswa : Noviana Hapsari
NIM : 41517010034
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang.

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



Noviana Hapsari

UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN LUARAN TUGAS AKHIR

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

Nama Mahasiswa : Noviana Hapsari
NIM : 41517010034
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang

Menyatakan bahwa :

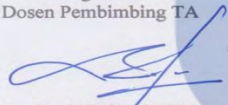
1. Luaran Tugas Akhir saya adalah sebagai berikut :

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 Teknik Informatika		
	ISSN :		
	Link Jurnal : http://journal.uinjkt.ac.id/index.php/ti/index		
	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



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

Jakarta, Januari 2021



Noviana Hapsari

UNIVERSITAS
MERCU BUANA



LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010034
Nama : Noviana Hapsari
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang

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 Januari 2021

(Dr. Ida Nurhaida, MT)

UNIVERSITAS
MERCU BUANA



LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010034
Nama : Noviana Hapsari
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk
Pemetaan Status Perekonomian Masyarakat Di
Kelurahan Tomang

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 Januari 2021

(Sabar Rudiarto, M.Kom)

UNIVERSITAS
MERCU BUANA



LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010034
Nama : Noviana Hapsari
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang

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 Januari 2021

(Hery Derajat Wijaya, S.Kom, MM)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : Noviana Hapsari
Nama : 41517010034
Judul Tugas Akhir : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang.

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

Menyetujui,



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

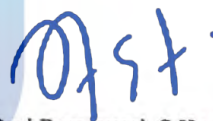
Dosen Pembimbing

Mengetahui,



(Diky Firdaus, S.Kom, MM)

Koord. Tugas Akhir Teknik Informatika



(Desi Ramayanti, S.Kom, MT)

Ka. Prodi Teknik Informatika

UNIVERSITAS
MERCU BUANA

ABSTRAK

Nama : Noviana Hapsari
NIM : 41517010034
Pembimbing TA : Dr. Devi Fitrianah, S.Kom., M.TI
Judul : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang

Status sosial ekonomi ialah gambaran tentang posisi dan kondisi seseorang. Tidak hanya di Indonesia tetapi diluar negeri, status sosial ekonomi juga sangat penting dalam kehidupan sosial. Latar belakang penelitian ini adalah adanya keperluan dari pengelola Kelurahan Tomang untuk dapat mengetahui profil status sosial ekonomi dari masyarakat yang berdomisili di area kelurahan tersebut. Dari sana, perlu dilakukan pemetaan status sosial ekonomi menggunakan pendekatan berbasis data *method driven* dengan teknik data mining khususnya clustering maka, akan didapatkan profil pemetaan status sosial ekonomi masyarakat. Metode clustering yang digunakan adalah algoritma Fuzzy C-Means, K-Means, dan K-Medoids. Dari ketiga algoritma clustering ini, akan dicari yang terbaik dan ditentukan mana yang lebih dapat menggambarkan profil status sosial ekonomi. Penelitian ini menggunakan perhitungan validasi *Dunn Index* untuk mengukur kekompakan cluster yang dihasilkan. Validasi terbaik dari hasil analisis cluster algoritma Fuzzy C-Means sebesar 0.91 dengan cluster sebanyak 6, algoritma K-Means mendapatkan validasi terbaik sebesar 0.95 dengan cluster sebanyak 5, dan algoritma K-Medoids mendapatkan validasi terbaik sebesar 0.98 dengan cluster sebanyak 6. Dari ketiga perbandingan algoritma tersebut, algoritma K-Medoids lebih dapat menggambarkan profil status sosial ekonomi pada data kependudukan yang dimiliki oleh Kelurahan Tomang karena mendapatkan hasil validasi tertinggi diantara kedua validasi yang lainnya dengan jumlah cluster sebanyak 6 cluster.

Kata kunci:
Fuzzy C-Means, K-Means, K-Medoids, Data Mining, Clustering

ABSTRACT

Name : Noviana Hapsari
Student Number : 41517010034
Counsellor : Dr. Devi Fitrianah, S.Kom, M.TI
Title : Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang

Social economic status is a description of a person's position and condition. Not only in Indonesia, but abroad social economic status is also very important in social life. The background of this research is that there is a need for the manager of Kelurahan Tomang to be able to find out the profile of the social economic status of the people who live in the village area. From there it is necessary to map the social economic status using a data-driven method driven approach with data mining techniques, especially clustering, so that a community social economic status mapping profile will be obtained. The clustering method used is the Fuzzy C-Means algorithm, K-Means, and K-Medoids. Of the three clustering algorithms, the best one will be sought and which one can better describe the profile of social economic status. This study uses the Dunn Index validation calculation to measure the cohesiveness of the resulting clusters. The best validation results from the cluster analysis of the Fuzzy C-Means algorithm are 0.91 with 6 clusters, K-Means algorithm gets the best validation of 0.95 with 5 clusters, and K-Medoids algorithm gets the best validation of 0.98 with 6 clusters. Algorithm K-Medoids algorithm can better describe the profile of social economic status on the population data owned by Kelurahan Tomang because it gets the highest validation results among the other two validations with a total of 6 cluster.

Key words:

Fuzzy C-Means, K-Means, K-Medoids, Data Mining, Clustering

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 Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang” 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.

Akhir kata, penulis berharap tugas akhir ini dapat bermanfaat bagi pembaca guna menambah pengetahuan dan wawasan serta pembaca dapat memberikan kritik maupun saran. 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.

Jakarta, Januari 2021

Penulis

xi

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	xi
NASKAH JURNAL	13
KERTAS KERJA	20
BAB 1. LITERATUR REVIEW	21
BAB 2. SOURCE CODE	26
BAB 3. DATASET	34
BAB 4. TAHAPAN EKSPERIMEN	36
BAB 5. HASIL SEMUA EKSPERIMEN	40
LAMPIRAN DOKUMEN HAKI	48
LAMPIRAN KORESPONDENSI	50

NASKAH JURNAL

Jurnal Teknik Informatika Vol. 13 No. 1, April 2020

| 1

ANALISA PERBANDINGAN ALGORITMA CLUSTERING UNTUK PEMETAAN STATUS PEREKONOMIAN MASYARAKAT DI KELURAHAN TOMANG

Noviana Hapsari¹, Devi Fitriyah²

^{1,2} Informatika, Ilmu Komputer
Universitas Mercu Buana
Jakarta Barat, Indonesia

novianahapsari17@gmail.com, devi.fitriyah@mercubuana.ac.id

ABSTRACT

Social economic status is a description of a person's position and condition. Not only in Indonesia, but abroad social economic status is also very important in social life. The background of this research is that there is a need for the manager of Kelurahan Tomang to be able to find out the profile of the social economic status of the people who live in the village area. From there it is necessary to map the social economic status using a data-driven method driven approach with data mining techniques, especially clustering, so that a community social economic status mapping profile will be obtained. The clustering method used is the Fuzzy C-Means algorithm, K-Means, and K-Medoids. Of the three clustering algorithms, the best one will be sought and which one can better describe the profile of social economic status. This study uses the Dunn Index validation calculation to measure the cohesiveness of the resulting clusters. The best validation results from the cluster analysis of the Fuzzy C-Means algorithm are 0.91 with 6 clusters, K-Means algorithm gets the best validation of 0.95 with 5 clusters, and K-Medoids algorithm gets the best validation of 0.98 with 6 clusters. Algorithm K-Medoids algorithm can better describe the profile of social economic status on the population data owned by Kelurahan Tomang because it gets the highest validation results among the other two validations with a total of 6 cluster.

Keywords: *Fuzzy C-Means, K-Means, K-Medoids, Data Mining, Clustering*

ABSTRAK

Status sosial ekonomi ialah gambaran tentang posisi dan kondisi seseorang. Tidak hanya di Indonesia tetapi diluar negeri, status sosial ekonomi juga sangat penting dalam kehidupan sosial. Latar belakang penelitian ini adalah adanya keperluan dari pengelola Kelurahan Tomang untuk dapat mengetahui profil status sosial ekonomi dari masyarakat yang berdomisili di area kelurahan tersebut. Dari sana, perlu dilakukan pemetaan status sosial ekonomi menggunakan pendekatan berbasis data *method driven* dengan teknik data mining khususnya clustering maka, akan didapatkan profil pemetaan status sosial ekonomi masyarakat. Metode clustering yang digunakan adalah algoritma Fuzzy C-Means, K-Means, dan K-Medoids. Dari ketiga algoritma clustering ini, akan dicari yang terbaik dan ditentukan mana yang lebih dapat menggambarkan profil status sosial ekonomi. Penelitian ini menggunakan perhitungan validasi *Dunn Index* untuk mengukur kekompakan cluster yang dihasilkan. Validasi terbaik dari hasil analisis cluster algoritma Fuzzy C-Means sebesar 0.91 dengan cluster sebanyak 6, algoritma K-Means mendapatkan validasi terbaik sebesar 0.95 dengan cluster sebanyak 5, dan algoritma K-Medoids mendapatkan validasi terbaik sebesar 0.98 dengan cluster sebanyak 6. Dari ketiga perbandingan algoritma tersebut, algoritma K-Medoids lebih dapat menggambarkan profil status sosial ekonomi pada data kependudukan yang dimiliki oleh Kelurahan Tomang karena mendapatkan hasil validasi tertinggi diantara kedua validasi yang lainnya dengan jumlah cluster sebanyak 6 cluster.

Kata Kunci: *Fuzzy C-Means, K-Means, K-Medoids, Data Mining, Clustering*

Nama Penulis, dkk: Judul Artikel

.....

p-ISSN 1979-9160 | e-ISSN 2549-7901

DOI: [http://dx.doi.org/10.15408/jti.v10i2.\(id article\)](http://dx.doi.org/10.15408/jti.v10i2.(id article))

I. PENDAHULUAN

Status sosial ekonomi ialah gambaran tentang posisi dan kondisi seseorang. Tidak hanya di Indonesia tetapi diluar negeri, status sosial ekonomi juga sangat penting dalam kehidupan sosial.

Status sosial ekonomi merupakan suatu situasi yang menunjukkan bahwa suatu keluarga mempunyai kemampuan finansial dan perlengkapan material. Situasi tersebut berstandar baik, cukup atau kurang[1].

Pekerjaan, pendidikan, kesehatan hingga pemenuhan kebutuhan hidup dalam rumah tangga merupakan aspek yang dapat melihat status sosial ekonomi suatu keluarga. Atas dasar inilah, kedudukan status sosial ekonomi dapat dibedakan menjadi status sosial ekonomi atas, menengah dan bawah[1].

Dari sana, perlu dilakukan pemetaan status sosial dengan menggunakan pendekatan berbasis data *method driven* dan menggunakan teknik data mining khususnya clustering, maka akan diperoleh profil pemetaan status sosial ekonomi masyarakat.

Penelitian ini membandingkan algoritma clustering Fuzzy C-Means, K-Means, dan K-Medoids untuk memetakan status perekonomian masyarakat di Kelurahan Tomang dengan menggunakan *Python* sebagai *tools*, dan *Dunn Index* sebagai validasi clustering untuk mengukur kekompakan cluster yang dihasilkan.

Tomang adalah kelurahan yang terletak di Kecamatan Grogol Petamburan, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, Indonesia. Kelurahan Tomang memiliki jumlah penduduk sebanyak 36.525 jiwa, terdiri dari 18.045 laki-laki dan 18.480 perempuan serta terdapat 16 Rukun Warga(RW) dan 174 Rukun Tetangga(RT).

Permasalahan yang sering dihadapi oleh pihak Kelurahan Tomang adalah kurangnya pemetaan dan pengidentifikasian kelompok masyarakat berdasarkan status perekonomiannya. Oleh karena itu, penelitian ini diharapkan dapat membantu pihak Kelurahan Tomang dalam memetakan status perekonomian masyarakat untuk dijadikan sebagai data masukan dalam pengambilan keputusan pemberian bantuan kepada masyarakat secara tepat.

Penelitian ini menggunakan data kuantitatif (data sekunder) yaitu data

kependudukan yang merupakan data statistik di Kelurahan Tomang pada tahun 2020 sebanyak 1161 data. Parameter yang digunakan pada penelitian ini, yaitu Jumlah Anggota Keluarga, Kepemilikan Rumah, Fasilitas BAB, Daya Listrik, Jenis Lantai, Jenis Dinding, Jenis Atap, Sumber Air Minum, Jumlah Pengeluaran Perbulan dalam bentuk Percentile, dan Aset Tidak Bergerak.

1.1 Data Mining

Data Mining menggunakan teknik statistik, kecerdasan buatan dan pembelajaran mesin dari berbagai database besar selama pemrosesan, untuk mengekstrak dan mengidentifikasi informasi[2].

Data Mining ialah proses penambahan repositori data besar dan menggunakan teknik pengenalan pola (seperti teknik statis dan matematika) untuk menemukan korelasi, pola, dan tren baru[3].

1.2 Analisis Cluster

Analisis cluster adalah teknik multivariat yang dirancang untuk mengelompokkan objek dengan karakteristik serupa[4].

Analisis cluster adalah analisis statistik yang bertujuan untuk mengklasifikasikan objek berdasarkan kesamaan karakteristik antar objek tersebut[5].

1.3 Pre-processing

Pre-processing merupakan proses perubahan data menjadi lebih sederhana, efektif dan sesuai dengan kebutuhan[6].

Pre-processing ialah proses perubahan data yang tidak terstruktur menjadi data terstruktur agar dapat digunakan sesuai kebutuhan[7].

1.4 Clustering

Clustering merupakan proses pengelompokan data. Data yang memiliki kesamaan dikumpulkan ke dalam kelompok atau cluster[8].

Clustering merupakan pengelompokan data dari karakteristik yang serupa. Pengetahuan yang berguna dapat dihasilkan dari kesamaan karakteristik masing-masing kelompok[9].

1.5 Algoritma Fuzzy C-Means

Fuzzy C-Means ialah algoritma yang mengandalkan nilai derajat keanggotaan untuk

menentukan suatu cluster dari setiap data[10]. Fuzzy C-Means adalah metode pengelompokan data yang paling umum digunakan karena keberadaan setiap titik data dalam kelompok bergantung pada derajat keanggotaannya[11].

1.6 Algoritma K-Means

K-Means menggunakan proses berulang untuk menghasilkan cluster data. Metode K-Means melakukan pemilihan nilai (k) sebagai titik awal centroid[12].

Algoritma K-Means merupakan algoritma clustering yang menggunakan pendekatan pengelompokan partisi. Metode K-Means mengelompokkan data menjadi cluster dengan karakteristik yang sama[13].

1.7 Algoritma K-Medoids

K-Medoids merupakan metode menemukan medoids dalam suatu kelompok. Medoids tersebut merupakan titik pusat dari suatu kelompok[14].

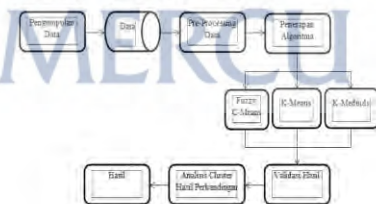
Metode K-Medoids merupakan metode meminimalisasi jarak antara medoids dengan urutan lainnya dalam kelompok[15].

1.8 Indeks Dunn

Dunn Index merupakan metode untuk mengukur validasi kekompakan cluster. Metode perhitungan diperoleh dengan menghitung nilai minimum dari perbandingan nilai perbedaan dua cluster sebagai intervalnya dan nilai maksimum dari diameter cluster sebagai keketatannya. Jumlah cluster optimal dari Dunn Index dinyatakan dengan semakin besar nilai yang didapatkan[16].

II. METODOLOGI

Dalam penelitian ini, penulis melakukan beberapa tahapan. Tahapan penelitian dapat dilihat pada Gambar 1 berikut ini:



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Pada tahap ini, data diperoleh langsung dari Kelurahan Tomang. Data yang digunakan sebanyak 1161 data. Algoritma Fuzzy C-Means, K-Means, dan K-Medoids akan digunakan untuk menyeleksi dan mengolah data yang telah terverifikasi oleh Kelurahan Tomang, Jakarta Barat. Data tersebut dapat

No	Persentase	Jumlah Aggregasi Rumah Tangga	Status Pemukiman	Jenis Fasilitas	Jenis Lantai	Jenis Dinding	Jenis Atap	Sumber Air Minum	Aspek Tidak Sejahtera
1	1	12	1	1	1	1	1	1	1
2	1	12	1	1	1	1	1	1	1
3	1	12	1	1	1	1	1	1	1
4	1	14	1	1	1	1	1	1	1
5	1	12	1	1	1	1	1	1	1
...
Total	45	135	1	1	1	1	1	1	1

dilihat pada Gambar 2 sebagai berikut:
Gambar 2. Data

2.2 Pre-Processing Data

Pada tahap ini, data yang diperoleh akan di *pre-processing* menggunakan *Min-Max Scaler* dengan transformasi linier atau memberi nilai skala dari 0 hingga 1 terhadap data asli yang akan digunakan[17].

Hasil *pre-processing* data dapat dilihat

No	Persentase	Jumlah Aggregasi Rumah Tangga	Status Pemukiman	Jenis Fasilitas	Jenis Lantai	Jenis Dinding	Jenis Atap	Sumber Air Minum	Aspek Tidak Sejahtera
1	0,00204	0,43975	0,34690	0,22173	0,33948	0,91607	0,25712	0,40544	0,91011
2	0,00408	0,87942	0,69379	0,44347	0,67896	0,83214	0,51424	0,81088	0,82022
3	0,00816	1,75884	1,38758	0,88694	1,35792	1,66428	1,02848	1,62176	1,64044
4	0,01632	3,51768	2,77516	1,77388	2,71584	3,32856	2,05696	3,24352	3,28088
5	0,03264	7,03536	5,55032	3,54776	5,43168	6,65712	4,11392	6,48704	6,56176
...
Total	0,15168	45,55032	35,55032	22,6176	33,948	41,664	25,712	40,544	40,511

pada Gambar 3.
Gambar 3. Pre-Processing Data

2.3 Penerapan Algoritma

2.3.1 Algoritma Fuzzy C-Means

Langkah-langkah dari algoritma Fuzzy C-Means adalah[18]:

1. Identifikasikan data yang akan di cluster dalam bentuk matriks dengan ukuran $n \times m$ (X_{ij} adalah data sampel ke- i ($i = 1,2,...,n$) dan atribut ke- j ($j = 1,2,...,m$))
2. Tentukan jumlah cluster (c), pangkat (w), iterasi maksimum ($maxiter$), error minimum yang diharapkan (ϵ), fungsi objektif awal ($p_0 = 0$) dan iterasi awal ($t = 1$)

3. Hasilkanlah bilangan acak μ_{ik} ($i = 1,2,\dots,n$) dan ($k= 1,2,\dots,c$) sebagai elemen dari matriks partisi awal u

4. Hitunglah jumlah tiap kolom dengan menggunakan rumus:

$$q_j = \sum_{k=1}^c \mu_{ik} \dots \dots \dots (1)$$

5. Hitunglah nilai elemen dari matriks partisi anggota himpunan u dengan menggunakan rumus:

$$\mu_{ik} = \frac{\mu_{ik}}{q_j} \dots \dots \dots (2)$$

6. Hitunglah nilai pusat cluster ke- k (v_{kj}) dimana $k=1,2,\dots,c$ dan $j=1,2,\dots,m$ dengan menggunakan rumus:

$$v_{kj} = \left(\mu_{ik} \right)^w \sum_{j=1}^m \mu_{ik}^m \dots \dots \dots (3)$$

7. Hitunglah fungsi objektif pada iterasi ke- t (pt):

$$pt = \sum_{i=1}^n \sum_{k=1}^c \left[\sum_{j=1}^m \left(x_{ij} - v_{kj} \right)^2 \right] \left(\mu_{ik} \right)^w \dots \dots \dots (4)$$

8. Hitunglah perubahan matriks partisi u :

$$\mu_{ik} = \hat{u}_{ik}$$

9. Periksa kondisi berhenti:

a. Jika $|pt - pt-1| < \epsilon$ atau $|t - t_{max}| < \epsilon$, maka iterasi berhenti

b. Jika $t = t + 1$, ulangi langkah ke-6 sampai langkah ke-9

2.3.2 Algoritma K-Means

Langkah-langkah dari algoritma K-Means adalah [19]:

1. Menentukan nilai (k) sebagai centroid awal
2. Membuat cluster dengan menempatkan tiap titik pada centroid terdekat
3. Hitunglah kembali centroid tiap cluster hingga centroid tidak berhenti

2.3.2 Algoritma K-Medoids

Langkah-langkah dari algoritma K-Medoids adalah [20]:

1. Inisialisasikan pusat cluster sebanyak jumlah cluster
2. Gunakan persamaan jarak *Euclidian* untuk menempatkan tiap data ke cluster terdekat
3. Pilihlah objek secara acak di setiap cluster sebagai calon medoids
4. Hitunglah jarak terdekat tiap objek di masing-masing cluster
5. Hitunglah total simpangan (S) dengan mengurangi jumlah jarak baru dengan jumlah jarak lama. Jika $S < 0$, gantilah objek dengan data cluster untuk membuat sekumpulan (k) objek baru sebagai medoids

6. Ulangi langkah ke-3 sampai langkah ke-5 hingga tidak ada perubahan pada medoids

2.4 Analisis Hasil Perbandingan

Analisis hasil perbandingan pada penelitian ini yaitu dengan membandingkan hasil validasi dari masing-masing algoritma dengan menggunakan nilai $k=2$ hingga $k=6$, kemudian membandingkan hasil validasi sehingga memperoleh hasil validasi tertinggi dari masing-masing algoritma tersebut. Setelah itu, membandingkan grafik dan hasil cluster yang dihasilkan.

III. HASIL DAN PEMBAHASAN

Pada penerapan algoritma Fuzzy C-Means, K-Means dan K-Medoids dari $k=2$ hingga $k=6$ diperoleh hasil validasi dari masing-masing algoritma tersebut. Hasil validasi dapat dilihat pada Gambar 4.

Gambar 4. Hasil Validasi

Algoritma	K	Validasi Dunn	Algoritma	K	Validasi Dunn	Algoritma	K	Validasi Dunn
Fuzzy C-Means	2	0.84	K-Means	2	0.84	K-Medoids	2	0.83
	3	0.84		3	0.88		3	0.83
	4	0.90		4	0.91		4	0.91
	5	0.84		5	0.95		5	0.98
	6	0.91		6	0.92		6	0.98

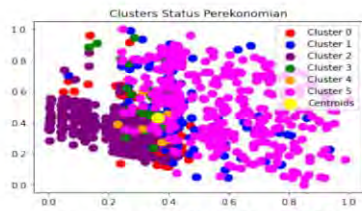
Pada Gambar 4, terlihat bahwa dengan menggunakan perhitungan validasi *Dunn Index* pada ketiga algoritma tersebut telah berada dikelompok yang tepat. Algoritma Fuzzy C-Means dan K-medoids terdapat pada $k=6$ sedangkan algoritma K-Means terdapat pada $k=5$, validasi terbaik dapat dilihat pada *Tabel 1*.

Tabel 1. Hasil Validasi Terbaik

Algoritma	K	Hasil Validasi Terbaik
Fuzzy C-Means	6	0.91
K-Means	5	0.95
K-Medoids	6	0.98

Dapat dilihat pada *Tabel 1* bahwa algoritma K-Medoids mendapatkan hasil validasi tertinggi diantara kedua validasi yang lainnya yaitu sebesar 0.98 dengan jumlah cluster sebanyak 6.

Hasil grafik yang diperoleh menggunakan $k=6$ pada algoritma Fuzzy C-Means dapat dilihat pada Gambar 5.



Gambar 5. Hasil Grafik Fuzzy C-Means

Pada Gambar 5, terlihat bahwa Cluster 0 terdiri dari 7 RW dan 17 RT, Cluster 1 terdiri dari 9 RW dan 17 RT, Cluster 2 terdiri dari 6 RW dan 21 RT, Cluster 3 terdiri dari 1 RW dan 1 RT, Cluster 4 tidak terdapat jumlah RW dan RT, dan Cluster 5 terdiri dari 13 RW dan 54 RT

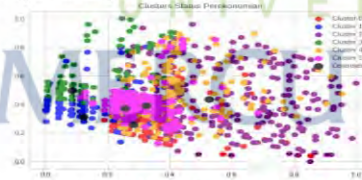
Hasil grafik yang diperoleh menggunakan k=5 pada Algoritma K-Means dapat dilihat pada Gambar 6.



Gambar 6. Hasil Grafik K-Means

Pada Gambar 6, terlihat bahwa Cluster 0 terdiri dari 8 RW dan 21 RT, Cluster 1 terdiri dari 4 RW dan 5 RT, Cluster 2 terdiri dari 6 RW dan 30 RT, Cluster 3 terdiri dari 8 RW dan 22 RT, dan Cluster 4 terdiri dari 8 RW dan 33 RT

Sedangkan hasil grafik yang diperoleh menggunakan k=6 pada Algoritma K-Medoids dapat dilihat pada Gambar 7.



Gambar 7. Hasil Grafik K-Medoids

Pada Gambar 7 dapat dilihat bahwa Cluster 0 terdiri dari 7 RW dan 10 RT, Cluster

1 terdiri dari 6 RW dan 9 RT, Cluster 2 terdiri dari 11 RW dan 50 RT, Cluster 3 terdiri dari 6 RW dan 8 RT, Cluster 4 terdiri dari 11 RW dan 24 RT, dan Cluster 5 terdiri dari 3 RW dan 8 RT

Dari perbandingan hasil validasi terbaik yang diperoleh, perbandingan hasil validasi tertinggi serta perbandingan ketiga hasil grafik dari algoritma Fuzzy C-Means, K-Means, dan K-Medoids diperoleh bahwa algoritma K-Medoids lebih dapat menggambarkan profil status sosial ekonomi pada data kependudukan yang dimiliki oleh Kelurahan Tomang karena mendapatkan hasil validasi tertinggi diantara kedua validasi yang lainnya yaitu sebesar 0.98 dengan jumlah cluster yang dihasilkan sebanyak 6 cluster.

Hasil analisa cluster yang diperoleh menggunakan algoritma K-Medoids pada penelitian ini ditunjukkan pada Tabel 2 berikut ini:

Tabel 2. Hasil Analisa Cluster

Cluster	RW	RT	Cluster	RW	RT
0	1	2	15	2,3	
	6	5	16	1,4,5,6,7,8,9	
	11	5	3	4	13
	13	4	6	14	
	14	1,2,3,4	11	10	
	15	1	12	7,8,9	
	16	2	13	8	
1	1	4	14	8	
	2	3,4	4	1	8
	3	7	3	1,3	
	5	3	4	5	
	11	1	5	1,4,9	
	12	1,2,3,4	6	1,4,6,7,13	
2	4	4,6,7,11,12,14	7	3,4	
	5	7,13	8	2	
	6	2,3,8,9,10,11	9	1,2,10	
	7	1,2,5,6,7,8,9,11	10	4,6,9	
		0,11,12,13,14			
	8	1,4,5,6	12	5,6	
	9	5	13	3	
	10	7	5	13	5
	11	2,3,4,6,7,9	14	5,6,7,9	
	13	2,6,7	15	4,5,10	

Analisa cluster yang dihasilkan yaitu Cluster 5 merupakan kumpulan wilayah yang memiliki nilai validasi terbaik dari ketiga algoritma, sehingga dapat dikelompokkan sebagai status perekonomian masyarakatnya sangat mampu. Cluster 4 merupakan kelompok wilayah status perekonomian masyarakatnya mampu. Cluster 3 merupakan kelompok wilayah status perekonomian masyarakatnya tidak mampu. Cluster 2 merupakan kelompok wilayah status perekonomian masyarakatnya cukup miskin. Cluster 1 merupakan kelompok wilayah status perekonomian masyarakatnya miskin, dan Cluster 0 merupakan kelompok wilayah status perekonomian masyarakatnya sangat miskin.

IV. PENUTUP

Berdasarkan beberapa parameter yaitu Jumlah Anggota Keluarga, Kepemilikan Rumah, Fasilitas BAB, Daya Listrik, Jenis Lantai, Jenis Dinding, Jenis Atap, Sumber Air Minum, Jumlah Pengeluaran Perbulan dalam Bentuk Percentile, dan Aset Tidak Bergerak diperoleh hasil perhitungan algoritma Fuzzy C-Means dengan validasi terbaik sebesar 0.91, algoritma K-Means mendapatkan validasi terbaik sebesar 0.95, dan algoritma K-Medoids mendapatkan validasi terbaik sebesar 0.98. Dalam penelitian ini, menunjukkan bahwa validasi terbaik menggunakan *Dunn Index* dari masing-masing algoritma telah berada dikelompok yang tepat. Dari ketiga perbandingan algoritma tersebut, algoritma K-Medoids lebih dapat menggambarkan profil status sosial ekonomi pada data kependudukan yang dimiliki oleh Kelurahan Tomang karena mendapatkan hasil validasi tertinggi diantara kedua validasi yang lainnya sebesar 0.98 dengan jumlah cluster yang dihasilkan sebanyak 6 cluster.

DAFTAR PUSTAKA

- [1] G. E. Wenas and H. Opod, "Hubungan Kebahagiaan Dan Status Sosial Ekonomi," *e-Biomedik (eBm)*, vol. 3, no. April, pp. 1–7, 2015.
- [2] R. Bena Siti Ashari, Steven Christ Otniel, "Perbandingan Kinerja K-Means Dengan DbSCAN Untuk Metode Clustering Data Penjualan Online Retail," vol. 5, no. 2, pp. 72–77, 2019.
- [3] F. Rahman, I. I. Ridho, M. Muffih, S. Pratama, M. R. Raharjo, and A. P. Windarto, "Application of Data Mining Technique using K-Medoids in the case of Export of Crude Petroleum Materials to the Destination Country," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 835, no. 1, 2020, doi: 10.1088/1757-899X/835/1/012058.
- [4] M. Gading Sadewo, A. Perdana Windarto, and A. Wanto, "KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer) PENERAPAN ALGORITMA CLUSTERING DALAM MENGELOMPOKKAN BANYAKNYA DESA/KELURAHAN MENURUT UPAYA ANTISIPASI/ MITIGASI BENCANA ALAM MENURUT PROVINSI DENGAN K-MEANS," vol. 2, pp. 311–319, 2018.
- [5] N. I. Febianto and N. Palasara, "Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 8, no. 2, p. 130, 2019, doi: 10.32736/sisfokom.v8i2.653.
- [6] S. Saifulhalla, M. Zarlis, Z. Zakaria, and R. W. Sembiring, "Analisa Terhadap Perbandingan Algoritma Decision Tree Dengan Algoritma Random Tree Untuk Pre-Processing Data," *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 1, no. 2, p. 180, 2017, doi: 10.30645/j-sakti.v1i2.41.
- [7] D. Steveson, H. Agung, and F. Mulia, "Aplikasi Pendeteksi Plagiarisme Tugas Dan Makalah Pada Sekolah Menggunakan Algoritma Rabin Karp," *J. Algorit. Log. dan Komputasi*, vol. 1, no. 1, pp. 12–17, 2018, doi: 10.30813/j-alu.v1i1.1104.
- [8] R. Venkat and K. S. Reddy, "Dealing big data using fuzzy c-means (FCM) clustering and optimizing with gravitational search algorithm (GSA)," *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, no. Icoei, pp. 465–467, 2019, doi: 10.1109/ICOEI.2019.8862673.
- [9] E. Elisawati, D. Wahyuni, and A. Arianto, "Analisa Clustering Pada Data Pelanggaran Lalulintas Di Pengadilan Negeri Dumai Dengan Menggunakan Metode K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 4, no. 2, pp. 76–87, 2019, doi: http://dx.doi.org/10.14421/jiska.2019.42-01.
- [10] N. Ulinnuha, "Provincial Clustering in Indonesia Based on Plantation Production Using Fuzzy C-Means," vol. 9, no. 1, pp. 8–12, 2020.
- [11] A. P. Pertiwi and R. Kurniawan, "Pengelompokan Daerah Rawan Bencana Banjir Di Indonesia Tahun 2013 Menggunakan Fuzzy C-Mean," pp. 677–687, 2017, doi: 10.31227/osf.io/e3r8s.
- [12] S. Suprihatin, Y. R. W. Utami, and D. Nugroho, "K-Means Clustering Untuk Pemetaan Daerah Rawan Demam

- Berdarah," *J. Teknol. Inf. dan Komun.*, vol. 7, no. 1, 2019, doi: 10.30646/tikomsin.v7i1.408.
- [13] C. D. Megawati, E. M. Yuniarno, and S. M. S. Nugroho, "Clustering of Female Avatar Face Features Consumers Choice using KMeans and SOM Algorithm," *Proc. - 2019 Int. Semin. Intell. Technol. Its Appl. ISITIA 2019*, pp. 366–370, 2019, doi: 10.1109/ISITIA.2019.8937279.
- [14] S. Sindi, W. R. O. Ningse, I. A. Sihombing, F. Ilmi R.H.Zer, and D. Hartama, "Analisis Algoritma K-Medoids Clustering Dalam Pengelompokan Penyebaran Covid-19 Di Indonesia," *Jti*, vol. 4, no. 1, pp. 166–173, 2020.
- [15] M. G. Johnson *et al.*, "A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering," *Syst. Biol.*, vol. 68, no. 4, pp. 594–606, 2019, doi: 10.1093/sysbio/syy086.
- [16] 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.
- [17] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, p. 78, 2019, doi: 10.24114/cess.v4i1.11458.
- [18] N. Butarbutar, A. P. Wndarto, D. Hertama, and Solikhun, "KOMPARASI KINERJA ALGORITMA FUZZY C-MEANS DAN K-MEANS DALAM PENGELOMPOKAN DATA SISWA BERDASARKAN PRESTASI NILAI AKADEMIK SISWA," *J. Ris. Sist. Inf. Tek. Inform.*, vol. 1, no. 2012, 2016.
- [19] E. Riksakomara, "SEGMENTASI SUPPLIER MENGGUNAKAN METODE K- MEANS CLUSTERING (STUDI KASUS : PTPN X PG MERITJAN) SUPPLIER SEGMENTATION K-MEANS CLUSTERING (CASE STUDY :
- [20] PTPN X PG MERITJAN) MEANS CLUSTERING," 2017.
I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “*Analisa Perbandingan Algoritma Clustering Untuk Pemetaan Status Perekonomian Masyarakat Di Kelurahan Tomang*”. 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 menjelaskan 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 meliputi penjelasan, cara perolehan data, dan jumlah parameter yang digunakan. Bagian IV menjelaskan mengenai tahapan eksperimen yang disajikan dalam gambar dengan penjelasan dari setiap tahapan. Bagian V merupakan bagian terakhir dari kertas kerja ini yaitu menjelaskan hasil keseluruhan dari eksperimen yang telah dilakukan, meliputi penjelasannya.

UNIVERSITAS
MERCU BUANA