



**STUDI KOMPARASI ALGORITMA CLUSTERING DALAM
MEMETAKAN STATUS PEREKONOMIAN MASYARAKAT DI
KELURAHAN CIPADU KOTA TANGERANG**

TUGAS AKHIR

Nensia Alendia
41517010116

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

2021



**STUDI KOMPARASI ALGORITMA CLUSTERING DALAM
MEMETAKAN STATUS PEREKONOMIAN MASYARAKAT DI
KELURAHAN CIPADU KOTA TANGERANG**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:
Nensia Alendia
41517010116

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2021

LEMBAR PERNYATAAN ORISINALITAS

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41517010116

Nama : Nensia Alendia

Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam Memetakan
Status Perekonomian Masyarakat Di Kelurahan Cipadu
Kota Tangerang

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, 8 Februari 2021



(Nensia Alendia)

UNIVERSITAS
MERCU BUANA

SURAT PERSETUJUAN PUBLIKASI TUGAS AKHIR

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

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

Nama Mahasiswa : Nensia Alendia.....
NIM : 41517010116.....
Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang.....
.....

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, 16 Februari 2021



SURAT PERNYATAAN LUARAN TUGAS AKHIR

SURAT PERNYATAAN LUARAN TUGAS AKHIR

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

Nama Mahasiswa : Nensia Alendia
NIM : 41517010116
Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	✓
		Jurnal Nasional Terakreditasi	
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
	Disubmit/dipublikasikan di :	Nama Jurnal : JURNAL SISTEM INFORMASI	
		ISSN : 2540 - 9719	
		Link Jurnal : http://sistemasi.ftik.unisi.ac.id/index.php/stmsi/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

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

Jakarta, 8 Februari 2021



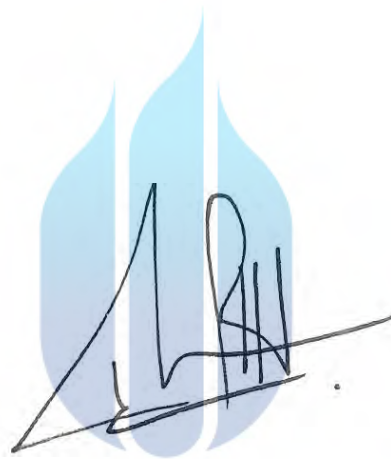
Nensia Alendia

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010116
Nama : Nensia Alendia
Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam
Memetakan Status Perekonomian Masyarakat Di
Kelurahan Cipadu Kota Tangerang

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, 8 Februari 2021



UNIVERSITAS (Dr. Mujiono Sadikin, MT)

MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010116
Nama : Nensia Alendia
Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

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, 8 Februari 2021



(Sabar Rudiarto, S.Kom., M.Kom)

U N I V E R S I T A S
M E R C U B U A N A

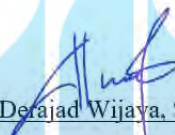
LEMBAR PERSETUJUAN PENGUJI

LEMBAR PERSETUJUAN PENGUJI

NIM : 41517010116
Nama : Nensia Alendia
Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam
Memetakan Status Perekonomian Masyarakat Di
Kelurahan Cipadu Kota Tangerang

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 Detajad Wijaya, S.Kom, MM)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

NIM : 41517010116
Nama : Nensia Alendia
Judul Tugas Akhir : Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

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, 8 Februari 2021

Menyetujui,


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

Mengetahui,

 
(Diky Firdaus, S.Kom., MM) (Desi Ramayanti, S.Kom., MT)
Koord. Tugas Akhir Teknik Informatika Ka. Prodi Teknik Informatika

ABSTRAK

Nama : Nensia Alendia
NIM : 41517010116
Pembimbing TA : Dr. Devi Fitrianah, S.Kom., M.TI
Judul : Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

Penelitian ini membahas mengenai perbandingan Algoritma Meanshift Clustering, K-Means dan K-Medoids dalam klasterisasi dan hasilnya akan diusulkan untuk menjadi bahan pertimbangan pemerintah setempat dalam mengelompokkan masyarakat berdasarkan status perekonomiannya. Pemetaan sosial ekonomi akan dilakukan dengan menggunakan pendekatan data mining khususnya menggunakan algoritma Clustering Mean Shift, K-Means dan K-Medoids ketiga algoritma ini akan dibandingkan untuk mendapatkan hasil yang terbaik. Hasil clustering nantinya akan dijadikan sebagai profil status yang ada dimasyarakat kelurahan Cipadu Kota Tangerang. Didapatkan data sejumlah 862 data dengan atribut Jumlah anggota rumah tangga, status pekerjaan, status bangunan, luas lantai, daya terpasang, jenis lantai, kondisi atap dan sumber minum. Penelitian ini menggunakan validasi Dunn Index. Berdasarkan hasil pengujian kualitas cluster dengan validasi Dunn Index di dapatkan nilai dari hasil analisis Algoritma Mean shift sebesar 0.887, kemudian di dapatkan hasil K-Means sebesar 0.995 dan K-Medoids sebesar 0.880 Maka dari ketiga algoritma tersebut dapat ditarik kesimpulan bahwa algoritma K-Means Clustering memiliki nilai yang paling baik karena paling mendekati angka 1.

Kata kunci: *Status Ekonomi, Meanshift, K-Means, K-medoids, clustering, Data Mining*

UNIVERSITAS
MERCU BUANA

ABSTRACT

Name : Nensia Alendia
Student Number : 41517010116
Counsellor : Dr. Devi Fitrianah, S.Kom., M.TI
Title : Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

This study discusses the comparison of the Meanshift Clustering Algorithm, K-Means and K-Medoids in clustering and the results will be proposed to be taken into consideration by the local government in grouping people based on their economic status. Socio-economic mapping will be carried out using a data mining approach, especially using the Clustering Mean Shift, K-Means and K-Medoids algorithms. These three algorithms will be compared to get the best results. The results of the clustering will later be used as a profile of existing status in the Cipadu village community, Tangerang City. The data obtained were 862 with the attributes of the number of household members, work status, building status, floor area, installed power, type of floor, roof condition and drinking source. This study uses Dunn Index validation. Based on the results of cluster quality testing with Dunn Index validation, the value of the Mean Shift Algorithm analysis is 0.887, then the K-Means results are 0.995 and K-Medoids is 0.880. So from the three algorithms it can be concluded that the K-Means Clustering algorithm has the best value because it is closest to number 1.

Key words:

Economic Status, Mean shift, K-Means, K-medoids, Clustering, Data Mining

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 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.....	vii
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT	x
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xii
NASKAH JURNAL	1
KERTAS KERJA.....	15
BAB 1. LITERATUR REVIEW	16
BAB 2. SOURCE CODE	25
BAB 3. DATASET.....	34
BAB 4. TAHAPAN EKSPERIMEN	36
BAB 5. HASIL SEMUA EKSPERIMEN.....	41
BAB 6. DAFTAR PUSTAKA.....	47
LAMPIRAN DOKUMEN HAKI.....	49
LAMPIRAN KORESPONDENSI	51

NASKAH JURNAL

SISTEMASI : Jurnal Sistem Informasi

STUDI KOMPARASI ALGORITMA *CLUSTERING* DALAM MEMETAKAN STATUS PEREKONOMIAN MASYARAKAT DI KELURAHAN CIPADU KOTA TANGERANG

¹Nensia Alendia¹, ²Devi Fitrihanah²,

^{1,2}Prodi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana,

Jl. Meruya Selatan No.1, Kembangan, Daerah Khusus Ibukota Jakarta

Email: 41517010116@student.mercubuana.ac.id, devifitrihanah16@mercubuana.ac.id

ABSTRACT

This study discusses the comparison of the MeanShift Clustering Algorithm, K-Means and K-Medoids in clustering and the results will be proposed to be taken into consideration by the local government in grouping people based on their economic status. Socio-economic mapping will be carried out using a data mining approach, especially using the Clustering Mean Shift, K-Means and K-Medoids algorithms. These three algorithms will be compared to get the best results. The results of the clustering will later be used as a profile of existing status in the Cipadu village community. Tangerang City. The data obtained were 862 with the attributes of the number of household members, work status, building status, floor area, installed power, type of floor, roof condition and drinking source. This study uses Dunn Index validation. Based on the results of cluster quality testing with Dunn Index validation, the value of the Mean Shift Algorithm analysis is 0.887, then the K-Means results are 0.995 and K-Medoids is 0.880. So from the three algorithms it can be concluded that the K-Means Clustering algorithm has the best value because it is closest to number 1.

Keywords: Economic Status, Mean shift, K-Means, K-medoids, Clustering, Data Mining

ABSTRAK

Penelitian ini membahas mengenai perbandingan Algoritma MeanShift Clustering, K-Means dan K-Medoids dalam klasterisasi dan hasilnya akan diusulkan untuk menjadi bahan pertimbangan pemerintah setempat dalam mengelompokkan masyarakat berdasarkan status perekonomiannya. Pemetaan sosial ekonomi akan dilakukan dengan menggunakan pendekatan data mining khususnya menggunakan algoritma Clustering Mean Shift, K-Means dan K-Medoids ketiga algoritma ini akan dibandingkan untuk mendapatkan hasil yang terbaik. Hasil clustering nantinya akan dijadikan sebagai profil status yang ada dimasyarakat kelurahan Cipadu Kota Tangerang. Didapatkan data sejumlah 862 data dengan atribut Jumlah anggota rumah tangga, status pekerjaan, status bangunan, luas lantai, daya terpasang, jenis lantai, kondisi atap dan sumber minum. Penelitian ini menggunakan validasi Dunn Index. Berdasarkan hasil pengujian kualitas cluster dengan validasi Dunn Index di dapatkan nilai dari hasil analisis Algoritma Mean shift sebesar 0.887, kemudian di dapatkan hasil K-Means sebesar 0.995 dan K-Medoids sebesar 0.880. Maka dari ketiga algoritma tersebut dapat ditarik kesimpulan bahwa algoritma K-Means Clustering memiliki nilai yang paling baik karena paling mendekati angka 1.

Kata Kunci: Status Ekonomi, MeanShift, K-Means, K-medoids, clustering, Data Mining

1 PENDAHULUAN

Kata ekonomi berasal dari bahasa Yunani yaitu oikonomia. Kata oikonomia berasal dari dua kata, oikos dan nomos. Oikos artinya keluarga, dan nomos artinya organisasi. Oleh karena itu, oikonomia berarti menata keluarga. Pembangunan ekonomi sudah menjadi ilmu, oleh karena itu ekonomi berarti penataan ilmu yang menyeluruh untuk mengatur keluarga. [1]

Deskripsi status sosial ekonomi adalah tentang kondisi seseorang atau masyarakat yang dapat dilihat dari ekonomi, pendidikan, penghasilan dan pekerjaan. Semua masyarakat menginginkan status ekonomi sosial yang lebih baik, namun sebenarnya masih banyak orang yang mengalami penurunan ekonomi, pada kenyataannya masih banyak masyarakat yang memiliki ekonomi yang rendah. [2]

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

Status sosial ekonomi seseorang tidak hanya berlaku di Indonesia tetapi juga di luar negeri, hal tersebut dapat mempengaruhi kehidupan sosial, pekerjaan bahkan pendidikan. Menurut Nasution, kedudukan atau status menentukan kedudukan seseorang dalam struktur sosial, yakni menentukan hubungan dengan orang lain. Status atau status seseorang, apakah ia berasal dari kelas atas atau kelas bawah, berasal dari status orang lain.[3] Dalam hal ini pemetaan sosial ekonomi dan lingkungan adalah proses penggambaran secara sistematis data sosial, ekonomi dan lingkungan dari suatu lokasi tertentu.[4]

Data mining memiliki banyak pandangan, seperti penemuan pengetahuan atau pengenalan pola. Kedua istilah ini sebenarnya memiliki keakuratannya masing-masing, istilah "penemuan kolaboratif" cocok karena digunakan untuk tujuan utama data mining yaitu untuk mendapatkan pengetahuan yang masih tersembunyi di dalam bongkahan data.[5] Penambangan data adalah metode untuk pemrosesan data skala besar. Oleh karena itu, data mining memegang peranan yang sangat penting dalam berbagai bidang kehidupan diantaranya industri, keuangan, cuaca, iptek. Dalam data mining, metode seperti klasifikasi, pengelompokan, regresi, pemilihan variabel, dan analisis keranjang pasar juga dapat digunakan.[6]

Analisis cluster adalah analisis statistik yang bertujuan untuk mengklasifikasikan objek berdasarkan kesamaan fiturnya. Objek-objek ini akan dibagi menjadi satu atau beberapa kelompok (cluster) sehingga objek-objek dalam satu kelompok serupa satu sama lain.[7] Analisis clustering merupakan analisis metode penambangan objek tanpa pengawasan, sedangkan analisis clustering K-means adalah metode analisis clustering non-hierarkis, yang berupaya membagi objek yang ada menjadi satu atau lebih cluster sesuai dengan karakteristiknya. Atau cluster. Kelompok objek. Oleh karena itu, objek dengan karakteristik yang sama akan dikelompokkan ke dalam cluster yang sama, dan objek dengan karakteristik berbeda akan dikelompokkan ke dalam cluster lain[8] Dengan adanya metode pengelompokan yang baik seseorang dapat mengidentifikasi anomali dalam data, memperoleh lebih banyak pengetahuan tentang data, menghasilkan hipotesis, menemukan rasio persamaan ikatan dalam instance, melakukan kompresi, dll[9] Clustering merupakan proses mengelompokkan data ke dalam beberapa cluster atau kelompok, sehingga data dalam satu cluster memiliki tingkat kemiripan yang sama.

Pengelompokan status ekonomi masyarakat merupakan salah satu cara untuk mengetahui karakteristik tingkat kesejahteraan masyarakat di masing-masing daerah, agar kebijakan dan strategi pembangunan dapat diambil dengan benar dan tepat sasaran untuk pembangunan serta kesejahteraan lainnya. Analisis cluster berguna dalam beberapa analisis pola eksplorasi, pengelompokan dan pengambilan keputusan.

Pada penelitian sebelumnya yang dilakukan oleh Dina Sunia mengenai Clusterisasi data penduduk miskin untuk mengelompokkan penerima bantuan dengan menggunakan metode K-Means dalam pengklasteran penduduk miskin. Hasil perhitungan menunjukkan, dari 512 sample data penduduk miskin di Kecamatan Jambi Selatan, dengan jumlah klaster sebanyak 5. Penulis melakukan iterasi sebanyak 3 kali, dan mendapatkan hasil 13 penduduk masuk sebagai anggota klaster 1, 153 penduduk berada di klaster 2, 129 penduduk di klaster 3, 138 penduduk berada di klaster 4, dan 79 penduduk berada di klaster 5. Berdasarkan kondisi tersebut perlu dilakukan pengklasteran untuk membantu Dinas Sosial Kota Jambi dalam pengelompokan keluarga miskin sehingga bantuan dapat tersalurkan dengan tepat.[10]

Penelitian terkait selanjutnya adalah tentang analisis clustering K-means. Penelitian ini menggunakan metode clustering K-Means untuk menganalisis situasi kemiskinan di Jawa Barat. Dengan menentukan dan mengukur indikator kemiskinan maka akan lebih mudah bagi pemerintah untuk menentukan tingkat kemiskinan suatu daerah. Karakteristik daerah dikelompokkan sesuai dengan indikator kemiskinan sehingga pemerintah dapat merumuskan kebijakan penanggulangan kemiskinan daerah dengan tepat dan cepat. Dalam hasil penelitian ini, berdasarkan skor indikator kemiskinan tertinggi dan terendah di Jawa Barat tahun 2018, dibentuk pemetaan karakteristik masing-masing kelompok.[5]

Pada penelitian ini dilakukan perbandingan algoritma *MeanShift Clustering*, *K-Means Clustering* dan *K-Medoids*. Tujuan dari kumpulan data ini adalah untuk menganalisis pola pengelompokan dan menentukan algoritma yang baik untuk mengolah data. Data yang akan Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

digunakan adalah data penduduk Kelurahan Cipadu Kota Tangerang tahun 2017. Data atribut yang akan digunakan terdiri dari 8 atribut yaitu : jumlah anggota keluarga, status pekerjaan, status bangunan, luas bangunan, daya terpasang, jenis lantai, kondisi atap dan sumber air minum. Percobaan akan dilakukan dengan menggunakan *Python* sebagai tools, dan nilai *Dunn Index* akan digunakan sebagai acuan dalam pengelompokan cluster. Kemudian pola pengelompokan yang diperoleh dapat dianalisa untuk mendapatkan sebuah informasi baru bagi pihak kelurahan yang nanti hasilnya dapat diusulkan untuk menjadi bahan pertimbangan pemerintah setempat dalam mengelompokkan masyarakat berdasarkan status perekonomian dan juga kedepan akan dijadikan sebagai profil status yang terdapat di masyarakat kelurahan Cipadu Kota Tangerang.

2 TINJAUAN PUSTAKA

Penelitian terkait algoritma Mean Shift yang bertujuan untuk menggunakan algoritma tersebut untuk menggantikan algoritma klasifikasi berbasis piksel untuk memetakan habitat bentik. Penelitian ini sangat penting artinya karena tidak semua data penginderaan jauh dapat memperoleh data kualitas spasial, radiasi dan spektral terbaik. Hasil penelitian menunjukkan bahwa data sub-kualitas penginderaan jauh masih dapat memetakan habitat bentik dengan relatif akurat.[11]

Pada penelitian terkait selanjutnya Penelitian ini menggunakan kuisioner sebanyak 50 responden untuk mendapatkan data yang diharapkan. Data yang dihasilkan dari kuisioner adalah data acak yang belum dikelompokkan berdasarkan status ekonomi masyarakat, maka dari itu menggunakan metode K-Means Clustering karena mampu mengimplementasikan clustering dengan melakukan pengelompokan data tanpa berdasarkan kelas data tertentu. Hasil penelitian yang dihasilkan adalah penulis berhasil menggunakan metode clustering untuk mengklasifikasikan yang mampu membentuk 3 kelompok status ekonomi yaitu status ekonomi kaya, status ekonomi menengah, status ekonomi miskin.[12]

Penelitian tentang realisasi algoritma clustering K-Means. Penelitian ini menggunakan algoritma clustering K-Means dalam mengelompokkan dan rekomendasi karir berdasarkan faktor demografi, usia dan jenis kelamin. Data yang digunakan penulis adalah data penduduk / tahun kelurahan Bungursari yang lahir antara tahun 1969 sampai dengan tahun 1969-1999 lulusan SLTA. Penerapan algoritma K-Means dapat membagi data penduduk kelurahan menjadi 4 cluster. C1 adalah rekomendasi pekerjaan untuk pekerja tidak tetap, C2 adalah rekomendasi pekerjaan untuk pengusaha, C3 adalah rekomendasi pekerjaan untuk karyawan swasta, dan C4 adalah rekomendasi pekerjaan untuk karyawan honorer. Nilai akurasi hasil 79%, dan nilai AUC tergolong sedang atau cukup.[13]

Penelitian selanjutnya, tujuan dari penelitian ini adalah untuk mengetahui hubungan antara status sosial ekonomi ibu rumah tangga di RW 02 Kelurahan Panggung Kidul Semarang Utara dengan intensitas komunikasi keluarga. Dengan menggunakan teknik stratified stratified random sampling, 73 subjek berpartisipasi dalam penelitian ini. Teknologi pengumpulan data menggunakan survei kuisioner untuk mengetahui status sosial ekonomi dan skala intensitas komunikasi keluarga yang meliputi 37 item ($\alpha = 0.886$). Hasil analisis regresi sederhana menunjukkan bahwa status sosial ekonomi berkorelasi positif dengan intensitas komunikasi ibu rumah tangga ($r = 0.327$; $p = 0.005$). Artinya terdapat hubungan positif antara status sosial ekonomi dengan intensitas komunikasi keluarga, dan status sosial ekonomi memberikan sumbangan efektif sebesar 10,7% terhadap intensitas komunikasi keluarga.[2]

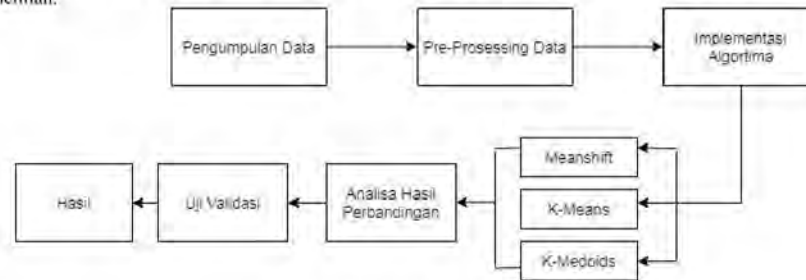
3 METODE PENELITIAN

Penelitian ini dilakukan untuk melakukan analisis terhadap data kependudukan di daerah Kelurahan Cipadu Kota Tangerang data yang digunakan adalah data tahun 2017, dengan menggunakan perbandingan algoritma *Meanshift Clustering*, *K-Means* dan algoritma *K-Medoids*. Penelitian ini dilakukan berdasarkan karakteristik pengelompokan penduduk miskin dengan atribut

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

Jumlah anggota rumah tangga, status pekerjaan, status bangunan, luas lantai, daya terpasang, jenis lantai, kondisi atap dan sumber minum.

Hasil analisis akan digunakan sebagai bahan pertimbangan pihak kelurahan Cipadu Kota Tangerang dalam meningkatkan kualitas pemetaan status ekonomi masyarakat yang ada di Kelurahan Cipadu. Adapun tahapan dari proses metode penelitian ini dapat diuraikan pada Gambar 1 Tahapan Penelitian.



Gambar 1. Tahapan Penelitian

3.1 Pengumpulan Data

Untuk melakukan penelitian ini, langkah pertama yang harus dilakukan adalah mengumpulkan data yang akan digunakan. Data yang dibutuhkan untuk penelitian ini adalah data kondisi perekonomian masyarakat di Desa Cipadu Kota Tangerang tahun 2017. Data yang diperoleh dalam penelitian ini adalah data penduduk yang berada di Kelurahan Cipadu Kota Tangerang. Data yang didapatkan berjumlah 862 data dan 40 atribut. Jumlah total data yang digunakan dalam penelitian ini berjumlah 862 dan dikonversikan didalam format csv. Data-data tersebut berupa data sensus penduduk yang telah dilakukan oleh pemerintah yang berada di Kelurahan Cipadu Kota Tangerang. Seperti terdapat pada contoh tabel 1 Dataset.

Tabel 1. Dataset

RW	Jum keluarga	Stat bangunan	Stat Lantai	Luas lantai	Jenis lantai	Jenis dinding	Jenis atap	Kondisi atap	Sumber minum	Sum penitiran	Daya terpasang	penggunaan fasilitas BAB	Stat Bekerja	Ada Komputer	Ada motor	Ada mobil	Asset Tak bergerak	Rumah Tempel Lain
3	1	2	2	30	2	1	5	2	2	1	3	1	1	4	4	2	2	4
3	1	1	1	70	2	1	4	2	5	1	2	1	1	4	3	2	2	4
4	1	1	1	30	2	1	4	2	5	1	2	1	1	4	3	2	2	4
4	1	2	2	25	4	1	5	2	2	1	2	1	1	4	4	2	2	4
4	1	1	1	25	4	1	4	2	5	1	2	1	1	4	4	2	2	4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4	1	1	1	25	4	1	4	2	5	1	2	1	1	4	4	2	2	4

3.2 Pre-Processing Data

Preprocessing data adalah proses yang dirancang untuk meningkatkan kualitas data dengan menghilangkan data yang tidak diperlukan dari data aslinya. Tidak semua persyaratan data yang diperoleh akan digunakan untuk pemrosesan, oleh karena itu data yang diperoleh akan dipilih terlebih dahulu. Hal ini dilakukan untuk menghilangkan data yang tidak relevan dan tidak valid, yang akan mempengaruhi hasil pemrosesan selanjutnya. Berdasarkan pemrosesan ini, 862 data untuk dieksekusi di pemrosesan berikutnya dihasilkan. Pada penelitian ini preprocessing meliputi:

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

3.2.1 Data Selection

Tahapan ini bertujuan untuk memilih atribut dalam data status ekonomi masyarakat. Atribut yang menjadi fokus penelitian adalah 8 atribut yang akan digunakan yaitu jumlah anggota keluarga, status pekerjaan, status bangunan, luas lantai, power supply terpasang, tipe lantai, kondisi atap dan sumber air minum, kemudian menghapus atribut yang tidak terpakai.

Tabel 2. Hasil data seleksi

No	RW	R T	Stat Bekerja	Jumlah_AR T	Stat bangunan	Luas lantai	Jenis lantai	Kondisi atap	Daya terpasang	Sumber minum
1	3	1	1	2	2	30	2	2	3	2
2	3	1	1	3	1	70	2	2	2	5
3	4	2	1	1	1	30	2	2	2	5
4	4	4	1	5	2	25	4	2	2	2
.....										
867			1	1	1	25	4	2	2	5

3.2.2 Data Transformation

Pada tahap ini data yang sudah siap diolah akan diolah terlebih dahulu. Proses preprocessing akan menggunakan min-max normalization. Normalisasi Min-Max adalah teknik normalisasi yang melakukan transformasi linier pada atribut data asli untuk melakukan rentang nilai yang sama. Normalisasi adalah teknik sederhana dimana teknik tersebut dapat secara khusus menyesuaikan data dalam batas yang telah ditentukan sebelumnya. Jika misalkan beberapa baris memiliki semua nilai identik, sehingga deviasi standar baris itu sama dengan nol maka semua nilai untuk baris itu disetel ke nol.[14] Salah satu dari banyak metode normalisasi data adalah metode min-max normalisasi Metode ini mengonversi nilai pada kumpulan data ke skala yang lebih kecil atau berkisar dari 0 hingga 1[15] Hasil konversi data ditunjukkan pada Tabel 4

Tabel 4. Data Hasil Preprocessing Min-Max

RW	RT	Jumlah_AR T	Stat Bekerja	Stat bangunan	Luas lantai	Jenis Lantai	Kondisi atap	Daya terpasan g
0.185319	0.111166	0.260948	0.191480	0.482727	0.776046	0.246443	0.241940	0.200138
0.175131	0.407470	0.215717	0.281897	0.361433	0.149355	0.161338	0.167904	0.146548
0.212152	0.111225	0.388517	0.073221	0.369940	0.134838	0.110019	0.060012	0.141999
0.374043	0.074340	0.499920	0.536182	0.227537	0.702995	0.248917	0.286227	0.376811
0.284978	0.074193	0.435451	0.076295	0.271442	0.099574	0.097200	0.000000	0.364597

3.3 Implementasi Algoritma

3.3.1 Mean Shift

Pergeseran rata-rata adalah proses penentuan *density function* atau fungsi kerapatan Kumpulan data [12]. Menurut penamaannya, fungsi pertama adalah rata-rata, artinya fungsi tersebut digunakan untuk menentukan rata-rata model atau data atau kumpulan data yang paling sering muncul. Sekaligus, fungsi kedua adalah untuk memindahkan gigi, yaitu berpindah ke mode area. Dalam algoritma pergeseran rata-rata, fungsi kerapatan atau kerapatan data sangat erat kaitannya dengan analisis ruang fitur pada citra.

Algoritma rata-rata pergeseran adalah teknik pengelompokan non-parametrik, yang tidak memerlukan jumlah klaster sebelumnya dan tidak membatasi bentuk klaster. Diberikan n titik data x_i pada ruang d -dimensional space R^d yaitu $i = 1, \dots, n$, kerapatan kernel multivariat yang diperoleh dengan kernel $K(x)$ dan radius jendela h diperkirakan sebagai

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \dots\dots\dots(1)$$

Untuk kernel yang simetris secara radial, itu cukup untuk menentukan profil kernel $k(x)$ memuaskan

$$K(x) = c_{k,d} k\left(\frac{\|x\|_2}{h}\right) \dots\dots\dots(2)$$

Dimana $c_{k,d}$ adalah konstanta normalisasi yang menjamin $K(x)$ terintegrasi ke 1. Mode fungsi kerapatan terletak di angka nol dari fungsi gradien $\nabla f(x) = 0$.

Gradien penduga kepadatan (1) is

$$\nabla f(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n x_i - x \left| g\left(\left\|\frac{x-x_i}{h}\right\|_2\right) \right| \dots\dots\dots(3)$$

Dimana $g(s) = -k'(s)$. Suku pertama sebanding dengan perkiraan kepadatan pada x yang dihitung dengan kernel $G(x) = c_{g,d} g\left(\left\|\frac{x}{h}\right\|_2\right)$ dan istilah kedua

$$m_h(x) = \sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|_2\right) \dots\dots\dots(4)$$

1.1.1 K-Means

K-Means clustering mempunyai tujuan untuk melakukan pembagian data sebanyak nilai k pada kelompok, yang berpusat di pusat atau centroid yaitu c_j , j adalah kelompok data ke-1 sampai dengan k . Cara yang dipakai pada algoritma ini ialah dengan menjauhkan sejauh mungkin antara satu pusat dengan pusat yang lain. [16] K-means adalah algoritma sederhana yang terkenal dengan efektivitasnya dalam clustering[9]

K-Means adalah algoritma untuk mengklasifikasikan atau mengelompokkan objek berdasarkan item atau fitur. K adalah bilangan positif. Proses pengelompokan yang meminimalkan jarak kuadrat antara data dan cluster centroid terkait. Oleh karena itu, tujuan utama dari k-means clustering adalah untuk mengelompokkan data. Efisiensi dan keaslian algoritma K-Means sangat bergantung pada centroid awal.

Pada data akan diproses dengan menggunakan algoritma k-means clustering. Adapun proses yang dilewati pada algoritma K-means adalah sebagai berikut:

1. Tentukan jumlah cluster yang akan digunakan. Menentukan clustering atau K merupakan hal pertama yang harus dilakukan dalam algoritma K-Means, dimana nilai K merupakan bilangan bulat positif [17] dalam penelitian ini nilai K yang digunakan adalah cluster $K = 2$ sampai $K = 6$
2. Kemudian untuk menentukan centroid, penentuan centroid dilakukan dengan memberikan nilai secara acak pada tiap atribut
3. Hitung jarak dari masing-masing data ke titik pusat tiap cluster centroid awal.
4. Hal ini untuk mengetahui cluster mana yang paling dekat dengan data.
5. Setelah memasukkan semua data ke dalam cluster terdekat, hitung ulang center cluster baru berdasarkan jumlah rata-rata anggota dalam cluster tersebut. Pada tahap ini, jika pusat massa belum berkumpul, iterasi akan berlanjut, dan hanya jika pusat massa baru berkumpul, Anda dapat menghentikan iterasi pusat massa lama.

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

1.1.2 K-Medoids

K-Medoids atau Partitioning Around Medoids (PAM) adalah metode pengelompokan kumpulan data yang dikembangkan oleh Kaufman dan Rosseuw, yang merupakan alternatif yang lebih baik untuk k-means. Pada metode ini sebelum menghitung jarak objek data ke clustering centroid, maka k clustering centroid dipilih secara acak dari n objek data sehingga objek tersebut dibuat berdasarkan kedekatan masing-masing objek pengelompokan untuk memulai partisi data. Objek yang paling terkonsentrasi dalam suatu cluster diambil sebagai titik referensi di sini yang sebenarnya adalah medoid dan bukan nilai rata-rata elemen dalam cluster. Prinsip dasar K-medoid adalah meminimalkan jumlah titik jarak yang berbeda dari partisi titik referensi yang harus dilakukan.[18] K-medoids atau partisi di sekitar algoritma medoids adalah variasi dalam algoritma k-means, dimana titik data dipilih sebagai medoid daripada memilih mean sebagai centroid di k-means.[19] Metode K-Medoids merupakan metode untuk meminimalkan jarak antara medoids dan urutan lainnya didalam grup atau kelompok[20]

Langkah-langkah dalam menyelesaikan K-Medoids adalah sebagai berikut:[21]

1. Inisialisasi pusat cluster, jumlahnya sama dengan jumlah cluster (k)
2. Tiap data atau objek dialokasikan ke dalam *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan rumus persamaan (3.1).
3. Pilih objek secara acak di setiap cluster sebagai kandidat medoid baru.
4. Hitung jarak setiap objek yang terdapat di setiap *cluster* dengan calon *medoid* baru.
5. Hitung total deviasi (S) dengan menghitung jumlah jarak baru-jumlah jarak lama. Jika $S < 0$, tukar objek dan data cluster untuk membuat kumpulan baru k objek sebagai cluster.
6. Ulangi langkah 3 sampai 5 sampai tidak ada perubahan pada *steroid*, untuk mendapatkan cluster dan anggota cluster masing-masing

Perbedaan antara algoritma K-Medoid, K-Means dan Meanshift adalah algoritma K-Medoids sangat berbeda dengan algoritma K-Means. K-Medoids memilih titik data sebagai pusatnya (*medoid*)[22] dari setiap *cluster* dan Algoritma K-Means membutuhkan rata-rata (*mean*) untuk menjadi pusat *cluster*.

3.4 Analisis Hasil Perbandingan

Pada tahapan ini analisa perbandingan yang akan dilakukan kepada ketiga algoritma yaitu Meanshift Clustering, K-Means dan K-Medoids, dari ketiga nya dengan melakukan perbandingan hasil nilai validasi Dunn Index, dari masing-masing cluster dengan mencari nilai $k=2$ s.d $k=6$, kemudian membandingkan hasil validasi yang telah didapatkan dengan mencari nilai tertinggi, dan dari ke tiga Algoritma tersebut akan dipilih mana nilai yang paling baik dengan nilai yang paling mendekati angka 1. Setelah itu membandingkan hasil grafik dari ke tiga Algoritma.

3.5 Uji Validasi

Validitas cluster adalah proses evaluasi pengujian, yang mencantumkan nilai standar kinerja berdasarkan pusat cluster. Ini berguna untuk memahami kinerja proses pengelompokan. [18]

Dalam studi ini evaluasi cluster menggunakan indeks Dunn pada perangkat Python. Hasil akhir yang diperoleh dalam proses evaluasi cluster adalah sebuah nilai, semakin besar nilai yang dihasilkan oleh Dunn Index maka semakin baik evaluasi cluster tersebut. Indeks Efektivitas Dunn (DN) menghitung rasio antara rasio antara nilai minimum perbedaan fungsional antara dua cluster (sebagai interval) dan rasio antara nilai maksimum diameter cluster (sebagai ketegasan). Jumlah cluster terbaik diwakili oleh semakin besar nilai DN (Dunn, J.C., 1974)[19].

4 HASIL DAN PEMBAHASAN

Setelah dilakukan *preprocessing* data, kemudian data diolah menggunakan *python* dan dilakukan validasi menggunakan *Dunn Index* untuk sebagai acuan dalam pengelompokan terbaik. Uji klaster dilakukan pada klaster $k = 2$ s.d $k = 6$ untuk mencari nilai yang klaster yang paling

baik pada ketiga Algoritma yaitu K-Means, K-Medoids dan Meanshift. Tabel 6 mencantumkan hasil pengolahan data berikut :

Tabel 6. Perbandingan Uji Validasi Ketiga Algoritma Menggunakan Dunn Index

Algoritma	K	Validasi	Algoritma	K	validasi	Algoritma	K	validasi
K-Medoids	2		K-Means	2		MeanShift	4	0.887
		0.861			0.887			
	3	0.879		3	0.995			
	4	0.823		4	0.976			
	5	0.880		5	0.923			
	6	0.880		6	0.955			

Berdasarkan perbandingan nilai validasi *Dunn Index*, menunjukan bahwa $k=3$ pada algoritma K-Means menjadi kluster terbaik dikarenakan nilai $k=3$ adalah 0.995 yaitu paling mendekati angka 1.

Untuk Algoritma Meanshift sendiri tidak perlu mencari nilai k seperti K-means dan K-Medoids dikarenakan Meanshift sendiri langsung menentukan kluster terbaiknya tanpa bantuan $n_clusters$ seperti K-Means, sehingga untuk algoritma Meanshift tidak diperlukan percobaan nilai k .

Tabel 7. Hasil Terbaik Dari Ketiga Algoritma Meanshift, K-Means, dan K-Medoid

Algoritma	Validasi Algoritma terbaik
K-Means	0.995
K-Medoids	0.880
Meanshift	0.887

Dapat dilihat pada dari ketiga Algoritma dan K-Medoids validasi yang cukup K-Means memiliki nilai kedua algoritma validasi K-Means validasi *Dunn Index* mendapatkan nilai tertinggi yaitu 0.995 sangat mendekati angka 1, yaitu dapat diartikan bahwa Algoritma K-Means yang paling baik.

Tabel 7. Bahwa hasil Meanshift, K-Means mendapatkan nilai tinggi dan Algoritma validasi terbaik dari lainnya. Karena nilai dengan menggunakan

Berikut ini adalah anggota dari setiap cluster dengan nilai cluster terbaik yang disertakan di setiap algoritma. Berikut adalah tabel anggota di setiap cluster:

Tabel 8. Jumlah Anggota Masing-masing Cluster pada Algoritma K-Means dengan nilai $K=3$

	Cluster 0	Cluster 1	Cluster 2
RT	364	267	231
RW	364	267	231
Jumlah ART	364	267	231
stat Bekerja	364	267	231
Luas lantai	364	267	231
Stat bangunan	364	267	231
Daya terpasang	364	267	231
Jenis lantai	364	267	231
kondisi atap	364	267	231
sumber minum	364	267	231

Pada Tabel 8. Ini merupakan hasil pengelompokan menggunakan algoritma K-means dengan nilai $K=3$. Hasil pengelompokan menunjukkan bahwa dibandingkan dengan cluster lain, anggota pada cluster 0 memiliki anggota yang paling banyak.

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

Tabel 9. Jumlah Anggota Masing-masing Cluster Pada Algoritma K-Medoids Dengan Nilai K=6.

	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
RT	139	160	163	193	101	106
RW	139	160	163	193	101	106
Jumlah ART	139	160	163	193	101	106
stat Bekerja	139	160	163	193	101	106
Luas lantai	139	160	163	193	101	106
Stat bangunan	139	160	163	193	101	106
Daya terpasang	139	160	163	193	101	106
Jenis lantai	139	160	163	193	101	106
kondisi atap	139	160	163	193	101	106
sumber minum	139	160	163	193	101	106

Pada Tabel 9. Merupakan hasil pengelompokan dengan menggunakan Algoritma K-Medoids dengan nilai K sebesar 6. Hasil dari pengelompokan tersebut didapatkan bahwa anggota pada cluster 3 memiliki anggota terbanyak dibandingkan cluster lainnya.

Tabel 10. Anggota Masing-masing Cluster Pada Algoritma Mean Shift Dengan Nilai Clustering=4

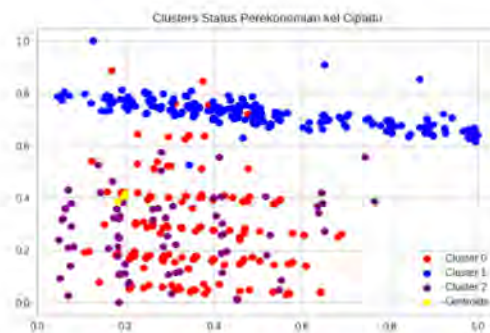
	Cluster 0	Cluster 1	Cluster 2	Cluster 3
RT	841	19	1	1
RW	841	19	1	1
Jumlah ART	841	19	1	1
stat Bekerja	841	19	1	1
Luas lantai	841	19	1	1
Stat bangunan	841	19	1	1
Daya terpasang	841	19	1	1
Jenis lantai	841	19	1	1
kondisi atap	841	19	1	1
sumber minum	841	19	1	1

Pada Tabel 10. Merupakan hasil pengelompokan dengan menggunakan Algoritma Mean Shift dengan nilai hasil clustering yang di dapat yaitu empat cluster. Hasil dari pengelompokan tersebut didapatkan bahwa anggota pada cluster 0 memiliki anggota terbanyak dibandingkan cluster lainnya.

Kemudian pada Gambar 1 dapat dilihat bahwa hasil $k = 3$ dapat digunakan pada grafik untuk menunjukkan sebaran cluster pada grafik.

UNIVERSITAS
MERCU BUANA

Gambar 1. Hasil Grafik K-Means



Hasil grafik yang di dapat untuk melihat hasil cluster menggunakan $k = 3$ pada Algoritma K-Means dapat dilihat pada Gambar 1. Berikut penjelasan hasil bagan K-Means dengan tiga cluster, direpresentasikan warna merah pada cluster 0, terdiri dari 8 RW (Rukun Warga) dan 23 RT (Rukun Tetangga) di cluster 1, direpresentasikan dengan warna biru itu adalah 8 RW dan 21 RT. Pada kelompok kedua yang digambarkan dengan warna ungu terdapat 7 RW dan 12 RT.

Gambar 2. Hasil Grafik K-Medoids

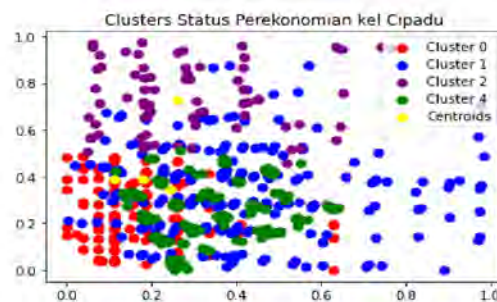


Pada Gambar 2. Seperti terlihat pada gambar, terdapat beberapa cluster di setiap cluster, dan cluster tersebut memiliki rentang nilai yang berbeda. Pada grafik berbentuk K-Medoids dengan 6 cluster. Berikut penjelasan hasil grafik K-Medoids 6 cluster, pada cluster 0 diwakili oleh warna merah yang terdiri dari 4 RW dan 13 RT, dan untuk cluster 1 diwakili oleh warna biru. Dapatkan 8 RW dan 24 RT, untuk cluster 2 diwakili oleh warna ungu mendapatkan 6 RW dan 22 RT, untuk cluster 3 yang berwarna hijau mendapatkan 5 RW dan 10 RT, untuk cluster 4 dengan warna orange mendapatkan 5 RW dan 7 RT, dan untuk cluster 5 warna pink mendapatkan 5 RW dan 8 RT.

MERCU BUANA

Gambar 3. Hasil Grafik Meanshift Clustering

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang



Pada Gambar 3, Merupakan grafik pengelompokan dengan algoritma meanshift terdapat empat cluster. Algoritma Meanshift sendiri dapat langsung menentukan cluster terbaik tanpa bantuan n clusters seperti K-Means, sehingga algoritma Meanshift tidak perlu menggunakan nilai k untuk eksperimen. Disini dapat dilihat bahwa untuk algoritma meanshift lebih banyak atau lebih dominan pada cluster 0.

Berikut penjelasan hasil peta cluster Meanshift dengan empat cluster, Pada cluster 0 direpresentasikan dengan warna merah meliputi 8 RW dan 56 RT. Untuk cluster 1 dideskripsikan dengan warna biru terdapat 6 RW dan 11, untuk cluster 2 dideskripsikan sebagai 1 RW dan 1 RT berwarna ungu, dan 1 RW dan 1 RT berwarna hijau untuk cluster 4.

Dari hasil pengujian Validasi dan membandingkan ketiga hasil grafik dari algoritma Meanshift Clustering, K-Means dan K-Medoids

Tabel 11. Analisa Cluster Hasil Uji Validasi Terbaik K-Means Clustering

Cluster	RW	RT
0	1	2,7,8
	2	1,3,6,7
	3	2,4
	4	2,3
	5	2,3,4
	6	1,6,7
	7	1,2,3
	8	1,3,4
1	1	1,4,5,6,9,10
	2	2,4,5
	3	1,5,6,7
	4	1,4,8
	5	5
	6	2,4,5
	7	5
	8	2,5,6
2	1	3,11
	2	10
	3	3,5

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

SISTEMASI : Jurnal Sistem Informasi

	5	1
	6	8
	7	4,6,7,8

Dari hasil analisa cluster algoritma K-Means diatas dapat dilihat bahwa:

1. Cluster 0 terdiri dari 8 Rukun Warga(RW) dan 23 Rukun Tetangga(RT)
2. Cluster 1 terdiri dari 8 RW dan 21 RT
3. Cluster 2 terdiri dari 7 RW dan 12 RT

Pada Tabel 11. Menunjukkan hasil analisa cluster dari algoritma K-Means yang mendapatkan nilai validasi *Dunn Index* terbesar yaitu 0.995. Dari analisa cluster yang telah dilakukan di dapatkan 3 Cluster pada algoritma K-means, yang merupakan kumpulan wilayah yang memiliki nilai validasi terbaik dari kedua algoritma pembandingan lainnya, sehingga dapat dikelompokkan sebagai status perekonomian masyarakat di Kelurahan Cipadu Kota Tangerang.

Adapun dari ketiga cluster yaitu cluster 0, cluster 1 dan cluster 2. Dari ketiganya dapat digambarkan sebagai berikut : Cluster 0 merupakan kelompok wilayah status perekonomian masyarakatnya Tidak Mampu terdiri dengan 8 RW dan 20 RT. Cluster 1 merupakan kelompok wilayah status perekonomian masyarakatnya Sedang terdiri dengan 8 RW dan 21 RT, dan Cluster 2 merupakan kelompok wilayah status perekonomian masyarakatnya Mampu terdiri dari 7 RW dan 12 RT.

5 KESIMPULAN

Pada penelitian ini, Algoritma Mean Shift, K-Means dan K-Medoids yang diimplementasikan pada kumpulan data kependudukan di Kelurahan Cipadu pada tahun 2017, untuk menganalisis pola pengelompokan dan menentukan algoritma yang baik untuk pengolahan data. Berdasarkan hasil eksperimen, algoritma K-Means memiliki performansi yang lebih baik dibandingkan dengan dua Algoritma pembandingan lainnya. Pola pengelompokan masing-masing cluster di peroleh berdasarkan jumlah anggota keluarga, status pekerjaan, status bangunan, luas bangunan, daya terpasang, tipe lantai, status atap dan sumber air minum.

Selanjutnya pada algoritma *K-Means* didapatkan hasil dengan validasi *Dunn Index* diperoleh nilai $k=3$ sebesar 0.995, untuk algoritma *K-Medoids* dengan validasi *Dunn Index* diperoleh hasil $k=3$ sebesar 0.880 dan juga untuk algoritma *Meanshift Clustering* didapatkan 4 cluster dengan validasi *Dunn Index* dengan nilai sebesar 0.887. Dari hasil uji validasi menggunakan validitas *Dunn Index* didapatkan Algoritma K-Means yang memiliki nilai paling baik diantara kedua algoritma pembandingan lainnya. Dari algoritma K-Means didapatkan 3 cluster terbaik yaitu cluster 0, cluster 1 dan cluster 2. Maka diinterpretasikan hasil Cluster 0 merupakan kelompok wilayah status perekonomian masyarakatnya Tidak Mampu, Cluster 1 merupakan kelompok wilayah status perekonomian masyarakatnya Sedang, dan Cluster 2 merupakan kelompok wilayah status perekonomian masyarakatnya Mampu.

Maka pada tahap pengolahan data ini dapat disimpulkan bahwa algoritma *K-Means* memiliki hasil yang paling baik diantara *Meanshift* dan *Kmedoids*. Hasil clustering nantinya akan dijadikan sebagai profil status yang ada dimasyarakat kelurahan Cipadu Kota Tangerang.

MERCU BUANA

REFERENSI

Alendia, *Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang*

- [1] M. M. Baldasaro, "No Title 血清及尿液特定蛋白检测在糖尿病肾病早期诊断中的意义," vol. 7, pp. 219–232, 2014.
- [2] E. S. Indrawati, "Status Sosial Ekonomi Dan Intensitas Komunikasi Keluarga Pada Ibu Rumah Tangga Di Panggung Kidul Semarang Utara," *J. Psikol. Undip*, vol. 14, no. 1, pp. 52–57, 2015, doi: 10.14710/jpu.14.1.52-57.
- [3] M. Thorir, Y. Soesatyo, and Harti, "Pengaruh Status Sosial Ekonomi," p. 150, 2016.
- [4] D. Rini and A. Zanuardi, "PEMETAAN TINGKAT KESEJAHTERAAN MASYARAKAT WILAYAH KAWASAN KAKI JEMBATAN SURAMADU (KKJS) DENGAN MODEL INDEKS INDIKATOR NON-INCOME Social Welfare Mapping of the Community in the Access Road of," no. April 2012, pp. 33–43, 2012.
- [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. Darma, S. Defit, D. Hartama, and W. Robiansyah, "Penerapan Metode K-Means Dalam Pengelompokan Jumlah Wisatawan Asing Di Indonesia," vol. 2, pp. 255–261, 2020.
- [7] L. Ramadhani, I. Purnamasari, and F. D. T. Amijaya, "Penerapan Metode Complete Linkage dan Metode Hierarchical Clustering Multiscale Bootstrap (Studi Kasus: Kemiskinan Di Kalimantan Timur Tahun 2016)," *Eksponensial*, vol. 9, no. 2016, pp. 1–10, 2018, [Online]. Available: [https://fmipa.unmul.ac.id/files/docs/\[1\] LISDA RAMADHANI 1307015041 Edit.pdf](https://fmipa.unmul.ac.id/files/docs/[1] LISDA RAMADHANI 1307015041 Edit.pdf).
- [8] Ediyanto, N. Mara, and N. S. Intisari, "Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis," *Bul. Ilm. Mat. Stat. dan Ter.*, vol. 02, no. 2, pp. 133–136, 2013.
- [9] S. A. Abbas, A. Aslam, A. U. Rehman, W. A. Abbasi, S. Arif, and S. Z. H. Kazmi, "K-Means and K-Medoids: Cluster Analysis on Birth Data Collected in City Muzaffarabad, Kashmir," *IEEE Access*, vol. 8, pp. 151847–151855, 2020, doi: 10.1109/ACCESS.2020.3014021.
- [10] K. dan P. A. J. Dina Sunia, "Penerapan Data Mining untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means," *J. Ilm. Mhs. Tek. Informatika*, vol. Vol 1 No 2, no. 2016, pp. 121–134, 2019.
- [11] P. Wicaksono and N. M. Farda, "Aplikasi Algoritma Klasifikasi Mean Shift untuk Pemetaan Habitat Bentik Studi Kasus Kepulauan Karimunjawa," *Pros. PIT MAPIN XX 2015*, no. November 2015, pp. 664–672, 2015, doi: 10.13140/RG.2.1.1404.6165.
- [12] 2011 Citra Kunia putri dan trisna insan Noor, "濟無 No Title No Title," *Anal. pendapatan dan tingkat Kesejaht. rumah tangga petani*, vol. 53, no. 9, pp. 1689–1699, 2013.
- [13] L. Listiani, Y. H. Agustin, and M. Z. Ramdhani, "Implementasi algoritma k-means cluster untuk rekomendasi pekerjaan berdasarkan pengelompokan data penduduk," pp. 761–769, 2017.
- [14] S. G. K. Patro and K. Kumar, "Normalization: A Preprocessing Stage,"
- [15] J. T. Hardinata, H. Okprana, A. P. Windarto, and W. Saputra, "Analisis Laju Pembelajaran dalam Mengklasifikasi Data Wine Menggunakan Algoritma Backpropagation," *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 3, no. 2, p. 422, 2019, doi: 10.30645/j-sakti.v3i2.161.
- [16] P. Algoritme, S. Mean, D. A. N. Pemilihan, F. Dasar, N. Ihromi, and H. Yat, "I ~," 2006.
- [17] S. Syahidatul Helma *et al.*, "Clustering pada Data Fasilitas Pelayanan Kesehatan Kota Pekanbaru Menggunakan Algoritma K-Means," *Puzzle Res. Data Technol. Fak. Sains dan Teknol.*, vol. 1, no. November, p. 4, 2019.
- [18] P. Kumar and D. Sirohi, "Comparative analysis of FCM and HCM algorithm on Iris data set," *Int. J. Comput. Appl.*, vol. 5, no. 2, pp. 33–37, 2010, doi: 10.5120/888-1261.
- [19] A. Alsayat and H. El-Sayed, "Efficient genetic K-Means clustering for health care knowledge discovery," *2016 IEEE/ACIS 14th Int. Conf. Softw. Eng. Res. Manag. Appl. SERA 2016*, pp. 45–52, 2016, doi: 10.1109/SERA.2016.7516127.
- [20] 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.
- [21] I. Kamila, U. Khairunnisa, and M. Mustakim, "Perbandingan Algoritma K-Means dan K-Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

SISTEMASI : Jurnal Sistem Informasi

- 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.
- [22] W. E. Susanto and D. Riana, "Komparasi Algoritma," *J. Speed*, vol. 8, no. 3, pp. 18–27, 2016.



UNIVERSITAS
MERCU BUANA

Alendia, Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat Di Kelurahan Cipadu Kota Tangerang

KERTAS KERJA

Ringkasan

Kertas kerja ini merupakan material kelengkapan artikel jurnal dengan judul “*Studi Komparasi Algoritma Clustering Dalam Memetakan Status Perekonomian Masyarakat di Kelurahan Cipadu Kota Tangerang*”. 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 literature 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, format awal, atribut data, dan penyesuaian data akhir yang siap untuk diolah. Bagian IV memuat tahapan eksperimen yang disajikan dalam bentuk blok diagram dengan penjelasan dari setiap tahapan. Bagian V merupakan bagian terakhir dari kertas kerja ini yang menjelaskan hasil keseluruhan dari eksperimen yang telah dilakukan, meliputi penjelasannya.

UNIVERSITAS
MERCU BUANA