

PERNYATAAN

Saya menyatakan bahwa skripsi ini masih mengandung plagiat dibawah batas yang diterapkan, dan apabila di kemudian hari terbukti terdapat plagiasi dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 27 Agustus 2023



Tia Mutiani

KATA PENGANTAR

Puji syukur kepada tuhan yang Maha Esa atas segala nikmatnya sehingga penulis dapat menyelesaikan studi dan menyusun proposal skripsi yang berjudul “Penerapan Variasi Metode K-Means Untuk Pengelompokan Kecamatan Di Kabupaten Bojonegoro Berdasarkan Data Kependudukan”. Ucapan terima kasih penulis haturkan kepada seluruh pihak yang terlibat dalam penulisan proposal skripsi ini berupa saran, motivasi, dan do’a, yaitu kepada :

1. Bapak K.H Jauharul Ma’arif, M.Pd.I selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Dr. Ridlwan Hambali, Lc., M.A selaku Wakil Rektor 1 Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
3. Bapak Dr. H. Yogi Prana Izza, Lc., M.A selaku Wakil Rektor II Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
4. Bapak Dr. Nurul Huda, M.H.I selaku Wakil Rektor III Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
5. Ibu Dr. Hj. Ifa Khoiria Ningrum, S.E., M.M selaku Wakil Rektor IV Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
6. Bapak Sunu Wahyudhi, M.Pd selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
7. Ibu Nita Cahyani, M.Stat selaku Ketua Program Studi Statistika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
8. Bapak Denny Nurdiansyah, M.Si selaku Dosen Pembimbing I, yang telah membimbing dengan sepenuh hati dan memberikan ilmu yang bermanfaat dalam menyelesaikan proposal skripsi ini.
9. Ibu Fetrika Angraini, M.Pd selaku Dosen Pembimbing II, yang telah memberikan ilmu yang bermanfaat dalam penyempurnaan penulisan pada proposal skripsi ini.
10. Terkhusus untuk kedua orang tua terkasih dan teman terdekat penulis yang senantiasa memberikan do’a, motivasi dan semangat kepada penulis sejak memulai perkuliahan hingga skripsi ini terselesaikan.

Penulis menyadari bahwa proposal skripsi ini masih jauh dari kata sempurna disebabkan kemampuan dan pengetahuan penulis, maka penulis meminta maaf yang sebesar-besarnya. Kritik dan saran yang membangun sangat penulis harapkan demi sempurnanya proposal skripsi ini dimasa yang akan datang. Penulis berharap agar proposal skripsi ini dapat berguna bagi semua orang.

Bojonegoro, 03 Maret 2023

Penulis,

Tia Mutiani



UNUGIRI

ABSTRACT

Mutiani, Tia. 2023. *Implementation of Variations of K-Means Method for Clustering Sub-Districts in Bojonegoro Regency Based on Population Data*. Undergraduate Thesis, Department of Statistics, Faculty of Science and Technology, Nahdlatul Ulama Sunan Giri University. Main Supervisor: Denny Nurdiansyah, M.Si.; Assistant Supervisor: Fetrika Anggraini, M.Pd. This study aims to implement variations of the K-Means method for clustering sub-districts in Bojonegoro Regency based on population data. Population data is crucial information for regional planning and development. Clustering sub-districts can provide insights into diverse population patterns within the area. In this context, data mining techniques are applied to identify patterns and relationships within population data. The K-Means method is a clustering technique that divides data into groups or clusters based on shared characteristics. In this research, variations of the K-Means method are applied, including K-Means Kernel, which involves mapping data to higher dimensions using mapping functions before the clustering process. Additionally, the K-Means Fast algorithm is employed to expedite model training time by addressing the re-centering issue of cluster centroids as data grows. The study's results indicate that the most optimal K-Means variant is K-Means Kernel with a cluster count (k) of 5. Evaluation is conducted by measuring the average distance within clusters. At $k = 5$, the coordinate patterns of data in K-Means Kernel already show a gradual trend, enabling clear visualization of clusters. The K-Means Kernel method utilized in this research is of the radial type, commonly known as radial kernel. Employing these variations of the K-Means method offers a better understanding of sub-district clustering in Bojonegoro Regency based on population data. The findings from this research can serve as a foundation for decision-making in regional planning and community development in Bojonegoro Regency.

Keywords: Population Data, K-Means, K-Means Kernel, K-Means Fast.

ABSTRAK

Mutiani, Tia. 2023. *Penerapan Variasi Metode K-Means Untuk Pengelompokan Kecamatan di Kabupaten Bojonegoro Berdasarkan Data Kependudukan*. Skripsi, Jurusan Statistika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri. Pembimbing utama Denny Nurdiansyah, M.Si. dan Pembimbing Pendamping Fetrika Anggraini, M.Pd. Penelitian ini bertujuan untuk menerapkan variasi metode K-Means dalam pengelompokan kecamatan di Kabupaten Bojonegoro berdasarkan data kependudukan. Data kependudukan merupakan informasi penting dalam perencanaan dan pengembangan wilayah. Pengelompokan kecamatan dapat memberikan wawasan tentang pola penduduk yang berbeda di wilayah tersebut. Dalam konteks ini, teknik data mining diterapkan untuk mengidentifikasi pola dan hubungan dalam data kependudukan. Metode K-Means adalah teknik pengelompokan yang membagi data ke dalam kelompok atau cluster berdasarkan kesamaan karakteristik. Pada penelitian ini, variasi dari metode K-Means diterapkan, termasuk K-Means Kernel, yang melibatkan pemetaan data ke dimensi yang lebih tinggi menggunakan fungsi pemetaan sebelum proses pengelompokan. Selain itu, juga digunakan algoritma K-Means Fast yang mempercepat waktu pelatihan model dengan mengatasi masalah penyesuaian ulang pusat cluster saat data bertambah. Hasil penelitian menunjukkan bahwa varian K-Means yang paling baik adalah K-Means Kernel dengan jumlah cluster (k) sebesar 5. Evaluasi dilakukan dengan mengukur jarak rata-rata dalam cluster. Pada $k = 5$, pola koordinat data pada K-Means Kernel sudah menunjukkan kecenderungan yang landai, sehingga cluster dapat dilihat dengan jelas. Metode K-Means Kernel yang digunakan dalam penelitian ini adalah K-Means Kernel jenis radial, atau lebih dikenal sebagai Kernel radial. Penggunaan variasi metode K-Means ini memberikan pemahaman yang lebih baik tentang pengelompokan kecamatan di Kabupaten Bojonegoro berdasarkan data kependudukan. Hasil penelitian ini dapat digunakan sebagai dasar untuk pengambilan keputusan dalam perencanaan wilayah dan pengembangan masyarakat di Kabupaten Bojonegoro.

Kata Kunci: Dokumen Kependudukan, K-Means, K-Means Kernel, K-Means Fast.

HALAMAN PERSETUJUAN

Yang bertandatangan di bawah ini Dosen Pembimbing dari:

Nama : Tia Mutiani

NIM : 2520190045

Program Studi/Fakultas : Statistika / Sains dan Teknologi

Judul Skripsi : Penerapan Variasi Metode *K-Means* Untuk
Pengelompokan Kecamatan di Kabupaten Bojonegoro
Berdasarkan Data Kependudukan

Menyatakan bahwa Mahasiswa tersebut telah disetujui dan memenuhi syarat untuk diajukan melakukan ujian skripsi.

Bojonegoro, 12 Agustus 2023

Pembimbing 1


Denny Nurdiansyah, M.Si.
NIDN. 0726058702

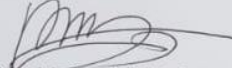
Pembimbing 2


Petrika Anggraini, M.Pd.
NIDN. 0718038803


HALAMAN PENGESAHAN

Nama : Tia Mutiani
Nim : 2520190045
Judul Skripsi : Penerapan Metode Variasi K-Means untuk Pengelompokan
Kecamatan di Kabupaten Bojonegoro Berdasarkan Data
Kependudukan
Telah diujikan dalam ujian skripsi pada tanggal 28 Agustus 2023


Dewan Penguji
Penguji I


Nur Mahmudah, M. Stat.
NIDN.0715039201


Tim Pembimbing
Pembimbing I


Denny Nurdiansyah, M.Si.
NIDN. 0726058702

Penguji II


Dr. H. M Ridlwan Hambali, Lc., M.A.
NIDN. 2117056803

Pembimbing II


Fetrika Anggraini, M. Pd.
NIDN. 0718038803

Mengetahui,
Dekan Fakultas Sains & Teknologi


Sunu Wahyudhi, M.Pd
NIDN. 0709058902
FST UNUGIRI

Mengetahui,
Ketua Program Studi


Nita Cahyani, M.Stat
NIDN. 0703038906
FST UNUGIRI
BOJONEGORO

MOTTO

إِنَّمَا أَمْرُهُ إِذَا أَرَادَ شَيْئًا أَنْ يَقُولَ لَهُ كُنْ فَيَكُونُ

(Sesungguhnya urusan-Nya menciptakan segala sesuatu sangatlah mudah bagi-Nya. Apabila dia menghendaki untuk menciptakan sesuatu, dia hanya berkata kepadanya, 'jadilah!' maka dengan serta-merta jadilah sesuatu yang dikehendaki-Nya itu. Qs. Yasin (36:82))

“Hidup yang tidak dipertaruhkan tidak akan pernah dimenangkan.”

(Sutan Sjahrir)

PERSEMBAHAN

Bapak dan Ibu

Skripsi ini saya persembahkan kepada kedua orang tua saya yang telah menjadi tempat inspirasi, motivasi, serta menjadi tujuan dari jalan kehidupan saya. Karena segala pintu berkahNya ada pada beliau berdua, Bapak dan Ibu.

Komunitas Sematta Sastra

Terimakasih telah kebersamai dalam proses penulisan skripsi ini. Dari teman – teman Komunitas Sematta saya belajar banyak hal baru dan baru. Komunitas Sematta layaknya embun rindu yang berpadu dalam jalan keilmuan serta kehidupan.

DAFTAR ISI

KATA PENGANTAR	iii
HALAMAN PENGESAHAN.....	Error! Bookmark not defined.
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Teoritis.....	5
1. Bagi Peneliti	5
3. Bagi Akademisi	5
4. Bagi Masyarakat.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Dasar Teori.....	6
2.2 Kerangka Teori.....	11
2.2.1 Data Kependudukan.....	11
2.2.2 Metode K-Means	12
2.2.3 Variasi Metode K-Means.....	14
BAB III METODOLOGI PENELITIAN.....	17
3.1 Sumber Data	17
3.2 Variabel Penelitian	17
3.3 Langkah Analisis.....	18
3.4 Diagram Alir.....	20
BAB 4 HASIL DAN PEMBAHASAN	22
4.1 Statistika Deskriptif Data Pengamatan.....	22
4.2 Metode Clustering Variasi K – Means.....	24
4.3 Metode <i>K-Means</i> Terbaik	29
4.4 Data Hasil <i>Clustering</i> Administrasi Kependudukan	36
4.5 Statistika Deskriptif Data Cluster.....	37
4.6 Histogram Clustering Pada Rapidminer Studio	40

.....	46
BAB 5 KESIMPULAN DAN SARAN	48
5.1 Kesimpulan.....	48
5.2 Saran.....	50



UNUGIRI

DAFTAR GAMBAR

Gambar 3. 1 Diagram Analisis Data dengan variasi metode K – Means... 20**Error!**

Bookmark not defined.

Gambar 4. 1 Clustering variasi K – Means.....	25
Gambar 4. 2 Grafik Clustering K-Means Clustering	26
Gambar 4. 3 Grafik Clustering K-Means Kernel.....	27
Gambar 4. 4 Grafik Clustering K-Means Fast	28
Gambar 4. 5 Histogram Data Kelahiran.....	40
Gambar 4. 6 Histogram Data Kematian.....	41
Gambar 4. 7 Histogram Data Pindah	42
Gambar 4. 8 Histogram Data Pendatang.....	43
Gambar 4. 9 Box Plot Data Kelahiran	44
Gambar 4. 10 Box Plot Data Kematian.....	45
Gambar 4. 11 Box Plot Data Pindah	46
Gambar 4. 12 Box Plot Data Pendatang.....	47

UNUGIRI