



**KLASTERISASI LINGKUNGAN PERUMAHAN DI JAKARTA SELATAN
MENGUNAKAN *MACHINE LEARNING***

SKRIPSI

ARIEF NUR ZAID

2010511061

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI SARJANA INFORMATIKA

2024



**KLASTERISASI LINGKUNGAN PERUMAHAN DI JAKARTA SELATAN
MENGUNAKAN *MACHINE LEARNING***

SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar

Sarjana Komputer

ARIEF NUR ZAID

2010511061

UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA

FAKULTAS ILMU KOMPUTER

PROGRAM STUDI SARJANA INFORMATIKA

2024

PERNYATAAN ORISINALITAS

Tugas skripsi ini adalah hasil karya sendiri, dan semua sumber yang dikutip maupun dirujuk telah saya nyatakan dengan benar.

Nama : Arief Nur Zaid

NIM : 2010511061

Tanggal : 29 Juli 2024

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan saya ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta , 30 Juli 2024

Yang Menyatakan



(Arief Nur Zaid)

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademika Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini

Nama : Arief Nur Zaid
NIM : 2010511061
Fakultas : Ilmu Komputer
Program Studi : Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Non Eksklusif (*Non-Exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

KLASTERISASI LINGKUNGAN PERUMAHAN DI JAKARTA SELATAN MENGUNAKAN MACHINE LEARNING

Dengan hak bebas royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam pangkalan data (basis data), merawat dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 30 Juli 2024
Yang Menyatakan,



Arief Nur Zaid

LEMBAR PENGESAHAN

Tugas Akhir ini diajukan oleh:

Nama : Arief Nur Zaid
NIM : 2010511061
Program Studi : S1 Informatika
Judul Tugas Akhir : KLASTERISASI LINGKUNGAN PERUMAHAN DI
JAKARTA SELATAN MENGGUNAKAN MACHINE
LEARNING

Telah berhasil dipertahankan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jakarta.


Dr. Bambang Sasas Yulistiawan, S.T., M.Kom.

Penguji I


Jayanta, S.Kom., M.Si.

Penguji II


Dr. Tjahjanto, S.Kom., M.M.

Pembimbing I


Ika Nurlaili, S.Kom, M.Sc.

Pembimbing II


Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM.

Dekan


Dr. Widya Cholil, M.I.T.

Kepala Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 27 Juni 2024

KLASTERISASI LINGKUNGAN PERUMAHAN DI JAKARTA SELATAN MENGUNAKAN *MACHINE LEARNING*

Arief Nur Zaid

ABSTRAK

Jakarta Selatan, salah satu kota metropolitan di Indonesia, adalah tempat yang dinamis dan terus berkembang. Ada berbagai lingkungan perumahan di daerah ini, mulai dari kawasan elit hingga pemukiman padat. Kota Jakarta Selatan telah berkembang pesat dengan banyak fasilitas, tetapi ada beberapa hal yang perlu diperbaiki untuk menjadi kota yang ideal dan tidak ada kesenjangan. Akses kesehatan, akses pendidikan, dan kualitas fasilitas umum adalah masalah yang sering menjadi perhatian utama bagi masyarakat dan pemerintah setempat. Lingkungan seseorang sangat dipengaruhi oleh lingkungan sekitarnya dan kualitas hidupnya. Oleh karena itu, memahami kondisi lingkungan perumahan di Jakarta Selatan bekerja sangat penting. Dengan bantuan teknologi modern seperti pembelajaran mesin, penulis dapat melakukan analisis mendalam untuk menemukan pola dan klusterisasi berdasarkan berbagai indikator. Dengan menggunakan klusterisasi, data yang dihasilkan dari metrik seperti akses pendidikan, akses transportasi, kesehatan, dan ruang hijau, dapat memberikan gambaran umum tentang kondisi lingkungan perumahan di Jakarta Selatan. Penelitian ini akan disajikan dalam bentuk *maps* dengan tiap kelurahan diberikan warna berdasarkan luas kelurahannya untuk menunjukkan klasternya dengan tiap kluster memiliki warna yang berbeda. Penelitian ini diharapkan dapat memberikan informasi yang bermanfaat bagi pemangku kepentingan dalam proses pengambilan keputusan strategis yang berkaitan dengan peningkatan dan perbaikan layanan publik.

Kata kunci : Pembelajaran mesin, perumahan, klusterisasi, ideal, keputusan, perbaikan

CLASSTERIZATION OF RESIDENTIAL ENVIRONMENTS IN SOUTH JAKARTA USING MACHINE LEARNING

Arief Nur Zaid

ABSTRACT

South Jakarta, one of the metropolitan cities in Indonesia, is a dynamic and continuously developing place. There are various residential neighborhoods in this area, ranging from elite areas to dense residential areas. The city of South Jakarta has developed rapidly with many facilities, but there are several things that need to be improved to become an ideal city and there are no gaps. Access to health, access to education and the quality of public facilities are issues that are often a major concern for the community and local government. A person's environment is greatly influenced by the surrounding environment and the quality of life. Therefore, understanding the conditions of the residential working environment in South Jakarta is very important. With the help of modern technology such as machine learning, the author can carry out in-depth analysis to find patterns and clustering based on various indicators. By using clustering, data generated from metrics such as access to education, access to transportation, health and green space, can provide a general picture of the conditions of residential neighborhoods in South Jakarta. This research will be presented in the form of a map with each sub-district given a color based on the area of the sub-district to show its cluster with each cluster having a different color. This research is expected to provide useful information for stakeholders in the process of making strategic decisions related to increasing and improving public services .

Keywords : Machine learning, housing, clustering, ideal, decision, improvement.

Kata Pengantar

Puji syukur saya panjatkan karena berkat kehendak tuhan yang maha esa Allah SWT. Skripsi saya yang berjudul “Klasterisasi Lingkungan Perumahan di Jakarta Selatan Menggunakan Machine Learning” dapat terselesaikan. Penyusunan skripsi ini bertujuan untuk memenuhi persyaratan meraih gelar sarjana komputer pada Fakultas Ilmu Komputer UPNVJ. Selain itu, skripsi yang saya buat bisa memberikan sumber literasi, referensi, dan wawasan baru untuk perkembangan ilmu pengetahuan. Penyelesaian skripsi ini tentunya melibatkan banyak pihak baik secara moril maupun materil. Saya mengucapkan terima kasih kepada :

1. Kedua orang tua saya yang selalu mendoakan, mendukung dengan tulus ikhlas demi masa depan setiap anaknya.
2. Keluarga besar yang selalu memberikan doa dan dukungan .
3. Prof. Dr. Ir. Supriyanto selaku Dekan Fakultas Ilmu Komputer.
4. Bapak Dr. Tjahjanto, S.Kom., M.M. dan Ibu Ika Nurlaili, S.Kom,M.Sc. selaku dosen pembimbing yang selalu suportif selama periode pengerjaan skripsi.
5. Bapak Muhammad Adrezo, S.Kom.,M.Sc selaku dosen pembimbing akademik yang membantu dari awal perkuliahan hingga saat ini.
6. Seluruh civitas akademika di lingkungan FIK UPNVJ.
7. Rekan-rekan LDF Ar-Risalah FIK UPNVJ, LDK Uswah UPNVJ, FORMASIKIP UPNVJ, UVFC UPNVJ, dan MPM UPNVJ yang telah menjadi rumah untuk tumbuh dan berkembang ,sebuah keberuntungan bisa mengenal kalian.
8. Rekan-rekan grup *mentoring* Uswah yang selalu membuat saya bersyukur karena dikelilingi orang-orang seperti kalian.
9. Kawan-kawan FIK 2020 UPNVJ yang sudah bersama-sama selama 4 tahun ini.
10. Kawan-kawan MRPBS yang selalu meluangkan waktunya di malam minggu untuk
11. sekedar bertemu dan menghabiskan waktu bersama.
12. Kawan-kawan SMA yang masih meluangkan waktunya untuk terus bertemu dan bermain.

Daftar Isi

KLASTERISASI LINGKUNGAN PERUMAHAN DI JAKARTA SELATAN MENGUNAKAN <i>MACHINE LEARNING</i>	i
PERNYATAAN ORISINALITAS.....	ii
PERNYATAAN PERSETUJUAN PUBLIKASI.....	iii
LEMBAR PENGESAHAN	iv
ABSTRAK.....	v
ABSTRACT.....	vi
Kata Pengantar	vii
Daftar Isi	viii
Daftar Gambar	xi
Daftar Tabel	xii
Daftar Lampiran.....	xiii
BAB I.....	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Permasalahan.....	2
1.3 Ruang Lingkup.....	2
1.4 Tujuan dan Manfaat Penelitian	3
1.5 Luaran yang Diharapkan	3
1.6 Sistematika Penulisan.....	4
BAB II.....	5
LANDASAN TEORI.....	5
2.1 DIKW	5
2.2 Data Driven Decision Making.....	5
2.3 Data Wragling	5
2.4 <i>Normalization</i>	6
2.5 <i>Skewness</i>	6
2.6 <i>Correlation</i>	6

2.7 <i>Exploratory Data Analysis (EDA)</i>	7
2.8 <i>Model selection</i>	7
2.9 <i>Model Evaluation</i>	7
2.11 Pembelajaran Mesin (<i>Machine Learning</i>):.....	7
2.12 Google Colaboratory	7
2.13 Library Python.....	8
2.15 Unsupervised Learning	8
2.16 K-means <i>Clustering</i>	8
2.17 Metode Elbow	10
2.18 Hierarical Clustering	11
2.19 Silhouette Score.....	12
2.20 Adjusted Rand Index (ARI)	12
2.20 V-Measure.....	13
2.21 Lingkungan Perumahan.....	13
BAB III	14
METODE PENELITIAN	14
3.1 Kerangka Berpikir	14
3.1.1 Menentukan topik	15
3.1.2 Identifikasi masalah.....	15
3.1.3 Studi Pustaka	15
3.1.4 Pengumpulan Data.....	15
3.1.5 Praproses Data	16
3.1.6 Metode K-Means	16
3.1.7 Metode Elbow.....	16
3.1.8 Visualisasi Data	17
3.2 Perangkat Penelitian	17
3.3 Jadwal Penelitian.....	19
BAB IV	20
HASIL DAN PEMBAHASAN	20
4.1 Pengumpulan Data	20

4.2 Penilaian Data	27
4.3 Pembersihan Data.....	29
4.4 Normalisasi Data.....	30
4.5 Mengukur Korelasi antar variabel.....	31
4.6 Melakukan Exploratory Data Analysis	33
4.7 K-Means	35
4.8 Metode Elbow	37
4.9 Hasil Klasterisasi.....	38
4.10 Analisis Pola Klaster 0	39
4.11 Analisis Pola Klaster 1	42
4.12 Analisis Pola Klaster 2	45
4.13 Analisis Pola Klaster 3	49
4.14 Urutan Klaster Terbaik hingga Terendah.....	52
4.15 Visualisasi Peta Jakarta Selatan	56
4.16 Hierarical Clustering	58
4.17 Membandingkan hasil Klaster K - Means dan Hierarical Clustering	61
4.18 Membuat Dashboard Menggunakan Streamlit.....	62
BAB V	64
KESIMPULAN DAN SARAN.....	64
5.1 Kesimpulan.....	64
5.2 Saran.....	64
Daftar Pustaka.....	65
LAMPIRAN.....	66

Daftar Gambar

Gambar 2.1 Data Wragling	6
Gambar 2.2 Proses Awal Kmeans	9
Gambar 2.3 Proses K Means.....	9
Gambar 2.4 Proses K Means.....	9
Gambar 2.5 Proses K Means.....	10
Gambar 2.6 Proses K Means.....	10
Gambar 2.7 Proses K Means.....	10
Gambar 2.8 Metode Elbow	11
Gambar 3.1 Kerangka Berpikir	14
Gambar 4.1 Web Satu Data	20
Gambar 4.2 dataset bank sampah sebelum dimodifikasi	21
Gambar 4.3 dataset akhir setelah dilakukan penggabungan dan modifikasi	25
Gambar 4.4 dataset akhir setelah dilakukan penggabungan dan modifikasi	27
Gambar 4.5 Mencari missing value pada dataset.....	28
Gambar 4.6 Mencari jumlah duplikasi pada dataset	28
Gambar 4.7 Melihat deskripsi deskriptif pada dataset.....	29
Gambar 4.8 dataset akhir setelah dilakukan penggabungan dan modifikasi	30
Gambar 4.9 Melakukan normalisasi pada dataset.....	31
Gambar 4.10 Matriks korelasi antar variabel.....	31
Gambar 4.11 Jumlah Bank Sampah Terbanyak pada Kelurahan Jagakarsa	35
Gambar 4.12 Jumlah Bank Sampah Tersedikit pada Kelurahan Pejaten Barat	35
Gambar 4.13 Potongan kode program Algoritma K Means	36
Gambar 4.14 Potongan Kode Program Metode Elbow.....	37
Gambar 4.15 Diagram Batang jumlah setiap kluster	38
Gambar 4.16 Visualisasi Kluster Peta.....	57
Gambar 4.17 Delete Kolom hasil kluster K means.....	58
Gambar 4.18 Impor fungsi dan jarak euclidean.....	58
Gambar 4.19 Menghitung matriks Linkage	59
Gambar 4.20 Visualisasi Dendogram	59
Gambar 4.21 Menentukan jumlah kluster	60
Gambar 4.22 Impor fungsi dan ARI Score	61
Gambar 4.23 V Measuer	61
Gambar 4.24 Cetak ARI dan V-Measure.....	62

Daftar Tabel

Tabel 3.1 Jadwal Penelitian	14
Tabel 4.1 Jumlah Bank Sampah setelah dimodifikasi	24
Tabel 4.2 Nama-nama kelurahan yang masuk kedalam klaster 0	39
Tabel 4.3 Tabel aspek kesehatan pada klaster 0	40
Tabel 4.4 Tabel aspek pendidikan pada klaster 0	40
Tabel 4.5 Tabel aspek fasilitas publik pada klaster 0	41
Tabel 4.6 Tabel aspek demografi penduduk pada klaster 0	41
Tabel 4.7 Tabel kejadian bencana banjir pada klaster 0	42
Tabel 4.8 Tabel nama-nama kelurahan pada klaster 1	42
Tabel 4.9 Tabel aspek kesehatan pada klaster 1	43
Tabel 4.10 Tabel aspek pendidikan pada klaster 1	43
Tabel 4.11 Tabel aspek fasilitas publik pada klaster 1	44
Tabel 4.12 Tabel aspek demografi penduduk pada klaster 1	44
Tabel 4.13 Tabel kejadian bencana banjir pada klaster 1	45
Tabel 4.14 Tabel kelurahan pada klaster 2	46
Tabel 4.15 Tabel aspek layanan kesehatan pada klaster 2	46
Tabel 4.16 Tabel aspek pendidikan pada klaster 2	47
Tabel 4.17 Tabel aspek fasilitas publik pada klaster 2	47
Tabel 4.18 Tabel aspek demografi penduduk pada klaster 1	48
Tabel 4.19 Tabel kejadian banjir pada klaster 2	48
Tabel 4.20 Tabel kelurahan yang termasuk klaster 3	49
Tabel 4.21 Tabel aspek layanan kesehatan pada klaster 3	50
Tabel 4.22 Tabel aspek pendidikan pada klaster 3	50
Tabel 4.23 Tabel aspek fasilitas publik pada klaster 3	51
Tabel 4.24 Tabel aspek demografi penduduk pada klaster 3	51
Tabel 4.25 Tabel kejadian banjir pada klaster 3	52
Tabel 4.26 Tabel perangkian klaster	53
Tabel 4.27 warna keterangan pada tabel 4.26	53

Daftar Lampiran

Lampiran 1	66
Lampiran 2	68
Lampiran 3	70
Lampiran 4	72
Lampiran 5	74
Lampiran 6	78
Lampiran 7	83
Lampiran 8	89
Lampiran 9	93
Lampiran 10	99