



**PENERAPAN ALGORITMA K-MEANS CLUSTERING  
MENGUNAKAN METODE ELBOW UNTUK ANALISA  
POLUSI UDARA DI KOTA YOGYAKARTA BERDASARKAN  
PARAMETER INDEKS STANDAR PENCEMAR UDARA  
(ISPU) PERIODE 2021**

**LAPORAN SKRIPSI**

**Muhammad Rizaldy Ekaputra**

**41518120119**

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

**2023**



**PENERAPAN ALGORITMA K-MEANS CLUSTERING  
MENGUNAKAN METODE ELBOW UNTUK ANALISA  
POLUSI UDARA DI KOTA YOGYAKARTA BERDASARKAN  
PARAMETER INDEKS STANDAR PENCEMAR UDARA  
(ISPU) PERIODE 2021**

**LAPORAN SKRIPSI**

**Muhammad Rizaldy Ekaputra**

**41518120119**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA JAKARTA  
2023**

## PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Rizaldy Ekaputra  
NIM : 41518120119  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma K-Means Clustering Menggunakan Metode Elbow untuk Analisa Polusi Udara di Kota Yogyakarta berdasarkan Parameter Indeks Standar Pencemar Udara (ISPU) Periode 2021

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 10 Juni 2023



Muhammad Rizaldy Ekaputra

## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Rizaldy Ekaputra  
NIM : 41518120119  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma K-Means Clustering Menggunakan Metode Elbow untuk Analisa Polusi Udara di Kota Yogyakarta berdasarkan Parameter Indeks Standar Pencemar Udara (ISPU) Periode 2021

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata I pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

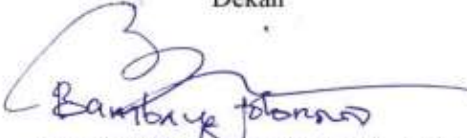
Disahkan oleh:

Pembimbing : Muhammad Rifqi, S.Kom, M.Kom (  )  
NIDN : 0301067101  
Ketua Penguji : Saruni Dwiasnati, ST, MM, M.Kom (  )  
NIDN : 0325128802  
Penguji I : Dr. Nungky Awang Chandra, MTI (  )  
NIDN : 0306117303

Jakarta, 15 Februari 2023

Mengetahui,

Dekan

  
Dr. Bambang Jekonowo, S.Si., MTI

Kepala Program Studi

  
Dr. Bagus Priambodo, ST., MTI

## KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof. Dr. Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana
2. Dr. Bambang Jokonowo, S.Si, M.Ti selaku Dekan Fakultas Ilmu Komputer
3. Bagus Priambodo, ST, M.Ti selaku Ketua Program Studi Teknik Informatika
4. Muhammad Rifqi, S.Kom, M.Kom selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.
5. Saruni Dwiasnati, ST, MM, M.Kom selaku Ketua Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
6. Dr. Nungky Awang Chandra selaku Dosen Penguji Tugas Akhir atas koreksi dan arahan serta masukannya.
7. Pihak lainnya yang telah memberikan dukungan sehingga terselesaikannya laporan tugas akhir ini.

Akhir kata, saya berharap Allah Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 10 Juni 2023

Muhammad Rizaldy Ekaputra

## PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Muhammad Rizaldy Ekaputra  
NIM : 41518120119  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma K-Means Clustering Menggunakan Metode Elbow untuk Analisa Polusi Udara di Kota Yogyakarta berdasarkan Parameter Indeks Standar Pencemar Udara (ISPU) Periode 2021

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan Laporan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 10 Juni 2023

Yang menyatakan,



Muhammad Rizaldy Ekaputra

## ABSTRAK

Nama : Muhammad Rizaldy Ekaputra  
NIM : 41518120119  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Penerapan Algoritma K-Means Clustering Menggunakan Metode Elbow untuk Analisa Polusi Udara di Kota Yogyakarta berdasarkan Parameter Indeks Standar Pencemar Udara (ISPU) Periode 2021  
Pembimbing : Muhammad Rifqi, S.Kom, M.Kom

Lingkungan yang sehat memiliki pengaruh terhadap kesehatan fisik makhluk hidup. Faktor krusial penunjang lingkungan yang sehat ialah kualitas udara yang memenuhi baku kesehatan. Indeks Standar Pencemar Udara yang selanjutnya disingkat ISPU adalah angka yang tidak mempunyai satuan yang menggambarkan kondisi mutu udara ambien di lokasi tertentu, yang didasarkan kepada dampak terhadap kesehatan manusia dan makhluk hidup lainnya. ISPU meliputi parameter seperti partikulat (PM10), partikulat (PM2.5), karbon monoksida (CO), nitrogen dioksida (NO<sub>2</sub>), sulfur dioksida (SO<sub>2</sub>) dan ozon (O<sub>3</sub>). Berdasarkan parameter ISPU tersebut, akan dilakukan pengelompokan (*clustering*) untuk menggali atau menemukan informasi yang belum diketahui sebelumnya. Algoritma K-means banyak digunakan dalam proses pengelompokan data karena memungkinkan menemukan pola dan korelasi dalam data dengan cara yang tidak diawasi. Pengelompokan yang dilakukan terbagi menjadi 6 kelompok (*cluster*). *Cluster 2* merupakan *cluster* dengan polusi tertinggi yang memiliki kadar PM<sub>2.5</sub> dan karbon monoksida (CO) dengan nilai rata-rata tertinggi. Disusul oleh *cluster 4* yang berkategori sedang sebesar 10,68% dengan kadar PM<sub>2.5</sub> yang cukup tinggi, namun *cluster 4* juga memiliki persentase kategori baik terbesar sebesar 32,87%. *Cluster 0, 3 dan 5* sama-sama memiliki kategori sedang di bawah 5% sedangkan *cluster 1* tidak memiliki kategori sedang.

**Kata Kunci** : Kualitas Udara, Indeks Standar Pencemar Udara, *Clustering*, K-Means, Metode Elbow

## ABSTRACT

Name : Muhammad Rizaldy Ekaputra  
NIM : 41518120119  
Study Program : Informatics Engineering  
Title Thesis : Implementation of K-Means Clustering Algorithm Using Elbow Method for Analysis of Air Pollution in Yogyakarta based on Pollutant Standards Index in 2021  
Counsellor : Muhammad Rifqi, S.Kom, M.Kom

*A healthy environment has an influence on the physical health of living things. The crucial factor supporting a healthy environment is air quality that meets health standards. The Air Pollution Standard Index, hereinafter abbreviated as ISPU, is a number that does not have a unit that describes the condition of ambient air quality in a certain location, which is based on the impact on human health and other living things. ISPU includes parameters such as particulates (PM10), particulates (PM2.5), carbon monoxide (CO), nitrogen dioxide (NO2), sulfur dioxide (SO2) and ozone (O3). Based on the ISPU parameters, clustering will be carried out to explore or find information that was not previously known. The K-means algorithm is widely used in the process of grouping data because it allows finding patterns and correlations in data in an unsupervised way. The grouping is divided into 6 groups (clusters). Cluster 2 is the cluster with the highest pollution having PM2.5 and carbon monoxide (CO) levels with the highest average value. Followed by cluster 4 which is in the moderate category at 10.68% with quite high levels of PM2.5, but cluster 4 also has the largest percentage of the good category at 32.87%. Clusters 0, 3 and 5 both have a moderate category below 5% while cluster 1 does not have a moderate category.*

**Keywords :** *Air Quality, Pollutant Standards Index, Clustering, K-Means, Elbow Method*



## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>I</b>
<b>PERNYATAAN KARYA SENDIRI .....</b>	<b>II</b>
<b>HALAMAN PENGESAHAN.....</b>	<b>III</b>
<b>KATA PENGANTAR.....</b>	<b>IV</b>
<b>PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS.....</b>	<b>V</b>
<b>ABSTRAK VI</b>	
<b>ABSTRACT .....</b>	<b>VII</b>
<b>DAFTAR ISI.....</b>	<b>VIII</b>
<b>DAFTAR TABEL .....</b>	<b>X</b>
<b>DAFTAR GAMBAR.....</b>	<b>XI</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah .....	2
1.3 Tujuan Penelitian .....	3
1.4 Manfaat Penelitian .....	3
1.5 Batasan Penelitian.....	3
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>4</b>
2.1 Penelitian Terdahulu .....	4
2.2 Teori Pendukung.....	13
2.2.1 <i>Clustering</i> .....	13
2.2.2 K-Means .....	14
2.2.3 Metode Elbow.....	14
2.2.4 Indeks Standar Pencemar Udara (ISPU).....	15
<b>BAB III METODE PENELITIAN .....</b>	<b>17</b>
3.1 Jenis Penelitian.....	17
3.2 Tahapan Penelitian.....	17
3.2.1 Pengumpulan data.....	17
3.2.2 <i>Data pre-processing</i> .....	18
3.2.3 Penerapan K-Means.....	18
3.2.4 Visualisasi Data .....	19
3.2.5 Analisa .....	19
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>20</b>
4.1 <i>Dataset</i> .....	20
4.2 Pre-Processing.....	20
4.2.1 Menghapus Fitur .....	20
4.2.2 <i>Missing Values</i> .....	21
4.2.3 <i>Resampling</i> .....	21
4.2.4 Normalisasi .....	22
4.3 Pembuatan Model .....	23
4.3.1 Menentukan Nilai K .....	23
4.3.2 Algoritma K-Means .....	23
4.4 Visualisasi Data .....	25
4.5 Analisis Hasil.....	27
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>31</b>
5.1 Kesimpulan .....	31

5.2 Saran .....	31
<b>DAFTAR PUSTAKA .....</b>	<b>32</b>
<b>LAMPIRAN.....</b>	<b>34</b>
Lampiran Bimbingan.....	34
Lampiran Luaran Tugas Akhir .....	35
Lampiran Bukti Submit Jurnal .....	36
Lampiran Naskah Artikel Jurnal .....	37
Lampiran Curriculum Vitae .....	46
Lampiran Surat Pernyataan HKI .....	48



UNIVERSITAS  
**MERCU BUANA**

## DAFTAR TABEL

Tabel 2.1	Daftar penelitian terkait .....	4
Tabel 2.2	Kategori rentang angka ISPU .....	15
Tabel 2.3	Contoh hasil pengukuran dan perhitungan ISPU .....	15
Tabel 4.1	<i>Dataset</i> pengukuran polusi udara .....	20
Tabel 4.2	Jumlah nilai yang hilang ( <i>missing value</i> ) .....	21
Tabel 4.3	Contoh hasil pengisian nilai yang hilang ( <i>missing value</i> ) .....	21
Tabel 4.4	Hasil <i>resampling</i> data .....	22
Tabel 4.5	Hasil normalisasi data .....	22
Tabel 4.6	<i>Centroid</i> awal .....	23
Tabel 4.7	Perhitungan Euclidean distance .....	24
Tabel 4.8	<i>Centroid</i> pada iterasi ke-6 .....	24
Tabel 4.9	Hasil pemberian kategori ISPU .....	27
Tabel 4.10	Kadar tiap-tiap parameter pada <i>cluster</i> 0 .....	28
Tabel 4.11	Kadar tiap-tiap parameter pada <i>cluster</i> 1 .....	28
Tabel 4.12	Kadar tiap-tiap parameter pada <i>cluster</i> 2 .....	29
Tabel 4.13	Kadar tiap-tiap parameter pada <i>cluster</i> 3 .....	29
Tabel 4.14	Kadar tiap-tiap parameter pada <i>cluster</i> 4 .....	29
Tabel 4.15	Kadar tiap-tiap parameter pada <i>cluster</i> 5 .....	30



## DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian.....	17
Gambar 4.1 Nilai SSE pada 2 hingga 20 <i>cluster</i> .....	23
Gambar 4.2 Jumlah data pada masing-masing <i>cluster</i> .....	25
Gambar 4.3 <i>Timeline</i> dari seluruh <i>cluster</i> .....	25
Gambar 4.4 Kadar masing-masing parameter pada setiap <i>cluster</i> .....	26
Gambar 4.5 <i>Timeline</i> setiap <i>cluster</i> berdasarkan kategori.....	28

