



**PENERAPAN ALGORITMA *CLUSTERING K-MEANS* PADA  
PENGELOMPOKAN DAERAH TEMPAT SAMPAH DI  
LINGKUNGAN KAMPUS PONDOK LABU UPN VETERAN  
JAKARTA**

**SKRIPSI**

**ALI HASYIMI**

**1610511044**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN  
JAKARTA FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI S-1 INFORMATIKA**

**2021**



**PENERAPAN ALGORITMA *CLUSTERING K-MEANS* PADA  
PENGELOMPOKAN DAERAH TEMPAT SAMPAH DI  
LINGKUNGAN KAMPUS PONDOK LABU UPN VETERAN  
JAKARTA**

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat Untuk Kelulusan Mata  
Kuliah Tugas Akhir**

**ALI HASYIMI**

**1610511044**

**UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN  
JAKARTA FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI S-1 INFORMATIKA**

**2021**

## LEMBAR PENGESAHAN

### LEMBAR PENGESAHAN

Dengan ini dinyatakan bahwa Tugas Akhir berikut:

Nama : Ali Hasyimi  
 NIM : 1610511044  
 Program Studi : Informatika  
 Judul : Penerapan Algoritma Clustering K-Means Pada Pengelompokan Daerah Tempat Sampah Di Lingkungan Kampus Pondok Labu UPN Veteran Jakarta

Telah berhasil dipertahankan dihadapan Tim Penguji dan diterima sebagai bagian dari persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika S.1, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.

Mengetahui,



Yuni Widiastiw, S.Kom., M.SI.

Penguji 1



Bayu Hananto, S.Kom., M.Kom.

Penguji 2



Dr. Ermatita, M.kom.

Pembimbing 1



Ika Nurlaili, S.Kom, M.Sc.

Pembimbing 2



Dr. Ermatita, M.kom.

Dekan



Yuni Widiastiw, S.Kom., M.SI.

Ketua Program Studi

Ditetapkan di : Jakarta

Tanggal Ujian : 29 Januari 2021



## PERNYATAAN ORISINALITAS

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

Nama : Ali Hasyimi  
NIM : 1610511044  
Tanggal : 15 Februari 2021

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

Jakarta, 17 Mei 2019

Yang Menyatakan,



(Ali Hasyimi)

## PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

---

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Ali Hasyimi  
NIM : 1610511044  
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:

### **Penerapan Algoritma *Clustering K-Means* Pada Pengelompokan Daerah Tempat Sampah Di Lingkungan Kampus Pondok Labu UPN Veteran Jakarta**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih media/formatkan, mengelola dalam bentuk pangkalan data (*database*), 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 : 15 Februari 2021

Yang Menyatakan,



(Ali Hasyimi)

# **Penerapan Algoritma *Clustering K-Means* Pada Pengelompokan Daerah Tempat Sampah Di Lingkungan Kampus Pondok Labu UPN Veteran Jakarta**

Ali Hasyimi

## **Abstrak**

Lingkungan yang bersih merupakan salah satu kebutuhan masyarakat dalam kehidupan. Upaya dalam mewujudkan lingkungan bersih terus dilakukan beriringan dengan berbagai metode – metode baru yang diterapkan dalam sistem pengolahan lingkungan dimasyarakat. Metode dan inovasi yang diterapkan dimasyarakat salah satunya adalah penggunaan *Data Mining* dan *Machine Learning* dalam mengolah limbah padat guna menciptakan sistem pengolahan limbah padat yang lebih efektif dan efisien. Penerapan teknologi yang fleksibel namun efisien ini membuat peningkatan kinerja serta mudahnya perolehan informasi dari berbagai macam aktivitas dan dari berbagai macam penyelesaian permasalahan. Hal ini membuat penulis berinisiatif melakukan penelitian terkait implementasi *Data Mining* dan *Machine Learning* terhadap persoalan lingkungan, khususnya terhadap pengolahan limbah padat. Penelitian ini menerapkan metode *Data Mining* yakni *Cluster Analysis* dengan algoritma *K-Means* untuk mengelola data persebaran tempat sampah pada lingkungan UPN Veteran Jakarta yang menghasilkan informasi dari *Dataset* dan visualisasi persebaran tempat sampah.

Kata Kunci : *Cluster Analysis, Data Mining, K-Means, Python.*

# **Implementation of the K-Means Clustering Algorithm in the Clusterization of Garbage Bins in the Pondok Labu Campus, UPN Veteran Jakarta**

Ali Hasyimi

## **Abstract**

*A clean environment is one of the needs of society in life. Efforts to create a clean environment continue to be carried out in tandem with various new methods applied in the environmental treatment system in the community. One of the methods and innovations applied in the community is the use of Data Mining and Machine Learning in processing solid waste in order to create a more effective and efficient solid waste treatment system. The flexible but efficient application of technology makes performance improvements and easier to obtain information from various activities and from various kinds of problem solving. This has led the author to take the initiative to conduct research related to the implementation of Data Mining and Machine Learning on environmental issues, especially on solid waste processing. This study applies the Data Mining method, namely Cluster Analysis with the K-Means algorithm to manage data on the distribution of trash bins in the UPN Veteran Jakarta environment which produces information from the dataset and visualizes the distribution of trash cans.*

**Keywords:** *Cluster Analysis, Data Mining, K-Means, Python.*

## KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada Allah SWT atas segala karunia-Nya, sehingga Skripsi ini berhasil diselesaikan. Penulis ingin mengucapkan terima kasih kepada:

1. Ibu Dr. Ermatita, M. Kom. Selaku Pembimbing 1 dan Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta.
2. Ika Nurlaili, S.Kom, M.Sc. Selaku Pembimbing 2 dan Pembimbing Akademis Universitas Pembangunan Nasional Veteran Jakarta.
3. Yuni Widiastiwi, S.Kom., M.SI. selaku Ketua Program Studi Informatika Universitas Pembangunan Nasional Veteran Jakarta.
4. Bapak/Ibu dosen Informatika Universitas Pembangunan Nasional Veteran Jakarta terima kasih atas ilmu-ilmu yang bermanfaat.
5. Kedua orang tua dan keluarga yang selalu memberikan dorongan dan doa kepada penulis sehingga dapat menyelesaikan penulisan skripsi ini.
6. Teman-teman Informatika 2016, 2015, 2017, rekan-rekan KSM Robotika, Group Apa Aja, Rekan Bimbinga Skripsi, Rekan KASMA, Medan Generasi Impian dan seluruh rekan yang tidak dapat penulis sebutkan namanya satu persatu, terima kasih atas saran dan dukungan yang telah diberikan.

Akhir kata, semoga skripsi ini dapat bermanfaat bagi para pembacanya.

Jakarta, 15 Februari 2021

Penulis



Ali Hasyimi



## DAFTAR ISI

<b>LEMBAR PENGESAHAN</b> .....	ii
<b>PERNYATAAN ORISINALITAS</b> .....	iv
<b>PERNYATAAN PERSETUJUAN PUBLIKASI</b> .....	v
<b>Abstrak</b> .....	vi
<b>Abstract</b> .....	vii
<b>KATA PENGANTAR</b> .....	viii
<b>DAFTAR ISI</b> .....	ix
<b>DAFTAR GAMBAR</b> .....	xi
<b>DAFTAR TABEL</b> .....	xii
<b>BAB I</b> .....	1
<b>1.1 Latar Belakang</b> .....	1
<b>1.2 Identifikasi Masalah</b> .....	3
<b>1.3 Rumusan Masalah</b> .....	3
<b>1.4 Batasan Masalah</b> .....	3
<b>1.5 Tujuan Penelitian</b> .....	4
<b>1.6 Manfaat Penelitian</b> .....	4
1.6.1 Manfaat Akademis .....	4
1.6.2 Manfaat Praktis .....	4
<b>1.8 Luaran Yang Diharapkan</b> .....	4
<b>1.9 Sistematika Penulisan</b> .....	4
<b>BAB II</b> .....	6
<b>2.1 Data Mining</b> .....	6
<b>2.3 Cluster Analysis</b> .....	7
<b>2.4 Algoritma K-Means</b> .....	7
2.3.1 <i>Euclidean Distance</i> .....	8
2.3.2. <i>K – Means ++</i> .....	9
2.3.3 <i>Elbow Plot</i> .....	10
<b>2.5 Python</b> .....	12
<b>2.6 Penelitian Terkait</b> .....	13
<b>BAB III</b> .....	20
<b>3.1 Tahapan Penelitian</b> .....	20
3.1.1 Identifikasi Masalah .....	20
3.1.2 Pra-Proses Data Keseluruhan .....	21

3.1.3 Pembangunan Program .....	23
3.1.4 Analisis Klaster .....	24
3.1.5 Dokumentasi .....	24
3.1.6 Lokasi dan Jadwal Penelitian .....	24
<b>3.2 Jadwal Penelitian .....</b>	<b>25</b>
<b>3.3 Alat dan Bahan Penelitian .....</b>	<b>25</b>
3.3.1 Perangkat Keras : .....	25
3.3.2 Perangkat Lunak : .....	25
<b>BAB IV .....</b>	<b>26</b>
<b>4.1 Pra-Proses Data Keseluruhan .....</b>	<b>27</b>
<b>4.2 Pembentukan Program .....</b>	<b>31</b>
4.2.1 Implementasi pada Data Pembangunan .....	31
4.2.2 Implementasi Pada Data Uji .....	39
<b>4.3 Pengambilan Kesimpulan .....</b>	<b>45</b>
<b>BAB V .....</b>	<b>55</b>
<b>5.1 Kesimpulan .....</b>	<b>55</b>
<b>5.2 Saran .....</b>	<b>55</b>
<b>DAFTAR PUSTAKA .....</b>	<b>56</b>
<b>RIWAYAT HIDUP .....</b>	<b>58</b>
<b>LAMPIRAN .....</b>	<b>59</b>

## DAFTAR GAMBAR

Gambar 1. 1 Visualisasi Jarak Antar Data .....	8
Gambar 2. 1 Tahapan Penelitian .....	20
Gambar 2. 2 Dataset Keseluruhan dari Google Maps .....	21
Gambar 3. 1 Tempat Sampah Kategori Besar .....	27
Gambar 3. 2 Tempat Sampah Kategori Sedang .....	28
Gambar 3. 3 Tempat Sampah Kategori Kecil .....	29
Gambar 3. 4 Visualisasi Dataset Pembangunan .....	31
Gambar 3. 5 Visualisasi Dataset Cluster Awal .....	34
Gambar 3. 6 Visualisasi Elbow Plot Dataset Pembangunan.....	36
Gambar 3. 7 Visualisasi Cluster Akhir.....	38
Gambar 3. 8 Visualisasi Dataset Utama Mentah.....	39
Gambar 3. 9 Visualisasi Dataset Utama Cluster Akhir .....	44
Gambar 3. 10 Grafik Kapasitas Total Tiap Cluster, visualisasi klaster akhir, dan prakiraan lokasi dari Google Maps.....	53

**DAFTAR TABEL**

Tabel 1. 1 Penelitian Terkait .....	13
Tabel 2. 1 Jadwal Penelitian .....	25
Tabel 3. 1 Dataset Pembangunan .....	29
Tabel 3. 2 Dataset Pembangunan dengan Cluster .....	32
Tabel 3. 3 Nilai WCSS Elbow Plot.....	35
Tabel 3. 4 Dataset Pembangunan dengan Cluster Akhir .....	37
Tabel 3. 5 Dataset Utama dengan Cluster .....	40
Tabel 3. 6 Tambahan Data dari Dataset Keseluruhan ke Dataset Utama .....	45
Tabel 3. 7 Kesimpulan Cluster 0 .....	49
Tabel 3. 8 Kesimpulan Cluster 1 .....	50
Tabel 3. 9 Kesimpulan Cluster 2 .....	50
Tabel 3. 10 Kesimpulan Cluster 3 .....	51
Tabel 3. 11 Kesimpulan Cluster 4 .....	51
Tabel 3. 12 Kesimpulan Cluster 5 .....	52