



**KLASTERISASI DATA PENERIMAAN IMUNISASI MENGGUNAKAN K-
MEANS PADA KECAMATAN DI PROVINSI DKI JAKARTA**

SKRIPSI

NABIL MUHAMMAD RAIHAN

NIM. 2010512012

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
2024**



**KLASTERISASI DATA PENERIMAAN IMUNISASI MENGGUNAKAN K-
MEANS PADA KECAMATAN DI PROVINSI DKI JAKARTA**

SKRIPSI

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

NABIL MUHAMMAD RAIHAN

NIM. 2010512012

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
2024**

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Nabil Muhammad Raihan

NIM : 2010512012

Judul : Klasterisasi Data Penerimaan Imunisasi Menggunakan K-Means Pada Kecamatan di Provinsi DKI Jakarta

Tanggal : 18 Juni 2024

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

Bekasi, 18 Juni 2024

Yang menyatakan,



Nabil Muhammad Raihan

PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

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

Nama : Nabil Muhammad Raihan

NIM : 2010512012

Fakultas : Ilmu Komputer

Program Studi : S1 Sistem Informasi

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

Klasterisasi Data Penerimaan Imunisasi Menggunakan K-Means Pada Kecamatan di Provinsi DKI Jakarta

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/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 sesungguhnya.

Dibuat di: Bekasi
Pada tanggal: 18 Juni 2024

Yang menyatakan,



Nabil Muhammad Raihan

LEMBAR PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Nabil Muhammad Raihan


NIM : 2010512012

Program Studi : S1 Sistem Informasi

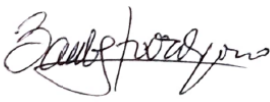
Judul Tugas Akhir : Klasterisasi Data Penerimaan Imunisasi Menggunakan K-Means Pada Kecamatan di Provinsi DKI Jakarta

Telah Berhasil dipertahankan di hadapan Tim Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana pada Program Studi S1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta.



Dr. Widya Cholil, S.Kom., M.I.T.
Dosen Pembimbing I


Muhammad Adrezo, S.Kom., M.Sc.
Dosen Pembimbing II


Dr. Bambang Saras Yulistiawan, S.T., M.Kom.
Dosen Penguji I


Bambang Triwahyono, S.Kom., M.Si.
Dosen Penguji II


Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM.
Dekan FIK UPN Veteran Jakarta


Anita Muliawati, S.Kom., M.TI
Kepala Program Studi Sistem Informasi

Ditetapkan di : Jakarta
Tanggal Ujian : 22 Mei 2024

KLASTERISASI DATA PENERIMAAN IMUNISASI MENGGUNAKAN K-MEANS PADA KECAMATAN DI PROVINSI DKI JAKARTA

ABSTRAK

Imunisasi merupakan suatu aktivitas yang dilakukan untuk menjaga tubuh dari penularan penyakit tertentu melalui pemberian vaksin. Berdasarkan data milik Direktorat Jenderal Pencegahan dan Pengendalian Penyakit (P2P) Kementerian Kesehatan, DKI Jakarta menjadi salah satu provinsi dengan cakupan imunisasi dasar lengkap pada bayi terendah nasional pada tahun 2021. DKI Jakarta menduduki posisi ke-6 dengan persentase sebesar 63,3% dengan rata-rata nasional sebesar 84,4%. Selain itu, dalam laporan statistik sektoral Provinsi DKI Jakarta dapat diketahui bahwa penerimaan imunisasi dasar lengkap belum merata. Oleh karena itu penelitian ini dilakukan untuk klusterisasi pada kecamatan di Provinsi DKI Jakarta berdasarkan data hasil penerimaan imunisasi untuk mempermudah Dinas Kesehatan DKI Jakarta dalam melakukan pemerataan penerimaan imunisasi. Penelitian ini menggunakan data penerimaan imunisasi dasar lengkap di Provinsi DKI Jakarta pada tahun 2022 yang telah didetailkan sampai level kecamatan. Penelitian ini menghasilkan jumlah cluster terbaik sebanyak dua cluster, yaitu cluster 0 (rendah) dan cluster 1 (tinggi). Cluster 0 sebagai cluster dengan karakteristik penerimaan imunisasi rendah mendominasi kecamatan di Provinsi DKI Jakarta dengan beranggotakan sebanyak 34 kecamatan, sedangkan cluster 1 dengan karakteristik lebih tinggi beranggotakan sebanyak 10 kecamatan.

Kata kunci: Imunisasi, klusterisasi, k-means, DKI Jakarta, kecamatan

CLASSTERIZATION OF IMMUNIZATION RECEIPT DATA USING K-MEANS IN DISTRICTS IN DKI JAKARTA PROVINCE

ABSTRACT

Immunization is an activity carried out to protect the body from transmission of certain diseases through administering vaccines. Based on data from the Directorate General of Disease Prevention and Control (P2P) of the Ministry of Health, DKI Jakarta is one of the provinces with the lowest national coverage of complete basic immunization for babies in 2021. DKI Jakarta is in 6th position with a percentage of 63.3% with an average -national average of 84.4%. Apart from that, in the sectoral statistical report for DKI Jakarta Province, it can be seen that receipt of complete basic immunization is not evenly distributed. Therefore, this research was carried out for clustering in sub-districts in DKI Jakarta Province based on data on immunization receipt results to make it easier for the DKI Jakarta Health Service to distribute immunization receipts evenly. This research uses data on receipt of complete basic immunization in DKI Jakarta Province in 2022 which has been detailed down to the sub-district level. This research produced the best number of clusters as two clusters, namely cluster 0 (low) and cluster 1 (high). Cluster 0 as a cluster with characteristics of low immunization acceptance dominates the sub-districts in DKI Jakarta Province with 34 sub-district members, while cluster 1 with higher characteristics has 10 sub-district members.

Key words: *Immunization, clustering, k-means, DKI Jakarta, sub-district*

KATA PENGANTAR

Dengan mengucapkan segala puji dan syukur kepada Allah Subhanahu wa Ta'ala yang telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan proposal tugas akhir/skripsi dengan judul “Klasterisasi Data Penerimaan Imunisasi Menggunakan K-Means Pada Kecamatan di Provinsi DKI Jakarta” sebagai syarat mengambil mata kuliah tugas akhir pada Program Studi S1 Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta. Penulis menyadari bahwa laporan akhir ini dapat terselesaikan karena adanya bantuan dari berbagai pihak. Maka penulis ingin menyampaikan terima kasih kepada:

1. Kedua orang tua penulis yang telah memberikan dukungan, doa, ilmu, serta fasilitas terbaik kepada penulis.
2. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM., selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Anita Muliawati, S.Kom., MTI. selaku ketua program studi S1 Sistem Informasi.
4. Ibu Dr. Widya Cholil, S.Kom., M.I.T. selaku dosen pembimbing 1 dan Bapak Muhammad Adrezo, S.Kom., M.Sc. selaku dosen pembimbing 2 yang telah memberikan arahan, saran, serta bimbingan yang telah diberikan kepada penulis sehingga proposal tugas akhir ini dapat terselesaikan sebaik mungkin.
5. Ibu Theresia Wati, S.Kom., MTI. selaku dosen pembimbing akademik.
6. Teman-teman penulis yang membantu menghadapi dan mengatasi masalah bersama selama perkuliahan.

Jakarta, 20 April 2024

Nabil Muhammad Raihan

DAFTAR ISI

PERNYATAAN ORISINALITAS.....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS.....	iv
LEMBAR PENGESAHAN	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	6
1.5 Ruang Lingkup Penelitian.....	6
1.6 Luaran yang Diharapkan	6
1.7 Sistematika Penulisan.....	7
BAB II TINJAUAN PUSTAKA	8
2.1 Klasterisasi	8
2.1.1 Klasterisasi Hierarki	8
2.1.2 Klasterisasi Partisi.....	9
2.2 Data Mining.....	10
2.3 Imunisasi	12
2.4 K-Means	15
2.5 Python.....	17
2.6 Metode Elbow	18
2.7 Penelitian yang Relevan	19
BAB III METODE PENELITIAN.....	25
3.1 Kerangka Berpikir	25
3.2 Identifikasi Masalah	25
3.3 Studi Pustaka	25
3.4 Pengumpulan Data	26
3.5 Pemahaman Data	26
3.6 Praproses Data	26
3.7 Pengolahan Data dengan Algoritma K-Means	27
3.8 Analisis Hasil dan Visualisasi Data	27
3.9 Alat Bantu Penelitian.....	27
3.10 Jadwal Penelitian	28
BAB IV HASIL DAN PEMBAHASAN.....	29
4.1 Pemahaman Data	29
4.2 Pre-processing Data.....	31
4.2.1 Pemilihan Atribut.....	31

4.2.2	Transformasi Data.....	34
4.2.3	Normalisasi Data.....	38
4.3	Pemodelan K-Means Clustering.....	40
4.3.1	Penentuan Jumlah Cluster K.....	40
4.3.2	Kalkulasi Jarak Data Ke Pusat <i>Cluster</i>	41
4.4	Analisis Hasil dan Visualisasi Data	51
BAB V	PENUTUP	54
5.1	Kesimpulan.....	54
5.2	Saran	54
DAFTAR	PUSTAKA	55

DAFTAR GAMBAR

Gambar 1.1 Grafik 10 Provinsi dengan Cakupan Imunisasi Dasar Lengkap pada Bayi Terendah Pada Lingkup Nasional Tahun 2021	2
Gambar 1.2 Grafik Cakupan Imunisasi Dasar Lengkap pada Bayi Usia 0-59 Bulan di Provinsi DKI Jakarta Tahun 2022	3
Gambar 2.1 Klasterisasi Hierarki	9
Gambar 2.2 Tahapan Klasterisasi Objek Menggunakan Metode K-Means	9
Gambar 2.3 Tahap-tahap Algoritma K-Means	16
Gambar 3. 1 Kerangka Berpikir	25
Gambar 4.1 Grafik Elbow	40
Gambar 4.2 Diagram Pencar Hasil Klasterisasi	51
Gambar 4.3 Visualisasi Kecamatan Berdasarkan Cluster	52

DAFTAR TABEL

Tabel 2.1 Tabel penelitian yang relevan	20
Tabel 3.1 Jadwal Penelitian	28
Tabel 4.1 Data Penerimaan Imunisasi Dasar Lengkap Provinsi DKI Jakarta Tahun 2022	29
Tabel 4.2 Keterangan Atribut Data	30
Tabel 4.3 Atribut Data Terpilih	31
Tabel 4.4 Data Hasil Pengurangan Dengan Target	32
Tabel 4.5 Hasil Deteksi Outlier Kolom "LK Selisih"	34
Tabel 4.6 Hasil Deteksi Outliers Kolom "PR Selisih"	35
Tabel 4.7 Data Hasil Transformasi	37
Tabel 4. 8 Tabel Hasil Normalisasi Data	39
Tabel 4. 9 Nilai Centroid Awal	41

Tabel 4.10 Hasil Klasterisasi Iterasi Ke-1	42
Tabel 4.11 Nilai Centroid Klasterisasi Iterasi Ke-1	43
Tabel 4.12 Hasil Klasterisasi Iterasi Ke-2.....	43
Tabel 4.13 Nilai Centroid Klasterisasi Iterasi Ke-2.....	45
Tabel 4.14 Hasil Klasterisasi Iterasi Ke-3.....	45
Tabel 4.15 Nilai Centroid Klasterisasi Iterasi Ke-3.....	46
Tabel 4.16 Hasil Klasterisasi Iterasi Ke-4.....	46
Tabel 4.17 Nilai Centroid Klasterisasi Iterasi Ke-4.....	48
Tabel 4.18 Hasil Klasterisasi Iterasi Ke-5.....	48
Tabel 4.19 Data Final Hasil Klasterisasi.....	49

DAFTAR PERSAMAAN

Persamaan 2.1 Persamaan Euclidian	16
Persamaan 2.2 Persamaan Mencari Centorid Baru	17
Persamaan 2.3 Persamaan Metode Elbow	19
Persamaan 4.1 Persamaan Interquartile Range (IQR).....	34
Persamaan 4.2 Formula Normalisasi Min-Max	38