

**ANALISIS PENGELOMPOKAN PROVINSI DI INDONESIA
BERDASARKAN TINGKAT PENGANGGURAN TERBUKA
TAHUN 2013-2023 MENGGUNAKAN ALGORITMA
*K-MEANS CLUSTERING***

SKRIPSI

Oleh:

Alif Izzuddin Ramadhan

202010225099



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2024**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas akhir : Analisis Pengelompokan Provinsi di Indonesia
: Berdasarkan Tingkat Pengangguran Terbuka
: Tahun 2013-2023 Menggunakan Algoritma K-Means Clustering

Nama Mahasiswa : Alif Izzuddin Ramadhan

Nomor Pokok Mahasiswa : 202010225099

Program Studi/Fakultas : Informatika / Ilmu Komputer

Jakarta, 05/07/2024

MENYETUJUI,

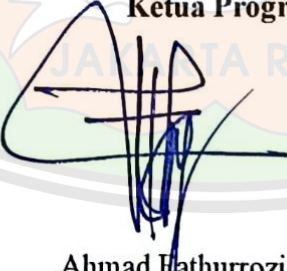
Pembimbing I

Pembimbing II


Prima Dina Atika, S.Kom., M.Kom.
NIDN. 0311037107


Khairunnisa Fadhillah Ramdhan, M.Si.
NIDN. 0328039201

Ketua Program Studi


Ahmad Fathurrozi, S.E., M.M.S.I.
NIP. 2012486

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya

2024

LEMBAR PENGESAHAN

Judul Tugas Akhir : Analisis Pengelompokan Provinsi di Indonesia
Berdasarkan Tingkat Pengangguran Terbuka
Tahun 2013-2023 Menggunakan Algoritma K-
Means Clustering

Nama Mahasiswa : Alif Izzuddin Ramadhan

Nomor Pokok Mahasiswa : 202010225099

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 26 Juni 2024

Jakarta, 05 Juli 2024

MENGESAHKAN,

Ketua Tim Penguji : Dr. Tb. Ai Munandar, S.Kom., MT.

NIDN : 0413098403

Penguji I : Nurfiyah, ST., M.Kom.

NIDN : 0312079403

Penguji II : Prima Dina Atika, S.Kom., M.Kom.

NIDN : 0311037107

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

Fakultas Ilmu Komputer

Ahmad Fathurrozi, S.E., M.M.S.I

NIP. 2012486

Dr. Dra. Tyastuti Sri Lestari, M.M

NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Alif Izzuddin Ramadhan
NPM : 202010225099
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Pengelompokan Provinsi di Indonesia
Berdasarkan Tingkat Pengangguran Terbuka Tahun 2013-
2023 Menggunakan Algoritma K-Means Clustering

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 05 Juli 2024

Penulis



Alif Izzuddin Ramadhan

ABSTRAK

Alif Izzuddin Ramadhan, 202010225099. Analisis Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Pengangguran Terbuka Tahun 2013-2023 Menggunakan Algoritma *K-Means Clustering*. Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2024.

Pengangguran merupakan masalah ekonomi yang kompleks dengan dampak signifikan terhadap kesejahteraan sosial, ketimpangan ekonomi, dan stabilitas politik di Indonesia. Penelitian ini bertujuan untuk menganalisis karakteristik tingkat pengangguran di provinsi-provinsi Indonesia menggunakan data dari Badan Pusat Statistik (BPS) periode 2013-2023 menggunakan algoritma *K-Means clustering* untuk mengelompokkan provinsi berdasarkan Tingkat Pengangguran Terbuka (TPT). Data sekunder dari BPS yang digunakan difokuskan pada TPT selama sepuluh tahun terakhir. Metode Elbow digunakan untuk menentukan jumlah cluster optimal, dan evaluasi dilakukan dengan menggunakan Davies-Bouldin Score, yang menghasilkan dua cluster utama yaitu cluster 0 dengan tingkat pengangguran yang lebih tinggi dari cluster 1 berdasarkan hasil analisis yang dilakukan. Cluster 0 terdiri dari 12 provinsi, sementara cluster 1 terdiri dari 22 provinsi. Evaluasi model menunjukkan skor Davies-Bouldin Index sebesar 0.7 yang menunjukkan kualitas clustering yang baik. Hasil clustering ini divisualisasikan dalam bentuk peta untuk memudahkan interpretasi dan pemahaman karakteristik pengangguran di berbagai provinsi. Penelitian ini diharapkan dapat menjadi alat bantu bagi pembuat kebijakan untuk merancang kebijakan yang lebih efektif dalam mengurangi tingkat pengangguran di Indonesia. Dengan menggunakan algoritma *K-Means clustering*, penelitian ini dapat mengidentifikasi pola-pola tertentu yang mungkin terlewatkan dengan analisis sederhana, memberikan wawasan yang lebih mendalam tentang perbedaan regional dalam hal pengangguran. Melalui hasil clustering yang divisualisasikan dalam bentuk peta, para pembuat kebijakan dapat dengan mudah mengidentifikasi provinsi-provinsi dengan tingkat pengangguran yang tinggi dan merancang intervensi yang lebih efektif serta tepat guna.

Kata kunci : Pengangguran Terbuka, *K-Means*, pemetaan *cluster*

ABSTRACT

Alif Izzuddin Ramadhan, 202010225099. *Clustering Analysis of Provinces In Indonesia Based On Open Unemployment Rate In 2013-2023 Using K-Means Clustering Algorithm.* Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2024.

Unemployment is a complex economic problem with significant impacts on social welfare, economic inequality, and political stability in Indonesia. This study aims to analyze the characteristics of the unemployment rate in Indonesian provinces using data from the Central Bureau of Statistics (BPS) for the period 2013-2023 using the K-Means clustering algorithm to group provinces based on the Open Unemployment Rate (TPT). The secondary data from BPS used focused on the TPT for the last ten years. The Elbow method was used to determine the optimal number of clusters, and evaluation was conducted using the Davies-Bouldin Score, which resulted in two main clusters, namely cluster 0 with a higher unemployment rate than cluster 1 based on the results of the analysis conducted. Cluster 0 consists of 12 provinces, while cluster 1 consists of 22 provinces. The model evaluation shows a Davies-Bouldin Index score of 0.7, which indicates good clustering quality. The clustering results are visualized in the form of maps to facilitate interpretation and understanding of unemployment characteristics in various provinces. This research is expected to be a tool for policy makers to design more effective policies in reducing the unemployment rate in Indonesia. By using the K-Means clustering algorithm, this research can identify certain patterns that may be missed by simple analysis, providing a deeper insight into regional differences in terms of unemployment. Through the clustering results visualized in the form of maps, policymakers can easily identify provinces with high unemployment rates and design more effective and appropriate interventions.

Keywords: *Open Unemployment, K-Means, cluster mapping*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Alif Izzuddin Ramadhan
NPM : 202010225099
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

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

Analisis Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Pengangguran Terbuka Tahun 2013-2023 Menggunakan Algoritma *K-Means Clustering*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 05 Juli 2024
Yang Menyatakan



Alif Izzuddin Ramadhan

KATA PENGANTAR

Puji dan syukur peneliti panjatkan kehadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga peneliti dapat menyelesaikan penyusunan skripsi yang berjudul **“Analisis Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Pengangguran Terbuka Tahun 2013-2023 Menggunakan Alogritma *K-Means Clustering*”**. Skripsi ini dibuat sebagai syarat untuk mencapai gelar Sarjana pada Universitas Bhayangkara Jakarta Raya Program Studi Informatika Fakultas Ilmu Komputer. Penyusunan skripsi ini dapat berjalan dengan baik berkat dukungan dan doa dari banyak pihak. Oleh karena itu peneliti mengucapkan terimakasih kepada Ibu saya, Rahma Dewi Suciati yang telah memberikan dukungan nya sehingga saya mampu menyelesaikan proposal skripsi ini dengan baik dan kepada:

1. Bapak Irjen. Pol. (Purn) Dr., Drs. H. Bambang Karsono, S.H, M.M., Ph.D., D.Crim (Honoriscausa) selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari., M.M. Selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya Bekasi.
3. Bapak Ahmad Fathurozi, S.E., M.M.S.I. Selaku Ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya Bekasi.
4. Ibu Prima Dina Atika, S.Kom., M.Kom. selaku Dosen pembimbing I pada penelitian ini.
5. Ibu Khairunnisa Fadhilla Ramdhania, M.Si, selaku Dosen pembimbing II pada penelitian ini.
6. Bapak Kusdarnowo Hantoro, S.Kom., M.Kom. selaku Dosen pembimbing akademik.
7. Rudiansyah, alm. Ayah saya yang telah mendidik saya hingga menjadi seorang laki-laki yang baik.
8. Shafa dhiya raihanah, adik saya yang telah menjadi adik, keluarga dan teman setia saya sampai sekarang.
9. Miftah Sa'diyyah, pasangan saya yang telah mendukung apapun yang saya lakukan untuk mengembangkan diri saya, dan telah menjadi tempat

saya untuk bercerita.

10. Bram Kahlil Romadhan, Fawwaz Aziz, Nafis Almajid dan Yanuar Ginting yang terus menjadi teman seperjuangan sampai pada tahap saat ini.

Peneliti memahami bahwa masih begitu banyak kekurangan dalam segi penyampaian, penulisan ataupun penelitian, untuk itu peneliti sangat terbuka terhadap kritik ataupun saran yang sifat nya membangun serta dapat meningkatkan skripsi ini jauh lebih baik lagi.

Jakarta, 8 Maret 2024

Peneliti

Alif Izzuddin Ramadhan



DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
<i>ABSTRACT</i>	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	4
1.3 Rumusan Masalah.....	4
1.4 Batasan Masalah	4
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
1.7 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Tinjauan Pustaka.....	7
2.2 Konsep Pengangguran Terbuka	14
2.3 <i>Data Mining</i>	14
2.4 <i>Cross Industry Standard Process for Data Mining (CRISP-DM)</i>	14
2.5 Algoritma	15
2.5.1 <i>Clustering</i>	16
2.5.2 <i>Metode K-Means Clustering</i>	16
2.6 <i>Davies Bouldin Index</i>	18
2.7 Python	19
BAB III METODOLOGI PENELITIAN	20
3.1 Objek Penelitian.....	20

3.2	Kerangka Penelitian	20
3.3	<i>Business Understanding</i>	21
3.4	<i>Data Understanding</i>	21
3.5	<i>Data Preparation</i>	21
3.6	<i>Modelling</i>	22
3.7	<i>Evaluation</i>	22
3.8	<i>Deployment</i>	22
BAB IV HASIL DAN PEMBAHASAN		23
4.1	Hasil Penelitian	23
4.2	<i>Business Understanding</i>	23
4.3	<i>Data Understanding</i>	24
4.4	<i>Data Preparation</i>	26
4.4.1	<i>Data Cleaning</i>	26
4.4.2	<i>Data Reduction</i>	31
4.4.3	<i>Data Integration</i>	33
4.5	<i>Modelling</i>	34
4.5.1	<i>Elbow Method</i>	35
4.5.2	<i>K-Means Modelling</i>	37
4.5.3	<i>Hasil Cluster</i>	51
4.6	<i>Evaluation</i>	58
4.7	Pembahasan	63
BAB V PENUTUP		64
5.1	Kesimpulan	64
5.2	Saran	65
DAFTAR PUSTAKA		66

DAFTAR TABEL

Tabel 2. 1 Tinjauan Penelitian Terdahulu	10
Tabel 4. 1 Detail Kolom Dataset	25
Tabel 4. 2 Hasil <i>Rename</i> Fitur	30
Tabel 4. 3 Penjelasan <i>Script Elbow Method</i>	36
Tabel 4. 4 Menentukan <i>Centroid</i> Awal	37
Tabel 4. 5 Hasil Iterasi Ke-1	40
Tabel 4. 6 <i>Centroid</i> Iterasi Ke-2	41
Tabel 4. 7 Hasil Perhitungan Iterasi Ke-2	44
Tabel 4. 8 <i>Centroid</i> Iterasi Ke-3	45
Tabel 4. 9 Hasil Perhitungan Iterasi Ke-3.....	47
Tabel 4. 10 <i>Centroid</i> Iterasi Ke-4	49
Tabel 4. 11 Hasil Analisis <i>Boxplot</i>	53
Tabel 4. 12 Hasil Akhir Pengelompokan	57
Tabel 4. 13 Jarak Antara <i>Data Point</i> Dan <i>Centroid Cluster</i> nya	59
Tabel 4. 14 Hasil Perhitungan Ssb.....	61
Tabel 4. 15 Hasil Perhitungan Rasio	61
Tabel 4. 16 Hasil Evaluasi Daveis Bouldin Index Tiap K	62

DAFTAR GAMBAR

Gambar 1. 1 Grafis TPT di Indonesia Tahun 2023	2
Gambar 3. 1 Kerangka Penelitian.....	20
Gambar 4. 1 <i>Dashboard</i> Akses Data Situs BPS	25
Gambar 4. 2 Contoh Tampilan Dataset Teratas TPT	26
Gambar 4. 3 <i>Script</i> untuk Mencari <i>Missing Values</i>	27
Gambar 4. 4 <i>Script</i> untuk Mengganti Nilai yang Hilang Menjadi NaN.....	27
Gambar 4. 5 Hasil Visualisasi Histogram pada df-2013.....	28
Gambar 4. 6 Hasil Visualisasi <i>Boxplot</i> pada df-2013	29
Gambar 4. 7 <i>Script</i> untuk Mengganti <i>Missing Values</i>	29
Gambar 4. 8 <i>Script</i> untuk Melakukan <i>Rename</i> pada Kolom TPT	30
Gambar 4. 9 <i>Script</i> untuk Mengganti Tipe Data pada Kolom	30
Gambar 4. 10 <i>Script</i> untuk Mengubah Format Nama Provinsi	31
Gambar 4. 11 <i>Dataset</i> Sebelum Dilakukan <i>Feature Reduction</i>	32
Gambar 4. 12 <i>Script</i> untuk Menghapus Kolom TPAK.....	32
Gambar 4. 13 <i>Dataset</i> Setelah Dilakukan <i>Feature Reduction</i>	33
Gambar 4. 14 <i>Script</i> untuk Menghapus Data yang Tidak Digunakan	33
Gambar 4. 15 <i>Script</i> untuk Melakukan <i>Merging Dataset</i>	33
Gambar 4. 16 Hasil <i>Merging Data</i>	34
Gambar 4. 17 Cara Kerja <i>K-Means Clustering</i>	35
Gambar 4. 18 <i>Script</i> Visualisasi <i>Elbow Method</i>	35
Gambar 4. 19 Plot Hasil <i>Elbow Method</i>	36
Gambar 4. 20 <i>Script</i> untuk Melakukan <i>Modelling K-Means Clustering</i>	49
Gambar 4. 21 <i>Script</i> untuk Melihat Jumlah Iterasi.....	50
Gambar 4. 22 <i>Script</i> untuk Melihat <i>Centroid</i> Akhir	51
Gambar 4. 23 <i>Script</i> dan Hasil untuk Melihat Analisis Statistik	51
Gambar 4. 24 <i>Script</i> untuk Membuat Visualisasi <i>Boxplot</i> Setiap Fitur.....	52
Gambar 4. 25 Contoh Hasil Visualisasi <i>Boxplot</i> Tiap Fitur.....	53
Gambar 4. 26 Hasil Visualisasi <i>Geo Map</i>	58
Gambar 4. 27 <i>Script</i> untuk Melihat Score DBI	62