

# PERNYATAAN

Saya menyatakan bahwa skripsi ini bebas dari plagiasi, dan apabila dikemudian hari terbukti terdapat plagiasi dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 03 Juli 2024



Agus Bima Saputra

NIM 210200396

...

# HALAMAN PERSETUJUAN

Nama : Agus Bima Saputra  
NIM : 2120200396  
Judul : Algoritma *K-Means* Untuk Mengelompokkan Tingkat Pengangguran  
Terbuka (TPT) Menurut Provinsi Di Indonesia

Telah disetujui dan juga memenuhi syarat untuk diajukannya dalam sidang skripsi.

Bojonegoro, 03 Juli 2024

Pembimbing I



Ucta Pradema Sanjaya, M.Kom.

NIDN. 0729128903

Pembimbing II



Ita Arista Sa'ida, M.Pd.


NIDN. 0708039101

## HALAMAN PENGESAHAN


Nama : Agus Bima Saputra  
NIM : 2120200396  
Program Studi : Teknik Informatika  
Judul : Algoritma *K-Means* Untuk Mengelompokkan Tingkat Pengangguran Terbuka (TPT) Menurut Provinsi Di Indonesia

Telah diujikan dalam sidang skripsi pada tanggal 12 Juli 2024


Dewan Penguji  
Penguji I

  
Nirma Ceisa Santi, M.Kom.  
NIDN.0730099402

Tim Pembimbing  
Pembimbing I

  
Ucta Rradema Sanjaya, M.Kom.  
NIDN. 0729128903

Penguji II

  
Dr. Nurul Huda, M.H.I  
NIDN. 2114067801

Pembimbing II

  
Ita Aristia Sa'ida, M.Pd.  
NIDN. 0708039101

Mengetahui,  
Dekan Fakultas Sains dan Teknologi

  
Muthajamad Jauhar Vikri, M.Kom.  
NIDN. 0712078803  
FST **UNUGIRI**

Mengetahui,  
Ketua Program Studi

  
Mula Azling Barata, S.S.T., M.Kom.  
NIDN. 0711049391  
TEKNIK INFORMATIKA  
FST **UNUGIRI**

## MOTTO DAN PERSEMBAHAN

### MOTTO

*"In Omnia Paratus (Siap Untuk Apapun Yang Terjadi)"*

### PERSEMBAHAN

Skripsi ini saya persembahkan untuk :

1. Kedua orangtua yang telah memberikan *support* dan senantiasa mendoakan yang terbaik bagi saya.
2. Kakak saya yang juga senantiasa memberikan semangat.
3. Teman-teman saya Zakky Septi, Agung Editya, dan Danu Aji yang selalu setia menemani dikala senang maupun susah, dan juga senantiasa meluangkan waktu untuk sekedar bercerita ataupun berdiskusi.

UNUGIRI

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT yang telah memberikan rahmat, nikmat, taufiq, hidayah dan inayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul ” Algoritma *K-Means* Untuk Mengelompokkan Tingkat Pengangguran Terbuka (TPT) Menurut Provinsi Di Indonesia”.

Skripsi ini diajukan dalam rangka menempuh ujian akhir untuk memperoleh gelar Sarjana Komputer (S. Kom) Program Sarjana (S1) Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri.

Penulis menyadari sepenuhnya bahwa penulisan Skripsi ini tidak akan terwujud tanpa adanya dukungan dari berbagai pihak. Ucapan terimakasih penulis haturkan kepada :

1. Kedua orangtua tercinta yang selama hidupnya telah mencurahkan do'a, ikhtiar dan cinta kasih untuk penulis.
2. Bapak M. Jauharul Ma'arif, M.Pd.I selaku Rektor Universitas Nahdlatul Ulama Sunan Giri.
3. Bapak M. Jauhar Vikri, M.Kom selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri.
4. Bapak Ucta Pradema Sanjaya, M. Kom selaku dosen pembimbing I Skripsi yang senantiasa memberikan ilmu, bimbingan, dukungan, dan motivasi untuk penulis menyelesaikan skripsi ini.
5. Ibu Ita Aristia Sa'ida, M. Pd Selaku dosen pembimbing II yang telah memberikan ilmu, arahan, dukungan, dan semangat dalam menyelesaikan skripsi ini.

Penulis menyadari masih kekurangan dalam penelitian dan penulisan skripsi ini. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan guna perbaikan penelitian selanjutnya.

Bojonegoro, 03 Juli 2024

Penulis

## ABSTRACT

Agus Bima Saputra 2024. *K-Means Algorithm for Clustering Open Unemployment Rate (TPT) by Province in Indonesia*. Thesis. Informatics Engineering Study Program. Faculty of Science and Technology. Nahdlatul Ulama Sunan Giri University. Main Supervisor Ucta Pradema Sanjaya, M.Kom and Assistant Supervisor Ita Aristia Sa'ida, M.Pd.

Unemployment is one of the main problems faced by many countries, including Indonesia. The Open Unemployment Rate (OER) is an important indicator used to measure the amount of labor force that is not absorbed in the labor market. This research aims to Cluster the provinces in Indonesia based on unemployment rate and school enrollment rate, so as to provide a clearer picture of the distribution of unemployment in different regions: The study identified three main Clusters: Cluster 1: Provinces with high unemployment rates, Cluster 2: Provinces with a medium unemployment rate, Cluster 3: Provinces with low unemployment rates. Distribution: The Clustering results show that 13 provinces are included in Cluster 1, 18 provinces in Cluster 2, and 3 provinces in Cluster 3. This study found that the K-Means algorithm is effective in Clustering provinces based on TPT and school enrollment rates. The Clustering results show significant variation between provinces, with some provinces having higher unemployment rates and lower school enrollment than others. This study successfully Clustered Indonesian provinces based on unemployment and school enrollment rates using the K-Means algorithm. The Clustering results provide valuable insights into the distribution of unemployment in Indonesia and can be used as a basis for more effective policy making.

**keywords :** TPT, APS, Data Mining, Clustering, K-means Algorithm

The logo of Nahdlatul Ulama Sunan Giri University (UNUGIRI) is a large, faint watermark in the background. It features a stylized emblem with a central figure and the text 'UNUGIRI' in large, bold, capital letters at the bottom.

## ABSTRAK

Agus Bima Saputra 2024. *Algoritma K-Means Untuk Mengelompokkan Tingkat Pengangguran Terbuka (TPT) Menurut Provinsi Di Indonesia*. Skripsi. Program Studi Teknik Informatika. Fakultas Sains dan Teknologi. Universitas Nahdlatul Ulama Sunan Giri. Pembimbing Utama Ucta Pradema Sanjaya, M.Kom dan Pembimbing Pendamping Ita Aristia Sa'ida, M.Pd.

Pengangguran merupakan salah satu masalah utama yang dihadapi oleh banyak negara, termasuk Indonesia. Tingkat Pengangguran Terbuka (TPT) adalah indikator penting yang digunakan untuk mengukur jumlah angkatan kerja yang tidak terserap di pasar tenaga kerja. Penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan tingkat pengangguran dan angka partisipasi sekolah, sehingga dapat memberikan gambaran yang lebih jelas tentang distribusi pengangguran di berbagai wilayah. *Cluster*: Penelitian ini berhasil mengidentifikasi tiga *Cluster* utama: *Cluster 1*: Provinsi dengan tingkat pengangguran tinggi. *Cluster 2*: Provinsi dengan tingkat pengangguran sedang. *Cluster 3*: Provinsi dengan tingkat pengangguran rendah. Distribusi: Hasil *Clustering* menunjukkan bahwa 13 provinsi termasuk dalam *Cluster 1*, 18 provinsi dalam *Cluster 2*, dan 3 provinsi dalam *Cluster 3*. Penelitian ini menemukan bahwa algoritma *K-Means* efektif dalam mengelompokkan provinsi berdasarkan TPT dan angka partisipasi sekolah (APS). Hasil *Clustering* menunjukkan adanya variasi yang signifikan antar provinsi, dengan beberapa provinsi memiliki tingkat pengangguran yang lebih tinggi dan partisipasi sekolah yang lebih rendah dibandingkan provinsi lainnya. Penelitian ini berhasil mengelompokkan provinsi-provinsi di Indonesia berdasarkan TPT dan angka partisipasi sekolah menggunakan algoritma *K-Means*. Hasil *Clustering* memberikan wawasan yang berharga mengenai distribusi pengangguran di Indonesia dan dapat digunakan sebagai dasar dalam pengambilan kebijakan yang lebih efektif.

**kata kunci** : TPT, APS, Data Mining, Clustering, Algoritma K-means

UNUGIRI

## DAFTAR ISI

|   | <b>Halaman</b> |
|---|----------------|
| <b>PERNYATAAN</b> .....                                   | <b>iii</b>     |
| <b>HALAMAN PERSETUJUAN</b> .....                          | <b>iv</b>      |
| <b>HALAMAN PENGESAHAN</b> .....                           | <b>v</b>       |
| <b>MOTTO DAN PERSEMBAHAN</b> .....                        | <b>vi</b>      |
| <b>KATA PENGANTAR</b> .....                               | <b>vii</b>     |
| <b>ABSTRACT</b> .....                                     | <b>viii</b>    |
| <b>ABSTRAK</b> .....                                      | <b>ix</b>      |
| <b>DAFTAR ISI</b> .....                                   | <b>x</b>       |
| <b>DAFTAR TABEL</b> .....                                 | <b>xiii</b>    |
| <b>DAFTAR GAMBAR</b> .....                                | <b>xv</b>      |
| <b>DAFTAR LAMPIRAN</b> .....                              | <b>xvi</b>     |
| <b>BAB I PENDAHULUAN</b> .....                            | <b>1</b>       |
| 1.1 Latar Belakang .....                                  | 1              |
| 1.2 Rumusan Masalah .....                                 | 4              |
| 1.3 Tujuan Penelitian.....                                | 4              |
| 1.4 Batasan Masalah .....                                 | 5              |
| 1.5 Manfaat Penelitian .....                              | 5              |
| <b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI</b> .....      | <b>6</b>       |
| 2.1 Penelitian Terkait.....                               | 6              |
| 2.2 Landasan Teori .....                                  | 19             |
| 2.2.1 <i>Data Mining</i> .....                            | 19             |
| 2.2.2 <i>Clustering</i> .....                             | 20             |
| 2.2.3 <i>K-Means</i> .....                                | 21             |
| 2.2.4 XAMPP .....   | 23             |
| 2.2.5 PHP ( <i>Hypertext Preprocessor</i> ) .....         | 23             |
| 2.2.6 <i>MySQL (My Struktur Query Language)</i> .....     | 24             |
| 2.2.7 <i>Software Development Life Cycle (SDLC)</i> ..... | 24             |
| 2.2.8 <i>Flowchart</i> .....                              | 25             |
| 2.2.9 <i>Unified Modeling Language (UML)</i> .....        | 26             |
| 2.2.10 <i>Kebutuhan Fungsional</i> .....                  | 28             |
| 2.2.11 <i>Kebutuhan Non-Fungsional</i> .....              | 28             |
| 2.2.12 <i>Blackbox Testing</i> .....                      | 28             |



|  |           |
|--|-----------|
| <b>BAB III METODE PENELITIAN .....</b>             | <b>29</b> |
| 3.1 Objek Penelitian .....                         | 29        |
| 3.2 Pengumpulan Data .....                         | 29        |
| 3.2.1 Atribut .....                                | 29        |
| 3.2.2 Data Penelitian .....                        | 29        |
| 3.3 Kerangka Penelitian .....                      | 30        |
| 3.4 Tahapan Penelitian .....                       | 31        |
| 3.5 Metode Yang Diusulkan .....                    | 31        |
| 3.5.1 Algoritma <i>K-Means</i> .....               | 32        |
| 3.6 Metode Pengembangan Sistem.....                | 36        |
| 3.7 Analisis Kebutuhan .....                       | 37        |
| 3.7.1 <i>System Request</i> .....                  | 37        |
| 3.7.2 <i>Analysis</i> .....                        | 37        |
| 3.7.3 Deskripsi Kebutuhan <i>User</i> .....        | 38        |
| 3.7.4 Kebutuhan Fungsional.....                    | 38        |
| 3.7.5 Kebutuhan Non Fungsional.....                | 40        |
| 3.7.6 Kebutuhan Perangkat Lunak .....              | 40        |
| 3.7.7 Kebutuhan Perangkat Keras .....              | 41        |
| 3.8 Desain Antar Muka .....                        | 41        |
| 3.9 Mockup .....                                   | 42        |
| 3.10 <i>Use Case</i> .....                         | 44        |
| 3.11 <i>Acitivity Diagram</i> .....                | 45        |
| 3.12 Rencana Pengujian .....                       | 48        |
| 3.12.1 Pengujian BlackBox .....                    | 48        |
| 3.13 <i>Timeline</i> Pengerjaan .....              | 50        |
| <b>BAB IV HASIL DAN PEMBAHASAN.....</b>            | <b>51</b> |
| 4.1 Implementasi Sistem .....                      | 51        |
| 4.1.1 Tampilan Halaman <i>Login</i> .....          | 51        |
| 4.1.2 Tampilan Halaman <i>Home/Dashboard</i> ..... | 51        |
| 4.1.3 Tampilan Halaman <i>Data</i> .....           | 52        |
| 4.1.4 Tampilan Halaman <i>Processing</i> .....     | 52        |
| 4.1.5 Tampilan Halaman <i>Result / Hasil</i> ..... | 53        |
| 4.1.6 Tampilan Halaman <i>Report</i> .....         | 53        |
| 4.1.7 Tampilan Halaman <i>Logout</i> .....         | 54        |
| 4.2 Implementasi Perhitungan Data .....            | 54        |
| 4.2.1 <i>Preprocessing Data</i> .....              | 55        |

|  |           |
|--|-----------|
| 4.2.2 Perhitungan Menggunakan <i>Microsoft Excel</i> .....       | 56        |
| 4.2.3 Uji Akurasi dengan <i>Davies-Bouldin Index (DBI)</i> ..... | 70        |
| 4.3 Pengujian <i>Blackbox</i> .....                              | 70        |
| <b>BAB V KESIMPULAN DAN SARAN</b> .....                          | <b>72</b> |
| 5.1 Kesimpulan .....   | 72        |
| 5.2 Saran .....  | 72        |
| <b>DAFTAR PUSTAKA</b> .....                                      | <b>73</b> |
| <b>LAMPIRAN</b> .....  | <b>77</b> |



**UNUGIRI**

## DAFTAR TABEL

|  | Halaman |
|--|---------|
| Tabel 2. 1 Tinjauan Pustaka .....  | 8       |
| Tabel 2. 2 Simbol <i>Flowchart</i> .....                                       | 25      |
| Tabel 2. 3 Simbol <i>Use Case Diagram</i> .....                                | 26      |
| Tabel 2. 4 Simbol <i>Activity Diagram</i> .....                                | 27      |
| Tabel 3. 1 <i>Dataset</i> Tingkat Pengangguran Terbuka (TPT).....              | 30      |
| Tabel 3. 2 <i>Dataset</i> Tingkat Pengangguran Terbuka (TPT) 2019-2023 .....   | 32      |
| Tabel 3. 3 <i>Dataset</i> Angka Partisipasi Sekolah (APS) 2019-2023 .....      | 32      |
| Tabel 3. 4 <i>Dataset</i> setelah di <i>Pre-Processing</i> .....               | 33      |
| Tabel 3. 5 Pusat <i>Cluster Random</i> .....                                   | 34      |
| Tabel 3. 6 Jarak <i>Euclidian Distance</i> dan WCV.....                        | 34      |
| Tabel 3. 7 Tabel <i>Between Cluster Variation</i> (BCV).....                   | 35      |
| Tabel 3. 8 Hasil <i>Cluster</i> yang telah konvergen .....                     | 35      |
| Tabel 3. 9 <i>System Request</i> .....   | 37      |
| Tabel 3. 10 <i>Analysis</i> .....  | 37      |
| Tabel 3. 11 Deskripsi Kebutuhan <i>User</i> .....                              | 38      |
| Tabel 3. 12 Kebutuhan Fungsional.....  | 39      |
| Tabel 3. 13 Kebutuhan Non Fungsional.....                                      | 40      |
| Tabel 3. 14 Kebutuhan Perangkat Lunak .....                                    | 40      |
| Tabel 3. 15 Kebutuhan Perangkat Keras .....                                    | 41      |
| Tabel 3. 16 Pengujian Perangkat Lunak Menggunakan Metode Blackbox .....        | 49      |
| Tabel 3. 17 <i>Timeline</i> Pengerjaan .....                                   | 50      |
| Tabel 4. 1 Data yang sudah di <i>preProcessing</i> .....                       | 55      |
| Tabel 4. 2 Pusat <i>Cluster Random</i> .....                                   | 56      |
| Tabel 4. 3 Hasil Perhitungan Jarak ke Pusat <i>Cluster</i> Iterasi 1 .....     | 56      |
| Tabel 4. 4 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 1 .....  | 57      |
| Tabel 4. 5 Hasil Perhitungan Rasio Iterasi 1 .....                             | 57      |
| Tabel 4. 6 Pusat <i>Cluster</i> Baru Iterasi 2.....                            | 58      |
| Tabel 4. 7 Hasil Perhitungan Jarak ke Pusat <i>Cluster</i> Iterasi 2 .....     | 58      |
| Tabel 4. 8 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 2 .....  | 59      |
| Tabel 4. 9 Hasil Perhitungan Rasio Iterasi 2 .....                             | 59      |
| Tabel 4. 10 Pusat <i>Cluster</i> Baru Iterasi 3.....                           | 59      |
| Tabel 4. 11 Hasil Perhitungan Jarak ke Pusat <i>Cluster</i> Iterasi 3.....     | 60      |
| Tabel 4. 12 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 3 ..... | 61      |
| Tabel 4. 13 Hasil Perhitungan Rasio Iterasi 3 .....                            | 61      |
| Tabel 4. 14 Pusat <i>Cluster</i> Baru Iterasi 4.....                           | 61      |
| Tabel 4. 15 Hasil Perhitungan Jarak ke Pusat <i>Cluster</i> Iterasi 4 .....    | 61      |
| Tabel 4. 16 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 4 ..... | 62      |
| Tabel 4. 17 Hasil Perhitungan Rasio Iterasi 4 .....                            | 63      |
| Tabel 4. 18 Pusat <i>Cluster</i> Baru Iterasi 5.....                           | 63      |
| Tabel 4. 19 Hasil Perhitungan Jarak ke Pusat <i>Cluster</i> Iterasi 5 .....    | 63      |
| Tabel 4. 20 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 5 ..... | 64      |
| Tabel 4. 21 Hasil Perhitungan Rasio Iterasi 5 .....                            | 64      |

|  |    |
|--|----|
| Tabel 4. 22 Pusat <i>Cluster</i> Baru Iterasi 6.....                           | 65 |
| Tabel 4. 23 Hasil Perhitungan Jarak Ke Pusat <i>Cluster</i> Iterasi 6.....     | 65 |
| Tabel 4. 24 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 6 ..... | 66 |
| Tabel 4. 25 Hasil Perhitungan Rasio Iterasi 6 .....                            | 66 |
| Tabel 4. 26 Pusat <i>Cluster</i> Baru Iterasi 7.....                           | 66 |
| Tabel 4. 27 Hasil Perhitungan Jarak Ke Pusat <i>Cluster</i> Iterasi 7.....     | 67 |
| Tabel 4. 28 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 7 ..... | 67 |
| Tabel 4. 29 Hasil Perhitungan Rasio Iterasi 7 .....                            | 68 |
| Tabel 4. 30 Pusat <i>Cluster</i> Baru Iterasi 8.....                           | 68 |
| Tabel 4. 31 Hasil Perhitungan Jarak ke Pusat <i>Cluster</i> Iterasi 8 .....    | 68 |
| Tabel 4. 32 Hasil Perhitungan Jarak Antar Pusat <i>Cluster</i> Iterasi 8 ..... | 69 |
| Tabel 4. 33 Hasil Perhitungan Rasio Iterasi 8.....                             | 69 |
| Tabel 4. 34 Tabel Uji <i>Blackbox</i> .....                                    | 70 |



**UNUGIRI**

## DAFTAR GAMBAR

|  | <b>Halaman</b> |
|--|----------------|
| Gambar 2. 1 <i>Flowchart</i> Algoritma <i>K-Means</i> .....                | 23             |
| Gambar 2. 2 Model <i>Waterfall</i> .....                                   | 24             |
| Gambar 3. 1 Kerangka Penelitian .....                                      | 30             |
| Gambar 3. 2 Pengujian <i>dataset</i> dengan algoritma <i>K-Means</i> ..... | 32             |
| Gambar 3. 3 Metode <i>Waterfall</i> .....                                  | 36             |
| Gambar 3. 4 <i>Mock-Up</i> Halaman <i>Login</i> .....                      | 42             |
| Gambar 3. 5 <i>Mock-Up</i> Halaman <i>Dashboard</i> .....                  | 42             |
| Gambar 3. 6 <i>Mock-Up</i> Menu <i>Dataset</i> .....                       | 43             |
| Gambar 3. 7 <i>Mock-Up</i> Menu Proses <i>Clustering Dataset</i> .....     | 43             |
| Gambar 3. 8 <i>Mock-Up</i> Menu Hasil <i>Clustering</i> .....              | 44             |
| Gambar 3. 9 <i>Use Case Diagram</i> .....                                  | 44             |
| Gambar 3. 10 <i>Activity Diagram Login</i> .....                           | 46             |
| Gambar 3. 11 <i>Activity Diagram Dashboard</i> .....                       | 46             |
| Gambar 3. 12 <i>Activity Diagram Dataset</i> .....                         | 47             |
| Gambar 3. 13 <i>Activity Diagram Perhitungan</i> .....                     | 47             |
| Gambar 3. 14 <i>Activity Diagram Hasil Perhitungan</i> .....               | 48             |
| Gambar 4. 1 Tampilan <i>Login</i> .....                                    | 51             |
| Gambar 4. 2 Tampilan <i>Home</i> .....                                     | 51             |
| Gambar 4. 3 Tampilan <i>Data</i> .....                                     | 52             |
| Gambar 4. 4 Tampilan <i>Processing</i> .....                               | 52             |
| Gambar 4. 5 Tampilan <i>Hasil</i> .....                                    | 53             |
| Gambar 4. 6 Tampilan <i>Report</i> .....                                   | 53             |
| Gambar 4. 7 Tampilan <i>Fitur Cetak</i> .....                              | 54             |
| Gambar 4. 8 Tampilan <i>Logout</i> .....                                   | 54             |
| Gambar 4. 9 Uji akurasi <i>rapidminer</i> .....                            | 70             |

UNUGIRI

## DAFTAR LAMPIRAN

|                                    | <b>Halaman</b> |
|------------------------------------|----------------|
| Lampiran 1 Hasil Cek Plagiasi..... | 77             |
| Lampiran 2 Uji Blackbox.....       | 78             |



# UNUGIRI