

**IMPLEMENTASI ALGORITMA K-MEANS UNTUK
CLUSTERING JAMAAH NGAJI TERAS DI TERAS DAKWAH**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Sistem Informasi



disusun oleh

MUHAMMAD SHAWQI RAHMAN

16.12.9010

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**IMPLEMENTASI ALGORITMA K-MEANS UNTUK
CLUSTERING JAMAAH NGAJI TERAS DI TERAS DAKWAH**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Sistem Informasi



disusun oleh

MUHAMMAD SHAWQI RAHMAN

16.12.9010

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI ALGORITMA K-MEANS UNTUK CLUSTERING JAMAAH
NGAJI TERAS DI TERAS DAKWAH**

yang disusun dan diajukan oleh

Muhammad Shawqi Rahman

16.12.9010

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Maret 2023

Dosen Pembimbing,

Eli Pujastuti, M.Kom

NIK. 190302227

HALAMAN PENGESAHAN

SKRIPSI

**IMPLEMENTASI ALGORITMA K-MEANS UNTUK CLUSTERING JAMAAH NGAJI
TERAS DI TERAS DAKWAH**

yang disusun dan diajukan oleh

Muhammad Shawqi Rahman

16.12.9010


Telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Februari 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Wiwi Widayani, M.Kom
NIK. 190302272



Irma Rofni Wulandari, S.Pd., M.Eng
NIK. 190302329

Eli Pujiastuti, M.Kom
NIK. 190302227

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 22 Februari 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 03 Mei 2023



Muhammad Shawqi Rahman

NIM. 16.12.9010

MOTTO

“Aku meminta kekuatan dan Allah memberikanku kesulitan untuk membuatku semakin kuat.”

”Jika ada berbagai macam kemungkinan, maka pilihlah kemungkinan yang paling berbeda dari semua kemungkinan-kemungkinan yang ada.”



PERSEMBAHAN

Banyak ucapan syukur penulis kepada berbagai pihak yang telah membantu penulis dalam menyelesaikan skripsi ini. Ucapan terima kasih dari penulis kepada:

1. Allah *Subhanahuwata'ala, Rabb* atas segala yang ada dan hidup di seluruh alam, atas segala karunia dan rahmat-Nya yang telah mengkaruniakan kemudahan dan kelancaran bagi penulis dalam menyelesaikan penulisan skripsi ini,
2. Kedua orang tua penulis yang senantiasa mengingatkan, mendoakan dan memberikan dukungan kepada penulis dalam bentuk apapun,
3. Ibu Eli Pujastuti M.Kom., selaku dosen pembimbing yang telah memberikan bimbingan dan arahan kepada penulis,
4. Seluruh dosen dan guru yang telah dan pernah memberikan ilmu pengetahuan dalam bentuk apapun. Yang baik secara sengaja maupun tidak sengaja telah menjadi motivasi dan pengetahuan penulis dalam menyelesaikan skripsi,
5. Bapak Achid Subiyanto selaku pengasuh Teras Dakwah, Bapak Ilyas Chairuddin selaku pimpinan 1 Teras dakwah, Bapak Rais Afif dan seluruh jajaran manajemen Teras Dakwah yang telah banyak memberikan dukungan secara moral kepada penulis dalam menyelesaikan skripsi ini yang tidak bisa penulis sebutkan satu persatu.

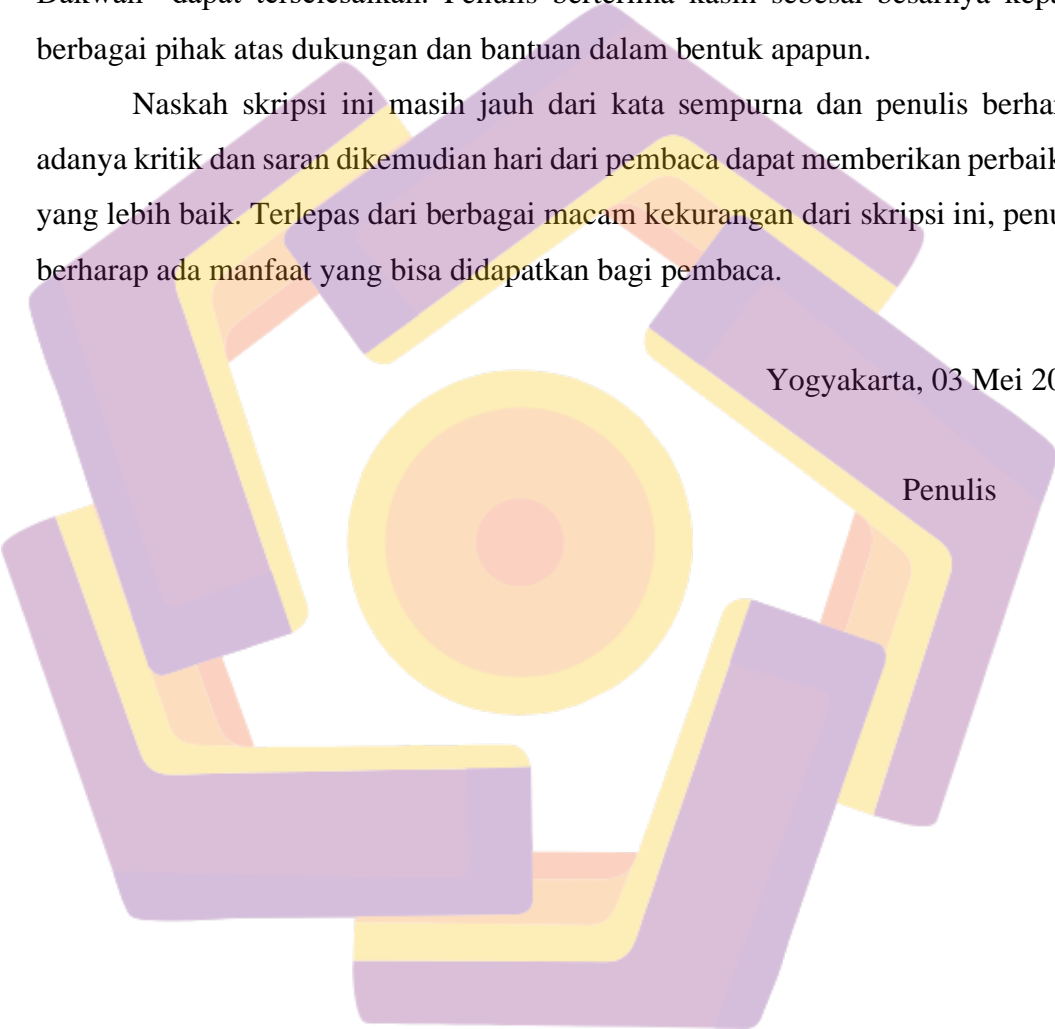
KATA PENGANTAR

Segala puji bagi Allah yang Maha Pengasih lagi Maha Penyayang yang telah mengkaruniakan rahmat dan hidayah-Nya sehingga skripsi berjudul, "Implementasi Algoritma K-Means untuk Clustering Jamaah Ngaji Teras di Teras Dakwah" dapat terselesaikan. Penulis berterima kasih sebesar-besarnya kepada berbagai pihak atas dukungan dan bantuan dalam bentuk apapun.

Naskah skripsi ini masih jauh dari kata sempurna dan penulis berharap adanya kritik dan saran dikemudian hari dari pembaca dapat memberikan perbaikan yang lebih baik. Terlepas dari berbagai macam kekurangan dari skripsi ini, penulis berharap ada manfaat yang bisa didapatkan bagi pembaca.

Yogyakarta, 03 Mei 2023

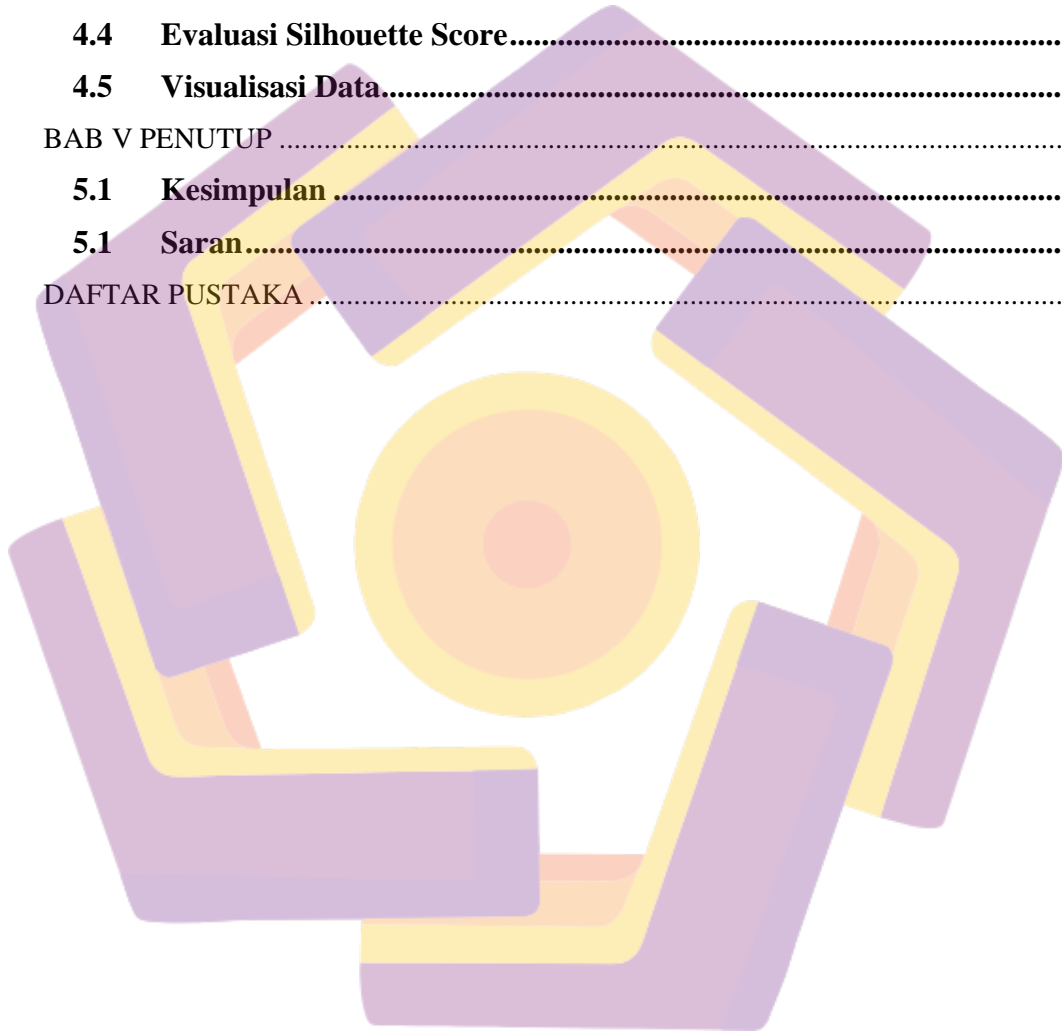
Penulis



DAFTAR ISI

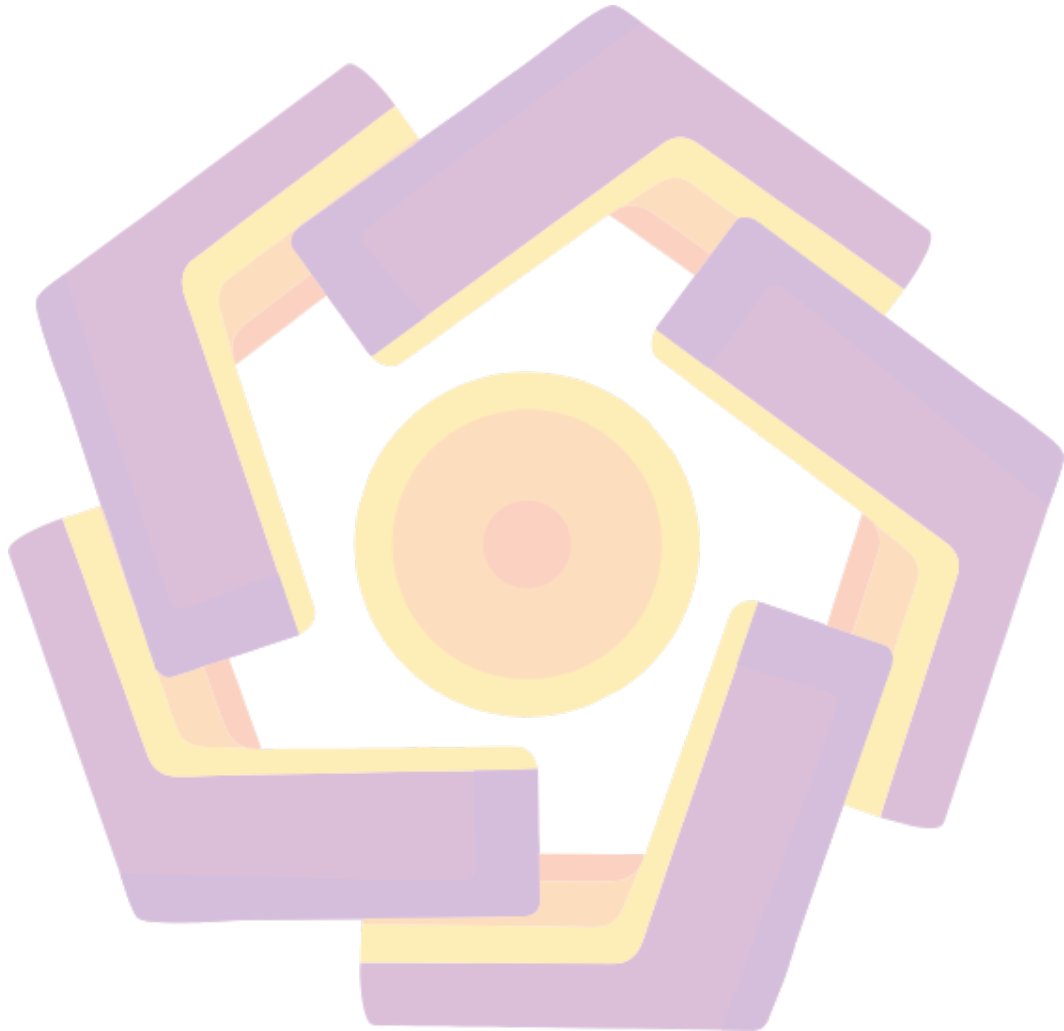
HALAMAN JUDUL.....	1
HALAMAN PERSETUJUAN.....	2
HALAMAN PENGESAHAN.....	3
PERNYATAAN	4
MOTTO	5
PERSEMBAHAN.....	6
KATA PENGANTAR.....	7
DAFTAR ISI.....	8
DAFTAR TABEL.....	10
DAFTAR GAMBAR.....	11
INTISARI	12
ABSTRACT.....	13
BAB I PENDAHULUAN.....	14
1.1 Latar Belakang	14
1.2 Rumusan Masalah.....	15
1.3 Batasan Masalah.....	15
1.4 Tujuan Penelitian	15
1.5 Manfaat Penelitian	16
1.6 Metode Penelitian	16
1.7 Sistematika Penulisan	18
BAB II TINJAUAN PUSTAKA.....	20
2.1 Tinjauan Pustaka.....	20
2.2 Dasar Teori.....	23
2.2.1 Definisi Cluster	23
2.2.2 Algoritma <i>K-Means</i>	24
2.2.3 Euclidean Distance	25
2.2.4 Metode evaluasi menggunakan <i>Silhouette score</i>.....	26
BAB III METODE PENELITIAN.....	27
3.1 Tinjauan umum	27
3.2 Akuisisi Dataset	28

3.3	<i>Preprocessing Data</i>	29
3.4	Alat.....	30
BAB IV HASIL DAN PEMBAHASAN.....		32
4.1	Koneksi File.....	32
4.2	Pre-processing.....	33
4.3	Implementasi Algoritman <i>K-Means Clustering</i>	37
4.4	Evaluasi Silhouette Score.....	37
4.5	Visualisasi Data.....	38
BAB V PENUTUP		40
5.1	Kesimpulan	40
5.1	Saran.....	40
DAFTAR PUSTAKA		42



DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Penelitian Terdahulu	21
Tabel 3.1 Konversi data	29
Tabel 4.1 Tabel Uji Coba Jumlah Cluster	38



DAFTAR GAMBAR

Gambar 2.1 Diagram Alir K-Means	25
Gambar 3.1 Alur implementasi Algoritma K-Means	28
Gambar 3.2 Tabel Data Peserta Ngaji Teras	29
Gambar 3.3 Tabel Data Setelah Preprocessing	30
Gambar 4.1 Pemanggilan Library dan Koneksi Google Colab dengan Google Drive	32
Gambar 4.2 Pemanggilan File dan Tampilan File	32
Gambar 4.3 Penyeleksian Dataset Duplikat	33
Gambar 4.4 Penyeleksian Dataset Berdasarkan Provinsi	33
Gambar 4.5 Pembuatan Dataframe Baru Dengan Variabel Tertentu	34
Gambar 4.6 Memastikan Tidak Ada Dataset yang Kosong	34
Gambar 4.7 Menseleksi Dataset	35
Gambar 4.8 Konversi Dataset ke Dalam Numerik	36
Gambar 4.9 Menampilkan Dataset yang Akan di Cluster	36
Gambar 4.10 Kode untuk Implementasi K-Means Clustering	37
Gambar 4.11 Visualisasi Dataset Menggunakan Matplotlib	37
Gambar 4.12 Hasil Evaluasi Clustering Menggunakan Silhouette Score	38
Gambar 4.13 Visualisasi Dataset Jumlah Umur Berdasarkan Cluster	39
Gambar 4.14 Visualisasi Dataset Jumlah Donatur Berdasarkan Umur	39

INTISARI

Yayasan yang bergerak di bidang sosial, terutama filantropi, sangat bergantung kepada dana donatur yang dikumpulkan dari masyarakat secara sukarela. Salah satu cara Teras Dakwah mendapatkan donatur adalah melalui kajian-kajian. Dengan memberikan formulir daftar hadir kepada jamaah, didapatkanlah data donatur tetap. Data yang didapatkan, kemudian dimasukkan kedalam form Microsoft Excel yang kemudian akan dihubungi kembali sebagai bentuk konfirmasi Ketika mengisi formulir.

Setelah mendapatkan konfirmasi, data tidak diolah lebih lanjut. Maka metode clustering digunakan untuk mengolah data yang telah ada dengan harapan dapat memberikan makna dan menemukan pengetahuan baru. Untuk melakukan clustering, maka algoritma yang dipilih dan akan digunakan adalah Algoritma K-Means. K-Means dipilih karena telah digunakan dalam penerapan-penerapan kepada data yang hampir serupa di objek yang berbeda. Data yang digunakan hanyalah data usia dan ketersediaan untuk menjadi donatur tetap.

Didapatkan keberhasilan dalam penerapan Algoritma K-Means pada penelitian ini. Hasil evaluasi yang dilakukan menggunakan Silhouette Score menunjukkan angka 0,646297267679837 dengan jumlah cluster sebanyak 2. Dimana cluster 1 mendapatkan rentang usia 15-30 dan cluster 2 dengan 31-46. Dalam penelitian kedepan, peneliti menyarankan untuk menjadikan data wilayah sebagai salah satu data yang diproses clustering untuk mendapatkan pengetahuan yang lebih mengenai hasil dari penerapan clustering di Teras Dakwah. Penentuan cluster di awal menggunakan Statistic Hopkins atau Probabilistic Sampling lebih disarankan untuk dijadikan sebagai bahan penelitian.

Keyword: K-Means, Clustering, Silhouette Score

ABSTRACT

Foundations that operate in the social and philanthropic sectors heavily rely on donations collected from the voluntary contributions of the community. One of the ways Teras Dakwah obtains donors is through study sessions. By providing attendance forms to the congregation, Teras Dakwah obtains data on regular donors. This data is then entered into a Microsoft Excel form, and the donors are contacted again as a confirmation when filling out the form.

After confirmation, the data is not further processed. Therefore, clustering methods are used to process the existing data with the hope of providing meaning and discovering new knowledge. To perform clustering, the selected algorithm is the K-Means Algorithm. K-Means is chosen because it has been used in similar applications to data from different objects. The data used only consists of age and willingness to become regular donors.

Successful implementation of the K-Means algorithm in this study was achieved. The evaluation conducted using the Silhouette Score showed a value of 0,646297267679837 with a total of 2 clusters. Cluster 1 obtained an age range of 15-30, and cluster 2 had an age range of 31-46. In future research, the researchers suggest including 'wilayah' data as one of the clustering data to obtain more knowledge about the results of clustering implementation in Teras Dakwah. The determination of clusters at the beginning using the Statistic Hopkins or Probabilistic Sampling is more recommended as a research material.

Keyword: *K-Means, Clustering, Silhouette Score*