

**PENERAPAN ALGORITMA K-MEANS PADA  
PENGELOMPOKAN DATA CALON SISWA BARU DI SEKOLAH  
MENENGAH KEJURUAN (STUDI KASUS : SMK MUHAMMADIYAH  
PAGUYANGAN)**

**SKRIPSI**

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



**Oleh :**

**NINGSI MULTI PURNAMASARI**

**42317003**

**JURUSAN SISTEM INFORMASI**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS PERADABAN**

**2022**

## PERNYATAAN PENULIS

JUDUL : PENERAPAN ALGORITMA *K-MEANS* PADA  
PENGELOMPOKAN DATA CALON SISWA BARU DI  
SEKOLAH MENENGAH KEJURUAN (STUDI KASUS : SMK  
MUHAMMADIYAH PAGUYANGAN)

NAMA : NINGSI MULTI PURNAMASARI

NIM : 42317003

“Saya menyatakan dan bertanggungjawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.

Bumiayu, 01 September 2022



Ningsi Multi Purnamasari

## HALAMAN PERSETUJUAN

JUDUL : PENERAPAN ALGORITMA *K-MEANS* PADA  
PENGELOMPOKAN DATA CALON SISWA BARU DI  
SEKOLAH MENENGAH KEJURUAN (STUDI KASUS : SMK  
MUHAMMADIYAH PAGUYANGAN)

NAMA : NINGSI MULTI PURNAMASARI

NIM : 42317003

Skripsi ini telah diperiksa dan disetujui

Bumiayu, 20 Agustus 2022

Pembimbing I



**Danar Ardian P, S.Si., M.Sc**

NIDN. 0614048802

Pembimbing II



**Achmad Syauqi, M.Kom**

NIDN: 0604068804

Ketua Jurusan,



**Mukrodin, M.Kom**  
NIDN. 0610038001

## PENGESAHAN SKRIPSI

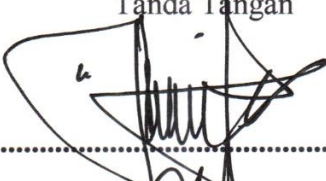



JUDUL : PENERAPAN ALGORITMA *K-MEANS* PADA  
PENGELOMPOKAN DATA CALON SISWA BARU DI  
SEKOLAH MENENGAH KEJURUAN (STUDI KASUS : SMK  
MUHAMMADIYAH PAGUYANGAN)

NAMA : NINGSI MULTI PURNAMASARI

NIM : 42317003

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 01 September 2022 Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk ditujukan penganugerahan gelar Sarjana Komputer (S.Kom)

Bumiayu, 08 September 2022

Nama Penguji	Tanda Tangan
1 <b><u>Mukrodin, M.Kom</u></b> NIDN. 0610038001	1 
2 <b><u>Fuaida Nabya, M.Kom</u></b> NIDN. 0625079301	2 
3 <b><u>Danar Ardian P, S.Si., M.Sc</u></b> NIDN. 0614048802	3 
4 <b><u>Achmad Syaqui, M.Kom</u></b> NIDN: 0604068804	4 

Mengetahui :

Dekan Fakultas Sains dan Teknologi  
Universitas Peradaban

  
**Dr. Apt. Pudiono, S.U**  
NUPN. 9990000424

Ketua Program Studi  
Sistem Informasi

  
**Mukrodin, M.Kom**  
NIDN. 0610038001

## **ABSTRACT**

*The choice of majors in vocational high schools (SMK) is one of the things that determines the success of students' learning. SMK Muhammadiyah Paguyangan is one of the schools of choice for prospective new students in continuing their education, with the aim of increasing the human resources that are needed. Often students are found who enter the wrong major because they choose a major only based on information from friends. In the process so far, an accurate clustering method has not been applied so that the level of accuracy resulting from the clustering of majors is unknown. One of the steps to achieve maximum results in the grouping of majors is data processing. One of them is by clustering using the K-Means algorithm. The K-Means algorithm is a clustering algorithm for data grouping based on the cluster center point (centroid) closest to the data. In this study, researchers applied the K-Means Clustering algorithm to group data for new students at SMK Muhammadiyah Paguyangan. This study aims to group data on prospective new students based on majors by using the variables of student values and interests, by applying the K-Means method. The results of this study are expected to assist in data processing for the determination of majors.*

*Keywords: Selection of Majors, Clustering, K-Means*

## ABSTRAK

Pemilihan jurusan di sekolah menengah kejuruan (SMK) merupakan salah satu hal yang menentukan keberhasilan belajar peserta didik. SMK Muhammadiyah Paguyangan merupakan salah satu sekolah yang menjadi pilihan untuk para calon siswa baru dalam melanjutkan pendidikannya, dengan tujuan untuk dapat meningkatkan sumber daya manusia yang sedang dibutuhkan. Seringkali dijumpai siswa yang salah masuk jurusan dikarenakan memilih jurusan hanya berdasarkan informasi dari teman. Dalam prosesnya selama ini belum diterapkan metode sebuah metode klasterisasi yang akurat sehingga membuat tingkat akurasi yang dihasilkan dari klasterisasi jurusan belum diketahui. Salah satu langkah untuk mencapai hasil maksimal dalam pengelompokan jurusan adalah dengan pengolahan data. Salah satunya dengan pengklasteran menggunakan algoritma *K-Means*. Algoritma *K-Means* merupakan algoritma klasterisasi pengelompokan data berdasarkan titik pusat *cluster (centroid)* terdekat dengan data. Pada penelitian ini peneliti menerapkan algoritma *K-Means Clustering* untuk mengelompokkan data siswa baru di SMK Muhammadiyah Paguyangan. Penelitian ini bertujuan untuk mengelompokkan data calon siswa baru berdasarkan jurusan dengan menggunakan variabel nilai dan minat siswa, dengan menerapkan metode *K-Means*. Hasil dari penelitian ini diharapkan dapat membantu dalam pengolahan data untuk penentuan jurusan.

**Kata Kunci :** *Pemilihan Jurusan, Klasterisasi, K-Means*

## KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT, yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul Penerapan Algoritma *K-Means* pada Pengelompokan Data Calon Siswa Baru di Sekolah Menengah Kejuruan.

Penyusunan skripsi ini merupakan salah satu syarat untuk menyelesaikan Program Studi Sistem Informasi S-1 pada Fakultas Sains dan Teknologi Universitas Peradaban Bumiayu, penulis banyak mendapatkan dukungan, bantuan dan bimbingan dari berbagai pihak. Sehubungan dengan itu maka penulis menyampaikan ucapan terima kasih kepada:

1. Dr. Muh. Kadarisman., S.H., M.Si, selaku Rektor Universitas Peradaban Bumiayu.
2. Bapak Dr. Pudjono, S.U., Apt, selaku Dekan Fakultas Sains dan Teknologi Universitas Peradaban.
3. Bapak Mukrodin, M.Kom, selaku Ketua Jurusan Sistem Informasi dan penguji yang telah memberikan banyak masukan dalam Skripsi saya.
4. Bapak Danar Ardian P, S.Si., M.Sc, selaku pembimbing pertama yang telah memberikan banyak pengarahan dalam menyelesaikan laporan Skripsi ini.
5. Bapak Achmad Syauqi, M.Kom, selaku pembimbing kedua yang telah memberikan banyak pengarahan dalam menyelesaikan laporan Skripsi ini.
6. Ibu Fuaida Nabyla, M.Kom, selaku penguji yang telah memberikan banyak masukan dalam Skripsi saya.
7. Bapak Alimudin, M.Pd selaku kepala sekolah SMK Muhammadiyah Paguyangan yang telah memberikan izin dalam penelitian ini.
8. Bapak M. Ardiyansah, M.Kom, selaku waka kurikulum SMK Muhammadiyah Paguyangan yang telah banyak membantu saya dalam penelitian ini.
9. Bapak / Ibu dosen dan staff di lingkungan Fakultas Sains dan Teknologi, khususnya Program Studi Sistem Informasi terima kasih atas pengalaman dan pengetahuan yang telah didapatkan penulis selama menyelesaikan studi ini.

10. Suami saya Oki Indra Saktian, S.Pd yang selalu memberi semangat, do'a dan banyak bantuan.
11. Kedua orang tua saya yang selalu mendoakan dan mendukung saya dalam keadaan apapun.
12. Adik-adik saya "Wili Juli Saputra" dan "Khaleysia Hava Nara" yang menjadi penyemangat saya.
13. Rekan Mahasiswa/i Program Studi S-1 Sistem Informasi khususnya angkatan 2017.

Penulis menyadari bahwa dalam penyusunan laporan ini masih banyak kekurangan, sehingga saran dan kritik yang membangun sangat penulis harapkan. Akhir kata penulis berharap penyusunan skripsi ini dapat membawa manfaat dan sumbangsih pengetahuan bagi masyarakat.



## DAFTAR ISI

HALAMAN JUDUL.....	i
PERNYATAAN PENULIS .....	ii
HALAMAN PERSETUJUAN.....	iii
PENGESAHAN SKRIPSI .....	iv
ABSTRACK .....	iv
ABSTRAK .....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiii
DAFTAR LAMPIRAN.....	14
BAB I PENDAHULUAN.....	<b>Error! Bookmark not defined.</b>
1.1 Latar Belakang .....	<b>Error! Bookmark not defined.</b>
1.2 Rumusan Masalah .....	<b>Error! Bookmark not defined.</b>
1.3 Tujuan Penelitian.....	<b>Error! Bookmark not defined.</b>
1.4 Manfaat Penelitian.....	<b>Error! Bookmark not defined.</b>
1.5 Ruang Lingkup .....	<b>Error! Bookmark not defined.</b>
1.6 Sistematika Penulisan.....	<b>Error! Bookmark not defined.</b>
BAB II TINJAUAN PUSTAKA.....	<b>Error! Bookmark not defined.</b>
2.1 Penelitian Terkait .....	<b>Error! Bookmark not defined.</b>
2.2 Landasan Teori .....	<b>Error! Bookmark not defined.</b>
2.2.1 Profil SMK Muhammadiyah Paguyangan .....	<b>Error! Bookmark not defined.</b>
2.2.2 Data Mining .....	<b>Error! Bookmark not defined.</b>
2.2.3 Clustering .....	<b>Error! Bookmark not defined.</b>
2.2.4 Algoritma .....	<b>Error! Bookmark not defined.</b>
2.2.5 Algoritma K-Means .....	<b>Error! Bookmark not defined.</b>
2.3 Tools.....	<b>Error! Bookmark not defined.</b>
2.3.1 Anaconda.....	<b>Error! Bookmark not defined.</b>
2.3.2 Jupyter Notebook .....	<b>Error! Bookmark not defined.</b>
2.3.3 Enterprise Architect .....	<b>Error! Bookmark not defined.</b>
2.3.4 Flowchart .....	<b>Error! Bookmark not defined.</b>

2.4	Library .....	<b>Error! Bookmark not defined.</b>
2.4.1	Scikit-Learn.....	<b>Error! Bookmark not defined.</b>
2.5	Kerangka Berfikir.....	<b>Error! Bookmark not defined.</b>
BAB III METODE PENELITIAN.....		<b>Error! Bookmark not defined.</b>
3.1	Tempat dan waktu penelitian .....	<b>Error! Bookmark not defined.</b>
3.2	Alat dan Bahan .....	<b>Error! Bookmark not defined.</b>
3.3	Tahapan Penelitian .....	<b>Error! Bookmark not defined.</b>
3.3.1	Identifikasi Masalah .....	<b>Error! Bookmark not defined.</b>
3.3.2	Studi literatur.....	<b>Error! Bookmark not defined.</b>
3.3.3	Pengumpulan Data .....	<b>Error! Bookmark not defined.</b>
3.3.4	Menganalisa data.....	<b>Error! Bookmark not defined.</b>
3.3.5	Penerapan Metode K-Means .....	<b>Error! Bookmark not defined.</b>
3.3.6	Pembuatan Laporan Hasil Penelitian .....	<b>Error! Bookmark not defined.</b>
3.4	Jadwal penelitian .....	<b>Error! Bookmark not defined.</b>
BAB IV PEMBAHASAN.....		<b>Error! Bookmark not defined.</b>
4.1	Persiapan Data .....	<b>Error! Bookmark not defined.</b>
4.2	Transformasi Data .....	<b>Error! Bookmark not defined.</b>
4.3	Clustering Menggunakan K-Means.....	<b>Error! Bookmark not defined.</b>
4.4	Implementasi pada Jupyter Notebook .....	<b>Error! Bookmark not defined.</b>
4.4.1	Clustering Jurusan TKR.....	<b>Error! Bookmark not defined.</b>
4.4.2	Clustering Jurusan TSM.....	<b>Error! Bookmark not defined.</b>
4.4.3	Clustering Jurusan RPL .....	<b>Error! Bookmark not defined.</b>
4.4.4	Hasil Cluster.....	<b>Error! Bookmark not defined.</b>
BAB V PENUTUP.....		<b>Error! Bookmark not defined.</b>
5.1	Kesimpulan.....	<b>Error! Bookmark not defined.</b>
5.2	Saran .....	<b>Error! Bookmark not defined.</b>
DAFTAR PUSTAKA .....		<b>Error! Bookmark not defined.</b>

## DAFTAR TABEL

Tabel 2. 1 Perbandingan Penelitian Terkait .....	<b>Error! Bookmark not defined.</b>
Tabel 2. 2 Kompetensi Keahlian .....	<b>Error! Bookmark not defined.</b>
Tabel 2. 3 Fasilitas .....	<b>Error! Bookmark not defined.</b>
Tabel 2. 4 Simbol penghubung atau alur .....	<b>Error! Bookmark not defined.</b>
Tabel 2. 5 Simbol Proses.....	<b>Error! Bookmark not defined.</b>
Tabel 2. 6 Input / Output Simbol .....	<b>Error! Bookmark not defined.</b>
Tabel 3. 1 Spesifikasi Komputer.....	<b>Error! Bookmark not defined.</b>
Tabel 3. 2 Jadwal Penelitian.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 1 Data Nilai .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 2 Data Minat Siswa .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 3 Pengkategorian Minat .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 4 Konversi Minat .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 5 Konversi Jurusan TKR.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 6 Konversi Jurusan TSM .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 7 Konversi Jurusan RPL .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 8 Centroid Awal .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 9 Jarak Cluster.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 10 Hasil Centroid .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 11 Jarak Cluster.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 12 Centroid Baru.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 13 Jarak Cluster.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 14 Centroid Baru.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 15 Jarak Cluster.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 16 Centroid Baru.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 17 Hasil Cluster TKR.....	<b>Error! Bookmark not defined.</b>
Tabel 4. 18 Hasil Cluster TSM .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 19 Hasil Cluster RPL .....	<b>Error! Bookmark not defined.</b>
Tabel 4. 20 Hasil Rekomendasi .....	<b>Error! Bookmark not defined.</b>

## DAFTAR GAMBAR

Gambar 2. 1 Tahapan Knowledge Discovery in Database ..	<b>Error! Bookmark not defined.</b>
Gambar 2. 2 Lambang Anaconda Navigation.....	<b>Error! Bookmark not defined.</b>
Gambar 2. 3 Lambang Jupyter .....	<b>Error! Bookmark not defined.</b>
Gambar 2. 4 Kerangka Berfikir.....	<b>Error! Bookmark not defined.</b>
Gambar 3. 1 Peta Lokasi SMK Muhammadiyah Paguyangan....	<b>Error! Bookmark not defined.</b>
Gambar 3. 2 Tahapan Penelitian .....	<b>Error! Bookmark not defined.</b>
Gambar 3. 3 Tahap Analisis Data .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 1 Dataset Jurusan TKR.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 2 Visualisai Dataset Jurusan TKR.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 3 Cluster awal.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 4 Distribusi Awal .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 5 Update Centroid .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 6 Hasil Cluster Akhir .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 7 Distribusi Akhir.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 8 Distribusi Cluter .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 9 Dataset Jurusan TSM .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 10 Visualisai Dataset Jurusan TSM .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 11 Cluster Awal.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 12 Distribusi Awal .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 13 Update Centroid .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 14 Hasil Cluster Akhir.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 15 Distribusi Akhir.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 16 Distribusi Cluster .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 17 Dataset Jurusan RPL .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 18 Visualisai Dataset Jurusan RPL .....	<b>Error! Bookmark not defined.</b>
Gambar 4. 19 Cluster Awal.....	<b>Error! Bookmark not defined.</b>
Gambar 4. 20 Distribusi Awal .....	<b>Error! Bookmark not defined.</b>

Gambar 4. 21 Update Centroid .....**Error! Bookmark not defined.**  
Gambar 4. 22 Hasil Cluster Akhir.....**Error! Bookmark not defined.**  
Gambar 4. 23 Distribusi Akhir.....**Error! Bookmark not defined.**  
Gambar 4. 24 Distribusi Cluster .....**Error! Bookmark not defined.**

## DAFTAR LAMPIRAN

Lampiran 1 Curriculum Vitae.....**Error! Bookmark not defined.**

Lampiran 2 Data Nilai Awal.....**Error! Bookmark not defined.**