

**PERBANDINGAN AKURASI METODE CLUSTERING  
ALGORITMA K-MEANS DENGAN ALGORITMA K-  
MEDOIDS DALAM PENGELOMPOKAN DATA MAHASISWA  
BARU UNTUK STRATEGI PROMOSI PROGRAM STUDI  
TEKNIK INFORMATIKA UNISNU JEPARA**



**SKRIPSI**

Diajukan sebagai Salah Satu Syarat untuk Memperoleh  
Gelar Sarjana Strata 1 (S.1) Program Studi Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Nahdlatul Ulama Jepara

Oleh :

**Aris Darussalam**

**NIM. 161240000480**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NAHDLATUL ULAMA JEPARA  
2021**

## PERSETUJUAN PEMBIMBING

*Assalamualaikum Wr. Wb.*

Setelah saya meneliti dan mengadakan perbaikan seperlunya, bersama ini saya kirim naskah Skripsi Saudara :

Nama : Aris Darussalam  
Nim : 161240000480  
Program Studi : Teknik Informatika  
Judul : Perbandingan Akurasi Metode Clustering Algoritma K-Means dengan Algoritma K-Medoids dalam Pengelompokan Data Mahasiswa Baru untuk Strategi Promosi Program Studi Teknik Informatika UNISNU Jepara.

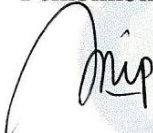
Skripsi ini telah disetujui pembimbing dan siap untuk dipertahankan dihadapan Dewan Penguji Sarjana Strata 1 (S1) Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama Jepara.

Demikian harap menjadi maklum.

*Wassalamualaikum Wr. Wb.*

Jepara, 02 Agustus 2021

Pembimbing 1,



Akhmad Khanif Zyen, M.Kom.

NIDN. 0621048602

Pembimbing 2,



Buang Budi Wahono, S.Si., M.Kom.

NIDN. 0603087802

Mengetahui,

Ketua Program Studi Teknik Informatika,



Gentur Wahyu Niyoto Wibowo, S.Kom., M.Kom

NIDN. 0623117902

## PENGESAHAN

Skripsi dengan judul “ Perbandingan Akurasi Metode Clustering Algoritma K-Means dengan Algoritma K-Medoids dalam Pengelompokan Data Mahasiswa Baru untuk Strategi Promosi Program Studi Teknik Informatika UNISNU Jepara”

Karya :

Nama : Aris Darussalam

NIM : 161240000480

Program Studi : Teknik Informatika


Telah diujikan dan dipertahankan dalam sidang oleh Dewan Penguji Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama Jepara dan dinyatakan lulus pada tanggal : 26 Agustus 2021.

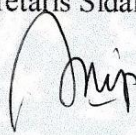
Selanjutnya dapat diterima sebagai syarat guna memperoleh gelar Sarjana Strata 1(S1) Program Studi Teknik Informatika pada Fakultas Sains dan Teknologi UNISNU Jepara Tahun Akademik 2020/2021.

Jepara, 26 Agustus 2021

Ketua Sidang,


Sekretaris Sidang,

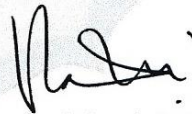
  
Buang Budi Wahono, S.Si., M.Kom.  
NIDN. 0603087802

  
Akhmad Khanif Zyen, M.Kom.  
NIDN. 0621048602

Penguji I,

Penguji II,


  
Sarwido, SE., M.M.  
NIDN. 0625016902

  
Nadia Annisa Maori, M.Kom.  
NIDN. 0626069201

Mengetahui,

Dekan Fakultas Sains dan Teknologi

UNISNU Jepara,

  
Dias Prihatmoko, ST. M.Eng  
NIDN. 0612128302

## PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini :

Nama : Aris Darussalam

NIM : 161240000480


Program Studi: Teknik Informatika

Saya menyatakan dengan penuh kejujuran dan tanggung jawab, bahwa Skripsi yang saya susun sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata 1 (S1) Universitas Islam Nahdlatul Ulama Jepara merupakan hasil karya saya sendiri yang jauh dari plagiarism dan belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar Sarjana dari Perguruan Tinggi lain.

Adapun bagian-bagian tertentu dalam penelitian Skripsi yang saya kutip dari karya orang lain telah dituliskan sumbernya secara jelas dengan norma, kaidah, dan etika penelitian ilmiah.

Selanjutnya saya bersedia menerima sanksi dari Fakultas Sains dan Teknologi UNISNU Jepara apabila dikemudian hari ditemukan ketidak benaran dari pernyataan ini.

Jepara, 02 Agustus 2021,



*Aris Darusslam*  
Aris Darusslam  
161240000480

## ABSTRAK

Aris Darussalam, 16124000480, Perbandingan Akurasi Metode Clustering Algoritma K-Means dengan Algoritma K-Medoids dalam Pengelompokan Data Mahasiswa Baru untuk Strategi Promosi Program Studi Teknik Informatika UNISNU Jepara, 2021, Akhmad Khanif Zyen, M.Kom, Buang Budi Wahono, S.Si., M.Kom, Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Nahdlatul Ulama Jepara.

Penelitian ini membahas mengenai perbandingan antara Algoritma *K-Means* dan Algoritma *K-Medoids* dalam klasterisasi, dan hasilnya bisa dijadikan sebagai pertimbangan dalam menangani strategi promosi prodi Teknik Informatika di UNISNU Jepara. Data yang digunakan dalam penelitian ini adalah data calon mahasiswa baru periode 2016 – 2019 yang didapat dari UPT Pusat Data UNISNU Jepara. Proses klasterisasi memerlukan beberapa data masukan meliputi asal kecamatan sekolah, asal sekolah, jenis kelamin dan nilai UN setelah dilakukan pembersihan dan perbaikan data. Klustering adalah metode pengelompokan data yang memiliki kemiripan yang sama dengan karakteristik yang sama dalam satu *cluster*. Penelitian ini menggunakan data sebanyak 478 data mahasiswa yang diolah menggunakan *Microsoft Excel* dan aplikasi *Rapidminer*. *Davies Boudin Index* (DBI) sebagai performance untuk pengujian validasi. Cluster yang terbentuk adalah 3 *cluster* dengan hasil akurasi algoritma *K-Means Clustering* sebesar 0,47 dan algoritma *K-Medoids* sebesar 0,416. Dan dapat di kategorikan cukup baik dalam menentukan strategi promosi prodi Teknik Informatika UNSNU Jepara.

**Kata Kunci** : Data Mining, K-Means Clustering, K-Medoids Clustering, Perbandingan algoritma, Rapid Miner.

## ABSTRACT

Aris Darussalam, 16124000480, Comparison of the Accuracy of the K-Means Algorithm Clustering Method with the K-Medoids Algorithm in New Student Data Grouping for the Promotion Strategy of the Informatics Engineering Study Program UNISNU Jepara, 2021, Akhmad Khanif Zyen, M.Kom, Buang Budi Wahono, S.Si ., M. Kom, Informatics Engineering, Faculty of Science and Technology, Jepara Nahdlatul Ulama Islamic University.

This study discusses the comparison between the K-Means Algorithm and the K-Medoids Algorithm in clustering, and the results can be used as a consideration in dealing with the promotion strategy of the Informatics Engineering study program at UNISNU Jepara. The data used in this study is data for prospective new students for the period 2016 – 2019 obtained from the UPT Data Center UNISNU Jepara. The clustering process requires some input data including the origin of the school district, school origin, gender and UN scores after cleaning and correcting the data. Clustering is a method of grouping data that have the same similarities with the same characteristics in one cluster. This study uses 478 student data which is processed using Microsoft Excel and the Rapidminer application. Davies Boudin Index (DBI) as a performance for validation testing. The clusters formed are 3 clusters with the results of the accuracy of the K-Means Clustering algorithm of 0.47 and the K-Medoids algorithm of 0.416. And it can be categorized quite well in determining the promotion strategy of the Informatics Engineering study program at UNSNU Jepara.

**Keywords:** Data Mining, K-Means Clustering, K-Medoids Clustering, Comparison of the algoritmh, Rapid Miner.

## **MOTTO**

“Ilmu adalah yang memberikan manfaat, bukan yang sekedar hanya dihafal.”

**(Imam Syafi’i).**

“Sistem pendidikan yang bijaksana setidaknya akan mengajarkan kita betapa sedikitnya yang belum diketahui oleh manusia, seberapa banyak yang masih harus ia pelajari.”

**(Sir John Lubbock).**

“Kesuksesan adalah saat persiapan dan kesempatan bertemu.”

**(Bobby Unser)**

## **PERSEMBAHAN**

Skripsi ini kupersembahkan kepada:

1. Diri sendiri yang sudah mau berjuang dan menerjang segala rintangan hingga penyusunan Skripsi ini selesai dengan baik.
2. Kedua orang tua yang selalu mendoakan setiap waktu, memberikan motivasi, semangat serta dukungan baik segi material maupun moril selama menempuh pendidikan di bangku perkuliahan hingga saat ini.
3. Keluarga besar, teman, dan semua orang yang tidak bisa disebutkan satu persatu yang telah membantu peneliti dalam menyelesaikan penyusunan Skripsi ini.



## KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat, taufiq hidayah, dan inayah-NYA, sehingga peneliti dapat menyelesaikan Skripsi dengan judul: “Perbandingan Akurasi Metode Clustering Algoritma K-Means dengan Algoritma K-Medoids dalam Pengelompokan Data Mahasiswa Baru untuk Strategi Promosi Program Studi Teknik Informatika UNISNU Jepara.” dengan baik.

Terlepas dari itu, peneliti juga menyadari adanya keterbatasan pengetahuan yang dimiliki, oleh karenanya tanpa keterlibatan berbagai pihak kemungkinan akan sulit bagi peneliti dalam menyelesaikan penyusunan Skripsi ini. Dengan segala kerendahan hati serta rasa syukur, peneliti mengucapkan terima kasih kepada:

1. Rektor Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, Bapak Dr. H. Sa’dullah Assaidi, M.Ag., yang telah memberikan ilmu pengetahuan sehingga menjadikan peneliti bersemangat dalam menempuh pendidikan.
2. Dekan Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, Bapak Dias Prihatmoko, ST. M.Eng, yang telah memfasilitasi dan memberikan kemudahan kepada peneliti sehingga dapat menyelesaikan perkuliahan dan Skripsi dengan baik.
3. Kaprodi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, Bapak Gentur Wahyu Nyipto Wibowo, S.Kom., M.Kom, yang telah memberikan arahan dan membimbing peneliti selama menempuh pendidikan.
4. Pembimbing I Bapak Akhmad Khanif Zyen, S.Kom., M.Kom. dan pembimbing II Bapak Buang Budi Wahono, S.Si., M.Kom. dengan penuh kesabaran dan tanggungjawab beliau yang berkenan membimbing peneliti dalam menyusun Skripsi ini.
5. Bapak dan Ibu Dosen Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, yang telah memberikan ilmu pengetahuan serta wawasannya kepada peneliti tanpa pamrih.

6. Teman angkatan tahun 2016 Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Nahdlatul Ulama (UNISNU) Jepara, yang ikut serta membantu peneliti.

Peneliti juga menyadari bahwa penulisan Skripsi ini masih jauh dari kata sempurna, maka dari itu kritik dan saran sangat diperlukan guna penyempurnaan. Skripsi ini nantinya. Peneliti berharap Skripsi ini dapat memberikan manfaat, baik untuk peneliti maupun pembaca.

Jepara, 02 Agustus 2021

Peneliti,

**Aris Darussalam**

**NIM. 161240000480**

## DAFTAR ISI

JUDUL.....	i
PERSETUJUAN PEMBIMBING.....	<b>Error! Bookmark not defined.</b>
PENGESAHAN.....	vi
PERNYATAAN KEASLIAN .....	<b>Error! Bookmark not defined.</b>
ABSTRAK .....	vii
MOTTO.....	x
PERSEMBAHAN.....	xi
KATA PENGANTAR.....	xii
DAFTAR ISI .....	xiv
DAFTAR TABEL.....	xvi
DAFTAR GAMBAR .....	xviii
DAFTAR LAMPIRAN .....	xix
BAB I PENDAHULUAN .....	1
1.1. Latar Belakang Masalah.....	1
1.2. Batasan Masalah .....	4
1.3. Rumusan Masalah.....	4
1.4. Tujuan Penelitian .....	5
1.5. Manfaat Penelitian .....	5
1.6. Sistematika Penulisan.....	5
BAB II LANDASAN TEORI.....	7
2.1. Tinjauan Studi .....	7
2.2. Tinjauan Pustaka .....	9
2.2.1. Definisi Strategi Promosi .....	9
2.2.2. Data Mining.....	13
2.2.3. Clustering .....	19
2.2.4. Algoritma Clustering K-Means. ....	20
2.2.5. Algoritma K-Medoids.....	22
2.2.6. Davies Bouldin Index .....	22
2.2.7. RapidMiner.....	24
2.2.8. Metode ROC ( <i>Receiver Operating Characteristic</i> ).....	26
2.2.9. Euclidean Distance .....	27
2.3. Kerangka Pemikiran .....	27
BAB III METODE PENELITIAN .....	29
3.1. Desain Penelitian.....	29

3.2. Metode Pengumpulan Data.....	30
3.3. Tahapan Pengolahan Awal Data .....	30
3.4. Model yang Diusulkan.....	31
3.4.1. Algoritma K-Means Clustering. ....	32
3.4.2. Algoritma <i>K-Medoids Clustering</i> .....	35
BAB IV HASIL PENELITIAN DAN PEMBAHASAN .....	39
4.1. Pre-Processing Data.....	39
4.1.1. Data Reduction .....	39
4.1.2. Data Cleaning .....	40
4.1.3. Data Transformation .....	40
4.1.4. Data Intregation .....	42
4.2. Penerapan Algoritma Menggunakan Aplikasi <i>Microsoft Excel</i> . ....	42
4.2.1. Penerapan Algoritma <i>K-Means</i> Menggunakan Aplikasi <i>Microsoft Excel</i> .....	42
4.2.2. Penerapan Algoritma <i>K-Medoids</i> Menggunakan Aplikasi <i>Microsoft Excel</i> ....	49
4.3. Penerapan Algoritma menggunakan aplikasi <i>Rapidminer Studio</i> . ....	56
4.4. Evaluasi.....	66
4.4.1. Pengujian Metode <i>Davies Bouldin Index</i> .....	66
4.4.2. Pengujian menggunakan Metode ROC .....	73
4.4.2. Perbandingan Hasil Perhitungan.....	74
BAB V_PENUTUP.....	76
5.1. Kesimpulan. ....	76
DAFTAR PUSTAKA .....	77
LAMPIRAN-LAMPIRAN .....	79

## DAFTAR TABEL

Tabel 2. 1. Tabel Metode ROC [20] .....	26
Tabel 3.1. Sampel Data Calon Mahasiswa Baru Teknik Informatika. ....	31
Tabel 4.1. Inisialisasi Data Asal Kecamatan Sekolah.....	41
Tabel 4. 2. Inisialisasi Data Jenis Kelamin .....	41
Tabel 4. 3. Inisialisasi Data Asal Sekolah.....	42
Tabel 4. 4. Titik Pusat Awal dari setiap Cluster .....	43
Tabel 4. 5. Centroid Baru Iterasi ke-1 .....	45
Tabel 4. 6. Centroid Baru Iterasi ke-2.....	46
Tabel 4.7. Centroid Baru Iterasi ke-3 (terakhir) .....	47
Tabel 4. 8. Hasil Cluster 0 Penerapan Algoritma K-Means (Microsoft Excel).....	47
Tabel 4. 9. Hasil Cluster 1 Penerapan Algoritma K-Means (Microsoft Excel).....	48
Tabel 4. 10. Hasil Cluster 2 Penerapan Algoritma K-Means (Microsoft Excel)....	49
Tabel 4. 11. Titik Pusat Awal dari setiap Cluster K-Medoids .....	50
Tabel 4. 12. Sampel Hasil Perhitungan Jarak ke Setiap Medoids. ....	51
Tabel 4. 13. Titik Pusat Baru (Centroid) Non Medoids Iterasi ke -1. ....	51
Tabel 4. 14. Sampel Hasil Perhitungan Jarak ke Setiap Medoids Iterasi ke -1.....	52
Tabel 4. 15. Titik Pusat Baru (Centroid) Non Medoids iterasi ke -2.....	53
Tabel 4. 16. Sampel Hasil Perhitungan Jarak ke Setiap Medoids Iterasi ke -2.....	53
Tabel 4. 17. Hasil Cluster 0 Penerapan Algoritma K-Medoidss (Microsoft Excel). .....	54
Tabel 4. 18. Hasil Cluster 1 Penerapan Algoritma K-Medoids (Microsoft Excel). 55	
Tabel 4. 19. Hasil Cluster 2 Penerapan Algoritma K-Medoidss (Microsoft Excel). .....	56
Tabel 4. 20. Hasil Cluster 0 Penerapan Algoritma K-Means (Rapid Miner).....	63
Tabel 4. 21. Hasil Cluster 1 Penerapan Algoritma K-Means (Rapid Miner).....	64
Tabel 4. 22. Hasil Cluster 2 Penerapan Algoritma K-Means (Rapid Miner).....	64
Tabel 4. 23. Hasil Cluster 0 Penerapan Algoritma K-Medoidss (Rapid Miner) ..	65
Tabel 4. 24. Hasil Cluster 1 Penerapan Algoritma K-Medoids (Rapid Miner).....	65
Tabel 4. 25. Hasil Cluster 2 Penerapan Algoritma K-Medoidss (Rapid Miner) ..	66
Tabel 4. 26. Nilai Centroid Iterasi Terakhir K-Means.....	66
Tabel 4. 27. Sampel hitung jarak anggota Cluster O hasil Clustering k-means....	67
Tabel 4. 28. Sampel hitung jarak anggota Cluster 1 hasil Clustering k-means.....	67
Tabel 4. 29. Sampel hitung jarak anggota Cluster 2 hasil Clustering k-means.....	68
Tabel 4. 30. Hasil Perhitungan SSB Algoritma K-means.....	69
Tabel 4. 31. Nilai Centroid Iterasi Terakhir K-Medoids.....	70
Tabel 4. 32. Sampel hitung jarak anggota Cluster O hasil Clustering k-medoids. 71	
Tabel 4. 33. Sampel hitung jarak anggota Cluster 1 hasil Clustering k-medoids. 71	
Tabel 4. 34. Sampel hitung jarak anggota Cluster 2 hasil Clustering k-medoids. 71	

Tabel 4. 35. Hasil Perhitungan SSB Algoritma K-medoids.....	72
Tabel 4. 36. Nilai Centroid Data PMB Periode 2016 – 2019 pada Algoritma K-Means Clustering. ....	73
Tabel 4. 37. Nilai Centroid Data PMB Periode 2016 – 2019 pada Algoritma K-Medoids Clustering. ....	74
Tabel 4. 38. Nilai Rasio .....	74
Tabel 4. 39. Hasil Perbandingan Perhitungan Microsoft Excel dan RapidMiner	75

## DAFTAR GAMBAR

Gambar 2. 1. Tahapan Knowledge Discovery Database .....	16
Gambar 2. 2. Konsep Clustering .....	20
Gambar 2. 3. Interface Rapidminer Studio .....	25
Gambar 2. 4. Kerangka Pemikiran .....	28
Gambar 3.1. Diagram Proses Pengolahan Data Algoritma K-means Clustering ..	33
Gambar 3.2. Flowchart Algoritma K-means Clustering .....	35
Gambar 3. 3. Diagram Proses Pengolahan Data Algoritma K-Medoids. ....	36
Gambar 3. 4. Flowchart Algoritma K-Medoids .....	38
Gambar 4. 1. Sampel Dataset Calon Mahasiswa Baru sebelum dilakukan Pre- processing Data.....	39
Gambar 4. 2. Sampel Dataset Setelah Tahapan Reduksi .....	40
Gambar 4. 3. Sampel Menghitung Jarak Data ke setiap Cluster pada Microsoft Excel.....	44
Gambar 4. 4. Sampel Pengelompokkan Data ke dalam Cluster iterasi ke -1 pada Microsoft Excel. ....	45
Gambar 4. 5. Sampel Pengelompokkan Data ke dalam Cluster iterasi ke -2 pada Microsoft Excel. ....	45
Gambar 4. 6. Sampel Pengelompokkan Data ke dalam Cluster iterasi ke -3 pada Microsoft Excel. ....	46
Gambar 4. 7. Sampel Hasil Pengklasteran Data dengan K-Medoids Clustering ..	54
Gambar 4. 8. Proses impor dataset ke Rapidminer Studio algoritma K-means. ...	57
Gambar 4. 9. Desain Clustering Menggunakan Algoritma K-Means Clustering..	58
Gambar 4. 10. Desain Clustering Menggunakan Algoritma K-Medoids Clustering. .....	58
Gambar 4. 11. Result History menggunakan Algoritma K-Means Clustering .....	59
Gambar 4. 12. Result History menggunakan Algoritma K-Medoids Clustering. .	59
Gambar 4. 13. Parameter Metode Davies Bouldin Indeks Algoritma Clustering .	59
Gambar 4. 14. Davies Bouldin Indeks (DBI) Algoritma K-Means Clustering .....	60
Gambar 4. 15. Davies Bouldin Indeks (DBI) untuk Performance Vector Algoritma K-Medoids Clustering.....	60
Gambar 4. 16. Hasil DBI dan Cenroid Distance Performance Vector Algoritma K- Means Clustering. ....	60
Gambar 4. 17. Hasil DBI dan Cenroid Distance Algoritma K-Medoids Clustering. .....	61
Gambar 4. 18. Penentuan Jumlah Cluster. ....	61
Gambar 4. 19. Hasil Cluster Algoritma K-Means Clustering. ....	61
Gambar 4. 20. Hasil Cluster Algoritma K-Medoids Clustering. ....	62
Gambar 4. 21. Tampilan Chart Bar Algoritma K-Means Clustering. ....	62
Gambar 4. 22. Tampilan Chart Bar Algoritma K-Medoids Clustering. ....	63

## DAFTAR LAMPIRAN

Lampiran 1. Buku Bimbingan Skripsi.....	80
Lampiran 2. Dataset Calon Mahasiswa Baru Periode 2016 - 2019.....	82