

DATA MINING SELEKSI SISWA BERPRESTASI UNTUK MENENTUKAN KELAS UNGGULAN MENGGUNAKAN METODE K-MEANS CLUSTERING (Studi Kasus di MTS Darul Fikri)

Reyhan Putra Primanda, Aslan Alwi, Dyah Mustikasari

Prodi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Ponorogo

E-mail Kosrespondensi : reyhanprimanda@gmail.com

History Artikel

Diterima : 18 Februari 2021 Disetujui : 04 Maret 2021 Dipublikasikan : 08 April 2021

Abstract

MTs Darul Fikri Bringin is a Tsanawiyah Madrasah located in Ponorogo. This madrasah has a good learning system to support the quality of madrasah, one of which is a superior class program. MTs Darul Fikri Bringin is always experiencing an increase in the acceptance of new student registrants every year. In determining the superior class, there are difficulties in class clustering based on students' abilities. This is due to the increasing number of registrants each year. Therefore, in this study the data mining method is applied to help cluster superior classes with student grade variables using the K-Means Clustering Algorithm and Rstudio Tools. Where the attributes used are the report card value, activeness and presence of class VII students. The K-Means Clustering Algorithm method is used to process these 12 attributes so as to produce 2 regular clusters and 1 superior cluster with the Rshiny interface in Rstudio using the R programming language. A in cluster 1 has 23 members. Classes B, C, D include regular classes which include class B in cluster 2 totaling 25 members and classes C and D in cluster 3 totaling 52 members. The application of the K-Means Clustering Algorithm and Rstudio Tools can cluster the determination and selection of superior classes at MTs Darul Fikri.

Keywords: *MTs Darul Fikri, Data Mining, k-means Clustering dan Tools Rstudio*

Abstrak

MTs Darul Fikri Bringin merupakan Madrasah Tsanawiyah yang berada di Ponorogo. Madrasah ini memiliki sistem pembelajaran yang baik untuk menunjang kualitas maderasah, salah satunya dengan program kelas unggulan. MTs Darul Fikri Bringin setiap tahun selalu mengalami peningkatan dalam penerimaan pendaftar siswa baru. Pada penentuan kelas unggulan terjadi kesulitan klasterisasi kelas yang berdasarkan pada kemampuan siswa. Hal ini disebabkan karena meningkatnya pendaftar di setiap tahunnya. Oleh karena itu pada penelitian ini metode data mining diterapkan untuk membantu klasterisasi kelas unggulan dengan variabel nilai siswa menggunakan Algoritma K-Means Clustering dan Tools Rstudio. Dimana atribut yang digunakan adalah nilai raport, keaktifan dan kehadiran siswa kelas VII. Metode Algoritma K-Means Clustering digunakan untuk mengolah 12 atribut tersebut sehingga menghasilkan 2 klaster regular dan 1 klaster unggulan dengan antar muka Rshiny di Rstudio menggunakan Bahasa pemograman R. Berdasarkan penelitian ini dari 100 data siswa didapatkan hasil bahwa siswa yang berada dalam kelas unggulan yaitu kelas A di klaster 1 berjumlah 23 anggota. Kelas B, C, D termasuk kelas regular yang meliputi kelas B di klaster 2 berjumlah 25 anggota dan kelas C dan D di klaster 3 berjumlah 52 anggota. Penerapan Algoritma K-Means Clustering dan Tools Rstudio dapat mengklaster penentuan dan penyeleksian kelas unggulan di MTs Darul Fikri.

Kata Kunci : *MTs Darul Fikri, Data Mining, k-means Clustering dan Tools Rstudio*

How to Cite : **Primanda, Reyhan Putra** (2021). *Data mining seleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan Metode K-Means Clustering (Studi kasus di MTs Darul Fikri).* KOMPUTEK : Jurnal Teknik Universitas Muhammadiyah Ponorogo, 5(1): Halaman 88-100

© 2021 Universitas Muhammadiyah Ponorogo. All rights reserved

ISSN 2614-0985 (Print)

ISSN 2614-0977 (Online)

1. PENDAHULUAN

Pada saat ini perkembangan data science berkembang sangat pesat seperti halnya kebutuhan manusia. Dalam hal ini data science memiliki fungsi signifikan terutama dalam hal pengolahan data. Pengolahan data berdasarkan data science dapat digunakan untuk mengklastering siswa kelas reguler dan unggulan sehingga mempermudah dalam penentuan kelas. Pondok Pesantren Darul Fikri didirikan sejak tahun 1991 M yang dikelola oleh KH. Ahmad Juhaimi, Lc dan diresmikan oleh Bapak Bupati Ponorogo kala itu yaitu Bapak Drs. Gatot Sumani. Pada awal berdirinya Pondok ini hanya terdiri dari Madrasah Tsanawiyah yang setara dengan SMP dan Madrasah Aliyah yang setara dengan SMA, akan tetapi sekarang ini sudah dikembangkan dengan adanya Tarbiyatul Atfal, Madrasah Ibtidaiyah (MI) yang setara dengan SD dan Sekolah Menengah Kejuruan otomotif. [1]

MTs Darul Fikri Bringin terletak di Desa Bringin tepatnya di Kecamatan Kauman, Kabupaten Ponorogo. MTs Darul Fikri saat ini di kepalai oleh Yanuri, S.Pd.I dengan Akreditasi A pada tahun 2015 dengan memiliki sistem pembelajaran yang baik dan program siswa kelas unggulan. Setiap tahunnya penerimaan pendaftaran Permasalahan tentang bertambahnya jumlah

siswa baru di MTs Darul Fikri meningkat dan menjadikan MTs Darul Fikri menjadi Madrasah tsanawiyah cukup terkenal di ponorogo.[1]

Bertambahnya jumlah siswa pendaftar menyebabkan kesulitan dalam klasterisasi kelas unggulan dan penyeleksian siswa berdasarkan kemampuannya. Oleh karena itu perlu dilakukan seleksi untuk mempermudah dalam pengelompokkan siswa kelas unggulan di gunakan penerapan Data Mining. Data mining merupakan bagian dari proses KDD (Knowledge Discovery in Database) KDD disebut juga sebagai proses menemukan pola-pola dalam data. proses ini dapat diklasifikasi, diprediksi, diperkirakan dan mendapatkan informasi baru dari kumpulan data dalam jumlah yang cukup besar yang di dukung dengan Algoritma K-Means Clustering menggunakan penerapan Data Mining [2].

Metode K-Means Clustering merupakan metode untuk mengatasi masalah pengolahan data menjadi sumber informasi strategis, untuk mendukung pengambilan keputusan dan juga K-Means merupakan Algoritma Clustering yang berulang-ulang [3]

pendaftar menyebabkan kesulitan untuk mengklaster serta menyeleksi siswa kelas

unggulan dan pengelompokan siswa berdasarkan kemampuan di MTs Darul Fikri tersebut. Penelitian ini akan menyeleksi siswa berprestasi untuk menentukan kelas unggulan menggunakan Metode K-Means Clustering dengan antar muka Rshiny di Rstudio dengan Bahasa pemrograman R. Metode ini akan memprediksi atau menyeleksi suatu masalah dengan klasterisasi data berdasarkan langkah yang sudah ada serta pemilihan atribut sesuai dengan kebutuhan. Berdasarkan penelitian ini, penulis mengangkat judul “Data Mining Seleksi Siswa Berprestasi Untuk Menentukan Kelas Unggulan Menggunakan Metode K-Means Clustering (Studi Kasus Di Mts Darul Fikri)”.

2. KAJIAN LITERATUR

a. Data Mining

Data Mining adalah pengambilan keputusan suatu informasi dengan pola yang terdapat pada sekumpulan data yang ada di database dengan jumlahnya cukup besar. Terdapat juga di suatu jurnal ilmiah, KDD adalah nama lain dari data mining. Data Mining terdapat bidang ilmu, yaitu Statistik, Database System, Data Warehousing, Information Retrieval, dan Machine Learning. [9]

b. Clustering

Clustering termasuk kelompok homogen, yaitu setiap bagian mempunyai

kesamaan satu sama lain. kualifikasi dari kluster harus memiliki data yang homogen dan tidak sama dengan data yang lainnya, sedangkan data yang berbeda akan berkumpul dengan kluster yang berbeda pula. Adapun contoh Aplikasi Clustering antara lain: Medicine Sciences, Bioinformatics, Social Sciences, Engineering, Data dan Web Mining.[15]

c. Algoritma K-Means

Pengelompokan data adalah teknik eksplorasi suatu kumpulan data yang memiliki objek karakteristik yang sama dimana di kelompokkan menjadi satu untuk mengklasifikasi pengolahan data lebih lanjut.[16]

Algoritma k-means adalah metode cluster analysis non hirarki yang yang mengelompokkan suatu data berdasarkan karakteristiknya kedalam satu atau lebih cluster, jadi data yang memiliki kesamaan dikelompokkan dalam suatu cluster yang sama dan sebaliknya data yang memiliki kesamaan yang tidak sama dikelompokkan kedalam cluster yang tidak sama pula.

Dalam *Algoritma K-Means* terdapat beberapa pendekatan dalam mengembangkan metode clustering, dua pendekatan utama yaitu *Algoritma Parsial* dan *Algoritma Hierarki*. *Algoritma Parsial* adalah algoritma yang Membuat beberapa partisi untuk di kelompokkan menjadi

beberapa objek berdasarkan kriteria tertentu sedangkan algoritma hierarki adalah algoritma yang Membuat suatu kelompok objek berdasarkan kesamaanya. Misalnya = tua-muda, merokok-tidak merokok. [17]

Berikut adalah rumus penghitungan jarak dalam Algoritma K-Means :

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \dots\dots(1)$$

dimana:

D (i,j) = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

d. Microsoft Excel 2016

Microsoft Corporation memproduksi dan menjual sebuah *Software Microsoft Office* yang di dalamnya terdapat *Software* berupa *Microsoft Excel, Microsoft Word, Microsoft Outlook, Microsoft PowerPoint,* dan lainnya. Dalam dunia kerja *Microsoft Excel* menjadi sebuah *Software* yang terkenal sebagai multifungsi dalam mengolah data berupa angka, di dalam *Microsoft Excel* juga terdapat kombinasi cell untuk memudahkan user dalam mengolah data. selain itu *Microsoft Excel* mendapat pengakuan dari berbagai profesi akan kemudahan dan kecepatan dalam menangani sebuah masalah yang di hadapi user dalam mengolah data berupa angka entah itu pekerjaan yang ringan maupun pekerjaan berat sekalipun untuk mendapatkan hasil yang maksimal.[21]

e. Pemograman R dan Rstudio

Pemograman R merupakan GNU komputasi statistik dan grafik yang dikembangkan oleh *Bell Laboratories* oleh *John Chamber* dan teman-temannya yang sebelumnya di buat pada tahun 1992 di Universitas Auckland, New Zealand oleh Ross Ihaka dan Robert Gentleman. Dalam pemograman R terdapat berbagai tools statistik linier, non linier, clustering, klasifikasi, uji statistik klasik dan Tool teknis grafis yang menampilkan data secara visual berbentuk grafik. Pemograman R juga merupakan program yang open source dalam mengolah data Statistika berbentuk grafik yang sangat berguna untuk penelitian maupun industri. [22]

3. METODE PENELITIAN

a. Metode Pengumpulan Data

Penelitian dan pengumpulan data ini dilaksanakan pada tanggal 4 juni 2020, penelitian ini mendapatkan data berupa nilai siswa, nilai keaktifan dalam proses belajar dan nilai kehadiran. Akan tetapi data-data yang telah didapatkan masih berupa data mentah yang masih perlu diseleksi untuk mendapatkan atribut yang sesuai dengan penelitan tersebut. Dalam pengumpulan data tersebut beberapa pihak yang terkait dalam penelitian yaitu kepala maderasah dan wali kelas VII

Dalam penelitian ini menggunakan

menberbagai sumber data yang diperlukan yaitu :

➤ **Data Primer**

Data primer didapatkan secara langsung dari sumber data yang berhubungan dengan penelitian yang dilakukan di MTs Darul Fikri.

➤ **Data Sekunder**

Data sekunder didapatkan dari berbagai buku, jurnal, yang bersumber dari instansi atau internet sebagai penunjang landasan penelitian. Pada tahap ini hal yang dilakukan adalah mencari, mengumpulkan, dan mempelajari data-data yang berkaitan dengan penyeleksian siswa kelas unggulan MTs Darul Fikri. Data yang didapatkan kemudian diolah menggunakan algoritma k-means clustering dengan mengambil nilai dari setiap atribut pada data untuk menentukan kelas siswa unggulan.

b. Analisis Data

Analisis data yaitu proses mencari dan menyusun suatu data yang diperoleh dari data primer dan data sekunder. Tidak semua data yang di peroleh di gunakan hanya data yang cocok untuk menunjang penelitian saja yang di gunakan.

Data yang digunakan untuk di proses pada tahap selanjutnya menggunakan atribut sebagai berikut :

Tabel 3. 1 Data Atribut Siswa Kelas VII

| | | |
|--|-------------------------|--|
| Aqidah akhlaq | Atribut nilai pelajaran | Aqidah akhlaq |
| Sejarah kebudayaan islam | Atribut nilai pelajaran | Sejarah kebudayaan islam |
| Al-quran hadist | Atribut nilai pelajaran | Al-quran hadist |
| Bahasa arab | Atribut nilai pelajaran | Bahasa arab |
| Bahasa Indonesia | Atribut nilai pelajaran | Bahasa Indonesia |
| Bahasa Inggris | Atribut nilai pelajaran | Bahasa Inggris |
| Ilmu pengetahuan alam | Atribut nilai pelajaran | Ilmu pengetahuan alam |
| Ilmu pengetahuan social | Atribut nilai pelajaran | Ilmu pengetahuan social |
| Matematika | Atribut nilai pelajaran | Matematika |
| Pendidikan pancasila dan kewarganegaraan | Atribut nilai pelajaran | Pendidikan Pancasila dan kewarganegaraan |
| keaktifan peroses belajar | Atribut Nilai | keaktifan peroses belajar |
| Kehadiran siswa | Atribut nilai | Kehadiran siswa |

c. Tahapan penelitian

Tahapan penelitian ini megunakan landasan sebagai berikut:

➤ **Studi Pustaka**

Mempelajari penelitian terdahulu mengenai data mining dan metodenya sebagai landasan untuk penelitian ini.

➤ **Pengumpulan data siswa**

Pengumpulan data diperoleh dengan cara wawancara, observasi, dll untuk mendapatkan data yang digunakan dalam penelitian ini.

- Nilai Pelajaran
- Nilai keaktifan dalam peroses belajar
- Kehadiran

➤ **Seleksi data**

Data yang diperoleh dari MTs Darul Fikri masih berupa data mentah yang perlu di lakukanya penyeleksian data.

➤ Perhitungan *Algoritma K-Means Clustering*

Algoritma k-means clustering iyalah suatu metode cluster analisis non hirarki yang mengelompokkan data berdasarkan karakteristiknya yang sama.

➤ Hasil Clustering

Memberikan hasil dari pengelompokan siswa kelas unggulan.

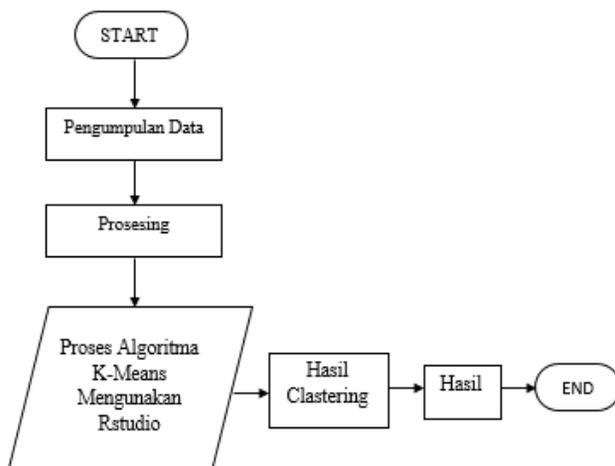
➤ Penyimpulan Hasil

Menyimpulkan hasil penelitian.

➤ Flowchart penelitian

Flowchart di buat bertujuan untuk mengetahui proses atau algoritma yang terjadi dalam penyeleksian siswa berprestasi dalam bentuk symbol – symbol.

Tabel 3. 4 Flowchart Penelitian Secara Umum



4. HASIL DAN PEMBAHASAN

a. Implementasi Algoritma K-Means

Dengan Rstudio

RStudio adalah software pengembangan yang terintegrasi (IDE) untuk melakukan eksekusi kode secara langsung dan menganalisis data untuk memberikan hasil yang di inginkan. penelitian ini menggunakan Algoritma K-Means Clustering untuk memecahkan suatu masalah. Perhitungan Algoritma K-Means memerlukan instalasi Packages Cluster, Faktoextra Dan Tidyverse untuk bisa menganalisis data.

➤ **Langkah Pertama**

Mengetikkan script berikut untuk menginstall dan memanggil packages dalam proses data mining.

```

Install.packages ("cluster")
Install.packages ("faktoextra")
Install.packages ("tidyverse")
# package cluster --- untuk clustering
library(cluster)
#package faktoextra --- untuk clustering dan visualisasi
library(faktoextra)
#package tidyverse -----untuk clustering
library(tidyverse)
  
```

➤ **Langkah Ke-Dua**

Mengetikkan script berikut untuk melakukan input data exel ke dalam R Studio.

```
#import data excel ke rstudio
table1=read.delim("clipboard")
#melihat data setelah import di
console
table1
#menampilkan table di source
view(table1)
```

Fungsi View (table1) adalah untuk menampilkan hasil data excel setelah di importkan ke dalam rstudio, berikut untuk hasilnya:

| NO | NIS | NAMA | AA | SKI | QH | BA | BIN | BSG | IPA | IPS | MM | PKN | Keaktifan | Kelulusan |
|----|---------|---------------------|----|-----|----|----|-----|-----|-----|-----|----|-----|-----------|-----------|
| 1 | 7130281 | ABYU NAUFALANDYAH | 86 | 88 | 83 | 81 | 86 | 84 | 87 | 87 | 85 | 84 | 80 | 80 |
| 2 | 7978983 | ACHEL PRIMA RIDIQ | 86 | 85 | 83 | 81 | 87 | 84 | 86 | 86 | 86 | 84 | 70 | 80 |
| 3 | 7948878 | AHMAD FAZI ABOLLAH | 86 | 86 | 85 | 82 | 87 | 84 | 88 | 89 | 87 | 84 | 80 | 80 |
| 4 | 8712182 | AHMAD FAUSAN | 87 | 87 | 83 | 87 | 88 | 82 | 81 | 86 | 84 | 70 | 80 | 80 |
| 5 | 7843420 | AHMAD NAFIS SJ | 86 | 87 | 87 | 81 | 83 | 84 | 84 | 84 | 86 | 85 | 70 | 80 |
| 6 | 8792824 | AHMAL QADRIYUL FA | 86 | 86 | 88 | 81 | 87 | 84 | 83 | 84 | 86 | 84 | 70 | 70 |
| 7 | 7394207 | ARY AZZAMA MUG | 86 | 85 | 87 | 81 | 88 | 84 | 86 | 86 | 86 | 84 | 70 | 80 |
| 8 | 7524867 | BENTANG NUSRAWATORD | 86 | 87 | 87 | 81 | 87 | 84 | 85 | 85 | 86 | 84 | 70 | 80 |
| 9 | 5368788 | FATYUR ROHMANN | 86 | 85 | 86 | 82 | 87 | 88 | 83 | 82 | 86 | 84 | 70 | 80 |
| 10 | 8348844 | GUNGGAM RIDLO S | 86 | 87 | 87 | 81 | 87 | 84 | 85 | 85 | 87 | 83 | 70 | 80 |

➤ Langkah Ke-Tiga

Mengetikkan script berikut untuk melakukan tranformasi data dengan scale (sekala), Atribut yang digunakan berbentuk numerik dengan 12 atribut terhitung dari sebelah kiri.

```
# transformasi data dengan scale
tablefix=scale(table1[, 4:15])
# melihat hasil scele
view(tablefix)
```

Fungsi View (tablefix) adalah untuk menampilkan hasil data setelah tranformasi secale (sekala).

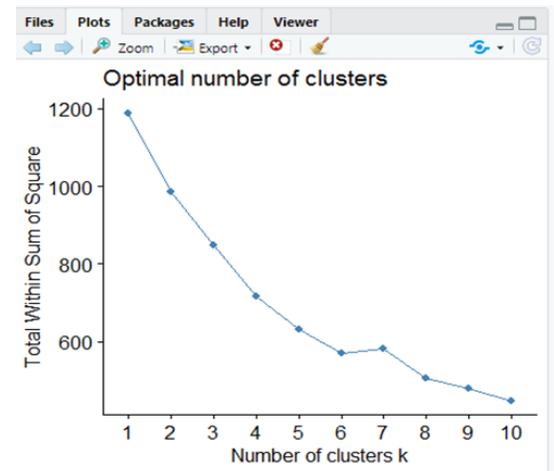
| AA | SKI | QH | BA | BIN | BSG | IPA | IPS | MM | PKN | Keaktifan | Kelulusan |
|------------|------------|------------|------------|------------|------------|------------|------------|-------------|-----------|------------|------------|
| -0.1537917 | 0.3385867 | -1.1951482 | -0.8779742 | -0.7897936 | -1.6052999 | 0.3765628 | 1.7007544 | -1.1242104 | -1.037346 | 1.3532435 | 0.4874937 |
| -0.1537917 | -0.4879307 | -1.1951482 | -0.8779742 | 0.05793843 | -1.6052999 | -0.2407526 | 1.0987811 | -0.08327786 | -1.037346 | -1.0224521 | 0.4874937 |
| -0.1537917 | 0.3385867 | -1.1951482 | -0.8779742 | 0.05793843 | -1.6052999 | 0.9938761 | 2.8927009 | 0.8978933 | -1.037346 | 1.3532435 | 0.4874937 |
| 0.5447202 | 1.5446461 | 0.7641111 | -0.8779742 | 0.05793843 | 0.5166889 | -2.7100201 | -1.9080853 | -0.08327786 | -1.037346 | -1.0224521 | 0.4874937 |
| -0.1537917 | 1.5446461 | 0.7641111 | -0.8779742 | -3.3528283 | -1.6052999 | -1.4731813 | -0.5021655 | -0.08327786 | -1.037346 | -1.0224521 | 0.4874937 |
| -0.1537917 | 0.3385867 | 1.7437428 | -0.8779742 | 0.05793843 | -1.6052999 | -2.0926957 | -0.5021655 | -0.08327786 | -1.037346 | -1.0224521 | -1.8899749 |
| -0.1537917 | -0.4879307 | 0.7641111 | -0.8779742 | 0.88563021 | -1.6052999 | -0.2407526 | 1.0987811 | -0.08327786 | -1.037346 | -1.0224521 | 0.4874937 |
| -0.1537917 | 1.5446461 | 0.7641111 | -0.8779742 | 0.05793843 | -1.6052999 | -0.8580670 | 0.4888078 | -0.08327786 | -1.037346 | -1.0224521 | 0.4874937 |
| -0.1537917 | -0.4879307 | -0.2355183 | -0.8779742 | 0.05793843 | 0.5166889 | -2.0926957 | -1.3041120 | -0.08327786 | -1.037346 | -1.0224521 | 0.4874937 |
| -0.1537917 | 1.5446461 | 0.7641111 | -0.8779742 | 0.05793843 | -1.6052999 | -0.8580670 | 0.4888078 | 0.8978933 | -1.728910 | -1.0224521 | -1.8899749 |

Gambar 4. 1 Tampilan Transformasi data menggunakan Scale (Sekala)

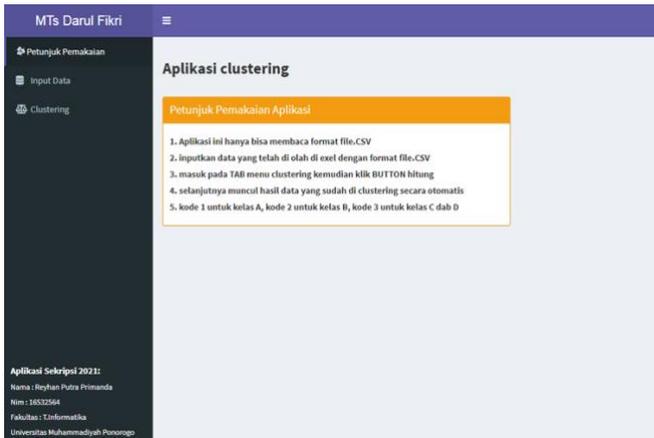
➤ Langkah Ke-Empat

Mengetikkan script berikut untuk menentukan jumlah cluster terbaik menggunakan Metode Elbow dan Metode Silhouette, informasi yang didapatkan berupa grafik yang berbentuk siku pada salah satu titiknya.

```
#menentukan cluster terbaik metode elbow
fviz_nbclust(tablefix,kmeans,method = "wss")
#menentukan cluster terbaik metode silhouette
fviz_nbclust(tablefix,kmeans)
```



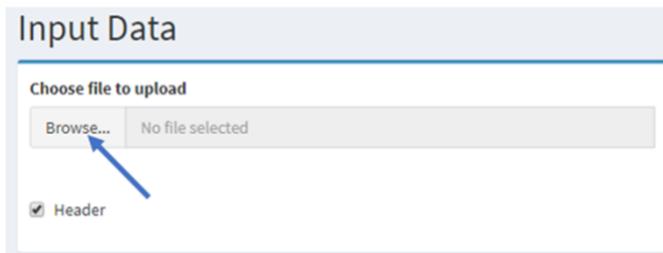
Gambar 4. 2 Tampilan hasil grafik Metode Elbow



Gambar 4. 5 Petunjuk aplikasi Rshiny dengan Algoritma K-Means

➤ Input data

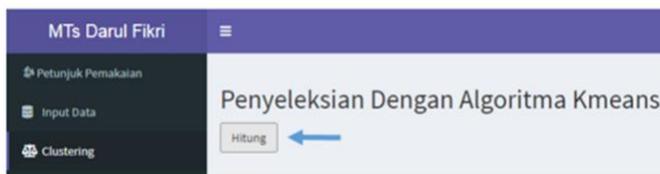
Tab ini merupakan tab untuk input file dalam format. CSV untuk data yang ingin di seleksi.



Gambar 4. 9 Input file yang akan di seleksi

➤ Hitung

Dalam tab ini menggunakan sebuah Action Button , yang digunakan untuk memulai proses perhitungan algoritma k-means.



Gambar 4. 10 Input file yang akan di seleksi

➤ Hasil penyeleksian

Berdasarkan data yang telah di seleksi sebelumnya menghasilkan hasil seleksi sebagai berikut :

| NO | NIS | NAMA | AA | SK | QH | BA | BN | BIG | IPA | IPS | MH | PKN | Krakitan | Kekhadiran | Kluster |
|----|----------|---------------------|----|----|----|----|----|-----|-----|-----|----|-----|----------|------------|---------|
| 1 | 71382831 | ABIRYU NAUFALANSYAH | 86 | 86 | 85 | 81 | 86 | 84 | 87 | 87 | 85 | 84 | 90 | 80 | 2 |
| 2 | 75798903 | ADHEL PRIHA RIZQI | 86 | 85 | 85 | 81 | 87 | 84 | 86 | 86 | 86 | 84 | 70 | 80 | 2 |
| 3 | 79498736 | AHMAD FAIZ ABOLLAH | 86 | 86 | 85 | 82 | 87 | 84 | 86 | 89 | 87 | 84 | 90 | 80 | 3 |
| 4 | 67105922 | AHMAD FAUSAN | 87 | 87 | 87 | 81 | 87 | 88 | 82 | 81 | 86 | 84 | 70 | 80 | 2 |
| 5 | 73434120 | AHMAD NAFIS S.J | 86 | 87 | 87 | 81 | 83 | 84 | 84 | 84 | 86 | 85 | 70 | 80 | 2 |
| 6 | 67903204 | AKMAL DOKIRYUL F.A | 86 | 86 | 86 | 81 | 87 | 84 | 83 | 84 | 86 | 84 | 70 | 70 | 2 |
| 7 | 73962207 | ARSY AZZAKA H.G | 86 | 85 | 87 | 81 | 88 | 84 | 86 | 86 | 86 | 84 | 70 | 80 | 2 |
| 8 | 73248871 | BINTANG NUSMANTORO | 86 | 87 | 87 | 81 | 87 | 84 | 85 | 85 | 86 | 84 | 70 | 80 | 2 |

Gambar 4. 11 Hasil penyeleksian data siswa

Berdasarkan penyeleksian dengan algoritma kmeans pada gambar 4.11 dari 100 data siswa kelas VII didapatkan hasil bahwa siswa yang menempati kelas A di kluster 1 berjumlah 23 anggota, kelas B di kluster 2 berjumlah 25 anggota dan kelas C dan D di kluster 3 berjumlah 52 anggota dengan pusat cluster di pilih secara random.

5. PENUTUP

a. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dapat ditarik kesimpulan sebagai berikut:

- Proses seleksi siswa berprestasi kelas VII di MTs Darul Fikri menggunakan metode Algoritma *k-means clustering* dengan antar muka *Rshiny di Rstudio* dengan Bahasa pemograman R berjalan dengan baik.
- Penelitian ini mendapatkan hasil dari 100 data siswa bahwa penempatan kelas unggulan adalah kelas A di kluster 1 berjumlah 23 anggota, sedangkan kelas regular terdiri dari kelas B,C, dan D meliputi kelas B di kluster 2 berjumlah 25 anggota dan

kelas C dan D di klaster 3 berjumlah 52 anggota.

- Penentuan kelas tidak pasti mengeluarkan hasil klaster 1 A, Klaster 2 B, Klaster 3 C dan D di karenakan penentuan pusat klaster di pilih secara random.
- Hasil olah data tersebut lebih memudahkan guru dalam menyeleksi kategori prestasi siswa dan memangkas waktu kerja mereka.

b. Saran

Berdasarkan penelitian yang telah dilakukan memuat Saran sebagai berikut:

- Penelitian selanjutnya proses data mining dapat dilakukan dengan lebih dari satu algoritma untuk mendapatkan clusterisasi yang lebih baik seperti decision tree dan naïve bayes.
- Perlunya penambahan icon edit jumlah variable di Rshiny tanpa harus masuk ke direktori Rstudio.

Perlunya penambahan icon download untuk memudahkan user dalam mendapatkan hasil clusterisasi.

DAFTAR PUSTAKA

[1], Pondok Pesantren Darul Fikri. (2018, 9 Februari). Pondok Pesantren Darul Fikri Bringin Kauman Ponorogo. Retrieved from 31 Juli 2020:

<http://sholikhatulwardani.blogspot.com/2018/02/pondok-pesantren-darul-fikri.html>

- [2],Mardi, Y. (n.d.). Data Mining: Klasifikasi Menggunakan Algoritma C4.5 . Jurnal Edik Informatika, V2.i2, 213-219.
- [3], Anita, Muharmi, Y., & Siska, S. T. (2018). Penentuan Tingkat Minat Belajar Online Melalui Media Sosial Menggunakan Metode Clustering K-Means. Rang Teknik Journal, Vol. 1 No.2, 126-134.
- [4], Wibowo, T. (2018). Penerapan Data Mining Pemilihan Siswa Kelas Unggulan. Skripsi, Universitas Muhammadiyah Surakarta.
- [5],Sari, V. N., Yupianti, & Maharani, D. (2018). Penerapan Metode K-Meas Clustering dalam Menentukan Predikat Kelulusan Mahasiswa untuk Menganalisa Kualitas Lulusan. JURTEKSI, Vol. IV No.2, 133-140.
- [6],Muhammad, A. (2016). Aplikasi Pembentukan Kelas dengan K-Means Clustering sebagai Alat Bantu Pemilihan Siswa Kelas Unggulan di MAN 3 Kediri. Artikel Skripsi, Universitas Nusantara PGRI Kediri.
- [7],Anita,, Muharmi, Y., & Siska, S. T. (2018). Penentuan Tingkat Minat Belanja Online Melalui Media Sosial Menggunakan Metode. Rang Teknik Journal, Vol. I No.2, 126-134.

- [8], Nugraha, R. A., & Kurniawan, D. (n.d.). Sistem Pendukung Keputusan Menggunakan Algoritma K-Means untuk Penilaian Kinerja Guru (PKG) di SMKN 1 Demak. IJCCS, 1-11.
- [9], Sugianto, C. A. (2015). Penerapan Teknik Data Mining untuk Menentukan Hasil Seleksi Masuk SMAN 1 Gibeber untuk Siswa Baru Menggunakan Decision Tree. TEDC, Vol 9 No. 1, 39-43.
- [10], Rufiyanto, A. (n.d.). Data Mining. Jurnal, Universitas Pandanaran, 1-9.
- [11], Konsep Data Mining. Diakses 31 Mei 2020, <https://ocw.upj.ac.id/files/Handout-TIF311-DM-1.pdf>
- [12], Bertalya. (2009). Konsep Data Mining. Universitas Gunadarma, 1-14.
- [13], Khomarudin, A. N. (2016). Teknik Data Mining: Algoritma. Diakses IlmuKomputer.com: <https://agusnkhom.wordpress.com/>
- [14], Han, J., & Kamber, M. (n.d.). Pengantar DM (Data Mining) 1.
- [15], Khotimah, T. (2014). Pengelompokan Surat Dalam Al-qur'an Menggunakan Algoritma K-Means. Jurnal SIMETRIS, 83-88.
- [16], Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2004). Selection of K in K-means clustering. Manufacturing Engineering Centre, Cardiff University, Vol. 219, 103-119.
- [17], Informatikalogi. (2020). Algoritma K-Means Clustering. Diakses dari INFORMATIKALOGI.COM: <https://informatikalogi.com/algoritma-k-means-clustering>
- [18], Asroni, R. A. (2015). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik. Jurnal Ilmiah Semesta Teknik, Vol. 18, No. 1, 76-82.
- [19], Tim Asprak Metkuan (2013) Kmeans Clustering. Diakses pada http://totoharyanto.staff.ipb.ac.id/files/2013/09/LKP-10-K_Means.pdf
- [20], Teknomo, K. (2007). K-Means Clustering Tutorial. Diakses pada <http://people.revoledu.com/kardi/tutorial/kMean/>
- [21], Patmawati, H., & Santika, S. (2016). Penggunaan Software Microsoft Excel sebagai Alternatif Pengolahan Data Statistika Penelitian Mahasiswa Tingkat Akhir. Seminar Nasional Matematika X Universitas Negeri Semarang, 124-129.
- [22], Supriyatna, A., & Mustika, W. P. (2018). Komparasi Algoritma Naive bayes dan SVM Untuk Memprediksi Keberhasilan Imunoterapi Untuk Memprediksi Keberhasilan Imunoterapi. Jurnal Sains Komputer & Informatika (J-SAKTI), Volume (2) No.2, 152-161.
- [23], Muharom, L. A. (2016). Rancang Bangun Data Warehouse dan R Studio Serta Pemanfaatannya dalam Peramalan Pola

Konsumsi Masyarakat di Kabupaten Jember.

JUSTINDO, Volume 01, Nomor 01, 17-25.

- [24], Salsabila, T. (2019, 8 Juli). K-Means Clustering using RStudio. Diakses K-Means Clustering using RStudio , 24 Juni 2020: <https://medium.com/@tianasalsabila.ts/k-means-clustering-using-rstudio-d6132830f6b>

