

IMPLEMENTASI ALGORITMA K-MEANS UNTUK MENENTUKAN KELAS UNGGULAN PADA SMPN 1 BOJONG

Azaria Bella Bernissa¹, Fitriyani²

¹Program Studi Sistem Infomasi/Universitas Adhirajasa Reswara Sanjaya
Jl. Sekolah Internasional NO. 1-2, Antapani, Bandung
e-mail: azariabella12@gmail.com

²Program Studi Sistem Infomasi/Universitas Adhirajasa Reswara Sanjaya
Jl. Sekolah Internasional NO. 1-2, Antapani, Bandung
e-mail: fitriyani@ars.ac.id

Abstrak

Sekolah umumnya memiliki siswa yang sedang menempuh pendidikan di sejumlah kelas, agar dapat mempunyai nilai berkualitas maka suatu sekolah perlu adanya kelas dengan siswa-siswi berprestasi yang kemudian dinamakan sebagai kelas unggulan. Pada kelas VII di SMPN 1 Bojong pihak sekolah kesulitan untuk menentukan nilai siswa-siswi yang dimasukkan ke kelas unggulan, maka digunakan teknik dari *data mining* dengan metode Algoritma *K-Means* yang dikelompokkan ke beberapa *cluster* berdasarkan kemiripan dari data tersebut. *Data mining* merupakan salah satu disiplin ilmu yang digunakan untuk menentukan suatu informasi tertentu dalam sekumpulan data sebagai pendukung pengambilan keputusan, *Clustering* adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa kelompok berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya, Algoritma *k-means* adalah algoritma yang merupakan pengelompokan data yang dibagi kedalam cluster-cluster yang memiliki kemiripan data pada satu cluster dan memiliki data ketidaksamaan pada cluster yang lain. Data yang diambil yaitu 108 data yang dibagi menjadi 3 *cluster*, pada *cluster* pertama dengan nilai rata-rata terendah masuk kedalam kelas C yaitu terdapat 36 siswa, pada *cluster* 2 dengan nilai rata-rata sedang akan dimasukkan kedalam kelas B yaitu terdapat 42 siswa, sedangkan pada *cluster* 3 dengan nilai rata-rata tertinggi akan dimasukkan kedalam kelas A yaitu terdapat 30 siswa.

Kata Kunci : Kelas Unggulan, *Data Mining*, *Clustering*, *K-Means*

Abstract

Schools generally have students who are currently studying in a number of classes, so that in order to have quality grades, a school needs a class with outstanding students which is then called the superior class.. In class VII at SMPN 1 Bojong the school had difficulty determining the grades of students who were assigned to the superior class, so data mining techniques were used with the K-Means Algorithm method which were grouped into several clusters based on the similarity of the data. Data mining is a discipline that is used to determine certain information in a set of data to support decision making, Clustering is a method used to divide a series of data into several groups based on similarities predetermined, the k-means algorithm is an Algorithm which is a grouping of data that is divided into clusters that have similar data in one cluster and have inequality data in another cluster. The data taken were 108 data which were divided into 3 clusters, in the first cluster with the lowest average value entered into class C, which is 36 students, in cluster 2 with a moderate average value will be included in class B, namely there are 42 students, while in cluster 3 with the highest average value will be included in class A, there are 30 students.

Keywords: Superior Class, *Data Mining*, *Clustering*, *K-Means*

1. Pendahuluan

Dunia pendidikan saat ini dituntut untuk mempunyai kemampuan yang sangat luas dengan memanfaatkan sumber daya pendidikan yang tinggi, maka dengan tingginya tingkat keberhasilan siswa dan rendahnya tingkat kegagalan siswa merupakan cermin kualitas dunia pendidikan. Maka pendidikan ini bisa dikatakan adalah salah satu kunci pembentukan siswa-siswi yang berkualitas (Sibuea & Safta, 2017) Pada dasarnya dunia pendidikan sangatlah penting bagi kehidupan, karna pendidikan adalah nomor satu bahkan pada siswa-siswi Sekolah Menengah Pertama yang sangatlah butuh pendidikan lebih baik dari sejak dini, maka dari itu (Fitriyani et al., 2017) memaparkan pentingnya pendidikan dalam perannya mencerdaskan kehidupan bangsa telah melahirkan banyak pihak untuk berkontribusi aktif dalam penyelenggaraannya, baik di bidang formal maupun non-formal.

Secara umum penilaian prestasi siswa yaitu dengan salah satu mata pelajaran baik teori maupun praktek, dan penilaian prestasi siswa tersebut dilakukan dengan memberi nilai oleh pengajar kepada semua siswa yang mengikuti pelajaran yang diajarkan dan Ekstrakurikuler yang diikutinya (Sibuea & Safta, 2017). Sehingga menurut Wardani dalam (Fitriyani et al., 2017) nilai menjadi salah satu hal penting dalam kegiatan belajar mengajar, sehingga diharapkan akan menjadi input yang akan menjadi output agar mengembangkan nilai siswa, dan nilai juga menjadi acuan untuk pengambilan keputusan.

Sekolah umumnya memiliki siswa yang sedang menempuh pendidikan di sejumlah kelas, agar dapat mempunyai nilai berkualitas maka suatu sekolah perlu adanya kelas yang dengan siswa-siswi berprestasi yang kemudian dinamakan sebagai kelas unggulan (Apriandala et al., 2017).

Kelas Unggulan adalah kelas yang diikuti oleh sejumlah siswa yang unggul dalam dua ranah penilaian dengan kecerdasan di atas rata-rata yang dikelompokkan secara khusus. Pengelompokan ini bertujuan untuk untuk memotivasi siswa-siswi dalam meningkatkan ilmu pengetahuan yang lebih luas dan optimal (Ramadani et al., 2017).

SMPN 1 BOJONG sekolah yang berada di Kota Purwakarta, dengan seiringnya waktu sekolah ini semakin

berkembang dengan sistem pembelajaran yang bagus, sehingga setiap tahunnya pendaftar sekolah semakin meningkat. Dengan meningkatnya pendaftar setiap tahunnya maka sekolah mengalami kesulitan dalam pemilihan siswa kelas unggulan dalam pengelompokan siswa kelas VII yang naik ke kelas VIII. Berdasarkan permasalahan yang ada oleh karena itu menggunakan penerapan *Data Mining* ini agar lebih memudahkan bagi kurikulum sekolah menyeleksi siswa-siswi masuk kelas unggulan maka penelitian ini menggunakan *Metode Clustering Algoritma K-Means* agar mengetahui hasil *cluster* pengelompokan siswa-siswi berprestasi karena algoritma ini merupakan algoritma yang mempunyai ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar (Kusuma & Aryati, 2019).

Ada beberapa penelitian sebelumnya yang telah dilakukan oleh peneliti lain mengenai kelas unggulan akan tetapi menggunakan algoritma yang lain seperti pada penelitian (Hutasuhut et al., 2019) dengan algoritma Iterative Dichotomiser 3 (ID3), lalu pada penelitian yang menggunakan algoritma *K-Means Clustering* dan dataset yang sama yaitu kelas unggulan, akan tetapi menggunakan Sistem Informasi Akademik secara terkomputerisasi pada penelitian (Kusuma & Aryati, 2019). Sedangkan penelitian yang menggunakan algoritma yang sama tetapi dengan dataset yang digunakan berbeda seperti algoritma *K-Means Clustering* pada penelitian (Sadewo et al., 2017). Algoritma *K-Means Clustering* pada penelitian (Yulianti et al., 2019). Algoritma *Fuzzy C-Means* dan *K-Means* pada penelitian (Butarbutar et al., 2016). Dan Algoritma *K-Means Clustering* pada penelitian (Praja et al., 2019).

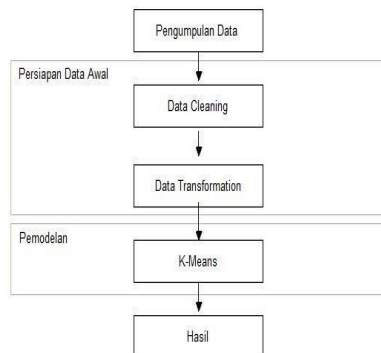
Maka berdasarkan latar belakang di atas penulis tertarik untuk membuat penelitian yang berjudul **“Implementasi Algoritma K-Means Untuk Menentukan Kelas Unggulan Berdasarkan Nilai Siswa SMPN 1 BOJONG”**.

2. Metodologi Penelitian

2.1. Tahapan Penelitian

Pada bagian ini, akan dijelaskan gambaran mengenai metodologi penelitian secara keseluruhan yang didasarkan pada CRISP-DM. Berikut adalah tahapan yang

dilakukan penelitian ini seperti yang terlihat pada gambar 1.



Gambar 1 Desain Penelitian

1. **Pengumpulan Data**
Pengumpulan data adalah salah satu langkah yang sangat penting dalam sebuah penelitian karena memiliki tujuan utama untuk mendapatkan data sebuah data.
 2. **Data Cleaning**
proses data yang perlu dilakukan dengan poses cleaning pada suatu data, proses cleaning ini antara lain membuang data yang tdak penting, tidak konsisten, dan memperbaiki kesalahan data.
 3. **Data Transformation**
data yang telah dipilih melalui proses data mining sehingga dalam proses KDD merupakan proses yang bergantung pada jenis atau pola yang dicari dalam basis data
 4. **K-Means**
Algoritma yang merupakan pengelompokan data yang dibagi kedalam cluster-cluster yang memiliki kemiripan data pada satu cluster dan memiliki data ketidaksamaan pada cluster yang lain
- 2.2. Instrumen Penelitian**
- Instrumen penelitian adalah suatu alat pengumpulan data atau mengukur objek suatu variabel penelitian, untuk memiliki data yang benar atau valid dan konsisten.(Yusup, 2018). Pada penelitian ini akan digunakan beberapa penelitian, antara lain sebagai berikut :
1. **Perangkat Lunak (Software)**
Perangkat lunak yang digunakan pada penelitian ini adalah *software rapidminer* versi 9.6
 2. **Dataset**
Penelitian ini menggunakan *dataset* data nilai siswa SMPN 1 Bojong tahun

sebanyak 108 data siswa yang telah melalui proses *cleaning* dan *transformation* terdiri dari atribut

2.3. Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data *public* yang di ambil langsung dari SMPN 1 BOJONG dengan cara melakukan riset kepada bagian kesiswaan.

A. Sampel Data

Sampel data yang digunakan untuk penelitian ini adalah data siswa kelas IX sebanyak 108 data siswa terdiri dari atribut NIS, Nama Siswa, Nilai mata pelajaran Bahasa Indonesia, Matematika, IPA, IPS

Tabel Atribut Data

NIS	Nomor Induk Siswa SMPN 1 BOJONG
Nama Siswa	Nama Siswa SMPN 1 BOJONG
Nilai	Nilai yang diambil dari rapor siswa

2.4. Metode Analisa Data

A. Pengolahan Data Awal

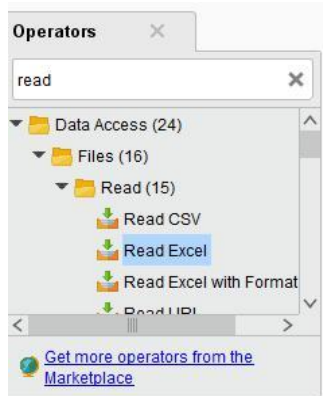
Pada tahap ini merupakan tahap untuk memastikan data siswa yang dipilih telah layak untuk dilakukan proses pengolahan *Data cleaning*.

Data Cleansing merupakan proses data yang tidak konsisten atau tidak relevan (Ridwan et al., 2013), data yang tidak konsisten, dan data yang tidak lengkap atau *missing value* dapat dihapus (Amalia & Evienna, 2017). Maka pada *dataset* ini hanya diambil 108 data dari 199 data, karena adanya data yang tidak bisa diolah datanya dan hanya mengambil 18 besar dari setiap kelas untuk diolah. Maka pada proses *cleaning* ini dilakukan penghapusan untuk menghilangkan *missing value* pada *dataset*.

B. Permodelan

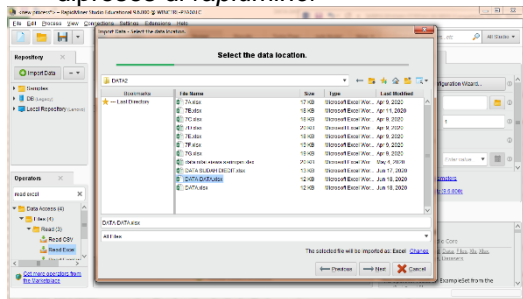
Pada tahap permodelan ini dilakukan dengan menggunakan software *rapidminer* 9.6 dengan langkah sebagai berikut :

1. Buka *software rapidminer*
2. Lalu pilih menu *file* kemudian *new proses*
3. Masuk kedalam panel *operator* seperti yang terlihat pda gambar 1 kemudian ketik *read excel*, lalu *double* klik atau *drag* pada *read axcel* agar *operator* tersebt muncul di panel *process* dan bisa digunakan untuk memasukan data *excel* yang telah melalui tahap data *cleansing* dan data *transformatio*



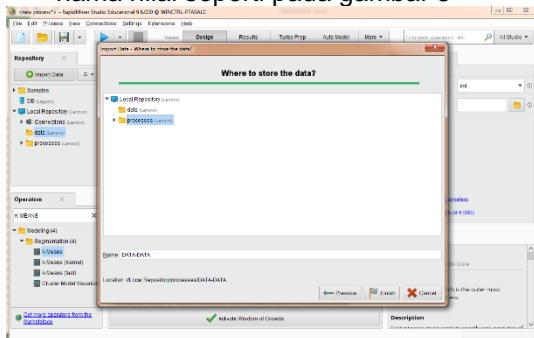
Gambar 1 Pilih Read Excel

4. Pilih *button import configuration wizard* agar bisa memilih file *dataset* yang akan diproses di *rapidminer*



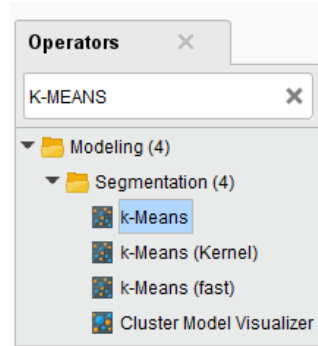
Gambar 2 Memilih Dataset

5. Kemudian simpan pada proses dengan nama nilai seperti pada gambar 3



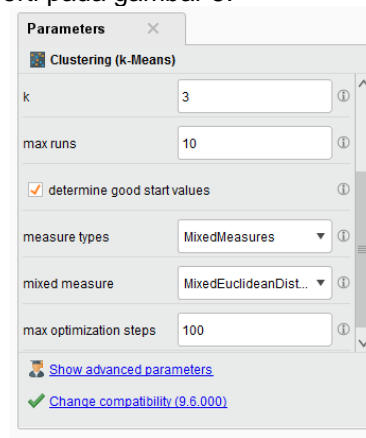
Gambar 3 Simpan Dataset

6. Langkah selanjutnya adalah memasukan model yang akan digunakan untuk proses *clustering data mining* dengan cara masuk ke *operator* kemudian pilih *K-Means* seperti pada gambar 4



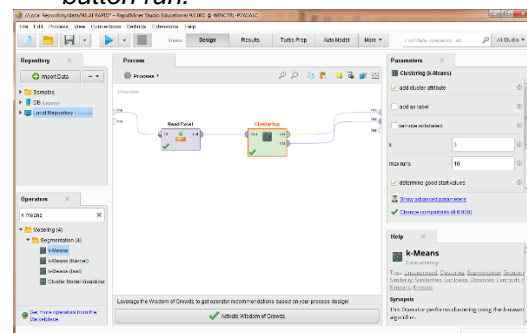
Gambar 4 Pilih Permodelan

7. Pada proses *clustering* tentukan terlebih dahulu jumlah klasternya menjadi $K=3$ dan *mixed measure* menjadi *mixed Euclidean distance* seperti pada gambar 5.



Gambar 5 Menentukan Jumlah K

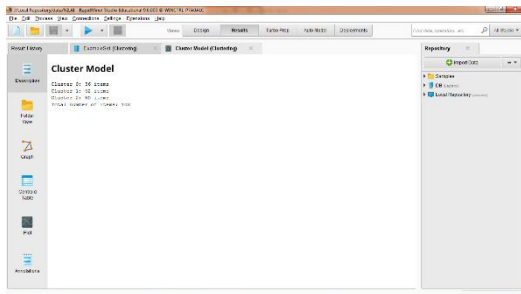
8. Langkah selanjutnya adalah menghubungkan *conector* masing – masing proses pada *main process* seperti pada gambar 6 selanjutnya klik *button run*.



Gambar 6 Menggabungkan Conector

9. Pada gambar 7 hasil pengujian *dataset* yang berjumlah 108 data menggunakan *software rapidminer 9.6* terbentuk 3 *clster*. Pada *cluster 0 (cluster pertama)*

terdapat 36 *items*, pada *cluster* 1 (*cluster* kedua) terdapat 42 *items*, dan pada *cluster* 2 (*cluster* ketiga) terdapat 30 *item*



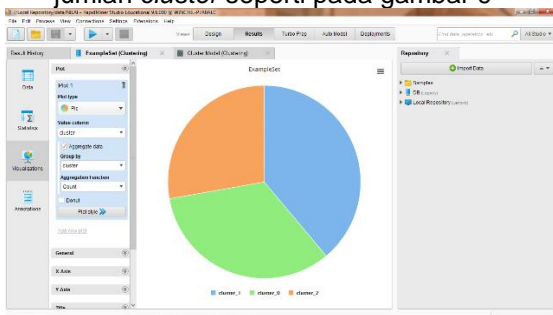
Gambar 7 Hasil Cluster

10. Dari 3 *cluster* yang terbentuk kemudian *centroid* akhir untuk setiap *cluster*, yaitu *cluster* 0 = 181907113, 81.306, 78.944, 81.250, 74.333 *cluster* 1 = 181907042, 81.595, 77.548, 83.143, 74.333 *cluster* 2 = 181907180, 83.200, 78.200, 82.200, 74.600 seperti pada gambar 8

Cluster	Centroid 0	Centroid 1	Centroid 2
X	181907113	181907042	181907180
Y	81.306	81.595	83.200
Z	78.944	77.548	78.200
W	81.250	83.143	82.200
V	74.333	74.333	74.600

Gambar 8 Centroid Tabel

11. Kemudian bisa dilihat *chart* dari setiap jumlah *cluster* seperti pada gambar 9

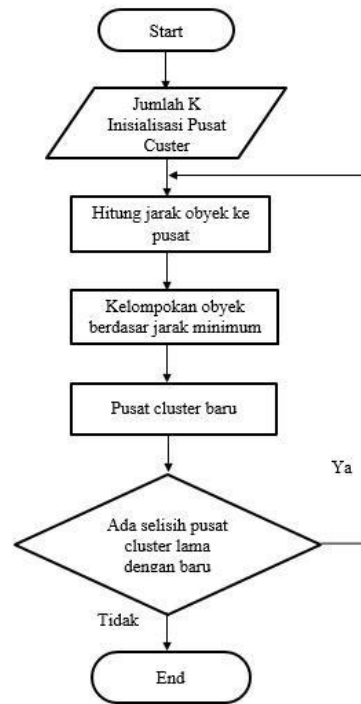


Gambar 9 Piechart Cluster

3. Hasil Penelitian dan Pembahasan

3.1. Hasil Perhitungan

A. Flowchat Pengolahan Data



Gambar 1 Flowchart K-Means

Untuk melakukan perhitungan manual seperti pada Flowchart diatas, maka penjelasannya sebagai berikut :

1. Siapkan data terlebih dahulu agar data bisa diolah/dihitung.
2. Tentukan jumlah K Inisialisasi Pusat Cluster sesuai yang dibutuhkan.
3. Setelah jumlah K Inisialisasi telah ditentukan maka hitung jarak ke pusat agar mengetahui hasil data yang akan dikelompokan pada setiap cluster
4. Setelah jarak objek ke pusat telah diketahui hasilnya, maka dapat dihitung jarak minimum data untuk dikelompokan.
5. Setelah data dikelompokan berdasarkan jarak minimum dan cluster, maka akan diketahui pusat cluster baru yang akan dihitung kembali
6. Setelah pusat cluster baru diketahui dihitung kembali, jika perhitungan antara cluster baru dan cluster lama ada selisih (berubah) data maka perhitungan diulang kembali, tetapi jika perhitungan antara cluster baru dan cluster lama sudah tidak ada selisih (tidak berubah) perhitungan bisa berhenti.

B. Perhitungan K-Means

Untuk menentukan banyaknya *cluster* k dapat dilakukan dengan beberapa pertimbangan seperti pertimbangan terioritis dan konseptual yang mungkin diusulkan untuk menentukan berapa banyak *cluster*. Penetapan jumlah *cluster* k pada penelitian ini yaitu berjumlah 3 *cluster*.

Titik Awal Tiap Cluster

Titik Pusat Awal	NIS	B.Indonesia	Matematika	IPA	IPS
Cluster 0	181907059	84	74	85	71
Cluster 1	181907108	76	75	79	71
Cluster 2	181907166	89	81	80	79

Hitung jarak setiap data ke pusat antara objek ke centroid dengan perhitungan jarak *Euclidean*. Perusahaan yang digunakan adalah:

$$(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2 + (p_4 - q_4)^2 + (p_5 - q_5)^2}$$

Persamaan ini digunakan karena atribut yang digunakan berjumlah 3.

Sebagai contoh, akan dihitung jarak dari data siswa pertama ke pusat cluster pertama dengan persamaan:

$$\begin{aligned} (1,1) &= \sqrt{(181907016 - 181907059)^2 + (79 - 84)^2 + (78 - 74)^2 + (80 - 85)^2 + (73 - 71)^2} \\ &= \sqrt{(-43)^2 + (-5)^2 + (4)^2 + (-5)^2 + (2)^2} \\ &= \sqrt{1.849 + 25 + 16 + 25 + 4} \\ &= \sqrt{1.919} \\ &= 43.80639 \end{aligned}$$

Dari hasil perhitungan di atas maka di dapatkan hasil bahwa jarak data nilai siswa pertama pada *cluster* pertama adalah 43.80639.

Pada data nilai siswa kedua dapat dihitung jarak pada setiap *cluster* dari nilai data siswa agar mengetahui hasil persamaan:

$$\begin{aligned} (2,1) &= \sqrt{(181907042 - 181907059)^2 + (77 - 84)^2 + (72 - 74)^2 + (83 - 85)^2 + (71 - 71)^2} \\ &= \sqrt{(-17)^2 + (-7)^2 + (-2)^2 + (-2)^2 + (0)^2} \\ &= \sqrt{289 + 49 + 4 + 4 + 0} \\ &= \sqrt{346} \\ &= 18.60108 \end{aligned}$$

Maka dari hasil perhitungan di atas di dapatkan hasil jarak bahwa jarak data nilai siswa kedua pada *cluster* pertama adalah 18.60108.

Pada data nilai siswa ketiga dapat dihitung jarak pada setiap *cluster* dari nilai data siswa agar mengetahui hasil persamaan:

$$\begin{aligned} (3,2) &= \sqrt{(181907059 - 181907108)^2 + (84 - 76)^2 + (74 - 75)^2 + (85 - 79)^2 + (71 - 71)^2} \\ &= \sqrt{(-49)^2 + (8)^2 + (-1)^2 + (6)^2 + (0)^2} \\ &= \sqrt{2.401 + 64 + 1 + 36 + 0} \\ &= \sqrt{2.502} \\ &= 50.01 \end{aligned}$$

Maka dari hasil perhitungan di atas di dapatkan hasil jarak bahwa jarak data nilai siswa ketiga pada *cluster* pertama adalah 0

Contoh Hasil Perhitungan Data ke Setiap Cluster

No	NIS	Nilai				Jarak			Cluster
		B.Indonesia	Matematika	IPA	IPS	C1	C2	C3	
1	181907016	79	78	80	73	43.80639	92.12491	150.4825	1
2	181907042	77	72	83	71	18.60108	66.19667	125.1958	1
3	181907059	84	74	85	71	0	50.01	107.758	1
4	181907073	79	78	80	73	16.30950	35.32703	93.77632	1
5	181907297	85	72	85	77	238.0861	189.4280	131.48	3
6	181907108	76	75	79	71	50.01	0	60.28267	2
7	181907129	81	82	80	72	70.70360	22.73763	38.50973	2
8	181907153	75	77	77	71	94.81561	45.099	21.21320	3
9	181907166	89	81	80	79	107.7589	60.39039	0	3
10	181907168	78	75	79	72	109.3389	60.04165	14.52583	3

Kelompokan data ke dalam *cluster* dengan jarak terdekat (minimal)

Tabel Dengan Centroid Awal

No	Cluster 1	Cluster 2	Cluster 3
1	1	0	0
2	1	0	0
3	1	0	0
4	1	0	0
5	0	0	1
6	0	1	0
7	0	1	0
8	0	0	1
9	0	0	1
10	0	0	1

Setelah semua data dimasukan kedalam *cluster* sesuai dengan jarak dan tempatnya, kemudian hitung kembali pusat *cluster* yang baru berdasarkan rata-rata anggota yang ada pada *cluster* tersebut sepusat dari setiap *cluster* tidak berubah lagi dan tidak ada lagi data yang berpindah dari *cluster* ke *cluster* yang lain

3.2. Hasil Penelitian

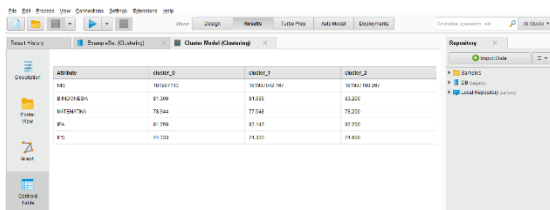
Hasil penelitian yang telah di uji ini maka hasil antara *Software Rapidminer* dan *Microsoft Excel* yaitu mendapatkan hasil yang sama. Pada eksperimen data di *Microsoft Excel* menemukan hasil yang tidak berubah yaitu pada iterasi 10 dengan 3

cluster dan beberapa atribut seperti Nama Siswa, NIS, Nilai (B.Indonesia, Matematika, IPA, IPS).

Tabel Centroid Baru

Titik Pusat Awal	NIS	B.Indonesia	Matematika	IPA	IPS
Cluster 1	181907113	81.31	78.94	81.25	74.33
Cluster 2	181907042	81.60	77.55	83.14	74.33
Cluster 3	181907180	83.20	78.20	82.20	74.60

Untuk memastikan hasil perhitungan *Software Rapidminer* 9.6 diperhitungan manual *Microsoft Excel* dengan hasil yang sama



Atribut	CLUSTER_3	CLUSTER_2	CLUSTER_1
NIS	181907113	181907042	181907180
B.INDONESIA	81.30556	81.59524	81.30556
MATEMATIKA	78.94444	77.54762	78.94444
IPA	81.25	83.14	81.25
IPS	74.33333	74.33333	74.6

Gambar IV. 2 Hasil Software Rapidminer

Hasil Microsoft Excel

Titik Pusat	NIS	B.Indonesia	Matematika	IPA	IPS
Cluster 1	181907113	81.30556	78.94444	81.25	74.33333
Cluster 2	181907042	81.59524	77.54762	83.14	74.33333
Cluster 3	181907180	83.2	78.2	82.2	74.6

Berdasarkan dari hasil perhitungan diatas yang telah dibuat, dari 108 data siswa di bagi menjadi kedalam 3 *cluster* berdasarkan rata-rata NIS, Nilai B.Indonesia, Nilai Matematika, Nilai IPA, dan Nilai IPS. Pada *cluster 1* yang telah dihitung berdasarkan *Microsoft excel* dan *Rapidminer* menghasilkan jumlah 36 siswa dengan nilai rata-rata 78.95833. Pada *cluster 2* yang telah dihitung berdasarkan *Microsoft excel* dan *Rapidminer* menghasilkan jumlah 42 siswa dengan nilai rata-rata 79.15471. Dan pada *cluster 3* yang telah dihitung berdasarkan *Microsoft excel*

dan *Rapidminer* menghasilkan jumlah 30 siswa dengan nilai rata-rata 79.55.

Dari 3 *cluster* yang terbentuk, yaitu *cluster 1* adalah *cluster* dengan nilai rata-rata terendah akan dimasukkan kedalam kelas C yaitu terdapat 36 siswa, pada *cluster 2* dengan nilai rata-rata sedang akan dimasukkan kedalam kelas B yaitu terdapat 42 siswa, sedangkan pada *cluster 3* dengan nilai rata-rata tertinggi akan dimasukkan kedalam kelas A yaitu terdapat 30 siswa.

4. Kesimpulan

Hasil dari penelitian yang telah diuji eksperimen data menggunakan *Microsoft Excel* dan *Software Rapidminer* dengan Algoritma *K-Means Clustering* yang bertujuan untuk mengetahui siswa berprestasi masuk kedalam kelas unggulan. Berdasarkan hasil yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Telah diterapkan metode Algoritma *K-Means Clustering* untuk mengelompokkan data nilai siswa masuk ke kelas unggulan berdasarkan nilai yang telah diperoleh oleh siswa-siswi kelas VII yang akan naik ke kelas VIII.
2. Pengolahan data yang dilakukan menghasilkan kelompok siswa berprestasi berdasarkan nilai, sehingga pihak kurikulum sekolah dapat lebih mudah dalam menyeleksi siswa-siswi masuk kelas unggulan dengan teknik *Data Mining*.
3. Setelah dilakukan eksperimen dapat diketahui 3 *cluster* yang di butuhkan oleh sekolah SMPN 1 Bojong. Yaitu pada *cluster 1* kelompok siswa berprestasi dengan nilai rata-rata terendah masuk kedalam kelas C dengan jumlah 36 siswa, *cluster 2* yaitu kelompok siswa berprestasi dengan nilai sebesar 50% dengan nilai rata-rata tertinggi masuk kedalam kelas B dengan jumlah 42 siswa, dan pada *cluster 3* yaitu kelompok siswa berprestasi dengan nilai rata-rata tertinggi masuk kedalam kelas A dengan jumlah 30 siswa.

5. Saran

Berikut ini adalah saran untuk pengembangan pada penelitian selanjutnya antara lain:

1. Pada penelitian selanjutnya penyeleksian kelas unggulan ini bisa menambahkan atribut lainnya

berdasarkan kategori nilai akademik dan non akademik.

2. Pada penelitian selanjutnya lebih baik bisa dibuatkan sebuah program atau aplikasi pengelompokan data agar lebih mudah untuk mengimplemetasikan data menyeleksi siswa-siswi masuk kedalam kelas unggulan.

Referensi

- Apriandala, R., Efendi, R., & Andreswari, D. (2017). *Pembagian Kelas Siswa Smarter Dan (Studi Kasus Smpn 1 Kota Bengkulu)*. 5(2), 209–219.
- Butarbutar, N., Perdana Windarto, A., Hartama, D., & Solikhun. (2016). Komparasi Kinerja Algoritma Fuzzy C-Means dan K-Means dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa (Studi Kasus : SMP Negeri 2 Pematangsiantar). *Jurnal Riset Informasi & Teknis Informatika*, 1(1).
- Fitriyani, F., Hadinata, N. S., & Sanjaya, R. (2017). Website Pengolahan Nilai Siswa Menggunakan Konsep Model-View-Controller pada LIA Pamulang. *Informatika*, 4(2), 2–7.
- Hilda Amalia, E. (2017). *Aplikasi Sistem Penunjang Keputusan Untuk Prediksi Persalinan Sesar*. 3(1), 121–126.
- Hutasuhut, M., Octavina, D., & Halim, J. (2019). *Penerapan Data Mining dalam Menganalisa Pola Kelayakan Siswa Pada Kelas Unggulan Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) pada*. 18(2), 154–160.
- Kusuma, A. S., & Aryati, K. S. (2019). Sistem Informasi Akademik Serta Penentuan Kelas Unggulan Dengan Metode Clustering Dengan Algoritama K-Means Di Smp Negeri 3 Ubud. *Jurnal Sistem Informasi Dan Komputer Terapan Indonesia (JSIKTI)*, 1(3), 143–152.
<https://doi.org/10.33173/jsikti.29>
- Praja, B. S., Kusuma, P. D., & Setianingsih, C. (2019). Penerapan Metode K-Means Clustering Dalam Pengelompokan Data Penumpang Dan Kapal Angkutan Laut Di Indonesia. *E-Proceeding of Engineering*, 06(1), 1442.
- Ramadani, S. F., Ekojono, E., & Santoso, N. (2017). Sistem Pendukung Keputusan Seleksi Siswa Kelas Unggulan Di Smp Negeri 7 Malang. *Jurnal Informatika Polinema*, 3(3), 27.
<https://doi.org/10.33795/jip.v3i3.30>
- Ridwan, M., Suyono, H., & Sarosa, M. (2013). *Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier*. 7(1), 59–64.
- Sadewo, M. G., Windarto, A. P., & Hartama, D. (2017). Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering. *InfoTekJar (Jurnal Nasional Informatika Dan Teknologi Jaringan)*, 2(1), 60–67.
<https://doi.org/10.30743/infotekjar.v2i1.164>
- Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *Jurteks*, 4(1), 85–92.
<https://doi.org/10.33330/jurteks.v4i1.28>
- Yulianti, Y., Utami, D. Y., Hikmah, N., & Hasan, F. N. (2019). Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Mengetahui Minat Customer Di Toko Hijab. *Jurnal Pilar Nusa Mandiri*, 15(2), 241–246.
<https://doi.org/10.33480/pilar.v15i2.650>
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23.
<https://doi.org/10.18592/tarbiyah.v7i1.2100>