

PENERAPAN DATA MINING TERHADAP MINAT SISWA DALAM MATA PELAJARAN MATEMATIKA DENGAN METODE K-MEANS

Sufajar Butsianto¹⁾, Nurhali Saepudin²⁾

Program Studi Teknik Informatika Fakultas Teknik
Universitas Pelita Bangsa
sufajar@pelitabangsa.ac.id

Disetujui, 30 Desember 2019

Abstrasi

Kemampuan matematika yang diperlukan untuk menguasai dan menciptakan teknologi di masa depan, menjadikan matematika yang kuat perlu dibina sejak dini. Tujuan penelitian ini adalah Menerapkan metode K-Means untuk mengelompokkan minat siswa terhadap mata pelajaran matematika dan untuk mendapatkan akurasi yang tepat dan cepat dalam mengelompokkan minat siswa terhadap mata pelajaran matematika menggunakan metode K-Means. Metode yang digunakan yaitu dengan teknik data mining menggunakan algoritma K-Means. Proses ini menghasilkan 2 cluster yaitu (minat) Matematika dan (kurang minat) matematika, dengan menggunakan teknik data mining menggunakan algoritma K-Means, dan akurasi diukur dengan Davies Bouldin Index. Pengujian menggunakan validasi DBI (Davies Bouldin Index) diperoleh nilai untuk tiap-tiap cluster. Untuk kelas 10 pengujian cluster 1 menghasilkan nilai DBI 0.941 dan cluster 2 nilai DBI 0.335, kelas 11 pengujian cluster 1 menghasilkan nilai DBI 0.660 dan cluster 2 nilai DBI 0.506, kelas 12 pengujian cluster 1 menghasilkan nilai DBI 0.271 dan cluster 2 nilai DBI 0.111. Dari perhitungan Davies Bouldin Index (DBI) dapat disimpulkan bahwa jika semakin kecil nilai Davies Bouldin Index (DBI) yang diperoleh (non negatif ≥ 0) maka cluster tersebut semakin baik.

Kata kunci : Data mining, *Clustering, K-Means*, Matematika, SMK Binamitra

Abstrak

The purpose of this mathematics is needed to design and develop future technology, to make mathematics that needs to be fostered early on. The purpose of this study is to apply the K-Means method to classify students' interest in students' mathematics lessons in mathematics using the K-Means method. The method used is data mining techniques using the K-Means algorithm. This process produces 2 clusters, namely (mathematics) mathematics and (less interest) mathematics, using data mining techniques using the K-Means algorithm, and processing accuracy with the Bouldin Davies Index. Tests using DBI (Davies Bouldin Index) validation obtain values for each cluster. For class 10, cluster 1 testing produces DBI value 0.941 and cluster 2, DBI value 0.335, class 11 testing cluster 1 produces DBI value 0.660 and cluster 2, DBI value 0.506, class 12 testing cluster 1 produces DBI value 0.271 and cluster 2, DBI value 0.111. From the calculation of Davies Bouldin Index (DBI) can determine the value if the smaller the value of Davies Bouldin Index (DBI) obtained (non negative ≥ 0), the better the cluster is.

Keywords : Data mining, *Clustering, K-Means*, Matematika, SMK Binamitra

1. Pendahuluan

Kemampuan matematika yang diperlukan untuk menguasai dan menciptakan teknologi di masa depan, menjadikan matematika yang kuat perlu dibina sejak dini. Masyarakat beranggapan bahwa matematika menjadi tolak ukur kecerdasan dan kepandaian anak dalam belajar. Namun demikian matematika masih dipandang sebagai salah satu bidang studi yang tidak disenangi atau bahkan bias jadi paling dibenci dan masih melekat kebanyakan siswa yang mempelajarinya sampai saat ini.

Hal itulah yang mempengaruhi minat seorang siswa dengan matematika, berbeda dengan siswa lainnya. Minat belajar menurut "Djamarah" dalam cenderung menghasilkan prestasi tinggi, sebaliknya minat belajar

yang kurang akan menghasilkan prestasi yang rendah. Hal itulah yang menyebabkan minat setiap siswa berbeda dengan siswa lainnya[1].

Berdasarkan Elea Tinggi, matematika berarti ilmu pengetahuan yang diperoleh dengan bernalar dan hasil yang jujur. Dalam hal ini dimaksudkan bukan berarti ilmu lain diperoleh tidak melalui penalaran, akan tetapi matematika lebih menekankan aktivitas dalam dunia rasio (penalaran), sedangkan dalam ilmu lain lebih menekankan hasil observasi atau eksperimen disamping penalaran[1].

Dalam metode clustering mempartisi data ke dalam kelompok, sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama. Jumlah kelompok yang dapat diidentifikasi tergantung pada banyak dan variasi data obyek. Tujuan dari pengelompokan sekumpulan data obyek kedalam beberapa kelompok yang mempunyai karakteristik tertentu dan dapat dibedakan satu sama lainnya adalah untuk analisis dan interpretasi lebih lanjut sesuai dengan tujuan penelitian yang dilakukan[1].

2. Tinjauan Studi

2.1 Observasi

Pengamatan langsung terhadap alur proses yang sedang berjalan pada SMK Binamitra Cikarang Timur untuk memperoleh informasi yang nantinya akan diolah kedalam perhitungan *k-means* untuk menentukan minat siswa dalam mata pelajaran matematika.

2.2 Landasan Teori

Setelah peneliti dan melakukan analisa terhadap beberapa penelitian, ada beberapa penelitian yang memiliki keterkaitan dengan penelitian yang kami lakukan[1].

2.3 Sistem K-Means

Data mining berkembang menjadi alat bantu untuk mencari pola-pola yang berharga dalam suatu basisdata yang sangat besar jumlahnya, sehingga tidak memungkinkan dicari secara manual. Beberapa teknik data mining dapat diklasifikasikan ke dalam kategori berikut, meliputi klasifikasi, clustering, penggalian kaidah asosiasi, analisa pola sekuensial, prediksi, visualisasi data dan lain sebagainya. Teknik clustering adalah teknik yang digunakan untuk menangani data yang besar dengan banyak atribut ke dalam sejumlah kelompok kecil. Clustering dilakukan dengan terlebih dahulu menganalisis bagian kecil dari data untuk menentukan klaster. Clustering merupakan pengelompokan record, observasi, atau kasus ke dalam kelas-kelas objek yang mirip. Clustering berbeda dengan klasifikasi dimana dalam clustering tidak terdapat variabel target. Salah satu algoritma clustering adalah K-Means. Clustering merupakan suatu teknik data mining yang membagi-bagikan data ke dalam beberapa kelompok (grup atau cluster atau segmen) yang tiap cluster dapat ditempati beberapa anggota bersama-sama. Setiap obyek dilewatan ke grup yang paling mirip dengannya. Ini menyerupai penyusunan binatang dan tumbuhan ke dalam keluarga – keluarga yang para anggotanya mempunyai kemiripan. Clustering tidak mensyaratkan pengetahuan sebelumnya dari grup yang dibentuk, juga dari para anggota yang harus mengikutinya[2].

Algoritma K-Means diperkenalkan oleh J.B. MacQueen pada tahun 1976, salah satu algoritma clustering sangat umum yang mengelompokkan data sesuai dengan karakteristik atau ciri-ciri bersama yang serupa. Grup data ini dinamakan sebagai cluster. Data di dalam suatu cluster mempunyai ciri-ciri (atau fitur, karakteristik, atribut, properti) serupa dan tidak serupa dengan data pada cluster lain[2].

Beberapa tahap dari algoritma K-Means dapat dilihat pada algoritma berikut :

1. Penentuan Awal *cluster*

Dalam menentukan n buah pusat cluster awal dilakukan pembangkitan bilangan random yang merepresentasikan urutan data input. Pusat awal cluster didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan random pusat awal dari data.

2. Perhitungan jarak dengan pusat cluster Untuk mengukur jarak antar data dengan pusat dengan cluster digunakan euclidian distance, algoritma perhitungan jarak data dengan pusat cluster :

- a. Pilih nilai data dan nilai pusat cluster
 - b. Hitung euclidian distance data dengan tiap pusat cluster
- $$(X_i, \mu_j) = \sqrt{((X_i - \mu_j))^2}$$

Penjelasan :

X_i : Data kriteria

μ_j : Centroid pada cluster ke j.

3. Pengelompokan Data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat cluster, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok

dengan pusat cluster terdekat.

Adapun cara pengelompokan data tersebut adalah :

- a. Pilih nilai jarak tiap pusat cluster dengan data.
- b. Cari nilai jarak terkecil.
- c. Kelompokkan data dengan pusat cluster yang memiliki jarak terkecil.

4. Penentuan Pusat *Cluster* baru

Untuk mendapatkan pusat cluster baru bisa dihitung dari rata-rata nilai anggota cluster dan pusat cluster. Pusat cluster yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen.

Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh user atau hasil yang dicapai sudah konvergen (pusat cluster baru sama dengan pusat cluster lama).

5. Algoritma Penentuan pusat cluster :

- a. Cari jumlah anggota tiap cluster
- b. Hitung pusat baru dengan rumus

$$\mu_j(t+1) = 1/Nsj \sum_{j \in S_j} X_j$$

Penjelasan :

$\mu_j(t+1)$: Centroid baru pada iterasi ke 1

Nsj : Banyak data pada cluster S_j

3. Desain Penelitian/Metodologi

3.1 DBI (Davies Bouldin Index)

Davies bouldin index (DBI) adalah metric untuk mengevaluasi atau mempertimbangkan hasil algoritma clustering. Pertama kali diperkenalkan oleh David L. Davies dan Donald W. Bouldin pada tahun 1979. Dengan menggunakan DBI suatu cluster akan dianggap memiliki skema clustering yang optimal adalah yang memiliki DBI minimal.

Langkah-langkah perhitungan Davies Bouldin Index adalah Sebagai berikut :

1. Sum Of Square Within- Cluster(SSW)

Untuk mengetahui kohesi dalam sebuah cluster ke-I salah satunya adalah dengan menghitung nilai dari Sum Of Square Within-Cluster (SSW). Dengan rumus sebagai berikut :

$d(x_j, c_i)$ = jarak setiap data ke *centroid* i yang dihitung menggunakan jara jarak *euclidiance*.

2. Sum Of Square Between-Cluster (SSB)

Perhitungan *Sum Of Square Between-Cluster (SSB)* bertujuan untuk mengetahui separasi atau jarak antar *cluster*. dengan rumus perhitungan sebagai berikut

$$SSB_{ij} = d(X_i, X_j)$$

Dimana :

$$d(X_i, X_j) = \text{jarak antara data ke } i \text{ dengan data ke } j \text{ di } cluster \text{ lain.}$$

3. Ratio (Rasio)

Perhitungan rasio ($R_{i,j}$) ini bertujuan untuk mengetahui nilai perbandingan antara cluster ke- i dan cluster ke- j untuk menghitung nilai rasio yang dimiliki oleh masing- masing cluster. indeks i dan j merupakan merepresentasikan jumlah cluster, dimana jika terdapat 4 cluster maka terdapat indeks sebanyak 4 yaitu i, j, k dan l . untuk menentukan nilai rasio dengan rumus sebagai berikut :

$$R_{ij, \dots, n} = \frac{SSW_i + SSW_j + \dots + SSW_n}{SSB_{i,j} + \dots + SSB_{nl, nj}} =$$

Dimana :

SSW_i = Sum Of Square Within- Cluster pada centroid i

$SSB_{i,j}$ = Sum of Square Between Cluster data ke i dengan j pada *cluster* yang berbeda

Pada rumus perhitungan 2.5 n akan berlanjut sejumlah *cluster* yang dipilih dengan syarat ni tidak sama dengan nj .

4. Davies Bouldin Index (DBI)

Nilai rasio yang diperoleh dari rumus 2.5 digunakan untuk mencari nilai DBI dengan menggunakan perhitungan sebagai berikut :

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (Ri, j, \dots k)$$

Dimana, Ri, j merupakan ratio dari nilai SSW dan SSB melalui perhitungan rumus 2.5 dari perhitungan 2.6 maka dapat diketahui k adalah jumlah *cluster*. Dari perhitungan *Davies Bouldin Index* (DBI) dapat disimpulkan bahwa jika semakin kecil nilai *Davies Bouldin Index* (DBI) yang diperoleh (non negatif ≥ 0) maka *cluster* tersebut semakin baik.

4. Hasil dan Pembahasan

4.1 Pengujian algoritma K-Means

Pada tahap pengujian ini dilakukan eksperimen menggunakan algoritma K-Means, untuk mengetahui hasil dari cluster data yang akan dikelompokkan dan hasil akurasi dari Algoritma K-Means.

Pada tahap pengujian ini dilakukan eksperimen menggunakan algoritma K-Means, untuk mengetahui hasil dari cluster data yang akan dikelompokkan dan hasil akurasi dari Algoritma K-Means. Pada penelitian ini jumlah data set yang digunakan sebanyak 334 data kelas 10, 11 dan 12 dari semua jurusan di SMK Binamitra. Dan Perhitungan akan di ambil dari kelas 10 dari 108 siswa saja untuk contoh.

Tabel 1. Tabel Siswa Kelas 10

No	Nama	Kelas 10 (Jurusan)	RT2 NILAI ULNGN HRN	NILAI ASLI PTS	NILAI AKHIR PTS
1	Ach Syaqui Al-Hakim	TKR	7	1	6
2	Ade Indra Febrian	TKR	5	1	5
3	Ade Priana	TKR	7	2	7
4	Agil Setiawan	TKR	7	1	6
5	Ananta Surya Jaya	TKR	7	2	7
6	Antoni Diyan Anugre	TKR	7	1	5
7	Arya Bagus Susanto	TKR	7	2	7
8	Didan Surul Mulyadi	TKR	7	1	6
9	Dimas Adhitia	TKR	5	1	5
10	Erwin	TKR	7	2	7
...
40	Intan Sopiyantri	AKUTANSI	8	5	8
41	Keywa Guardian	AKUTANSI	7	4	6
42	Kurtubi Ziafahrezi	AKUTANSI	7	2	5
43	Mia Aditya	AKUTANSI	8	4	7
79	Aldi Junaedi	TKJ	2	3	2
80	Syahfitri Azzahra	TKJ	6	7	6
106	Fitria Anggrayeni	TKJ	6	7	6
107	Putri Wulandari	TKJ	6	7	7
108	Cahyo Agil Pratama	TKJ	3	2	4

Tabel 2. Centroid awal kelas 10

	Kelas X (Jurusan)	RT2 NILAI ULNGN HRN	NILAI ASLI PTS	NILAI AKHIR PTS
C1	2	7	1	6
C2	2	5	1	5

1. Hitung jarak data dengan centroid Eucliden Distance dengan rumus sebagai berikut:

$$(X_i, \mu_j) = \sqrt{(X_i - \mu_j)^2}$$

a. C1 (data 1, Centroid 1)

$$= \sqrt{(2-2)^2 + (7-7)^2 + (1-1)^2 + (6-6)^2}$$

$$= 0$$

b. C2 (data 1, centroid 2)

$$= \sqrt{(2-2)^2 + (7-5)^2 + (1-1)^2 + (6-5)^2}$$

$$= 2.236068$$

Tabel 3 .Hasil Perhitungan Jarak Pusat Cluster kelas 10

No	C1	C2	JARAK MINIMAL
1	0	2,236068	0
2	2,236068	0	0
3	1,414214	3	1,414214
4	0	2,236068	0
5	1,414214	3	1,414214
6	1	2	1

	5,477226	3,605551	3,605551
104	6,082763	4,242641	4,242641
105	6,480741	6,403124	6,403124
106	6,164414	6,244998	6,164414
107	6,244998	6,480741	6,244998
108	4,690416	2,645751	2,645751

Untuk centroid baru didapatkan dari jumlah seluruh data sebuah atribut pada satu centroid dibagi dengan jumlah data dan berlaku untuk seterusnya untuk semua atribut centroid. Sebagai contoh untuk atribut umur pada centroid pertama.

Proses kembali lagi ke langkah 2 yaitu dengan menggunakan centroid baru dari iterasi pertama yang dihitung dari nilai rata-rata tiap kelompok cluster. Untuk centroid baru didapatkan dari jumlah seluruh data sebuah atribut pada satu centroid dibagi dengan jumlah data dan berlaku untuk seterusnya untuk semua atribut centroid. Sebagai contoh untuk atribut Jurusan pada kelas 10 pada centroid pertama :

$$\frac{\text{jumlah seluruh jurusan pada kelas 10 cluster pertama}}{\text{banyak data pada cluster pertama}} = \frac{158}{73} = 2,164$$

Tabel 4. Data Cluster Pertama Kelas 10

C1	No	Kelas X (Jurusan)	RT2 NILAI ULNGN HRN	NILAI ASLI PTS	NILAI AKHIR PTS
	1	2	7	1	6
	2	2	7	2	7
	3	2	7	1	6
	4	2	7	2	7
	5	2	7	1	5
	6	2	7	2	7
	7	2	7	1	6
	8	2	7	2	7
	9	2	7	1	6
	10	2	7	1	7

	71	1	6	7	6
	72	1	6	7	7
	73	1	3	2	4

Dimana perhitungan pengulangan pada penelitian ini dengan centroid baru berhenti pada perhitungan iterasi ke 2. Dimana *centroid* terakhir yang tidak mengalami perubahan.

Tabel 5. Centroid Terakhir kelas 10

	Kelas X (Jurusan)	RT2 NILAI ULNGN HRN	NILAI ASLI PTS	NILAI AKHIR PTS
C1	2,076923077	7,153846	4,384615	6,794872
C2	1,133333333	3,366667	3,366667	3,066667

Tabel 6. Hasil Iterasi Terakhir kelas 10

No	Nama	Kelas 10 (Jurusan)	RT2 NILAI ULNGN HRN	NILAI ASLI PTS	NILAI AKHIR PTS	C1	C2	JARAK MINIMAL
1	Ach Syaqui Al-Hakim	2	7	1	6	3,471448	5,418692	3,471448
2	Ade Indra Febrian	2	5	1	5	4,411486	8,664545	3,664545
3	Ade Priana	2	7	2	7	2,391155	5,694637	2,391155
4	Agil Setiawan	2	7	1	6	3,471448	5,418692	3,471448
5	Ananta Surya Jaya	2	7	2	7	2,391155	5,694637	2,391155
6	Antoni Diyan Anugrah	2	7	1	5	3,826316	4,949298	3,826316
7	Arya Bagus Susanto	2	7	2	7	2,391155	5,694637	2,391155
8	Didan Surul Mulyadi	2	7	1	6	3,471448	5,418692	3,471448
9	Dimas Adhitha	2	5	1	5	4,411486	8,664545	3,664545
10	Erwin	2	7	2	7	2,391155	5,694637	2,391155
11	Farhan Rahmansyah	2	7	1	6	3,471448	5,418692	3,471448
12	Hendi Hadwijaya	2	7	1	7	3,385441	6,019044	3,385441
13	Herdi Eka Prasetya	2	7	1	7	3,385441	6,019044	3,385441
14	Herfila Ikhsan Ade Putra	2	7	1	6	3,471448	5,418692	3,471448
...
105	Adelia Cahyati	1	5	7	7	3,605506	5,618027	3,605506
106	Fitria Anggrayani	1	6	7	6	3,190479	5,406375	3,190479
107	Putri Wulandari	1	6	7	7	3,096678	6,007958	3,096678
108	Cahyo Agil Pratama	1	3	2	4	5,683957	8,701633	1,701633

3.1 Evaluasi Davies Bouldin Index (DBI) =

$$\begin{aligned}
 SSW_1 &= \frac{3.471+2.391+3.471+2.391+3.826+2.391+\dots+3.190+3.096}{78} \\
 &= \frac{403.4772}{78} \\
 &= 5.17278 \\
 SSW_2 &= \frac{3.666+3.664+3.6444+4.650+1.869+\dots+1.167+1.701}{30} \\
 &= \frac{403.4772}{30} \\
 &= 18.39951
 \end{aligned}$$

Setelah mengetahui nilai SSW maka selanjutnya dilakukan perhitungan Sum of square between-cluster (SSB). Untuk menghitung nilai SSB maka diperlukan centroid terakhir pada iterasi terakhir pada. Perhitungan SSB dilakukan sebagai berikut:

$$\begin{aligned}
 SSB_{1,2} &= \sqrt{(2.079 - 1.133)^2 + (7.153 - 3.366)^2 + (4.384 - 3.336)^2 + (6.794 - 3.066)^2} \\
 &= 5.492
 \end{aligned}$$

Setelah nilai SSW dan nilai SSB telah dihitung dan mendapatkan hasilnya, maka selanjutnya adalah mencari nilai rasio antar cluster dengan perhitungan sebagai berikut :

$$\begin{aligned}
 R_1 &= \frac{SSW}{SSB} \\
 R_1 &= \frac{5.17278}{5.492} = 0.941 \\
 R_2 &= \frac{SSW}{SSB} \\
 R_2 &= \frac{1.839951}{5.492} = 0.335
 \end{aligned}$$

Setelah mengetahui rasio antar cluster, kemudian hitunglah nilai DBI sebagai berikut :

$$\begin{aligned}
 DBI &= \frac{R_1+R_2}{K} \text{ atau } \frac{R_1+R_2}{2} \\
 &= \frac{0.941+0.335}{2} \\
 &= 0.638
 \end{aligned}$$

Setelah melakukan pengklasteran data dan melakukan pengujian menggunakan sistem *k-means* maka dapat disimpulkan bahwa Kelas 10 dari 2 *cluster* yaitu C1 (Minat Matematika) dengan jumlah siswa 78 siswa, *cluster* C2 (Tidak Minat Matematika) dengan jumlah siswa 30, dapat dilihat jumlah siswa kelas 10 minat terhadap mata pelajaran matematika lebih banyak dibandingkan kurang minat. Berdasarkan hasil analisa persentase dari, 2 *cluster* didapatkan hasil 82% yang minat dan 18% untuk siswa yang kurang minat.

Tabel 7. Hasil Clustering Kelas 10

Cluster 1 (C1)	78	80%	Minat Matematika
Cluster 2 (C2)	30	20%	Kurang Minat
Total	108	100%	

Evaluasi hasil dari Davies bouldin index (DBI) dari tiap kelas dan mendapatkan hasil nilai sebagai berikut :

Tabel 8. Hasil Clustering Kelas 10

KELAS	R1(C1)	R2 (C2)	DBI
10	0.941	0.335	0.638
11	0.660	0.506	0.583
12	0.271	0.111	0.191

Evaluasi hasil dari *Davies Bouldin Index* mendekati angka 0 mengartikan bahwa masing-masing anggota yang ada didalam *cluster* berada dalam jarak yang berdekatan. Semakin kecil nilai *davies bouldin index* maka semakin baik *cluster* yang diperoleh dari pengelompokan menggunakan metode *clustering*.

Hasil perhitungan dari kelas 10 menggunakan algoritma *k-means* menunjukkan nilai 0.638 angka tersebut memiliki hasil yang cukup baik karena mendekati angka 0. Dari perhitungan *Davies Bouldin Index* (DBI) dapat disimpulkan bahwa jika semakin kecil nilai *Davies Bouldin Index* (DBI) yang diperoleh (non negatif ≥ 0) maka *cluster* tersebut semakin baik.

5. Kesimpulan

Dari hasil penelitian yang telah dilakukan oleh peneliti , maka dapat diperoleh kesimpulan sebagai berikut:

- 1) Berdasarkan data yang telah diperoleh untuk membantu pengelompokan siswa dan siswi, untuk pengelompokan minat atau tidak minat dalam mata pelajaran matematika menggunakan metode *K-Means* dapat diimplementasikan di Smk Binamitra.
- 2) Dari Data kelas 10 dengan 2 *Cluster* di dapatkan hasil *cluster* 1 yaitu (Minat) Dengan Total Jumlah 78 data siswa, *cluster* 2 yaitu (Tidak Minat) Dengan Total Jumlah 30 data siswa, dengan jumlah siswa 108 siswa. Pengujian menggunakan validasi DBI (*Davies Bouldin Index*) diperoleh nilai untuk tiap-tiap *cluster*. pengujian *cluster* 1(Minat) menghasilkan nilai DBI 0.941, *cluster* 2 menghasilkan nilai DBI 0.335. Karena nilai DBI dari *cluster* 2 lebih kecil maka *cluster* tersebut bisa disebut optimal.
- 3) Data kelas 11 dengan 2 *Cluster* di dapatkan hasil *cluster* 1 yaitu (Minat) Dengan Total Jumlah 86 data siswa, *cluster* 2 yaitu (Tidak Minat) Dengan Total Jumlah 20 data siswa. Dengan jumlah siswa 106 siswa. Pengujian menggunakan validasi DBI (*Davies Bouldin Index*) diperoleh nilai untuk tiap-tiap *cluster*. pengujian *cluster* 1(Minat) menghasilkan nilai DBI 0.660, *cluster* 2 menghasilkan nilai DBI 0.506. Karena nilai DBI dari *cluster* 2 lebih kecil maka *cluster* tersebut bisa disebut optimal.
- 4) Data kelas 12 dengan 2 *Cluster* di dapatkan hasil *cluster* 1 yaitu (Minat) Dengan Total Jumlah 42 data siswa, *cluster* 2 yaitu (Tidak Minat) Dengan Total Jumlah 78 data siswa. Dengan jumlah siswa 120 siswa.

Penguji menggunakan validasi DBI (*Davies Bouldin Index*) diperoleh nilai untuk tiap-tiap *cluster*. pengujian *cluster* 1(Minat) menghasilkan nilai DBI 0.271, *cluster* 2 menghasilkan nilai DBI 0.111. Karena nilai DBI dari cluster 2 lebih kecil maka *cluster* tersebut bisa disebut optimal.

Daftar Pustaka

- Marwan, "Pengertian Penjualan – Jenis, Tujuan, Faktor, Pasar, Para Ahli," [Online]. Available: <https://www.gurupendidikan.co.id/pengertian-penjualan/>. [Accessed 28 Juli 2019].
- M. Badrul, "Algoritma Asosiasi dengan Algoritma Apriori untuk Analisa Data Penjualan," *Jurnal Pilar Nusa Mandiri*, vol. XII, no. PP. 121-129, 2016.
- D. A. Nurdin, "Penerapan Data Mining Untuk Menganalisis Penjualan Barang Dengan Menggunakan Metode Apriori Pada Supermarket Sejahtera Lhokseumawe," *Techsi*, vol. 6, no. 1, April 2015.
- L. Landau, "Hakikat Matematika," *Zhurnal Eksp. i Teor. Fiz.*, pp. 12– 65, 1937.
- U. Ayyad, *Advances In Knowledge Discovery and Data Mining*, MIT Press, 1996. [9] K. d. E. T. Luthfi, "Algoritma Data Mining," 2009.
- R. A. Djamro, "Penerapan Algoritma K-Means Dalam Memilih Tanah Yang Tepat Untuk Tanaman Padi," vol. VII, no. 1, pp. 12–20, 2018.
- Y. O. a. D. D. D. Haryanto, "Implementasi Analisa Keranjang Belanja dengan Aturan Asosiasi Menggunakan Algoritma Apriori pada penjualan suku cadang Sepeda Motor," *Jurnal Buana Informatika*, pp. pp. 81-94, 2011.