



Penerapan Algoritma K-Means Clustering untuk Melihat Hubungan Kegiatan Tahfiz dengan Hasil Belajar

Asri Hidayat^{1✉}, Sarjon Defit², Sumijan³
^{1,2,3}Universitas Putra Indonesia YPTK Padang

asrihidayat333@gmail.com

Abstract

The purpose of this study is to evaluate whether Tahfiz activities and learning outcomes are effective or not. The data processed in this study were data on tahfiz activities and data on student learning outcomes in class XI (eleven) totaling 42 data sourced from memorization of tahfiz, tahfiz grades, and student grades in Madrasah Aliyah Negeri 1 Bukittinggi. Based on the analysis of the data, this classification uses one of the methods of the Data Mining algorithm, K-Means Clustering. K-Means Clustering algorithm works based on the grouping method. In this data mining technique consists of data testing and training data with the input of the number of memorization of tahfiz, and the value of tahfiz as well as learning outcomes. The results of this study the school can determine how influential this activity tahfiz on student grades.

Keywords: Tahfiz Activity, Learning Outcomes, K-means Clustering, Data Mining, Student.

Abstrak

Tujuan dari penelitian ini untuk mengevaluasi kegiatan tahfiz dan hasil belajar apakah sudah efektif atau belum. Data yang diolah dalam penelitian ini adalah data kegiatan tahfiz dan data hasil belajar siswa kelas XI (sebelas) sebanyak 42 data yang bersumber dari data hafalan tahfiz, nilai tahfiz, dan nilai siswa di Madrasah Aliyah Negeri 1 Bukittinggi. Berdasarkan analisis terhadap data, klasifikasi ini menggunakan salah satu metode algoritma *Data Mining* yaitu *K-Means Clustering*. Algoritma *K-Means Clustering* bekerja berdasarkan metode pengelompokan. Dalam teknik data mining ini terdiri dari data testing dan data training dengan inputan jumlah hafalan tahfiz, dan nilai tahfiz serta hasil belajar. Hasil penelitian ini pihak sekolah dapat menentukan seberapa berpengaruhnya kegiatan tahfiz ini terhadap nilai siswa.

Kata Kunci: Kegiatan Tahfiz, Hasil belajar, *K-means Clustering*, *Data Mining*, Siswa.

© 2020 JSisfotek

1. Pendahuluan

Saat ini pemanfaatan teknologi komputer sebagai alat bantu dalam mendukung kegiatan dalam bidang pendidikan untuk memudahkan guru dalam mengolah data siswa secara cepat, tepat, akurat sehingga tujuan dari suatu pekerjaan tersebut dapat tercapai secara efektif dan efisien oleh Triyansyah dan Fitriana, (2018) [1]. Pada proses pengelompokan data dan terjadi kesalahan dalam pengelompokan yang dilakukan. Melalui penelitian ini, dilakukan sebuah analisa terhadap kegiatan tahfiz dan nilai siswa dengan menerapkan metode *K-Means Clustering*.

Berdasarkan masalah tersebut, perlu adanya solusi yang tepat untuk permasalahan yang sedang dihadapi dalam melihat hubungan kegiatan tahfiz dengan hasil belajar siswa yang tepat. solusi yang digunakan untuk mewujudkannya adalah *Data Mining*. Menurut Irfiani dan Rani *Data Mining* merupakan sebuah cara untuk menemukan pola yang menarik dan pengetahuan dalam jumlah datayang besar. Pada data mining menggunakan teknik matematika, statistik, artificial intelegent serta machine learning untuk mengidentifikasi dan mengekstraksi menjadi informasi atau pengetahuan yang bermanfaat yang diambil dari berbagai database yang

besar [2]. Metode yang digunakan adalah metode *Clustering*. Menurut Sulastri dan Gufroni metode *Clustering* adalah suatu proses pengelompokan objek data yang mirip satu sama lain kedalam kluster yang sama dan berbeda dengan objek yang ada di kluster lain [3]. Algoritma yang digunakan dalam penelitian ini untuk menentukan startegi marketing adalah algoritma *K-Means Clustering*.

Menurut Rosmini, dkk. (2018) K-Means merupakan salah satu metode clustering atau pengelompokan non hirarki. Teknik pengelompokan datanya sederhana dan cepat. Ada banyak pendekatan untuk membuat cluster, diantaranya adalah membuat aturan yang mendikte keanggotaan dalam grup yang sama berdasarkan tingkat persamaan diantara anggota-anggotanya. Pendekatan lainnya adalah dengan membuat sekumpulan fungsi yang mengukur beberapa properti dari pengelompokan tersebut sebagai fungsi dari beberapa parameter dari sebuah clustering [4], dan jika menggunakan metode k-means, harus melakukan transmisi data agar dapat diolah, maka data yang berjenis data nominal seperti jurusan, data setoran ayat, nilai rapor semester, data nilai sikap siswa dan data nilai keterampilan siswa harus diinisialisasikan terlebih

dahulu kedalam bentuk angka [5]. dan jika data sudah dalam bentuk angka, nantinya akan diproses menggunakan *tools Rapid Miner*.

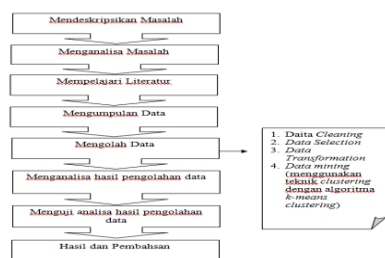
Dengan dilakukannya penelitiannya ini diharapkan dapat membantu pihak Hoyweapstore dalam menentukan strategi marketing yang tepat untuk meningkatkan laba penjualan dan mampu bersaing dengan kompetitor.

Penelitian terdahulu menurut Asroni, dkk (2018) pengujian yang dilakukan pada data penerimaan calon mahasiswa baru Fakultas Kedokteran dan Ilmu Keperawatan pada setiap *cluster* adalah sebagai berikut:

1. *Cluster 0* dengan jurusan Ilmu Keperawatan, sebanyak 3795 pendaftar dari jumlah 25000 calon mahasiswa baru (15%);
2. *Cluster 1* dengan jurusan Pendidikan Dokter, sebanyak 8296 pendaftar dari jumlah 25000 calon mahasiswa baru (33%);
3. *Cluster 2* dengan jurusan Kedokteran Gigi, sebanyak 7427 pendaftar dari jumlah 25000 calon mahasiswa baru (30%);
4. *Cluster 3* dengan jurusan Pendidikan Dokter, sebanyak 5482 pendaftar dari jumlah 25000 calon mahasiswa baru (22%)[5].

2. Metodologi Penelitian

Tahapan-tahapan yang dilakukan dalam penelitian adalah seperti pada gambar dibawah ini:



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja peneltitian pada gambar 1, maka dapat menjelaskan beberapa kerangka kerja yang akan dilakukan dalam penelitian ini, yaitu sebagai berikut :

2.1 Mendeskripsikan Masalah

Mendeskripsikan masalah adalah menjelaskan tentang masalah dalam penelitian secara terstruktur dan sistematis untuk mendukung dalam pengambilan keputusan yang lebih baik dan tepat. Masalah yang dideskripsikan dalam penelitian ini adalah melihat hubungan kegiatan tahfiz dengan hasil belajar.

2.2 Menganalisa Masalah

Menganalisa masalah adalah tahap di mana peneliti menganalisa permasalahan yang ada sebagai

langkah untuk dapat memahami masalah yang ada. Untuk melihat hubungan kegiatan tahfiz dengan hasil belajar yang ada selama ini dilakukan secara manual, sehingga proses pengolahan datanya belum optimal. Dengan penerapan *data mining* teknik *clustering* menggunakan algoritma *K-means clustering* membantu dalam menentukan kelompok siswa yang kegiatan tahfiz dan kelompok kegitan hasil belajar, yang mana akan diliat hubungan antara keduanya.

2.3 Mempelajari Literatur

Pada tahap ini akan dicari dan dikumpulkan serta dipelajari sejumlah literature mengenai teori dan konsep yang mendukung penyelesaian permasalahan dalam penelitian. Literatur yang digunakan berupa buku referensi atau buku penunjang, jurnal internasional dan nasional sertakonsep-konsep yang mendukung dalam menyelesaikan penelitian ini. Selain itu juga dilakukan studi kepustakaan yaitu dengan membaca buku-buku yang menunjang dalam melakukan penganalisisan terhadap data dan informasi yang didapat. Kemudian literatur-literatur yang dipelajari tersebut diseleksi untuk dapat ditentukan literatur-literatur mana yang akan digunakan dalam penelitian.

2.4 Mengumpulkan Data

Dalam pengumpulan data dilakukan observasi yaitu pengamatan secara langsung di tempat penelitian sehingga permasalahan yang ada dapat diketahui secara jelas. Kemudian dilakukan *interview* atau wawancara dengan guru atau personalia yang mengurus bagian penerima setoran ayat dari siswa MAN 1 Bukittinggi, yang bertujuan untuk mendapatkan informasi tentang hubungan antara kegitan tahfiz dengan hasil belajar serta data lain yang dibutuhkan. Data yang dikumpulkan adalah data siswa MAN 1 Bukittinggi semester genap TP. 2018/2019.

2.5 Mengolah data

Setelah data yang diperlukan dalam penelitian selesai dikumpulkan, langkah selanjutnya adalah melakukan pengolahan data dengan beberapa tahap diantaranya adalah :

2.5.1 *Data cleaning*, dilakukan untuk menghilangkan data yang tidak konsisten atau tidak relevan.

2.5.2 *Data selection*, dilakukan untuk mengambil data yang sesuai untuk dianalisa. Pada penelitian ini hanya beberapa atribut data saja yang digunakan, atribut tersebut yaitu NISN siswa, nama siswa, jurusan, data setoran ayat, nilai rapor semester, data nilai sikap siswa, data nilai keterampilan siswa.

2.5.3 *Data transformation*, agar data dapat diolah menggunakan algoritma *k-means clustering*, maka data yang berjenis data huruf seperti jurusan, data setoran ayat, nilai rapor semester, data nilai sikap siswa dan data nilai

keterampilan siswa harus di inialisasikan terlebih dahulu menjadi bentuk angka.

2.5.4 *Data mining*, setelah data ditransformasikan maka data tersebut dapat diproses menggunakan algoritma *k-means clustering*.

2.6 Menganalisa hasil pengolahan data

Pada tahap ini dilakukan analisa terhadap hasil pengolahan data berupa *cluster* hubungan kegiatan tahfiz dengan nilai siswa yang telah diolah dengan menggunakan algoritma *k-means clustering*. Analisa hasil pengolahan data terdiri dari analisa setiap cluster yang meliputi analisa jumlah siswa setiap cluster dengan rincian jumlah siswa per jurusan, data setoran ayat, nilai rapor semester, data nilai sikap siswa dan data nilai keterampilan siswa. Analisa terhadap hasil setiap *cluster* dapat digunakan untuk menentukan calon penerima beasiswa kurang mampu di MAN 1 Bukittinggi.

2.7 Menguji analisa hasil pengolahan data

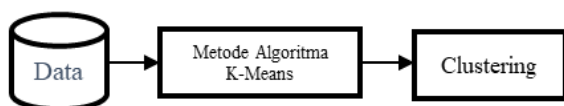
Hasil analisa pengolahan data yang dilakukan dengan teknik *clustering* menggunakan algoritma *k-means clustering* selanjutnya dilakukan pengujian. Pengujian dilakukan terhadap data yang diolah secara manual berdasarkan langkah-langkah algoritma *k-means clustering*, kemudian disbanding kandungan hasil perhitungan menggunakan Rapid Miner. Pengujian dilakukan terhadap data sampel dan seluruh data yang meliputi pengujian iterasi dan hasil *cluster*.

2.8 Menarik kesimpulan dan rekomendasi

Langkah terakhir setelah semua proses selesai, maka penulis dapat menarik suatu kesimpulan terhadap masalah yang diteliti. Selanjutnya diberikan rekomendasi dengan tujuan kelemahan-kelemahan dapat dihilangkan dan solusi kedepannya dapat diimplementasikan sehingga tujuan yang diharapkan dapat dicapai.

3. Hasil dan Pembahasan

Pada tahap ini data yang telah dikumpulkan akan dianalisis menggunakan metode *K-Means Clustering*. Analisis ini bertujuan untuk memperoleh pengetahuan untuk melihat hubungan kegiatann tahfiz dengan hasil belajarr siswa. Hasil prediksi ini akan digunakan oleh pihak MAN 1 Bukittinggi untuk mengevaluasi nilai tahfiz dengan hasil belajar siswa. Berdasarkan pembahasan dari sebelumnya dapat dibentuk bagan alir dari sistem yang akan dirancang seperti Gambar 2.



Gambar 2 Arsitektur Sistem Yang Dirancang

Data yang digunakan dalam penelitian ini adalah data yang didapat dari MAN 1 Bukittinggi. Peneliti meminta kepada waka kurikulum dengan menyerahkan surat

penelitian yang diperoleh dari pihak sekolah. Sehingga pihak MAN 1 Bukittinggi dengan tidak keberatan bersedia menyerahkan data yang diperlukan oleh peneliti yakni nilai tahfiz, jumlah hafalan, dan hasil belajar tahun ajaran 2018-2019. Berikut data yang diberikan oleh pihak MAN 1 Bukittinggi dapat dilihat pada Tabel 1.

Tabel 1. Data Awal

No	Nama	Nilai Angka	Rata-Rata
1.	Zulfitra	98	88,83
2.	Muhammad Wahyudi Arafah	84	85,38
3.	Amirul Mukminin	84	85,48
4.	Olivia Levan's	62	87,23
5.	Annisa Yulia Dasman	95	87,18
6.	Dzurriyyatul'afifah	95	87,73
7.	Afifah Athohirah	79	83,83
8.	Amamatur Rusydia Mufid	90	83,83
9.	Khairani	90	84,48
10.	Jerri Dwi Alkarizmi	64	80,78
11.	Yusril Yamanda	95	84,8
12.	Fikri Estiawan	76	85,88
13.	Nadia Yustika R.	95	84,73
14.	Alhusna Melfi	92	87,68
15.	Zahra Zakiah Raihanisa	79	84,72
16.	Andini	92	86,14
17.	Haura Fadhila Dzamri	79	85,25
18.	Farzana Quratul Aini	94	85,86
19.	Rangga	79	82,28
20.	Yaumul Khair Djumardi	92	89,31
21.	Feri Gunawan	84	83,39
22.	Fauziah Az Zahro	95	87,89
23.	Muhammad Farhab Shiddiq	64	85,53
24.	Wilza Lolisya Shandori	79	84,44
25.	Bogy Advantage	84	85,22
26.	Habibatul Khairiah	95	88,81
27.	Miftahus Sa'adah	84	90,28
28.	Annisa Fadhillah	98	84,81
29.	Nahda Navilah Daulai	59	87,28
30.	Nurmishbah Suduri	95	87,69
31.	Mufidah Asri Azizah	75	85,5
32.	Wardah Hafifah	64	82,81
33.	Megita Irza Mulyani	96	86,31
34.	Rizki Ananta	84	85,06
35.	Yolanda Agustina	84	82,47
36.	Yudi Alfikri	64	80,08
37.	Abdurrazaq	92	81,61
38.	Najla Amatullah	64	84,53
39.	M. Imam Faisal	84	81,11
40.	Fadhlan Habib	79	82,67
41.	M. Alfattah	64	82,92
42.	Fadila Khairunnisa	76	83,19
43.	Dwi Fitri R.	95	84,11
44.	Teguh Sabrian T.	84	79,83
45.	Muhammad Akthar	84	80,86

Data sampel yang telah diperoleh selajutnya akan diolah dengan menggunakan metode *K-Means Clustering* agar mendapatkan hasil berupa pengelompokan data yang diharapkan sebelumnya. Adapun algoritma yang terdapat pada *K-Means* adalah sebagai berikut:

a. Menentukan jumlah Cluster

Dalam penelitian ini, peneliti mengklasterisasi data menjadi 2 Cluster, yaitu C0 dan C1. C0 untuk kelompok yang tidak berpengaruh, C1 untuk kelompok berpengaruh.

b. Menentukan Titik Pusat Cluster Centroid Awal

Dalam melakukan penentuan nilai *centroid* awal ini dapat diambil secara acak dari data sampel yang digunakan dalam penelitian. *Centroid* awal ini merupakan titik pusat *cluster* pertama. *Centroid* untuk C0 diambil dari data sampel ke 3 dan C1 diambil dari data sampel ke 42. Berikut ini adalah bentuk dari *centroid* awal yang telah dituangkan dalam bentuk Tabel 2.

Tabel 2. Titik Pusat / Centroid Awal

No	Nama	Nilai Angka	Rata-Rata
3	Amirul Mukminin	84	85,48
42	Fadila Khairunnisa	76	83,19

c. Menghitung Jarak dari Centroid awal

Formula yang dipakai untuk menghitung jarak antara titik *Centroid* dengan titik tiap objek yaitu menggunakan *Euclidian Distance*. Adapun perhitungan *Centroid* awal secara manual dengan rumus:

$$D(m,n) = \sqrt{(x_{1m} - x_{1n})^2 + (x_{2m} - x_{2n})^2 + \dots + (x_{km} - x_{kn})^2} \dots\dots(1)$$

1. Perhitungan jarak dari data ke-1 terhadap Cluster

$$DC0 = \sqrt{(98 - 84)^2 + (88,83 - 85,48)^2} = 14,395$$

$$DC1 = \sqrt{(98 - 76)^2 + (88,83 - 84,11)^2} = 22,711$$

2. Perhitungan jarak dari data ke-1 terhadap Cluster

$$DC0 = \sqrt{(84 - 84)^2 + (85,38 - 85,48)^2} = 0,100$$

$$DC = \sqrt{(84 - 76)^2 + (85,38 - 84,11)^2} = 8,294$$

3. Perhitungan jarak dari data ke-1 terhadap Cluster

$$DC0 = \sqrt{(84 - 84)^2 + (85,48 - 85,48)^2} = 0$$

$$DC1 = \sqrt{(84 - 76)^2 + (85,48 - 84,11)^2} = 8,321$$

4. Perhitungan jarak dari data ke-1 terhadap Cluster

$$DC0 = \sqrt{(62 - 84)^2 + (87,23 - 85,48)^2} = 22,069$$

$$DC1 = \sqrt{(62 - 76)^2 + (87,23 - 84,11)^2} = 14,571$$

5. Perhitungan jarak dari data ke-1 terhadap Cluster

$$DC0 = \sqrt{(95 - 84)^2 + (87,18 - 85,48)^2} = 11,131$$

$$DC1 = \sqrt{(95 - 76)^2 + (87,18 - 84,11)^2} = 19,414$$

d. Hasil perhitungan dataset dari jarak iterasi pertama

Setelah hasil perhitungan *dataset* dari jarak data ke-1,2,3,4,5.....45 dengan data *Centroid* yang didapat, lalu dilakukan alokasi atau menempatkan anggota *Cluster* berdasarkan nilai terkecil. Alokasi anggota *Cluster* dikelompokkan dengan memberi kode "C0" apabila bilangan hasil perhitungan jarak terdekatnya terletak pada kelompok pertama, pemberian kode "C1" diberikan apabila bilangan terdapat pada kelompok yang kedua. Berikut ini adalah bentuk dari pengelompokan anggota *cluster* berdsarkan jarak terdekat yang telah dituangkan dalam bentuk Tabel 3.

Tabel 3. Pengelompokan Anggota Cluster Berdsarkan Jarak Terdekat

No	Nama	DC0	DC1	Cluster
1.	Zulfitra	14,395	22,711	C0
2.	Muhammad Wahyudi Arafah	0,100	8,294	C0
3.	Amirul Mukminin	0	8,321	C0
4.	Olivia Levan's	22,069	14,571	C1
5.	Annisia Yulia Dasman	11,131	19,414	C0
6.	Dzurriyyatul'afifah	11,228	19,535	C0
7.	Afifah Athohirah	5,265	3,068	C1
8.	Amamatur Rusydia Mufid	6,223	14,015	C0
9.	Khairani	6,083	14,059	C0
10.	Jerri Dwi Alkarizmi	20,545	12,240	C1
11.	Yusril Yamanda	11,021	19,068	C0
12.	Fikri Estiawan	8,010	2,690	C1
13.	Nadia Yustika R.	11,026	19,062	C0
14.	Alhusna Melfi	8,297	16,618	C0
15.	Zahra Zakiah Raihanisa	5,057	3,368	C1
16.	Andini	8,027	16,270	C0
17.	Haura Fadhila Dzamri	5,005	3,639	C1
18.	Farzana Quratul Aini	10,007	18,197	C0
19.	Rangga	5,936	3,135	C1
20.	Yaumul Khair Djumardi	8,870	17,131	C0
21.	Feri Gunawan	2,090	8,002	C0
22.	Fauziah Az Zahro	11,261	19,573	C0
23.	Muhammad Farhab Shiddiq	20,000	12,226	C1
24.	Wilza Lolisya Shandori	5,107	3,250	C1
25.	Bogy Advantage	0,260	8,254	C0
26.	Habibatul Khairiah	11,493	19,814	C0
27.	Miftahus Sa'adah	4,800	10,690	C0
28.	Annisia Fadhillah	14,016	22,060	C0
29.	Nahda Navilah Daulai	25,065	17,485	C1
30.	Nurmisbah Suduri	11,220	19,526	C0
31.	Mufidah Asri Azizah	9,000	2,517	C1
32.	Wardah Hafifah	20,177	12,006	C1
33.	Megita Irza Mulyani	12,029	20,242	C0
34.	Rizki Ananta	0,420	8,216	C0
35.	Yolanda Agustina	3,010	8,032	C0
36.	Yudi Alfikri	20,716	12,396	C1
37.	Abdurrazq	8,887	16,078	C0
38.	Najla Amatullah	20,023	12,075	C1
39.	M. Imam Faisal	4,370	8,266	C0
40.	Fadhlhan Habib	5,736	3,045	C1
41.	M. Alfattah	20,163	12,003	C1
42.	Fadila Khairunnisa	8,321	0	C1
43.	Dwi Fitri R.	11,085	19,022	C0
44.	Teguh Sabrian T.	5,650	8,677	C0
45.	Muhammad Akthar	4,620	8,332	C0

Dari hasil iterasi pertama telah dikelompokkan perhitungan dataset terdekat di mana :

DC0 = Hasil perhitungan dengan *Centroid* 0;

DC1 = Hasil perhitungan dengan *Centroid* 1.

Pengelompokan iterasi pertama pada *Cluster* C0 ada 28 item dalam kategori tidak berpengaruh dapat dilihat pada Tabel 4.

Tabel 4. Anggota Cluster C0 dengan Tidak Berpengaruh

No	Nama	DC0	DC0	Cluster
----	------	-----	-----	---------

1.	Zulfitra	14,395	22,711	C0
2.	Muhammad	0,100	8,294	C0
3.	Wahyudi Arafah	0	8,321	C0
4.	Annisa Yulia	11,131	19,414	C0
5.	Dasman	11,228	19,535	C0
6.	Dzurriyyatul'afifah	6,223	14,015	C0
7.	Amamatur Rusydiah	6,083	14,059	C0
8.	Mufid	11,021	19,068	C0
9.	Khairani	11,026	19,062	C0
10.	Yusril Yamanda	8,297	16,618	C0
11.	Nadia Yustika R.	8,027	16,270	C0
12.	Alhusna Melfi	10,007	18,197	C0
13.	Andini	8,870	17,131	C0
14.	Farzana Quratul Aini	2,090	8,002	C0
15.	Yaumil Khair	11,261	19,573	C0
16.	Djumardi	0,260	8,254	C0
17.	Feri Gunawan	11,493	19,814	C0
18.	Fauziah Az Zahro	4,800	10,690	C0
19.	Bogy Advantage	14,016	22,060	C0
20.	Habibatul Khairiah	11,220	19,526	C0
21.	Miftahus Sa'adah	12,029	20,242	C0
22.	Annisa Fadhillah	0,420	8,216	C0
23.	Nurmishbah Suduri	3,010	8,032	C0
24.	Megita Irza Mulyani	8,887	16,078	C0
25.	Rizki Ananta	4,370	8,266	C0
26.	Yolanda Agustina	11,085	19,022	C0
27.	Abdurrazq	5,650	8,677	C0
28.	M. Imam Faisal	4,620	8,332	C0
29.	Dwi Fitri R.			C0
30.	Teguh Sabrian T.			C0
31.	Muhammad Akthar			C0

Cluster C1 terdiri 17 item sebagai Cluster dengan kategori berpengaruh yang nantinya siswa yang berpengaruh tersebut semoga dapat ditingkatkan lagi kegiatan tahfizny dan hasil belajarnya, dari pengelompokan Cluster 1 dapat dilihat pada Tabel 5.

Tabel 5 Anggota Cluster C1 dengan Kategori Berpengaruh

No	Nama	DC0	DC0	Cluster
1	Olivia Levan's	22,069	14,571	C1
2	Afifah Athohirah	5,265	3,068	C1
3	Jerri Dwi Alkarizmi	20,545	12,240	C1
4	Fikri Estiawan	8,010	2,690	C1
5	Zahra Zakiah Raihanisa	5,057	3,368	C1
6	Haura Fadhila Dzamri	5,005	3,639	C1
7	Rangga	5,936	3,135	C1
8	Muhammad Farhab Shiddiq	20,000	12,226	C1
9	Wilza Lolisya Shandori	5,107	3,250	C1
10	Nahda Navilah Daulai	25,065	17,485	C1
11	Mufidah Asri Azizah	9,000	2,517	C1
12	Wardah Hafifah	20,177	12,006	C1
13	Yudi Alfikri	20,716	12,396	C1
14	Najla Amatullah	20,023	12,075	C1
15	Fadhlan Habib	5,736	3,045	C1
16	M. Alfattah	20,163	12,003	C1
17	Fadila Khairunnisa	8,321	0	C1

e. Pengelompokan Anggota Cluster

Tiap-tiap anggota Cluster telah dikelompokan, kemudian akan dilakukan iterasi dengan membangkitkan nilai *Centroid* baru berdasarkan penjumlahan anggota Cluster dibagi dengan total anggota Cluster, dengan persamaan sebagai berikut:

$$\mu = \frac{\sum_{i=1}^n x_i}{n}; i = 1, 2, 3 \dots n \dots \dots \dots (2)$$

Tabel 6 Perhitungan Mendapatkan Nilai Centroid Baru

Cluster	Nilai Angka	Rata-Rata
C0	98+84+84+95+95+9	88,83+85,38+85,48+87,18
	0+90+95+95+92+92	+87,73+83,83+84,48+84,8
	94+92+84+95+84+9	+84,73+87,68+86,14+85,8
	5+84+98+95+96+84	6+89,31+83,39+87,89+85,
	84+92+84+95+84+8	22+88,81+90,28+84,81+8
C0	4/28	7,69+86,31+85,06+82,47+
		81,61+81,11+84,11+79,83
		+80,86/28
	62+79+64+76+79+7	87,23+83,83+80,78+85,88
	9+79+64+79+59+75	+84,72+85,25+82,28+85,5
C0	+64+64+64+79+64+	3+84,44+87,28+85,5+82,8
	76/17	1+80,08+84,53+82,67+82,
		92+83,19/17

Dari hasil perhitungan penjumlahan pada Tabel 6 didapatkan pusat Cluster baru sebagaimana dapat dilihat pada Tabel 7.

Tabel 7. Centroid baru iterasi kedua

Cluster	Nilai Angka	Rata-Rata
C0	90,5	85,389
C1	70,941	84,054

Setelah mendapatkan hasil dengan *Centroid* baru maka lakukan lagi perhitungan jarak antara titik *Centroid* baru dengan titik tiap objek dataset kembali yang mana data dataset nya pada tabel 1

Untuk mendapatkan hasil eterasi kedua maka gunakan kembali rumus *Euclidion Distance* untuk menghitung jarak Cluster dengan *Centroid* baru yang telah didapatkan hasil dari perhitungan iterasi pertama. Hasil perhitungan dataset dari jarak ietrassi kedua. Setelah hasil perhitungan dataset dari jarak data ke-1,2,3,4,5.....45 dengan data *Centroid* baru yang didapat, lalu dilakukan alokasi atau menempatkan anggota Cluster berdasarkan nilai terkecil. Iterasi selanjutnya melakukan perhitungan lagi dengan menggunakan titik *Centroid* yang baru iterasi ke-2. Adapun hasil perhitungan iterasi ke-2 dapat dilihat pada Tabel 8.

Tabel 8. Perhitungan Jarak dan Pengelompokan Data Iterasi Ke-2

No	Nama	DC0	DC1	Cluster	No	Nama	DC0	DC0	Cluster
1.	Zulfitra	8,252	27,477	C0	1.	Zulfitra	8,252	27,477	C0
2.	Muhammad Wahyudi Arafah	6,500	13,126	C0	2.	Muhammad Wahyudi Arafah	6,500	13,126	C0
3.	Amirul Mukminin	6,501	13,137	C0	3.	Amirul Mukminin	6,501	13,137	C0
4.	Olivia Levan's	28,559	9,488	C1	4.	Annisa Yulia Dasman	4,843	24,261	C0
5.	Annisa Yulia Dasman	4,843	24,261	C0	5.	Dzurriyyatul'afifah	5,073	24,338	C0
6.	Dzurriyyatul'afifah	5,073	24,338	C0	6.	Amamatur Rusydia Mufid	1,637	19,060	C0
7.	Afifah Athohirah	11,605	8,062	C1	7.	Khairani	1,037	19,064	C0
8.	Amamatur Rusydia Mufid	1,637	19,060	C0	8.	Yusril Yamanda	4,538	24,071	C0
9.	Khairani	1,037	19,064	C0	9.	Nadia Yustika R.	4,548	24,068	C0
10.	Jerri Dwi Alkarizmi	26,898	7,674	C1	10.	Alhusna Melfi	2,738	21,369	C0
11.	Yusril Yamanda	4,538	24,071	C0	11.	Andini	1,677	21,162	C0
12.	Fikri Estiawan	14,508	5,378	C1	12.	Farzana Quratul Aini	3,532	23,130	C0
13.	Nadia Yustika R.	4,548	24,068	C0	13.	Yaumil Khair Djumardi	4,198	21,705	C0
14.	Alhusna Melfi	2,738	21,369	C0	14.	Feri Gunawan	6,800	13,076	C0
15.	Zahra Zakiah Raihanisa	11,519	8,086	C1	15.	Fauziah Az Zahro	5,148	24,363	C0
16.	Andini	1,677	21,162	C0	16.	Bogy Advantage	6,502	13,111	C0
17.	Haura Fadhlila Dzamri	11,501	8,147	C1	17.	Habibatul Khairiah	5,653	24,525	C0
18.	Farzana Quratul Aini	3,532	23,130	C0	18.	Miftahus Sa'adah	8,135	14,467	C0
19.	Rangga	11,913	8,252	C1	19.	Annisa Fadhillah	7,522	27,070	C0
20.	Yaumil Khair Djumardi	4,198	21,705	C0	20.	Nurmisbah Suduri	5,054	24,332	C0
21.	Feri Gunawan	6,800	13,076	C0	21.	Megita Irza Mulyani	5,577	25,160	C0
22.	Fauziah Az Zahro	5,148	24,363	C0	22.	Rizki Ananta	6,508	13,098	C0
23.	Muhammad Farhab Shiddiq	26,500	7,096	C1	23.	Yolanda Agustina	7,125	13,155	C0
24.	Wilza Lolisa Shandori	11,539	8,068	C1	24.	Abdurrazq	4,066	21,200	C0
25.	Bogy Advantage	6,502	13,111	C0	25.	M. Imam Faisal	7,782	13,387	C0
26.	Habibatul Khairiah	5,653	24,525	C0	26.	Dwi Fitri R.	4,678	24,059	C0
27.	Miftahus Sa'adah	8,135	14,467	C0	27.	Teguh Sabrian T.	8,553	13,725	C0
28.	Annisa Fadhillah	7,522	27,070	C0	28.	Muhammad Akthar	7,922	13,444	C0
29.	Nahda Navilah Daulai	31,557	12,369	C1					
30.	Nurmisbah Suduri	5,054	24,332	C0					
31.	Mufidah Asri Azizah	15,500	4,309	C1					
32.	Wardah Hafifah	26,625	7,052	C1					
33.	Megita Irza Mulyani	5,577	25,160	C0					
34.	Rizki Ananta	6,508	13,098	C0					
35.	Yolanda Agustina	7,125	13,155	C0					
36.	Yudi Alfikri	27,027	7,998	C1					
37.	Abdurrazq	4,066	21,200	C0					
38.	Najla Amatullah	26,514	6,957	C1					
39.	M. Imam Faisal	7,782	13,387	C0					
40.	Fadhlan Habib	11,817	8,177	C1					
41.	M. Alfattah	26,615	7,033	C1					
42.	Fadila Khairunnisa	14,666	5,132	C1					
43.	Dwi Fitri R.	4,678	24,059	C0					
44.	Teguh Sabrian T.	8,553	13,725	C0					
45.	Muhammad Akthar	7,922	13,444	C0					

Setelah perhitungan pada iterasi kedua dilakukan dan diperoleh anggota *Cluster* yang sama seperti iterasi yang pertama seperti pada Tabel 4.3. maka proses iterasi dihentikan. Algoritma *K-Means* yang telah digunakan dalam perhitungan diatas adalah algoritma yang mempartisi data kedalam *Cluster-cluster* yang sama dan data yang memiliki ketidaksamaan berada pada *Cluster* yang lain. Adapun analisa hasil dari pengelompokan 2 anggota *Cluster* yang diperoleh yang dapat dilihat pada Tabel 9.

Tabel 9. Anggota *Cluster* C0 dengan Tidak Berpengaruh

C0 terdiri dari 28 item sebagai *Cluster* dengan kategori tidak berpengaruh yang nantinya dipertahankan lagi, karena pada *Cluster* "C0" yang nantinya siswa-siswa yang kegiatan tahfizny tidak mempengaruhi hasil belajar mereka. Oleh karena itu pada *cluster* "C0" apat disimpulkan bahwa hubungan kegiatan tahfiz dengan hasil belajar siswa tidak mempengaruhi, untuk lebih jelasnya dapat dilihat pada Tabel 10.

Tabel 10. Anggota *Cluster* C1 dengan Kategori Berpengaruh

No	Nama	DC0	DC0	Cluster
1.	Olivia Levan's	28,559	9,488	C1
2.	Afifah Athohirah	11,605	8,062	C1
3.	Jerri Dwi Alkarizmi	26,898	7,674	C1
4.	Fikri Estiawan	14,508	5,378	C1
5.	Zahra Zakiah Raihanisa	11,519	8,086	C1
6.	Haura Fadhlila Dzamri	11,501	8,147	C1
7.	Rangga	11,913	8,252	C1
8.	Muhammad Farhab Shiddiq	26,500	7,096	C1
9.	Wilza Lolisa Shandori	11,539	8,068	C1
10.	Nahda Navilah Daulai	31,557	12,369	C1
11.	Mufidah Asri Azizah	15,500	4,309	C1
12.	Wardah Hafifah	26,625	7,052	C1
13.	Yudi Alfikri	27,027	7,998	C1
14.	Najla Amatullah	26,514	6,957	C1
15.	Fadhlan Habib	11,817	8,177	C1
16.	M. Alfattah	26,615	7,033	C1
17.	Fadila Khairunnisa	14,666	5,132	C1

C1 terdiri dari 17 item sebagai *Cluster* dengan kategori berpengaruh yang nantinya pada *Cluster* "C1" siswa-siswa yang nilai tahfizny kurang atau belum mencapai target agar lebih ditingkatkan lagi jumlah hafalannya sehingga siswa tersebut mendapatkan nilai yang mencapai target yang diinginkan. Untuk lebih jelasnya dapat dilihat pada tabel 4.10.

4. Kesimpulan

Dalam penelitian ini dapat diambil kesimpulan bahwa hasil perhitungan dengan metode *K-Means Clustering* untuk melihat hubungan kegiatan tahfiz mendapatkan hasil 28 siswa yang tidak berpengaruh, dan 17 siswa yang berpengaruh terhadap hasil belajar siswa. Dengan perhitungan total semua pengujian ada 45 siswa. Berdasarkan dari perbandingan dengan data Real kegiatan tahfiz dengan hasil belajar terdapat 2 siswa yang tingkat keakuratan dengan perhitungan *K-Means Clustering*, sehingga mendapatkan hasil perbandingan 95,56%.

Daftar Rujukan

- [1] Triyansyah, D., & Fitriana, D. (2018). Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing. *Jurnal Telekomunikasi dan Komputer*, 8(3), 162-182. DOI: <http://dx.doi.org/10.22441/incomtech.v8i3.4174>.
- [2] Irfiani, E., & Rani, S. S. (2018). Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 6(4), 165-172. DOI: <http://dx.doi.org/10.26418/justin.v6i4.29024>.
- [3] Sulastri, H., & Gufroni, A. I. (2017). Penerapan Data Mining dalam Pengelompokan Penderita Thalassaemia. *Jurnal Nasional Teknologi & Sistem Informasi*, 3(2), 299-305. DOI: <https://doi.org/10.25077/TEKNOSI.v3i2.2017.299-305>.
- [4] Rosmini, R., Fadlil, A., & Sunardi, S. (2018). Implementasi Metode K-Means dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah. *IT Journal Research and Development*, 3(1), 22-31. DOI: [https://doi.org/10.25299/itjrd.2018.vol3\(1\).1773](https://doi.org/10.25299/itjrd.2018.vol3(1).1773).
- [5] Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknika* 21(1), 60-64.