
Implementasi Algoritma K-Means Clustering Dalam Pengelompokan Uang Kuliah Tunggal Mahasiswa pada STAIN Mandailing Natal

Fitri Ida Sari *Muhammad Syahril **, Ahmad Calam**

* Program Studi Sistem Informasi, STMIK Triguna Dharma

** Program Studi Sistem Informasi Dosen Pembimbing, STMIK Triguna Dharma

Article Info

Article history:

-

Keyword:

UKT

, Data mining, KMEANS

ABSTRACT

Uang Kuliah Tunggal (UKT) merupakan sistem pembayaran uang kuliah yang diterapkan oleh pemerintah pada perguruan tinggi negeri. Uang kuliah tunggal berkaitan dengan biaya pendidikan yang harus dibayar oleh setiap mahasiswa. Setiap tahunnya nominal uang pangkal mengalami kenaikan, sehingga terasa sangat memberatkan mahasiswa dalam proses pembayaran saat menjadi mahasiswa baru. Proses pembayaran uang kuliah didasarkan indikator sosial ekonomi. Data mahasiswa harus sesuai besaran uang kuliah karena akan menjadi tumpang tindih bagi mahasiswa, dan pengelompokan data yang tidak seimbang akan berpengaruh pada besaran yang diterima mahasiswa jika tidak tepat sasaran. Berdasarkan masalah diatas maka diperlukan suatu aplikasi untuk mengelompokan uang kuliah tunggal dengan menggunakan metode Algoritma K-means Clustering. Yang nantinya dapat digunakan pihak kampus STAIN Mandailing Natal dalam mengelompokan uang kuliah tunggal. Hasil dari klasterisasi dari pengelompokan data uang kuliah tunggal diharapkan dapat mengontrol uang kuliah tunggal setiap mahasiswa pada kampus STAIN Mandailing Natal dengan mudah..

First Author : Fitri Ida Sari

Kampus :STMIK Triguna Dharma

Program Studi : Sistem Informasi

E-Mail : fitriidanasty@gmail.com

1. PENDAHULUAN

Kuliah merupakan proses pembelajaran tingkat lanjut untuk membentuk karakter dan pengembangan diri sehingga dapat memperluas wawasan pengetahuan. Dalam dunia perkuliahan terdapat banyak kampus favorit yang tersebar diberbagai daerah, terhusus di kabupaten Mandailing Natal terdapat sebuah Sekolah Tinggi Agama Islam Negeri Mandailing Natal (STAIN MADINA) yang merupakan satu-satunya perguruan tinggi negeri yang berada di Kabupaten Mandailing Natal Sumatera Utara sehingga menjadi pilihan utama untuk melanjutkan kejenjang yang lebih tinggi. Dalam dunia perkuliahan tidak luput dari biaya kuliah, di Sekolah Tinggi Agama Islam Negeri Mandailing Natal(STAIN MADINA) memiliki program yang bernama Uang Kuliah Tunggal(UKT).

Uang kuliah Tunggal (UKT) merupakan sistem pembayaran uang kuliah yang diterapkan oleh pemerintah pada perguruan tinggi negeri. Uang kuliah tunggal berkaitan dengan biaya pendidikan yang harus dibayar oleh setiap mahasiswa. Setiap tahunnya nominal uang pangkal mengalami kenaikan, sehingga terasa sangat memberatkan mahasiswa dalam proses pembayaran saat menjadi mahasiswa baru. Biaya kuliah Perguruan Tinggi Keagamaan Islam Negeri diatur melalui Keputusan Menteri Agama Republik Indonesia nomor 151 tahun 2019 tentang uang kuliah tunggal tahun akademik 2019 – 2020. [1].

Dengan adanya pengelompokan Uang Kuliah Tunggal (UKT) maka akan mempermudah mahasiswa dalam proses pembayaran uang kuliah didasarkan indikator sosial ekonomi. Data mahasiswa harus

sesuai besaran uang kuliah akan menjadi tumpang tindih bagi mahasiswa, dan pengelompokan data yang tidak seimbang akan berpengaruh pada besaran yang diterima mahasiswa jika tidak tepat sasaran. Dalam sebuah penelitian sebelumnya mengungkapkan bahwa “data untuk dikelompokkan berdasarkan pola kemiripan atribut-atributnya. Pada proses ini akan menghasilkan 5 (lima) pengelompokan besaran UKT mahasiswa berdasarkan kondisi ekonomi orang tua mahasiswa. Kelompok inilah nantinya akan menjadi dasar penentuan kategori besaran UKT calon mahasiswa baru”[2].

Agar proses pengelompokan Uang Kuliah Tunggal (UKT) mendapatkan hasil yang baik, maka digunakan bidang ilmu Data Mining dengan algoritma K-means Clustering. Didalam Data Mining memiliki konsep dasar yakni untuk menemukan pola pola yang tersembunyi maupun hubungan- hubungan yang terdapat dalam data base yang besar dan menghasilkan aturan-aturan yang digunakan[3].

Didalam bidang ilmu data mining terdapat beberapa algoritma di dalam nya, dalam konteks ini algoritma yang sesuai dengan masalah pengelompokan UKT(uang kuliah tunggal) adalah K-means Clustering.K-Means merupakan salah satu metode data clustering non hierarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyaikarakteristik yang berbeda dikelompokkan ke dalam kelompok yang lainnya[4]

2. TINJAUAN PUSTAKA

2.1 Data mining

Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database (KDD). Dengan data mining, kita dapat melakukan pengklasifikasian, memprediksi, memperkirakan dan mendapatkan informasi lain yang bermanfaat dari kumpulan data dalam jumlah yang besar[4].

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi didalam database. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika kecerdasan buatan, dan machine learning untuk mengekstrasi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan didalam database yang ukurannya besar[5].

2.2 Algoritma K-Means

Algoritma K-Means merupakan metode analisis kelompok yang mengarah pada pemartisian objek pengamatan ke dalam kelompok (cluster) dimana setiap objek pengamatan dimiliki oleh sebuah

kelompok dengan mean (rata-rata) terdekat, mirip dengan algoritma Expectation-Maximization untuk Gaussian Mixture dimana keduanya mencoba untuk menemukan pusat dari kelompok dalam data sebanyak iterasi perbaikan yang dilakukan oleh kedua algoritma.

Berikut ini adalah langkah-langkah algoritma K-Means[10]:

1. Tentukan K sebagai jumlah cluster yang dibentuk.
2. Tentukan pusat (centroid) cluster awal. digunakan rumus sebagai berikut.

$$C_i = \frac{1}{m} \sum_j^m = 1x_j$$

3. METODOLOGI PENELITIAN

Metode penelitian merupakan cara atau langkah yang harus dilakukan untuk mengumpulkan suatu informasi yang berisikan data yang kita peroleh dari seorang pakar atau ahli dalam bidangnya

1. Data Collecting

Dalam teknik pengumpulan data terdapat beberapa yang dilakukan di antaranya yaitu sebagai berikut:

a.Observasi

Dalam observasi penelitian melakukan pra-riiset terlebih dahulu untuk mencari mencari masalah yang terjadi di STAIN Mandailing Natal dalam melakukan pengelompokan uang kuliah tunggal. Dari masalah tersebut akan dirumuskan

dalam penelitian ini sehingga menemukan apa saja yang perlu dipersiapkan untuk bagaimana menyelesaikan masalah tersebut.

b.Wawancara

Dalam mendapatkan data yang baik, dalam hal ini penelitian melakukan wawancara kepada stakeholderatau pihak-pihak yang terlibat dalam mendukungnya penelitian ini. Dalam hal ini peneliti melakukan wawancara kepada pihak yang terkait tentang uang kuliah tunggal mahasiswa pada STAIN Mandailing Natal yang telah ditunjuk (data terlampir). Selain itu juga, peneliti mencoba mencari data sekunder dengan melakukan surfing di mesin pencarian terkait hal-hal penting dalam pengelompokan uang kuliah tunggal pada STAIN Mandailing Natal.

1. Studi Literatur

Di dalam studi literatur, penelitian ini banyak menggunakan jurnal-jurnal baik jurnal internasional, jurnal nasional, jurnal lokal maupun buku sebagai sumber referensi. Dari komposisi yang ada jumlah literatur yang digunakan sebanyak 21 jurnal. Diharapkan dengan literatur tersebut dapat membantu peneliti dalam menyelesaikan yang terjadi di STAIN Mandailing Natal terkait pengelompokan

pengelompokan uang kuliah tunggal di STAIN Mandailing Natal.

3.1 Metode perancangan Sistem

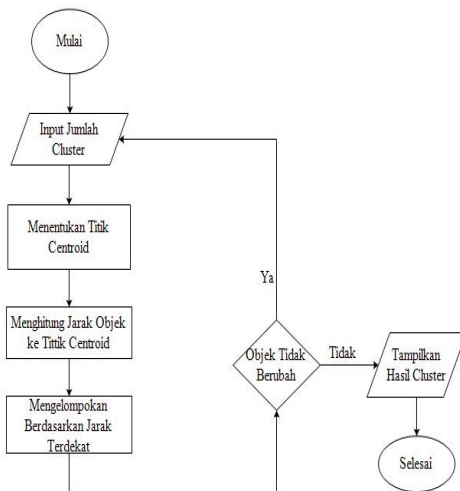
Metode perancangan sistem merupakan salah satu unsur penting dalam penelitian. Dalam metode perancangan sistem, khususnya software atau perangkat lunak, dapat diadopsi beberapa metode, yang diantaranya adalah algoritma water fall algorithm.

3.2 Algoritma Sistem

Algoritma sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan sistem penerapan data mining dalam pengelompokan uang kuliah tunggal menggunakan algoritma K-Means. Hal ini dibutuhkan untuk mempermudah mahasiswa dalam proses pembayaran uang kuliah didasarkan indikator sosial ekonomi, sehingga memenuhi point– point pada bab 1. Hasil dari proses perhitungan Algoritma K-Means nantinya akan dibentuk dalam Kategori keadaan Tinggi (C1), Sedang (C2) dan Rendah (C3).

3.2.1 Flowchart Sistem

Berikut ini adalah flowchart dari algoritma K-Means yaitu sebagai berikut.



Gambar 3.2 Flowchart metode KMEANS

3.3 Proses Perhitungan K-Means

Tahap ini dilakukan penerapan algoritma K-Means dengan rumus :

$$D(x,y) = \|x-y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1,2,3,\dots,n$$

Penerapan jumlah *cluster* (*K*) yaitu 3 cluster, Setelah menetapkan jumlah *cluster*, Tentukan titik pusat awal *cluster* (*Centroid*).

Untuk dapat melakukan pengelompokan data-data tersebut menjadi beberapa *cluster* perlu dilakukan beberapa langkah, yaitu :

1. Tentukan jumlah *cluster* yang diinginkan. Dalam penelitian ini data yang ada akan dikelompokkan menjadi tiga *cluster*.
2. Tentukan titik pusat awal dari setiap *cluster*. Dalam penelitian ini titik pusat awal ditentukan secara *random* dan didapat titik pusat dari setiap *cluster* dapat dilihat pada tabel 3.6.

Tabel 3. 1 Titik Pusat (*Centroid*) Awal Setiap *Cluster*

<i>Centroid</i>	Nama Mahasiswa	K1	K2
M1	S9	1	1
M2	S15	4	3
M3	S29	5	4

3. Menghitung *Distance* (jarak) antara variable dari stiap sample data dengan *centroid* yaitu:

- a. Hitunglah setiap object dengan M1 (3, 2)

- Jarak antara data ke-1 dengan titik pusat centroid M1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2 - 3)^2 + (2 - 2)^2}$$

$$= 1$$

- Jarak antara data ke-2 dengan titik pusat centroid M1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(2 - 3)^2 + (2 - 2)^2}$$

$$= 1$$

- Jarak antara data ke-3 dengan titik pusat centroid M1

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$= \sqrt{(3 - 3)^2 + (3 - 2)^2}$$

$$= 1$$

Lakukan proses yang sama sampai dengan data ke-35 dan M1.

- b. Hitunglah setiap object dengan M2 (1, 2)

- Jarak antara data ke-1 dengan titik pusat centroid M2

$$= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

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

$$= 1$$

- Jarak antara data ke-2 dengan titik pusat centroid M2

$$= \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

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

- Jarak antara data ke-3 dengan titik pusat centroid M2

$$= \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$= \sqrt{(3 - 1)^2 + (3 - 2)^2} = 2$$

Lakukan proses yang sama sampai dengan data ke-35 dan M2.

c. Hitunglah setiap object dengan M (1, 2)

- Jarak antara data ke-1 dengan titik pusat centroid M1

$$= \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$= \sqrt{(2 - 3)^2 + (2 - 2)^2}$$

$$= 1$$

- Jarak antara data ke-2 dengan titik pusat centroid M1

$$= \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$= \sqrt{(2 - 3)^2 + (2 - 2)^2}$$

$$= 1$$

- Jarak antara data ke-3 dengan titik pusat centroid M1

$$= \sqrt{\sum_{i=1}^n (xi - yi)^2}$$

$$= \sqrt{(3 - 3)^2 + (3 - 2)^2}$$

$$= 1$$

Lakukan proses perhitungan yang sama dengan objek ke 35, Adapun hasil dari perhitungan iterasi 1 dapat dilihat pada tabel dibawah ini. Dimana jarak terdekat dilihat dari perhitungan yang paling dekat ke pusat cluster. Sementara WCV (*Within Cluster Variatio*) adalah keadaan dari hasil pangkat dari perhitungan jarak terdekat ke pusat cluster.

11	S11	0	3.61	5	0	C1
12	S12	3.61	0	1.41	0	C2
13	S13	0	3.61	5	0	C1
14	S14	1	3.16	4.47	1	C1
15	S15	3.61	0	1.41	0	C2
16	S16	2.24	1.41	2.83	1.41	C2
17	S17	1	3.16	4.47	1	C1
18	S18	0	3.61	5	0	C1
19	S19	0	3.61	5	0	C1
20	S20	0	3.61	5	0	C1
21	S21	1	2.83	4.24	1	C1
22	S22	1	3.16	4.47	1	C1
23	S23	0	3.61	5	0	C1
24	S24	1.41	2.24	3.61	1.41	C1
25	S25	0	3.61	5	0	C1
26	S26	1	3.16	4.47	1	C1
27	S27	2.83	1	2.24	1	C2
28	S28	1	3.16	4.47	1	C1
29	S29	5	1.41	0	0	C3
30	S30	1	3.16	4.47	1	C1
31	S31	2.24	1.41	2.83	1.41	C2
32	S32	2.24	1.41	2.83	1.41	C2
33	S33	0	3.61	5	0	C1
34	S34	0	3.61	5	0	C1
35	S35	2.83	1	2.24	1	C2

Tabel 3. 2 Hasil Perhitungan Iterasi Ke-1

No	Nama Mahasiswa	M1	M2	M3	Jarak Terdekat	Keterangan
1	S1	1.41	2.24	3.61	1.41	C1
2	S2	1.41	2.24	3.61	1.41	C1
3	S3	2.83	1	2.24	1	Maka akan
4	S4	0	3.61	5	0	Cluster
5	S5	1	2.24	3.61	1	Centroid
6	S6	5	1.41	0	0	centroid
7	S7	1	2.83	4.24	1	M1 = (1, 1),
8	S8	5.66	2.24	1	1	centroid M2 = (4, 3) dan
9	S9	0	3.61	5	0	centroid M3 = (5, 4).
10	S10	2.24	1.41	2.83	1.41	

Keterangan :

- C1 memiliki 23 Anggota, yaitu data {S1, S2, S4, S5, S7, S9, S11, S13, S14, S17, S18, S19, S20, S21, S22, S23, S24, S25, S26, S28, S30, S33, S34}.
- C2 memiliki 9 Anggota, yaitu data {S3, S10, S12, S15, S16, S27, S31, S32, S35}.
- C3 memiliki 3 anggota yaitu data {S8, S29}.

Maka akan dihitung rasio besaran BCV (*Between Cluster Variation*) dengan WCV (*Within Cluster Variation*) :

Karena centroid M1 = (1, 1), centroid M2 = (4, 3) dan centroid M3 = (5, 4).

$$d(m1,m2) = \sqrt{(1 - 4)^2 + (1 - 3)^2} = 3,60$$

$$d(m1,m3) = \sqrt{(1 - 5)^2 + (1 - 4)^2} = 5$$

$$d(m2,m3) = \sqrt{(4 - 5)^2 + (3 - 4)^2} = 1,41$$

Nilai BCV = d(m1,m2) + d(m1,m3) + d(m2,m3)

$$= 3.60 + 5 + 1.41$$

=10.01

WCV adalah memilih jumlah jarak terkecil dipangkat dua antara data dengan *centroid* pada masing-masing *cluster*.

$WCV = 1.41^2 + 1.41^2 + 1^2 + 0^2 + 1^2 + \dots + 1^2$
 =26.91

Sehingga besar rasio = $BCV/WCV = 10.01/26.91 = 0.37$

Karena ini merupakan langkah iterasi 1 maka lanjutkan ke langkah berikutnya.

Setelah proses perhitungan selesai, maka akan terbentuk tabel *cluster* baru dengan tiap-tiap anggota *cluster* seperti tabel dibawah ini :

Tabel 3. 3 Anggota Cluster 1

No	Nama Mahasiswa	K1	
1	S1	2	
2	S2	2	
3	S4	1	
4	S5	2	
5	S7	2	
6	S9	1	
7	S11	1	
8	S13	1	
9	S14	1	

Tabel 3.8 Anggota Cluster 1 (Lanjutan)

No	Nama Mahasiswa	K1	
10	S17	1	
11	S18	1	
12	S19	1	
13	S20	1	
14	S21	2	
15	S22	1	
16	S23	1	
17	S24	2	
18	S25	1	
19	S26	1	
20	S28	1	
21	S30	1	
22	S33	1	
23	S34	1	
Total		29	
Average		1.26	

Tabel 3. 4 Anggota Cluster 2

No	Nama Mahasiswa	K1	
1	S3	3	
2	S10	3	
3	S12	4	
4	S15	4	
5	S16	3	

Tabel 3.9 Anggota Cluster 2 (Lanjutan)

No	Nama Mahasiswa	K1	
6	S27	3	
7	S31	3	
8	S32	3	
9	S35	3	
Total		29	
Average		3.22	

Tabel 3. 5 Anggota Cluster 3

No	Nama Mahasiswa	K1	
1	S6	5	
11	S8	5	
2	S29	5	
3			
Total		15	
Average		5	

Nilai pada baris *Average* merupakan hasil perhitungan rata-rata dari tiap-tiap kolom titik *cluster*. Nilai *average* ini nantinya menjadi titik pusat untuk *cluster* 2 yang baru.

1. Menghitung kembali Pusat-pusat baru dari kelompok (*cluster*) yang baru terbentuk.

Tabel 3. 6 Titik Pusat (*Centroid*) Baru Setiap *Cluster* Pada Iterasi 2

	K1	K2
Cluster 1	1.26	1.43
Cluster 2	3.22	2.56
Cluster 3	5	4.33

2. Menghitung kembali jarak antara *variable* dari setiap sampel data dengan centroidnya.

a. Hitunglah setiap object dengan M1 (1.26, 1.43)

- Jarak antara data ke-1 dengan titik pusat centroid M1

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(2 - 1.26)^2 + (2 - 1.43)^2} \\
 &= 0.93
 \end{aligned}$$

- Jarak antara data ke-3 dengan titik pusat centroid M1

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(2 - 1.26)^2 + (2 - 1.43)^2} \\
 &= 0.93
 \end{aligned}$$

- Jarak antara data ke-3 dengan titik pusat centroid M1

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(3 - 1.26)^2 + (3 - 1.43)^2} \\
 &= 2.34
 \end{aligned}$$

Lakukan proses yang sama sampai dengan data ke-35 dan M1.

b. Hitunglah setiap object dengan M2 (3.22, 2.56)

- Jarak antara data ke-1 dengan titik pusat centroid M2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(2 - 3.22)^2 + (2 - 2.56)^2} \\
 &= 1.34
 \end{aligned}$$

- Jarak antara data ke-1 dengan titik pusat centroid M2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(2 - 3.22)^2 + (2 - 2.56)^2} \\
 &= 1.34
 \end{aligned}$$

- Jarak antara data ke-1 dengan titik pusat centroid M2

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(3 - 3.22)^2 + (3 - 2.56)^2} \\
 &= 0.49
 \end{aligned}$$

Lakukan proses yang sama sampai dengan data ke-35 dan M2.

c. Hitunglah setiap object dengan M3 (5, 4.33)

- Jarak antara data ke-1 dengan titik pusat centroid M3

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(2 - 5)^2 + (2 - 4.33)^2} \\
 &= 3.80
 \end{aligned}$$

- Jarak antara data ke-1 dengan titik pusat centroid M3

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(2 - 5)^2 + (2 - 4.33)^2} \\
 &= 3.80
 \end{aligned}$$

- Jarak antara data ke-1 dengan titik pusat centroid M3

$$\begin{aligned}
 &= \sqrt{\sum_{i=1}^n (xi - yi)^2} \\
 &= \sqrt{(3 - 5)^2 + (3 - 4.33)^2} \\
 &= 2.40
 \end{aligned}$$

Lakukan proses yang sama sampai dengan data ke-35 dan M3.

adapun hasil dari perhitungan iterasi ke2 dapat dilihat pada table dibawah ini:

Tabel 3. 7 Hasil Perhitungan Iterasi Ke-2

No	Nama Mahasiswa	M1	M2	M3	Jarak
1	S1	0.93	1.34	3.80	
2	S2	0.93	1.34	3.80	
3	S3	2.34	0.49	2.40	
4	S4	0.50	2.71	5.20	
5	S5	0.93	1.34	3.80	
6	S6	4.54	2.29	0.33	
7	S7	0.86	1.98	4.48	
8	S8	5.17	3.02	0.67	
9	S9	0.50	2.71	5.20	
10	S10	1.83	0.60	3.07	
11	S11	0.50	2.71	5.20	
12	S12	4.06	0.90	1.66	
13	S13	0.50	2.71	5.20	
14	S14	0.63	2.29	4.63	
15	S15	3.16	0.90	1.66	
16	S16	1.83	0.60	3.07	
17	S17	0.63	2.29	4.63	
18	S18	0.50	2.71	5.20	
19	S19	0.50	2.71	5.20	
20	S20	0.50	2.71	5.20	
21	S21	1.24	1.98	4.48	

Tabel 3.12 Hasil Perhitungan Iterasi Ke-2 (Lanjutan)

No	Nama Mahasiswa	M1	M2	M3	Berdasarkan hasil metode <i>k-means clustering</i> jarak terdekat	Keterangan
22	S22	0.63	2.29	4.63	0.63	iterasi ke-2, dimana titik pusatnya tidak lagi berubah atau data tidak lagi berpindah antar cluster
23	S23	0.50	2.71	5.20	0.50	C1
24	S24	0.93	1.34	3.80	0.93	C1
25	S25	0.50	2.71	5.20	0.50	C1
26	S26	0.63	2.29	4.63	0.63	C1
27	S27	2.34	0.49	2.40	0.49	C2
28	S28	0.63	2.29	4.63	0.63	C1
29	S29	4.54	2.29	0.33	0.33	C3
30	S30	0.63	2.29	4.63	0.63	C1
31	S31	1.83	0.60	3.07	0.60	C2
32	S32	1.83	0.60	3.07	0.60	C2

3.3.1 Interpretation atau Evaluation

Pada tahap ini C1 dapat diketahui hasil klusterisasi menggunakan algoritma *K-Means Clustering* untuk pengelompokan uang kuliah tunggal. Dari hasil dibentuk kategori Tinggi, Sedang dan Rendah untuk pengelompokan uang kuliah tunggal.

Tabel 3.8 Hasil Kategori Rendah

No	Nama Mahasiswa	Inisialisasi	Cluster	Keterangan
1	Ahmad Rizal	S1	1	Rendah
2	Ahmad Zaidi	S2	1	Rendah
3	Bela Nirwana Lubis	S4	1	Rendah
4	Devina Sari	S5	1	Rendah
5	Fery Laso	S7	1	Rendah
6	Latifah	S9	1	Rendah
7	Mey Waryuni Sumanja	S11	1	Rendah
8	Miskah Fitriani	S13	1	Rendah
9	Mufidah	S14	1	Rendah
10	Muhammad Sobwan	S17	1	Rendah
11	Musa Azhari	S18	1	Rendah
12	Musthafa Ardiyansyah Matondang	S19	1	Rendah
13	Nelly Aida	S20	1	Rendah
14	Nur Hapipah	S21	1	Rendah
15	Nur Hasanah	S22	1	Rendah
16	Nur Hidayah Lubis	S23	1	Rendah

Keterangan :

- C1 memiliki 23 Anggota, yaitu data {S1, S2, S4, S5, S7, S9, S11, S13, S14, S17, S18, S19, S20, S21, S22, S23, S24, S25, S26, S28, S30, S33, S34}.
- C2 memiliki 9 Anggota, yaitu data {S3, S10, S12, S15, S16, S27, S31, S32, S35}.
- C3 memiliki 3 anggota yaitu data {S6, S8, S29}.

Maka akan dihitung rasio besaran BCV (*Between Cluster Variation*) dengan WCV (*Within Cluster Variation*) :

Karena centroid M1 = (1.26, 1.43), centroid M2 = (3.22, 2.56) dan centroid M3 = (5, 4.33).

$$D(m1,m2) = \sqrt{(1.26 - 3.22)^2 + (1.43 - 2.56)^2} = 2.262$$

$$D(m1,m3) = \sqrt{(1.26 - 5)^2 + (1.43 - 4.33)^2} = 3.732$$

$$D(m2,m3) = \sqrt{(3.22 - 5)^2 + (2.56 - 4.33)^2} = 2.510$$

$$\text{Nilai BCV} = d(m1,m2) + d(m1,m3) + d(m2,m3) = 2.262 + 3.732 + 2.510 = 8.504$$

WCV adalah memilih jumlah jarak terkecil dipangkat dua antara data dengan centroid pada masing-masing cluster.

$$\text{WCV} = 0.93^2 + 0.93^2 + 0.49^2 + 0.50^2 + 0.93^2 + \dots + 0.49^2 = 15.31$$

Sehingga besar rasio = BCV/WCV = 8.504/15.31 = 0.55

17	Nur Hijjah	S24	1	Rendah
18	Purnama Maulida	S25	1	Rendah
19	Rahmi Anggina	S26	1	Rendah
20	Riza Erpina	S28	1	Rendah
21	Rosna Sari	S30	1	Rendah
22	Suci Rahmina	S33	1	Rendah
23	Vikram Ansori	S34	1	Rendah

Tabel 3. 9 Hasil Kategori Sedang

No	Nama Mahasiswa	Inisialisa si	Clust er	Keterang an
1	Ana Angriana Sari	S3	2	Sedang
2	Madina Putri Nst	S10	2	Sedang
3	Milfah Fitriah Nasution	S12	2	Sedang

Tabel 3.14 Hasil Kategori Sedang (Lanjutan)

No	Nama Mahasiswa	Inisialisa si	Clust er	Keterang an
4	Muhamm ad Ruly Ripandi Hsb	S15	2	Sedang
5	Muhamm ad Saleh Nst	S16	2	Sedang
6	Risqa	S27	2	Sedang
7	Sangkot Layla Fitri	S31	2	Sedang
8	Soripada Mulia	S32	2	Sedang
9	Yuli Anti	S35	2	Sedang

Tabel 3. 10 Hasil Kategori Tinggi

No	Nama Mahasiswa	Inisialisa si	Clust er	Keterang an
----	----------------	---------------	----------	-------------

1	Eli Wahyuni	S6	3	Tinggi
2	Khairul Anwar Tanjung	S8	3	Tinggi
3	Rizki Ananda	S29	3	Tinggi

Berdasarkan data pengelompokan uang kuliah tunggal maka dapat disimpulkan sebagai berikut :

1. *Cluster* 1 adalah mahasiswa dengan kategori uang kuliah rendah dengan tarif Rp.400.000 – Rp. 800.000 dengan jumlah anggota 23 mahasiswa.
2. *Cluster* 2 adalah mahasiswa dengan kategori uang kuliah sedang dengan tarif Rp.1.000.000 – Rp. 1.500.000 dengan jumlah anggota 9 mahasiswa.
3. mahasiswa dengan kategori uang Kuliah tinggi dengan tarif Rp. 2.400.000 dengan jumlah anggota 3 mahasiswa.

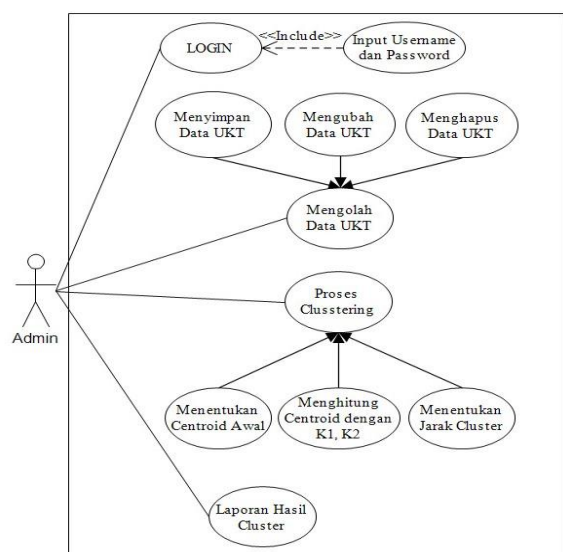
Dari pembahasan diatas pengelompokan uang kuliah tunggal dapat dikategorikan menjadi 3 kelompok yaitu kelompok rendah, sedang, dan tinggi yang berjumlah 35 mahasiswa, sehingga pihak STAIN Mandailing Natal dapat mengelompokan mahasiswa dengan kategori masing-masing. Dan pihak kampus dapat menambahkan data variabel untuk mendapatkan data mahasiswa yang lebih banyak.

4. PEMODELAN

4.1 Pemodelan Sistem

4.1.1 Use case diagram

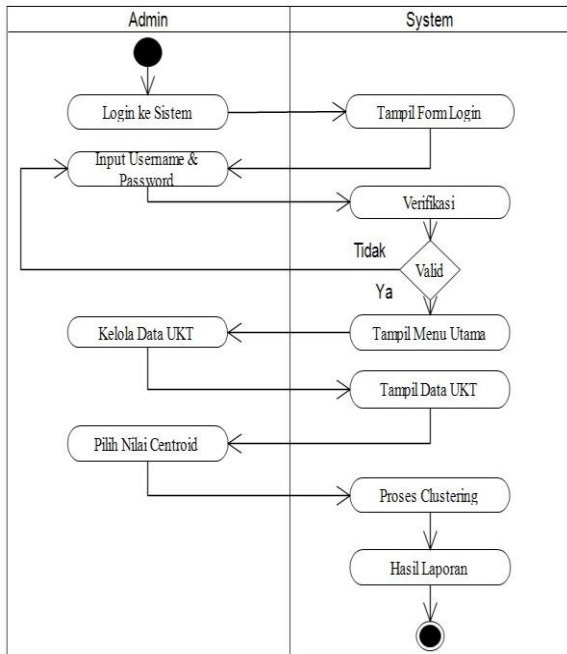
Berdasarkan skenario di atas berikut ini adalah gambar dari Use Case Diagramnya yaitu sebagai berikut:



Gambar 4.1 Use Case Diagram Sistem

4.1.2 Activity diagram

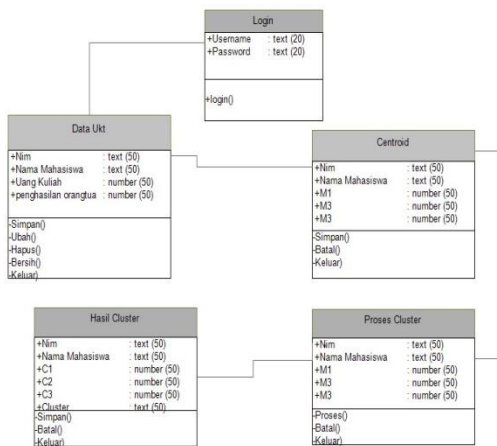
Berdasarkan skenario di atas berikut ini adalah gambar dari Activity Diagramnya yaitu sebagai berikut:



Gambar 4.2 Activity Diagram Sistem

4.1.3 Class Diagram

Class Diagram merupakan suatu diagram yang dapat menggambarkan seluruh hubungan dari setiap class pada suatu sistem:



Gambar 4.3 Class Diagram Sistem

5. HASIL DAN PEMBAHASAN

5.1 Kebutuhan Sistem

Dalam membangun sistem yang telah dirancang didalam data mining ini membutuhkan 2 buah

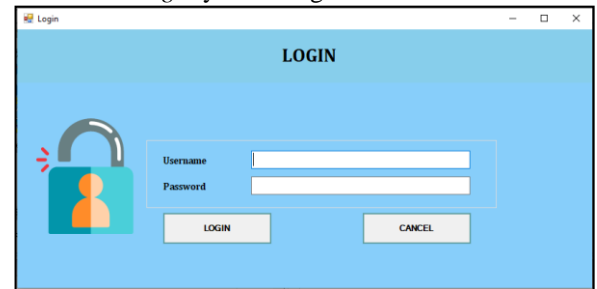
perangkat yaitu, perangkat lunak dan perangkat keras. Adapun perangkat keras dan perangkat lunak yang dibutuhkan adalah sebagai berikut :

1. Perangkat Lunak
 - a. Sistem Operasi (OS) Minimum Windows 7
 - b. Microsoft Visual Basic
 - c. Microsoft Access
 - d. Crystal Report
2. Perangkat Keras
 - a. Komputer dengan Processor minimal Dual Core
 - b. Random Access Memory (RAM) minimal 4 GB
 - c. Hard Disk Minimal 500 GB
 - d. Mouse, Keyboard dan Monitor

5.2 Hasil Tampilan Antarmuka

Tampilan antar muka adalah gambar hasil tampilan halaman dari aplikasi sistem yang telah dibuat yang disesuaikan dengan perancangan interface pada Perancangan sebelumnya. Adapun gambar tampilan antarmukanya adalah sebagai berikut

1. Form Login
Form Login merupakan halaman untuk menginput *username* dan *password* dari aplikasi data mining ini. Berikut ini adalah tampilan dari Form Login yaitu sebagai berikut :



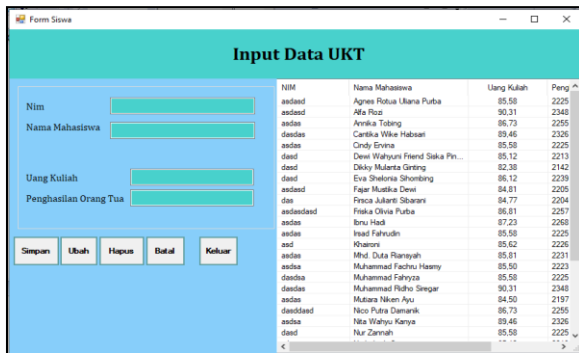
Gambar 5.1 Tampilan Form Login

2. Form Menu Utama
Form Menu Utama adalah halaman utama dari data mining ini. Berikut ini adalah tampilan antarmuka dari Form Menu Utama dari aplikasi data mining ini :



Gambar 5.2 Tampilan Form Menu Utama

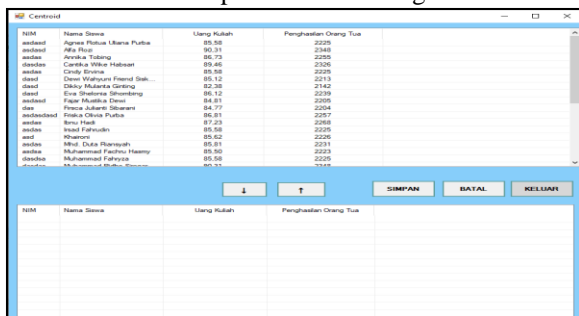
3. Form Data UKT
Berikut ini adalah tampilan antarmuka dari Form Data UKT dari aplikasi data mining ini :



Gambar 5.3 Tampilan Form Data UKT

4. Form Data Centroid

Berikut ini adalah tampilan antarmuka dari Form Data Centroid dari aplikasi data mining ini :



Gambar 5.4 Tampilan Form Data centroid

5. Form data proses KMEANS

Berikut ini adalah tampilan antarmuka dari Form proses KMEANS dari aplikasi data mining ini :



Gambar 5.5 Tampilan Form proses KMEANS

6. Laporan

Berikut ini adalah tampilan antarmuka Laporan dari aplikasi data mining ini :



Gambar 5.6 Tampilan Laporan

5.3 Kelebihan dan Kekurangan Sistem

Setelah melakukan proses implementasi dan pengujian terhadap sistemnya, terdapat beberapa kelebihan dan kekurangan dari sistem yang dirancang, berikut ini adalah kelebihan dan kekurangannya yaitu sebagai berikut :

1.Kelebihan Sistem

Adapun kelebihan dari data mining ini yaitu sebagai berikut :

- a.Sistem ini dapat membantu pihak STAIN Mandailing Natal dalam pengelompokan uang kuliah.
- b.Dengan Sistem ini dapat lebih memudahkan dalam memproses data alternatif yang banyak dan lebih menghemat waktu dibandingkan memproses data secara manual.
- c.Sistem ini memiliki user friendly.

2.Kekurangan Sistem

Adapun kekurangan dari sistem ini adalah

- a.Data mining yang dirancang terbatas dalam hal penyelesaian masalah terkait pengelompokan uang kuliah.
- b.Aplikasi ini belum dilengkapi dengan keamanan data yang baik, aman dan akurat karena tidak menggunakan algoritma pengamanan data.
- c.Sistem ini hanya tersedia untuk tampilan dekstop dan tidak bisa diakses dari mana saja..

6. KESIMPULAN DAN SARAN

6.1 Kesimpulan

Setelah dilakukan penelitian, Dan berdasarkan rumusan masalah yang telah dijelaskan pada Bab I sebelumnya maka kesimpulan dari penelitian ini yaitu sebagai berikut:

- 1.Algoritma K-means telah dapat dipakai dalam mengelompokan data uang kuliah tunggal dan data yang digunakan dalam sistem ini adalah data yang didapat dari hasil observasi pada Kampus STAIN Mandailing Natal.

- 2.Sistem dan aplikasi yang dibangun untuk mengelompokan uang kuliah tunggal menggunakan rancangan UML yang didalamnya terdapat rancangan Use Case Diagram, Activity, Class Diagram.

3. Aplikasi dibangun memiliki keluaran laporan terkait pengelompokan uang kuliahdan sudah layak digunakan pada kantor kampus yang bersangkutan

6.2 Saran

Adapun saran dari penelitian ini yaitu:

1. Perangkat lunak dapat dikembangkan lebih lanjut dengan menambahkan jumlah atribut yang digunakan dalam proses perhitungan

2. Diharapkan peneliti berikutnya dapat menggunakan metode lain sebagai studi banding dan pengembangan khasanah keilmuan.


3. Diharapkan peneliti berikutnya juga dapat membangun aplikasi lain seperti aplikasi berbasis web dan aplikasi berbasis mobile baik Android maupun IOS.

REFERENSI

- [1] “Undang-undang No. 55 Tahun 2013,” *Peratur. Menteri Pendidik. dan Kebud. No. 55*, 2013.
- [2] Haris Kurniawan, Sarjon Defit, and Sumijan, “Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal,” *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 80–89, 2020, doi: 10.52158/jacost.v1i2.102.
- [3] L. Suriani, “Pengelompokan Data Kriminal Pada Poldasu Menentukan Pola Daerah Rawan Tindak Kriminal Menggunakan Data Mining Algoritma K-Means Clustering,” *J. Sist. Komput. dan Inform.*, vol. 1, no. 2, p. 151, 2020, doi: 10.30865/json.v1i2.1955.
- [4] B. M. Metisen and H. L. Sari, “Analisis clustering menggunakan metode K-Means dalam pengelompokan penjualan produk pada Swalayan Fadhila,” *J. Media Infotama*, vol. 11, no. 2, pp. 110–118, 2015.
- [5] Y. Irawan, “Penerapan data mining untuk evaluasi data penjualan menggunakan metode clustering dan algoritma hirarki divisive,” *Jtiulm*, vol. 3, no. 1, pp. 13–20, 2019.
- [6] Y. Mardi, “Data Mining: Klasifikasi Menggunakan Algoritma C4.5,” *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [7] B. A. B. Ii and L. Teori, “Input Proses Output,” vol. 2, pp. 9–26, 2009.
- [8] M. Effendi, N. A. Latifah, U. Islam, N. Sayyid, and A. Rahmatullah, “Jurnal Pendidikan dan Pembelajaran,” vol. 2, no. 2, pp. 127–142, 2021.
- [9] A. K. Muchsin and M. Sudarma, “Penerapan Fuzzy C-Means Untuk Penentuan Besar Uang Kuliah Tunggal Mahasiswa Baru,” *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 6, no. 3, p. 175, 2015, doi: 10.24843/lkjiti.2015.v06.i03.p04.
- [10] Y. D. Darmi and A. Setiawan, “Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk,” *J. Media Infotama*, vol. 12, no. 2, pp. 148–157, 2017, doi: 10.37676/jmi.v12i2.418.
- [11] T. Y. Tecnológico, “No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析Title,” vol. 1, pp. 1–26, 2018.
- [12] R. Rosmini, A. Fadlil, and S. Sunardi, “Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah,” *It J. Res. Dev.*, vol. 3, no. 1, pp. 22–31, 2018, doi: 10.25299/itjrd.2018.vol3(1).1773.
- [13] D. Menggunakan and M. Centroid, “Pengelompokan usia konsumen berdasarkan minat berbelanja dengan menggunakan metode centroid.”
- [14] B. Poerwanto and R. Y. Fa’rifah, “Analisis Cluster K-Means dalam Pengelompokan Kemampuan Mahasiswa,” *J. Sci. Pinisi*, vol. 2, no. 2, pp. 92–96, 2016.
- [15] M. R. Muttaqin and M. Defriani, “Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa,” *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [16] W. M. P. Duhita, “Clustering Menggunakan Metode K-Means Untuk,” *J. Inform.*, vol. 15, no. 2, pp. 160–174, 2016.
- [17] M Teguh Prihandoyo, “Unified Modeling Language (UML) Model Untuk Pengembangan Sistem Informasi Akademik Berbasis Web,” *J. Inform. J. Pengemb. IT*, vol. 3, no. 1, pp. 126–129, 2018.
- [18] I. Akil, P. Studi, M. Administrasi, and J. Timur, “REKAYASA PERANGKAT LUNAK DENGAN MODEL UNIFIED PROCESS STUDI KASUS: SISTEM INFORMASI JOURNAL,” no. 1, pp. 1–11, 2016.
- [19] Y. Heriyanto, “Perancangan Sistem Informasi Rental Mobil Berbasis Web Pada PT.APM RENT CAR,” vol. 2, no. 2, pp. 64–77, 2018.
- [20] P. Pt and A. P. M. Rent, “No Title,” vol. 2, no. 2, pp. 64–77, 2018.
- [21] G. Urva, H. F. Siregar, J. Prof, M. Y. Kisaran,

- and S. Utara, "Pemodelan UML E- Marketing Minyak Goreng," no. 9, pp. 92–101, 2015.
- [22] H. Purwanto and B. Desktop, "Rancangan Sistem Informasi Penjualan Barang Koperasi Xyz," *J. Sist. Inf. Univ. Suryadarma*, vol. 6, no. 1, 2014, doi: 10.35968/jsi.v6i1.278.
- [23] E. Iswandy, D. S. T. M. I. Komputer, and S. J. Padang, "Sistem Penunjang Keputusan Untuk Menentukan Dan Santunan Sosial Anak Nagari Dan Penyaluran Bagi Mahasiswa Dan Pelajar Kurang Mampu," *J. TEKNOIF*, vol. 3, no. 2, 2015, doi: 2338-2724.
- [24] F. Rahman and Santoso, "Aplikasi Pemesanan Undangan Online," vol. 1, pp. 78–87, 2015.
- [25] A. F. Ninuk Wiliani, "Digital menu pada x cafe berbasis desktop graphical user Interface dengan visual basic 2010 dan microsoft access 2007," vol. 6, no. 1, pp. 71–82, 2017.
- [26] A. Tahir, "OTOMATISASI PENGISIAN TANGKI AIR DENGAN VISUALISASI MENGGUNAKAN PEMROGRAMAN VISUAL BASIC," vol. 10, no. 1, pp. 330–338.
- [27] R. Irviani *et al.*, "APLIKASI PERPUSTAKAAN PADA SMA N1 KELUMBAYAN BARAT MENGGUNAKAN VISUAL BASIC," vol. 8, no. 1, 2017.
- [28] F. R. Firdaus, "Sistem informasi penjualan toko dan monitoring pendapatan/pengeluaran/stok belanja dengan menggunakan visual basic dan crystal report," pp. 127–131, 2018.

BIBLIOGRAFI PENULIS

	<p>Nama : Fitri Idasari Nirm : 2016020941 Program Studi : Sistem Informasi Deskripsi : Mahasiswa stambuk 2016. Saat ini sedang menempuh pendidikan Strata-1 (S1) di STMIK Triguna Dharma.</p>
	<p>Nama Lengkap : Muhammad Syahril, SE M.Kom Tempat dan Tanggal Lahir : Medan 6 Nopember 1978 Jenis Kelamin : Laki Laki Jabatan Fungsional : Lektor Pendidikan Tertinggi : Magister Komputer Status : Dosen Tetap Program Studi : Sistem Informasi NIP/NIDN : Alamat Email : msyahril@trigunadharm.ac.id No tlpn/HP : 082161333968 Bidang Keilmuan: 1. Disain Web 2. Manajemen Basis Data 3. Datamining beliau aktif sebagai beliau aktif sebagai Dosen Pembimbing 1 saya</p>
	<p>Dr. Ahmad Calam, S.Ag., M.A Program Studi : Sistem Informasi Deskripsi : Dosen Tetap STMIK Triguna Dharma yang aktif mengajar pada bidang PPKN. Prestasi : - beliau aktif sebagai beliau aktif sebagai Dosen Pembimbing 2 saya</p>