

## MEMANFAATKAN ALGORITMA K-MEANS DALAM MENENTUKAN PEGAWAI YANG LAYAK MENGIKUTI ASESSMENT CENTER UNTUK CLUSTERING PROGRAM SDP

Page | 87

Iin Parlina<sup>1</sup>, Agus Perdana Windarto<sup>2</sup>, Anjar Wanto<sup>3</sup>, M.Ridwan Lubis<sup>4</sup>

<sup>1,4</sup>Dosen Amik Tunas Bangsa Pematangsiantar, <sup>2,3</sup>Dosen STIKOM Tunas Bangsa

Jl.Jendral Sudirman Blok A No.1, 2, 3 Pematangsiantar

iin@amiktunasbangsa.ac.id<sup>1</sup>, agus@stikomtunasbangsa.ac.id<sup>2</sup>, anjar@stikomtunasbangsa.ac.id<sup>3</sup>,  
Ridwan@amiktunasbangsa.ac.id<sup>4</sup>

**Abstrak** — *Data mining merupakan teknik pengolahan data dalam jumlah besar untuk pengelompokan. Teknik Data mining mempunyai beberapa metode dalam mengelompokkan salah satu teknik yang dipakai penulis saat ini adalah K-Means. Dalam hal ini penulis mengelompokkan data daftar program SDP tahun 2017 untuk mengetahui manakah pegawai yang layak lolos dalam program SDP sehingga dapat melakukan Registrasi Aessment Center. Pengelompokan tersebut berdasarkan kriteria – kriteria data Program SDP. Pada penelitian ini, penulis menerapkan algoritma K-Means Clustering untuk pengelompokan data Program SDP di PT.Bank Syariah. Dalam hal ini, pada umumnya untuk memamasuki program SDP tersebut disesuaikan dengan ketentuan dan parameter Program SDP saja, namun dalam penelitian ini pengelompokan disesuaikan dengan kriteria – kriteria Program SDP seperti kedisiplinan pegawai, Target Kerja Pegawai, Kepatuhan Program SDP. Penulis menggunakan beberapa kriteria tersebut agar pengelompokan yang dihasilkan menjadi lebih optimal. Tujuan dari pengelompokan ini adalah terbentuknya kelompok SDP pada Program SDP yang menggunakan algoritma K-Means clustering. Hasil dari pengelompokan tersebut diperoleh tiga kelompok yaitu kelompok Lolos, Hampir Lolos dan Tidak Lolos. Terdapat pusat cluster dengan Cluster-1= 8;66;13, Cluster-2= 10;71;14 dan Cluster-3=7;60;12. Pusat cluster tersebut didapat dari beberapa iterasi sehingga menghasilkan pusat cluster yang optimal.*

**Kata kunci** — *K-Means, Data Mining, SDP*

### I. PENDAHULUAN

*Clustering* merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*) dan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik antara satu data dengan data lain[9]. Menurut kategori kekompakan, pengelompokan terbagi menjadi dua, yaitu komplet dan parsial. Jika semua data dapat bergabung menjadi satu, dapat dikatakan semua data kompak menjadi satu kelompok[2]. Pada *clustering* ini terdapat beberapa algoritma pengelompokan untuk mengelompokkan data secara mudah[1]. Salah satunya adalah algoritma *K-Means* yang merupakan metode analisis kelompok yang mengarah pada partisian  $N$  objek pengamatan ke dalam  $K$  kelompok, di mana setiap objek pengamatan sebuah kelompok data dengan mean (rata-rata) terdekat[2]. Pengelompokan pada umumnya diterapkan untuk mengelompokkan dokumen atau benda yang tidak tersusun dengan rapi dan tidak sesuai susunannya pada tempatnya[7]. Namun fungsi dari pengelompokan tidak hanya sekedar mengelompokkan dokumen atau benda[4]. Pengelompokan (*clustering*) dapat diterapkan dalam hal penentuan program SDP agar program SDP tersebut sesuai dengan kualifikasi dan memenuhi syarat untuk lanjut *assessment center*

pada program SDP[1]. Banyak Perusahaan seperti Perusahaan dalam bidang Perbankan selama ini untuk promosi-promosi jabatan hanya menggunakan nilai fisik sebagai tolak ukur untuk memilih yang mengikuti program *assessment center* lanjutan, sehingga pada akhirnya Pelaksanaan Program SDP ini menjadi salah mengambil keputusan sehingga data yang terdaftar di program SDP tidak memenuhi syarat administratif. Dalam hal ini, untuk pengelompokan data yang masuk ke daftar pegawai dalam program SDP lanjutan adalah pegawai yang memenuhi kualifikasi sesuai persyaratan administrasi. Data yang dapat menjadi tolak ukur untuk pengelompokan program SDP yaitu Prestasi pegawai (Target Kerja 75%), loyalitas pegawai (Kepatuhan 15%) dan Kedisiplinan 10%. Dengan demikian mengelompokkan program SDP yang dilakukan akan lebih efektif bagi Program SDP itu sendiri. Algoritma *K-Means* dalam hal ini akan mengelompokkan data tersebut sesuai kriteria – kriteria yang dipilih dari Program SDP.

### II. TINJAUAN PUSTAKA

#### A. Data Mining

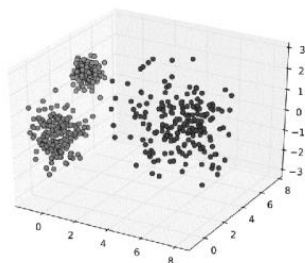
Istilah data mining mulai dikenal sejak tahun 1990, ketika pekerjaan pemanfaatan data menjadi

sesuatu yang penting dalam berbagai bidang, mulai dari bidang akademik, bisnis hingga edisi[6]. Munculnya data mining didasarkan pada jumlah data yang tersimpan dalam basis data semakin besar[8]. Dalam berbagai literatur, teori-teori pada data mining sudah ada sejak lama seperti antara lain K-Means Clustering dan text mining. Data mining disebut juga dengan *knowledge discovery in database (KDD)* ataupun *pattern recognition*[8]. Istilah KDD atau disebut penemuan pengetahuan data karena tujuan utama data mining adalah untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna[8]. Sedangkan istilah *pattern recognition* atau disebut pengenalan pola mempunyai tujuan pengetahuan yang akan digali dari dalam bongkahan data yang sedang dihadapi

### B. Dasar Cluster

Analisis kelompok (*cluster analysis*) adalah mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan diantaranya[1]. Analisis *Cluster* sebagai metodologi untuk klasifikasi data secara otomatis menjadi beberapa kelompok dengan menggunakan ukuran asosiasi, sehingga data yang sama berada dalam satu kelompok yang sama dan data yang berbeda berada dalam kelompok data yang tidak sama[9].

Masukan (*input*) untuk sistem analisis *cluster* adalah seperangkat data dan kesamaan ukuran (atau perbedaan) antara dua data. Sedangkan keluaran (*output*) dari analisis *cluster* adalah sejumlah kelompok yang membentuk sebuah partisi atau struktur partisi dari kumpulan data[10]. Salah satu hasil tambahan dari analisis *cluster* adalah deskripsi umum dari setiap *cluster* dan hal itu sangat penting untuk analisis lebih dalam dari karakteristik data set tersebut[9]. Ada saatnya di mana set data yang akan diproses dalam data mining belum diketahui label kelasnya. Pengelompokan data dilakukan dengan menggunakan algoritma yang sudah ditentukan dan selanjutnya data akan diproses oleh algoritma untuk dikelompokkan menurut karakteristik alaminya[6].



Gbr. 1 Pengelompokan dengan clustering

### C. Algoritma K-Means

Algoritma *K-Means* merupakan algoritma pengelompokan iteratif yang melakukan partisi set data ke dalam sejumlah *K cluster* yang sudah

ditetapkan di awal[3]. Algoritma *K-Means* sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi, umum Penggunaannya dalam praktek[2]. *K-Means* dapat diterapkan pada data yang direpresentasikan dalam *r*-dimensi ruang tempat[4]. *Kmeans* mengelompokkan set data *r*-dimensi,  $X = \{x_i/i=1, \dots, N\}$ [5].

Algoritma *K-Means* pengelompokan semua titik data dalam *X* sehingga setiap titik  $x_i$  hanya jatuh dalam satu *K* partisi[5]. Tujuan pengelompokan ini adalah untuk meminimalkan fungsi objek yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok[3].

Parameter yang hrsus dimasukkan ketika menggunakan algoritma *K-Means* adalah nilai *K*[5]. Nilai *K* yang digunakan pada umumnya didasarkan pada informasi yang diketahui sebelumnya mengenai sebenarnya berapa banyak *cluster* yang muncul dalam *X*, berapa banyak yang digunakan untuk penerapannya, atau jenis *cluster* dicari dengan melakukan percobaan dengan beberapa nilai *K*. Set representatif *cluster* dinyatakan  $C = \{c_{j/j=1, \dots, K}$ . sejumlah *K* representatif *cluster* tersebut sebagai *cluster centroid* (titik pusat *cluster*)[10]. Untuk set data dalam *X* dikelompokkan berdasarkan konsep kedekatan atau kemiripan, namun kuantitas yang digunakan untuk mengukurnya adalah ketidak miripan[2]. Metrik yang umum digunakan untuk ketidak miripan tersebut adalah *Euclidean*[1]. Secara umum algoritma *K-Means* memiliki langkah-langkah dalam pengelompokan, diantaranya:

1. Inisialisasi: menentukan nilai *K centroid* yang diinginkan dan metrik ketidakmiripan (jarak) yang diinginkan.
2. Memilih *K* data dari set *X* sebagai *centroid*. Untuk menentukan *centroid* dapat menggunakan persamaan (1).

$$\frac{\text{Jumlah Data}}{\text{Jumlah Class} + 1} \quad (1)$$

3. Mengalokasikan semua data ke *centroid* terdekat dengan matrik jarak yang telah ditetapkan
4. Menghitung kembali *centroid C* berdasarkan data yang mengikuti *cluster* masing – masing.
5. Mengulangi langkah 3 dan 4 hingga kondisi konvergen tercapai.

Berikut ini adalah rumus untuk menentukan jumlah *cluster*:

$$K = \sqrt{\frac{N}{2}} \quad (2)$$

Keterangan:

K = klaster

N = jumlah data

Menghitung jarak pada ruang jarak *Euclidean* menggunakan formula:

$$D(x_2, x_1) = \|x_2 - x_1\|_2 = \sqrt{\sum_{j=1}^p |x_{2j} - x_{1j}|^2}$$

(3)

Keterangan:

$D = euclidean\ distance$

$x =$  banyaknya objek

$\Sigma p =$  jumlah data *record*

### III. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengelompokan data Program SDP baru sekolah menengah program SDP dengan menggunakan algoritma *K-Means*. Agar tercapai tujuan tersebut, penulis akan melakukan pengujian dengan menggunakan data pada salah satu Perusahaan Perbankan di Pematangsiantar yaitu PT.Bank Syariah Mandiri area-2 tahun 2017. Adapun pengelompokan tersebut menggunakan langkah – langkah sebagai berikut:

#### A. Sumber Data

Dalam penelitian ini, sumber data diambil dari PT. Bank Syariah Mandiri Pematangsiantar dimana data tersebut merupakan data pegawai yang masuk ke daftar Program SDP baru tahun 2017.

TABEL 1  
DATA DAFTAR PROGRAM SDP TAHUN 2017

NO	NIP	NAMA	JABATAN SAAT INI	UNIT KERJA SAAT INI	UNIT KERJA PROVENSI	AREA
1	108273387	Rena Ridwan	Junior Consumer Banking Relationship Manager	KCP Satran Center	Area Batam	Area Batam
2		Muhammad Khalaf	Business Banking Verification Staff	Retail Financing Recovery Group	KCP Batam	Area Batam
3	110775207	Muhammad	Financing Compliance & Legal Administration Staff	Financing Operation Group	Area Batam	Area Batam
4	098679707	Yuni Restihasari	Junior Consumer Banking Relationship Manager	KCP Batam Indo	Area Batam	Area Batam
5	118679651	M. Irfan Anhar	Junior Consumer Banking Relationship Manager	KCP Asamhas	Area Batam	Area Batam
6	117879655	Arenn Teva	Junior Consumer Banking Relationship Manager	KCP Batam Sangkong	Area Batam	Area Batam
7	120712113	Andika Septian	Junior Consumer Banking Relationship Manager	KC Langsa	KC Langsa	Area Medan Gajah Mada
8	129412394	Muhammad Kurniawan	Account Maintenance Staff	Retail Financing Recovery Group	Retail Financing Region	Area Aceh
9	118579195	Aryandi	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Area Batam
10	108679118	Adhira Darmadi	Financing Compliance & Legal Administration Staff	Financing Operation Group	Financing Operation	Area Batam
11	120911296	Rischa Alia	Customer Service	KCP Medan Sukarama	KCP Medan Sukarama	Area Medan Almam' Yasi
12	120211805	Sulartono	Account Maintenance Staff	Retail Financing Recovery Group	Retail Financing Region	Area Medan Almam' Yasi
13	108675347	Dermawan Hamonangan H	Financing Compliance & Legal Administration Staff	Financing Operation Group	Financing Operation	Area Pekanbaru
14	108579825	Rony Octonangga	Account Maintenance Staff	Retail Financing Recovery Group	Retail Financing Region	Area Pematangsiantar
15	108277816	Andriani Lubis	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Area Pematangsiantar
16	108579854	Rudy Bahaji Muchtar	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Area Aceh
17	118579634	Sari Diana	Financing Compliance & Legal Administration Staff	Financing Operation Group	Financing Operation	Area Aceh
18	108579834	Indri	Account Maintenance Staff	Retail Financing Recovery Group	Retail Financing Region	Area Medan Gajah Mada
19	118779861	Debi Kurnia	Loan Processing & Document Custody Staff	Financing Operation Group	Retail Financing Region	Area Pekanbaru
20	118679219	Andi Saputra	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Area Pekanbaru
21	118779813	Areni Kurniawan	Financing Compliance & Legal Administration Staff	Financing Operation Group	Financing Operation	Area Pekanbaru
22	120810764	Iqbal Darmawan	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Area Pematangsiantar
23	118579191	Winni Wardana	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Area Pematangsiantar
24	108579119	Dibi Irvan	Financing Compliance & Legal Administration Staff	Financing Operation Group	Financing Operation	Region I / Sumatera I
25	078270284	Amita Rahma	Human Capital Administration Staff	Region I / Sumatera I	Region I / Sumatera I	Region I / Sumatera I
26	108779315	Taufik Kholid Pullang	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Region I / Sumatera I
27	098574211	Nural Adh Sucanto	Loan Processing & Document Custody Staff	Financing Operation Group	Financing Operation	Region I / Sumatera I

Gbr.2. Data Daftar Program SDP

#### B. Pengolahan Data

Sebelum data Program SDP dikelompokan ke dalam Seleksi SDP yang sesuai kriteria. Data mentah tersebut akan di transformasi dengan cara menginisialisasi data ke dalam bentuk angka yang dapat diolah dalam pengelompokan.

TABEL 2  
RANGE PENGELOMPOKAN PROGRAM SDP

Cluster	ΣMin	ΣMax	Cluster
1	6	18	<b>Kedisiplinan</b>
2	10	13	<b>Kinerja</b>
3	27	0	<b>Kepatuhan</b>

Berdasarkan *range* yang telah ditentukan pada pengelompokan jurusan, maka dapat Disimpulkan bahwa pengelompokan data Program SDP ke dalam *Assessment Center* tanpa menggunakan metode. Pada data tersebut, tidak semua kriteria yang dapat menjadi patokan. Dalam hal ini, kriteria yang diinisialisasikan adalah Kedisiplinan diubah menjadi X1, Kinerja diubah menjadi X2, dan Kepatuhan diubah menjadi X3.

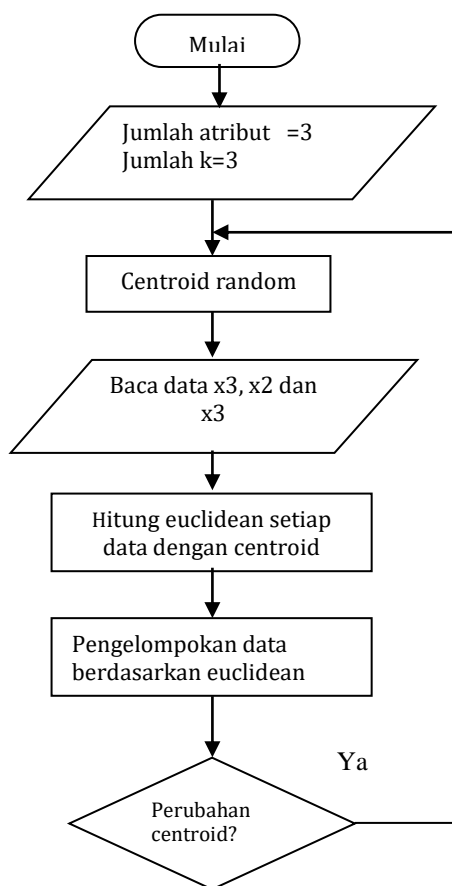
TABEL 3  
HASIL INISIALISASI DATA SDP

Alternatif	Kriteria			Jumlah	Cluster
	C1	C2	C3		
A1	11	72	14	97	3
A2	11	73	15	98	5
A3	11	73	15	98	2
A4	11	72	14	97	3
A5	11	73	15	98	3
A6	10	72	14	97	3
A7	10	72	14	97	3
A8	10	71	14	95	4
A9	10	71	14	95	1
A10	11	72	14	97	2
A11	7	62	12	82	5
A12	6	61	12	79	4
A13	5	60	12	78	2
A14	6	56	11	73	4
A15	9	60	12	81	1
A16	9	61	12	82	1
A17	8	60	12	80	2
A18	9	61	12	82	4
A19	10	60	12	82	1
A20	9	60	12	82	1
A21	9	65	13	87	2
A22	9	64	13	86	1
A23	8	64	13	85	1

A24	9	69	14	91	2
A25	9	69	14	91	5
A26	10	70	14	94	1
A27	10	69	14	92	1

Page | 90 C. Proses Data

Setelah data diolah, maka langkah selanjutnya adalah data diproses untuk membentuk pengelompokan data ke dalam daftar program SDP sesuai kriteria yang telah ditentukan. Data yang telah ditransformasi tersebut akan diproses dengan menggunakan sebuah algoritma pengelompokan (*cluster*) yaitu algoritma *K-means*. Berikut ini adalah sebuah *flowchart* dalam pengelompokan program SDP pada Program SDP di HCG Pada PT. Bank Syariah Mandiri Pematangsiantar.



Gbr. 2 Flowchart Pengclusteran K-Means

Pada tahapan ini yaitu menentukan *centroid K-means* dari data Program SDP tahun 2017. Dalam menentukan target *K-means* tersebut bertujuan untuk mendapatkan target data atau jarak kelompok yaitu titik pusat (*centroid*) kelompok awal untuk menghitung algoritma *K-means*. Pada tabel 4 merupakan hasil perhitungan *K-means* untuk data sampel Program SDP baru periode tahun 2017 pada

HCG di PT.Bank Syariah Mandiri Pematangsiantar yang berdasarkan kriteria-kriteria yang telah ditentukan untuk setiap kelompok (*cluster*).

Dalam hal ini, pengelompokan yang terbentuk terjadi sebanyak 3 kelompok (*cluster*). Untuk representasi perhitungan secara lengkap dapat dilihat pada tabel – tabel berikut:

TABEL 4  
TARGET K-MENS DATA SAMPEL PROGRAM SDP UNTUK C1

NIP	KRITERIA		
	C1	C2	C3
108675347	10	71	14
098673907	9	60	12
118679651	9	61	12
118578195	10	60	12
108277016	9	60	12
118778861	9	64	13
118679219	8	64	13
108775315	10	70	14
098574221	10	69	14
<b>cluster 1</b>	<b>9</b>	<b>65</b>	<b>13</b>

Setelah pengelompokan untuk kelompok pertama (*cluster-1*) telah terkelompok. Maka dapat dilakukan untuk pengelompokan data kelompok kedua (*cluster- 2*).

TABEL 5  
TARGET K-MENS DATA SAMPEL PROGRAM SDP UNTUK C2

NIP	KRITERIA		
	C1	C2	C3
108576825	11	73	15
118579634	11	72	14
078273284	5	60	12
117879655	8	60	12
108576854	9	65	13
128810764	9	69	14
128412364	11	72	14
108576834	11	72	14
118510236	11	73	15
128911296	10	72	14
1107775207	10	72	14
<b>Cluster 2</b>	<b>10</b>	<b>69</b>	<b>14</b>

Data yang dikelompokkan lainnya akan masuk pada kelompok 3 (*cluster 3*) yaitu terdapat tujuh record data pada kelompok 3.

TABEL 6  
TARGET K-MENS DATA SAMPEL PROGRAM SDP UNTUK C3

NIP	KRITERIA		
	C1	C2	C3
108675118	10	71	14
108575319	6	61	12
108275387	6	56	11
128712213	9	61	12
128211825	11	73	15
118378913	7	62	12
118579191	9	69	14
<b>Cluster 3</b>	<b>8</b>	<b>65</b>	<b>13</b>

Page | 91

Setelah semua data telah terkelompok, maka untuk menentukan titik pusat pada cluster dapat diambil nilai rata – rata dari setiap. Adapun tahapan proses algoritma *K-means* adalah sebagai berikut:

1. Menentukan nilai *k* dari jumlah cluster yang akan dibentuk. Pada tahap ini cluster yang terbentuk sebanyak 3 cluster.
2. Menentukan titik pusat (*centroid*) awal dari setiap cluster.

Dalam penelitian ini titik pusat awal ditentukan dengan menggunakan range dari program SDP.

TABEL 7  
CENTROID AWAL CLUSTER

Cluster	C1	C2	C3
X1	9	65	13
X2	10	69	14
X3	8	65	13

Setelah titik pusat cluster ditentukan, maka tahap selanjutnya adalah menghitung jarak terdekat atau *euclidean* dengan menggunakan persamaan 3, yaitu menghitung jarak dari data Program SDP pertama ke titik pusat cluster.

$$D(1,1) = \sqrt{(10 - 11)^2 + (69 - 72)^2 + (13 - 14)^2} = 16,8$$

Dari hasil perhitungan di atas didapatkan bahwa jarak terdekat bernilai 16,8. Selanjutnya perhitungan jarak data pertama untuk cluster kedua.

$$D(1,2) = \sqrt{(9 - 11)^2 + (64 - 73)^2 + (13 - 15)^2} = 0,3$$

Dari hasil perhitungan tersebut didapatkan jarak terdekat bernilai 0,3. Kemudian perhitungan jarak data pertama untuk cluster ketiga.

$$D(1,3) = \sqrt{(8 - 5)^2 + (64 - 2)^2 + (13 - 6)^2} = 2,0$$

Dari hasil perhitungan tersebut didapatkan jarak terdekat bernilai 2,0. Berdasarkan perhitungan *euclidean* untuk data pertama diperoleh jarak terdekat cluster adalah 16,8 yang terkelompok pada cluster kesatu (C1) dan untuk data kedua diperoleh jarak terdekat cluster adalah 0,3 yang terkelompok pada cluster kedua (C2) dan untuk data ketiga diperoleh jarak terdekat cluster adalah 2,0.

TABEL 8  
HASIL PROSES PERHIRUNGAN JARAK TERDEKAT (EUCLIDEAN)

NIP	KRITERIA			CLUSTER			Jarak Terdekat
	C1	C2	C3	X1	X2	X3	
098574221	10	69	14	16,8	0,2	16,7	0,2
118778861	9	64	13	0,3	22,2	0,3	0,3
118679219	8	64	13	2,0	24,8	0,3	0,3
108576854	9	65	13	0,6	16,3	0,4	0,4
118579191	9	69	14	16,6	0,5	15,7	0,5
128810764	9	69	14	16,9	1,3	15,1	1,3
108775315	10	70	14	31,5	1,5	32,6	1,5
108675118	10	71	14	47,6	5,5	47,2	5,5
108675347	10	71	14	47,6	5,5	47,2	5,5
118378913	7	62	12	10,0	52,5	7,5	7,5
128911296	10	72	14	58,5	9,7	58,7	9,7
1107775207	10	72	14	58,5	9,7	58,7	9,7
128412364	11	72	14	59,9	10,6	60,9	10,6
108576834	11	72	14	59,9	10,6	60,9	10,6
118579634	11	72	14	59,9	10,6	60,9	10,6
118679651	9	61	12	12,8	65,9	14,3	12,8
128712213	9	61	12	12,8	65,9	14,3	12,8
128211825	11	73	15	70,8	15,3	71,5	15,3
108576825	11	73	15	70,8	15,3	71,5	15,3
118510236	11	73	15	70,8	15,3	71,5	15,3
108277016	9	60	12	17,9	76,8	20,6	17,9
118578195	10	60	12	18,1	76,5	21,6	18,1
098673907	9	60	12	18,1	77,6	20,0	18,1
108575319	6	61	12	22,9	78,9	19,2	19,2
117879655	8	60	12	25,1	91,3	26,4	25,1
078273284	5	60	12	31,4	94,1	27,2	27,2
108275387	6	56	11	80,6	180,4	79,1	79,1

Setelah data dikelompokkan pada iterasi pertama, maka langkah selanjutnya adalah membentuk titik pusat baru dengan menentukan nilai rata-rata dari



setiap data yang sudah membentuk *cluster* untuk melanjutkan perhitungan jarak terdekat iterasi kedua dan akan menunjukkan *cluster* yang terbentuk selanjutnya dapat membentuk *cluster* yang konvergen.

TABEL 9  
CENTROID BARU PERTAMA

Cluster	C1	C2	C3
1	8,69	66,24	13,25
2	9,50	69,62	13,92
3	8,99	64,98	13,00

Jika *cluster* belum konvergen, maka *centroid* akan dibangkitkan kembali dan menghitung kembali *euclidean* dari setiap data Program SDP. Langkah perhitungan jarak terdekat seperti langkah ketiga sebelumnya. Berikut perhitungan jarak terdekat untuk data kedua.

$$D(2,1) = \sqrt{(3-8,69)^2 + (3-66,24)^2 + (3-13,25)^2} = 2.0518$$

Dari hasil perhitungan tersebut didapatkan jarak terdekat bernilai 2.0518. Untuk perhitungan *euclidean* dapat dilihat secara lengkap pada tabel di bawah ini.

TABEL 10  
PERHITUNGAN EUCLIDEAN ITERASI-2

NIP	KRITERIA			CLUSTER			Jarak Terdekat ke Cluster
	C1	C2	C3	X1	X2	X3	
108576854	9	65	13	1,2	21,5	0,2	0,2
118778861	9	64	13	3,2	28,3	0,4	0,4
108775315	10	70	14	16,9	0,9	27,3	0,9
098574221	10	69	14	6,2	1,2	13,6	1,2
118579191	9	69	14	5,7	1,4	13,3	1,4
118679219	8	64	13	4,2	30,7	1,9	1,9
128810764	9	69	14	5,5	2,0	13,4	2,0
108675118	10	71	14	27,8	3,0	42,1	3,0
108675347	10	71	14	27,8	3,0	42,1	3,0
128911296	10	72	14	36,8	6,5	52,5	6,5
1107775207	10	72	14	36,8	6,5	52,5	6,5
128412364	11	72	14	38,5	7,5	54,0	7,5
108576834	11	72	14	38,5	7,5	54,0	7,5
118579634	11	72	14	38,5	7,5	54,0	7,5
128211825	11	73	15	47,1	11,2	64,3	11,2
108576825	11	73	15	47,1	11,2	64,3	11,2
118510236	11	73	15	47,1	11,2	64,3	11,2
118378913	7	62	12	18,6	60,9	11,5	11,5
118679651	9	61	12	27,7	76,5	15,9	15,9
128712213	9	61	12	27,7	76,5	15,9	15,9

108277016	9	60	12	35,6	88,4	21,7	21,7
098673907	9	60	12	35,4	89,1	21,9	21,9
118578195	10	60	12	36,1	88,2	22,0	22,0
108575319	6	61	12	35,5	88,8	25,4	25,4
117879655	8	60	12	44,4	103,6	29,4	29,4
078273284	5	60	12	46,0	104,8	34,4	34,4
108275387	6	56	11	110,0	196,6	87,3	87,3

Pada perhitungan jarak terdekat di atas *centroid* baru yang dibangkitkan ternyata belum konvergen, sehingga iterasi harus dilanjutkan. Dalam penelitian ini, iterasi *clustering* data Program SDP terjadi sebanyak 3 kali iterasi. Pada pengelompokan data di atas dan diperoleh titik *cluster* baru dengan nilai sebagai berikut:

TABEL 11  
CENTROID BARU KEDUA

Cluster	C1	C2	C3
1	8,97	66,86	13,37
2	10,41	71,70	14,34
3	7,65	60,34	12,07

*Centroid* baru kedua yang dibangkitkan ternyata sudah konvergen, sehingga perhitungan jarak terdekat *clustering* berhenti pada iterasi-3. Adapun perhitungan secara lengkapnya sebagai berikut:

TABEL 12  
PERHITUNGAN EUCLIDEAN ITERASI-3

NIP	KRITERIA			CLUSTER			Jarak Terdekat ke Cluster
	C1	C2	C3	X1	X2	X3	
108576854	9	65	13	3,1	47,8	25,1	3,1
118778861	9	64	13	6,0	57,5	18,7	6,0
108775315	10	70	14	11,8	3,3	102,9	3,3
098574221	10	69	14	3,3	11,1	73,6	3,3
118579191	9	69	14	3,0	12,1	72,1	3,0
118679219	8	64	13	7,5	61,6	17,8	7,5
128810764	9	69	14	3,1	13,6	71,0	3,1
108675118	10	71	14	21,2	0,4	129,5	0,4
108675347	10	71	14	21,2	0,4	129,5	0,4
128911296	10	72	14	29,0	0,1	147,7	0,1
1107775207	10	72	14	29,0	0,1	147,7	0,1
128412364	11	72	14	30,5	0,3	150,4	0,3
108576834	11	72	14	30,5	0,3	150,4	0,3
118579634	11	72	14	30,5	0,3	150,4	0,3

128211825	11	73	15	38,2	1,1	167,3	1,1
108576825	11	73	15	38,2	1,1	167,3	1,1
118510236	11	73	15	38,2	1,1	167,3	1,1
118378913	7	62	12	25,0	102,2	5,3	5,3
118679651	9	61	12	34,8	120,4	1,5	1,5
128712213	9	61	12	34,8	120,4	1,5	1,5
108277016	9	60	12	43,2	134,4	2,0	2,0
098673907	9	60	12	43,4	135,9	0,9	0,9
118578195	10	60	12	43,5	133,4	3,4	3,4
108575319	6	61	12	44,2	137,6	3,7	3,7
117879655	8	60	12	53,5	154,1	0,6	0,6
078273284	5	60	12	55,8	157,4	4,9	4,9
108275387	6	56	11	124, 7	266,0	19,8	19, 8

Pada iterasi-3 tersebut, titik pusat dari setiap *cluster* sudah tidak berubah dan tidak ada lagi terdapat data yang berpindah dari satu *cluster* ke *cluster* yang lain. Maka dari itu, data yang telah terkelompok sesuai dengan kriteria akan melanjutkan proses ujian Assesmen Center pada bagian HCG PT. Bank Syariah Mandiri Pematang siantar menggunakan model algoritma *K-Means*.

#### IV. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, maka penulis dapat menarik beberapa kesimpulan, diantaranya sebagai berikut:

1. Pengujian yang dilakukan dalam penelitian ini, iterasi *clustering* pada data daftar Program SDP terjadi sebanyak 3 kali iterasi.
2. Berdasarkan hasil *cluster* dengan menerapkan beberapa kriteria dari daftar Program SDP menggunakan algoritma *K-Means* dapat diambil pengelompokan dengan rata-rata data program SDP yang dapat melakukan Assessment center lanjutan adalah yang lolos dan hasil klasifikasi program SDP yang Hampir lolos harus memperbaiki Administrasi seperti Kedisiplinan dari bulan juni sampai bulan oktober agar dapat mengikuti Assessment Center lanjutan, sedangkan hasil klasifikasi daftar data program SDP yang Tidak lolos harus memperbaiki data kedisiplinannya selama 1 Tahun.

#### REFERENSI

- [1] I Parlina, H Mawengkang, S Efendi InfoTekJar (2017) "Analisis Kinerja Algoritma Clustering Fuzzy Tsukamoto Dengan Fuzzy C-Means" (Jurnal Nasional Informatika dan Teknologi Jaringan) 1 (2), 90-94
- [2] Nur, F., Zarlis, M., & Nasution, B. B. (2017). *Penerapan Algoritma K-Means Pada Siswa Baru Sekolahmenengah*

- Kejuruan Untuk Clustering Jurusan. InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, 1(2), 100-105.
- [3] N Butarbutar, AP Windarto, D Hartama, S Solikhun Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika) 1 (1), 46-55
  - [4] AP Windarto International Journal of artificial intelligence research 1 (2), 26-33
  - [5] MG Sadewo, AP Windarto, D Hartama InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan) 2 (1), 60-67
  - [6] AP Windarto "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method" International Journal of Artificial Intelligence Research 1 (2)
  - [7] N Butarbutar, AP Windarto, D Hartama, S Solikhun "Komparasi Kinerja Algoritma Fuzzy C-Means Dan K-Means Dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai Akademik Siswa" Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika) 1 (1), 46-55
  - [8] MG Sadewo, AP Windarto, D Hartama "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
  - [9] RW Sembiring, JM Zain, A Embong "Clustering high dimensional data using subspace and projected clustering algorithms" arXiv preprint arXiv:1009.0384
  - [10] PA Zizwan, M Zarlis, EB Nababan "Performance Analysis Of Combined Methods Of Genetic Algorithm And K-Means Clustering In Determining The Value Of Centroid" Journal of Physics: Conference Series 930 (1), 012008