



## Pengelompokan Mahasiswa Potensial Drop Out menggunakan Metode Clustering K-Means.

Ieannoal Vhallah<sup>a</sup>, Sumijan<sup>b</sup>, Julius Santony<sup>c</sup>

<sup>a</sup>Fakultas Ilmu Komputer, Universitas Putra Indoensia YPTK Padang, ieannoal@gmail.com

<sup>b</sup>Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, sumijan@upi.ptk.ac.id

<sup>c</sup>Fakultas Ilmu Komputer, Universitas Putra Indonesia YPTK Padang, julius\_santony@yahoo.com

### Abstract

*Clustering K Means is used for grouping. The K-Means method seeks to group the existing data into several unique groups, where data in one group have the same characteristics with each other and have different characteristics than the data exists in the other group. To perform student grouping the potential drop out required attributes. Total Semester Credit System, Comulative Achievement Index, and Total Semester. Clustering process K- Mean is done by determining the nearest initial centroid point in a group of potential drop out students. Clustering results K-Mean by Total Credit System semester, Comulative graduate point, and Total Semester. Results Clustering of potential drop out students for class of 2014 is in cluster 0 are 4 students or 30.77% of 13 Samples, class of 2015 in cluster 1 are 4 students and cluster 2 are to 2 students or 66.7% of 9 samples, the force of 2016 is in cluster 0 amounting to 2 students and cluster 1 is 10 students or 50% from 24 samples, and force of 2017 is in cluster 2 strength 4 student or 22,22% from 18*

*Keywords: Data Mining, Clustering, K-Means, Potensial Drop Out.*

### Abstrak

Clustering K Means digunakan untuk melakukan pengelompokan. Metode K-Means berusaha mengelompokkan data yang ada dalam beberapa kelompok yang unik, dimana data dalam satu kelompok memiliki karakteristik yang sama satu sama lainnya dan memiliki karakteristik yang berbeda dengan data di kelompok yang lain. Oleh sebab itu cara ini digunakan untuk melakukan pengelompokan mahasiswa potensi drop out dengan beberapa atribut yaitu Sistem Kredit Semester Kumulatif, Indek Prestasi Kumulatif, dan Semester Kumulatif. Proses Clustering K- Mean dilakukan dengan menentukan titik centroid awal secara acak dalam satu kelompok mahasiswa. Sehingga diperoleh hasil dengan metode K Mean mahasiswa yang potensial drop out, Urntuk angkatan 2014 berada pada cluster 0 berjumlah 4 orang mahasiswa atau 30,77% dari 13 Sampel, angkatan 2015 berada pada cluster 1 berjumlah 4 mahasiswa dan cluster 2 berjumlah 2 mahasiswa atau 66,7% dari 9 sampel, angkatan 2016 berada pada cluster 0 berjumlah 2 mahasiswa dan cluster 1 berjumlah 10 mahasiswa atau 50% dari 24 sampel, dan angkatan 2017 berada pada cluster 2 kekuatan 4 mahasiswa atau 22,22% dari 18 sampel.

Kata Kunci: Data Mining, Klastering, K-Means, Potential Drop Out

© 2018 Jurnal RESTI

### 1. Pendahuluan

Pemanfaatan sistem informasi berbasis teknologi komputeryang sudah dimanfaatkan selama bertahun-tahun menghasilkan jumlah data yang cukup besar. Data yang dihasilkan tersimpan dalam sistem komputer yang dirancang agar cepat dan akurat baik dalam mengoperasikan maupun administrasinya. Kumpulan data Prestasi Akademik tersebut memiliki informasi untuk pengembangan institusi. Perkembangan pesat

*Internet-of-Things (IoT)* berpotensi menghasilkan sejumlah besar data deret waktu[1].

Data prestasi akademik mahasiswa yang potensial DO merupakan sebuah data yang sangat penting dalam nenentukan langkah apa yang akan dilakukan sebuah institusi perguruan tinggi. Pola informasi data yang dihasilkanberfokus pada pengembangan berbagai cara untuk menganalisis jenis data unik di lingkungan pendidikan [2].

Diterima Redaksi : 19-03-2018 | Selesai Revisi : 01-06-2018 | Diterbitkan Online : 02-08-2018

Data mining adalah langkah analisis dari penemuan pengetahuan dalam database atau *Knowledge Discovery Database* (KDD). Untuk menerapkan mengolah data dengan terlebih dahulu harus ditentukan data apa saja yang diperlukan untuk mendapatkan data mahasiswa berpotensi Drop Out (DO).

Teknik data mining dengan metode algoritma clustering K Means dapat menemukan karakteristik-karakteristik dari nilai prestasi mahasiswa dan menggunakan karakteristik tersebut untuk melakukan pengelompokan (*Clustering*).

Menggunakan data mining untuk membantu pengembangan kurikulum yang baru dan membantu mahasiswa teknik untuk menseleksi bidang utamanya. [4]. Teknik pemilihan atribut pertambangan adalah pendekatan pilihan atribut berbasis objektif dibandingkan untuk pendekatan seleksi atribut subyektif, yang mengandalkan kajian masa lalu, akal sehat, dan konsultasi berbasis teori, [5].

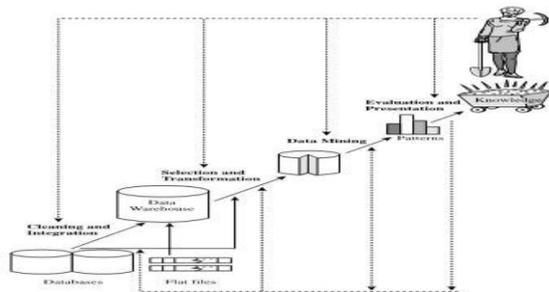
Perkembangan ilmu pengetahuan dan teknologi yang sangat pesat, akan mendorong sebuah institusi untuk memanfaatkan teknologi tersebut dalam segala bidang, salah satunya system pengolahan data yang sangat cepat dan akurat. Untuk memperoleh data yang akurat tersebut Fakultas Pertanian Universitas Riau memerlukan sebuah sistem pengolahan data mahasiswa yang potensial DO.

## 2. Tinjauan Pustaka

### 2.1 Penambangan Data ( *Data Mining* )

*Data Mining* merupakan proses untuk menggali nilai tambah dari sekumpulan data berupa pengetahuan yang selama ini tersembunyi atau tidak diketahui secara manual [3]. Data mining juga dapat digunakan untuk mencari pola- pola data yang dapat dijadikan informasi berharga dengan mengekstrak dan menarik informasi dari basis data yang diperoleh tersebut.

Proses untuk mencari nilai tambah dari sekumpulan data dikenal juga sebagai penemuan pengetahuan dari pangkalan data (*Knowledge Discovery in Databases = KDD*) yaitu tahap-tahap yang dilakukan dalam menggali pengetahuan dari sekumpulan data. Tahap-tahap yang dimaksud terlihat pada Gambar 1.



Gambar 1. Tahapan *Data mining*[3]

Tahap-tahap data mining dapat dijelaskan sebagai berikut:

1. Pembersihan data (*data cleaning*)
2. Integrasi data (*data integration*) yaitu penggabungan data dari berbagai sumber.
3. Seleksi data (*data selection*) yaitu memilih data yang sesuai dengan jenis penelitian.
4. Transformasi data (*data transformation*)
5. Proses *mining* Evaluasi pola yang ditemukan (untuk menemukan informasi dan pengetahuan yang menarik).
6. Proses Pola (*pattern evaluation*).

### 2.2 K-Means Clustering

K-Means Clustering merupakan teknik dalam klaster data yang sangat terkenal karena kecepatannya dalam mengklasterkan data. Akan tetapi K-Means Clustering memiliki kelemahan didalam memproses data yang berdimensi banyak. Khususnya untuk masukan yang bersifat *non-linierly separable*.

Data akademik mahasiswa bersal dari nilai mahasiswa tiap semester dapat di lakukan clustering dengan etode K-Mean[7]. K-Means clustering juga tidak mampu mengrupkan data yang bertipe kategorikal dan juga data campuran (numeric dan kategorikal) [8].Kenyataan didunia nyata data yang tersedia atau yang diperoleh memiliki dimensi yang banyak dan juga bersifat campuran. Untuk mengatasi permasalahan ini, telah banyak diusulkan oleh parapeneliti metode-metode yang dapat mengatasi kelemahan ini, salah satudi antaranya adalah Kernel K-Means Clustering [6]

Adapun langkah algoritma *K-means* adalah sebagai berikut :

1. Tentukan nilai *k* sebagai jumlah klaster yang ingin dibentuk.
2. Inisialisasi *k* sebagai centroid yang dapat dibangkitkan secara random.

$$V = \frac{\sum_{i=1}^n x_i}{n} \dots \dots \dots (1)$$

Dimana  
*V*= Centroid  
*x<sub>i</sub>* = Objek ke *i*  
*n*= Banyaknya Objek yang menjadi anggota cluster

3. Hitung jarak setiap data ke masing-masing centroid menggunakan persamaan *Euclidean Distance* yaitu sebagai berikut.

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j, C_j)^2} \dots \dots \dots (2)$$

Dimana  
*d* = Jarak  
*j* = Banyak Data  
*c* = Centroid  
*x* = Data

4. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan centroidnya.
5. Tentukan posisi *centroid* baru ( $k$ )

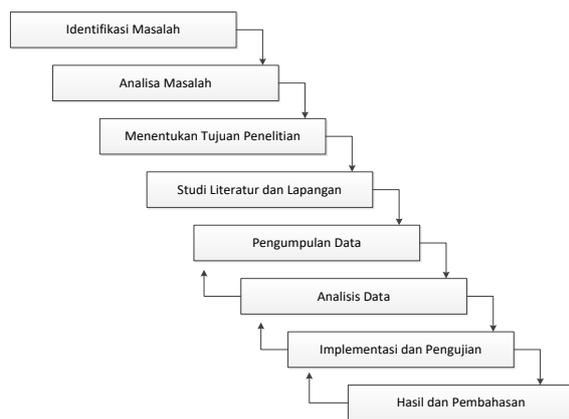
Kembali ke langkah 3 jika posisi *centroid* baru dengan *centroid* lama tidak sama.

### 3. Metodologi Penelitian

Pada tahap awal penelitian dilakukan dengan cara mengambil data mahasiswa pada Bagian Akademik Fakultas Pertanian Universitas Riau melalui persetujuan Dekan, setelah mengirim surat terlebih dahulu untuk mendapatkan data tersebut, data tersebut akan menjadi subjek penelitian, yaitu berupa Indeks Prestasi Kumulatif, SKS Kumulatif dan Tahun Masuk. Kemudian data tersebut di olah dengan menggunakan metode *Clustering K Means*.

#### 3.1 Kerangka Kerja Penelitian.

Untuk memudahkan dalam melakukan penelitian perlu dibuat alur kerangka kerja yang akan dilakukan pada penelitian ini. Alur kerangka kerja penelitian digambarkan dalam diagram aktivitas pada ( Gambar 2)



Gambar 2. Kerangka Penelitian.

Berdasarkan kerangka kerja penelitian di atas, maka masing-masing langkahnya dapat diuraikan seperti berikut ini:

1. Identifikasi masalah  
Pada tahap ini dilakukan proses penentuan masalah dan mendefinisikan batasan masalah yang akan diteliti, sehingga tujuan penelitian dapat tercapai dengan baik.
2. Menganalisis Masalah  
Pada tahap ini dilakukan analisis masalah untuk dapat memahami masalah yang telah ditentukan ruang lingkup atau batasannya. dengan menganalisis masalah yang telah ditentukan tersebut, maka diharapkan masalah dapat dipahami dengan baik.
3. Menentukan tujuan penelitian  
Pada tahap ini dikemukakan tujuan yang ingin dicapai melalui proses penelitian. Tujuan penelitian harus jelas dan tegas. Tujuan penelitian adalah

suatu target yang akan dicapai untuk mengatasi masalah-masalah yang ada.

4. Studi literatur  
Setelah melakukan identifikasi masalah, kemudian menganalisa data dan menentukan tujuan, langkah selanjutnya adalah mempelajari literatur yang berhubungan dengan judul. Sumber literatur didapatkan dari buku, jurnal, artikel, yang membahas tentang Metode K Mean Clustering, data mining dan bahan bacaan lain yang mendukung penelitian.
5. Pengumpulan data.

Dalam tahap pengumpulan data untuk proses K-Means dilakukan beberapa cara yaitu :

- a. Melakukan pengambilan data dari data base Akademik Mahasiswa yang ada di Fakultas Pertanian Universitas Riau melalui persetujuan Dekan.
- b. Melakukan studi pustaka dengan membaca buku-buku dan jurnal ilmiah yang berhubungan dengan permasalahan untuk dapat menganalisis data dan informasi yang dibutuhkan.

#### 3.2 Analisis Data

Analisis data diperlukan untuk mendapatkan data yang sesuai dengan kriteria- kriteria untuk pengelompokan mahasiswa yang potensial drop out dengan metode clustering K Means. Data akademik yaitu data nilai yang bersumber dari Hasil studi mahasiswa setiap semester [20]

Untuk dapat melakukan analisis dan pengelompokan dari mahasiswa yang berpotensi drop out diperlukan data sampel, dimana data sampel yang akan digunakan dalam penelitian ini diperoleh dari bagian administrasi dan akademik Fakultas Pertanian Universitas Riau.

Metode pengumpulan data yang digunakan yaitu : wawancara, mengambil data akademik Mahasiswa pada bagian Administrasi Akademik data berupa Sistem Kredit Semester Total ( SKS Total), Indeks Prestasi (IP) dan Semester untuk mahasiswa Angkatan 2014,2015, 2016 dan 2017 periode semester Ganjil 2017/2018 yang memiliki  $IPK \leq 2,50$  dari 1696 Mahasiswa Aktif.

Dari seluruh data akademik mahasiswa diambil mahasiswa yang memiliki IPK Kurang dari 2,5 Sehingga didapat data akademik mahasiswa sebanyak 64 mahasiswa yang terdiri dari berbagai program studi

Adapun data yang dijadikan objek penelitian ini adalah data Akademik Mahasiswa angkatan 2014, 2015,2016 dan 2017 seperti terlihat pada tabel 1,2,3 dan 4

Tabel 1 . Data Akademik Mahasiswa Angkatan 2014.

NO	Kode Program Studi	NIM	SKS Total	IPK	Semester
1	54211	1406110606	132	2,5	7
2	54211	1406111097	116	2,10	7
3	41231	1406111311	130	2,41	7
4	41231	1406112089	114	1,89	7
5	54211	1406112491	125	2,27	7
6	41231	1406113572	128	2,29	7
7	54211	1406114014	121	2,06	7
8	41231	1406118121	108	2,05	7
9	41231	1406119543	140	2,41	7
10	54211	1406119994	130	2,50	7
11	41231	1406120577	129	2,28	7
12	41231	1406120661	132	2,37	7
13	41231	1406122333	133	2,08	7

Tabel4. Data Akademik Mahasiswa Angkatan 2017.

NO	Kode Program Studi	NIM	SKS Total	IPK	Semester
1	54211	1706110039	23	2,15	1
2	54211	1706110053	23	2,45	1
3	54211	1706110199	23	2,47	1
4	54201	1706111300	23	0,52	1
5	54211	1706111469	23	2,4	1
6	54211	1706113534	23	1,76	1
7	54251	1706113761	23	2,5	1
8	54211	1706113899	23	2,32	1
9	54211	1706114189	23	2,5	1
10	41211	1706114360	23	0,78	1
11	54211	1706121775	23	2,46	1
12	54251	1706122128	23	0,85	1
13	54201	1706122256	23	2,5	1
14	54201	1706122514	23	2,37	1
15	41231	1706122543	23	2,28	1
16	41211	1706123003	23	1,09	1
17	54211	1706123213	23	2,24	1
18	54201	1706195298	24	2,4	1

Tabel2 .Data Akademik Mahasiswa Angkatan 2015.

NO	Kode Program Studi	NIM	SKS Total	IPK	Semester
1	41231	1506110776	96	2,47	5
2	54211	1506111917	96	2,3	5
3	54211	1506115433	96	2,28	5
4	54211	1506115758	95	2,05	5
5	41231	1506115915	92	2,45	5
6	54201	1506120376	94	2,50	5
7	54211	1506121197	99	2,39	5
8	54211	1506121882	101	2,28	5
9	54211	1506122428	102	2,46	5

Tabel3.Data Akademik Mahasiswa Angkatan 2016.

NO	Kode Program Studi	NIM	SKS Total	IPK	Semester
1	54211	1606110158	61	2,46	3
2	54211	1606110854	62	2,44	3
3	54201	1606111155	59	1,81	3
4	54201	1606111741	55	1,52	3
5	41231	1606112063	62	2,4	3
6	54201	1606112142	59	2,45	3
7	54211	1606112190	62	2,27	3
8	54201	1606114469	62	2,46	3
9	54251	1606116026	59	2,08	3
10	54211	1606121705	62	2,36	3
11	54251	1606122632	53	1,12	3
12	41231	1606123114	59	1,93	3
13	54201	1606123607	56	2,2	3
14	54211	1606123807	65	2,5	3
15	41211	1606136722	56	2,35	3
16	41211	1606136744	50	1,7	3
17	41211	1606136940	59	2,29	3
18	41211	1606136945	62	2,38	3
19	41211	1606136949	62	2,35	3
20	41211	1606136952	62	2,4	3
21	41211	1606136959	56	2,13	3
22	41211	1606137090	65	2,19	3
23	41211	1606137278	62	1,78	3
24	41231	1606197494	59	2,24	3

#### 4. Hasil dan Pembahasan

Pengujian hasil nilai Clustering K Means mahasiswa yang Potensial Drop Out dilakukan dengan menggunakan Software.

1. Windows 10 64 Byte.
2. Microsoft Office Excel 2016 digunakan sebagai pengolah bahan penelitian
3. *Rapid Miner Studio 8.0* merupakan tool untuk pengujian metode *Clustering K Means*.

Setelah dilakukan perhitungan secara manual dengan metode *Clustering K- Means*, maka dilakukan pengujian dengan melakukan langkah atau mekanisme pengujian sebagai berikut :

- a. Membuat suatu kasus uji yaitu pengumpulan data yang dibutuhkan dalam pengujian dengan perangkat lunak.
- b. Menentukan hasil yang diharapkan yaitu hasil klasifikasi dengan melakukan proses perhitungan secara manual.
- c. Melakukan evaluasi dilakukan dengan cara membandingkan hasil yang didapatkan pada tahap implementasi sistem dengan hasil yang dibuat secara manual.

##### 4.1 Hasil *Clustering K-Means*.

Hasil dari proses clustering K-Means dengan menggunakan persamaan *Euclidean Distanced* dengan jumlah Cluster 3 dan Centroid Awal  $C_0 = ( 114; 1,890; 7)$ ,  $C_1 = ( 140; 2,410; 7 )$ ,  $C_3 = ( 129; 2,280; 7 )$  sehingga didapatkan hasil seperti pada Tabel 5

Tabel 5. Iterasi 2 Angkatan 2014.

NIM	SKS Total	IPK	Semester	Jarak Terdekat
1406110606	132	2,50	7	cluster2
1406111097	116	2,10	7	cluster 0
1406111311	130	2,41	7	cluster 2
1406112089	114	1,89	7	cluster 0
1406112491	125	2,27	7	cluster 2
1406113572	128	2,29	7	cluster 2
1406114014	121	2,06	7	cluster 0
1406118121	108	2,05	7	cluster 0
1406119543	140	2,41	7	cluster 1
1406119994	130	2,50	7	cluster 2
1406120577	129	2,28	7	cluster 2
1406120661	132	2,37	7	cluster 2
1406122333	133	2,08	7	cluster 2

Dari tabel 5 dapat disimpulkan bahwa mahasiswa yang memiliki IPK Paling Kecil atau SKS Paling sedikit adalah mahasiswa yang potensial Drop Out, mahasiswa yang potensial drop out berada pada Cluster 0.

Hasil dari proses clustering K-Means dengan Menggunakan persamaan *Euclidean Distanced* dengan jumlah Cluster 3 dan Centroid Awal  $C_0 = ( 101; 2,28; 5 )$ ,  $C_1 = ( 96; 2,28; 5 )$ ,  $C_2 = ( 94; 2,45; 5 )$  sehingga didapatkan hasil seperti pada Tabel 6

Tabel 6. Tabel Iterasi 2 Angkatan 2015.

NIM	SKS Total	IPK	Semester	Jarak Terdekat
1506110776	96	2,47	5	Cluster 1
1506111917	96	2,30	5	Cluster 1
1506115433	96	2,28	5	Cluster 1
1506115758	95	2,05	5	Cluster 1
1506115915	92	2,45	5	Cluster 2
1506120376	94	2,50	5	Cluster 2
1506121197	99	2,39	5	Cluster 0
1506121882	101	2,28	5	Cluster 0
1506122428	102	2,46	5	Cluster 0

Dari Tabel 6 dapat disimpulkan bahwa mahasiswa yang memiliki IPK paling kecil atau SKS Paling Sedikit adalah mahasiswa yang potensial Drop Out, Mahasiswa yang potensial drop out berada pada Cluster 1 dan Cluster 2.

Hasil dari proses clustering K-Means dengan Menggunakan persamaan *Euclidean Distanced* dengan jumlah Cluster 3 dan Centroid Awal  $C_0 = ( 53,00; 1,12; 3 )$ ,  $C_1 = ( 56,00; 2,35; 3 )$ ,  $C_2 = ( 62,00; 2,38; 3 )$  sehingga didapatkan hasil seperti pada Tabel 7

Tabel 7. Tabel Iterasi 3 Angkatan 2016.

NIM	SKS Total	IPK	Semester	Jarak Terdekat
1606110158	61	2,46	3	Cluster 2
1606110854	62	2,44	3	Cluster 2
1606111155	59	1,81	3	Cluster 1
1606111741	55	1,52	3	Cluster 1

NIM	SKS Total	IPK	Semester	Jarak Terdekat
1606112063	62	2,40	3	Cluster 2
1606112142	59	2,45	3	Cluster 1
1606112190	62	2,27	3	Cluster 2
1606114469	62	2,46	3	Cluster 2
1606116026	59	2,08	3	Cluster 1
1606121705	62	2,36	3	Cluster 2
1606122632	53	1,12	3	Cluster 0
1606123114	59	1,93	3	Cluster 1
1606123607	56	2,20	3	Cluster 1
1606123807	65	2,50	3	Cluster 2
1606136722	56	2,35	3	Cluster 1
1606136744	50	1,70	3	Cluster 0
1606136940	59	2,29	3	Cluster 1
1606136945	62	2,38	3	Cluster 2
1606136949	62	2,35	3	Cluster 2
1606136952	62	2,40	3	Cluster 2
1606136959	56	2,13	3	Cluster 1
1606137090	65	2,19	3	Cluster 2
1606137278	62	1,78	3	Cluster 2
1606197494	59	2,24	3	Cluster 1

Dari Tabel 7 dapat disimpulkan bahwa mahasiswa yang memiliki IPK paling kecil atau SKS Paling Sedikit adalah mahasiswa yang potensial Drop Out, Mahasiswa yang potensial drop out berada pada Cluster 0 dan Cluster 1.

Hasil dari proses clustering K-Means dengan menggunakan persamaan *Euclidean Distanced* dengan jumlah Cluster 3 dan Centroid Awal  $C_0 = ( 24; 2,40; 1 )$ ,  $C_1 = ( 23,00; 2,32; 1 )$ ,  $C_2 = ( 23,00; 0,85; 1 )$  sehingga didapatkan hasil seperti pada Tabel 8

Tabel 8. Tabel Iterasi 2 Angkatan 2017.

NIM	SKS Total	IPK	Semester	Jarak Terdekat
1706110039	23	2,15	1	Cluster 1
1706110053	23	2,45	1	Cluster 1
1706110199	23	2,47	1	Cluster 1
1706111300	23	0,52	1	Cluster 2
1706111469	23	2,40	1	Cluster 1
1706113534	23	1,76	1	Cluster 1
1706113761	23	2,50	1	Cluster 1
1706113899	23	2,32	1	Cluster 1
1706114189	23	2,50	1	Cluster 1
1706114360	23	0,78	1	Cluster 2
1706121775	23	2,46	1	Cluster 1
1706122128	23	0,85	1	Cluster 2
1706122256	23	2,50	1	Cluster 1
1706122514	23	2,37	1	Cluster 1
1706122543	23	2,28	1	Cluster 1
1706123003	23	1,09	1	Cluster 2
1706123213	23	2,24	1	Cluster 1
1706195298	24	2,40	1	Cluster 0

Dari Tabel 8 dapat disimpulkan bahwa mahasiswa yang memiliki IPK paling kecil atau SKS Paling Sedikit

adalah mahasiswa yang potensial Drop Out, Mahasiswa yang potensial drop out berada pada Cluster 2.

## 5. Kesimpulan.

### 5.1 Simpulan

Berdasarkan hasil analisis dan pengujian yang dilakukan terhadap prestasi akademik mahasiswa berupa SKS Total, IPK dan Semester terhadap mahasiswa potensial drop out pada Fakultas Pertanian Universitas Riau, maka dapat disimpulkan bahwa

1. Dengan Menerapkan Metode *Clustering K-Means* dapat diketahui mahasiswa potensial drop out lebih dini, sehingga dapat dilakukan pembinaan untuk peningkatan prestasi akademik mahasiswa agar terhindar dari drop out.
2. Dengan mengelompokkan mahasiswa dengan beberapa cluster maka dapat diketahui bahwa hasil clustering angkatan 2014 berada pada cluster 0 berjumlah 4 orang mahasiswa atau 30,77 % dari 13 Sampel, angkatan 2015 berada pada cluster 1 berjumlah 4 mahasiswa dan cluster2 berjumlah 2 mahasiswa atau 66,7% dari 9 sampel, angkatan 2016 berada pada cluster0 berjumlah 2 mahasiswa dancluster 1 berjumlah 10 mahasiswa atau 50% dari 24 sampel, dan angkatan 2017 berada pada cluster 2 berjumlah 4 mahasiswa atau 22,22% dari 18 sampel dan angkatan 2017 terdapat 4 mahasiswa yang potensial drop out.
3. Dengan menggunakan Metode Clustering K- Mean dan software tool dengan menggunakan atribut SKS Total, IPK dan Semester dapat diketahui mahasiswa yang potensial drop out dapat diketahui lebih dini.

### 5.2 Saran

Diarankan agar dilakukan penelitian lebih lanjut terhadap mahasiswa potensial drop out berdasarkan

prestasi akademik dengan atribut SKS Total, IPK dan Semester, dan membandingkan hasil metode ini dengan metode hierarchical clustering, partitional clustering, single linkage, complete linkage, average linkage, DBSCAN, Fuzzy C-Means, Self-Organizing Map, K-Modes dan lain-lain.

## Daftar Rujukan

- [1] Xiaowei X., Dewen Z., Wenayo X., Yiyu S., dan Yu H., 2017 "An Efficient Memristor-based Distance Accelerator for Time Series Data Mining on Data Centers.", IEEE Explore No. 17241385.
- [2] Rui M., Ding L., Yue L., Lu L., Shujun X., Qian W dan Wenbin F, 2016" Discovery of Acupoints and Combinations for the Moxibustion Treatment of Cervical Radiculopathy: A Data Mining Analysis ", IEEE Explore, NO. 17027672
- [3] Han,J. and Kamber,M. "Data mining: Concepts and Techniques", 2nd edition. The Morgan Kaufmann series in Data Management System, Jim Grey, series Editor. 2006.
- [4] Sajadin S, Embong. A, Mohammad, M. A, Furqan.M, 2009 " Improving Student Akademik Performance Using Data Mining Techniques". Proceeding The 5th IMTGT International Conference on Mathematic, Statistics and Their Application (ICMSA 2009), ISBN 978-602-95343-0-6, Hal 390-394.
- [5] Getachew B. D., Tsegaye T., dan Yared B., 2017 "Data Mining Attribute Selection Approach For Drought Modelling: A Case Study For Greater Horn Of Africa", IJDKP, Volume 7, No 4 Hal 1-15
- [6] I.S. Dhillon, Y. Guan, and J. Kogan, 2002. "Iterative clustering of High Dimensional text data augmented by local search. In Proceedings of the 2002 IEEE International Conference on Data Mining Pages, 131-138.
- [7] Asroni., dan Ronald A., 2015, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang": Jurnal Ilmiah Semesta Teknika Volume 18 No 1: H 76-82
- [8] Govinda R.S, 2015 "Performance Validation of the Modified KMeans Clustering Algorithm Clusters Data " ,International Journal of Scientific & Engineering Research Volume 6 Issue 10 Hal 726-730