

Penerapan Algoritma K-Means Untuk Analisis Prestasi Akademik Mahasiswa Fakultas Ilmu Komputer Universitas Klabat

Application of K-Means Algorithm for Academic Achievement Analysis of Faculty of Computer Science Universitas Klabat

¹Green F Mandias, ²Green A Sandag, ³Susi Susanti, dan ⁴Haryanto Reza Musak

^{1,2,3}Program Studi Sistem Informasi, Universitas Klabat, Airmadidi

e-mail: ¹green@unklab.ac.id, ²greensandag@unklab.ac.id,

³s21410127@student.unklab.ac.id, ⁴s21410079@student.unklab.ac.id

Abstrak

Universitas Klabat (UNKLAB) adalah salah satu perguruan tinggi swasta yang berada dibawah naungan organisasi Gereja Masehi Advent Hari Ketujuh, yang bertempat di Airmadidi, Sulawesi Utara. Universitas Klabat termasuk universitas yang sangat dikenal di Sulawesi utara, yang di dalamnya memiliki 1 program pascasarjana, 6 fakultas dan 1 akademik. Penelitian ini dilakukan untuk mengetahui pencapaian prestasi mahasiswa fakultas ilmu komputer yang berada pada tingkat 4 yang memiliki 52 mahasiswa yang aktif dengan memanfaatkan metode data mining. Berdasarkan data mahasiswa fakultas ilmu computer, penelitian ini dilakukan untuk mencari tahu berapa banyak mahasiswa yang memiliki prestasi akademik dibidang databases, networking dan programming dengan menggunakan algoritma data mining. Penelitian ini menggunakan algoritma K-Means dalam menganalisis prestasi akademik mahasiswa fakultas llmu komputer di Universitas Klabat. Data yang dianalisis dikelompokan terlebih dahulu agar terstruktur serta data yang dianalisis memiliki kejelasan hasil yang lebih dalam. Hasil yang didapat dari 52 mahasiswa tersebut adalah, 33% mahasiswa memiliki nilai prestasi di bidang database, 42% mahasiswa pada bidang networking dan 25% mahasiswa di bidang programming.

Keywords : UNKLAB, Algoritma K-Means ,WEKA, Cluster, Data Mining.

Abstract

Universitas Klabat (UNKLAB) is one of the private universities under the auspices of the Seventh-day Adventist Church organization, located in Airmadidi, North Sulawesi. The University of Klabat includes a well-known university in northern Sulawesi, where it has 1 postgraduate program, 6 faculty and 1 academic. This research was conducted to find out the achievement of computer science faculty students who are at level 4 which has 52 active students by utilizing data mining method. Based on data from computer science faculty students, this research is done to find out how many students have academic achievement in databases, networking and programming using data mining algorithm. This research uses K-Means algorithm in analyzing academic achievement of computer science faculty student at Universitas Klabat. The data analyzed are grouped first so that the structured and the analyzed data have clearer result. The result is 52% of students have achievement score in the database field, 42% of students in networking and 25% of students in the field programming.

Keywords : UNKLAB, K-Means Algorithm ,WEKA, Cluster, Data Mining.

1. PENDAHULUAN

Universitas Klabat (UNKLAB) adalah salah satu perguruan tinggi swasta yang berada dibawah naungan organisasi Gereja Masehi Advent Hari Ketujuh, yang bertempat di Airmadidi, Minahasa Utara, Sulawesi Utara. Universitas Klabat termasuk universitas yang tergolong dalam universitas swasta, yang di dalamnya memiliki 1 program pascasarjana, 6 fakultas dan 1 akademik yaitu: Fakultas Ilmu Komputer, Fakultas Ilmu Keperawatan, Fakultas Ekonomi, Fakultas Keguruan dan Ilmu Pendidikan, Fakultas Filsafat, Fakultas Pertanian dan Akademik Sekertaris.

Fakultas Ilmu Komputer adalah salah satu fakultas yang memiliki 408 Mahasiswa yang aktif di Fakultas Ilmu Komputer. Dengan begitu banyak mahasiswa yang ada saat ini, sehingga peneliti memanfaatkan data mahasiswa untuk dapat mengetahui mahasiswa yang memiliki prestasi dimata kuliah yang telah dikontrak dan telah lulus pada matakuliah yang sudah diambil. Untuk dapat membantu Fakultas Ilmu Komputer dalam memberikan informasi dibidang akademik khususnya mata kuliah *Database, Networking, dan Programming*. Penelitian yang akan peneliti terapkan merupakan suatu dasar pengetahuan teknologi dibidang data mining sehingga dapat menerapkan algoritma *K-means* untuk menganalisis prestasi akademik mahasiswa khususnya mahasiswa Fakultas Ilmu Komputer dan mengetahui dimana mahasiswa tersebut memiliki pengetahuan, keterampilan dan potensi selama masa perkuliahan di Universitas Klabat.

Sesuai dengan visi dan misi dari Fakultas Ilmu Komputer yang memiliki indeks prestasi di masing-masing mata kuliah yang dapat memberikan mutu dan kualitas dari mahasiswa tersebut yang bisa menjadi tolak ukur bagi Fakultas Ilmu Komputer saat ini. Untuk menunjang kinerja yang dimiliki oleh dosen-dosen dengan menerapkan ilmu yang bisa membantu bagi mahasiswa untuk diberikan kegiatan yang bisa menunjang akan perkembangan di bidang mata kuliah yang banyak diminati oleh mahasiswa agar dapat memberikan bimbingan terhadap matakuliah yang dijadikan pusat pengetahuan yang dimiliki. Hal ini menjadi suatu tolak ukur bagi fakultas dalam memberikan materi, pelajaran dan bimbingan serta memberikan pelajaran tambahan. Berupa seminar, kegiatan-kegiatan yang menyangkut mata kuliah yang dapat memberikan mutu yang baik dan bermanfaat dalam menerapkan ilmu yang di dapat dan bisa menghasilkan mahasiswa yang memiliki indeks prestasi.

K-means merupakan salah satu metode data *cluster* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster/* kelompok [1]. Salah satu sebab fakultas tidak mengetahui mahasiswa yang memiliki nilai yang tinggi dan kemampuan dalam mata kuliah untuk dapat diberikan rekomendasi untuk mengikuti perlombaan di bidang mata kuliah *Database, Networking, dan Programming*. Oleh karena itu akan di ukur menggunakan Algoritma *K-Means* dengan cara penerapan algoritma. Agar mengetahui prestasi yang di miliki mahasiswa tersebut, bagi Fakultas Ilmu Komputer agar terwujudnya visi dan misi dari Fakultas Ilmu Komputer. Fungsi lain dari penerapan ini agar menjadi satu penghargaan bagi mahasiswa yang memiliki prestasi di bidang akademik maupun non akademik yang baik, agar dapat membantu fakultas dalam menilai pencapaian mahasiswa dalam pengetahuan yang didapat. Dengan penerapan ini, maka di butuhkan nilai dari keahlian *Database, Networking, Programming* yang sudah dikontrak oleh mahasiswa, agar bisa membantu peneliti dalam melakukan analisis kemampuan yang di miliki mahasiswa Fakultas Ilmu Komputer dalam mencapai prestasi.

2. METODE PENELITIAN

Metode penelitian yang digunakan pada penelitian memperhatikan dengan teliti tahap-tahap yang dilalui dan teknik pengumpulan data yang digunakan dalam mengumpulkan data untuk analisis Prestasi Akademik Mahasiswa Fakultas Ilmu Komputer di Universitas Klabat dengan wawancara dan pengambilan data secara langsung dengan mahasiswa terkait.

2.1 Desain Penelitian

Dalam penelitian ini peneliti menggunakan Algoritma *K-Means* dalam menganalisis Prestasi Akademik Mahasiswa Fakultas Ilmu Komputer di Universitas Klabat. Data yang dianalisis akan di kelompokkan supaya bisa lebih terstruktur serta data yang akan dianalisis bisa lebih jelas.

2.2 Instrumen Penelitian

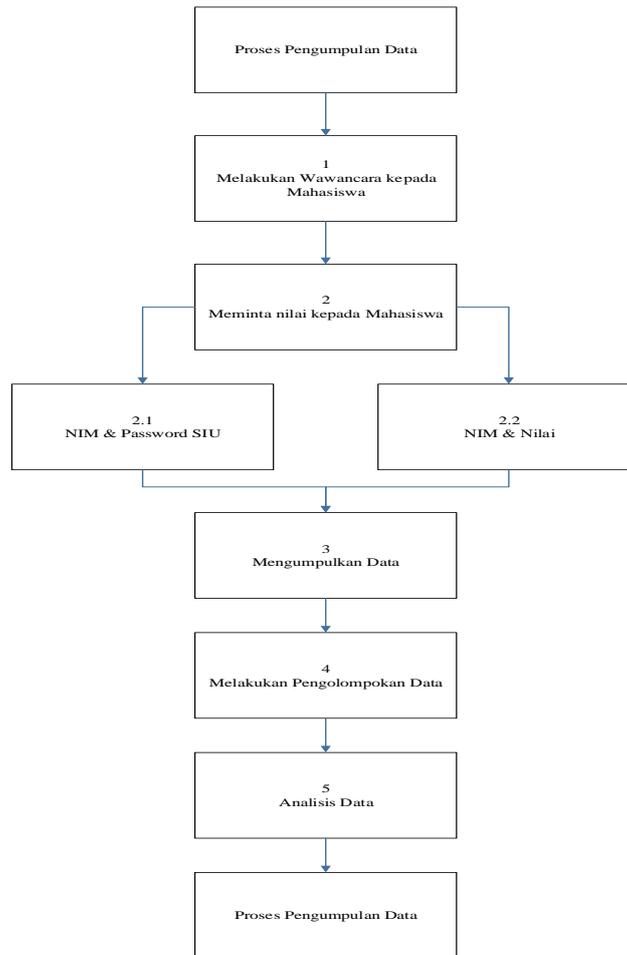
Instrumen yang akan peneliti gunakan untuk mendapatkan data yang diperlukan adalah:

1. Data: data yang diperoleh dan digunakan oleh peneliti adalah data primer dan data sekunder. Data primer yang dikumpulkan berupa hasil wawancara peneliti pada pihak mahasiswa dan IT *staff* yang ada di Universitas Klabat. Data sekunder berupa data-data dari mahasiswa Fakultas Ilmu Komputer yang diperlihatkan.
2. Teknik Pengumpulan Data: dalam pengumpulan data, teknik yang digunakan peneliti yaitu mengumpulkan data dari proses wawancara dan pengambilan data secara langsung dengan pihak yang terkait yaitu mahasiswa.
3. Prosedur Pengumpulan Data

Dalam penelitian ini, langkah-langkah yang dilakukan peneliti adalah sebagai berikut:

1. Peneliti melakukan wawancara terlebih dahulu kepada mahasiswa FIK.
2. Peneliti meminta nilai kepada mahasiswa yang sudah mengambil mata kuliah.
3. Peneliti melakukan pengumpulan data dari nilai mahasiswa yang diminta.
4. Peneliti melakukan pengelompokan dari data yang sudah dikumpulkan berdasarkan mata kuliah.
5. Peneliti melakukan analisis dari data yang sudah di kelompokkan terlebih dahulu.
6. Dari hasil analisis maka akan di dapatkan hasil akhir.

Gambar 1 menunjukkan proses dalam pengumpulan data mahasiswa yang digunakan sebagai dasar dalam pembentukan *cluster* pada penelitian ini.



Gambar 1 Proses Pengumpulan Data

3. HASIL DAN PEMBAHASAN

Data dibagi menjadi beberapa bagian sebelum diolah dengan *tool* Weka. Pembagian ini perlu dilakukan untuk mendapatkan hasil analisis dan susunan yang tepat sehingga menghasilkan *output* yang diharapkan.

3.1 Hasil Weka

Peneliti menggunakan data mahasiswa untuk melaksanakan penelitian yang sesuai prosedur untuk mengetahui prestasi sesuai dengan data, dimana data ini adalah data mahasiswa yang sudah mengambil mata kuliah tersebut.

```
Number of iterations: 11
Within cluster sum of squared errors: 2.781454442025083
Missing values globally replaced with mean/mode
```

Gambar 2 Iteration Data

Pada Gambar 2 menunjukkan saat *attribute* yang sudah dimasukkan kemudian menjalankannya di *software* Weka hasilnya akan langsung keluar. Pada gambar ini merupakan hasil bahwa jumlah iterasi yang didapat dalam melakukan *clustering* nilai mahasiswa sebanyak 11 iterasi.

Attribute	Full Data (52)	Cluster#		
		0 (17)	1 (22)	2 (13)
Database	83.8654	88.5294	85.0909	75.6923
Networking	81.3462	89.9412	79.4091	73.3846
Programming	82.5769	89.3529	81.4091	75.6923

Gambar 3 Hasil *Clustering*

Gambar 3 merupakan hasil *cluster* dengan menggunakan algoritma K-means di *software* Weka. Dari hasil ada tiga *attribute* yaitu Database, Networking, Programming. Dimana rata-rata dari setiap keahlian ini sudah ditampilkan dalam *full data* dengan 52 mahasiswa, pada gambar ini ada 3 *cluster* yaitu *cluster 0*, *cluster 1*, *cluster 2*. Dari hasil yang didapatkan, ditetapkan bahwa dalam *cluster 0* merupakan kumpulan dari mahasiswa yang berjumlah 17 orang dengan nilai keahlian yang paling tinggi, untuk *cluster 1* dengan mahasiswa 22 orang merupakan kumpulan mahasiswa yang mempunyai nilai tengah, sedangkan untuk *cluster 2* merupakan kumpulan dari mahasiswa dengan nilai dari keahlian yang rendah. Jika dilihat dari *cluster 0* maka dari tiga keahlian tersebut bisa dipastikan bahwa keahlian Networking merupakan keahlian dengan nilai paling tinggi.

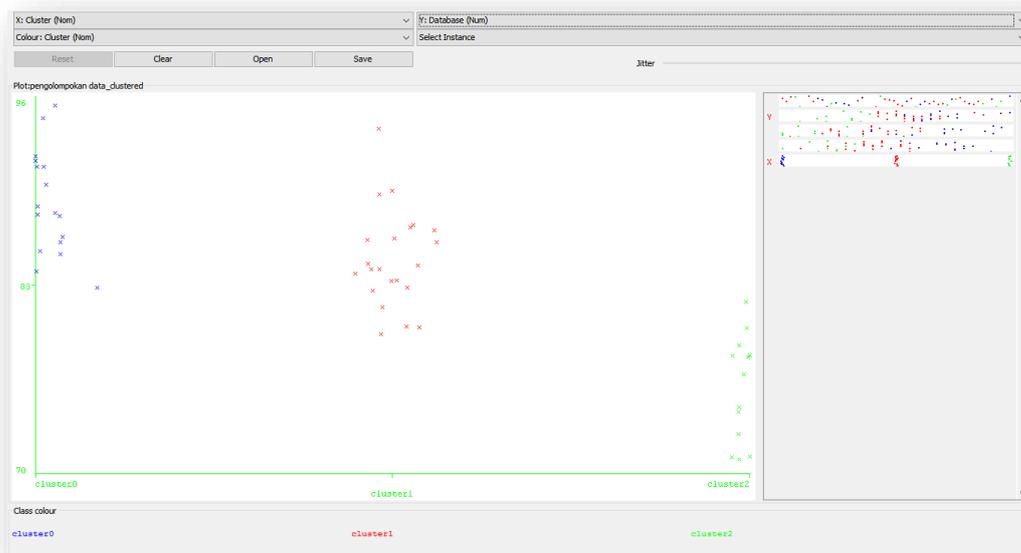
```
Time taken to build model (full training data) : 0.02 seconds
=== Model and evaluation on training set ===

Clustered Instances

0      17 ( 33%)
1      22 ( 42%)
2      13 ( 25%)
```

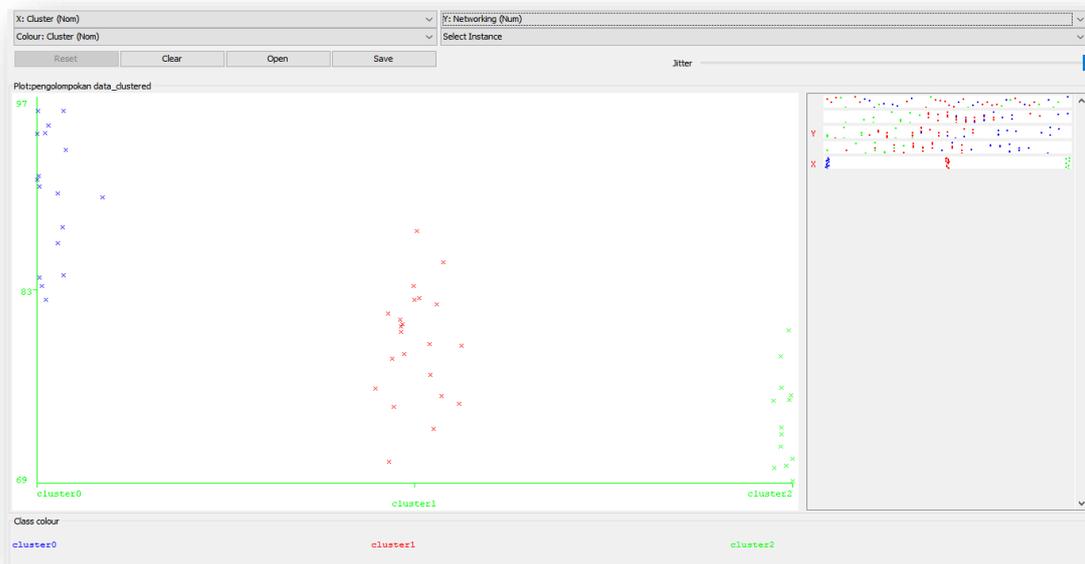
Gambar 4 *Clustered Instances*

Pada Gambar 4 merupakan jumlah mahasiswa yang tergabung dalam *cluster0*, *cluster1*, dan *cluster2*. Dari jumlah mahasiswa yang tergabung kemudian dapat dilihat berapa persen dari mahasiswa yang ada di tiap *cluster*.



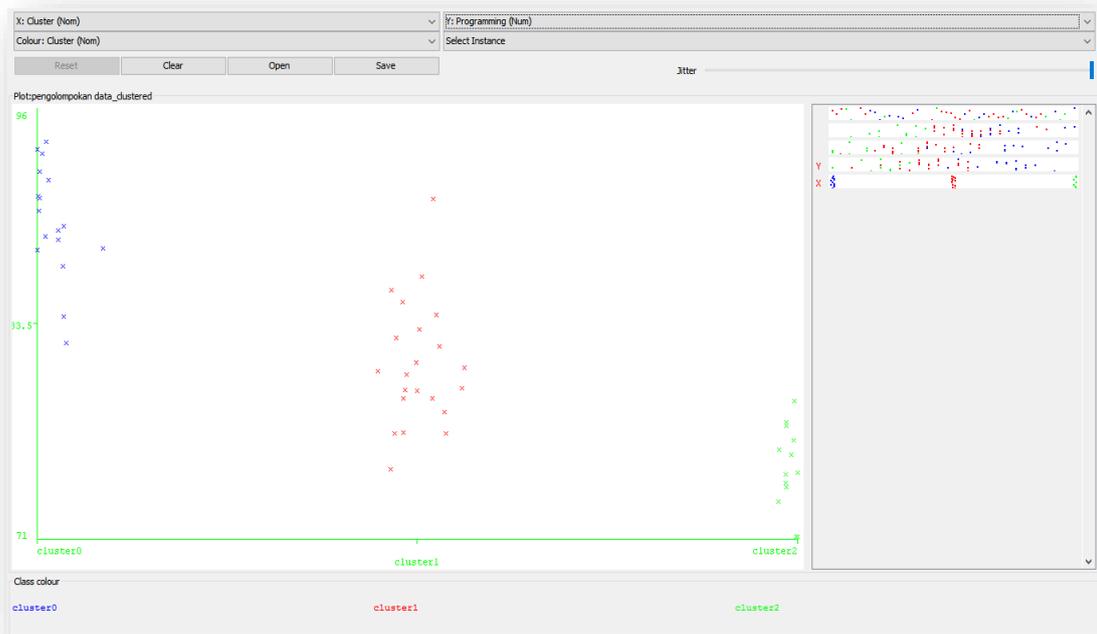
Gambar 5 Clustering Databases

Pada Gambar 5 menunjukkan *cluster* dari mahasiswa dengan keahlian Database. Pada gambar ini mahasiswa yang mempunyai nilai-nilai tertinggi ada pada *cluster* 0, dimana mahasiswa lebih dominan hampir ke nilai tengah.



Gambar 6 Clustering Networking

Pada Gambar 6 merupakan *cluster* mahasiswa terhadap keahlian Networking. Dapat dilihat bahwa mahasiswa yang mempunyai nilai tertinggi masih dalam *cluster* 0, dimana mahasiswa lebih dominan ke nilai tertinggi dan sebagian juga ke nilai tengah.



Gambar 7 Clustering Programming

Pada Gambar 7 mahasiswa yang mempunyai nilai tertinggi di bidang keahlian Programming lebih dominan sebagian ke nilai tertinggi dan sebagian nilai tengah.



Gambar 8 Visualisasi 3 Cluster

Pada Gambar 8 merupakan hasil dari *clustering*, dimana dalam grafik tersebut memperlihatkan bahwa ada berapa jumlah mahasiswa yang termasuk dalam nilai *cluster 0*, *cluster 1*, *cluster 2*.

Tabel 1 Pengelompokan Data *Cluster*

Relation: pengelompokan data_clustered					
No.	Instance_number Numeric	Database Numeric	Networking Numeric	Programming Numeric	Cluster Nominal
6	5.0	89.0	89.0	88.0	cluster0
9	8.0	95.0	86.0	88.0	cluster0
10	9.0	84.0	91.0	89.0	cluster0
11	10.0	86.0	84.0	94.0	cluster0
13	12.0	90.0	89.0	86.0	cluster0
14	13.0	83.0	97.0	90.0	cluster0
15	14.0	87.0	89.0	91.0	cluster0
16	15.0	86.0	85.0	90.0	cluster0
17	16.0	94.0	97.0	96.0	cluster0
19	18.0	96.0	96.0	90.0	cluster0
29	28.0	89.0	94.0	91.0	cluster0
32	31.0	87.0	83.0	90.0	cluster0
33	32.0	89.0	80.0	92.0	cluster0
39	38.0	90.0	90.0	89.0	cluster0
43	42.0	85.0	95.0	83.0	cluster0
48	47.0	87.0	89.0	88.0	cluster0
52	51.0	88.0	95.0	84.0	cluster0

Relation: pengelompokan data_clustered					
No.	Instance_number Numeric	Database Numeric	Networking Numeric	Programming Numeric	Cluster Nominal
1	0.0	81.0	86.0	78.0	cluster 1
2	1.0	86.0	79.0	81.0	cluster 1
3	2.0	81.0	80.0	80.0	cluster 1
7	6.0	83.0	80.0	82.0	cluster 1
8	7.0	81.0	82.0	81.0	cluster 1
18	17.0	83.0	83.0	81.0	cluster 1
20	19.0	89.0	84.0	84.0	cluster 1
22	21.0	84.0	70.0	85.0	cluster 1
24	23.0	83.0	75.0	80.0	cluster 1
25	24.0	93.0	80.0	84.0	cluster 1
26	25.0	87.0	75.0	80.0	cluster 1
27	26.0	86.0	82.0	81.0	cluster 1
28	27.0	85.0	76.0	79.0	cluster 1
30	29.0	82.0	81.0	85.0	cluster 1
34	33.0	88.0	74.0	85.0	cluster 1
35	34.0	84.0	76.0	84.0	cluster 1
36	35.0	82.0	76.0	90.0	cluster 1
37	36.0	84.0	79.0	73.0	cluster 1
41	40.0	92.0	85.0	82.0	cluster 1
42	41.0	85.0	86.0	82.0	cluster 1
44	43.0	85.0	79.0	78.0	cluster 1
45	44.0	88.0	79.0	76.0	cluster 1

4	3.0	79.0	77.0	74.0	duster2
5	4.0	75.0	79.0	75.0	duster2
12	11.0	78.0	69.0	71.0	duster2
21	20.0	70.0	72.0	76.0	duster2
23	22.0	77.0	71.0	76.0	duster2
31	30.0	72.0	77.0	75.0	duster2
38	37.0	78.0	69.0	77.0	duster2
40	39.0	80.0	75.0	79.0	duster2
46	45.0	74.0	73.0	76.0	duster2
47	46.0	74.0	71.0	75.0	duster2
49	48.0	79.0	71.0	77.0	duster2

Pada tabel 1 merupakan hasil akhir dari melakukan *cluster* dengan Weka, dimana dari hasil ini dapat dilihat bahwa mahasiswa mana saja yang tergabung dalam tiap *cluster*.

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan, dapat dipastikan bahwa mahasiswa mana saja yang termasuk dalam *cluster 0*, *cluster 1*, *cluster 2*. Hasilnya sebagai berikut:

1. Mahasiswa yang memiliki prestasi di bidang matakuliah yang dilakukan dengan beberapa *attribute* mata kuliah dengan sampel 52 mahasiswa yang sudah mengambil matakuliah yaitu; Mahasiswa yang termasuk dalam nilai tertinggi atau *cluster 0*, ada sebanyak 17 mahasiswa dari 52 mahasiswa (33%), Mahasiswa yang termasuk dalam nilai tengah atau *cluster 1*, sebanyak 22 mahasiswa dari 52 mahasiswa (42%), Mahasiswa yang termasuk dalam nilai rendah atau *cluster 2*, sebanyak 13 mahasiswa dari 52 mahasiswa (25%)
2. Dari data mahasiswa yang peneliti gunakan ada banyak mahasiswa yang termasuk mahasiswa yang memiliki prestasi dibidang mata kuliah tertentu dan dapat dikelompokkan dengan menggunakan algoritma k-means yang peneliti gunakan dan memiliki hasil yang tepat untuk menjadi sumber informasi bagi fakultas ilmu komputer.

5. SARAN

Setelah melakukan pengumpulan data mahasiswa dan dianalisis prestasi mahasiswa dengan menggunakan algoritma yang peneliti gunakan saat ini di Fakultas Ilmu Komputer, diharapkan kepada Universitas Klabat dan terlebih khusus Fakultas Ilmu Komputer agar dapat mempertahankan hasil prestasi dan mengembangkan mahasiswa yang memiliki prestasi dan memberikan perhatian kepada mahasiswa yang memiliki nilai mata kuliah yang di bawah rata-rata untuk dapat membimbing agar bisa mencapai prestasi dibidang matakuliah, agar meningkatkan mutu Universitas Klabat khususnya Fakultas Ilmu Komputer. Saran juga untuk penelitian selanjutnya agar supaya dapat melakukan penelitian yang sama tetapi dengan jumlah atau sampel mahasiswa yang lebih banyak lagi, agar supaya hasil yang didapatkan bisa lebih baik.

DAFTAR PUSTAKA

- [1] G. Heribertus, "Penerapan Algoritma *Clustering* K-means, K-medoid, Gath Geva," Universitas Gajah Mada, Yogyakarta.
- [2] E. Untari, "Pengaruh Kreativitas dan Minat Belajar Mahasiswa Terhadap Prestasi Belajar Mahasiswa," vol. XVIII, 2016.
- [3] Sumarno, "Prestasi Belajar Ditinjau dari Sikap Belajar Mahasiswa dan Cara Dosen Mengajar Dalam Mata Kuliah Kewirausahaan Program Studi Pendidikan Akuntansi FKIP UMS," 2012.

- [4] D. E. Turban, *Decision Support System and Intelligent Systems*, Yogyakarta: Andi Offset, 2005.
- [5] L. T. E. Kusriani, *Algoritma dan Data Mining*, Yogyakarta: Andi, 2009.
- [6] U. Fayyad, *Advanced in Knowledge Discovery and Data Mining*, Yogyakarta: Andi, 1996.
- [7] P. Berkhin, "A Survey of *Clustering* Data Mining Techniques. In: *Grouping Multidimensional Data*," 2006, pp. 25-71.
- [8] R. A. Asroni, "Penerapan Metode K-means Untuk *Clustering* Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus pada Jurusan Teknik Informatika UMM Magelang," vol. 18, pp. 76-82, 2015.
- [9] Y. Agusta, "K-means Penerapan, Permasalahan dan Metode Terkait," *Jurnal Sistem dan Informatika*, pp. 47-60, 2007.
- [10] "Game Analytics," Mei 2014. [Online]. Available: <http://www.gameanalytics.com/blog/introducing-clustering-behavioral-profiling-game-analytics.html>. [Accessed Rabu Juni 2017].
- [11] S. Jiménez, "Text Classification and Klustering with WEKA: WEKA A guided example," Universidad Nacional De Colombia.
- [12] D. Sondakh, "Data Mining for Healthcare Data: A Comparison of Neural Networks Algorithms," *Cogito Smart Journal*. vol. 1, pp. 10-19, 2017