

Perancangan Prediksi Prestasi Nilai Akademik Mahasiswa Menggunakan Metode K-Means Clustering

Abimanyu Agung Saputro¹, Risa Helilintar²

^{1,2}Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

E-mail: ¹ab1m4nyu45@gmail.com, ²risa.helilintar@gmail.com

Abstrak – *Prediksi Prestasi Nilai Akademik Mahasiswa Menggunakan Metode K-Means Clustering.* Penelitian ini dilatar belakangi, dosen wali belum mempunyai suatu indikator yang dapat digunakan untuk memprediksi prestasi akademik mahasiswa. Prediksi prestasi mahasiswa dapat digunakan untuk memberikan gambaran kepada mahasiswa tentang hasil akhir prestasi akademik yang akan diperolehnya. Prediksi prestasi akademik juga dapat digunakan sebagai peringatan awal kepada mahasiswa agar mahasiswa menjadi lulusan yang berkualitas yang mempunyai nilai prestasi akademik sesuai dengan standar. Pada penelitian ini, peneliti mencoba untuk menerapkan data mining dengan menggunakan metode clustering pada pengelompokan mahasiswa berdasarkan prestasi akademik yang diperoleh, pengelompokan ini diharapkan menjadi suatu indikator bagi Dosen wali untuk memberikan bimbingan kepada mahasiswa. Prediksi prestasi akademik juga dapat digunakan sebagai peringatan awal kepada mahasiswa agar mahasiswa menjadi lulusan yang berkualitas yang mempunyai nilai prestasi akademik sesuai dengan standar, Berdasarkan penelitian tersebut dapat disimpulkan bahwa penelitian ini akan menghasilkan sebuah sistem yang memudahkan admin melakukan prediksi nilai berdasarkan kemampuan mahasiswa. Dengan penggunaan Algoritma K-Means Clustering hasil perhitungan antara cluster cukup baik, Dan dari penelitian ini diharapkan dapat dijadikan acuan untuk meningkatkan prestasi dari nilai akademik masing-masing mahasiswa.

Kata Kunci — *K-means Clustering, Prediksi, Data mining*

1. PENDAHULUAN

Data mining merupakan analisis dari peninjauan kumpulan data untuk menentukan hubungan yang tidak diduga dan meringkas data dengan cara berbeda sebelumnya, yang dapat dipahami dan bermanfaat bagi pemilik data. Pada saat ini data mining sering digunakan sebagai alat pendukung dalam pengambilan keputusan, tentang contoh suatu kasus penentuan nasabah yang aman atau nasabah yang beresiko dalam pengajuan kredit pada suatu bank dengan menggunakan metode clustering yang merupakan salah satu dari enam metode yang terdapat pada data mining. Sedangkan pada dunia pendidikan terutama perguruan tinggi, data mining dapat digunakan untuk membantu perguruan tinggi dalam meningkatkan kualitas lulusan yang dihasilkan. Seperti halnya contoh kasus penentuan nasabah yang aman dan nasabah yang beresiko, metode pada contoh kasus tersebut dapat diterapkan dalam menganalisa dan memprediksi mahasiswa yang akan berprestasi dan yang kurang berprestasi.

Data mining merupakan suatu metode yang digunakan untuk mendapatkan suatu pengetahuan dari sejumlah data berukuran besar yang dapat digunakan untuk suatu kebutuhan tertentu. Data mining merupakan salah satu proses dalam

Knowledge Discovery from Data (KDD). KDD terdiri dari proses sebagai berikut [1][2]:

- 1) Data Cleaning (membersihkan data yang tidak valid/ lengkap).
- 2) Data Integration (proses integrasi data).
- 3) Data Selection (proses pemilihan data yang layak untuk dianalisis).
- 4) Data Transformation (proses transformasi data supaya lebih mudah dalam melakukan data mining).
- 5) Data Mining (proses penggalian informasi dengan algoritma-algoritma yang menunjang sesuai kebutuhan).
- 6) Pattern Evaluation (proses evaluasi hasil penggalian informasi).
- 7) Knowledge Presentation (proses visualisasi dan penyampaian dari hasil penggalian informasi).

Clustering atau klasifikasi adalah metode yang digunakan untuk rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. Cluster adalah sekelompok atau sekumpulan objek-objek data yang similar satu sama lain dalam cluster yang sama dan disimilar terhadap objek-objek yang berbeda cluster. Objek akan di kelompokkan ke dalam satu atau lebih cluster sehingga objek-objek yang berada dalam satu cluster

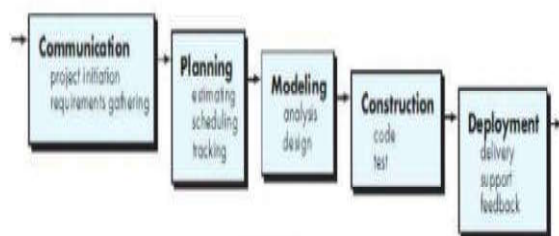
akan mempunyai kesamaan yang tinggi antara satu dengan yang lainnya [3].

Namun dalam melakukan pengawasannya, dosen wali belum mempunyai suatu indikator yang dapat digunakan untuk memprediksi prestasi akademik mahasiswa. Prediksi prestasi mahasiswa dapat digunakan untuk memberikan gambaran kepada mahasiswa tentang hasil akhir prestasi akademik yang akan diperolehnya. Prediksi prestasi akademik juga dapat digunakan sebagai peringatan awal kepada mahasiswa agar mahasiswa menjadi lulusan yang berkualitas yang mempunyai nilai prestasi akademik sesuai dengan standar.

Pada penelitian ini, peneliti mencoba untuk menerapkan data mining dengan menggunakan metode clustering pada pengelompokan mahasiswa berdasarkan prestasi akademik yang diperoleh, pengelompokan ini diharapkan menjadi suatu indikator bagi Dosen wali untuk memberikan bimbingan kepada mahasiswa.

2. METODE PENELITIAN

Model waterfall adalah model klasik yang bersifat sistematis, berurutan dalam membangun software. Nama model ini sebenarnya adalah "Linear Sequential Model". Model ini sering disebut juga dengan "classic life cycle" atau metode waterfall. Model ini termasuk ke dalam model generic pada rekayasa perangkat lunak dan pertama kali diperkenalkan oleh Winston Royce sekitar tahun 1970 sehingga sering dianggap kuno, tetapi merupakan model yang paling banyak dipakai dalam Software Engineering (SE). Model ini melakukan pendekatan secara sistematis dan berurutan. Disebut dengan waterfall karena tahap demi tahap yang dilalui harus menunggu selesainya tahap sebelumnya dan berjalan berurutan. Fase-fase dalam [4]:



Gambar 1. Model Metode Waterfall

2.1 Communication (Project Initiation & Requirements Gathering)

Sebelum memulai pekerjaan yang bersifat teknis, sangat diperlukan adanya komunikasi dengan

customer demi memahami dan mencapai tujuan yang ingin dicapai. Hasil dari komunikasi tersebut adalah inisialisasi proyek, seperti menganalisis permasalahan yang dihadapi dan mengumpulkan data-data yang diperlukan, serta membantu mendefinisikan fitur dan fungsi software. Pengumpulan data-data tambahan bisa juga diambil dari jurnal, artikel, dan internet.

2.2 Planning (Estimating, Scheduling, Tracking)

Tahap berikutnya adalah tahapan perencanaan yang menjelaskan tentang estimasi tugas-tugas teknis yang akan dilakukan, resiko-resiko yang dapat terjadi, sumber daya yang diperlukan dalam membuat sistem, produk kerja yang ingin dihasilkan, penjadwalan kerja yang akan dilaksanakan, dan tracking proses pengerjaan sistem.

2.3 Modeling (Analysis & Design)

Tahapan ini adalah tahap perancangan dan permodelan arsitektur sistem yang berfokus pada perancangan struktur data, arsitektur software, tampilan interface, dan algoritma program. Tujuannya untuk lebih memahami gambaran besar dari apa yang akan dikerjakan.

2.4 Construction (Code & Test)

Tahapan Construction ini merupakan proses penerjemahan bentuk desain menjadi kode atau bentuk/bahasa yang dapat dibaca oleh mesin. Setelah pengkodean selesai, dilakukan pengujian terhadap sistem dan juga kode yang sudah dibuat. Tujuannya untuk menemukan kesalahan yang mungkin terjadi untuk nantinya diperbaiki.

2.5 Deployment (Delivery, Support, Feedback)

Tahapan Deployment merupakan tahapan implementasi software ke customer, pemeliharaan software secara berkala, perbaikan software, evaluasi software, dan pengembangan software berdasarkan umpan balik yang diberikan agar sistem dapat tetap berjalan dan berkembang sesuai dengan fungsinya [4].

2.6 K-Means

K-Means adalah salah satu metode dalam fungsi clustering atau pengelompokan. Clustering mengacu pada pengelompokan data, observasi atau kasus berdasar kemiripan objek yang diteliti. Sebuah cluster adalah suatu kumpulan data yang mirip dengan lainnya atau tidakmiripan data pada kelompok lain [5]. Sedangkan (Xu & Wunsch II, 2009) menjelaskan bahwa clustering adalah membagi objek data (bentuk, entitas, contoh,

ketaatan, unit) ke dalam beberapa jumlah kelompok (grup, bagian atau kategori)[6]. Tujuan proses clustering adalah meminimalkan terjadinya objective function yang diset dalam proses clustering, yang pada umumnya digunakan untuk meminimalisasikan variasi dalam suatu cluster dan memaksimalkan variasi antar cluster [5].

3. HASIL DAN PEMBAHASAN

3.1. Simulasi Algoritma

Dalam Algoritma K-means ini digunakan dalam menentukan pengelompokan data predikat prestasi mahasiswa Berikut adalah perhitungan manual algoritma k-means:

Kriteria nilai yang digunakan di Prodi Teknik Informatika adalah sebagai berikut:

- a. Nilai
- b. Sks
- c. Mata

1) Contoh Dataset

Tabel 2.3 merupakan tabel dataset dari 4 mahasiswa yang memprogramkan mata kuliah masing-masing. Dari 4 mahasiswa tersebut akan dikelompokkan menjadi 4 bagian yaitu kelompok sangat baik, baik, cukup dan kurang.

Tabel 1. Contoh Dataset

Nama	Alpro II	Prak. AlproII	Aljabar Linier	Pemrograman Visual	Arsikom	Basis Data I
Aan Agus	9	2.5	8	6	10	10.5
Aan Saputra	12	3.5	14	8	14	10.5
Aan Suhartiri	7.5	3	12	8	12	9
Achmad Faisal	7.5	3	14	7	14	10.5
Achmad Fauzi	12	4	12	7	12	9

2) Setelah menentukan dataset, maka perlu menentukan jumlah cluster yang akan dibentuk. Adapun cluster yang akan dibentuk antara lain :

- a.Cluster 1 (C1) = Sangat Baik
- b.Cluster 2 (C2) = Baik
- c.Cluster 3 (C3) = Cukup
- d.Cluster 4 (C4) =Kurang

- 3) Tetapkan C pusat cluster awal secara random
- 4) Alokasikan semua data/obyek ke dalam cluster terdekat. Berikut hasil dari alokasi data ke jarak cluster.

Tabel 2. Contoh Jarak Cluster

Keterangan	kategori	Alpro	Prak.Alpro	Aljabar Linier	Pemrograman Visual	Arsikom	Basis Data 1	Prak. Basis data 1
Cluster 1	sangat baik	12	4	14	8	14	10.5	4
Cluster 2	baik	9	3.5	12	7	12	9	3
Cluster 3	cukup	7.5	3	8	6	10	7.5	2
Cluster 4	kurang	6	2	4	5	8	6	1

$$d(1,1)=\sqrt{(12-9)^2+(4-2.5)^2+(14-8)^2+(8-6)^2+(14-10)^2+(10.5-10.5)^2+(4-3)^2}$$

$$= 8.3$$

$$d(1,2)=\sqrt{(12-12)^2+(4-3.5)^2+(14-14)^2+(8-8)^2+(14-14)^2+(10.5-10.5)^2+(4-4)^2}$$

$$= 0.5$$

$$d(1,3)=\sqrt{(12-7.5)^2+(4-3)^2+(14-12)^2+(8-8)^2+(14-12)^2+(10.5-9)^2+(4-3)^2}$$

$$=5.7$$

$$d(1,4)=\sqrt{(12-7.5)^2+(4-3)^2+(14-14)^2+(8-7)^2+(14-14)^2+(10.5-10.5)^2+(4-4)^2}$$

$$=4.7$$

$$d(1,5)=\sqrt{(12-12)^2+(4-4)^2+(14-12)^2+(8-7)^2+(14-12)^2+(10.5-9)^2+(4-3)^2}$$

$$=3.5$$

$$d(2,1)=\sqrt{(9-9)^2+(3.5-2.5)^2+(12-8)^2+(7-6)^2+(12-10)^2+(9-10.5)^2+(3-3)^2}$$

$$=4.9$$

$$d(2,2)=\sqrt{(9-12)^2+(3.5-3.5)^2+(12-14)^2+(7-8)^2+(12-14)^2+(9-10.5)^2+(3-4)^2}$$

$$= 4.6$$

$$d(2,3)=\sqrt{(9-7.5)^2+(3.5-3)^2+(12-12)^2+(7-8)^2+(12-12)^2+(9-9)^2+(3-3)^2}$$

$$=1.9$$

$$d(2,4)=\sqrt{(9-7.5)^2+(3.5-3)^2+(12-14)^2+(7-7)^2+(12-14)^2+(9-10.5)^2+(3-4)^2}$$

$$=3.7$$

$$d(2,5)=\sqrt{(9-12)^2+(3.5-4)^2+(12-12)^2+(7-7)^2+(12-12)^2+(9-9)^2+(3-3)^2}$$

$$=3.0$$

$$d(3,1)=\sqrt{(7.5-9)^2+(3-2.5)^2+(8-8)^2+(6-6)^2+(10-10)^2+(7.5-10.5)^2+(2-3)^2}$$

$$=3.5$$

$$d(3,2)=\sqrt{(7.5-12)^2+(3-3.5)^2+(8-14)^2+(6-8)^2+(10-14)^2+(7.5-10.5)^2+(2-4)^2}$$

$$=9.5$$

$$d(3,3)=\sqrt{(7.5-7.5)^2+(3-3)^2+(8-12)^2+(6-8)^2+(10-12)^2+(7.5-9)^2+(2-3)^2}$$

$$=5.7$$

$$d(3,4)=\sqrt{(7.5-7.5)^2+(3-3)^2+(8-14)^2+(6-7)^2+(10-14)^2+(7.5-10.5)^2+(2-4)^2}$$

$$=8.1$$

$$d(3,5)=\sqrt{(7.5-12)^2+(3-4)^2+(8-12)^2+(6-7)^2+(10-12)^2+(7.5-9)^2+(2-3)^2}$$

$$=6.7$$

$$d(4,1)=\sqrt{(6-9)^2+(2-2.5)^2+(4-8)^2+(5-6)^2+(8-10)^2+(6-10.5)^2+(1-3)^2}$$

$$=7.4$$

$$d(4,2)=\sqrt{(6-12)^2+(2-3.5)^2+(4-14)^2+(5-8)^2+(8-14)^2+(6-10.5)^2+(1-4)^2}$$

$$=14.6$$

$$d(4,3)=\sqrt{(6-7.5)^2+(2-3)^2+(4-12)^2+(5-8)^2+(8-12)^2+(6-9)^2+(1-3)^2}$$

$$=22.5$$

$$d(4,4)=\sqrt{(6-7.5)^2+(2-3)^2+(4-14)^2+(5-7)^2+(8-14)^2+(6-10.5)^2+(1-4)^2}$$

$$=13.1$$

$$d(4,5)=\sqrt{(6-12)^2+(2-4)^2+(4-12)^2+(5-7)^2+(8-12)^2+(6-9)^2+(1-3)^2}$$

$$=11.7$$

Setelah melakukan perhitungan maka didapat hasil seperti berikut ini :

Tabel 3. Hasil Perhitungan

C1	C2	C3	C4
8.3	4.9	3.5	7.4
0.5	4.6	9.5	14.6
5.7	1.9	5.2	22.5
4.7	3.7	8.1	13.1
3.5	3.0	6.7	11.7

Tentukan kembali titik pusat cluster yang baru berdasarkan rata-rata

Cluster baru tersebut didapat dari rumus = nilai hasil / banyak hasil

Tabel 4 Cluster Baru

Cluster 1	12.0	3.8	14.0	8.0	14.0	9.8	3.5
Cluster 2	9.3	3.3	12.3	7.2	12.7	9.3	3.2
Cluster 3	9.0	2.5	8.0	6.0	10.0	10.5	3.0
Cluster 4	6.0	2.5	4.0	5.0	8.0	6.0	1.0

Lakukan kembali langkah 4 hingga titik pusat dari setiap cluster tidak berubah

Berikut hasil yang didapat sesuai dengan langkah ke 4

Tabel 5. Hasil Baru Perhitungan

C1	C2	C3	C4
8.2	5.5	0.0	8.1
1.0	3.8	8.2	15.5
5.5	2.2	5.4	11.2
4.8	3.2	7.5	14.1
12.2	2.9	5.9	12.5
3.1	2.9	5.9	12.5
6.2	3.0	5.8	10.5

Hasil dari tahapan yang pertama dan kedua tidak berubah, maka hasil sudah sesuai dengan

pengelompokkan kluster. Berikut adalah hasil dari pengelompokkan tersebut.

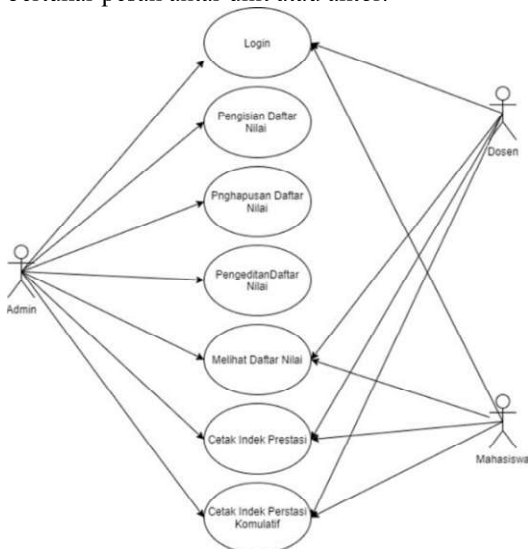
Tabel 6. Hasil Pengelompokkan Clustering

C1 = Nilai nya <u>sangat baik</u>	
2	Aan Saputra
C2 = Nilai nya <u>Baik</u>	
3	Aan Suhartri
4	Achmad Faisal
5	Achmad Fauzi
C3 = Nilai nya <u>Cukup</u>	
1	Aan Agus Hendra Al Azis
C4 = Nilai nya <u>Kurang</u>	
-	-

Desain proses bertujuan untuk menjelaskan dan menerangkan mengenai sistem yang akan dibangun secara keseluruhan. Desain sistem memberikan gambaran bagi para user atau pemakai sistem, dalam menerangkan proses yang terjadi pada suatu sistem sehingga memberikan kemudahan bagi mereka. Desain sistem dapat digambarkan dengan menggunakan.

1) Use Case Diagram

Use case diagram merupakan fungsionalitas yang disediakan sistem sebagai unit – unit yang saling bertukar pesan antar unit atau aktor.



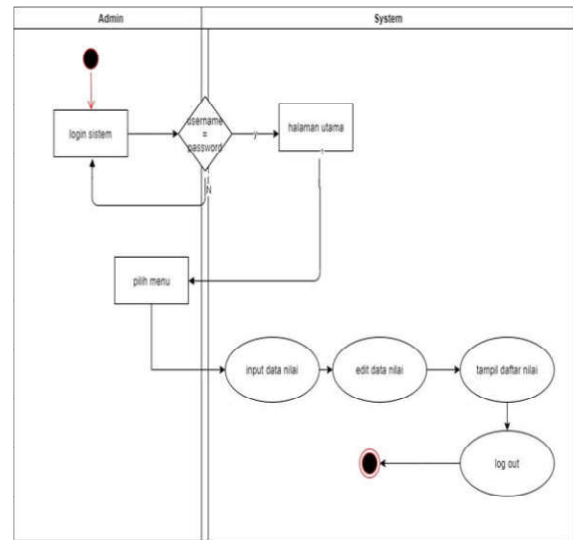
Gambar 2. Use Case Diagram

- Pada gambar diatas memiliki tiga entitas
- Login merupakan aktifitas pertama yang dilakukan oleh Admin,Dosen,dan Mahasiswa
 - Admin melakukan aktifitas yang meliputi pengisian daftar nilai,penghapusan daftar nilai,pengeditan daftar nilai,melihat nilai matakuliah,cetak indeks prestasi,cetak indeks prestasi kumulatif.

- Dosen melakukan aktifitas meliputi meliputi pengisian daftar nilai,penghapusan daftar nilai,pengeditan daftar nilai,melihat nilai matakuliah.
- Mahasiswa dapat melakukan aktifitas yang meliputi melihat nilai matakuliah,cetak indeks prestasi,cetak indeks prestasi kumulatif.

2) Activity Diagram

Activity diagram adalah memodelkan alur kerja (workflow) sebuah proses bisnis dan urutan aktivitas dalam suatu proses.



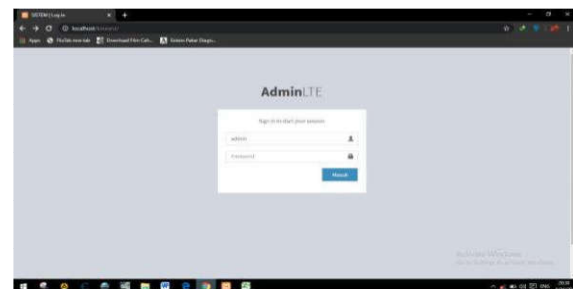
Gambar 3. Activity Diagram Admin

Pada activity diagram gambar 3 Admin membuka sistem dengan memasukkan user name dan password. Sistem menampilkan halaman utama, kemudian admin memasukkan daftar nilai, lalu mengedit nilai, selanjutnya sistem memproses data dan menampilkan hasil data yang di inputkan di halaman utama yang nanti nya akan proses menjadi daftar nilai.

3.2. Hasil

1) Tampilan Login

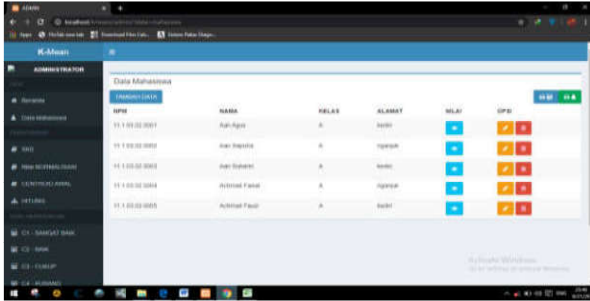
Menu Login adalah form yang pertama kali muncul saat program dijalankan. Form masuk digunakan untuk masuk ke dalam sistem. Tampilan Masuk dapat dilihat pada Gambar 4 sebagai berikut :



Gambar 4. Tampilan Login

2) Tampilan Data Mahasiswa

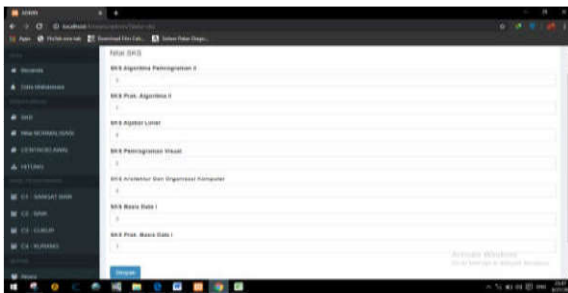
Dalam form ini dapat dilihat oleh admin dan digunakan menambah, menghapus dan mengedit data mahasiswa. Tampilan form dapat dilihat pada Gambar 5 sebagai berikut :



Gambar 5. Tampilan Data Mahasiswa

3) Tampilan Nilai Sks

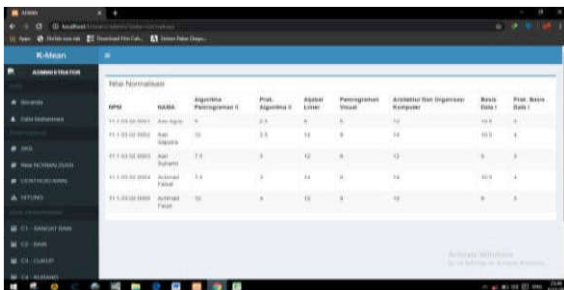
Dalam menu form nilai sks ini admin dapat mengisi nilai sks mahasiswa berdasarkan kriteria masing-masing jurusan. Tampilan form dapat dilihat pada Gambar 6 sebagai berikut :



Gambar 6. Tampilan Nilai Sks

4) Tampilan Nilai Normalisasi

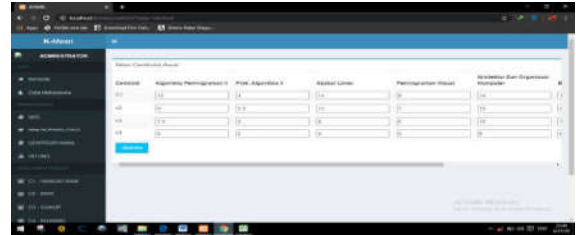
nilai normalisasi ini admin dapat melihat hasil normalisasi nilai mahasiswa. Tampilan form dapat dilihat pada Gambar 7 sebagai berikut :



Gambar 7. Tampilan Nilai Normalisasi

5) Tampilan Nilai Centroid awal

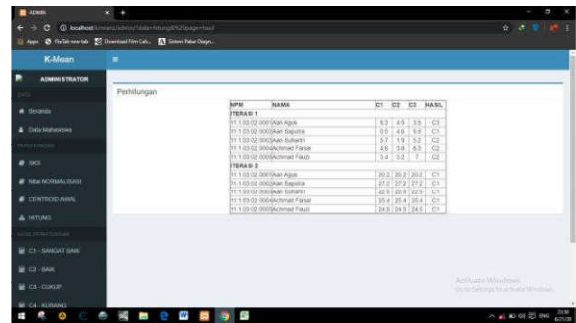
Dalam form centroid awal ini admin dapat mengisi nilai cluster yang telah ditentukan dari nilai normalisasi. Tampilan form dapat dilihat pada Gambar 8 sebagai berikut :



Gambar 8. Tampilan Nilai Centroid Awal

6) Tampilan Hasil Perhitungan

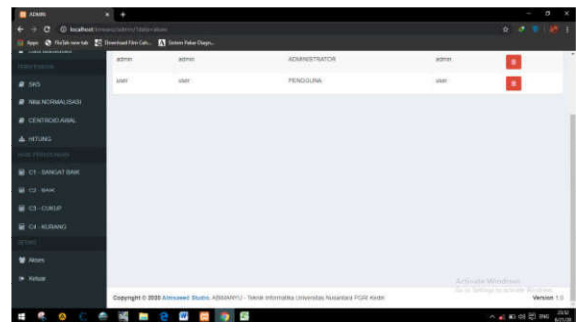
Dalam form ini admin dapat melihat hasil perhitungan. Tampilan form dapat dilihat pada Gambar 9 sebagai berikut :



Gambar 9. Tampilan Hasil Perhitungan

7) Tampilan Akses

Dalam form ini admin dapat menambahkan siapa saja yang bisa mengakses web. Tampilan form dapat dilihat pada Gambar 3.7 sebagai berikut:



Gambar 10. Tampilan Akses

4. SIMPULAN

Dari hasil penelitian, perancangan, pembuatan dan pengujian aplikasi Prediction of Student Academic Value Achievement Using the K-Means Clustering Method Web Based dengan menggunakan metode Algoritma K-Means Clustering didapatkan kesimpulan sebagai berikut:

- 1) Aplikasi Prediction of Student Academic Value Achievement Using the K-Means Clustering Method Web Based dengan menggunakan metode Algoritma K-Means Clustering berhasil sesuai rencana dibuat dengan tampilan yang sangat memudahkan penggunaannya.

- 2) Dengan penggunaan Algoritma K-Means Clustering hasil perhitungan antara cluster cukup baik

5. SARAN

Dari hasil penelitian, perancangan, pembuatan dan pengujian aplikasi Prediksi Prestasi Nilai Akademik Mahasiswa Menggunakan Metode K-Means Clustering Berbasis yang telah di buat masih perlu dikembangkan. Disarankan untuk menambahkan sampel nilai yang lebih banyak agar hasil yang didapat lebih maksimal.

DAFTAR PUSTAKA

- [1] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques Third Edition*, San Fransisco: Morgan Kauffman, 2012.
- [2] H. Olson and Y. Shi, *Introduction to Bussiness Data Mining*, New York: McGraw-Hill, 2007.
- [3] Widodo. 2004. *Psilogi Belajar*. Jakarta: Rineka Cipta.
- [4] Pressman, R.S. 2015. *Rekayasa Perangkat Lunak: Pendekatan Partisi Buku I*. Yogyakarta : Andi
- [5] Agusta, Y. (2007). K-Means - Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika*, 3, 47-60.
- [2] Xu, R., & Wunsch II, D. C. (2009). *Clustering*. Kanada: IEEE Press.