

# Perancangan Perangkat Lunak Data Mining Pemilihan Jurusan pada Sekolah Kristen Makedonia Ngabang

Fourensia Labat Paskharida<sup>\*1</sup>, Utin Kasma<sup>2</sup>

<sup>1,2</sup>Jurusan Teknik Informatika; STMIK Pontianak. Jl. Merdeka No.372 Pontianak, 0561-735555  
e-mail: <sup>\*1</sup>fourensia.labat@gmail.com, <sup>2</sup>utin.kasma@stmikpontianak.ac.id

## **Abstrak**

*Perkembangan teknologi informasi pada saat ini memang sangat pesat dan berperan penting dalam segala aspek kehidupan termasuk dalam instansi pendidikan. Dengan semakin berkembangnya teknologi pada masa kini tidak bisa dipungkiri bahwa teknologi menjadi kebutuhan sekunder bagi para penggunanya. Sekolah Kristen Makedonia merupakan salah satu yang berperan dalam bidang pendidikan. SMA ini memiliki jumlah murid yang tidak sedikit dan tidak sedikit juga jumlah siswa yang akan melanjutkan ke jenjang selanjutnya. Clustering merupakan metode data mining yang paling banyak digunakan untuk mengelompokkan data berdasarkan nilai tertentu dan algoritma yang digunakan yaitu algoritma K-Means. Algoritma K-Means merupakan algoritma yang sering digunakan dalam melakukan pengelompokkan karena dari algoritma ini dapat menentukan data untuk masuk kedalam kelompoknya berdasarkan jumlah rata-rata dari nilai tersebut. Tujuan dalam penelitian ini yaitu, untuk mengetahui bagaimana penerapan Data Mining database dalam menentukan jurusan bagi siswa Sekolah Kristen Makedonia. Metode perancangan yang digunakan peneliti adalah metode RAD dan menggunakan aplikasi berbasis desktop dengan bahasa pemrograman Visual Basic 2010. Hasil dari penelitian ini berupa Aplikasi Pemilihan Jurusan yang dapat membantu siswa dalam menentukan jurusan apa yang akan diambil nantinya.*

**Kata kunci : Sekolah Kristen Makedonia, Data Mining, Visual Basic 2010, K-Means**

## **Abstract**

*The development of information technology current days is very fast and has an important part in all of life aspect includes education institute. Technology has become secondary needs for the users. Makedonia Christian Schools that has a role in education field. This senior high school has large amount of students that will continue their education to the next level. Clustering is the one of data mining method that mostly use to classify the data according to certain score and algorithm tha being used is K-Means Algorithm. K-Means algorithm is an algotihm which can determine the data to its own group according to score mean. The purpose af this research is to know how to apply data mining database in order to determine student's major in Makedonia Christian School. Methodology that being used in this research is Rapid Application Development(RAD) method and desktop application with visual basic 2010 program. The result of this research in the form of "pemilihan jurusan" application which can help student to determine the major that they will take.*

**Keywords : Macedonia Christian School, Data Mining, Visual Basic 2010, K-Means**

## 1. PENDAHULUAN

Perkembangan teknologi informasi pada saat ini memang sangat pesat dan berperan penting dalam segala aspek kehidupan, sehingga orang banyak dapat mengetahui kegiatan apa saja yang dijalankan dan keakuratan informasi pun sangat dibutuhkan. Kemampuan teknologi

---

informasi dalam menyimpan, mengumpulkan data dalam jumlah yang sangat besar dan bahkan dapat mengambil informasi dalam sistem yang sudah ada.

Persekolahan Kristen Makedonia Ngabang merupakan suatu lembaga yang bergerak dibidang dunia pendidikan yang memiliki jumlah murid tidak sedikit. Pada saat ini lulusan dari Sekolah Kristen Makedonia sudah banyak mencetak para sarjana, mulai dari bidang kesehatan, keguruan, teknik hingga yang lainnya. Namun sejujurnya hal dalam menentukan jurusan jurusan sedikit membingungkan bagi para siswa khususnya siswa kelas 3 SMA, dimana pengetahuan dan informasi yang sangat terbatas tentang pemilihan jurusan kuliah. Untuk dapat membantu para siswa dalam memilih keputusan untuk pemilihan jurusan, maka sekolah berusaha membantu para siswa. Tetapi dikarenakan jumlah siswa yang banyak maka akan menyulitkan pihak sekolah terutama guru BK yang dimana memiliki peran untuk membantu para siswa jika ada keluhan dan juga sebagai pemberi solusi bagi masalah yang disampaikan para siswa. Namun bagi anak kelas 3 SMP yang akan melanjutkan ke tingkat SMA tidak terlalu sulit untuk menentukan jurusan mana yang akan diambil karena dibantu oleh guru BK dalam menentukan jurusan. Oleh karena itu sekolah sangat membutuhkan sistem yang dimana nantinya dapat membantu para siswa dalam memilih jurusan sehingga para murid tidak kebingungan lagi dengan pemilihan jurusan untuk melanjutkan study ke jenjang yang lebih tinggi terutama siswa kelas 3 SMA. Jurusan untuk melanjutkan pendidikan sangat lah beragam sehingga dapat membingungkan para siswa dalam memilih. Adanya data mining dapat memberi informasi kepada para siswa jurusan mana yang paling banyak peminatnya dan paling mana yang sedikit peminatnya, sehingga para siswa dapat memilih jurusan mana yang diambil sesuai dengan nilai mereka.

Pengelompokkan berdasarkan nilai akademik siswa dapat membantu dalam menentukan jurusan mana yang paling banyak diambil jika dilihat dari nilai-nilai itu, dengan itu dapat ditemukan suatu solusi untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam kelanjutan penelitian ini. Salah satu teknik yang digunakan dalam penelitian kali ini adalah teknik K-Means yang dimana teknik ini merupakan teknik sederhana dalam analisa clustering.

Data mining dapat diartikan sebagai proses untuk mendapatkan informasi yang berguna dari tempat penyimpanan data yang besar yang membantu dalam pengambilan keputusan. Salah satu teknik dalam data mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam tempat penyimpanan. Kebutuhan untuk memprediksi juga dapat memanfaatkan teknik ini. Data mining bertujuan dapat mengetahui pola universal data-data yang ada dan selain itu juga dapat digunakan untuk mendukung kegiatan perusahaan atau suatu lembaga pendidikan dan pemerintahan. Salah satu metode yang ada pada data mining ini adalah metode clustering yang dimana metode ini melakukan pengelompokkan data-data ke dalam sejumlah kelompok berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada.

Algoritma K-Means mengambil data dengan menggunakan metode Clustering untuk menghitung nilai cluster dari setiap kelompok yang telah ditentukan dengan menggunakan metode clustering. Metode clustering terlebih dahulu sudah menentukan jumlah cluster dan nilai-nilai yang dapat masuk kedalam cluster masing-masing dengan perhitungan menggunakan algoritma K-Means. Algoritma K-Means menggunakan 2 kali iterasi dan bisa lebih untuk menentukan hasil yang lebih akurat.

Dalam dunia pendidikan, khususnya Universitas, menurut para pengembang untuk menentukan siswa-siswa yang mengikuti lomba, pengelompokkan studi dan memilih jurusan dengan menganalisis menggunakan nilai yang ada. Untuk menemukan kesamaan dan perbedaan sehingga dapat menentukan siswa yang mengikuti lomba dengan melihat dari nilai siswa[1].

Untuk menggolongkan atas beberapa kriteria yang memiliki karakter yang sama, sehingga dapat menghasilkan penentuan dalam pengelompokan studi pada universitas[2]. Untuk memberikan informasi jurusan keahlian kepada mahasiswa sehingga mahasiswa dapat menentukan jurusan keahliannya[3].

Dengan menggunakan algoritma K-Means dalam data mining dengan metode Clustering dapat ditentukan jurusan yang baik dan diambil siswa berdasarkan nilai yang siswa miliki berdasarkan hasil yang diperoleh dengan dilakukannya iterasi sebanyak 2 kali. Dengan mengetahui hasilnya maka akan memudahkan siswa dalam mengambil atau menentukan jurusan apa yang cocok dan pas untuk melanjutkan pendidikan ke jenjang selanjutnya.

## 2. METODE PENELITIAN

Penelitian ini berbentuk studi literature dan perancangan eksperimen, sebagai bahan untuk mengumpulkan dan melakukan perancangan. Teknik pengumpulan data dilakukan dengan melakukan studi literatur dan observasi. Hasil dari observasi dikumpulkan menjadi data lebih yang akan digunakan untuk membangun algoritma K-Means untuk pemilihan jurusan. Metode perancangan perangkat lunak yang digunakan menggunakan Rapid Application Development (RAD) karena proses perkembangan perangkat lunak ini berjalan satu arah yaitu dari awal sampai akhir dan bersifat inkremental terutama untuk waktu pengerjaan yang pendek[4].

Metode perancangan perangkat lunak yang digunakan dalam penelitian ini yaitu menggunakan *Rapid Application Development (RAD)*.

Adapun *fase-fase Rapid Application Development (RAD)* model sebagai berikut :

1. **Pemodelan Bisnis**  
Pemodelan yang dilakukan untuk memodelkan fungsi bisnis untuk mengetahui informasi yang terkait proses bisnis, informasi apa yang harus dibuat, siapa yang harus membuat informasi itu, bagaimana alur informasi itu, proses apa saja yang terkait informasi itu.
2. **Pemodelan Data**  
Memodelkan data apa saja yang dibutuhkan berdasarkan pemodelan bisnis dan mendefinisikan atribut-atributnya beserta relasinya dengan data-data yang lain.
3. **Pemodelan proses**  
Mengimplementasikan fungsi bisnis yang sudah didefinisikan terkait dengan pendefinisian data.
4. **Pembuatan aplikasi**  
Mengimplementasikan pemodelan proses dan data menjadi program. Model RAD sangat menganjurkan pemakaian komponen yang sudah ada jika dimungkinkan.
5. **Pengujian dan pergantian**  
Menguji komponen-komponen yang dibuat. Jika sudah teruji maka tim pengembang komponen dapat beranjak untuk mengembangkan komponen berikutnya.

Metode Clustering merupakan pengelempokan sejumlah data/obyek ke dalam cluster (group) sehingga setiap cluster akan berisi data yang semirip mungkin. Metode ini berusaha untuk menempatkan obyek yang mirip (jaraknya dekat) dalam cluster dan membuat jarak cluster sejauh mungkin. Ini berarti obyek dalam satu cluster sangat mirip satu sama lain dan berbeda dengan obyek dalam cluster-cluster yang lain. Dalam metode ini tidak diketahui sebelumnya berapa jumlah cluster dan bagaimana pengelompokannya[5].

K-Means merupakan metode analisis kelompok yang mengarah pada pemartisan N objek pengamatan ke dalam K kelompok (*cluster*) diman setiap obyek pengamatan dimiliki oleh sebuah kelompok dengan mean (rata-rata) terdekat. Tujuan pengelompokan ini untuk

meminimalkan fungsi objektif yang diset dalam proses pengelompokkan, yang berusaha meminimalkan variasi dalam suatu kelompok dan memaksimalkan variasi antar kelompok[6].

Adapun pengelompokkan dengan metode K-Means sebagai berikut :

- a. Tentukan jumlah kelompok
- b. Alokasikan data ke dalam kelompok secara acak
- c. Hitung pusat kelompok (sentroid/rata-rata) dari data yang ada masing-masing kelompok
- d. Alokasikan masing-masing data ke sentroid/rata-rata terdekat
- e. Kembali ke langkah 3, apabila masih ada data yang berpindah kelompok, atau apabila ada perubahan nilai sentroid di atas nilai ambang yang ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih diatas nilai ambang yang ditentukan.

Rumus yang digunakan dalam melakukan perhitungan yaitu rumus Euclidean Distance dimana Euclidean menjadi pilihan jika kita ingin memberikan jarak terpendek antara dua titik dan nilai jarak keduanya bersifat mutlak.

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

Dalam merancang sebuah sistem, mengatur kompleksitas adalah salah satu alasan utama mengapa harus membuat model. Pemodelan membantu para pengembang untuk dapat fokus, dapat mendokumentasikan, menangkap keseluruhan sistem dan mengkomunikasikan aspek-aspek penting dalam sistem yang sedang dirancang. UML tepat digunakan untuk memodelkan informasi sistem untuk perusahaan, bahkan untuk sistem yang rumit sekalipun. UML (Unified Modelling Language) adalah bahasa pemodelan untuk sistem atau perangkat lunak yang berparadiagram berorientasi objek[7].

Metode *White Box* merupakan sebuah filosofi perancangan *test-case* yang menggunakan struktur kontrol yang dijelaskan sebagai bagain dari perancangan perangkat komponen untuk menghasilkan *test-case*[8].

### 3. HASIL DAN PEMBAHASAN

Penelitian yang dilakukan peneliti dalam pengembangan aplikasi data mining ini menggunakan metode *Rapid Application Development* (RAD). Pendekatan pengembangan ini memudahkan peneliti dalam merancang aplikasi, sebab dalam tahapan pengembangan Rapid Application Development (RAD) memiliki tahapan pengembangan yaitu permodelan bisnis, permodelan data, permodelan proses, pembuatan aplikasi, pengujian dan pengantian.

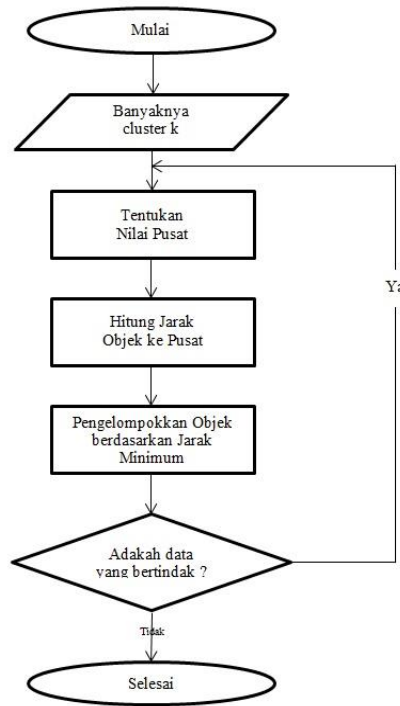
#### 3.1 Perancangan Permodelan Perangkat Lunak

Perancangan permodelan perangkat lunak yang digunakan dalam membangun perancangan aplikasi data mining ini adalah *flowchart*, dan diagram UML yaitu *use case diagram*, *sequence diagram*, dan *activity diagram*.

*Flowchart* digunakan untuk memudahkan dalam pemahaman akan cara kerja algoritma k-means. Pada awalnya ditentukan beberapa *cluster* yang akan dibentuk. Sembarang obyek atau elemen pertama dalam *cluster* dapat dipilih untuk dijadikan sebagai titik tengah (centroid point) cluster. Algoritma K-means selanjutnya akan melakukan pengulangan langkah-langkah berikut

---

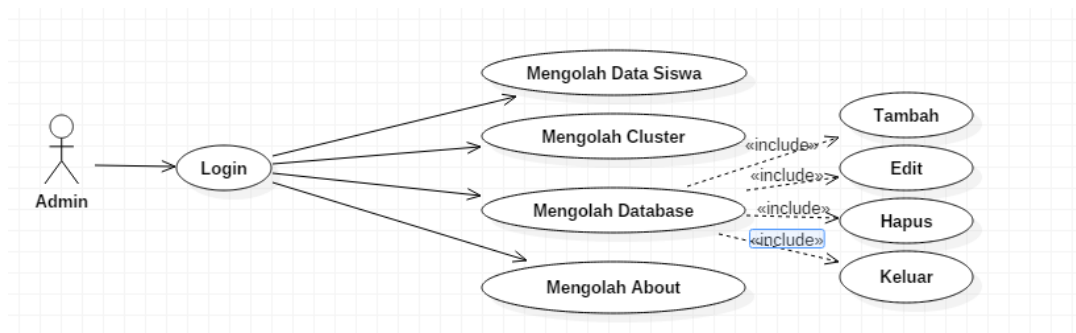
sampai terjadi kestabilan atau tidak ada obyek yang dapat dipindahkan, maka perhitungan selesai.



Gambar 1 Flowchart Algoritma K-Means

### 3.1.1 Use Case Diagram

Use case diagram menjelaskan fungsi atau manfaat sistem jika menurut orang yang berada diluar sistem atau user. Diagram ini menunjukkan fungsional suatu sistem atau kelas dari bagaimana sistem berinteraksi dengan dunia luar. Perancangan proses yang terjadi pada perancangan aplikasi data mining menggunakan Use Case Diagram sebagai berikut :



Gambar 2 Use Case Diagram Perancangan Perangkat Lunak

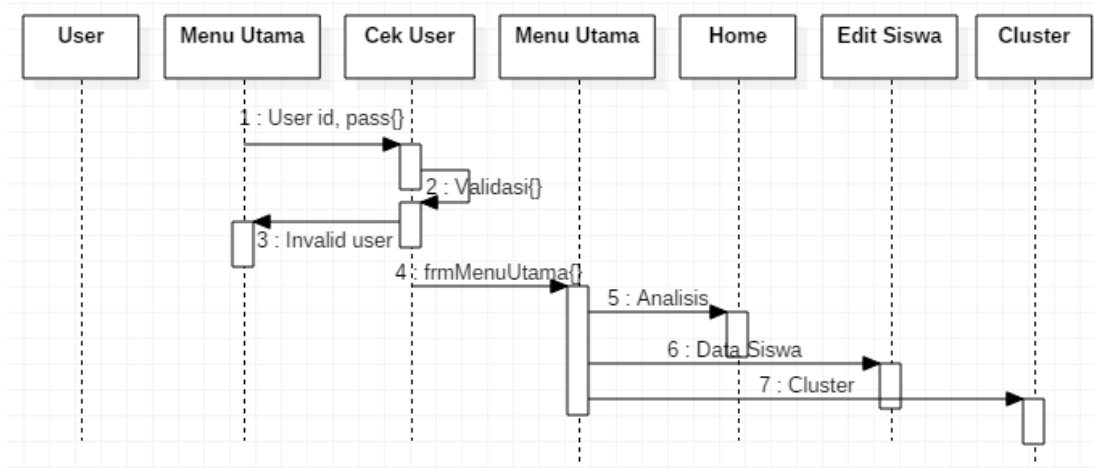
Use Case Diagram pengolahan perancangan perangkat lunak terdiri dari admin. Admin bertugas untuk memanajemen isi dari perancangan perangkat lunak secara keseluruhan.

Actor:	Admin	
--------	-------	--

	1.	Use Case	Login Admin
Deskripsi:		<ol style="list-style-type: none"> <li>1. Admin harus melakukan login untuk kebutuhan pengolahan data.</li> <li>2. Username dan password harus diinputkan.</li> <li>3. Apabila username dan password sesuai maka akan masuk ke menu utama.</li> <li>4. Jika username dan password salah maka admin diminta mengisikan ulang.</li> </ol>	
	2	Use Case	Mengolah Data Siswa
Deskripsi:		Admin dapat menambahkan data siswa, mengedit data siswa dan menghapus data siswa.	
	3	Use Case	Mengolah Cluster
Deskripsi:		Admin mengcluster data umat yang sebelumnya sudah tersimpan didalam database.	
	4	Use Case	Input Database
Deskripsi:		Admin menginput data siswa dan memasukkan kedalam database.	

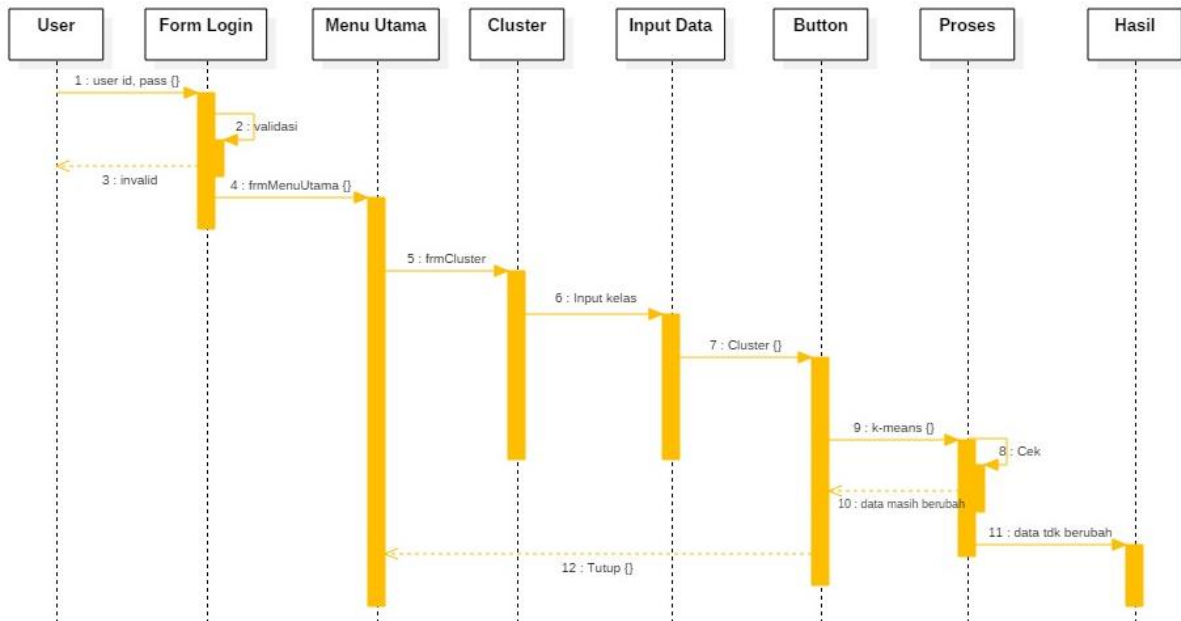
### 3.1.2 Sequence Diagram

*Sequence Diagram* menggambarkan objek pada use case dengan mendeskripsikan waktu hidup objek dan pesan yang dikirimkan dan diterima antar objek. Membuat diagram sequence juga dibutuhkan untuk melihat skenario yang ada pada use case. Banyaknya diagram sequence yang harus digambar adalah minimal sebanyak pendefinisian *use case* yang memiliki proses sendiri atau yang penting semua *use case* yang telah didefinisikan interaksi jalannya pesan sudah dicakup pada diagram sequence sehingga semakin banyak *use case* yang didefinisikan maka diagram sequence yang harus dibuat harus semakin banyak. Diawali dari apa men-trigger aktivitas tersebut, proses dan perubahan apa saja yang terjadi secara internal dan output apa yang dihasilkan. Masing-masing objek, termasuk actor, memiliki *;-lifeline vertical*. Message digambarkan sebagai garis berpanah dari satu objek ke objek lainnya. Pada fase desain berikut, message akan dipetakan menjadi operasi atau metode dari class.



Gambar 3 Sequence Diagram Utama

Gambar 2 menunjukkan User/Admin memilih Menu lalu akan dilakukan proses penginputan username dan password kemudian setelah valid maka user/admin dapat memilih open Home sistem akan mencari About dan Keluar, jika memilih Keluar maka sistem akan kembali ke Form Utama tetapi tidak dapat mengakses Edit Siswa dan Cluster ataupun yang lainnya. User juga dapat memilih menu Data Siswa/Edit Siswa dan dapat melakukan proses Cluster.

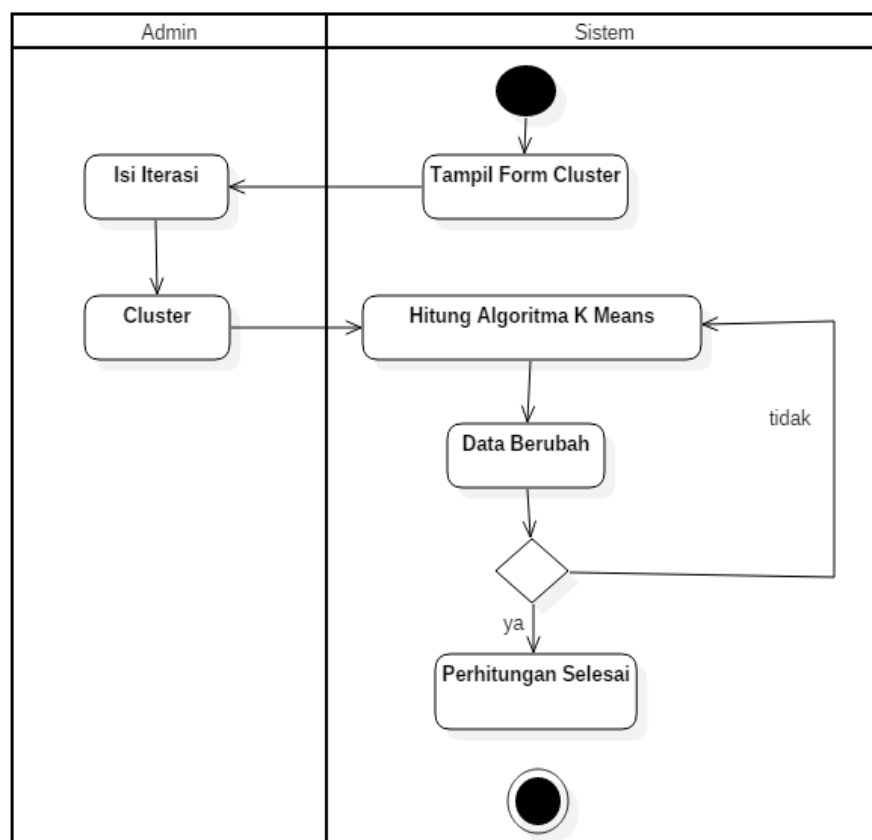


Gambar 4 Sequence Diagram Cluster

Gambar 3 menunjukkan user/admin memilih cluster, kemudian admin/user menginputkan kelas yang akan dilakukan cluster kemudian admin harus mengklik tombol proses maka cluster akan melakukan perhitungan sehingga dapat menghasilkan hasil dari perhitungan dan langsung dalam kelompok yang ada.

### 3.1.3 Activity Diagram

*Activity Diagram* menggambarkan berbagai alur aktivitas dalam sistem yang sedang dirancang, bagaimana masing-masing alur berawal, desicion yang mungkin terjadi, dan bagaimana mereka berakhir. *Activity Diagram* juga dapat menggambarkan proses parallel yang mungkin terjadi pada beberapa eksekusi. *Activity Diagram* merupakan state diagram khusus, dimana sebagian besar state adalah action dan sebagai besar transisi di-trigger oleh selesainya state sebelumnya (*internal processing*). Oleh karena itu *Activity Diagram* tidak menggambarkan behavior internal sebuah sistem (dan interaksi antar subsistem) secara eksak, tetapi lebih menggambarkan proses-proses dan jalur-jalur aktivitas dari level atas secara umum.



Gambar 5 *Activity Diagram* Clustering K-Means

Merupakan alur kerja sistem pada form cluster, pada saat user mengisi cluster yaitu menentukan kelas mana yang akan di cluster lalu akan mengklik tombol cluster maka sistem akan menghitung dan menghasilkan proses hasil yang dihitung.

### 3.2 Spesifikasi File Database

Spesifikasi file database merupakan serangkaian tabel database yang menjadi media penyimpanan rancangan sistem yang diusulkan. Berikut ini akan diuraikan spesifikasi dari tabel-tabel yang terdapat dalam perancangan aplikasi perangkat lunak data mining.



Tabel 1 Tabel Login

Nama Tabel	Nama Field	Tipe Data	Size	Primary Key
tbl_login	username	Varchar	50	*
	password	Varchar	50	

Tabel 2 Tabel Data Siswa

Nama Tabel	Nama Field	Tipe Data	Size	Primary Key
Tbl_12ipaa	Kelas	Varchar	10	
	NISN	Varchar	20	*
	Nama	varchar	100	
	Nilai_mtk	Varchar	10	
	Nilai_jrs	varchar	10	

Tabel data siswa yang ditampilkan cuma 1 dikarenakan tabel nya berisikan atribut yang sama.

### 3.3 Tampilan Hasil

Tabel 4 Tabel Nilai Siswa

1401151	89	77	1401176	90	74
1401152	80	75	1401177	75	80
1401153	78	95	1401178	87	82
1401154	88	89	1401179	79	85
1401155	80	77	1401180	76	82
1401156	75	76	1401181	93	89
1401157	78	78	1401182	91	71
1401158	77	80	1401183	90	87
1401159	82	78	1401184	80	84
1401160	81	79	1401185	82	82
1401161	80	81	1401186	81	80
1401162	79	82	1401187	79	83
1401163	90	76	1401188	78	81
1401164	92	98	1401189	71	80
1401165	89	73	1401190	70	80
1401166	87	80	1401191	77	83
1401167	93	93	1401192	79	85
1401168	79	82	1401193	74	78
1401169	98	86	1401194	80	79
1401170	85	90	1401195	84	75
1401171	91	73	1401196	81	76
1401172	78	98	1401197	87	72
1401173	80	89	1401198	85	77
1401174	77	78	1401199	82	76
1401175	75	87	1401200	90	78

Tabel 5 Tabel Cluster

Titik Pusat Awal	Nilai Matematika	Nilai Jurusan
Cluster 1	75	75
Cluster 2	80	80
Cluster 3	85	85

Persamaan yang penulis gunakan untuk menghitung jarak setiap data cluster antara object ke centroid adalah dengan jarak Euclidean  $d(p,q) = \sqrt{(p_1-p_2)^2 + (p_2-q_2)^2}$  Persamaan ini digunakan karena atribut yang digunakan berjumlah 2.

Keterangan : p = nilai dari nilai rata-rata siswa (nilai matematika dan jurusan)

q = nilai dari titik pusat

jadi perhitungan cluster k-means untuk data siswa pertama ke pusat cluster pertama dengan persamaan

$$\begin{aligned}d(1,1) &= \sqrt{(89-75)^2 + (77-75)^2} \\ &= \sqrt{(14)^2 + (2)^2} \\ &= \sqrt{196+4} = 14.1421\end{aligned}$$

perhitungan cluster k-means untuk data siswa pertama ke pusat cluster kedua

$$\begin{aligned}d(1,2) &= \sqrt{(89-80)^2 + (77-80)^2} \\ &= \sqrt{(9)^2 + (-3)^2} \\ &= \sqrt{81+9} = 9.4868\end{aligned}$$

perhitungan cluster k-means untuk data siswa pertama ke pusat cluster ketiga

$$\begin{aligned}d(1,3) &= \sqrt{(89-85)^2 + (77-85)^2} \\ &= \sqrt{(4)^2 + (-8)^2} \\ &= \sqrt{16+64} = 8.9442\end{aligned}$$

Setelah melakukan cluster dengan titik pusat pertama maka akan didapatkan hasil sesuai dengan kelompok-kelompok masing. Maka setelah itu ditentukanlah titik pusat baru berdasarkan kelompok masing-masing.

No	Nomor Induk	Nama	Nilai			c1	c2	c3		
			Matematika	Jurusan						
1	1401151	Afgan Dharma	89	77	14,14213562	9,486832981	8,94427191	8,94427191	c3	
2	1401152	Angelina Cici	80	75		5	11,18033989		c2	
3	1401153	Anggun Paramita	78	95	20,22374842	15,13274595	12,20655562	12,20655562	c3	
4	1401154	Anita	88	89	19,10497317	12,04159458	5		c3	
5	1401155	Anjani	80	77	5,385164807		3	9,433981132	c2	
6	1401156	Benedikta Barada	75	76		1	6,403124237	13,45362405	c1	
7	1401157	Benediktus Paskalis	78	78	4,242640687	2,828427125	9,899494937	2,828427125	c2	
8	1401158	Benny Anugrah	77	80	5,385164807		3	9,433981132	c2	
9	1401159	Bernice Grazinia	82	78	7,615773106	2,828427125	7,615773106	2,828427125	c2	
10	1401160	Brian Panjaitan	81	79	7,211102551	1,414213562	7,211102551	1,414213562	c2	
11	1401161	Cindy Ivanna Nainggolan	80	81	7,810249676		1	6,403124237	c2	
12	1401162	Cheza Daris	79	82	8,062257748	2,236067977	6,708203932	2,236067977	c2	
13	1401163	Cici Margaretha	90	76	15,03329638	10,77032961	10,29563014	10,29563014	c3	
14	1401164	Cornelius Barage	92	98	28,60069929	21,63330765	14,76482306	14,76482306	c3	
15	1401165	Dede Purnama	89	73	14,14213562	11,40175425	12,64911064	11,40175425	c2	
16	1401166	Diana Putri	87	80		13	7	5,385164807	c3	
17	1401167	Dody Firmansyah	93	93	25,45584412	18,38477631	11,3137085	11,3137085	c3	
18	1401168	Eleazer Kurniawan	79	82	8,062257748	2,236067977	6,708203932	2,236067977	c2	
19	1401169	Erwin Pardamean	98	86	25,49509757	18,97366596	13,03840481	13,03840481	c3	
20	1401170	Endah Kusuma Sari	85	90	18,02775638		11,18033989	5	c3	
21	1401171	Eka Kusuma Wijaya	91	73	16,1245155	13,03840481	13,41640786	13,03840481	c3	
22	1401172	Firmandika Syaputra	78	98	23,19482701	18,11077028	14,76482306	14,76482306	c3	
23	1401173	Fivian Glory	80	89	14,86606875		9	6,403124237	c3	
24	1401174	Felix Geraldo	77	78	3,605551275	3,605551275	10,63014581	3,605551275	c2	
25	1401175	Gema Putri Siregar	75	87		12	8,602325267	10,19803903	c2	
26	1401176	Genta Putra	90	74	15,03329638	11,66190379	12,08304597	11,66190379	c2	
27	1401177	Gladysia Nainggolan	75	80		5	11,18033989	5	c2	
28	1401178	Harry Andika	87	82	13,89244399	7,280109889	3,605551275	3,605551275	c3	
29	1401179	Hermansyah Rafael	79	85	10,77032961	5,099019514	6	5,099019514	c2	
30	1401180	Hendro Simatupang	76	82	7,071067812	4,472135955	9,486832981	4,472135955	c2	

Gambar 6 Hasil Cluster Pertama

Setelah hasil cluster pertama telah ditemukan maka pengelompokkan dilakukan berdasarkan nilai terkecil yang telah dilakukan oleh sistem. Nilai yang mendekati angka 0 maka akan masuk kedalam kelompok tersebut. setelah masuk kedalam kelompoknya masing-masing maka sistem akan melakukan pencarian centroid baru untuk menentukan nilai centroid yang pertama. Menentukan centroid baru harus berdasarkan nilai rata-rata dari kelompok masing-masing.

$$C11 = (75+71+70+74)/4 = 72,5$$

$$C12 = (76+80+80+78)/4 = 78,5$$

Untuk cluster 1 nilai tengah yang akan digunakan dalam menghitung cluster kedua yaitu 72,5 dan 78,5

$$C21=(80+80+78+77+82+81+80+79+89+79+77+75+90+75+79+76+91+80+82+81+79+78+79+74+77+80+84+81+87+85)/30 = 80,5$$

$$C22=(75+77+78+80+78+79+81+82+73+82+78+87+74+80+85+82+71+84+82+80+83+81+85+78+83+79+75+76+72+77)/30 = 79,23333333$$

Untuk cluster kedua nilai tengah yang akan digunakan dalam menghitung cluster kedua yaitu 80,5 dan 79,23333333

## Perancangan Perangkat Lunak Data Mining Pemilihan Jurusan pada Sekolah Kristen Makedonia Ngabang

No	Nomor Induk	Nama	Nilai				c3	
			Matematika	Jurusan	c1	c2		
1	1401151	Afgan Dharma	89	77	16,56804153	8,787656115	8,609326338	c3
2	1401152	Angelina Cici	80	75	8,276472679	4,259448321	13,44100071	c2
3	1401153	Anggun Paramita	78	95	17,39252713	15,96693145	13,95064515	c3
4	1401154	Anita	88	89	18,72164523	12,31677312	3,432273299	c3
5	1401155	Anjani	80	77	7,64852927	2,285366491	11,9306538	c2
6	1401156	Benedikta Barada	75	76	3,535533906	6,378314824	16,38293319	c1
7	1401157	Benediktus Paskalis	78	78	5,522680509	2,786198126	12,78047339	c2
8	1401158	Benny Anugrah	77	80	4,74341649	3,583699206	12,59366904	c1
9	1401159	Bernice Grazinia	82	78	9,513148795	1,939819579	9,849898477	c2
10	1401160	Brian Panjaitan	81	79	8,514693183	0,550363516	9,820412415	c2
11	1401161	Cindy Ivanna Naiggol	80	81	7,90569415	1,839266158	9,47103479	c2
12	1401162	Cheza Daris	79	82	7,38241153	3,150063491	9,955927882	c2
13	1401163	Cici Margaretha	90	76	17,67766953	10,0340869	9,731418191	c3
14	1401164	Cornelius Barage	92	98	27,57716447	22,01278946	12,96227218	c3
15	1401165	Dede Purnama	89	73	17,39252713	10,53863843	12,60001984	c2
16	1401166	Diana Putri	87	80	14,57737974	6,545448801	5,727172077	c3
17	1401167	Dody Firmansyah	93	93	25,10975906	18,5973896	8,788657463	c3
18	1401168	Eleazer Kurniawan	79	82	7,38241153	3,150063491	9,955927882	c2
19	1401169	Erwin Pardamean	98	86	26,58006772	18,7638722	9,719079174	c3
20	1401170	Endah Kusuma Sari	85	90	16,98528775	11,6723134	5,510036298	c3
21	1401171	Eka Kusuma Wijaya	91	73	19,30025907	12,20913183	12,86858578	c2
22	1401172	Firmandika Syaputra	78	98	20,26079959	18,93575718	16,12887163	c3
23	1401173	Fivian Glory	80	89	12,9034879	9,782785902	8,967747766	c3
24	1401174	Felix Geraldo	77	78	4,527692569	3,709838271	13,59854772	c1
25	1401175	Gema Putri Siregar	75	87	8,860022573	9,519606084	13,36564626	c1
26	1401176	Genta Putra	90	74	18,069311	10,84448708	11,70557559	c2
27	1401177	Gladysia Nainggolan	75	80	2,915475947	5,553638447	14,41389954	c1
28	1401178	Harry Andika	87	82	14,91643389	7,065613915	3,805325216	c3
29	1401179	Hermansyah Rafael	79	85	9,192388155	5,961786645	9,308087881	c2
30	1401180	Hendro Simatupang	76	82	4,949747468	5,284212335	12,80080076	c1

Gambar 7 Hasil Cluster Kedua

Untuk c1 itu merupakan siswa yang cocok masuk kedalam jurusan non tehnik, c2 merupakan siswa yang cocok masuk delama jurusan tehnik dan c3 merupakan siswa yang bebas memilih masuk ke jurusan apa saja. Dapat di tarik kesimpulan melalui nilai yang siswa punya dan melalui perhitungan algoritma k-means. Jadi dengan siswa yang memperoleh nilai 80-85 maka siswa tersebut dapat masuk kedalam jurusan tehnik dan nilai 85 keatas masuk kedalam 2 jurusan tersebut, ini didapat untuk memberikan referensi bagi siswa agar dapat memilih jurusan sesuai dengan nilai yang mereka miliki.

#### 4. KESIMPULAN

Setelah menyelesaikan Pengujian dan Perancangan Perangkat Lunak Data Mining Sekolah Kristen Makedonia Ngabang, penulis telah mengambil kesimpulan yaitu bahwa siswa yang memiliki nilai dengan rentang 70-80 memiliki peluang untuk masuk ke jurusan non tehnik, siswa dengan rentang 80-85 memliki peluang untuk masuk ke jurusan tehnik, dan siswa dengan nilai diatas 85 ke atas dapat masuk ke kedua jurusan ini. Namun terlepas dari hasil perhitungan diatas semua kembali lagi ke para siswa dan orang tua ingin melanjutkan ke dalam jurusan yang mana. Penelitian ini juga dilakukan atas dasar tujuan yaitu untuk menghasilkan aplikasi yang bisa memudahkan guru di Sekolah Kristen Makedonia Ngabang untuk mengetahui jurusan apa yang nantinya cocok bagi para siswa nya jika ingin melanjutkan pendidikannya sesuai dengan nilai yang mereka miliki.

## 5. SARAN

Penulis menyadari, dalam penelitian masih terdapat banyak kekurangan pada perancangan perangkat lunak data mining dalam menentukan pemilihan jurusan. Untuk itu, penulis memberikan saran untuk pengembangan selanjutnya agar kedepannya lebih baik lagi. Penulis menyarankan agar aplikasi ini tidak hanya dapat digunakan oleh guru saja tetapi juga dapat diakses oleh para siswa sehingga para siswa dapat langsung mengetahui jurusan apa yang cocok sesuai dengan nilai yang mereka miliki dan juga dapat memotivasi mereka untuk belajar lebih giat demi jurusan yang akan mereka tempuh nantinya. Dan untuk pengembangan aplikasi ini lebih lanjut, dapat diperbaiki pada penyederhanaan koding program algoritma K-Means dan memperbaiki dalam tampilan desain program. Dalam program ini proses pengulangan hanya dilakukan 2 kali saja dan untuk pengembangan lebih lanjutnya dapat diperbaiki sehingga proses pengulangan dapat dilakukan sampai 5 atau bahkan selebihnya sehingga dapat menghasilkan nilai yang jauh lebih akurat.

## DAFTAR PUSTAKA

- [1] Asroni, Adrian 2015. Penerapan Metode K-Means untuk Clustering Mahasiswa Berdasarkan Nilai Akademik dengan Weka Interface Studi Kasus pada Jurusan Teknik Informatika UMM Magelang.
- [2] Fina, Surya 2015. Penerapan K-Means Clustering pada Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Potensi Utama).
- [3] Andini, Akbar, Helda 2013. Penerapan Data Mining Untuk Mengolah Informasi Konsentrasi Keahlian Dengan Metode Clustering Pada Universitas Bina Darma.
- [4] Rosa, A.S., Shalahuddin, M. 2013. Rekayasa Perangkat Lunak : Terstruktur dan Berorientasi Objek. Informatika. Bandung.
- [5] Grossman, David A., Ophir Frieder. 2014. *Information Retrieval – Algorithms and Heuristics Second Edition*. Springer. The Netherlands.
- [6] Prasetyo, Eko. 2013. *Data Mining: Konsep dan Aplikasi menggunakan Matlab*, Yogyakarta: Andi Publisher.
- [7] Nugroho, A., 2010, *Rekayasa Perangkat Lunak Berbasis Objek dengan Metode USDP*, Andi, Yogyakarta.
- [8] Pressman, Roger. 2012. *Rekayasa Perangkat Lunak (Pendekatan Praktisi) Edisi 7 : Buku 1*. Yogyakarta : Andi Publishers.