

CLUSTERING DATA SISWA BERPRESTASI DENGAN MENGGUNAKAN METODE K-MEANS

(Studi Kasus di SMP Negeri 2 Kwandang)

Oleh

SUSANTO PARAMATA

T3116324

SKRIPSI

**Untuk memenuhi salah satu syarat ujian
guna memperoleh gelar sarjana**



**PROGRAM SARJANA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2020**

PENGESAHAN SKRIPSI

CLUSTERING DATA SISWA BERPRESTASI DENGAN MENGGUNAKAN METODE K-MEANS

(Studi Kasus di SMP Negeri 2 Kwandang)

Oleh

SUSANTO PARAMATA

T3116324

SKRIPSI

Untuk memenuhi salah satu syarat ujian
Guna memperoleh gelar Sarjana
Program Studi Teknik Informatika
Ini telah disetujui oleh Tim Pembimbing

Gorontalo, 07 Desember 2020

Pembimbing Utama

Pembimbing Pendamping

Sudirman Melangi S.Kom, M.Kom
NIDN : 0908017702

Sumarni S.Kom, M.Kom
NIDN : 0926018604

PERSETUJUAN SKRIPSI

CLUSTERING DATA SISWA BERPRESTASI DENGAN MENGGUNAKAN METODE K-MEANS

(Studi Kasus di SMP Negeri 2 Kwandang)

OLEH
SUSANTO PARAMATA
T3116324

Diperiksa oleh Panitia Ujian Strata Satu (S1)
Universitas Ichsan Gorontalo

1. Ketua Penguji
Irma Surya Kummala, S.Kom, M.Kom
2. Anggota
Suhardi Rustam, S.Kom, M.Kom
3. Anggota
Sarlis Mooduto, S.Kom, M.Kom
4. Anggota
Sudirman Melangi, S.Kom, M.Kom
5. Anggota
Sumarni, S.Kom, M.Kom

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun di perguruan tinggi lainnya.
2. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan dari Tim pembimbing.
3. Dalam karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan dalam naskah dan dicantumkan pula dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo.

Gorontalo, 07 Desember 2020

Yang Membuat Pernyataan

Susanto Paramata

ABSTRAK

Sekolah Menengah Pertama merupakan pendidikan formal pada jenjang pendidikan dasar. Pendidikan dan pembelajaran ditingkat SMP memberikan penekanan peletakan pondasi dalam menyiapkan generasi agar menjadi manusia yang mampu menghadapi era yang semakin berat. Perkembangan teknologi yang ada di SMP Negeri 2 Kwandang belum memiliki program aplikasi khusus untuk menunjang atau meningkatkan mengenai pengelompokan data siswa berprestasi. Berdasarkan latar belakang diatas penulis melakukan pengolahan data dengan metode *Clustering* dengan *Algoritma K-Means* sebagai pembanding, sehingga data tersebut dapat membantu pihak sekolah SMP Negeri 2 Kwandang. Dalam penelitian ini digunakan variabel dan variabel yang digunakan dalam perhitungan Algoritma K-Means yaitu NISN Siswa, Mata Pelajaran IPA, IPS, Matematika, Bahasa Indonesia, dan Bahasa Inggris. Sehingga dapat diperoleh hasil penelitian didapat 5 Cluster, sehingga dapat dikatakan penelitian ini berhasil untuk pengelompokan Data Siswa Berprestasi hal ini dibuktikan dengan pengujian *Whitebox* dimana $V(G) = 2 + 1 = 3$ CC = R1, R2, R3, R4, 45, R6, R7, R8, R9, R10, R11

Kata Kunci : Siswa, Nilai, Clustering K-Means, dan WhiteBox

ABSTRACT

Junior high school is formal education in basic education. Education and learning at the junior high school level emphasizes laying the foundation in preparing generations to become humans capable of facing an increasingly tough era. Technological developments in SMP Negeri 2 Kwandang do not have a special application program to support or improve the grouping of outstanding student data. Based on the above background, the authors perform data processing with the clustering method with the K-Means algorithm as a comparison, so that the data can help the SMP Negeri 2 Kwandang school. In the study, variables and variables used in the calculation of the K-Means algorithm are used, namely : NISN students, science subjects, social studies, mathematics, Indonesian, and English. So it can be obtained the results of the study obtained 5 clusters, so it can be said that this study was successful for grouping outstanding student data, this is evidenced by white box testing where $V(G) = 2 + 1 = 3$ $CC = R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, R11$

Keywords : Students, Grades, K-Means Clustering and Whitebox

KATA PENGANTAR

Puji syukur penulis panjatkan kepada kehadiran ALLAH S.W.T, karena atas berkat dan Rahmat-nya penulis dapat menyelesaikan Skripsi ini dapat diselesaikan pada waktunya.

Penulis menyadari bahwa dalam penulisan Skripsi ini masih jauh dari kesempurnaan. Oleh karena itu dengan segala kerendahan hati penulis sangat mengharapkan kritik dan saran yang sifatnya membangun guna perbaikan dan penyempurnaannya.

Pada kesempatan yang sangat berharga ini penulis haturkan ucapan terima kasih yang setinggi-tingginya kepada :

1. Bapak Muhamad Ichsan Gaffar Latjokke, S.E M.Ak., selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo.
2. Bapak Dr. Abdul Gaffar Latjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo.
3. Ibu Zohrahayati S.Kom, M.Kom selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.
4. Bapak Irvan Abraham Salihi S.Kom, M.Kom selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.
5. Bapak Sudirman Melangi S.Kom, M.Kom selaku Wakil Dekan III di bidang kemahasiswaan, Serta dosen pembimbing utama, yang juga telah membimbing dan telah memberikan saran pada pembuatan usulan penelitian ini.
6. Ibu Sumarni S.Kom, M.Kom selaku dosen pendamping, yang juga telah memberikan banyak motivasi kepada penulis.
7. Kedua orang tua, Ayah dan ibu, serta kakak dan adik, yang juga telah memberikan doa, motivasi, dan dukungan.
8. Segenap dosen Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.
9. Rekan-rekan Program Study Teknik Informatika, Kampus 3 Universitas Ichsan Gorontalo.

10. Teman-teman dari Squad Goroho dan X-9 Squad yang telah memberikan motivasi dan kritikan kepada penulis sehingga usulan penelitian ini bisa selesai.
11. Dan semua pihak yang tidak dapat disebut satu per satu yang telah membantu penulisan usulan penelitian ini.

Penulis juga menyadari bahwa Skripsi ini masih banyak kekurangan baik didalam penyusunan maupun penulisan. Oleh karena itu saran dan kritik sangat penulis harapkan demi kesempurnaan penelitian berikutnya.

Akhirnya penulis hanya berharap semoga Skripsi ini dapat bermanfaat bagi kita semua . Amin

Gorontalo,07 Desember 2020

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
HALAMAN PERNYATAAN.....	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masaalah	4
1.3 Rumusan Masaalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat penelitian	5
1.5.1 Manfaat Teoritis.....	5
1.5.2 Manfaat Praktis	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Tinjauan Studi.....	6
2.2 Tinjauan Pustaka.....	8
2.2.1 Siswa Berprestasi	8
2.2.2 <i>Clustering</i>	9
2.2.3 <i>Data Mining</i>	9
2.2.4 <i>Algoritma K-Means</i>	12
2.2.5 Penerapan <i>Algoritma K-Means</i>	13
2.2.6 Siklus pengembangan sistem	17
2.2.7 Analisa Sistem	18
2.2.8 Desain sistem	19

2.2.9 Perancangan konseptual.....	23
2.2.10 Perancangan fisik.....	23
2.2.11 Implementasi sistem.....	29
2.2.12 Operasi dan pemeliharaan	30
2.2.13 Teknik pengujian sistem	31
2.3 Framework	33
2.4 Tools Pendukung	34
2.5 Kerangka Pikir	35
BAB III METODOLOGI PENELITIAN	36
3.1 Objek penelitian	36
3.2 Metode penelitian.....	36
3.2.1 Tahap analisis.....	36
3.2.2 Tahap desain	38
3.2.3 Tahap pembuatan	39
3.2.4 Tahap Pengujian sistem	40
3.2.5 Tahap implementasi	40
BAB IV HASIL PENELITIAN.....	41
4.1 Hasil Pengumpulan Data.....	41
4.2 Hasil Pemodelan	41
4.2.1 Iterasi 1.....	46
4.2.2 Iterasi 2.....	47
4.3 Analisa Sistem	47
4.3.1 Sistem yang berjalan	47
4.3.2 Sistem yang diusulkan	48
4.4 Hasil Pengembangan Sistem.....	48
4.5 Kamus Data.....	52
4.6 Desain Database	55
4.7 Arsitektur Sistem	56
4.8 Interface Desain Mekanisme User	56
4.9 Interface desain Input.....	57
4.10 Tampilan Window Input Data Centroid	57

4.11 Program Desain.....	58
4.12 Hasil Konstruksi Sistem.....	59
4.13 Kode Program Untuk Pengujian Whitebox	59
4.14 Flowchart Program untuk pengujian Whitebox.....	60
4.15 Flowgraph program untuk pengujian Whitebox	61
4.16 Perhitungan CC pada perhitungan Whitebox	61
4.17 Path pada pengujian Whitebox	62
4.18 Hasil Pengujian Blackbox.....	63
BAB V PEMBAHASAN PENELITIAN	65
5.1 Pembahasan Sistem.....	65
5.1.1 Hasil Tampilan Window Login Admin	65
5.1.2 Hasil Tampilan Window Beranda Admin.....	65
5.1.3 Hasil Tampilan Window Beranda Pimpinan	66
5.1.4 Hasil Tampilan Window Data Nilai Siswa.....	66
5.1.5 Hasil Tampilan Window Input Data Centroid	67
5.1.6 Hasil Tampilan Window Iterasi <i>K-Means</i>	67
5.1.7 Hasil Tampilan Window Hasil <i>Clustering</i>	68
BAB VI PENUTUP	69
6.1 Kesimpulan	69
6.2 Saran	69
DAFTAR PUSTAKA	70
LAMPIRAN	72

DAFTAR GAMBAR

Gambar 2.1 Tahapan <i>Knowledge Discovery In Database</i>	10
Gambar 2.2 Siklus hidup pengembangan sistem	18
Gambar 2.3 Contoh notasi kesatuan luar	28
Gambar 2.4 Contoh notasi arus data	28
Gambar 2.5 Contoh notasi proses	28
Gambar 2.6 Contoh notasi simpanan data	29
Gambar 2.7 Contoh Diagram alir	32
Gambar 2.8 Kerangka Pemikiran	35
Gambar 3.1 Sistem yang diusulkan	38
Gambar 4.1 Bagan Alir Dokumen	47
Gambar 4.2 Bagan Alir Sistem Yang Diusulkan	48
Gambar 4.3 Diagram Konteks	48
Gambar 4.4 Diagram Berjenjang	49
Gambar 4.5 Arus Data DAD Level 0	49
Gambar 4.6 Arus Data DAD Level 1 Proses 1	50
Gambar 4.7 Arus Data DAD Level 1 Proses 2	51
Gambar 4.8 Arus Data DAD Level 1 Proses 3	51
Gambar 4.9 Tampilan Login Admin	55
Gambar 4.10 Desain Input Data	56
Gambar 4.11 Flowchart Proses Pengujian Whitebox	60
Gambar 4.12 Flowgraph Program Pengujian Whitebox	61
Gambar 5.1 Hasil Tampilan Window Login Admin	65
Gambar 5.2 Hasil Tampilan Beranda Admin	65
Gambar 5.3 Hasil Tampilan Beranda Pimpinan	66
Gambar 5.4 Hasil Tampilan Data Nilai Siswa	66
Gambar 5.5 Tampilan Window Input Data Centroid	67
Gambar 5.6 Tampilan Window <i>Iterasi K-Means</i>	67
Gambar 5.7 Tampilan Window Hasil Cluster	68

DAFTAR TABEL

Tabel 1.1 Data Set Siswa Berprestasi	3
Tabel 2.1. <i>State Of The Art</i>	6
Tabel 2.2 Data Seleksi.....	14
Tabel 2.3. Data Hasil Perhitungan Jarak.....	15
Tabel 2.4 Hasil Perhitungan jarak Terdekat.....	16
Tabel 2.5 Hasil Perhitungan Pusat Cluster baru	17
Tabel 2.6 Bagan Alir Sistem	25
Tabel 2.7 Tools Pendukung.....	34
Tabel 4.1 Hasil Pengumpulan Data.....	41
Tabel 4.2 Data Nilai Siswa	42
Tabel 4.3 Contoh Sampel Data Nilai Siswa	42
Tabel 4.4 Hasil Iterasi 1	46
Tabel 4.5 Hasil Iterasi 2	47
Tabel 4.6 Rancangan File Secara Umum.....	56
Tabel 4.7 Interface Design	56
Tabel 4.8 Tampilan Input Window Input Login	57
Tabel 4.9 Desain Program.....	57
Tabel 4.10 Path Pengujian <i>White Box</i>	62
Tabel 4.11 Hasil Pengujian <i>Black Box</i>	63
Tabel 5.1 Hasil Clustering.....	71

BAB I

PENDAHULUAN

1.1 Latar Belakang

Suatu bangsa menginginkan kemajuan dibidang pendidikan, maka harus ada upaya untuk mengembangkan potensi dan bakat dari peserta didik. Untuk mengembangkan potensi dan bakat peserta didik, dilakukan melalui proses pembelajaran. Dengan proses pembelajaran yang baik dan berkualitas, peserta didik akan mendapatkan pengetahuan, keterampilan serta bekal untuk menghadapi berbagai kemajuan dan tantangan zaman. Seiring dengan kemajuan zaman, berkembang pula teori-teori pembelajaran. Teori pembelajaran ini, dapat digunakan sebagai bekal oleh guru untuk memperbaiki proses pembelajaran yang dilakukan sehingga akan terciptanya pembelajaran yang lebih efektif.

SMP (Sekolah Menengah Pertama) merupakan pendidikan formal pada jenjang pendidikan dasar. Pendidikan dan pembelajaran di tingkat SMP memberikan penekanan peletakan pondasi dalam menyiapkan generasi agar menjadi manusia yang mampu menghadapi era yang semakin berat. Menurut Undang-Undang Sistem Pendidikan Nasional no 2 tahun 2003 pasal 17 tentang pendidikan dasar disebutkan bahwa pendidikan dasar terdiri dari SD (Sekolah Dasar)/Sederajat dan SMP (Sekolah Menengah Pertama)/sederajat.

Mengetahui tingkat kemampuan dan pengetahuan seseorang dalam ilmu pengetahuan umum seperti mata pelajaran IPA, IPS, Matematika, Bahasa Indonesia, dan Bahasa Inggris bukan hal yang mudah bagi dirinya sendiri, orang tua, maupun pihak lembaga. Mengingat banyaknya siswa yang sepiantas memiliki kemampuan yang sama dengan siswa yang lain, serta adanya berbagai macam tingkatan kemampuan dan pertimbangan serta perhitungan nilai terhadap aspek-aspek dari beberapa hasil uji kompetensi, maka hal tersebut menjadi menarik untuk dikaji sehingga berdasarkan kondisi itulah muncul suatu inspirasi untuk membangun suatu upaya sederhana dan cepat untuk mengetahui siswa mana saja yang memiliki prestasi unggul hingga rendah terhadap mata pelajaran umum.

Metode analisis *Clustering* adalah salah satu metode analisis utama dalam sebuah data *Mining*. Metode *Algoritma Cluster* akan mempengaruhi hasil *Clustering* secara berlangsung. Penelitian ini membahas *Algoritma Clustering K-Means* standar dan menganalisis kekurangan dari *Algoritma K-Means* standar. Seperti *Algoritma Cluster K-Means* harus menghitung jarak antara setiap objek data dan semua pusat kluster dalam setiap iterasi yang membuat efisiensi pengelompokan tidak tinggi. Penelitian ini mengusulkan *Algoritma K-Means* yang ditingkatkan untuk mengatasi ini. (Shi Na, Liu Xumin, Guan Yong. *Algoritma Clustering K-Means* 2010).

Istilah *Data Mining* memiliki beberapa pandangan seperti *Knowledge Discover* ataupun *Pattern Recognition*. Kedua istilah tersebut sebenarnya memiliki ketepatannya masing-masing, istilah *Knowledge Discovery* atau pengetahuan tepat karena digunakan tujuan utama dari data *Mining* memang untuk mendapatkan pengetahuan yang masih tersembunyi didalam bongkahan data. (Berry, M. W and Browne, M. 2006. *Lecture notes in data mining*. World Scientific. (Susanto, S. And Suryadi, D. , 2010. *Pengantar data mining : menggali pengetahuan dari bongkahan data*).

Metode *K-Means* berusaha mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Cynthia Paramanda, *Analisis Data Konsumen Dengan Menggunakan Metode Clustering Pada Toko Bangunan Rahayu*, 2015).

Aturan dan metode *Algoritma K-Means* dapat diterapkan pada sebuah program bantu untuk mengelompokkan data nilai siswa. Langkah-langkah *Algoritma K-Means* diterjemahkan dalam bahasa pemrograman untuk melakukan tugas kluster data. Data set nilai siswa dimasukan ke input program, kemudian program melakukan pengolahan sesuai langkah *Algoritma K-Means*, dan hasilnya berupa *Clustering* data. Hasil berupa data *Clustering* inilah digunakan sebagai bahan pertimbangan untuk pembuatan kebijakan pembelajaran disekolah.

Tabel 1.1 Dataset Siswa Berprestasi

No	NIS	IPA	IPS	Matematika	Bahasa Indonesia	Bahasa Inggris
1	0039293047	88	86	84	86	86
2	0041208584	88	89	90	88	87
3	3046559563	84	82	87	80	86
4	0055463313	90	93	92	91	87
5	3052009270	80	80	85	86	86
6	0048739784	87	85	81	82	87
7	0053578343	92	92	93	91	86
8	0052753461	90	91	90	87	88
9	0053577836	88	88	82	90	92
10	0056570925	89	88	90	93	87
...	
460		88	85	88	89	87

(Sumber : SMP N 2 Kwandang 2016 – 2018).

Dalam penelitian ini akan melakukan pengelompokan data siswa dengan menggunakan *Algoritma K-Means*, *Algoritma K-Means* dipilih dalam melakukan pengelompokan karena algoritma tersebut dapat mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya, sehingga nantinya pihak sekolah SMP Negeri 2 Kwandang dapat membantu dalam pengelompokan data siswa berprestasi. Kelebihan dari *Algoritma K-Means* itu sendiri yaitu mudah di implementasikan. Adapun variabel yang digunakan dalam penelitian ini yaitu : Mata Pelajaran IPA, IPS, Matematika, Bahasa Indonesia, dan Bahasa Inggris.

Pada penelitian sebelumnya yang dilakukan oleh Nisia Yuanita 2016 membahas tentang pengelompokan kelas di MTs Sunan Ampel dengan menggunakan *Algoritma K-Means*. Pada penelitian ini dalam membagi kelas siswa masih menggunakan sistem manual sehingga pembagian kelas menjadi tidak

merata. Kesimpulan penelitian ini adalah dihasilkan sebuah aplikasi untuk pengelompokan kelas murid menggunakan metode *K-Means*. (Nisia Yuanita 2016 Implementasi *K-Means Clustering* Untuk Pembagian Kelas Siswa).

Berdasarkan uraian diatas, maka di anggap perlu untuk melakukan penelitian mengenai proses yang berjalan diatas, dengan judul **“Clustering Data Siswa Berprestasi dengan Menggunakan Metode K-Means (Studi Kasus SMP Negeri 2 Kwandang)”**.

1.2 Identifikasi Masaalah

1. Pihak Sekolah SMP Negeri 2 Kwandang masih sulit dalam mengelompokan siswa berprestasi
2. Belum adanya metode yang digunakan dalam pengelompokan data siswa berprestasi pada SMP Negeri 2 Kwandang.

1.3 Rumusan Masaalah

Berdasarkan uraian diatas, dapat dirumuskan masaalah pokok yang berkaitan yaitu :

1. Bagaimana membangun metode dalam pengklusteran siswa berprestasi menggunakan *Algoritma K-Means*.
2. Bagaimana mengimplementasikan *Algoritma K-Means* dalam pengelompokan data siswa berprestasi pada SMP Negeri 2 Kwandang.

1.4 Tujuan Penelitian

1. Untuk membangun metode dalam pengelompokan data menggunakan *Algoritma K-Means*.
2. Untuk mengimplementasikan *Algoritma K-Means* dalam pengelompokan siswa berprestasi pada SMP Negeri 2 Kwandang.

1.5 Manfaat Penelitian

Penelitian ini diharapkan mempunyai manfaat yaitu :

1.5.1 Manfaat teoritis

Memberi masukan pada perkembangan ilmu pengetahuan dan teknologi, khususnya pada bidang kajian data *Mining*, yaitu berupa pemutakhiran metode *K-Means*.

1.5.2 Manfaat Praktis

Sebagai sumbangan pemikiran atau solusi bagi *Software Developer* guna mendukung pengambilan keputusan dalam rangka menghasilkan *Software* yang berkualitas.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Penelitian tentang pengelompokan menggunakan *Algoritma K-Means* telah dilakukan oleh beberapa peneliti antara lain :

Tabel 2.1 *State Of The Art*

No	Nama Peneliti	Judul Penelitian	Tahun	Metode	Hasil
1	Nisa Yuanita	Implementasi <i>K-Means Clustering</i> Untuk Pembagian Kelas Siswa	2016	<i>K-Means</i>	Berdasarkan penelitian yang dilakukan di MTs Sunan Ampel, maka dihasilkan sebuah aplikasi untuk pengelompokan kelas siswa Variabel yang digunakan adalah nilai UTS pada semester gasal meliputi bidang studi matematika, bahasa indonesia, bahasa inggris,

					dan ilmu pengetahuan alam
2	Teguh Hariyadi	Penerapan <i>Algoritma K-Means</i> Untuk Pengelompokan Data Nilai Siswa	2013	<i>K-Means</i>	Program bantu <i>Algoritma K-Means</i> dapat digunakan untuk mengelompokkan data menjadi beberapa klaster dan hasil pengelompokan tersebut dapat digunakan sebagai pengetahuan untuk mengambil suatu kebijakan terkait nilai siswa
3	Ernie Kurniawan	Penerapan <i>Algoritma K-Means</i> Untuk <i>Clustering</i> dokumen jurnal STIMIK GI MDP	2016	<i>K-Means</i>	Menghasilkan kelompok-kelompok yang sesuai dengan <i>Query</i> . Dalam prosesnya dilakukan tahap <i>Preprocessing</i> yaitu

					<i>Tokennization</i> penghilangan <i>Stopwords</i> dan <i>Stemming</i> .
--	--	--	--	--	---

2.2 Tinjauan Pustaka

2.2.1 Siswa Berprestasi

Prestasi merupakan hasil sebuah usaha yang tidak selamanya identik dengan hasil baik. Misalnya mendapatkan nilai lima, bisa dikatakan memperoleh prestasi buruk atau rendah. Prestasi seseorang sangat dipengaruhi oleh dua macam faktor yang berasal dari luar dirinya, yaitu :

1. faktor dari dalam diri, diantaranya bakat atau potensi, kependian, minat, kebiasaan, motivasi, pengalaman, kesehatan dan emosi.
2. faktor dari luar, misalnya keluarga, sekolah, masyarakat, sarana prasarana, fasilitas, gizi, dan tempat tinggal.

Kedua jenis faktor tersebut mendukung satu sama lain. Prestasi biasanya akan muncul bila kedua macam faktor diatas ter[enuhi secara baik.

Orang yang berprestasi adalah orang yang dianggap sukses dalam bidang tertentu, karena pada kenyataan iya memiliki kelebihan-kelebihan yang tidak dimiliki oleh orang lain.

Berprestasi adalah idaman setiap banyak orang, baik itu berprestasi dalam bidang pekerjaan, pendidikan, seni, sosial, politik, budaya dan lain-lain. Dengan adanya prestasi yang pernah diraih oleh seseorang akan menumbuhkan suatu semangat baru untuk menjalani sautu aktifitas.

Berikut adalah ciri-ciri seseorang siswa yang berprestasi :

1. Selalu mempunyai tujuan dan rencana
2. Mengambil resiko yang wajar dan diperhitungkan
3. Bertanggung jawab secara pribadi
4. Mengerjakan sesuatu yang kreatif dan inovatif
5. Selalu menggunakan umpan-balik
6. Merasa dikejar waktu

7. Menyukai situasi yang serba mungkin/beraneka ragam
8. Berinisiatif dan suka menyelidiki lingkungan
9. Berhubungan dengan orang lain, tidak hanya berteman

2.2.2 Clustering

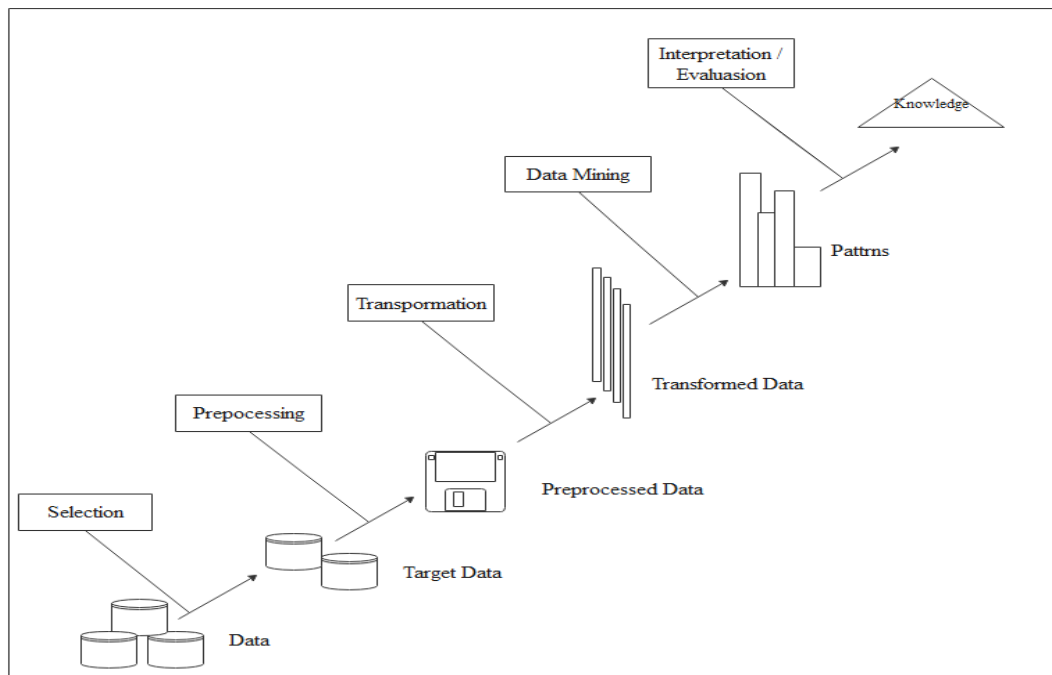
Pengclustering data adalah teknik data *Mining* yang digunakan untuk meredaksi data atau menyederhanakan data kedalam kelompok-kelompok yang mempunyai sifat / kareteristik yang mirip. Pengelompokan data telah banyak digunakan dalam berbagai bidang dengan tujuan yang bermacam-macam seperti pengelompokan pola, images, analisis, dan sebagainya metode yang dikelompokkan adalah *K-means* dimana metode ini akan mengalokasikan data kedalam kelompok-kelompok sesuai banyak kelompok yang diinputkan oleh pengguna. (Tinus Septiko, *Aplikasi K-Means Clustering Pada Data rumah tangga* 2013).

Pembentukan *Cluster* merupakan salah satu teknik yang digunakan dalam mengekstrak pola kecenderungan suatu data. Teknik ini digunakan dalam proses *Knowledge Discovery Database* (KDD). Data *Mining* biasanya identik dengan proses penggalian data-data yang cukup besar dan dikelompokkan menjadi data yang tersusun rapi. (Fina Natasari, *penerapan K-means Clustering Pada Data Penerimaan Mahasiswa Baru* (studi kasus : Universitas Potensi Utama, 2015).

2.2.3 Data Mining

Data Mining yang juga memiliki batasan waktu data tak berguna, data yang di simpan dengan jarak terpisah dari data operasional sebagai data Warehouse (Gudang Data). Sehingga jenis data ini dapat kategorisasi dalam management data di bagi dua adalah OLTP (*On Line Transaction Processing*) yaitu management data operational kemudian data dimanfaatkan merupakan data terkini untuk yang akan datang. (Suhardi Rustam, *Penerapan Optimalisasi Jumlah Kluster Pada K-Means Untuk Pengelompokan Kelas Mata Kuliah Kosentrasi Mahasiswa Semester Akhir*, 2020)

Istilah data mining dan *knowledge discovery in databases* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersumbunyi dari basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahap dalam keseluruhan proses KDD adalah *data mining*. Proses KDD (*Knowledge discovery in database*) secara garis besar dapat dijelaskan sebagai berikut. Larose, D. T. *Discovering Knowledge in data : an introduction to data mining*, Hoboken, New Jersey: John Wiley & Sons, Inc. 2005).



Gambar 2.1 Tahapan *Knowledge Discovery in Database*

(Sumber : Larose, D. T. Hoboken, New Jersey. John Wiley & Sons, Inc. 2005)

1. *Data selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge data discovery* (KDD) dimulai. Data hasil seleksi yang digunakan untuk proses data mining, disimpan dalam bentuk berkas, terpisah dari basis data operasional.

2. *Preprocessing* atau *cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *knowledge data discovery*. Proses *cleaning* mencakup proses antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak juga dilakukan proses *enrichment*, yaitu proses memperkaya data yang sudah ada dengan data atau informasi yang lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi.

3. *Transformation*

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam *knowledge data discovery* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining*

Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation* atau *evaluation*

Pola informasi yang dihasilkan dari proses *data mining* perlu di tampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada pada sebelumnya.

Data mining di bagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Teguh Hariyadi, Penerapan Algoritma *K-Means* Untuk Pengelompokan Data Nilai Siswa, 2013) :

1. Deskripsi

Deskripsi adalah menggambarkan pola yang kecenderungan yang terdapat dalam data yang memungkinkan memberikan penjelasan dari suatu pola atau kecenderungan tersebut.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik dari pada kategori. Model dibangun menggunakan record lengkap yang menyediakan variabel target sebagai nilai prediksi.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, akan tetapi dalam prediksi nilai dari hasil di masa mendatang.

4. Clustering

Clustering adalah proses untuk menemukan model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep dengan tujuan memprediksikan kelas untuk data yang tidak diketahui kelasnya.

5. Pengklusteran

Pengklusteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam cluster lain.

6. Asosiasi

Asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

2.2.4 Algoritma K-Means

Algoritma K-means merupakan *Algoritma Clustering* yang paling mudah dan paling banyak digunakan. *Algoritma K-means* adalah algoritma pengelompokan sederhana iteratif yang membagi data set yang diberikan menjadi sejumlah kluster K yang ditentukan oleh pengguna. Algoritma ini sederhana untuk diterapkan dan dijalankan, relatif cepat, mudah diadaptasi dan umum digunakan dalam praktek. Secara historis, algoritma *K-means* merupakan salah satu algoritma yang paling penting dalam *data mining*. (Larose, D. T. *Discovering Knowledge in data : an introduction to data mining*, Hoboken, New Jersey: John Wiley & Sons, Inc. 2005).

Adapun rumus perhitungan jarak didefinisikan sebagai berikut :

$$d(x,y) = \sqrt{(x_i - y_i)^2 + (x_i - y_i)^2}$$

Keterangan : D = titik dokumen

X = data *record*

Y = data *centroid*

Tahapan pada *Algortima K-means* :

1. Tentukan beberapa banyak klaster K dari dataset yang akan dibagi.
2. Tetapkan secara acak data K menjadi pusat awal lokasi klaster.
3. Untuk masing-masing data, temukan pusat klaster terdekat. Dengan demikian berarti masing-masing pusat klaster memiliki sebuah subset dari dataset, sehingga mewakili bagian dari dataset. Oleh karena itu, telah terbentuk klaster k : C1, C2, C3, Ck.r
4. Untuk masing-masing klaster , temukan pusat luasan klaster, dan perbaharui lokasi dari masing-masing pusat klaster ke nilai baru dari pusat luasan.

Ulangi langkah ke-3 dan ke-5 sehingga data-data pada tiap klaster menjadi terpusat atau selesai.

2.2.5 Penerapan *Algoritma K-Means*

Dalam hal ini penelitian oleh (Amri Muhamad, 2016 , tentang pengelompokan siswa unggulan dengan *Metode K-Means Clustering*) adapun atribut yang digunakan untuk menyeleksi siswa adalah Nilai matematika, Nilai IPA, Nilai bahasa inggris, Nilai bahasa indonesia, Nilai agama:

Diketahui :

Jumlah *Cluster* = 3

Jumlah Data = 13

Jumlah atribut = 5

Tabel 2.2 Data seleksi

No	Matematika	IPA	Bahasa Inggris	Bahasa Indonsia	Agama	Status
1	87	83	92	87	93	Lolos
2	88	75	87	90	89	Tidak
3	92	91	86	88	95	Lolos
4	75	82	90	89	92	Lolos
5	73	72	87	80	92	Tidak
6	74	80	88	81	90	Tidak
7	75	80	88	89	90	Lolos
8	87	79	83	87	91	Tidak
9	91	80	90	89	92	Lolos
10	81	72	86	75	92	Tidak
11	73	86	88	88	93	Tidak
12	73	78	85	85	90	Tidak
13	76	72	82	77	88	Tidak

(Sumber : Amri Muhammad, 2016)

a. Menentukan jumlah kelas dan pusat *Cluster*

Kelas yang akan digunakan berjumlah 3 kelas dan pusat *Cluster* akan ditentukan secara acak.

b. Penentuan nilainya centroid :

1. $c_1 = (83, 80, 84, 84, 93)$

2. $c_2 = (81, 85, 87, 92, 92)$

3. $c_3 = (75, 97, 95, 94, 93)$

c. menghitung jarak data yang ada pada setiap pusat *Cluster*

Adapun rumus *K-Means Clustering* sebagai berikut :

$$d(x,y) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots (p_n - q_n)^2}$$

1. C1,1

$$= \sqrt{(87 - 83)^2 + (83 - 80)^2 + (87 - 84)^2 + (92 - 84)^2 + (93 - 93)^2}$$
$$= 0$$

2. C1, 2

$$= \sqrt{(92 - 83)^2 + (91 - 80)^2 + (88 - 84)^2 + (86 - 84)^2 + (95 - 93)^2}$$
$$= 11.9$$

3. C1,3

$$= \sqrt{(75 - 83)^2 + (85 - 80)^2 + (89 - 84)^2 + (90 - 84)^2 + (92 - 93)^2}$$
$$= 6.24$$

Dari hasil perhitungan *Cluster* tersebut, sebagai berikut :

Tabel 2.3 Hasil perhitungan jarak

No	c1	c2	c3
1	9.899495	6.4031242	20.19901
2	15.27889	14.391355	20.9311
3	12.28821	6.6332496	14.03567
4	10.68	9.0863909	19.0148
5	11.22992	11.392005	24.2235
6	12.56096	10.556725	17.73258
7	8.062258	8.4852814	19.46792
8	10.11187	8.3815273	21.33659
9	11.57584	12.767145	19.0263
10	18.43306	14.813657	16.4452
11	16.55295	14.035669	19.49359
12	3.31625	10.488088	24.02082
13	14.10674	8.3666003	15.45962

Menemukan *Cluster* dengan jarak terdekat pada masing-masing data.

Adapun hasilnya tampilkan pada tabel dibawah ini :

Tabel 2.4 Hasil perhitungan jarak terdekat

No	C1	C2	C3	C1	C2	C3
1	9.899495	6.4031242	20.19901		*	
2	15.27889	14.391355	20.9311		*	
3	12.28821	6.6332496	14.03567		*	
4	10.68	9.0863909	19.0148		*	
5	11.22992	11.392005	24.2235	*		
6	12.56096	10.556725	17.73258		*	
7	8.062258	8.4852814	19.46792	*		
8	10.11187	8.3815273	21.33659		*	
9	11.57584	12.767145	19.0263	*		
10	18.43306	14.813657	16.4452		*	
11	16.55295	14.035669	19.49359		*	
12	3.316625	10.488088	24.02082	*		
13	14.10674	8.3666003	15.45962		*	

Menghitung pusat *Cluster* baru

Untuk menghitung *cluster* selanjutnya, ditentukan dari nilai dari setiap mata pelajaran dibagi dengan jumlah anggota yang ada di C1 untuk pusat *Cluster* 1, C2 untuk pusat *Cluster* 2 dan C3 untuk pusat *cluster*. Adapun cara menghitungnya adalah sebagai berikut :

$$C1, 1 = (75+80+70+78+79+81+70+72+69+70+68+70+)/19$$

$$C2, 1 = (68+69+70+68+68+65+69+68+70+70+68+)/62$$

$$C3, 1 = (83+82+80+79+81+82+85+78+85+86+88+87+80+)/9$$

Hasilnya sebagai berikut :

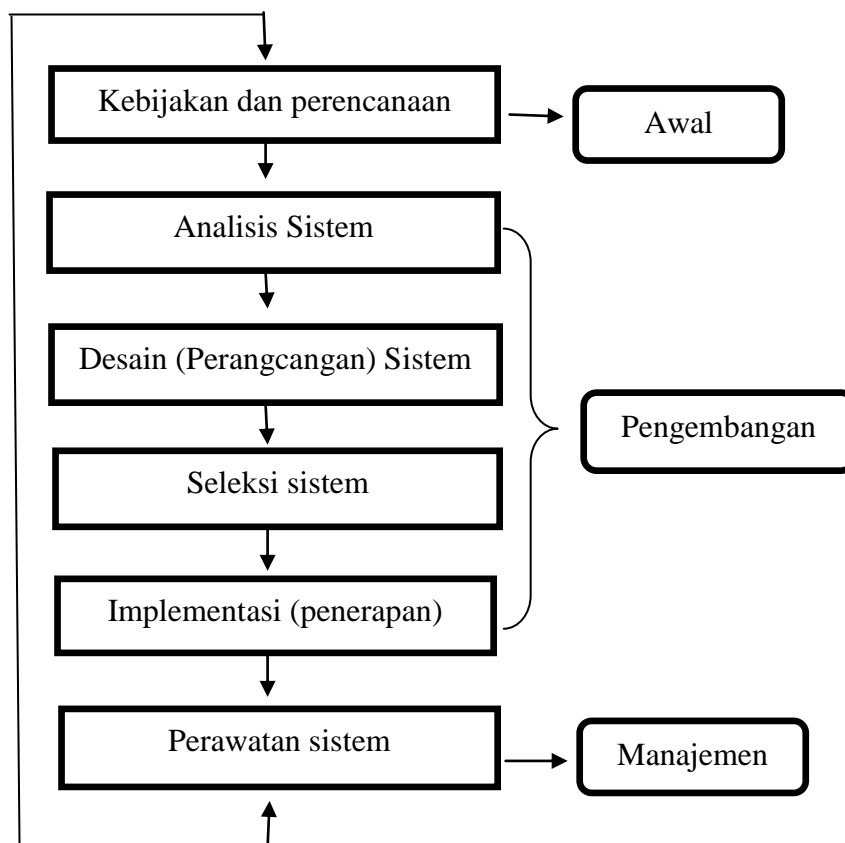
Tabel 2.5 Hasil perhitungan pusat *cluster* baru.

c1	83.5263	82.3649	87.0526	84.8947	91.631578
c2	82.0806	86.4715	87.7983	89.7258	91.887096
c3	77.4444	95.1644	91.6666	87.8888	93.888888

Ulangi langkah 3 (Menghitung jarak data yang ada terhadap setiap pusat *Cluster*) 4 (Menentukan *cluster* dengan jarak terdekat pada masing-masing data) 5 (Menghitung pusat *cluster* baru) sampai posisi data pada setiap *cluster* sudah tidak mengalami perubahan.

2.2.6 Siklus Hidup Pengembangan Sistem

Pengembangan sistem informasi yang berbasis komputer dapat merupakan tugas kompleks yang membutuhkan banyak sumber daya yang dapat memakan waktu berbulan-bulan bahkan bertahun-tahun untuk menyelesaikannya. Proses pengembangan sistem melewati beberapa tahapan dari mulai sistem itu direncanakan sampai dengan sistem tersebut diterapkan, dioperasikan, dan dipelihara. Bila operasi sistem yang telah dikembangkan masih timbul kembali permasalahan-permasalahan yang kritis serta tidak dapat diatasi dalam tahap pemeliharaan sistem untuk mengatasinya dari proses ini kembali ketahap yang pertama, yaitu tahap perencanaan sistem. Siklus ini disebut dengan siklus hidup suatu sistem (*systems life cycle*). Siklus hidup pengembangan sistem dengan langkah-langkah utamanya yang akan digunakan sebagai berikut. (Jogianto, HM, 1999 Analisis Dan Desain Sistem Informasi : Pendekatan terstruktur teori dan praktek aplikasi bisnis edisi II Yogyakarta : Andi Yogyakarta).



(Sumber : Jogianto, HM. 1999)

Gambar 2.2 Siklus Hidup Pengembangan Sistem

2.2.7 Analisa Sistem

Analisa Sistem sebagai berikut : “Analisa Sistem (*System analysis*) sebagai penguraian dari suatu sistem informasi yang utuh kedalam bagian-bagian komponennya yang dimaksud untuk mengidentifikasi dan mengevaluasi permasalahan-permasalahan, kesempatan-kesempatan, hambatan-hambatan yang terjadi dan kebutuhan-kebutuhan yang diharapkan sehingga dapat diusulkan perbaikan-perbaikannya. (Nisia Yuanita 2016 Implementasi *K-Means Clustering* Untuk Pembagian Kelas Siswa).

Tahap analisis merupakan tahap yang kritis dan sangat penting, karena kesalahan didalam tahap ini akan menyebabkan juga kesalahan ditahap selanjutnya. Tahap analisa sistem mencakup studi kelayakan dan analisis kebutuhan.

Didalam tahap analisis sistem terdapat langkah-langkah dasar yang harus dilakukan oleh analisis sistem, yaitu sebagai berikut :

1. *Identify*, yaitu mengidentifikasi masalah.

Mengidentifikasi (Mengenai) masalah merupakan langkah pertama yang dilakukan dalam tahap analisis sistem. Masalah (*Problems*) dapat didefinisikan sebagai suatu pertanyaan yang diinginkan untuk dipecahkan.

2. *Understand*, yaitu memahami kerja dari sistem yang ada.

Langkah kedua dari analisis sistem adalah memahami kerja dari sistem yang ada. Langkah ini dapat dilakukan dengan mempelajari secara terinci bagaimana sistem yang beroperasi. Untuk mempelajari operasi dari sistem ini diperlukan data yang dapat diperoleh dengan cara melakukan penelitian.

3. *Analyze*, yaitu menganalisis sistem tanpa *report*.

Langkah ini dilakukan berdasarkan data yang telah diperoleh dari hasil penelitian yang telah dilakukan.

4. *Report*, yaitu membuat laporan hasil analisis

Tujuan utama dari pembuatan laporan hasil analisis :

- a. Pelaporan bahwa analisis telah selesai dilakukan
- b. Meluruskan kesalahan pengertian mengenai apa yang telah di temukan dan dianalisis oleh analisis sistem tetapi tidak sesuai menurut manajemen.

2.2.8 Desain Sistem

Setelah tahap analisis sistem selesai dilakukan, maka analisis sistem telah mendapatkan gambaran dengan jelaskan apa yang harus dikerjakan. Tiba waktunya sekarang bagi analisis sistem untuk memikirkan bagaimana membentuk sistem tersebut. Tahap ini disebut dengan desain sistem (*System design*).

Desain sistem adalah tahap setelah analisis dari siklus pengembangan sistem ; pendefinisian dari kebutuhan-kebutuhan fungsional dan dari persiapan untuk rancang bangun untuk implementasi menggambarkan bagaimana suatu

sistem dibentuk. (Nisia Yuanita 2016 Implementasi *K-Means Clustering* Untuk Pembagian Kelas Siswa).

Tahap desain sistem mempunyai dua tujuan utama yaitu :

- a. Untuk memenuhi kebutuhan kepada pemakai sistem.
- b. Untuk memberikan gambar yang jelas dan rancang bangun yang lengkap kepada pemrogram komputer dan ahli-ahli teknis lainnya.

Desain sistem dapat dibagi dalam dua bagian yaitu desain sistem secara umum (*general system design*) dan desain sistem secara terinci (*detailed system design*).

1. Desain sistem secara umum (*General System Design*).

Tujuan dari desain sistem secara umum adalah untuk memberikan gambaran secara umum kepada *user* tentang sistem yang baru, yang mana merupakan persiapan dari desain sistem secara rinci. Desain secara umum dilakukan oleh analisis sistem untuk mengidentifikasi komponen-komponen sistem informasi yang akan didesain secara rinci oleh pemrogram komputer dan ahli teknik lainnya.

Pada tahap ini, komponen-komponen sistem informasi dirancang dengan tujuan untuk dikomunikasikan kepada *user*. Komponen sistem informasi yang didesain adalah model, *Input*, *database*, *output*, teknologi dan kontrol.

2. Desain sistem secara rinci (*detailed system design*)

- a. Desain *Input* terinci

Masukan merupakan awal dimulainya proses informasi bahan mental dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan oleh organisasi. Data hasil transaksi merupakan masukan untuk sistem informasi. Hasil dari sistem informasi tidak lepas dari data yang dimasukan. Desain *input* terinci dimulai dari desain dokumen dasar sebagai penangkap *input* yang pertama kali. Jika dokumen dasar tidak didesain dengan baik, kemungkinan *input* yang tercatat dapat salah bahkan kurang.

Fungsi dokumen dasar dalam penanganan arus data :

- 1). Dapat menunjukan macam dari data yang harus dikumpulkan.
- 2). Data dapat dicatat dengan jelas, konsisten, dan akurat.

3). Dapat mendorong lengkapnya data disebabkan data yang dibutuhkan disebutkan satu persatu didalam dokumen dasarnya.

b. Desain *output* terinci.

Desain *output* terinci dimaksudkan untuk mengetahui bagaimana dan seperti apa bentuk *output-output* dari sistem yang baru. Desain *output* terinci terbagi atas dua yaitu desain *output* berbentuk laporan media kertas dan desain *output* dalam bentuk dialog pada layar terminal.

➤ Desain *output* dalam bentuk laporan

Desain ini dimaksudkan untuk menghasilkan *output* dalam bentuk laporan dimedia kertas. Bentuk laporan yang paling banyak digunakan adalah dalam bentuk tabel dan berbentuk grafik atau bagan.

➤ Desain *output* dalam bentuk dialog layar terminal

Desain ini merupakan rancang bangun dari percakapan antara pemakai sistem (*user*) dengan komputer. Percakapan ini dapat terdiri dari proses masukan data ke sistem, menampilkan *output* informasi kepada *user* atau keduanya.

Beberapa strategi dalam membuat layar dialog terminal :

1. Dialog pernyataan / jawaban
2. Menu.

Menu banyak digunakan karena merupakan jalur pemakai yang mudah dipahami dan mudah

Basis data (*database*) merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan disimpanan luar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. *Database* merupakan salah satu komponen yang penting digunakan. Menu berisi beberapa alternatif atau option atau pilihan yang disediakan kepada *user*. Pilihan menu akan lebih baik bila dikelompokkan sesuai fungsinya.

- ✓ Desain *database* terinci disistem informasi karena berfungsi sebagai basis penyedia informasi bagi para pemakainya. Penerapan *database* dalam sistem informasi disebut *database system*.

Sistem basis data (*database system*) adalah satu sistem informasi yang mengintegritaskan kumpulan dari data yang saling berhubungan satu dengan yang lainnya dan membuatnya tersedia untuk beberapa aplikasi yang bermacam-macam di dalam suatu organisasi. Dalam sistem basis data, tiap-tiap orang atau bagian dapat memandang *database* dari beberapa sudut pandang yang berbeda.

Pada tahap ini, desain *database* dimaksudkan untuk mendefinikan isi atau struktur dari tiap-tiap file yang telah didefinisikan didesain secara umum.

✓ Desain teknologi.

Tahap desain teknologi terbagi atas dua yaitu desain teknologi secara umum dan terinci. Pada tahap ini kita menentukan teknologi yang akan dipergunakan dalam menerima *input*, menjalankan model, ,menyimpan dan mengakses data menghasikan dan mengirimkan keluaran dan membantu mengendalikan dari sistem secara keseluruhan. Teknologi yang dimaksud meliputi :

- perangkat keras (*hardware*), yang terdiri dari alat masukan, alat pemroses, alat *output* dan simpanan luar.
- Perangkat lunak (*software*), yang terdiri dari perangkat lunak sistem operasi (*operating system*), perangkat lunak bahasa (*language software*) dan perangkat lunak (*application software*)
- Sumber daya manusia (*brainware*), misalnya operator komputer, pemrogram, spesialis telekomunikasi, sistem analisis dan sebagainya.

Desain teknologi sangat dibutuhkan pada tahap implementasi dan pengujian untuk membuktikan bahwa sistem dapat berjalan secara semestinya.

✓ Desain model.

Tahap desain model terbagi menjadi dua yaitu desain model secara umum dan terinci. Tahap desain model secara umum berupa desain sistem secara fisik dan logika. Desain fisik dapat digambarkan dengan bagan alir dokumen.

2.2.9 Perancangan konseptual

Perancangan konseptul sering kali disebut dengan perancangan logis, pada perancangan ini kebutuhan pemakai dan pemecah masaalah yang teridentifikasi selama tahap analisis sistem mulai dibuat untuk diimplementasikan. Ada tiga langkah penting yang dilakukan dalam perancangan konseptual, yaitu evaluasi alternatif rancangan, penyiapan spesifikasi rancangan dan penyimpanan laporan rancangan sistem secara konseptual.

Setelah alternatif rancangan dipilih, tahap selanjutnya adalah penyimpanan spesifikasi rancangan yang elemen-elemen sebagai berikut :

a) Keluaran

Rancangan laporan mencakup frekuensi laporan (harian, mingguan, dsb), isi laporan, bentuk laporan dan laporan cukup ditampilkan pada layar atau perlu dicetak.

b) Menyimpan Data

Dalam hal ini, semua data yang diperlukan untuk membentuk laporan ditemukan lebih detail, termasuk ukuran data dan ukurannya dalam berkas.

c) Masukan

Rancangan masukan meliputi data yang perlu dimasukan kedalam sistem.

d) Prosedur Pemrosesan dan Operasi

Rancangan ini menjelaskan bagaimana data masukan diproses dan disimpan dalam rangka untuk menghasilkan laporan.

Langkah berikutnya adakah menyiapkan laporan rancangan sistem konseptual. Berdasarkan laporan inilah, perancangan sistem secara fisik dibuat.

2.2.10 Perancangan fisik

Pada rancangan ini, rancangan yang masih bersifat konsep diterjemahkan dalam bentuk fisik sehingga terbentuk spesifikasi lengkap tentang modul sistem dan antarmuka antar modul serta rancangan basis data secara fisik.

Berapa hasil akhir setelah tahap perancangan fisik berakhir :

1. Rancangan keluaran

Rancangan keluaran berupa bentuk laporan dan rancangan dokumen.

2. Rancangan masukan

Rancangan masukan berupa rancangan layar untuk pemasukan data.

3. Rancangan antarmuka pemakai dan sistem.

Rancangan ini berupa rancangan interaksi antarmuka dan sistem, misalnya berupa menu, icon dan lain-lain.

4. Rancangan *Platform*.

Rancangan ini berupa rancangan yang melakukan *Hardware* dan *Software* yang akan digunakan.

5. Rancangan basis data.

Rancangan ini berupa rancangan-rancangan berkas dalam basis data termasuk penentuan kapasitas masing-masing.

6. Rancangan modul.

Rancangan berupa rancangan program yang dilengkapi dengan *Algoritma* (cara modul / program kerja).

7. Rancangan kontrol.

Rancangan ini berupa rancangan kontrol-kontrol yang digunakan dalam sistem seperti validasi, otoritas dan audit data.

8. Dokumentasi

Berupa hasil dokumentasi hingga tahap perancangan fisik.

9. Rancangan pengujian








Berupa rancangan yang dipakai untuk menguji sistem.










10. Rancangan konversi.

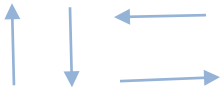

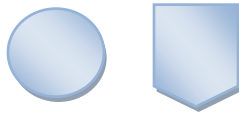
Berupa rancana untuk menerapkan sistem baru terhadap sistem lama.

Bagan alir sistem merupakan bagan yang menunjukkan arus pekerjaan secara keseluruhan dari sistem bagan alir sistem di gambarkan dengan simbol-simbol sebagai berikut.

Tabel 2.6 Bagan Alir Sistem

No	NAMA SIMBOL	SIMBOL	KETERANGAN
1.	Simbol Dokumen		Menunjukkan dokumen <i>Output</i> dan <i>Input</i> baik itu proses manual, mekanik, atau komputer.
2.	Simbol Kegiatan Manual		Menunjukkan pekerjaan manual.
3.	Simbol Simpanan Offline		Menunjukkan file non-komputer yang di arsip urut angka (<i>Numerical</i>), huruf (<i>Alphabetical</i>), atau tanggal (<i>Chronological</i>).
4.	Simbol Kartu Plong		Menunjukkan <i>Input</i> dan <i>Output</i> yang menggunakan kartu plong (<i>Punched Card</i>).
5.	Simbol Proses		Menunjukkan kegiatan proses dari operasi program komputer.
6.	Simbol Operasi Luar		Menunjukkan operasi yang dilakukan diluar proses operasi komputer.
7.	Simbol Pengurutan Offline		Menunjukkan proses urut data diluar proses komputer, operasi luar menunjukkan operasi yang dilakukan diluar proses operasi

			komputer.
8.	Simbol Pita Mengetik		Menunjukkan <i>Input</i> dan <i>Output</i> menggunakan pita <i>Magnetic</i> .
9.	Simbol Hardisk		Menunjukkan <i>Input</i> dan <i>Output</i> menggunakan <i>Hardisk</i> .
10.	Simbol Diskette		Menunjukkan <i>Input</i> dan <i>Output</i> menggunakan <i>Diskette</i> .
11.	Simbol Drum Magnetik		Menunjukkan <i>Input</i> dan <i>Output</i> menggunakan <i>Magnetic</i> .
12.	Simbol Pita Kertas Berlubang		Menunjukkan <i>Input</i> dan <i>Output</i> menggunakan pita keras berlubang
13.	Simbol Keyboard		Menunjukkan <i>Input</i> menggunakan <i>On-line keyboard</i> .
14.	Simbol Display		Menunjukkan <i>Output</i> yang ditampilkan di monitor.
15.	Simbol Pita Kontrol		Menunjukkan penggunaan pita kontrol (<i>Control Tape</i>) dalam <i>Batch Control</i> total untuk pencocokan di proses <i>Batch Processing</i> .
16.	Simbol Hubungan Komunikasi		Menunjukkan proses transmisi data melalui <i>Channel</i> komunikasi.

17.	Simbol Garis Alir		Menunjukkan arus dari proses.
18.	Simbol Penjelasan		Menunjukkan penjelasan dari suatu proses.
19.	Simbol Penghubung		Menunjukkan perhubungan ke halaman yang masih sama atau ke halaman yang lain.

Sumber : (Jogiyanto HM, 2005)

Untuk mempermudah penggambaran suatu sistem yang ada atau di sistem yang baru yang akan di kembangkan secara logika tanpa memperhatikan lingkungan fisik di mana data tersebut mengalir atau lingkungan fisik di mana data tersebut akan disimpan, maka digunakan Diagram Arus Data (DAD) atau *Data Flow Diagram* (DFD). Dalam menggambarkan sistem yang perlu dilakukan pembentukan simbol, berikut ini simbol-simbol yang sering digunakan dalam DAD :

1. *External Entity* (kesatuan luar) atau *Boundary* (batas sistem).

Setiap sistem pasti mempunyai batas sistem (*Boundary*) yang memisahkan suatu sistem dengan lingkungan luarnya. Sistem akan menerima *Input* dan menghasilkan *Output* pada lingkungan luarnya. Kesatuan luar (*external Entity*) merupakan kesatuan dilingkungan luar sistem yang dapat berupa orang, organisasi atau sistem yang lain berada dilingkungan luarnya yang akan memberikan *Input* serta menerima *Output* dari sistem.

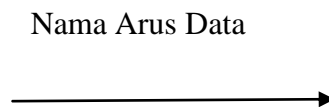


Gambar 2.3 Contoh Notasi kesatuan luar
(Jogiyanto HM, 2005)

2. *Data Flow* (arus data)

Arus data ini menunjukkan arus atau aliran yang dapat berupa masukan untuk sistem atau hasil dari proses sistem.

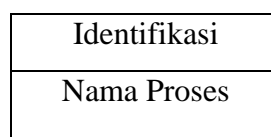
Contoh Notasi Arus Data :



Gambar 2.4 Contoh arus data
(Jogiyanto HM, 2005)

3. *Process* (Proses)

Suatu proses adalah kegiatan atau kerja yang dilakukan orang, mesin atau komputer dari hasil suatu arus data yang akan keluar dari proses.



Gambar 2. 5 Contoh Notasi Proses
(Jogiyanto HM, 2005)

4. *Data Store* (simpanan data)

Simpanan data pada DFD dapat disimbolkan dengan sepasang garis horizontal paralel yang tertutup disalah satu ujungnya.

Media	Data Store
-------	------------

Gambar 2.6 Contoh Notasi Simpanan Data
(Jogiyanto HM, 2005).

2.2.11 Implementasi Sistem

Sistem dianalisa dan didesain secara rinci dan teknologi yang diseleksi dan dipilih. Tiba saatnya sekarang sistem untuk diimplementasikan (diterapkan). Tahap implementasi sistem merupakan tahap meletakkan sistem supaya siap untuk dioperasikan. Tahap implementasi sistem dapat terdiri dari langkah-langkah sebagai berikut :

1. Menerapkan rencana implementasi

Rencana implementasi merupakan kegiatan awal dari tahap implementasi sistem. Rencana implementasi dimaksudkan terutama untuk mengatur biaya dan waktu yang dibutuhkan selama tahap implementasi.

2. Melakukan kegiatan Implementasi

Kegiatan implementasi dilakukan dengan dasar kegiatan yang telah direncanakan dalam rencana implementasi. Kegiatan-kegiatan yang dapat dilakukan dalam tahap implementasi ini adalah sebagai berikut :

- Persiapan tempat dan instalasi perangkat keras dan perangkat lunak

Jika peralatan baru akan dimiliki, maka tempat atau ruangan untuk peralatan ini perlu juga dipertimbangkan. Sistem komputer yang besar membutuhkan tempat dengan lingkungan yang lebih harus diperhitungkan. Langkah selanjutnya setelah persiapan fisik tempat adalah penginstalasi perangkat lunak yang sudah ada.

- Pemograman dan pengetesan sistem

Pemograman merupakan kegiatan menulis kode program yang akan dieksekusi oleh komputer. Kode program yang ditulis oleh pemogram harus berdasarkan dokumentasi yang harus disediakan oleh analisis sistem hasil dari desain sistem secara rinci. Sebelum program diterapkan, maka program harus terlebih dahulu terbebas dari kesalahan-kesalahan. Oleh

sebab itu, program harus diuji untuk menemukan kesalahan-kesalahan yang mungkin dapat terjadi. Program dites untuk tiap-tiap modul dan dilanjutkan dengan pengetesan untuk semua modul yang dirangkai.

➤ **Pengetesan sistem**

Pengetesan sistem biasanya dilakukan setelah pengetesan program. Pengetesan sistem dilakukan untuk memeriksa kekompakan antar komponen sistem yang diimplementasikan. Tujuan utama dari pengetesan sistem ini adalah untuk memastikan bahwa elemen-elemen atau komponen-komponen dari sistem telah berfungsi sesuai dengan yang diharapkan.

2.2.12 Operasi dan pemeliharaan

Setelah masa sistem berjalan sepenuhnya menggantikan sistem lama, sistem memasuki pada tahapan operasi dan pemeliharaan. Zwass (1999) membagi pemeliharaan perangkat lunak menjadi tiga macam yaitu :

- ✓ **Pemeliharaan perfektif** ditujukan untuk memperbaharui sistem lama sebagai tanggapan atas perubahan kebutuhan pemakai dan kebutuhan organisasi, meningkatkan efisiensi sistem, dan memperbaiki dokumentasi.
- ✓ **Pemeliharaan adaptif**
Pemeliharaan adaptif berupa perubahan aplikasi untuk menyesuaikan diri terhadap lingkungan perangkat keras dan perangkat lunak baru. Sebagai contoh pemeliharaan ini dapat berupa perubahan aplikasi dari *Mainframe* ke lingkungan *Client / Server* atau mengkonversi dari sistem berbasis berkas ke lingkungan basis data.
- ✓ **Pemeliharaan korektif**
Pemeliharaan korektif berupa pembetulan atas kesalahan-kesalahan yang ditemukan pada saat sistem berjalan.

2.2.13 Teknik Pengujian Sistem

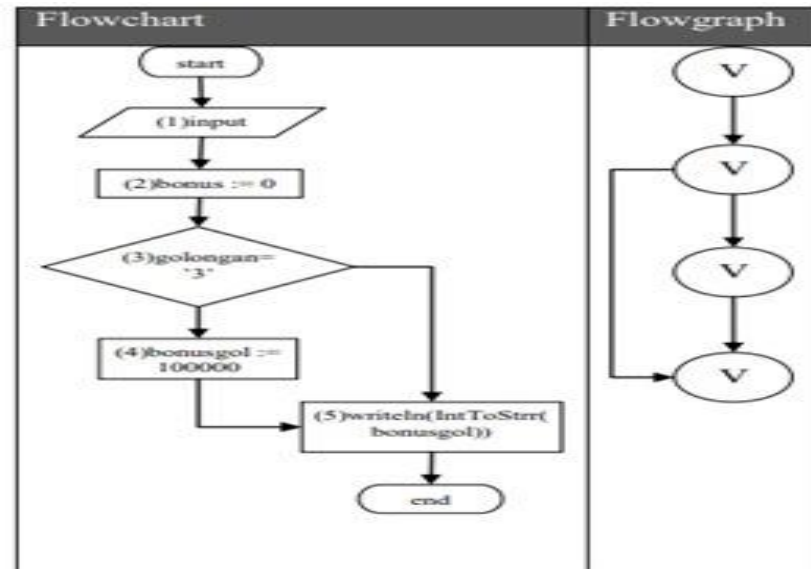
2.2.13.1 *White Box Testing*

Pengujian perangkat lunak adalah element kritis dari jaminan kualitas perangkat lunak dan merepresentasikan kajian pokok dari spesifikasi, desain dan pemodelan.

Pengujian sistem / perangkat lunak memiliki sejumlah aturan yang berfungsi sebagai sasaran pengujian, diantaranya sebagai berikut :

- Pengujian adalah proses eksekusi suatu program dengan maksud menentukan kesalahan.
- *Test Case* yang baik adalah *Test Case* yang memiliki probabilitas tinggi untuk menemukan kesalahan yang belum pernah ditemukan sebelumnya.
- Pengujian yang sukses adalah pengujian yang mengungkap semua kesalahan yang belum pernah ditemukan sebelumnya.

Pengujian *White Box*, adalah metode pengujian yang menggunakan struktur kontrol desain prosedur untuk memperoleh *Test Case*. Dengan menggunakan metode *White Box*, perekayasa sistem dapat melakukan *Test Case* yang memberikan jaminan bahwa semua jalur independen pada suatu modul telah digunakan paling tidak satu kali, menggunakan semua keputusan logis pada sisi *True* dan *False*, mengeksekusi semua *Loop* pada batasan mereka dan pada batasan operasional mereka, dan menggunakan struktur data internal untuk menjamin validitasnya. Pengujian *Basis Path* adalah teknik pengujian *White Box* yang diusulkan pertama kali oleh Tom McCabe, metode *Basis Path* ini menginginkan desainer *Test Case* mengukur kompleksitas logis dari desain prosedural dan menggunakan sebagai pedoman untuk menetapkan basis data set dari jalur eksekusi.



Gambar 2.7 Contoh Diagram Alir (*Flowchart*) dan (*Flowgraph*)
(Sumber ; Meylanie Oliviya, 2018)

Bagan alir digunakan untuk menggambarkan struktur kontrol program. Dan untuk menggambarkan grafik alir, harus memperhatikan representasi desain prosedural pada bagan alir, pada gambar ini, bagan alir memetakan bagan tersebut kedalam alir yang sesuai (dengan mengasumsikan bahwa tidak ada kondisi senyawa yang di eksekusikan didalam diamond keputusan dari bagan alir tersebut. Masing-masing lingkaran yang disebut *Simpul* grafik alir, mempresentasikan satu atau lebih statemen prosedural. Urutan kotak proses dan permata keputusan dapat memetakan simpul tunggal. Anak panah tersebut yang disebut *Edges* atau *Links*, mempresentasikan aliran kontrol dan analog dengan anak panah bagan alir. *Edges* harus berhenti pada suatu simpu, meskipun bila simpul tersebut tidak mempresentasikan statemen prosedural. (Roger S. Pressman, 2002).

2.2.13.2 *Black Box Testing*

Pengujian *Black Box* berusaha menemukan kesalahan dalam kategori :

- Pada struktur data (pengaksesan basis data)
- Kesalahan fungsi tidak benar atau hilang
- Kesalahan antarmuka

- Kesalahan inisialisasi dan akhir program
- Kesalahan performansi.

Pengujian ini berfokus pada persyaratan fungsional perangkat lunak dan merupakan komplemen dari pengujian *White Box*. Hal tersebut dapat dilakukan melalui :

- a. Pengujian *Graph-Based* : dimulai dengan membuat grafik sekumpulan *Mode* yang mempresentasikan objek (misal *New File*, layar baru dengan atributnya), link (hubungan antar objek), *Node-weight* (misal nilai data tertentu seperti atribut layar , perilaku) dan *Link-weight* (karakteristik suatu *Link* misal menu *Select*).
- b. *Equivalence Partitioning* : membagi domain *Input* untuk pengujian agar diperoleh kelas-kelas kesalahan (misal kelompok data karakter atau atribut yang lain).
- c. Analisis Nilai Batas : pengujian berdasarkan nilai batas domain *Input*.
- d. Pengujian perbandingan : disebut juga pengujian *Back To Back* yang diterapkan pada suatu versi perangkat lunak atau perangkat lunak redundan untuk memastikan konsistensinya.

2.3 Framework

Framework merupakan perangkat lunak yang mulai jadi pilihan untuk membuat aplikasi (Andersta, 2008). Kemudahan- kemudahan yang diberikan menarik orang-orang untuk menggunakannya. Hal ini tidak terlepas dari tingkat efektifitas dan efisiensinya yang lebih baik dalam proses pengembangan suatu perangkat lunak.

Framework adalah sekumpulan perintah/fungsi dasar yang dapat membantu dalam menyelesaikan proses-proses yang lebih kompleks (Visikom, 2009). *Framework* adalah suatu aplikasi yang dapat digunakan ulang untuk membuat bermacam-macam aplikasi (Jhonson, 2009). *Framework* merupakan kumpulan beberapa kelas abstrak pada domain tertentu, sehingga pengembangan

yang menggunakan *framework* harus melengkapi kelas abstrak tersebut menjadi perangkat lunak yang diinginkan (Andresta. 2008)

Framework merupakan rancangan sistem yang dapat digunakan ulang. Didalamnya terdapat interaksi kumpulan objek tertentu. *Framework* mendeskripsikan bagaimana hubungan dan interaksi objek-objek tersebut beserta antar muka dan aliran kembali antar objek tersebut.

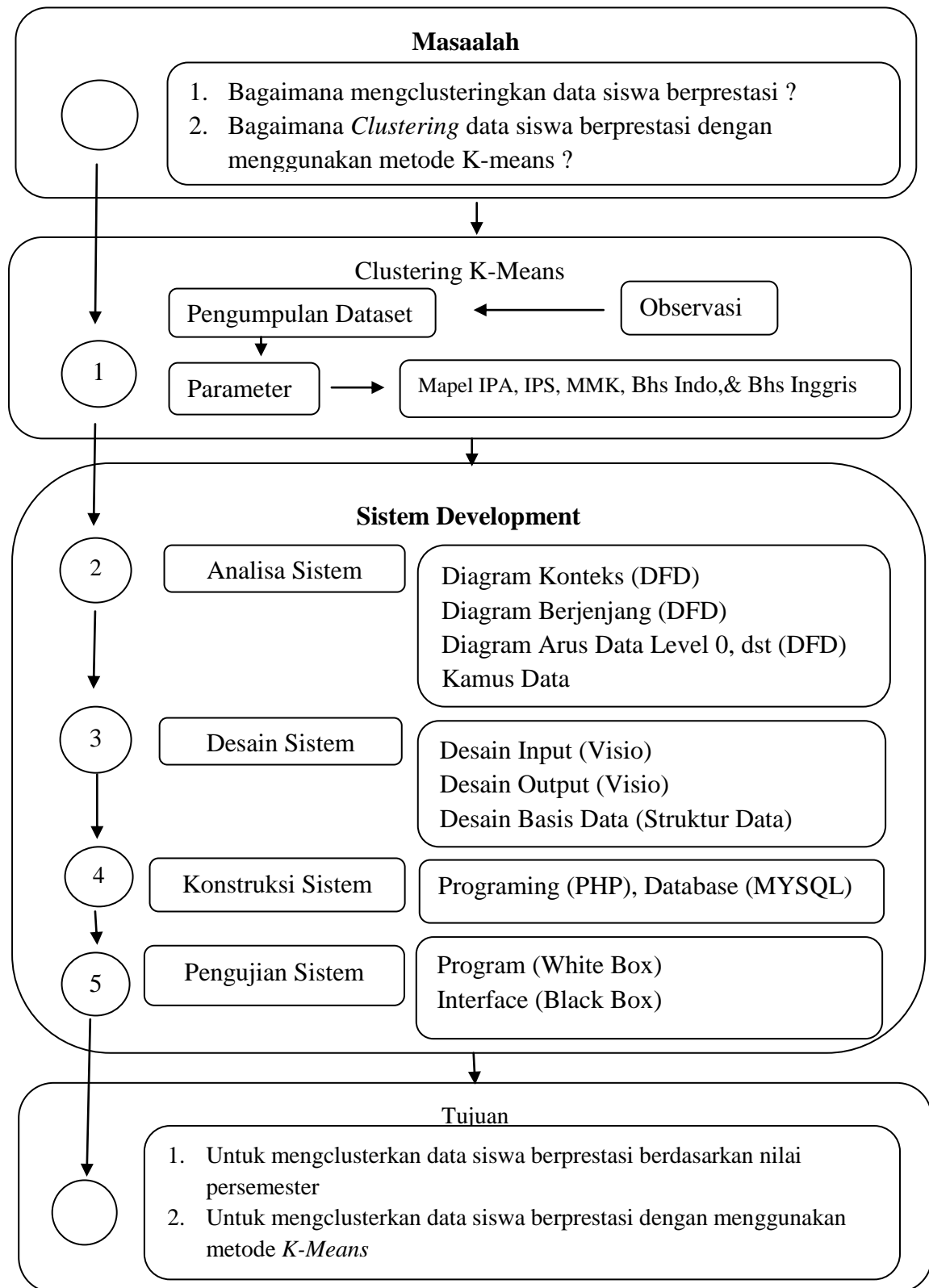
2.4 TOOLS Pendukung

Perangkat lunak pendukung yang digunakan penulis dalam membangun sistem ini yaitu PHP dan *Xampp*. Seperti pada tabel dibawah ini :

Tabel 2.7 Tools Pendukung

NO	TOOLS	KEGUNAAN
1	PHP	PHP (<i>Hypertext Preprocessor</i>) adalah suatu server-side scripting yang menyatu dengan HTML (<i>Hyper Text Markup Language</i>) untuk membuat website yang dinamis, server-side scripting sendiri berarti suatu pemogramaan yang pengeksesksiannya berada disetiap server, seluruh proses didalam bahasa pemograman PHP dilakukan disebuah server tetapi hasil akan ditampilkan di browser.
2	Xampp	XAMPP adalah <i>software</i> adalah <i>web server apache</i> yang didalamnya tertanam server MySQL yang didukung dengan bahasa pemograman PHP untuk membuat <i>website</i> yang dinamis, XAMPP sendiri mendukung dua sistem operasi yaitu windows dan linux, untuk linux dalam proses penginstalannya menggunakan interface gratis sehingga lebih mudah dalam penggunaan XAMPP di windws di banding dengan linux, dalam XAMPP ada 3 komponen utama yang di tanam didalamnya yaitu tab server Apache, PHP dan MySQL

2.5 Kerangka Fikir



Gambar 2.8 Kerangka Pemikiran.

BAB III

METODE PENELITIAN

3.1 Objek Penelitian

Berdasarkan latar belakang dan kerangka pemikiran seperti yang telah diuraikan dalam Bab I dan Bab II, maka menjadi objek penelitian adalah **Clustering Data Siswa Berprestasi**.

3.2 Metode Penelitian

Metode yang digunakan dalam penelitian untuk masalah *Clustering Data Siswa Berprestasi* yaitu metode *Deskriptif* dimana metode ini bertujuan untuk mengungkapkan kebenaran yang *Objektif*. Dengan kata lain menjabarkan suatu keadaan atau fenomena yang terjadi saat ini dengan menggunakan prosedur ilmiah untuk menjawab masalah secara aktual.

3.2.1 Tahap Analisis

Analisis untuk *Clustering Data Siswa Berprestasi* menggunakan *Metode K-Means* sebagai berikut :

a) Analisis Sistem Berjalan

SMP Negeri 2 Kwandang merupakan salah satu sekolah yang ada di Kabupaten Gorontalo Utara, dalam operasionalnya sekolah mengumpulkan informasi melalui sistem komputerisasi menggunakan aplikasi *Microsoft Office* yang berguna untuk menampung dan mengelola data transaksi, kemudian nantinya data tersebut di olah sehingga dapat diketahui tingkat pada volume suatu data pada suatu waktu tertentu. Terkadang hasil dari pengolahan data dengan cara sederhana menggunakan aplikasi *Microsoft Office*. Kesulitan untuk menganalisis data.

b) Analisis Sistem yang di usulkan

- Data Sekunder

Data sekunder merupakan data yang diperoleh dari studi literatur yang dilakukan. Sumber dari studi literatur yaitu jurnal, makalah ilmiah atau buku yang membahas tentang penelitian *Algoritma K-Means*.

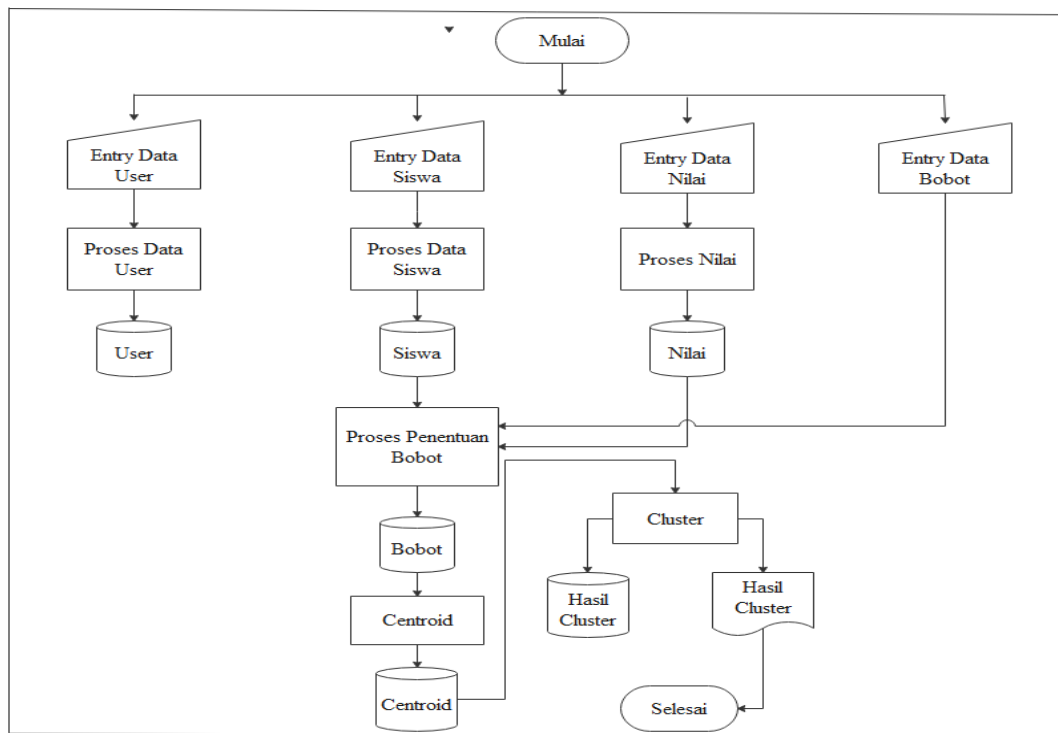
- Data Primer

Data primer merupakan data yang diperoleh dari penelitian. Data primer pada penelitian ini yaitu berupa data siswa, seperti NISN, mata pelajaran IPA, mata pelajaran IPS, mata pelajaran Matematika, mata pelajaran Bahasa Indonesia dan mata pelajaran Bahasa Inggris

Berikut ini merupakan tahap dalam melakukan pengelompokan data siswa berprestasi :

1. Entri Data : Mata pelajaran IPA, IPS, Matematika, Bahasa Indonesia, dan Bahasa Inggris.
2. Proses : Clustering
3. Laporan : Laporan Hasil Clustering.

Sistem yang diusulkan dapat digambarkan menggunakan flowchart dokumen.



Gambar 3.1 Sistem yang Diusulkan

3.2.2 Tahap Desain

Desain sistem menggunakan pendekatan berorientasi procedural/structural

- *Desain Output*

Desain Output dimaksudkan untuk mengetahui bagaimana dan seperti bentuk *output-output* dari sistem yang akan dibuat. *Desain Output* terinci terbagi atas dua, yaitu *desain Output* berbentuk laporan di media kertas dan *desain output* dalam bentuk dialog di layar terminal (*Monitor*).

- *Desain Input*

Masukan merupakan awal dimulainya proses pengolahan informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan oleh konsumen. Data hasil dari transaksi tidak lepas dari data yang dimasukan. *Desain Input* terinci dimulai dari desain dokumen dasar dari

perangkat *input* yang pertamakali jika dokumen dasar tidak didesain dengan baik, kemungkinan *input* yang tercatat dapat salah bahkan kurang.

- Desain *Database*

Basis data (*database*) merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan disimpan luar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. *Database* merupakan salah satu komponen yang penting disistem informasi, karena berfungsi sebagai basis penyedia informasi bagi para pemakainya, penerapan *database* dalam aplikasi disebut *database system*.

- Desain Teknologi

Pada tahap ini kita menentukan teknologi yang akan dipergunakan dalam menerima input, menjelaskan model, menyimpan dan mengakses data, menghasilkan dan mengirimkan kelurahan dan membantu pengendalian dari sistem secara keseluruhan.

- Desain Program

Pada tahap ini menggunakan alat bantu visio dalam bentuk Pseudoce program pada proses penerapan algoritma K-Means.

3.2.3 Tahap Pembuatan

Merupakan tahapan dimana kita melakukan pengembangan, melakukan tahap produksi sistem hasil analisa dan desain sistem sebelumnya. Termasuk didalamnya membangun sebuah sistem Klasifikasi Data Siswa Berprestasi Menggunakan *Metode K-Means* menginstal paket tambahan untuk menjalankan program, menulis *Listing* program dan membangunnya dalam bentuk sebuah formuli, antarmuka atau integrasi sistem-sistem program yang terdiri dari *Input*, *Proses*, dan *Output*, yang tersusun dalam sebuah sistem menu sehingga dapat dijalankan oleh pengguna sistem. Dalam tahapan ini penulis menggunakan bahasa pemrograman PHP, dan *DataBase MySQL Server*.

3.2.4 Tahap Pengujian

Tahap ini dilakukan setelah semua modul selesai dibuat, dan program dapat berjalan, dimana seluruh perangkat lunak, program tambahan dan semua program yang terlibat dalam pembangunan sistem yang diuji untuk memastikan sistem dapat berjalan sesuai dengan rencangan atau belum. Pengujian yang dilakukan dengan dua teknik, yaitu :

a. *White Box*

Dalam pengujian *White Box* ini dengan membuat bagan alir program, *Listing* program, grafik alir, pengujian *BasisPath* serta perhitungan *Cyclomatic Complexity*.

b. *Black Box*

Pengujian *Black Box* yang termasuk dalam tahap ini yaitu menguji antarmuka sistem, apakah sebuah sistem setelah diberikan kepengguna dapat dioperasikan atau tidak.

3.2.5 Tahap Implementasi

Tahap implementasi sistem (*Sistem Implementation*) merupakan tahap meletakkan sistem supaya siap untuk dioperasikan di Sekolah SMP Negeri 2 Kwandang.

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Berikut data penelitian yang telah diperoleh dari lokasi penelitian sebagai berikut :

Tabel 4.1 Hasil Pengumpulan Data

No	NIS	IPA	IPS	Matematika	Bahasa Indonesia	Bahasa Inggris
1	0039293047	88	86	84	86	86
2	0041208584	88	89	90	88	87
3	3046559563	84	82	87	80	86
4	0055463313	90	93	92	91	87
5	3052009270	80	80	85	86	86
6	0048739784	87	85	81	82	87
7	0053578343	92	92	93	91	86
8	0052753461	90	91	90	87	88
9	0053577836	88	88	82	90	92
10	0056570925	89	88	90	93	87
...	
460	0041251161	88	85	88	89	87

(Sumber : SMP N 2 Kwandang 2016-2018).

4.2 Hasil Pemodelan

Berikut Tahapan Algoritma *K-Means* :

1. Penentuan pusat awal Cluster
2. Perhitungan jarak pusat Cluster, untuk mengukur antara data dengan pusat Cluster digunakan *Euclidian distance* kemudian akan didapatkan matrik jarak sebagai berikut : $d(x,y) = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + \dots + (x_5-y_5)^2}$
3. Menentukan Cluster dengan jarak terdekat pada masing-masing data
4. Menghitung pusat Cluster baru

Berikut contoh tabel dengan menggunakan rumus diatas :

Tabel 4.2 Data Nilai Siswa

No	NIS	IPA	IPS	Matematika	Bahasa Indonesia	Bahasa Inggris
1	0039293047	88	86	84	86	86
2	0041208584	88	89	90	88	87
3	3046559563	84	82	87	80	86
4	0055463313	90	93	92	91	87
5	3052009270	80	80	85	86	86
6	0048739784	87	85	81	82	87
7	0053578343	92	92	93	91	86
8	0052753461	90	91	90	87	88
9	0053577836	88	88	82	90	92
10	0056570925	89	88	90	93	87

Langkah 1 : Tanyakan pada user, kesepuluh data tersebut hendak dibagi kedalam berapa cluster ? misalkan hendak dibagi kedalam 5 cluster, jadi $K = 5$

Langkah 2 : Tetapkan secara sembarang, $K = 5$ rekord diantaranya sebagai *Initial Cluster Centre Locations*, dan Hitung *BCV*nya.

Tabel 4.3 Contoh Sampel Data Nilai Siswa

Data Pilihan					
	IPA	IPS	Matematika	Bahasa Indo	Bahasa Inggris
C1	66.5	78	80.5	93.5	92.5
C2	71	72.625	93.5	81	91.5
C3	63	69	77.5	81.5	92
C4	70	76	88.5	85.5	91.5
C5	69	79	78	76	90

Diketahui rumus :

$$C1 = \sqrt{((x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2 + (x4 - y4)^2 + (x5 - y5)^2}$$

$$C2 = \sqrt{((x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2 + (x4 - y4)^2 + (x5 - y5)^2}$$

$$C3 = \sqrt{((x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2 + (x4 - y4)^2 + (x5 - y5)^2}$$

$$C4 = \sqrt{((x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2 + (x4 - y4)^2 + (x5 - y5)^2}$$

$$C5 = \sqrt{((x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2 + (x4 - y4)^2 + (x5 - y5)^2}$$

Keterangan :

X = Data *Record*

Y = Data *Centroid*

Berikut cara kerja perhitungan manual Sbb:

$$\begin{aligned} C1 \quad (1) &= \sqrt{((x1 - y1)^2 + (x2 - y2)^2 + (x3 - y3)^2 + (x4 - y4)^2 + (x5 - y5)^2} \\ &= \sqrt{(66,5 - 66,5)^2 + (78 - 78)^2 + (80,5 - 80,5)^2 + (93,5 - 93,5)^2 + (92,5 - 92,5)^2} \\ &= \sqrt{0 + 0 + 0 + 0 + 0} \\ &= 0 \end{aligned}$$

$$\begin{aligned}
\text{C1} \quad (2) &= \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
&\quad \sqrt{(66-66,5)^2 + (76-78)^2 + (76,5-80,5)^2 + (93,5-93,5)^2 + (93-92,5)^2} \\
&\quad \sqrt{(-0,5)^2 + (-2)^2 + (-4)^2 + (0)^2 + (0,5)^2} \\
&\quad \sqrt{(0,25) + (4) + (16) + (0) + (0,25)} \\
&\quad \sqrt{20,5} \\
&= 4,57269257 \\
\\
\text{C2} \quad (3) &\quad \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
&\quad \sqrt{((69-71)^2 + (66,5-72,625)^2 + (87-93,5)^2 + (78-81)^2 + (92-91,5)^2)} \\
&\quad \sqrt{(-2)^2 + (-6,125)^2 + (-6,5)^2 + (-3)^2 + (0,5)^2} \\
&\quad \sqrt{4 + 37,515625 + 42,25 + 9 + 0,25} \\
&\quad \sqrt{93,015625} \\
&= 9,64446085 \\
\\
\text{C2} \quad (4) &\quad \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
&\quad \sqrt{(67-71)^2 + (76-72,625)^2 + (88,5-93,5)^2 + (77,5-81)^2 + (91-91,5)^2} \\
&\quad \sqrt{(-4)^2 + 3,375^2 + (-5)^2 + (-3,5)^2 + (0,5)^2} \\
&\quad \sqrt{16 + 11,390625 + 25 + 12,25 + 0,25} \\
&\quad \sqrt{64,890625} \\
&= 8,05547174 \\
\\
\text{C3} \quad (5) &\quad \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_2-y_2)^2)} \\
&\quad \sqrt{((63-63)^2 + (69-69)^2 + (77,5-77,5)^2 + (81,5-81,5)^2 + (92-92)^2)} \\
&\quad \sqrt{0 + 0 + 0 + 0 + 0} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\text{C3} \quad (6) \quad & \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
& \sqrt{(69-63)^2 + (79-69)^2 + (78,-77,5)^2 + (76-81,5)^2 + (90-92)^2} \\
& \sqrt{6^2 + 10^2 + 0,5^2 + (-5,5)^2 + (-2)^2} \\
& \sqrt{36 + 100 + 0,25 + 30,25 + 4} \\
& \sqrt{170,5} \\
& = 13,0575649 \\
\\
\text{C4} \quad (7) \quad & \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
& \sqrt{(69,5-70)^2 + (66-76)^2 + (79,5-88,5)^2 + (77,5-85,2)^2 + (91-91,5)^2} \\
& \sqrt{(-0,5)^2 + (-10)^2 + (-9)^2 + (-8)^2 + (-0,5)^2} \\
& \sqrt{0,25 + 100 + 81 + 64 + 0,25} \\
& \sqrt{245,5} \\
& = 15,6684396 \\
\\
\text{C4} \quad (8) \quad & \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
& \sqrt{(61,5-70)^2 + (82-76)^2 + (80,5-88,5)^2 + (88,5-88,5)^2 + (90,5-91,5)^2} \\
& \sqrt{(-8,5)^2 + (6)^2 + (-8)^2 + (3)^2 + (-1)^2} \\
& \sqrt{72,25 + 36 + 64 + 9 + 1} \\
& \sqrt{182,25} \\
& = 13,5 \\
\\
\text{C5} \quad (9) \quad & \sqrt{((x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 + (x_4-y_4)^2 + (x_5-y_5)^2)} \\
& \sqrt{(70-69)^2 + (76-79)^2 + (88,5-78)^2 + (88,5-76)^2 + (91,5-90)^2} \\
& \sqrt{(1)^2 + (-3)^2 + (10,5)^2 + (9,5)^2 + (1,5)^2}
\end{aligned}$$

$$\begin{aligned}
& \sqrt{1 + 9 + 110,25 + 90,25 + 2,25} \\
& \sqrt{212,75} \\
& = 14,5859521
\end{aligned}$$

C5 (10)
$$\begin{aligned}
& \sqrt{((x1-y1)^2 + (x2-y2)^2 + (x3-y3)^2 + (x4-y4)^2 + (x5-y5)^2)} \\
& \sqrt{(71-69)^2 + (72,625-79)^2 + (93,5-78)^2 + (81-76)^2 + (91,5-90)^2)} \\
& \sqrt{(2)^2 + (-6,375)^2 + (15,5)^2 + (5)^2 + (1,5)^2)} \\
& \sqrt{4 + 40,640625 + 240 + 25 + 2,25} \\
& \sqrt{312,140625} \\
& = 17,6675019
\end{aligned}$$

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat Cluster. jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster terdekat.

4.2.1 Iterasi 1

Tabel 4.4 Hasil Iterasi 1

JARAK KE PUSAT CLUSTER							Cluster	min jarak
NO	NISN	C1	C2	C3	C4	C5		
1	2888	0	3,391165	1,322876	3,905125	2	C1	0
2	2808	3,391165	0	2,783882	0,866025	5,049752	C2	0
3	2878	1,322876	2,783882	0	3,082207	2,397916	C3	0
4	2837	3,905125	0,866025	3,082207	0	5,408327	C4	0
5	2958	2	5,049752	2,397916	5,408327	0	C5	0
6	2916	6,689544	9,785193	7,314369	10,22252	5,17204	C5	5,17204
7	2850	10,03743	12,87439	10,29563	13,20984	8,291562	C5	8,291562
8	2957	5,454356	8,645808	6	9,027735	3,968627	C5	3,968627
9	2910	9,205976	12,0727	9,591663	12,4298	7,53326	C5	7,53326
10	2787	7,937254	10,36822	8,170067	10,61838	6,63325	C5	6,63325

4.2.2 Iterasi 2

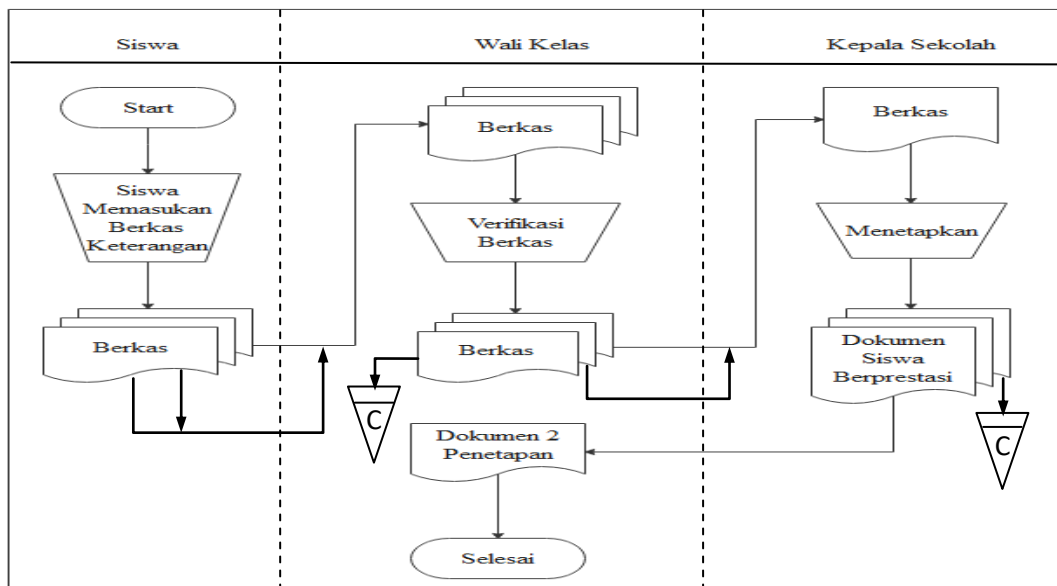
Tabel 4.5 Hasil Iterasi 2

JARAK KE PUSAT CLUSTER							ANGGOTA	min jarak
NO	NISN	C1	C2	C3	C4	C5		
1	2888	0	3,391165	1,322876	3,905125	6,572248	C1	0
2	2808	3,391165	0	2,783882	0,866025	9,549578	C2	0
3	2878	1,322876	2,783882	0	3,082207	6,97814	C3	0
4	2837	3,905125	0,866025	3,082207	0	9,909311	C4	0,866025
5	2958	2	5,049752	2,397916	5,408327	4,901814	C1	2
6	2916	6,689544	9,785193	7,314369	10,22252	1,333333	C5	6,689544
7	2850	10,03743	12,87439	10,29563	13,20984	4,323707	C5	10,03743
8	2957	5,454356	8,645808	6	9,027735	2,068279	C5	5,454356
9	2910	9,205976	12,0727	9,591663	12,4298	2,728451	C5	9,205976
10	2787	7,937254	10,36822	8,170067	10,61838	3,407508	C5	7,937254

Dengan Kesimpulan hasil iterasi ke-2 maka iterasinya telah berhenti karena nilainya sama dengan nilai Sebelumnya atau sama dengan iterasi 1.

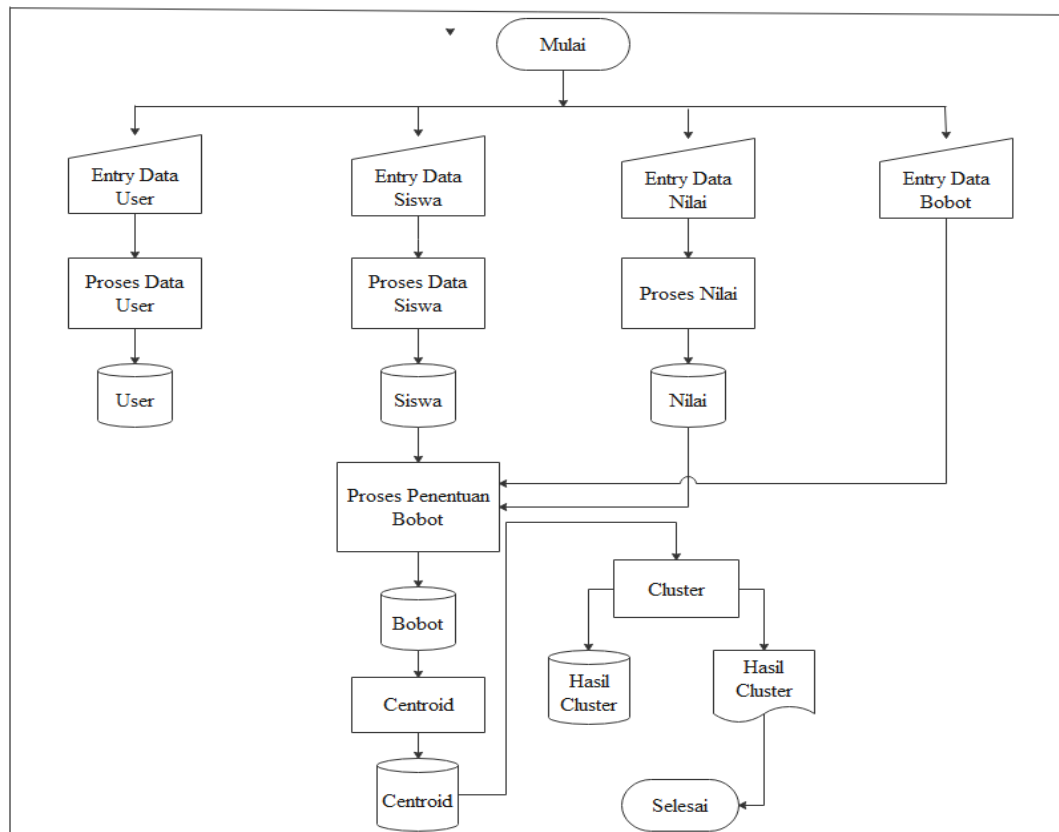
4.3 Analisa Sistem

4.3.1 Sistem yang berjalan



Gambar 4.1 Bagan Alir Dokumen

4.3.2 Proses sistem yang diusulkan

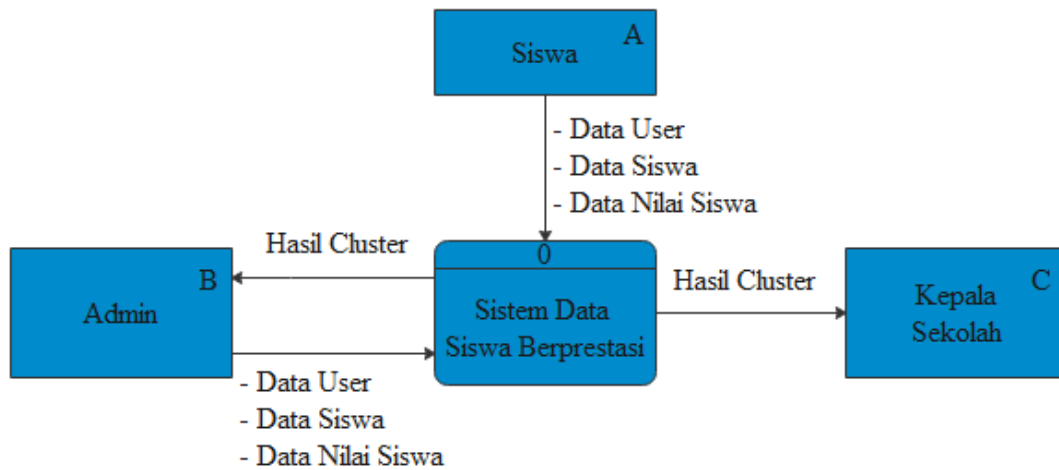


Gambar 4.2 Bagan Alir Sistem Yang Diusulkan

4.4 Hasil Pengembangan Sistem

4.4.1 Diagram Konteks

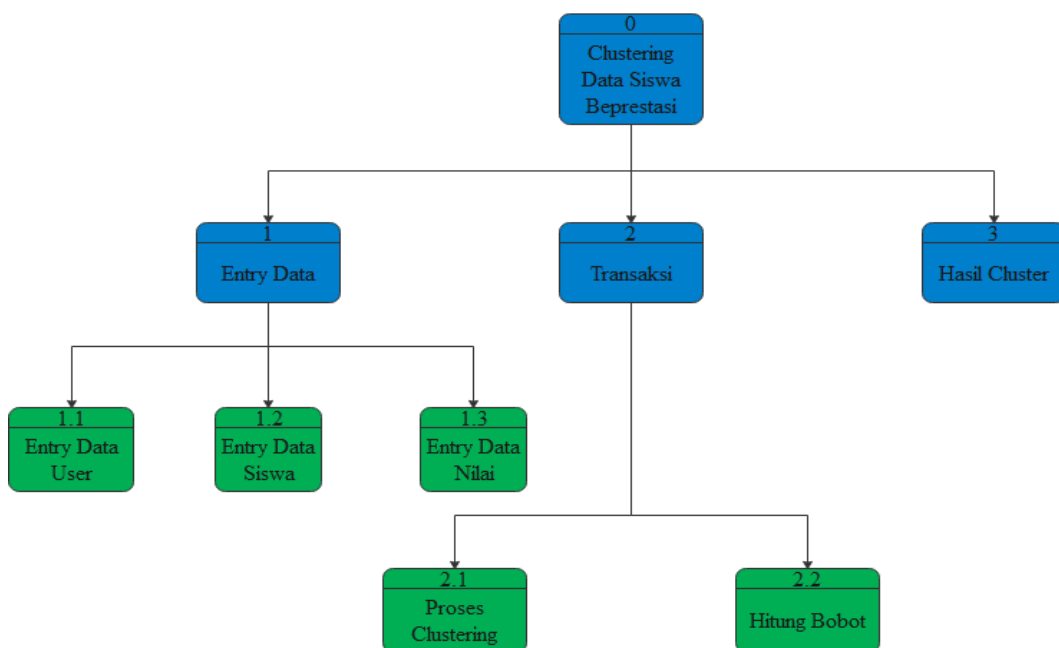
Diagram Konteks adalah sebuah bagian level dari *Data Flow Diagram* yang digunakan untuk menetapkan konteks serta batasan-batasan sistem pada sebuah pemodelan. Hal ini termasuk hubungan dengan entitas entitas diluar syste itu sendiri, seperti sistem, kelompok organisasi, penyimpanan data eksternal lain.



Gambar 4.3 Diagram Konteks

4.4.2 Diagram Berjenjang

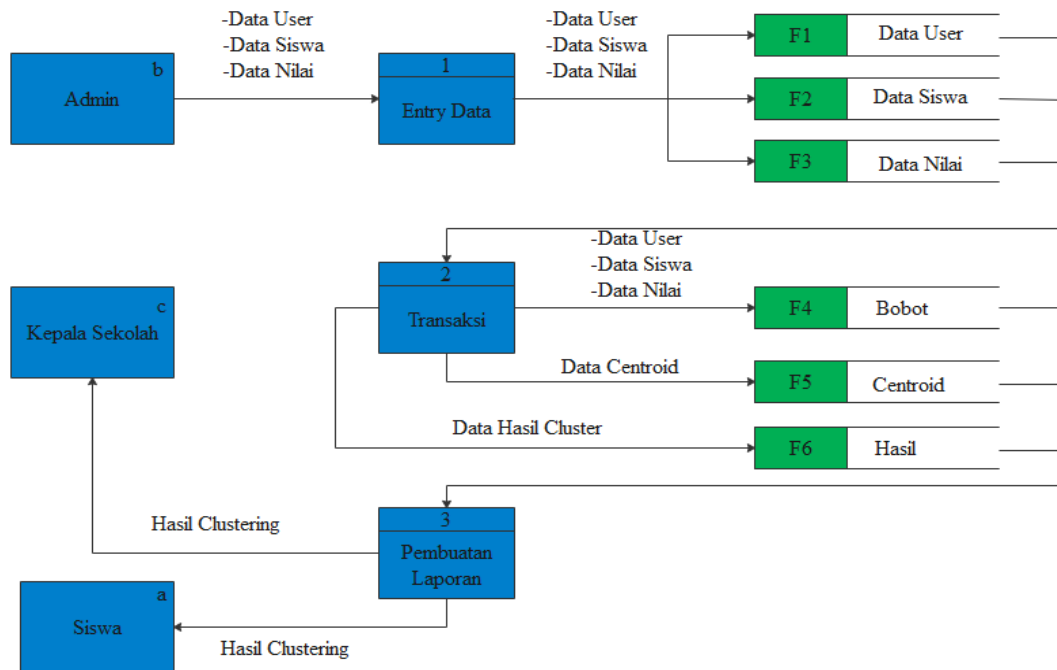
Menggambarkan struktur dari sistem berupa suatu bagan berjenjang yang menggambarkan semua proses yang ada disistem. Dipergunakan untuk mempersiapkan penggambaran DFD ke level lebih bawah lagi.



Gambar 4.4 Diagram Berjenjang

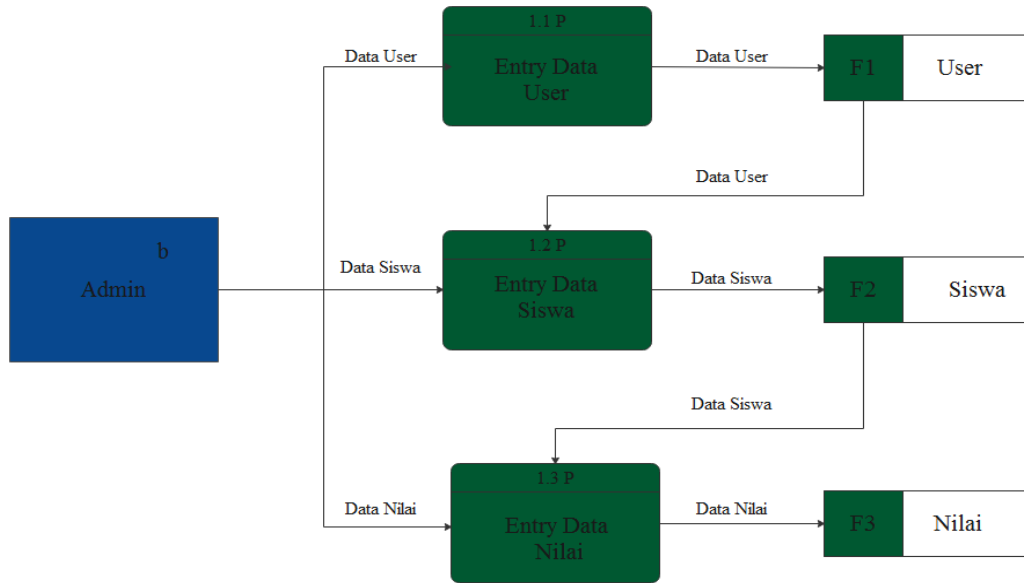
Dari gambar diatas dapat dijelaskan sebagai berikut : dari sistem clustering data siswa berprestasi terdapat beberapa proses diantaranya entry data, transaksi dan hasil cluster. Pada proses entry data terdapat beberapa proses diantaranya entry data user, entry data siswa dan entry data nilai. Pada proses transaksi memproses transaksi clustering data hitung bobot. Proses hasil cluster adalah hasil yang diserahkan pada pihak sekolah.

4.4.3 Diagram Arus Data Level 0



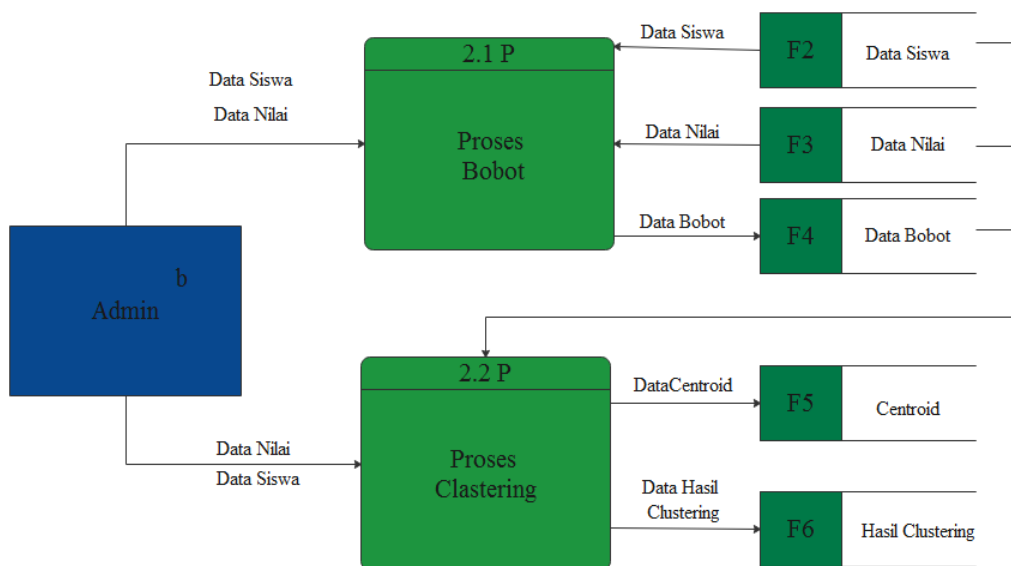
Gambar 4.5 Arus Data DAD Level 0

4.4.4 Diagram Arus Data Level 1 Proses 1



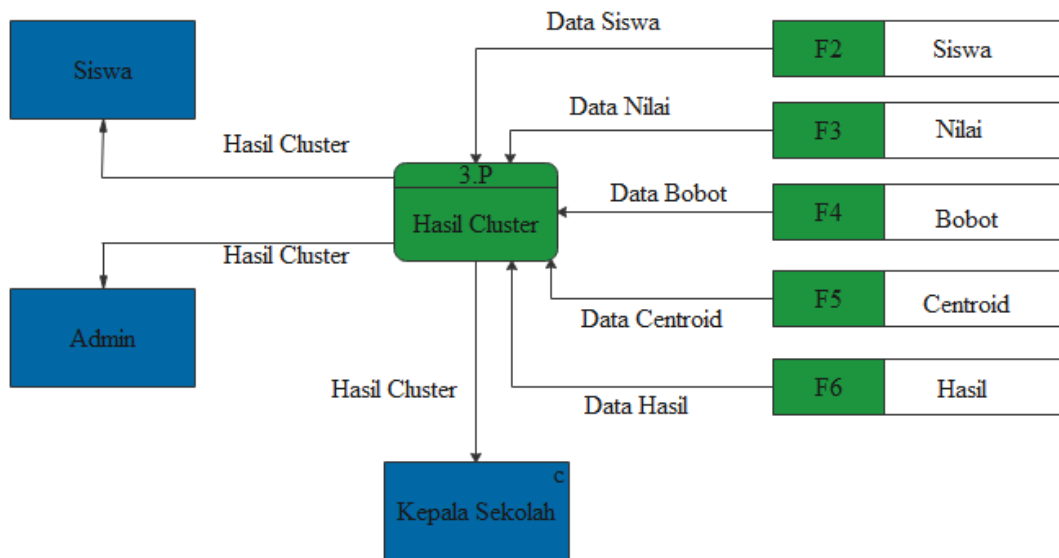
Gambar 4.6 Diagram Arus Data Level 1 Proses 1

4.4.5 Diagram Arus Data Level 1 Proses 2



Gambar 4.7 Diagram Arus Data Level 1 Proses 2

4.4.6 Diagram Arus Data Level 1 Proses 3



Gambar 4.8 Arus Data DAD Level 1 Proses 3

4.5 Kamus Data

Kamus Data atau *Data Dictionary* adalah katalog fakta tentang data dan kebutuhan-kebutuhan informasi dari suatu sistem informasi. Kamus data digunakan untuk merancang input, file-file database dan output, kamus data dibuat berdasarkan arus yang mengalir pada DAD, dimana di dalamnya terdapat struktur dari arus data secara detail.

4.5.1 Kamus Data Terinci

Kamus Data : Data User				
Nama Arus Data : Data User			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data user			Arus Data : b-1, 1-f1, f1-2,	
Periode : Setiap ada penambahan data user			b-1.1p, 1.1p-f1, f1-1 – 1.2p	
Rekap nilai				
No	Nama Item Data	Type	Width	Ket
1	Id_Admin	N	5	Kode Admin
2	Id_Pimpinan	N	8	Kode Pimpinan
3	Id_Pengguna	N	8	Kode Pengguna

Kamus Data : Data siswa				
Nama Arus Data : Data siswa			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data siswa			Arus Data : b-1, 1-f2, f2-2, b-1.2p, 1.2p-f2, f2-1.3p, b-2.1p, b-2.2p, f2-3p	
Periode : Setiap ada penambahan data siswa				
Rekap nilai				
No	Nama Item Data	Type	Width	Ket
1	NISN	N	8	Kode NISN
2	Name siswa	C	25	Nama Siswa
3	Kelas	C	5	Kelas
4	Umur	N	2	Umur
5	Alamat	C	20	Alamat

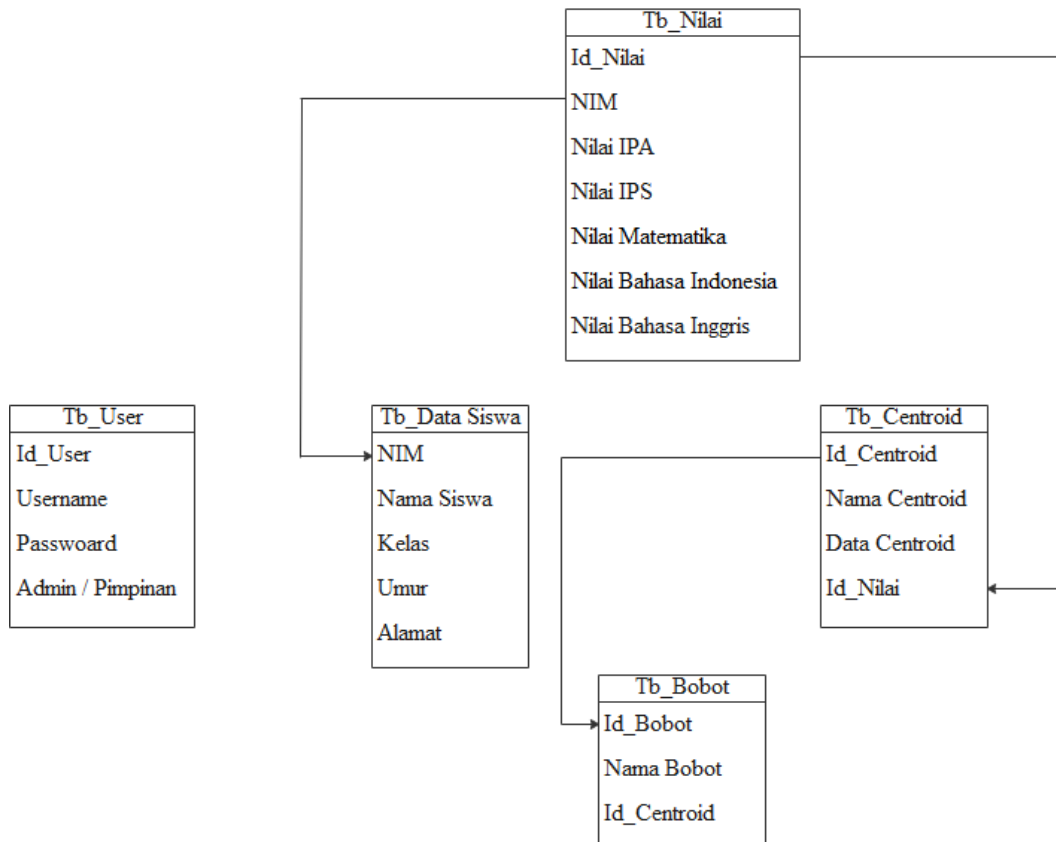
Kamus Data : Data nilai				
Nama Arus Data : Data nilai			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data nilai			Arus Data : b-1, 1-f3, f3-2, b-1.3p, 1.3p-f3, f3-3p	
Periode : Setiap ada penambahan data nilai				
Rekap nilai				
No	Nama Item Data	Type	Width	Ket
1	Id_Nilai	N	10	Kode Nilai
2	Nilai IPA	N	5	Nilai IPA
3	Nilai IPS	N	5	Nilai IPS
4	Nilai Matematika	N	5	Nilai Matematika
5	Nilai Bahasa Indo	N	5	Nilai Bhs Indo
6	Nilai Bahasa Inggris	N	5	Nilai Bhs Inggris

Kamus Data : Data centroid				
Nama Arus Data : Data centroid			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data centroid			Arus Data : b-1, 1-f3, f3-2, 2-f5, b-2.1p, 2.1p-f4, f4-2.2p, 2.2p-f5, f5-3.p	
Periode : Setiap ada penambahan data centroid				
Rekap nilai				
No	Nama Item Data	Type	Width	Ket
1	Id_Centroid	N	20	Id_Centroid
2	Nama Centroid	N	20	Nama Centroid
3	Data Centroid	N	20	Data Centroid

Kamus Data : Data bobot				
Nama Arus Data : Data bobot			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data bobot			Arus Data : b-1, 1-f3, f3-2, 2-f5, b-2.1p, 2.1p-f4, f4-3.p	
Periode : Setiap ada penambahan data bobot				
Rekap nilai				
No	Nama Item Data	Type	Width	Ket
1	Id_Bobot	N	5	Kode bobot
2	Nama Bobot	C	30	Nama bobot

Kamus Data : Data hasil cluster				
Nama Arus Data : Data hasil cluster			Bentuk Data : Dokumen	
Penjelasan : Berisi data-data hasil cluster			Arus Data : b-1, 1-f3, f3-2, 2-f6, b-2.1p, 2.1p-f4, f4-2.2p, 2.2p – f6, f6-3.p	
Periode : Setiap ada penambahan data hasil cluster				
Rekap nilai				
No	Nama Item Data	Type	Width	Ket
1	Cluster 1	N	30	Hasil C1
2	Cluster 2	N	30	Hasil C2
3	Cluster 3	N	30	Hasil C3
4	Cluster 4	N	30	Hasil C4
5	Cluster 5	N	30	Hasil C5

4.6 Desain Database



4.6.1 Daftar input yang didesain

Daftar input yang didesain			
Untuk : SMP N 2 Kwandang			
Tahap : Desain Sistem			
Kode Input	Nama Input	Simbol Input	Periodik
1-001	Data User	Admin	Non Periodik
2-002	Data Siswa	Admin	Non Periodik
3-003	Data Nilai	Admin	Non Periodik
4-004	Data Bobot	Admin	Non Periodik
5-005	Data Centroid	Admin	Non Periodik
6-006	Data Hasil	Admin	Non Periodik

4.6.2 Desain file secara umum

Tabel 4.6 Rancangan file secara umum

Kode File	Nama File	Field	Type	File
F1	User	30	C	Id_User
F2	Siswa	30	C	Id_Siswa
F3	Nilai	5	N	Id_Nilai
F4	Bobot	5	N	Id_Bobot
F5	Centroid	10	N	Id_Centroid
F6	Hasil	20	N	Id_Hasil

4.7 Arsitektur Sistem

Spesifikasi hardware dan software yang direkomendasikan yaitu :

- a. Processor : Intel Core i3 atau Lebih
- b. RAM : 4 GB atau lebih
- c. VGA : 256MB atau lebih
- d. Harddisk : 40 GB atau lebih
- e. Operating System : Windows 10
- f. Tools : Notepad ++ atau Dreamwaver

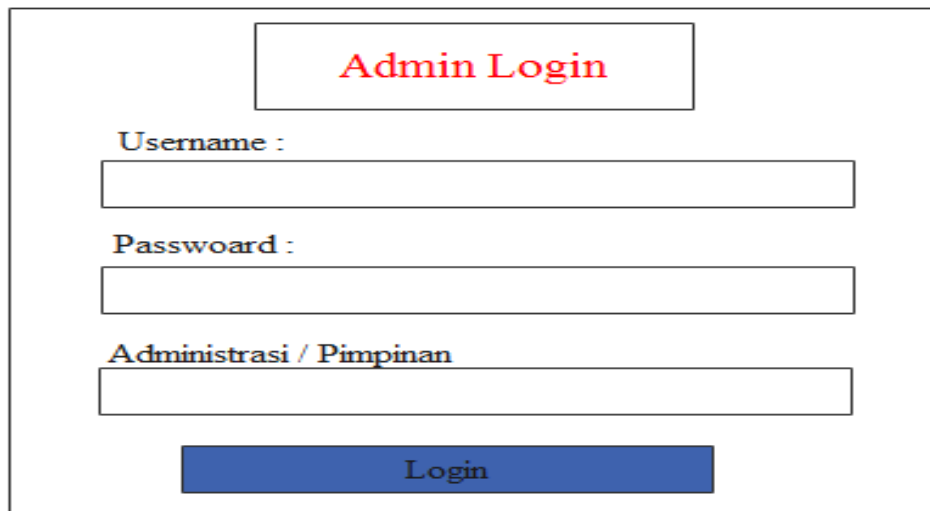
4.8 Interface Desagin Mekanisme User

Tabel 4.7 Interface Desagin

USER	KATEGORI	AKSES INPUT	AKSES OUTPUT
Admin	Administrator	All	All
User	User	Tidak Ada	Hasil <i>Cluster</i>

4.9 Interface Design Input

4.8 Tampilan Window Input Login



The image shows a login window titled "Admin Login" in red text. Below the title, there are three input fields: "Username :", "Password :", and "Administrasi / Pimpinan". At the bottom, there is a blue "Login" button.

Admin Login

Username :

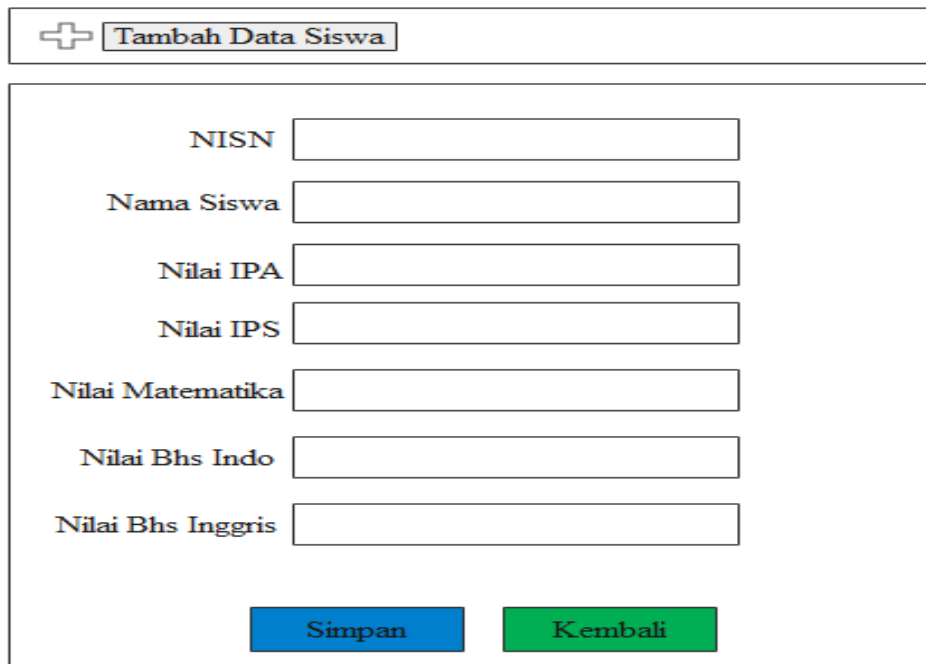
Password :

Administrasi / Pimpinan

Login

Gambar 4.9 Tampilan Login Admin

4.10 Tampilan Window Input Data Centroid



The image shows a window titled "Tambah Data Siswa" with a plus icon. Below the title, there are seven input fields: "NISN", "Nama Siswa", "Nilai IPA", "Nilai IPS", "Nilai Matematika", "Nilai Bhs Indo", and "Nilai Bhs Inggris". At the bottom, there are two buttons: "Simpan" (blue) and "Kembali" (green).

+ Tambah Data Siswa

NISN

Nama Siswa

Nilai IPA

Nilai IPS

Nilai Matematika

Nilai Bhs Indo

Nilai Bhs Inggris

Simpan **Kembali**

Gambar 4.10 Tampilan Window Input Data Centroid

4.11 Program Desain

Tabel 4.9 Desain Program

Class / Type	Attributes[Type]	Methods[Events/or Type]
Menu Index	Login[Menu]	Login[Click]
	DataMining[Menu]	DataMining[Click]
	Cluster[Menu]	Cluster[Click]
Menu Index	K-Means[Menu]	K-Means[Click]
	SelamatDatang[Menu]	SelamatDatang[View]
Menu Login	UserName[TextBox]	UserName[TextBox]
	Password[TextBox]	Password[TextBox]
	Login[Button]	Login[Button]
Menu Home	MenuAdmin[Menu]	MenuAdmin[Load]
	HomePage[Menu]	HomePage[Click]
	HasilClustering[Menu]	HasilClustering[Click]
	Logout[Menu]	Logout[Click]
Menu HomePage	HomePage[Load]	HomePage[View]
Menu Semua Data	NamaObjek[TextBox]	NamaObjek[TextBox]
	Data[TextBox]	Data[TextBox]
	Cluster[Load]	Cluster[wiew]
	CentroidAwal[Textboxt]	CentroidAwal[Textboxt]
	Input[TextBoxt]	Input[TextBoxt]
	Tambah[Buton]	Tambah[Click]
	EditData[Buton]	EditData[Click]
	Hapus[Buton]	Hapus[Click]
Menu Hasil	Hasil[Load]	Hasil[Hasil]
Menu Diagram	Diagram[Buton]	Diagram[Click]

4.12 Hasil Konstruksi Sistem

Pada tahap konstruksi sistem, hasil dari analisis dan desain sistem kemudian diterjemahkan kekonstruksi sistem/software dengan menggunakan bahasa pemrograman PHP. Adapun alat bantu yang digunakan pada tahap ini adalah :

1. **PHP** untuk pemrogramannya
2. **Mysql** untuk tempat penyimpanan databasenya
3. **Notepad++** Untuk Editor Webnya

4.13 Kode Program untuk pengujian *WhiteBox*

```
//centroid baru 1.a
$jum = 0;
$arr = array();
for ($i = 0; $i < count($arr_c1); $i++) {
    $arr[$i] = $arr_c1_temp[$i] * $arr_c1[$i];
    if ($arr_c1[$i] == 1) {
        $jum++;
    }
}
$c1a_b = array_sum($arr) / $jum;

//centroid baru 1.b
$jum = 0;
$arr = array();
for ($i = 0; $i < count($arr_c2); $i++) {
    $arr[$i] = $arr_c2_temp[$i] * $arr_c1[$i];
    if ($arr_c1[$i] == 1) {
        $jum++;
    }
}
$c1b_b = array_sum($arr) / $jum;

//centroid baru 1.c
$jum = 0;
$arr = array();
for ($i = 0; $i < count($arr_c3); $i++) {
    $arr[$i] = $arr_c3_temp[$i] * $arr_c1[$i];
    if ($arr_c1[$i] == 1) {
        $jum++;
    }
}
$c1c_b = array_sum($arr) / $jum;

//centroid baru 1.d
$jum = 0;
```



```

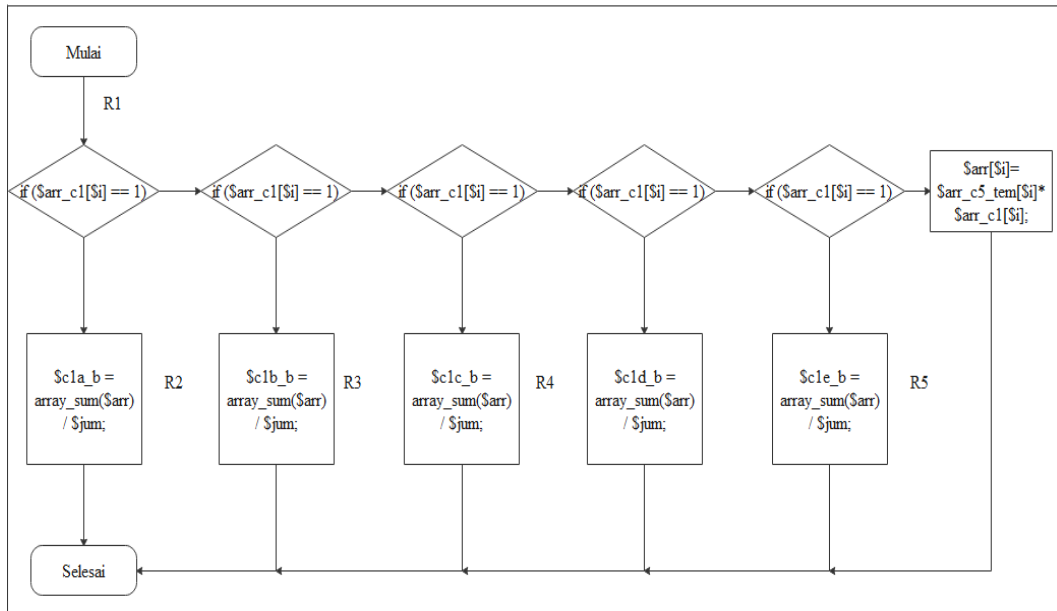
$arr = array();
for ($i = 0; $i < count($arr_c4); $i++) {
    $arr[$i] = $arr_c4_temp[$i] * $arr_c1[$i];
    if ($arr_c1[$i] == 1) {
        $jum++;
    }
}
$clb_b = array_sum($arr)/$jum;

//centroid baru 1.e
$jum = 0;
$arr = array();
for ($i = 0; $i < count($arr_c5); $i++) {
    $arr[$i] = $arr_c5_tem[$i] * $arr_c1[$i];
    if ($arr_c1[$i] == 1) {
        $jum++;
    }
}
$cle_b = array_sum($arr)/$jum;

```

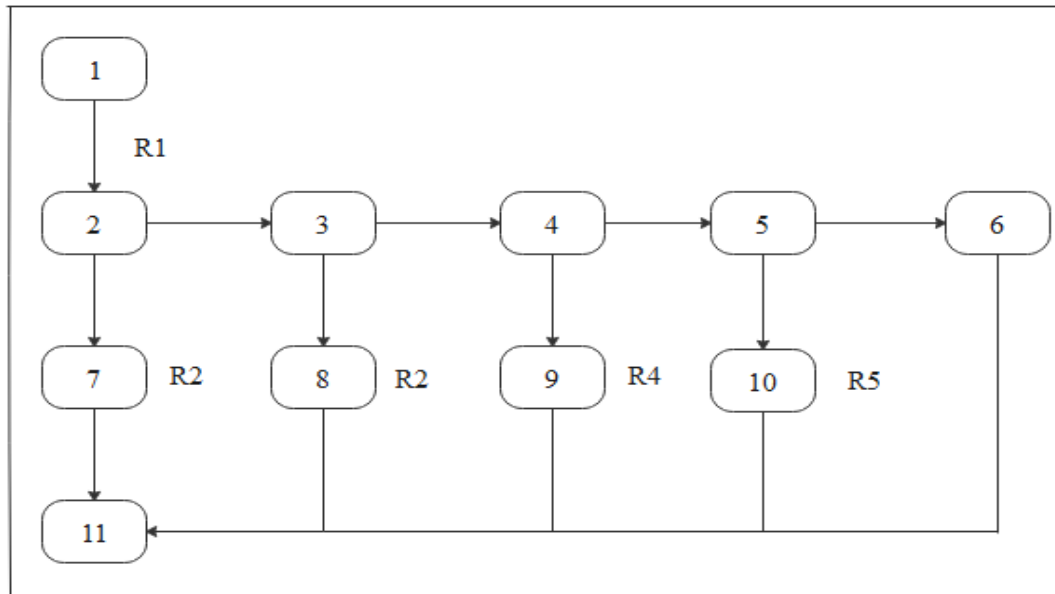
</table>
</div>
</div>

4.14 Flowchart Program untuk Pengujian White Box



Gambar 4.11 : Flowchart Program pengujian white box

4.15 Flowgraph Program untuk pengujian WhiteBox



Gambar 4.12 : Flowgraph program pengujian Whitebox

4.16 Perhitungan CC pada pengujian WhiteBox

Dari Flowgraph yang sudah tersedia, *cyclomatic complexity* dari sebuah program dapat dibuat dengan menggunakan rumus dibawah ini :

$$V(G) = E - N + 2$$

Keterangan :

$V(G)$: *cyclomatic complexity*

E : Total Jumlah Edge

N : Total Jumlah Node

P : Predikat Node = 1

Pada contoh *Flowgraph* diatas dapat dihitung *cyclomatic complexity* nya sebagai berikut :

Diketahui :	Region (R)	= 5
	Node (N)	= 11
	Edge (E)	= 12

Predicate Node (P) = 4

$$V(G) = (E - N) + 2$$

$$V(G) = P + 1$$

$$V(G) = (12 - 11) + 2 = 3$$

$$V(G) = 2 + 1 = 3$$

CC = R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, R11

4.17 Path Pada Pengujian White Box

Tabel 4.10 Path Pengujian White Box

No	Path	Keterangan
1	1-2-7-11	OK
2	1-2-3-8-.....11	OK
3	1-2-3-4-9-.....11	OK
4	1-2-3-4-5-10-.....11	OK
5	1-2-3-4-5-6-11-...11	OK

4.18 Hasil Pengujian *BlackBox*

Tabel 4.11 Pengujian *Black Box*

Input/Event	Fungsi	Hasil	Hasil Uji
Login	Login dengan memasukkan username dan password	<ul style="list-style-type: none"> - Jika Password Salah, maka ulangi memasukkan username dan password - Jika Password Benar, akan masuk ke sistem 	Sesuai
Menu Home Admin	Menampilkan Halaman Admin	Halaman Admin ditampilkan	Sesuai
Pilih Menuu Home Page	Menampilkan Halaman Home Page	Halaman Home Page ditampilkan	Sesuai
Pilih Menu Profil Perusahaan	Menampilkan Halaman Profil Perusahaan	Halaman Profil Perusahaan ditampilkan	Sesuai
Pilih menu Semua data	<ul style="list-style-type: none"> - Menampilkan Halaman tabel input data - Menampilkan Halaman Input Cluster 	<ul style="list-style-type: none"> - Halaman tabel input data ditampilkan - Halaman Tabel Input Cluster ditampilkan 	Sesuai
Pilih Menu Hasil Clustering	Menampilkan Halaman Hasil Cluster	Halaman Hasil Cluster ditampilkan	Sesuai
Pilih menu Hasil Diagram	Menampilkan Halaman Hasil Diagram	Halaman Hasil Diagram ditampilkan	Sesuai

Pilih menu Logout	Kembali ke halaman Login	Halaman Login ditampilkan	Sesuai
Pilih Menu Edit	Menampilkan Halaman Edit Data	Halaman Edit Data ditampilkan	Sesuai
Pilih Menu Hapus	Menampilkan Halaman Hapus Data	Halaman Hapus Data ditampilkan	Sesuai
Pilih Menuh Tambahkan	Menampilkan Halaman Tambah Data	Halaman Tambah Data ditampilkan	Sesuai

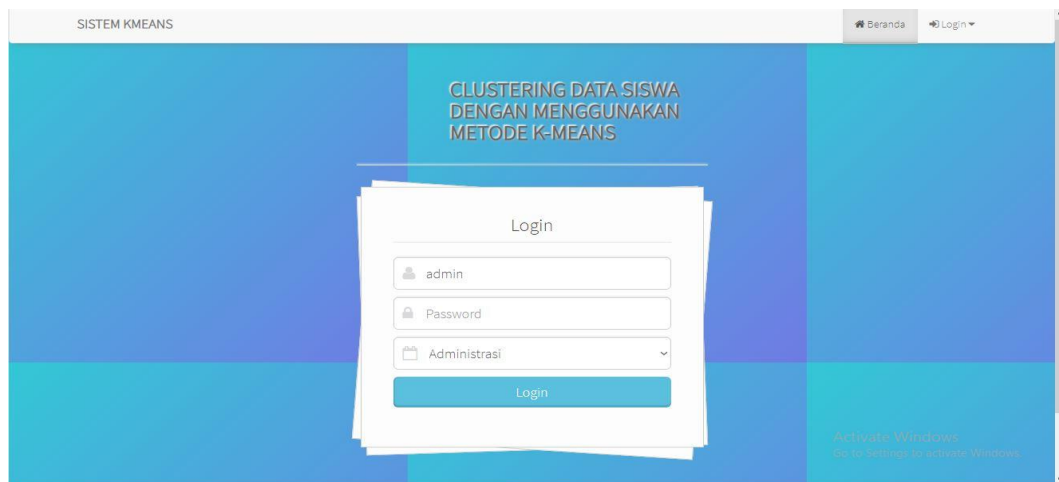
BAB V

PEMBAHASAN PENELITIAN

5.1 Pembahasan Sistem

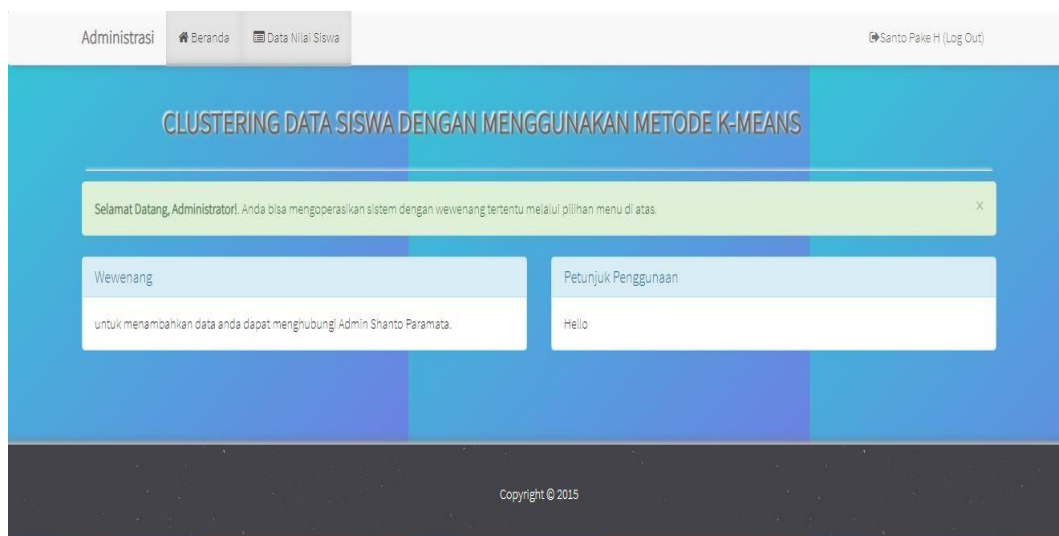
Berikut adalah tampilan sistem Clustering Penerapan Algoritma K-Means Untuk Pengelompokan Data Siswa Berprestasi.

5.1.1 Hasil Tampilan Window Login Admin



Gambar 5.1 Tampilan Window Login Admin

5.1.2 Hasil Tampilan Window Beranda Admin



Gambar 5.2 Tampilan Window Beranda Admin

5.1.3 Hasil Tampilan Window Beranda Pimpinan



Gambar 5.3 Tampilan Window Beranda Pimpinan

5.1.4 Hasil Tampilan Window Data Nilai Siswa

Administrasi

Beranda

Data Nilai Siswa

Santo Paik H (Log Out)

CLUSTERING DATA SISWA DENGAN MENGGUNAKAN METODE K-MEANS

Data Nilai Siswa

Tambah Data

10

Baris data tiap halaman

Cari

Masukkan Kata Kunci

NISN	Nama Siswa	Nilai IPA	Nilai IPS	Nilai Matematika	Nilai Bahasa Indonesia	Aksi
2609	Aditya Hau	88.5	88	86	87	<div>Edit</div> <div>Hapus</div>
2617	Ahmad I Ajunu	85.5	85.5	85.5	86	<div>Edit</div> <div>Hapus</div>
2630	Adriyanto Mukmin	87.5	87	86.5	86.5	<div>Edit</div> <div>Hapus</div>
2634	Arnold Latereng	85.5	85	86	85.5	<div>Edit</div>

Activate Windows

Go to Settings to activate Windows.

Gambar 5.4 Hasil Tampilan Data Nilai Siswa

5.1.5 Hasil Tampilan Window Input Data Centroid

CLUSTERING DATA SISWA DENGAN MENGGUNAKAN METODE K-MEANS

+ Tambah Data Siswa

NISN

Nama Siswa

Nilai IPA

Nilai IPS

Nilai Matematika

Nilai Bahasa Indonesia

Simpan
Kembali

Gambar 5.5 Tampilan Window Input Data Centroid

5.1.6 Hasil Tampilan Window Iterasi K-Means

CLUSTERING DATA SISWA DENGAN MENGGUNAKAN METODE K-MEANS

Data Awal

Proses Iterasi Selanjutnya

NISN	Nama Siswa	IPA	IPS	Matematika	Bahasa Indonesia	Centroid 1				Centroid 2				Centroid 3				Centroid 4				C1	C2	C3	C4
						81	65	65	67	65	81	65	76	65	65	81	65	78	67	88	89				
	Aditya Hau	88.5	88	86	87	0.000000				1.581139				1.414214				1.802776				1	0	0	0
	Ahmad I Ajunu	85.5	85.5	85.5	86	4.062019				2.738613				4.949747				3.278719				0	1	0	0
	Adriyanto Mukmin	87.5	87	86.5	86.5	1.581139				0.000000				2.345208				1.118034				0	1	0	0
	Arnold Laterreng	85.5	85	86	85.5	4.500000				3.041381				5.315073				3.605551				0	1	0	0
	Arpan Masliu	88.5	89	87	87	1.414214				2.345208				0.000000				2.291288				0	0	1	0
	Delis Mohamad	90	92	88.5	90.5	6.062178				7.158911				5.074446				6.519202				0	0	1	0
	Eka Putri M	93.5	92	88.5	90.5	7.713624				9.000000				6.964194				8.500000				0	0	1	0

Gambar 5.6 Tampilan Window Iterasi K-Means

5.1.7 Hasil Tampilan Window Hasil Clustering

Iterasi ke-2				
C1	C2	C3	C4	C5
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	0	1	0
1	0	0	0	0
0	0	0	0	1
0	0	0	0	1
0	0	0	0	1
0	0	0	0	1
0	0	0	0	1
0	0	1	0	0
0	0	0	0	1
0	0	0	0	1
0	0	0	0	1
0	0	0	0	1

Gambar 5.7 Tampilan Window Hasil Clustering

BAB VI

PENUTUP PENELITIAN

6.1 Kesimpulan

Berdasarkan hasil penelitian yang dilakukan pada Sekolah Menengah Pertama Negeri 2 Kwandang dan pembahasan yang telah diuraikan sebelumnya maka dapat ditarik suatu kesimpulan bahwa :

1. Aplikasi data mining untuk Pengelompokan Data Siswa Berprestasi Dengan Menggunakan Metode *K-Means Clustering* dapat direkayasa.
2. Dapat diketahui bahwa aplikasi data mining untuk Mengelompokan Data Siswa Berprestasi Dengan Menggunakan Metode *K-Means Clustering* yang dirancang dapat digunakan. Hal ini dibuktikan dengan hasil pengujian yang dilakukan dengan metode *White Box Testing* dan *Black Box Testing* yang menggambarkan kebenaran sebuah logika sehingga didapat bahwa logika *Flowchart* benar dan menghasilkan aplikasi data mining untuk *Clustering* yang tepat dan dapat digunakan.

6.2 Saran

Setelah melakukan Penelitian dan Perancangan Aplikasi Mengelompokan Data Siswa Berprestasi pada Sekolah Menengah Pertama dengan menggunakan Metode *K-Means Clustering*, ada beberapa saran yang harus diperhatikan untuk mencapai tujuan yang diharapkan, yaitu sebagai berikut :

1. Agar penelitian ini dapat dikonfigurasi Algoritma Komputasi dan perlu dilakukan Eksperimen terhadap Algoritma lain untuk mendapatkan hasil *Clustering* yang lebih baik lagi.
2. Penulis mengharapkan agar dilakukan penambahan data set dan pusat Cluster agar hasil Cluster lebih banyak.

DAFTAR PUSTAKA

- Amri Muhammad, Aplikasi Pemebeentukan Kelas Dengan K-Means Clustering Sebagai Alat Bantu Pemilihan Siswa Kelas Unggulan Di Man 3 Kediri, 2016.
- Agustinus Purwanto, Pengelompokan dalam kelas yang sama berdasarkan kemampuan akademik, 2015. Amri Muhammad, Aplikasi Pemebeentukan Kelas Dengan K-Means Clustering Sebagai Alat Bantu Pemilihan Siswa Kelas Unggulan Di Man 3 Kediri, 2016.
- Berry, M. W and Browne, M. 2006. Lecture notes in data mining. World Scientific. (Susanto, S. And Suryadi, D. , 2010. Pengantar data mining : menggali pengetahuan dari bongkahan data.)
- Cynthia Paramanda, Analisis Data Konsumen Dengan Menggunakan Metode Clustering Pada Toko Bangunan Rahayu, 2015.
- Ernie Kurniawan, Penerapan Algoritma *K-Means* Untuk *Clustering* dokumen jurnal STIMIK GI MDP, 2016.
- Fina Natasari, *penerapan K-means Clustering Pada Data Penerimaan Mahasiswa Baru* (studi kasus : Universitas Potensi Utama) 2015.
- Jogianto, HM, 1999 Analisis Dan Desain Sistem Informasi : Pendekatan terstruktur teori dan praktek aplikasi bisnis edisi II Yogyakarta : Andi Yogyakarta.
- Larose, D. T. *Discovering Knowledge in data : an introduction to data mining*, Hoboken, New Jersey: John Wiley & Sons, Inc. 2005.
- Nisia Yuanita 2016 Implementasi *K-Means Clustering* Untuk Pembagian Kelas Siswa.Teguh Hariyadi, Penerapan Algoritma *K-Means* Untuk Pengelompokan Data Nilai Siswa, 2013
- Pressman, R.S *Rekayasa Perangkat Lunak : Pendekatan Praktis (Buku 1)* Yogyakarta : Andi Yogyakarta, 2002.
- Rini Agustiningih, Implementasi Prinsip-Prinsip Penilaian Dalam Evaluasi Peelaajaran di TK Negeri 2 Yogyakarta, 2015.

Rohita, Pelaksanaan Penelitian Pembelajaran di Taman Kanak-Kanak (Study Deskriptif pada Taman Kanak-Kanak di Jakarta), 2017.

Shi Na, Liu Xumin, Guan Yong. Algoritma Clustering *K-Means* 2010.

Suhardi Rustam, (*Penerapan Optimasi Jumlah Kluster Pada K-Means Untuk Pengelompokan Kelas Mata Kuliah Kosentrasi Mahasiswa Semester Akhir*), 2020

Teguh Hariyadi, Penerapan Algoritma *K-Means* Untuk Pengelompokan Data Nilai Siswa, 2013.

Teguh Hariyadi, Penerapan Algoritma *K-Means* Untuk Pengelompokan Data Nilai Siswa, 2013. Tinus Septiko, *Aplikasi K-Means Clustering Pada Data rumah tangga* 2013.