

**IMPLEMENTASI DATA MINING UNTUK
PENGELOMPOKAN DATA PENERIMAAN
SISWA BARU MENGGUNAKAN *K-MEANS***

(Studi Kasus Di Madrasah Aliyah Anggrek)

Oleh

ASNUN S LAGARUSU

T3117236

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar Sarjana



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2021**

PENGESAHAN SKRIPSI

**IMPLEMENTASI DATA MINING UNTUK
PENGELOMPOKAN DATA PENERIMAAN
SISWA BARU MENGGUNAKAN *K-MEANS***

(Studi Kasus Di Madrasah Aliyah Anggrek)

Oleh

ASNUN S LAGARUSU

T3117236

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar Sarjana

Program Studi Teknik Informatika,

Ini telah disetujui oleh Tim Pembimbing

Gorontalo, 05 Juni 2021

Pembimbing Utama



Suhardi Rustam S.kom, M.Kom

NIDN 0915088403

Pembimbing Pendamping



Sumarni S.kom, M.Kom

NIDN 0926018604

PERSETUJUAN SKRIPSI

IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN DATA PENERIMAAN SISWA BARU MENGGUNAKAN *K-MEANS*

(Studi Kasus Di Madrasah Aliyah Anggrek)

Oleh

ASNUN S LAGARUSU

T3117236

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

1. Ketua Penguji

Zohrahayaty, M.Kom

2. Anggota

Irma Surya Kumala, M.Kom

3. Anggota

Sarlis Mooduto, M.Kom

4. Anggota

Suhardi Rustam, M.Kom

5. Anggota

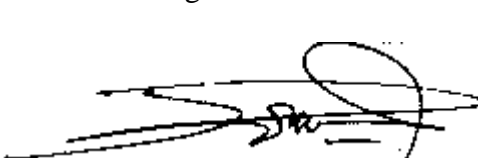
Sumarni, M.Kom

Mengetahui

Dekan Fakultas Ilmu Komputer


Zohrahayaty, M.Kom
NIDN.0912117702

Ketua Program Studi


Irvan Abraham Salihi, M.Kom
NIDN.0928028101

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 diperguruan tinggi lalannya.
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 tilis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis 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 yang telah diperoleh karena karya tulis ini, serta sanksi lainnya dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo

Kwandang, 05 Juni 2021

Yang Membuat Pernyataan



Asnun S Lagarusu

ABSTRACT

Asnun S Lagarusu. T3117236. Implementation Of Data Mining For New Student Admission Data Grouping Using K-Means

This study has the purpose of classifying data on new student admissions and providing convenience for the school in managing student data. The grouping of new student data for the selection of Science and Social Studies classes is very important for students when enrolling for a Senior High School or equivalent level. In the selection of Science and Social Studies classes, it is intended that new students are not wrong in choosing classes or majors. Choosing classes or majors helps students so that they do not get low grades due to wrong class choices in the future. The computational method capable of solving this group classification is the K-means algorithm. K-means algorithm is a clustering algorithm that can group data. The variables used in the calculation of the K-means Algorithm are the values of Mathematics, Indonesian, Science, and Social Studies.

Keywords: data mining, new student admissions, K-means, clustering

ABSTRAK

Asnun S Lagarusu. T3117236. Implementasi Data Mining Untuk Pengelompokkan Data Penerimaan Siswa Baru Menggunakan *K-Means*

Penelitian ini bertujuan untuk mengelompokkan data penerimaan siswa baru serta memberikan kemudahan bagi pihak sekolah dalam mengelola data siswa. Pengelompokkan data siswa baru untuk pemilihan kelas IPA dan IPS sangat penting bagi para siswa yang mendaftarkan diri disekolah menengah atas atau sederajat (SMA). Dalam pemilihan kelas ipa dan ips ini bertujuan agar siswa baru tidak salah dalam pemilihan kelas atau jurusan, dalam pemilihan kelas atau jurusan ini juga dapat membantu siswa agar kedepannya tidak mendapatkan nilai yang rendah akibat salah dalam pemilihan kelas. Metode komputasi yang mampu untuk menyelesaikan pembagian kelompok ini adalah *Algoritma K-means*. *Algoritma K-means* merupakan algoritma *clustering* yang dapat mengelompokkan data. Variabel yang digunakan dalam perhitungan *Algoritma K-means* adalah nilai Matematika, Bahasa Indonesia, Ipa, dan Ips.

Kata kunci: *data mining*, penerimaan siswa baru, *K-means*, *clustering*

KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat dan hidayahnya sehingga penulis dapat menyelesaikan proposal ini dengan tepat waktu. Proposal dengan judul ” Implementasi Data Mining Untuk Pengelompokkan Data Penerimaan Siswa Baru Menggunakan Kmeans Di MA Anggrek ” adapun tujuan dari judul proposal ini adalah untuk merancang sistem dalam mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS menggunakan metode *K-means* serta menerapkan hasil pengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS menggunakan metode *K-means*.

Dalam proses penyusunan proposal ini penulis menjumpai banyak hambatan, namun berkat dukungan dari berbagai pihak akhirnya penulis dapat menyelesaikan tugas ini dengan baik, maka pada kesempatan ini penulis menyampaikan banyak terimakasih kepada:

1. Bapak Muhammad Ichsan Gaffar, SE, M. Ak, Selaku Ketua Yayasan Pembangunan Ilmu Pengetahuan dan Teknologi(YPIPT) Ichsan Gorontalo;
2. Bapak Dr Abdul Gafar Latjokke, M.Si, selaku Rektor Universitas Ichsan Gorontalo;
3. Ibu Zohrahayaty, S.Kom, M.kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
4. Bapak Sudirman S. Panna, S.Kom, M.Kom, selaku Pembantu Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;

5. Ibu Irma Surya Kumala Idris, S.Kom, M.Kom, selaku Pembantu Dekan II Bidang Administrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
6. Bapak Sudirman Melangi, S.Kom, M.Kom, selaku Pembantu Dekan III Bidang Kemahasiswaan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
7. Irvan Abraham Salihi, S.Kom, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
8. Bapak Suhardi Rustam, S.Kom, M.Kom, selaku Pembimbing I;
9. Ibu Sumarni, S.Kom, M.Kom, selaku Pembimbing II;
10. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah membantu dan mendidik serta mengajarkan berbagai disiplin ilmu kepada penulis;
11. Kedua orang tua saya yang sangat saya cintai, atas kasih sayang, jerih payah, doa serta restunya penulis sampai pada titik ini;
12. Rekan rekan saya seperjuangan yang telah banyak membantu dan memberikan dukungan yang sangat besar kepada penulis;
13. Kepada semua pihak yang telah banyak membantu dalam penyelesaian proposal/skripsi ini yang tidak sempat saya sebutkan satu persatu

Semoga apa yang telah diberikan mereka kepada penulis, akan mendapatkan limpahan serta imbalan dari Allah SWT. Penulis menyadari bahwasanya apa yang telah dicapai ini masih jauh dari kata sempurna dan masih banyak kekurangan. Oleh karena itu penulis sangat mengharapkan kritik dan saran. Akhir kata penulis penulis berharap hasil yang telah dicapai ini dapat bermanfaat bagi kita semua, Amin

Gorontalo, 05 Juni 2021

Penulis

DAFTAR ISI

HALAMAN SAMPUL	
HALAMAN JUDUL	i
HALAMAN PENGESAHAN SKRIPSI.....	ii
HALAMAN PERSETUJUAN SKRIPSI.....	iii
HALAMAN PERNYATAAN SKRIPSI.....	iv
ABSTRACT	v
ABSTRAK	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 Masalah	4
1.3 Rumusan Masalah	4
1.4 Tujuan penelitian.....	5
1.5 Manfaat penelitian.....	5
BAB II LANDASAN TEORI	6
2.1 Tinjauan Studi	6
2.2 Tinjauan pustaka	8
2.2.1 siswa.....	8

2.2.2 Dataset.....	8
2.2.3 Data Mining	9
2.2.4 Clustering	10
2.2.5 Algoritma K-means.....	10
2.2.6 Contoh Penerapan Algoritma K-means	12
2.2.7 Siklus Hidup Pengembangan Sistem	19
2.2.8 Perencanaan Sistem.....	20
2.2.9 Analisis Sistem.....	20
2.2.10 Desain Sistem.....	21
2.2.11 Konstruksi Sistem	22
2.2.12 Pemeliharaan Sistem	27
2.2.13 Teknik Pengujian	28
2.2.13.1 White Box	28
2.2.13.2 Black Box	28
2.3 Framework	29
2.4 Pengujian Perangkat Lunak.....	29
2.5 Kerangka Fikir	30
BAB III METODE PENELITIAN	31
3.1 Jenis, Metode, Subjek, Waktu, dan Lokasi Penelitian	31
3.2 Pengumpulan Data	31
3.3 Pemodelan/Abstraksi.....	32
3.3.1 Pengembangan Model.....	32
3.3.2 Evaluasi Model.....	32
3.4 Pengembangan Sistem	33
3.4.1 Analisis sistem	34
3.4.2 Desain sistem	34
3.4.3 Konstruksi sistem.....	35
3.4.4 Pengujian sistem.....	35
BAB IV HASIL PENELITIAN.....	37
4.1 Hasil Pengumpulan Data.....	37
4.1.1 Penerapan Metode K-means	38

4.2 Hasil Pemodelan.....	38
4.2.1 Penjelasan Algoritma	38
4.2.2 Perhitungan Algoritma	40
4.3 Hasil Analisis	48
4.3.1 UML (Unified Modeling Language	48
4.4 Arsitektur Sistem.....	54
4.5 Interface Design	54
4.5.1 Mekanisme User	54
4.5.2 Mekanisme Navigasi.....	54
4.5.3 Mekanisme Login	55
4.5.4 Mekanisme Input Data Siswa	55
4.5.5 Mekanisme Output.....	56
4.6 Data Design.....	56
4.6.1 Struktur Data.....	56
4.7 Hasil Pengujian Sistem	60
4.7.1 Pengujian White Box	60
4.7.2 Flowchart	61
4.7.3 Flowgraph	62
4.7.4 Hasil Pengujian Black Box	63
BAB V HASIL DAN PEMBAHASAN	65
5.1 Pembahasan Sistem.....	65
5.1.1 Tampilan Halaman Login	65
5.1.2 Tampilan Halaman Data Siswa.....	66
5.1.3 Tampilan Halaman Iterasi K-means	66
5.1.4 Tampilan Halaman Cluster	67
BAB VI PENUTUP	69
6.1 Kesimpulan	69
6.2 Saran.....	69
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 2.1 Siklus Hidup Pengembangan Sistem	19
Gambar 2.2 Bagan Alir	28
Gambar 2.3 Kerangka Pikir	30
Gambar 3.1 Sistem Yang Diusulkan	33
Gambar 4.1 Use Case <i>Diagram Admin Clustering</i> Siswa.....	48
Gambar 4.2 <i>Activity Diagram Login Admin</i>	49
Gambar 4.3 <i>Activity Diagram Login</i> Keoala Sekolah.....	50
Gambar 4.4 <i>Activity Diagram</i> Data Siswa baru	51
Gambar 4.5 <i>Class Diagram</i>	52
Gambar 4.6 <i>Sequence Diagram Login</i>	53
Gambar 4.7 <i>Sequence Diagram</i> Data Siswa	53
Gambar 4.8 Mekanisme <i>Nvigasi Home Admin</i>	54
Gambar 4.9 Mekanisme <i>Login</i>	55
Gambar 4.10 Mekanisme <i>Input</i> Data Siswa.....	55
Gambar 4.11 Mekanisme <i>Output</i>	56
Gambar 4.12 Gambar <i>Flowchart</i>	61
Gambar 4.13 Gambar <i>Flowgraph</i>	62
Gambar 5.1 Halaman <i>Login</i>	65
Gambar 5.2 Halaman Data Siswa	66
Gambar 5.3 Halaman <i>Iterasi K-means</i>	66
Gambar 5.4 Halaman Hasil <i>Clustering</i>	67

DAFTAR TABEL

Table 2.1 Data Set	1
Tabel 2.2 Penelitian Terkait	6
Tabel 2.3 Data Set.....	8
Tabel 2.4 Data Mahasiswa Baru Yang Diperoleh	12
Tabel 2.5 Data Hasil <i>Transformasi</i>	14
Tabel 2.6 Tiitk Pusat Awal <i>Cluster</i>	15
Tabel 2.7 Hasil Perhitungan Setiap Data Kesetiap <i>Cluster</i> iterasi 1	16
Tabel 2.8 Titik Pusat Iterasi-1 Setelah <i>Cluster</i>	18
Tabel 2.9 Hasil Perhitungan Kesetiap Cluster Pada Iterasi	18
Tabel 2.10 <i>Simbol Case Diagram</i>	22
Tabel 2.11 Simbol <i>Actifity Diagram</i>	24
Tabel 2.12 Simbol <i>Class Diagram</i>	25
Tabel 2.13 Simbol <i>Sequence Diagram</i>	27
Tabel 2.14 Simbol Fungsi PHP Berhubungan Dengan Database MySQL	29
Tabel 3.1 <i>Atribut Data</i>	32
Tabel 4.1 Hasil Pengumpulan <i>Dataset</i>	37
Tabel 4.2 Sempel Data Set Siswa Baru.....	40
Tabel 4.3 Penentuan Awal <i>Cluster</i>	41
Tabel 4.4 Menacari Jarak Terdekat.....	44
Tabel 4.5 Kelompok Pembagian Data 1	45
Tabel 4.6 Penentuan <i>Cluter</i> Baru.....	47
Tabel 4.7 Kelompok Pembagian Data 2	47

Tabel 4.8 Mekanisme <i>User</i>	54
Tabel 4.9 Data Nilai Matematika	56
Tabel 4.10 Data Nilai Bahasa Indonesia	57
Tabel 4.11 11 Data Nilai IPA.....	57
Tabel 4.12 11 Data Nilai IPS	58
Tabel 4.13 11 Data Siswa.....	58
Tabel 4.14 Asal Sekolah	59
Tabel 4.15 Proses Design.....	59
Tabel 4.16 <i>Bsis Path</i>	63
Tabel 4.17 Pengujian <i>Black Box Admin</i>	63
Tabel 4.17 Pengujian <i>Black Box Kepala Sekolah</i>	64

BAB I

PENDAHULUAN

1.1 Latar Belakang

Siswa merupakan seorang pelajar, yang belajar agar memperoleh ilmu pengetahuan untuk dapat menggapai mimpi dan cita-citanya. Siswa itu sendiri merupakan seseorang yang menempuh pembelajaran didunia pendidikan dengan tujuan mendapatkan ilmu pengetahuan serta menjadi manusia yang berakhlak mulia dan madiri.

Dengan adanya kurikulum 2013, disini tidak semua sekolah sudah menarapkan kurikulum 2013, hanya sebagian kecil sekolah yang sudah menerapkan kurikulum 2013 tersebut. Seperti Madrasah Aliyah Anggrek, sekolah ini sudah menerapkan kurikulum 2013. Dimana siswa baru kelas 1 atau X yang tadinya penjuruan/pengelompokkan kelas IPA dan IPS dilakukan di kelas 2 atau XI, sekarang sudah mulai diterapkan ketika masuk kelas X atau siswa baru.

Tabel 2.1 Data set

NO	NAMA SISWA	Asal Sekolah	MATEMATIKA	BHS INDONESIA	IPA	IPS
1	Alfi Raja	MTs Muh.Monano	82,22	80	86,6	85,16
2	Fahri Sulubani	MTs Muh.Monano	86,52	81,51	84,5	86,5
3	Farut Lakoro	SMP N 3 Monano	81,4	83,8	86,16	89,83
4	Iston Lakoro	SMP N 3 Monano	83	87,33	85	88,16
5	Jufrin S. Mada	SMP N 3 Monano	87,16	88,33	86,66	88,66
6	Mohamad Walidi S. Payu	MTs Muh.Monano	84,33	87,33	88,33	87,66
7	Adelia Hasu	MTs Muh.Monano	86,5	89,33	88	90
8	Desmita Gobel	MTs Muh.Monano	87,5	88	88,5	87,5
9	Ismi S. Otoluwa	MTs Muh.Monano	84,66	89,66	86,33	88,33

10	Kadita Putri Latif	MTs Muh.Monano	90,66	87,66	86,16	86,83
11	Kristin Raja	MTs Muh.Monano	85,33	89	84	86,16
12	Nurlela Otoluwa	SMP N 3 Monano	86,16	88,16	88,33	88,33
13	Puspita Kai	MTs Muh.Monano	86,66	89,5	87,66	90,5
14	Rini Anggriani Paramata	MTs Muh.Monano	84,16	88	89	86,33
15	Siska T. Dou	SMP N 2 Monano	80	87,6	88,16	87,83
16	Siti Azizah Pakaya	MTs Muh.Monano	86,5	88	89,5	87
17	Yulinda H. Kuka	SMP N 3 Monano	88,33	87,5	88	90,33
18	Abdul Wahid Arafa	SMP N 1 Anggrek	85,83	88,33	87,33	90,16
19	Aldi Lakoro	SMP N 3 Monano	87,83	86,83	87,5	87,16
20	Fandi J. Masita	SMP N 4 Monano	89,33	85,33	88	86,66

sumber: Arsip Madrasah Aliyah Anggrek)

Dari tabel data set diatas menunjukkan bahwa kita bisa mengelompokkan data penerimaan siswa baru dalam Pemilihan Kelas IPA dan IPS berdasarkan nilai rata-rata rapor mata pelajaran Matematika, Bahasa Indonesia, IPA dan IPS. dilakukannya pengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS pada siswa baru adalah untuk mencegah nilai siswa yang menjadi rendah kedepannya diakibatkan salah dalam pemilihan kelas. Sehingga penulis tertarik untuk memberikan suatu solusi dengan membangun suatu sistem untuk mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS berdasarkan nilai rata-rata rapor untuk mata pelajaran Matematika, Bahasa Indonesia, IPA dan IPS siswa baru menggunakan Algoritma K-means, di mana metode ini sangat dibutuhkan pihak sekolah dalam mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS.

Algoritma yang digunakan dalam penelitian ini adalah algoritma K-means, yang dimana algoritma kmeans ini merupakan salah satu algoritma dalam *data mining* yang bisa digunakan untuk melakukan pengelompokan/*clustering* suatu data. Kelebihan dari algoritma *K-means* adalah mudah untuk diimplementasikan dan dijalankan .waktuyang dibutuhkan untuk menjalankan pembelajaran relatif cepat, mudah diaptasi, umum digunakan.

Berdasarkan penelitian yang dilakukan oleh Fitri Yunita (2018) dengan judul ”Penerapan *Data Mining* Menggunakan *Algoritma K-means Clustering* Pada Penerimaan Mahasiswa Baru” berdasarkan penelitian ini dapat dilihat bahwa hasil cluster algoritma *K-means* dapat mengelompokkan data penerimaan mahasiswa baru serta strategi promosi bagi mahasiswa baru, dan untuk dapat meningkatkan minat studi yang akan di ambil oleh masing masing sekolah.

Selanjutnya penelitian yang dilakukan oleh Fina Nasari (2015)dengan judul “Penerapan *K-means Clustering* Pada Data Penerimaan Mahasiswa Baru” berdasarkan pada penelitian ini menjelaskan bahwa pengelompokan data yang digunakan adalah *algoritma K-means Clustering*, *K-means Clustering* mampu

mengelompokkan data pada kelompok yang sama dan data yang berbeda pada kelompok yang berbeda.

Selantunya penelitian yang dilakukan oleh Juliana Tandy (2019) dengan judul “Analisis Dan Rancangan *Clustering* Siswa Baru Menggunakan Metode *K-means* Pada SMK Negeri 1 Kota Jambi” dalam penelitian ini menjelaskan bahwa dalam proses penyeleksian dan penerimaan membutuhkan aplikasi yang lebih efisien yaitu menggunakan aplikasi, merubah sistem manual menjadi sistem komputerisasi.

Penelitian ini dilaksanakan di Madrasah Aliyah Anggrek yang beralamatkan di desa monas kecamatan monano. Untuk penelitian ini adalah seluruh data siswa baru Tahun Ajaran 2015/2016 di Madrasah Aliyah anggrek desa monas kecamatan monano. Berdasarkan uraian diatas, dalam penelitian ini penulis melakukan penelitian dengan judul **“Implementasi Data Mining Untuk Pengelompokan Data Penerimaan Siswa Baru Menggunakan K-means”**.

1.2 Identifikasi Masalah

- a. Sekolah masih kesulitan dalam mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS.
- b. Belum adanya sistem untuk mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS.

1.3 Rumusan Masalah

Berikut rumusan masalah berdasarkan uraian latar belakang dan identifikasi masalah diatas yaitu:

- a. Bagaimana cara merancang sistem dalam mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS menggunakan metode *K-means*?
- b. Bagaimana cara penerapan sistem dalam mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS dengan metode *K-means*.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

- a. Untuk merancang sistem dalam mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS menggunakan metode *K-mean*.
- b. Untuk menerapkan hasil pengelompokan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS menggunakan metode *K-means*.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian adalah:

- a. Penelitian ini diharapkan dapat memberikan masukan terhadap pengembangan ilmu pengetahuan dibidang teknologi komputer dan menerapkan data mining khususnya metode *K-means*.
- b. Diharapkan aplikasi ini dapat bermanfaat bagi pihak sekolah dan ini dapat digunakan oleh guru yang ada di Madrasah Aliyah Anggrek Kecamatan Monano Kabupaten Gorontalo Utara untuk mengetahui cara mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS menggunakan metode *K-means*.
- c. Sebagai bahan referensi dan masukan bagi peneliti lain yang mengadakan penelitian selanjutnya tentang pengelompokan data menggunakan metode *K-means*.

BAB II

LANDASAN TEORI

2. 1. Tinjauan Studi

Terdapat beberapa penelitian yang terkait dengan penelitian tentang pengelompokan data siswa baru menggunakan metode kmean yang juga menjadi acuan bagi penulis diantaranya.

Table 2.2 Penelitian Terkait

NO	PENELITI	JUDUL	TAHUN	METODE	HASIL
1.	Fitri Yunita	Penerapan data mining menggunakan algoritma k-means clustering pada penerimaan mahasiswa baru (studi kasus : universitas islam indragiri)	2018	K-means Clustering	Dengan menggunakan Metode K-means clustering tersebut maka dapat ditentukan beberapa strategi promosi dalam penerimaan mahasiswa baru yang dapat dilakukan oleh pihak universitas islam Indragiri dalam melakukan promosi program studi dimasing masing sekolah

NO	PENELITI	JUDUL	TAHUN	METODE	HASIL
2.	Vina Nasari	Penerapan k-means clustering pada data penerimaan mahasiswa baru (studi kasus : universitas potensi utama)	2015	K-means Clustering	Dengan menggunakan Metode K-means clustering hasil yang didapatkan berfungsi sebagai penimbang penerimaan mahasiswa baru tiap tahun ajaran dan menjadikan sistem kerja yang efisien bagi pihak sekolah
3.	Juliana Tandi	Analisis dan perancangan an clustering siswa baru menggunakan metode k-means pada SMK negeri 1 kota jambi	2019	K-means Clustering	Dengan menggunakan Metode K-means Clustering tersebut maka hasil yang didapatkan berfungsi sebagai penimbang penerimaan siswa baru tiap tahun ajaran dan menjadikan sistem kerja yang efisien bagi pihak sekolah

2. 2. Tinjauan Pustaka

2.2.1 Siswa

Siswa merupakan seorang pelajar yang menempuh pendidikan untuk belajar serta untuk memperoleh ilmu pengetahuan dengan tujuan untuk menggapai mimpi dan cita-citanya. Selain itu siswa juga adalah seorang pelajar yang berusaha mengembangkan potensi diri melalui proses pembelajaran dengan jalur pendidikan sehingga menjadi manusia yang berkualitas sesuai dengan tujuan pendidikan nasional.

2.2.2 Data set

Adapun atribut yang digunakan dalam pengimputan data penerimaan siswa baru ditunjukkan pada Tabel dibawah ini.

Tabel 2.3 Data set

NO	NAMA SISWA	Asal Sekolah	MATEMATIKA	BHS INDONESIA	IPA	IPS
1	Alfi Raja	MTs Muh.Monano	82,22	80	86,6	85,16
2	Fahri Sulubani	MTs Muh.Monano	86,52	81,51	84,5	86,5
3	Farut Lakoro	SMP N 3 Monano	81,4	83,8	86,16	89,83
4	Iston Lakoro	SMP N 3 Monano	83	87,33	85	88,16
5	Jufrin S. Mada	SMP N 3 Monano	87,16	88,33	86,66	88,66
6	Mohamad Walid S. Payu	MTs Muh.Monano	84,33	87,33	88,33	87,66
7	Adelia Hasu	MTs Muh.Monano	86,5	89,33	88	90
8	Desmita Gobel	MTs Muh.Monano	87,5	88	88,5	87,5
9	Ismi S. Otoluwa	MTs Muh.Monano	84,66	89,66	86,33	88,33
10	Kadita Putri Latif	MTs Muh.Monano	90,66	87,66	86,16	86,83
11	Kristin Raja	MTs Muh.Monano	85,33	89	84	86,16
12	Nurlela Otoluwa	SMP N 3 Monano	86,16	88,16	88,33	88,33

13	Puspita Kai	MTs Muh.Monano	86,66	89,5	87,66	90,5
14	Rini Anggriani Paramata	MTs Muh.Monano	84,16	88	89	86,33
15	Siska T. Dou	SMP N 2 Monano	80	87,6	88,16	87,83
16	Siti Azizah Pakaya	MTs Muh.Monano	86,5	88	89,5	87
17	Yulinda H. Kuka	SMP N 3 Monano	88,33	87,5	88	90,33
18	Abdul Wahid Arafa	SMP N 1 Anggrek	85,83	88,33	87,33	90,16
19	Aldi Lakoro	SMP N 3 Monano	87,83	86,83	87,5	87,16
20	Fandi J. Masita	SMP N 4 Monano	89,33	85,33	88	86,66

(sumber: Arsip Madrasah Aliyah Anggrek)

2.2.3 Data Mining

Data mining yaitu proses yang paling penting dimana akan dilakukan berbagai teknik yang diaplikasikan untuk mengekstrak berbagai pola-pola potensial untuk mendapatkan data yang berguna. *Data mining dan Knowledge discovery in database* (KDD) seringkali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu bisnis data yang besar. Dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining (Nasari et al, 2015).

Proses KDD secara garis besar dapat dijelaskan sebagai berikut :

1. Data Selection, pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining disimpan dalam satu berkas, terpisah dari basis data operasional.
2. *Pre-processing / cleaning*, sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses pembersihan pada data yang menjadi fokus KDD. Proses pembersihan mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (typografi).

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

2.2.4 Clustering

Clustering merupakan proses pengelompokan data, Pengelompokan (*clustering*) merupakan bagian dari ilmu data mining yang bersifat tanpa arahan (*unsupervised*). Clustering adalah proses pembagian data ke dalam kelas atau cluster berdasarkan tingkat kesamaannya. Dalam *clustering*, data yang memiliki kesamaan dimasukkan ke dalam cluster yang sama, sedangkan data yang tidak memiliki kesamaan dimasukkan dalam cluster yang berbeda (Khotimah, Teknik, Studi, Informatika, & Kudus, 2014).

2.2.5 Algoritma K-means

Algoritma kmeans ini merupakan salah satu algoritma dalam *data mining* yang bisa digunakan untuk melakukan pengelompokan/*clustering* suatu data. *K-means Clustering* adalah K dimaksudkan sebagai konstanta jumlah cluster yang diinginkan, *Means* dalam hal ini berarti nilai suatu rata-rata dari suatu group data yang dalam hal ini didefinisikan sebagai cluster, sehingga *K-mean clustering* adalah suatu metode penganalisaan atau metode data mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan system partisi. Metode *K-mean* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada didalam kelompok yang lain. *Algoritma K-mean* merupakan algoritma yang membutuhkan parameter input sebanyak k dan membagi sekumpulan n objek kedalam k cluster sehingga tingkat kemiripan antar anggota dalam suatu cluster tinggi sedangkan tingkat kemiripan dengan anggota cluster lain sangat rendah. Kemiripan anggota terhadap cluster diukur dengan kedekatan objek terhadap nilai mean pada *cluster*

atau dapat disebut sebagai *centroid cluster* atau pusat massa (Khotimah et al.,2014).

Berikut adalah rumus untuk menentukan jarak data dari masing-masing centroid :

::

$$d(P, Q) = \sqrt{\sum_{j=1}^p (x_j(P) - x_j(Q))^2}$$

Keterangan :

D = titik dokumen

P = data record

Q = data centroid

Jarak yang terpendek antara centroid dengan dokumen menentukan posisi cluster suatu dokumen.

$$C(i) = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x}$$

Keterangan :

x1 = Nilai data record ke-1

x2 = Nilai data record ke-2

$\sum x$ = jumlah data record

Eksperimen ini menggunakan algoritma yang paling umum digunakan dalam clustering yaitu algoritma *K-Means*. Algoritma ini populer karena mudah diimplementasikan dan kompleksitas waktunya linear. Kelemahannya adalah algoritma ini sensitive terhadap inisialisasi cluster (Langgeni, Baizal, & W, 2010).

Dasar algoritmanya adalah sebagai berikut :

1. Inisialisasi cluster
2. Masukkan setiap dokumen ke *cluster* yang paling cocok berdasarkan ukuran kedekatan dengan centroid. *Centroid* adalah vector term yang dianggap sebagai titik tengah cluster.
3. Setelah semua dokumen masuk ke *cluster*. Hitung ulang centroid cluster berdasarkan dokumen yang berada didalam cluster tersebut.
4. Jika centroid tidak berubah (dengan threshold tertentu) maka stop. Jika tidak kembali ke langkah 2.

2.2.6 Contoh penerapan *Algoritma K-means*

Berikut ini merupakan contoh penerapan algoritma K-means yang penulis ambil dari peneliti sebelumnya yaitu Fina Nasari dengan judul “Penerapan K-means Clustering Pada Data Penerimaan Mahasiswa Baru”. Table 2.3 merupakan data uji yang digunakan sebagai data pendukung sistem berjumlah 20 data.

Table 2.4 Data Mahasiswa Baru Yang Diperoleh

No	NAMA	FAKULTAS	PRODI	ASAL SEKOLAH	NILAI UAN
1.	Ade Saputra	TIK	Teknik Sipil	MA Nurul Huda	44,6
2.	Kamariddin	TIK	Teknik Sipil	SMAN 1 Kota Gasib	34,8
3.	Reza Try Aditya	TIK	Teknik Sipil	SMAN 1 Tembilahan	33,4
4.	M. Yoga Ali Akbar	TIK	Teknik Sipil	SMAN 1 Keritang	46,4
5.	Andi Septiyawan	TIK	Teknik Sipil	SMAN 2 Tembilahan	36,4
6.	Zulkifli	TIK	Teknik Sipil	BLM Ada Ijazah	45,7
7.	Aprizal	TIK	Teknik Sipil	SMAN 2 Tembilahan	33,2

8.	Fitria Nurma Yunita	TIK	Sistem Informasi	SMAN 1 Tembilahan Hulu	45,7
9.	Felby Irsyahdi	TIK	Sistem Informasi	SMAN 1 Tembilahan Hulu	33,9
10.	Ari Anggara	TIK	Sistem Informasi	BLM Ada Ijazah	47,4
11.	Muhammad Budiman	TIK	Sistem Informasi	SMAN 1 Tembilahan	42,5
12.	Ruzimah	TIK	Sistem Informasi	SMAN 1 Tembilahan Hulu	33,6
13.	Hariyadi	TIK	Sistem Informasi	BLM Ada Ijazah	56,3
14.	Mahmuddin	TIK	Sistem Informasi	MA Ummul Quro Al-Islami Bogor	45,3
No	NAMA	FAKULTAS	PRODI	ASAL SEKOLAH	NILAI UAN
15.	Irwandi Rizki Putra	TIK	Sistem Informasi	SMAN Dharma Pendidikan Kempas	32,4
16.	Mahmud Rizky Armi	TIK	Sistem Informasi	SMKN 2 Tembilahan	51,3
17.	Firman Iswandi	TIK	Sistem Informasi	SMAN 1 Pariaman	44,6
18.	Teja Sekti Wicaksomo	TIK	Sistem Informasi	SMAN 1 Tembilahan Hulu	34,8
19.	Muhammad Ilyas	TIK	Sistem Informasi	SMAN 1 Tanah Merah	33,4
20.	M. Ansyari Saputra		Sistem Informasi	SMAN 1 Tembilahan Hulu	46,4

(sumber: Fitri Yunita, 2018)

1. Transformasi Data

- Program studi sistem informasi ditransformasikan menjadi nilai 1
- Program studi teknik sipil ditransformasikan menjadi nilai 2
- Program studi teknik informatika ditransformasikan menjadi nilai 3
- Asal sekolah SMA ditransformasikan menjadi nilai 1
- Asal sekolah SMK ditransformasikan menjadi nilai 2
- Asal sekolah selain SMA dan SMK ditransformasikan menjadi nilai 3
- Nilai UAN dengan rata-rata ≤ 5 ditransformasikan menjadi nilai 1
- Nilai UAN dengan rata-rata < 5 ditransformasikan menjadi nilai 2
- Nilai UAN dengan rata-rata ≤ 7 ditransformasikan menjadi nilai 3

Berikut adalah data hasil transformasi

Table 2.5 Data Hasil Transformasi

Id	Prodi	Asal Sekolah	Nilai UAN
K1	2	3	2
K2	2	1	2
K3	2	1	2
K4	2	1	2
K5	2	1	2
K6	2	3	2
K7	2	1	2
K8	1	1	2
K9	1	1	1
K10	1	2	1
K11	1	1	2
K12	1	1	1
K13	1	2	1
K14	1	3	2
K15	1	1	1
K16	1	2	1
K17	1	2	1

K18	1	1	2
K19	1	1	2
K20	1	1	1

(sumber: Fitri Yunita, 2018)

Fitri Yunita(2018) perhitungan berdasarkan data uji yang ada dijabarkan sebagai berikut :

2. Penentuan pusat awal *cluster*

Table 2.6 Titik Pusat Awal Cluster

Titik Pusat			
Centroid 1	2	1	2
Centroid 2	2	3	2
Centroid 3	2	1	1

(sumber: Fitri Yunita, 2018)

3. Perhitungan jarak claster

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2}$$

persamaan diatas digunakan karena atribut yang digunakan berjumlah 3.

Sebagai contoh, akan dihitung jarak dari data mahasiswa pertama ke pusat *cluster* pertama dengan persamaan:

$$\begin{aligned} d(1,1) &= \sqrt{(p_1 - q_1)^2 + (p_2 - p_2)^2 + (p_3 - p_3)^2} \\ &= \sqrt{(2 - 2)^2 + (3 - 1)^2 + (2 - 2)^2} \\ &= 2 \end{aligned}$$

Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data mahasiswa baru pertama dengan *cluster* pertama adalah 2. Jarak data mahasiswa baru pertama ke pusat *cluster* kedua dengan persamaan:

$$\begin{aligned} d(1,2) &= \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \\ &= \sqrt{(2 - 2)^2 + (3 - 3)^2 + (2 - 2)^2} \\ &= \sqrt{(0)^2 + (0)^2 + (0)^2} = 0 \end{aligned}$$

Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data mahasiswa baru pertama ke *cluster* kedua adalah 0. Jarak data mahasiswa baru pertama ke pusat *cluster* ketiga dengan persamaan:

$$\begin{aligned} d(1,3) &= \sqrt{(p1 - q1)^2 + (p2 - q2)^2 + (p3 - q3)^2} \\ &= \sqrt{(2 - 2)^2 + (1 - 1)^2 + (2 - 1)^2} \\ &= \sqrt{(0)^2 + (0)^2 + (1)^2} = 2,2 \end{aligned}$$

Dari hasil perhitungan diatas didapatkan hasil bahwa jarak data mahasiswa baru pertama ke *cluster* ketiga adalah 2,2.

Table 2.7 Hasil Perhitungan Setiap Data Kesetiap *Cluster* iterasi 1

Data ke-i	Jarak Ke Centroid			Jarak Terdekat	Cluster diikuti
	1	2	3		
K1	2	0	2,2	0	2
K2	0	2	1	0	1
K3	0	2	1	0	1
K4	0	2	1	0	1
K5	0	2	1	0	1
K6	2	0	2,2	0	2
K7	0	2	1	0	1
K8	1	2,2	1,4	1	1
K9	1,4	2,4	1	1	3
K10	1,7	1,7	1,4	1,4	3
K11	1	2,2	1,4	1	1
K12	1,4	2,4	1	1	3
K13	1,7	1,7	1,4	1,4	3
K14	2,2	1	2,4	1	2
K15	1,4	2,4	1	1	3
K16	1,7	1,7	1,4	1,4	3
K17	1,7	1,7	1,4	1	3
K18	1	2,2	1,4	1	1
K19	1	2,2	1,4	1	1

K20	1,4	2,4	1	1	3
-----	-----	-----	---	---	---

(sumber: Fitri Yunita, 2018)

4. Pengelompokan data

Untuk mengelompokan hasil claster pada iterasi pertama yang belum kovergen. Untuk membangkitkan kembali centroid baru dengan rumus berikut:

$$C = \frac{\sum m}{n}$$

Dimana:

C: centroid data

M: anggota data yang termasuk kedalam *centroid* tertentu

N: jumlah data yang menjadi anggota *centroid* tertentu

Contoh perhitungan pada cluster 1 adalah sebagai berikut:

$$\frac{2+2+2+2+2+1+1+1+1+1+3+3}{12} = 1,5$$

$$\frac{1+1+1+1+1+1+1+1+1+1+1+1}{12} = 1$$

$$\frac{2+2+2+2+2+2+2+2+2+2+2+2}{12} = 2$$

Contoh perhitungan pada claster 2 adalah sebagai berikut:

$$\frac{2+2+1+3+3}{5} = 2,2$$

$$\frac{3+3+3+3+3}{5} = 3$$

$$\frac{2+2+2+1+2}{5} = 1,8$$

Contoh perhitungan pada cluster adalah sebagai berikut:

$$\frac{2+2+2+2+2+2+1+1+1+3+3}{10} = 1,4$$

$$\frac{1+2+1+2+1+2+2+1+2+2}{10} = 1,6$$

$$\frac{1+1+1+1+1+1+1+1+1+1}{10} = 1$$

Pada pengelompokan data dengan menggunakan rumus diatas didapatkan titik pusat cluster dengan nilai sebagai berikut:

Table 2.8 Titik pusat iterasi-1 setelah cluster

Titik Pusat			
Centroid 1	1,5	1	2
Centroid 2	1,2	3	1,8
Centroid 3	1,4	1,6	1

(sumber: Fitri Yunita, 2018)

Table 2.9 Hasil perhitungan dan ke setiap cluster pada iterasi

Data ke-i	Jarak Ke Centroid			Jarak Terdekat	Cluster diikuti
	1	2	3		
K1	2,06	0,3	1,82	1,0969	2
K2	0,5	2	1,31	0,1863	1
K3	0,5	2	1,31	0,1863	1
K4	0,5	2	1,31	0,1863	1
K5	0,5	2	1,31	0,1863	1
K6	2,06	0,3	1,82	0,3	2
K7	0,5	2	1,31	0,5	1
K8	0,5	2,3	1,23	0,5	1
K9	1,12	2,5	0,72	0,72	3
K10	1,5	1,8	0,57	0,57	3
K11	0,5	2,3	1,23	0,5	1
K12	1,12	2,5	0,72	0,72	3
K13	1,5	1,8	0,57	0,57	3
K14	2,06	1,2	1,77	1,2	2
K15	1,12	2,5	0,72	0,72	3
K16	1,5	1,8	0,57	0,57	3
K17	1,5	1,8	0,57	0,57	3
K18	0,5	2,3	1,23	0,5	1
K19	0,5	2,3	1,23	0,5	1
K20	1,12	2,5	0,72	0,72	3

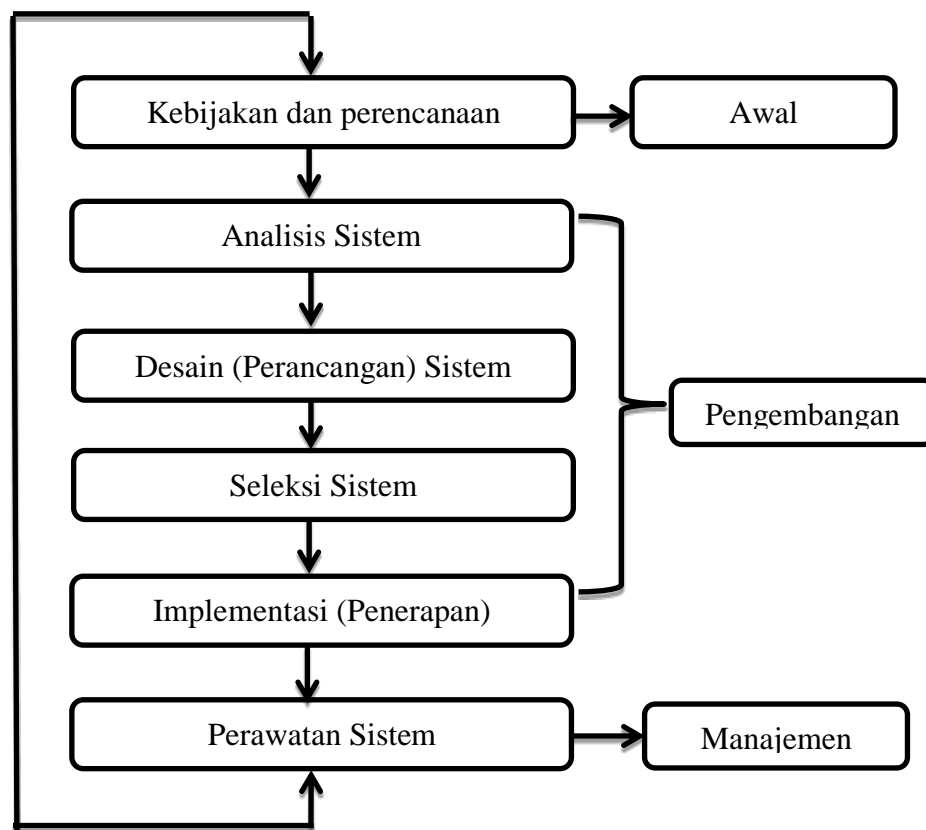
K21	0,5	2,3	1,23	0,5	1
K22	2,06	1,5	1,65	1,5	2
K23	1,5	2,2	1,98	1,5	1
K24	2,69	1,1	2,13	1,1	2
K25	1,5	2,2	1,98	1,5	1

(sumber: Fitri Yunita, 2018)

Pada iterasi ketiga, titik pusat dari setiap *cluster* sudah tidak berubah dan tidak ada lagi data yang berpindah dari satu *cluster* ke *cluster* yang lain.

2.2.7 Siklus Hidup Pengembangan sistem

Siklus hidup pengembangan sistem yaitu dalam rekayasa sistem, rekayasa perangkat lunak merupakan proses pembuatan dan perubahan sistem serta model dan metodologi yang digunakan untuk mengembangkan sistem tersebut. Siklus hidup pengembangan sistem dengan langkah-langkah yang akan digunakan adalah sebagai berikut:



Gambar 2.1 Siklus Hidup Pengembangan Sistem

Sumber : Jogianto, HM, (2005 : 52)

2.2.8 Perencanaan Sistem

Perencanaan sistem merupakan kata dari sebuah konsep, dimana dalam pengembangan suatu sistem/perangkat lunak konseptualisasi ini dilakukan dengan maksud tujuan tertentu :

Heriyanto (2004 : 353) mengungkapkan :

“Tujuan konseptualisasi adalah untuk menghabiskan spesifikasi perilaku system yang disepakati antara pembeli dan pengembang, pemakai dan stakeholder lain serta merupakan kontrak resmi pengembangan dan client, juga menjadi dokumen yang menuntut program dalam implementasi sistem”. Yang mana pada tahap perencanaan ini pengembang melakukan observasi untuk mengenai calon pengguna dari sistem informasi/perangkat lunak yang akan dikembangkan nantinya. Pada pengembangan sistem/perangkat lunak berorientasi objek menggunakan UML sebagai kelas (tool), semua permasalahan dimodelkan sebagai user case untuk menggambarkan seluruh kebutuhan-kebutuhan pengguna.

2.2.9 Analisis Sistem

Analisis sistem atau System Analysis adalah suatu teknik atau metode pemecahan masalah dengan cara menguraikan system ke dalam komponen-komponen tersebut bekerja dan saling berinteraksi satu sama lain untuk mencapai tujuan system.

Whitten, et, al. (2004 : 33) mengungkapkan “System analisis adalah studi domain masalah bisnis untuk merekomendasikan perbaikan dan menspesifikasikan persyaratan dan prioritas bisnis untuk solusi”

Didalam menganalisis sistem ada beberapa langkah dasar yang harus diperhatikan dan dilakukan oleh seorang analisis sistem yaitu :

1. *Identify*; yaitu mengidentifikasi masalah yang merupakan langkah pertama, yang dilakukan dalam tahap analisis sistem. Masalah (problems)

dapat diidentifikasi sebagai suatu pertanyaan yang diinginkan untuk dipecahkan.

2. *Understand*; yaitu memahami kerja dari sistem yang ada, dari tahap analisis sistem adalah harus memahami kerja dari sistem yang ada, Langkah ini dapat dilakukan dengan mempelajari secara terinci bagaimana sistem yang ada beroperasi. Untuk lebih mempelajari operasi dari sistem ini diperlukan data yang dapat diperoleh dengan cara melakukan penelitian.
3. *Analyze*; yaitu dengan 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 yakni pelaporan, bahwa analisis telah selesai dilakukan. Yang kedua meluruskan kesalahan pengertian mengenai apa yang telah ditemukan dan dianalisis oleh sistem tetapi tidak sesuai menurut manajemen.

2.2.10 Desain Sistem

Desain sistem dapat didefinisikan sebagai penggambaran, perencanaan, dan pembuatan sketsa atau pengaturan dari beberapa elemen yang terpisah kedalam satu kesatuan yang utuh dan berfungsi.

Whitten, et, al. (2004 : 34) mengungkapkan : “System design adalah spesifikasi atau indtruksi solusi yang teknis dan berbasis computer dan persyaratan bisnis yang didefinisikan dalam analisis sistem”.

Desain sistem dibagi menjadi dua bagian, yakni desain sistem secara umum dan desain sistem secara terinci:

1. Desain sistem secara umum, pada tahap ini yakni tahap desain sistem secara umum, komponen-komponen sistem informasi yang dirancang dengan tujuan dikomunikasikan kepada *user* bukan untuk pemrograman. Komponen sistem informasi yang didesain adalah *mode*, *output*, *input*, *database*, *teknologi* dan *control*.

- Desain model secara umum, analisis sistem dapat mendesain model dari sistem informasi yang diusulkan dalam bentuk physical sistem dan logical model. Bagan alir sistem merupakan alat yang tepat digunakan untuk menggambarkan physical sistem. Logical model dapat digambarkan dengan diagram arus data.


2.2.11 Konstruksi Sistem



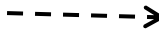
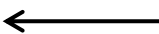

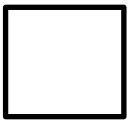

Konstruksi sistem adalah pengembangan sistem, instalasi sistem, dan pengujian komponen sistem. Pada tahap konstruksi di penelitian ini, penulis menggunakan UML (Unified Modeling Language) sebagai alat bantu. UML atau (Unified Modeling Language) adalah suatu metode permodelan secara visual yang berfungsi sebagai sarana perancangan sistem berorientasi objek. Unified Modeling Language digunakan untuk berbagai diagram khusus yang dapat dipahami dan digambarkan oleh analisis dan pengguna akhir yang digunakan dalam proyek pengembangan sistem.

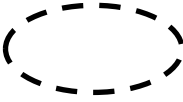
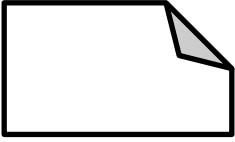
Berikut ini adalah model komponen sistem yang menggunakan Unified Modeling Language (UML) yaitu :

1. *Use Case Diagram* adalah diagram yang menggambarkan hubungan antara actor dengan sistem. *Use case* diagram bisa mendeskripsikan sebuah interaksi antar satu atau lebih actor dengan sistem yang akan dibuat. *Use case* diagram juga bisa digunakan untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem dan bisa juga mempresentasikan sebuah interaksi actor dengan sistem.

Table 2.10 *Simbol Case Diagram*

NO	NAMA	GAMBAR	KETERANGAN
1.	<i>Actor</i>		Menspesifikasikan himpunan peran yang pengguna mainkan ketika berinteraksi dengan <i>use case</i> .

2.	<i>Depedency</i>		Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (<i>independent</i>) akan mempengaruhi elemen yang bergantung padanya elemen yang tidak mandiri.
NO	NAMA	GAMBAR	KETERANGAN
3.	<i>Generalization</i>		Hubungan dimana objek anak (<i>descendent</i>) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk (<i>ancestor</i>).
4.	<i>Include</i>		Menspesifikasikan bahwa <i>use cases</i> sumber secara eksplisit.
5.	<i>Extend</i>		Menspesifikasikan bahwa <i>use case</i> target memperluas perilaku dari <i>use case</i> sumber pada suatu titik yang diberikan.
6.	<i>Assocition</i>		Apa yang menghubungkan antara objek satu dengan objek lainnya.
7.	<i>System</i>		Menspesifikasikan paket yang menaplikan sistem secara terbatas.
8.	<i>Use Case</i>		Deskripsi dari urutan aksi – aksi yang ditampilkan sistem menghasilkan suatu hasil yang terukur bagi suatu <i>actor</i> .



9.	<i>Collaboration</i>		Interaksi aturan – aturan dan elemen lain yang bekerja sama untuk menyediakan perilaku yang lebih besar dari jumlah dan elemen – elemennya (sinergi).
NO	NAMA	GAMBAR	KETERANGAN
10.	<i>Note</i>		Elemen fisik yang eksis saat aplikasi dijalankan dan mencerminkan suatu sumber daya komputasi.




(Sumber : <https://www.coursehero.com>)

2. Activity Diagram

Mahdiana(2011) menyatakan, “*activity Diagram* adalah salah satu untuk memodelkan event-event yang terjadi dalam suatu *use case*.”

Table 2.11 Simbol *Actifity Diagram*

NO	NAMA	GAMBAR	KETERANGAN
1.	<i>Actifity</i>		Memperlihatkan bagaimana masing – masing kelas antarmuka saling berinteraksi satu sama lain.
2.	<i>Action</i>		State dari sistem yang mencerminkan eksekusi dari suatu aksi.


3.	<i>initian Node</i>		Bagaimana objek dibentuk atau diawali.
4.	<i>Actifity Find Node</i>		Bagaimana objek dibentuk dan dihancurkan.
NO	NAMA	GAMBAR	KETERANGAN
5.	<i>Fork Node</i>		Satu aliran yang pada tahap tertentu berubah menjadi beberapa aliran.

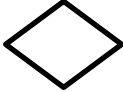
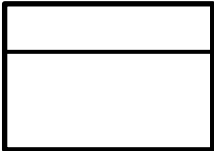
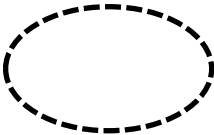



(Sumber : elib.unikom.ac.id)

3. Class Diagram

Class diagram adalah sebuah spesifikasi yang jika diinstansikan akan menghasilkan sebuah objek dan merupakan inti dari pengembangan dan desain berorientasi objek. Class menggambarkan keadaan (atribut / properti) suatu sistem, sekaligus menawarkan layanan untuk memanipulasi keadaan tersebut (metode / fungsi). *Class diagram* menggambarkan struktur dan deskripsi *class*, *package* dan objek beserta hubungan satu sama lain seperti *containment*, pewarisan, asosiasi, dan lain-lain (Mahdiana, 2011).

Tabel 2.12 Simbol Class Diagram

NO	NAMA	GAMBAR	KETERANGAN
1.	Generalization		Hubungan dimana objek anak (<i>descendent</i>) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk (<i>ancestor</i>).


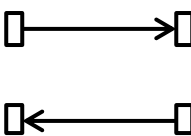
2.	Nary Association		Upayah untuk menghindari asosiasi dengan lebih dari 2 objek.
3.	Class		Himpunan dari objek-objek yang terbagi atribut serta operasi yang sama.
NO	NAMA	GAMBAR	KETERANGAN
4.	Collaboration		Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan suatu hasil yang terukur bagi suatu actor.
5.	Realization		Operasi yang benar-benar dilakukan oleh suatu objek.
6.	Dependency		Hubungandimana perubahan yang terjadi pada suatu elemen mandiri (<i>independent</i>) akan mempengaruhi elemen yang bergantung padanya elemen yang tidak mandiri.
7.	Association		Apa yang menghubungkan antara objek satu dengan objek lainnya.

(Sumber : elib. Unikom.ac.id)

4. *Sequence Diagram*

Sequence diagram merupakan sebuah diagram yang menunjukkan eksekusi *operation* disebuah objek yang melibatkan pemanggilan *operation* diobjek lain.

Table 2.13 Simbol *Sequence Diagram*

NO	NAMA	GAMBAR	KETERANGAN
1.	Life Line		Objek entry, antarmuka yang saling berinteraksi
2.	Message		Spesifikasi dari komunikasi antar objek yang memuat informasi – informasi tentang aktifitas yang terjadi.

(Sumber : elib.unikom.ac.id)

2.2.12 Pemeliharaan Sistem

Pemeliharaan sistem adalah proses pengubahan sistem setelah beroperasi dan digunakan. (Hryanto 2004 : 603).

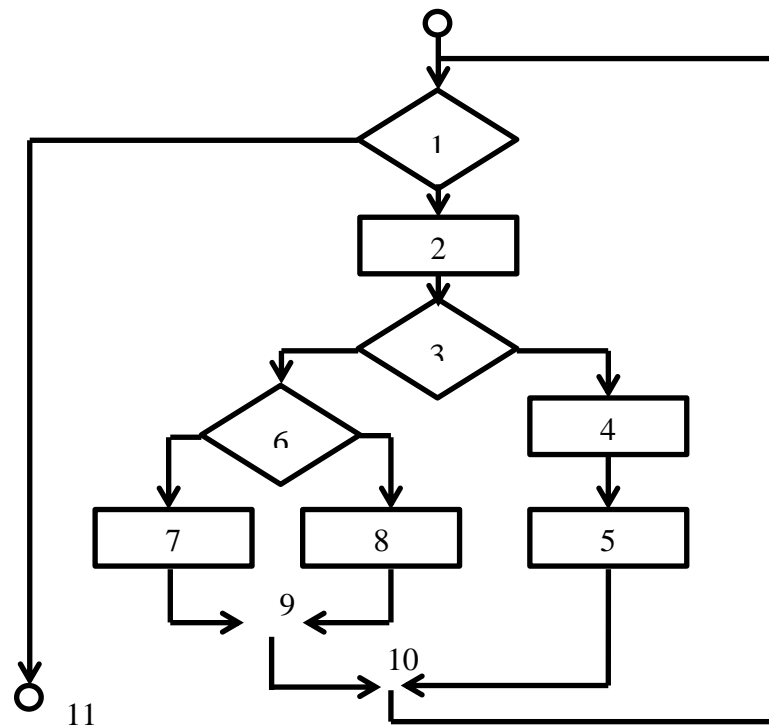
Pemeliharaan sistem adalah tahap dimana kita mulai pengoperasian sistem dan, jika diperlukan melakukan perbaikan-perbaikan kecil. (A. Nugroho, 2010 : 7).

Dari kedua pendapat diatas memiliki pandangan sama dimana setelah sistem dioperasikan dapat saja dilakukan perubahan, apakah itu sebagian kecil dari sistem taukah secara keseluruhan dengan harapan untuk dapat memenuhi kebutuhan dari pengguna sistem tersebut.

2.2.13 Teknik Pengujian

2.2.13.1 White box Testing

White Box Testing adalah pengujian perangkat lunak dimana struktur internal diketahui untuk menguji siapa yang akan menguji perangkat lunak, pengujian ini membutuhkan pengetahuan internal tentang kemampuan sistem dan pemrogram. Untuk melakukan pengujian *Test Case* terlebih dahulu akan dilakukan penerjemahan *Flowchart* kedalam notasi.



Gambar 2.2 Bagan Alir

2.2.13.2 Black Box Testing

Black Box Testing adalah pengujian fungsional yang merupakan pengujian yang menggunakan perangkat lunak tanpa mengetahui struktur internal kode atau program, dalam pengujian ini tester menyadari apa yang harus dilakukan oleh program tetapi tidak memiliki pengetahuan tentang bagaimana melakukannya.

2. 3. *Framework*

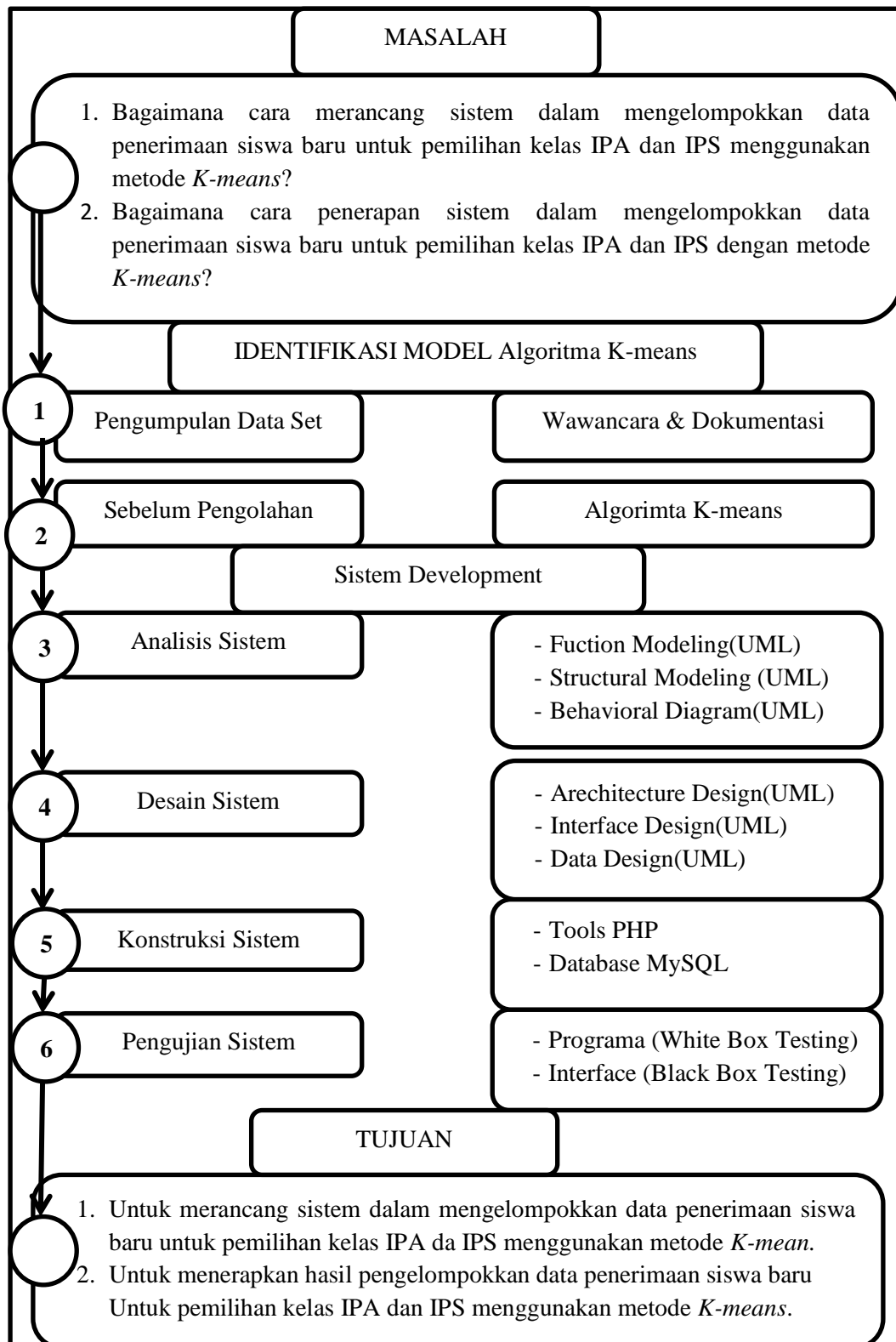
Framework adalah salah satu library yang telah diorganisasikan pada sebuah rancangan arsitektur untuk memberikan kecepatan, ketepatan, kemudahan, dan konsisten dalam pengembangan suatu aplikasi (Siena 2009). Framework mempunyai arti struktur dasar yang melandasi sebuah system. Sehingga system bergerak berdasarkan kerangka struktur dasar yang telah dibentuk sebelumnya.

2. 4. Pengujian Perangkat Lunak

Table 2.14 Fungsi PHP Yang Berhubungan Dengan Database MySQL

NO	TOOLS	KETERANGAN
1	PHP	PHP merupakan bahasa pemrograman yang digunakan secara luas untuk penanganan pembuatan dan pengembangan sebuah situs web dan bisa digunakan bersama dengan HTML
2	MySQL	MySQL merupakan software system manajemen basis data SQL (<i>striktire query language</i>) atau DBMS yang <i>multi thread</i> dan <i>multi user</i> . PHP dan MySQL seolah pasangan sejati yang tak terpisahkan. Keduanya paling sering disandingkan dalam pembuatan aplikasi berbasis web (<i>web application development</i>)

2. 5. Kerangka Fikir



Gambar 2.3 Kerangka Pikir

BAB III

METODE PENELITIAN

3.1 Jenis, Metode, Subjek, Waktu Dan Lokasi Penelitian

Dalam tingkat penerapannya, penelitian ini merupakan penelitian terapan dimana penelitian ini bertujuan untuk memberikan solusi atas permasalahan yang terjadi. Dari jenis informasi yang didapatkan dan diolah, maka penelitian ini merupakan penelitian kuantitatif dimana penelitian ini menggunakan data yang berupa angka sebagai alat untuk menganalisis keterangan mengenai apa yang ingin diketahui.

Dari perlakuan terhadap data, maka penelitian ini merupakan penelitian konfirmatori karena penelitian ini merupakan salah satu bentuk analisis faktor juga khususnya dalam penelitian sosial.

Metode penelitian ini menggunakan metode studi kasus. Dengan demikian jenis penelitian ini adalah penelitian deskriptif. Subjek dari penelitian ini adalah pengelompokan data penerimaan siswa baru. Penelitian ini di mulai pada bulan Oktober 2020 sampai dengan Desember 2020 yang berlokasi di Madrasah Aliyah Angrek Kabupaten Gorontalo Utara.

3.2 Pengumpulan Data

Data primer adalah data yang diperoleh secara langsung dari sumbernya. Data primer yang digunakan dalam penelitian ini bersumber dari Madrasah Aliyah Angrek, yaitu tentang pengelompokan data penerimaan siswa baru tahun ajaran 2016/2017 sampai 2019/2020.

Adapun variabel atau atribut dengan tipe datanya masing-masing pada tabel berikut :

Table 3.1 Atribut Data

NO	NAME	TYPE	VALUE	KETERANGAN
1.	Nama	String	1-50	Variabel Input
2.	Asal Sekolah	String	1-50	Variabel Input
3.	Kelas	String	1-20	Variabel Input
4.	Nilai Rata-Rata Rapor MP IPA	Integer	1-10	Variabel Input
5.	Nilai Rata-Rata Rapor MP IPS	Ineger	1-10	Variabel Input
6.	Nilai UN	Integer	1-10	Variabel Input
7.	Kelompok Kelas IPA dan IPS	String	1-50	Variabel Output

3.3 Pemodelan/Abstraksi

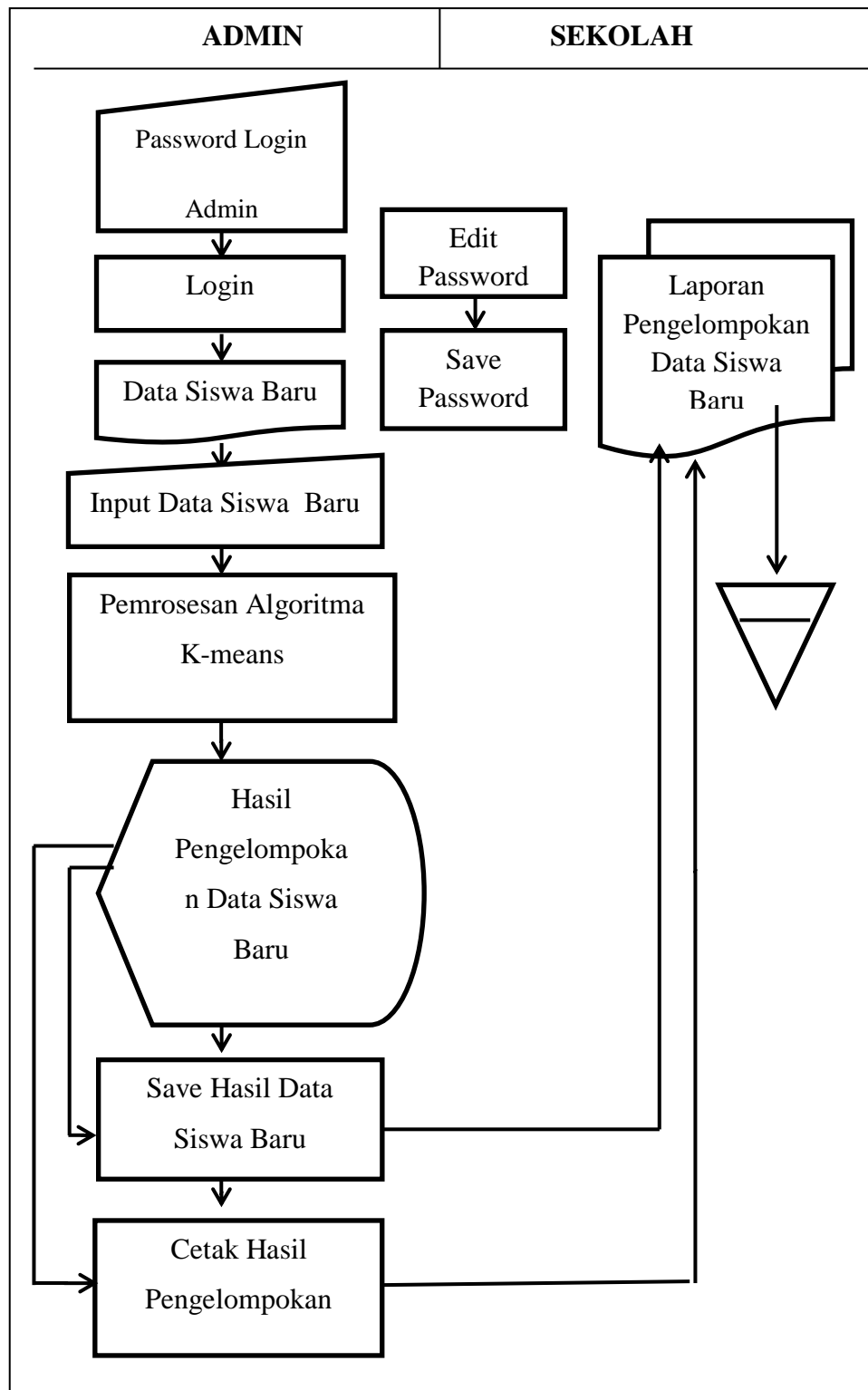
3.3.1 Pengembangan Model

Dalam tahap ini prosedur atau langkah-langkah pokok dalam implementasi data mining menggunakan algoritma K-means untuk pengelompokan data penerimaan siswa baru di Madrasah Aliyah Anggrek desa monas kecamatan monano, dengan tahap ini digunakan beberapa alat bantu antara lain yaitu PHP dan MySQL serta dalam tahap pengujian menggunakan White Box Testing dan Black Box Testing untuk menguji kinerja sistemnya.

3.3.2 Evaluasi Model

Model yang telah dihasilkan kemudian dievaluasi dengan menggunakan Mean Absolute Percentage Error (MAPE).

3.3.3 Pengembangan Sistem



Gambar 3.1 Sistem Yang Diusulkan

3.3.4 Analisa Sistem

Analisis sistem menggunakan pendekatan berorientasi objek yang digambarkan dalam bentuk :

- a) *Function Modelling*, menggunakan alat bantu UML, dalam bentuk :
 - *Use Case Diagram*
 - *Activity Diagram*
- b) *Structural Modelling*, menggunakan alat bantu UML, dalam bentuk :
 - *Class Diagram*
- c) *Behavioral Modelling*, menggunakan alat bantu UML, dalam bentuk :
 - *Sequence Diagram*

Pada tahap ini analisis sistem yang diusulkan dalam pengelompokan data siswa baru yakni terdiri dari :

1. Entry Data :
 - Nama
 - Asal Sekolah
 - Matematika
 - Bahasa Indonesia
 - IPA
 - IPS
2. Proses *Clustering*
3. Laporan :
 - Kelompok Kelas IPA dan Kelas IPS

3.3.5 Desain Sistem

Desain sistem ini menggunakan pendekatan berorientasi objek yang digambarkan dalam bentuk :

- a) *Architecture Design*, menggunakan alat bantu UML, dalam bentuk :
 - Model jaringan dari sistem adalah stand alone
 - Spesifikasi hardware dan software yang direkomendasikan adalah:
 1. Sistem Operasi : Windows 10
 2. Prosesor Dengan Kecepatan Minimal 1,6GHz
 3. Harddisk Free space 500 GB
 4. RAM : 4 GB

b) *Interface design*, menggunakan alat bantu UML, dalam bentuk :

- Mekanisme user
- Mekanisme navigasi
- Mekanisme input (Form)
- Mekanisme output (report)

c) *Data design*, menggunakan alat bantu UML, dalam bentuk :

- Format data yang digunakan [File SQL]
- Struktur data
- *Database diagram*

d) *Progres design*, menggunakan alat bantu UML, dalam bentuk :

- *Class*
- *Attribute*
- *Method*
- *Event*

3.3.6 Konstruksi Sistem

Pada tahap ini dilakukan tahap pembuatan sistem menggunakan *tools PHP* dan Database *MySQL* serta *White Box Testing* dan *Black Box Testing* untuk menguji kinerja sistem dan pengukuran akurasi menggunakan *Confusion Matrix*.

3.3.7 Pengujian Sistem

Pengujian sistem merupakan pengujian program perangkat lunak yang lengkap dan terintegrasi, perangkat lunak atau yang sering dikenal dengan sebutan software hanyalah satuan elemen dari sistem berbasis komputer yang lebih besar, biasanya perangkat lunak dihubungkan dengan perangkat lunak dan perangkat keras lainnya. Pengujian yang dilakukan dengan menggunakan teknik pengujian perangkat lunak yaitu :

a) *White Box Testing*

White Box Testing merupakan software yang sudah atau yang telah di rekayasa kemudian diuji dengan metode white box testing pada kode program proses penerapan metodeya/modelnya. Kode program tersebut dibuatkan flowchart programnya, kemudian dipetakan kedalam bentuk *flowgraph*, (bagan alir kontrol) yang tersusun dari beberapa node dan *edge*. Berdasarkan *flowgraph*, di tentukan jumlah region dan *Cyclomatic Complexity* (CC). apabila *Independent Path* = $V(G) = (CC) = \text{Region}$, di mana setiap Path hanya dieksekusi sekali dan sudah benar, maka sistem di nyatakan efisien dari segi kelayakan logika pemrogram.

b) *Black Box Testing*

Black Box Testing pengujian Black Box melalui program PHP dan *database MySQL*. Selanjutnya software diuji pula dengan metode black box testing yang focus pada keperluan fungsional dari software, serta berusaha untuk menemukan kesalahan dalam beberapa kategori diantaranya adalah 1) fungsi yang salah atau hilang, 2) kesalahan interface, 3) kesalahan dalam struktur data atau akses basis data eksternal, 4) kesalahan performa, 5) kesalahan inisialisasi dan terminasi.

BAB IV

HASIL PENELITIAN

4.1. Hasil Pengumpulan Data

Penelitian ini dilakukan dimadrasah Aliyah Anggrek dimana sekolah ini merupakan sekolah yang dibangun pada tahun 2006. masalah yang terjadi dimadrasah Aliyah Anggrek dimana sekolah tersebut masih kesulitan dalam mengelompokkan data siswa baru untuk pemilihan kelas IPA dan IPS dikarenakan nilai siswa yang rendah akibat salah dalam memilih kelas serta belum adanya sistem untuk mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS. Sehingga penulis tertarik memberikan suatu solusi dengan membangun suatu sistem untuk mengelompokkan data penerimaan siswa baru untuk pemilihan kelas IPA dan IPS berdasarkan nilai rapor semester 1 sampai semester 6 dan nilai UN.

Berdasarkan hasil pengumpulan data, maka diperoleh data primer sebagai berikut:

Tabel 4.1 Hasil Pengumpulan Dataset

NO	NAMA SISWA	Asal Sekolah	MATEMATIKA	BHS INDONESIA	IPA	IPS
1	Alfi Raja	MTs Muh.Monano	82,22	80	86,6	85,16
2	Fahri Sulubani	MTs Muh.Monano	86,52	81,51	84,5	86,5
3	Farut Lakoro	SMP N 3 Monano	81,4	83,8	86,16	89,83
4	Iston Lakoro	SMP N 3 Monano	83	87,33	85	88,16
5	Jufrin S. Mada	SMP N 3 Monano	87,16	88,33	86,66	88,66
6	Mohamad Walidi S. Payu	MTs Muh.Monano	84,33	87,33	88,33	87,66
7	Adelia Hasu	MTs Muh.Monano	86,5	89,33	88	90
8	Desmita Gobel	MTs Muh.Monano	87,5	88	88,5	87,5
9	Ismi S. Otoluwa	MTs Muh.Monano	84,66	89,66	86,33	88,33

10	Kadita Putri Latif	MTs Muh.Monano	90,66	87,66	86,16	86,83
11	Kristin Raja	MTs Muh.Monano	85,33	89	84	86,16
12	Nurlela Otoluwa	SMP N 3 Monano	86,16	88,16	88,33	88,33
13	Puspita Kai	MTs Muh.Monano	86,66	89,5	87,66	90,5
14	Rini Anggriani Paramata	MTs Muh.Monano	84,16	88	89	86,33
15	Siska T. Dou	SMP N 2 Monano	80	87,6	88,16	87,83
16	Siti Azizah Pakaya	MTs Muh.Monano	86,5	88	89,5	87
17	Yulinda H. Kuka	SMP N 3 Monano	88,33	87,5	88	90,33
18	Abdul Wahid Arafa	SMP N 1 Anggrek	85,83	88,33	87,33	90,16
19	Aldi Lakoro	SMP N 3 Monano	87,83	86,83	87,5	87,16
20	Fandi J. Masita	SMP N 4 Monano	89,33	85,33	88	86,66

4.1.1 Penerapan Metode K-means

4. 2. Hasil Pemodelan

4.2.1 Penjelasan Algoritma

Algoritma kmeans ini merupakan salah satu algoritma dalam *data mining* yang bisa digunakan untuk melakukan pengelompokan/*clustering* suatu data. *K-means Clustering* adalah K dimaksudkan sebagai konstanta jumlah cluster yang diinginkan, *Means* dalam hal ini berarti nilai suatu rata-rata dari suatu group data yang dalam hal ini didefinisikan sebagai cluster, sehingga *K-mean clustering* adalah suatu metode penganalisaan atau metode data mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan system partisi. Metode *K-mean* berusaha mengelompokkan data yang ada kealam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada didalam kelompok yang lain. *Algoritma K-mean* merupakan algoritma yang membutuhkan

parameter input sebanyak k dan membagi sekumpulan n objek kedalam k cluster sehingga tingkat kemiripan antar anggota dalam suatu cluster tinggi sedangkan tingkat kemiripan dengan anggota cluster lain sangat rendah. Kemiripan anggota terhadap cluster diukur dengan kedekatan objek terhadap nilai mean pada *cluster* atau dapat disebut sebagai *centroid cluster* atau pusat massa (Khotimah et al.,2014).

Berikut adalah rumus untuk menentukan jarak data dari masing-masing centroid :

::

$$d(P, Q) = \sqrt{\sum_{j=1}^p (x_j(P) - x_j(Q))^2}$$

Keterangan :

D = titik dokumen

P = data record

Q = data centroid

Jarak yang terpendek antara centroid dengan dokumen menentukan posisi cluster suatu dokumen.

$$C(i) = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x}$$

Keterangan :

x_1 = Nilai data record ke-1

x_2 = Nilai data record ke-2

$\sum x$ = jumlah data record

Eksperimen ini menggunakan algoritma yang paling umum digunakan dalam clustering yaitu algoritma *K-Means*. Algoritma ini populer karena mudah diimplementasikan dan kompleksitas waktunya linear. Kelemahannya adalah algoritma ini sensitive terhadap inisialisasi cluster (Langgeni, Baizal, & W, 2010). Dasar algoritmanya adalah sebagai berikut :

5. Inisialisasi cluster
6. Masukkan setiap dokumen ke *cluster* yang paling cocok berdasarkan ukuran kedekatan dengan centroid. *Centroid* adalah vector term yang dianggap sebagai titik tengah cluster.
7. Setelah semua dokumen masuk ke *cluster*. Hitung ulang centroid cluster berdasarkan dokumen yang berada didalam cluster tersebut.
8. Jika centroid tidak berubah (dengan threshold tertentu) maka stop. Jika tidak kembali ke langkah 2.

4.2.2 Perhitungan Algoritma

Berikut adalah perhitungan algoritma K-means dengan judul “Implementasi Data Mining Untuk Pengelompokan Data Penerimaan Siswa Baru”. Tabel 4.2 merupakan data yang akan dihitung dengan jumlah data.

Diketahui :

Jumlah Cluster 4

Jumlah Data 10

Jumlah Atribut 4

Tabel 4.2 Sempel Data Set Siswa Baru

NO	NAMA SISWA	Asal Sekolah	MATEMATIKA	BHS INDONESIA	IPA	IPS
1	Alfi Raja	MTs Muh.Monano	82,22	80	86,6	85,16
2	Fahri Sulubani	MTs Muh.Monano	86,52	81,51	84,5	86,5
3	Farut Lakoro	SMP N 3 Monano	81,4	83,8	86,16	89,83
4	Iston Lakoro	SMP N 3 Monano	83	87,33	85	88,16
5	Jufrin S. Mada	SMP N 3 Monano	87,16	88,33	86,66	88,66
6	Mohamad Walid S. Payu	MTs Muh.Monano	84,33	87,33	88,33	87,66
7	Adelia Hasu	MTs Muh.Monano	86,5	89,33	88	90

8	Desmita Gobel	MTs Muh.Monano	87,5	88	88,5	87,5
9	Ismi S. Otoluwa	MTs Muh.Monano	84,66	89,66	86,33	88,33
10	Kadita Putri Latif	MTs Muh.Monano	90,66	87,66	86,16	86,83

Iterasi ke-1

1. Menentukan nilai centroid

Penentuan nilai awal centroid dapat diambil secara acak. Berikut data centroid yang dipilih secara acak :

Tabel 4.3 Penentuan Awal Cluster

1	Alfi Raja	MTs Muh.Monano	82,22	80	86,6	85,16
5	Jufrin S. Mada	SMP N 3 Monano	87,16	88,33	86,66	88,66
7	Adelia Hasu	MTs Muh.Monano	86,5	89,33	88	90
10	Kadita Putri Latif	MTs Muh.Monano	90,66	87,66	86,16	86,83

2. Perhitungan jarak pada cluster

Berikut rumus yang digunakan dengan menggunakan persamaan *Euclidean Distance Space* :

$$C1 = \sqrt{(x1-y1)^2 + (x2-y2)^2 + (x3-y3)^2 + (x4-43)^2}$$

$$C2 = \sqrt{(x1-y1)^2 + (x2-y2)^2 + (x3-y3)^2 + (x4-43)^2}$$

$$C3 = \sqrt{(x1-y1)^2 + (x2-y2)^2 + (x3-y3)^2 + (x4-43)^2}$$

Keterangan :

x = data record

y = data centroid

berikut cara kerja perhitungan manual sebagai berikut :

$$\begin{aligned} C1,(1) \quad & \sqrt{(82,22-82,22)^2 + (80-80)^2 + (86,6-86,6)^2 + (85,16-85,16)^2} \\ & \sqrt{0} = 0 \end{aligned}$$

$$\begin{aligned} C1,(2) \quad & \sqrt{(86,52-82,22)^2 + (81,51-80)^2 + (84,5-86,6)^2 + (86,5-85,16)^2} \\ & \sqrt{(4,3)^2 + (1,51)^2 + (-2,1)^2 + (1,34)^2} \\ & \sqrt{18,49 + 2,2801 + 4,41 + 1,7956} \\ & \sqrt{26,9757} \\ & = 5,1938136278 \end{aligned}$$

$$\begin{aligned} C1,(3) \quad & \sqrt{(81,4-82,22)^2 + (83,8-80)^2 + (86,16-86,6)^2 + (89,83-85,16)^2} \\ & \sqrt{(0,28)^2 + (3,8)^2 + (0,44)^2 + (4,67)^2} \\ & \sqrt{0,0784 + 14,44 + 0,1936 + 21,8089} \\ & \sqrt{36,5209} \\ & = 6,0432524356 \end{aligned}$$

$$\begin{aligned} C2,(4) \quad & \sqrt{(83-87,16)^2 + (87,33-88,33)^2 + (85-86,66)^2 + (88,16-88,66)^2} \\ & \sqrt{(-4,16)^2 + (-1)^2 + (-1,66)^2 + (-0,5)^2} \\ & \sqrt{17,3056 + 1 + 1,3456 + 0,25} \\ & \sqrt{19,9012} \\ & = 4,4610761034 \end{aligned}$$

$$\begin{aligned} C2,(5) \quad & \sqrt{(87,16-87,16)^2 + (88,33-88,33)^2 + (86,66-86,66)^2 + (88,66-88,66)^2} \\ & \sqrt{0} = 0 \end{aligned}$$

$$\begin{aligned}
C2,(6) \quad & \sqrt{(84,33-87,16)^2 + (87,33-88,33)^2 + (88,33-86,66)^2 + (87,66-88,66)^2} \\
& \sqrt{(-2,83)^2 + (-1)^2 + (1,67)^2 + (-1)^2} \\
& \sqrt{8,0089 + 1 + 2,7889 + 1} \\
& \sqrt{12,7978} \\
& = 3,5774012914
\end{aligned}$$

$$\begin{aligned}
C3,(7) \quad & \sqrt{(86,5-86,5)^2 + (89,33-89,33)^2 + (88-88)^2 + (90-90)^2} \\
& \sqrt{0} = 0
\end{aligned}$$

$$\begin{aligned}
C3,(8) \quad & \sqrt{(87,5-86,5)^2 + (88-89,33)^2 + (88,5-88)^2 + (87,5-90)^2} \\
& \sqrt{(1)^2 + (-1,33)^2 + (0,5)^2 + (-2,5)^2} \\
& \sqrt{1 + 1,7689 + 0,25 + 6,25} \\
& \sqrt{9,2689} \\
& = 3,0444868205
\end{aligned}$$

$$\begin{aligned}
C4,(9) \quad & \sqrt{(84,66-90,66)^2 + (89,66-87,66)^2 + (86,33-86,16)^2 + (88,33-86,83)^2} \\
& \sqrt{(-6)^2 + (2)^2 + (0,17)^2 + (1,5)^2} \\
& \sqrt{36 + 4 + 0,0289 + 2,25} \\
& \sqrt{42,2789} \\
& = 6,5022226969
\end{aligned}$$

$$\begin{aligned}
C4,(10) \quad & \sqrt{(90,66-90,66)^2 + (87,66-87,66)^2 + (86,16-86,16)^2 + (86,83-86,83)^2} \\
& \sqrt{0} = 0
\end{aligned}$$

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.

Tabel 4.4 Mencari Jarak Terdekat

NO	NAMA SISWA	CENTRO ID 1	CENTRO ID 2	CENTRO ID 3	CENTRO ID 4	JARAK TERDEKAT	KELOMPOK DATA
1	Alfi Raja	0	10,29786 871	11,43472 343	11,5278 6624	0	C1
2	Fahri Sulubani	5,193813 628	7,500213 33	9,254879 794	7,60438 0317	5,19381362 8	C1
3	Farut Lakoro	6,092199 931	7,437566 806	7,746315 253	10,4712 5589	6,09219993 1	C1
4	Iston Lakoro	8,117715 196	4,616405 528	5,351224 159	7,86759 1754	4,61640552 8	C2
5	Jufrin S. Mada	10,29786 871	0	2,242052 631	4,03705 3381	0	C2
6	Mohamad Waldi S. Payu	8,211205 758	3,577401 291	3,780661 318	6,75097 0301	3,57740129 1	C2
7	Adelia Hasu	11,43472 343	2,242052 631	0	5,79042 3128	0	C3
8	Desmita Gobel	10,04808 439	2,226140 157	3,044486 82	4,00321 1211	2,22614015 7	C2
9	Ismi S. Otoluwa	10,45901 525	2,869965 157	3,012025 896	6,50222 2697	2,86996515 7	C2
10	Kadita Putri Latif	11,52786 624	4,037053 381	5,790423 128	0	0	C4

3. Pengelompokan data

Berikut ini akan ditampilkan data matriks pengelompokan *cluster*, nilai satu berarti data tersebut berada dalam *group* (kelompok data).

Tabel 4.5 Kelompok Pembagian Data 1

No	C1	C2	C3	C4
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	0	1	0	0
5	0	1	0	0
6	0	1	0	0
7	0	0	1	0
8	0	1	0	0
9	0	1	0	0
10	0	0	0	1

4. Menentukan pusat cluster baru

$$\text{Matematika C1} = \frac{82,22 + 86,52 + 81,4}{3} = \frac{250,14}{3} = 83,38$$

$$\text{Bahasa Indonesia C1} = \frac{80 + 81,51 + 83,8}{3} = \frac{245,31}{3} = 81,77$$

$$\text{Ipa C1} = \frac{86,6 + 84,5 + 86,16}{3} = \frac{257,26}{3} = 85,7533$$

$$\text{Ips C1} = \frac{85,16 + 86,5 + 89,83}{3} = \frac{261,49}{3} = 87,1633$$

$$\text{Matematika C2} = \frac{83 + 87,16 + 84,33 + 87,5 + 84,66}{5} = \frac{426,65}{5} = 85,33$$

$$\text{Bahasa Indonesia C2} = \frac{87,33 + 88,33 + 87,33 + 88 + 89,66}{5} = \frac{440,65}{5} = 88,13$$

$$\text{Ipa C2} = \frac{85 + 86,66 + 88,33 + 88,5 + 86,33}{5} = \frac{434,82}{5} = 86,964$$

$$\text{Ips C2} = \frac{88,16 + 88,66 + 87,66 + 87,5 + 88,33}{5} = \frac{440,31}{5} = 88,062$$

$$\text{Matematika C3} = \frac{86,5}{1} = 86,5$$

$$\text{Bahasa Indonesia C3} = \frac{89,33}{1} = 89,33$$

$$\text{Ipa C3} = \frac{88}{1} = 88$$

$$\text{Ips C3} = \frac{90}{1} = 90$$

$$\begin{aligned}
 \text{Matematika C4} &= \frac{90,66}{1} = 90,66 \\
 \text{Bahasa Indonesia C4} &= \frac{87,66}{1} = 87,66 \\
 \text{Ipa C4} &= \frac{86,16}{1} = 86,16 \\
 \text{Ips C4} &= \frac{86,83}{1} = 86,83
 \end{aligned}$$

Iterasi ke-2

Ulangi langkah ke 2 hingga posisi data tidak mengalami perubahan

Tabel 4.6 Penentuan Cluster Baru

No	Centroid	MATEMATIKA	BHS INDONESIA	IPA	IPS
1	C1	83,38	81,77	85,7533	87,1633
2	C2	85,33	88,13	86,964	88,062
3	C3	86,5	89,33	88	90
4	C4	90,66	87,66	86,16	86,83

Tabel 4.7 Kelompok Pembagian Data 2

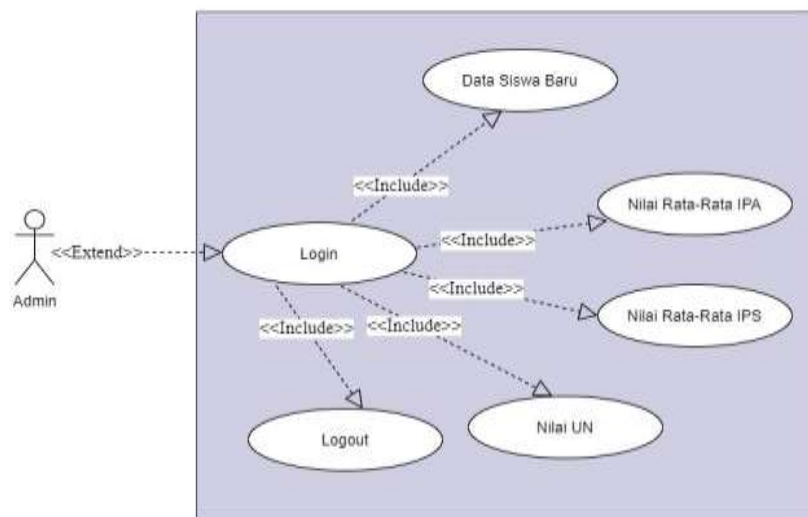
No	C1	C2	C3	C4
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	0	1	0	0
5	0	1	0	0
6	0	1	0	0
7	0	0	1	0
8	0	1	0	0
9	0	1	0	0
10	0	0	0	1

Pada perhitungan iterasi ini telah berhenti pada iterasi ke-2 karena kelompok data ke-1 dan hasil clustering telah mencapai stabil dan konvergen.

4. 3. Hasil Analisis

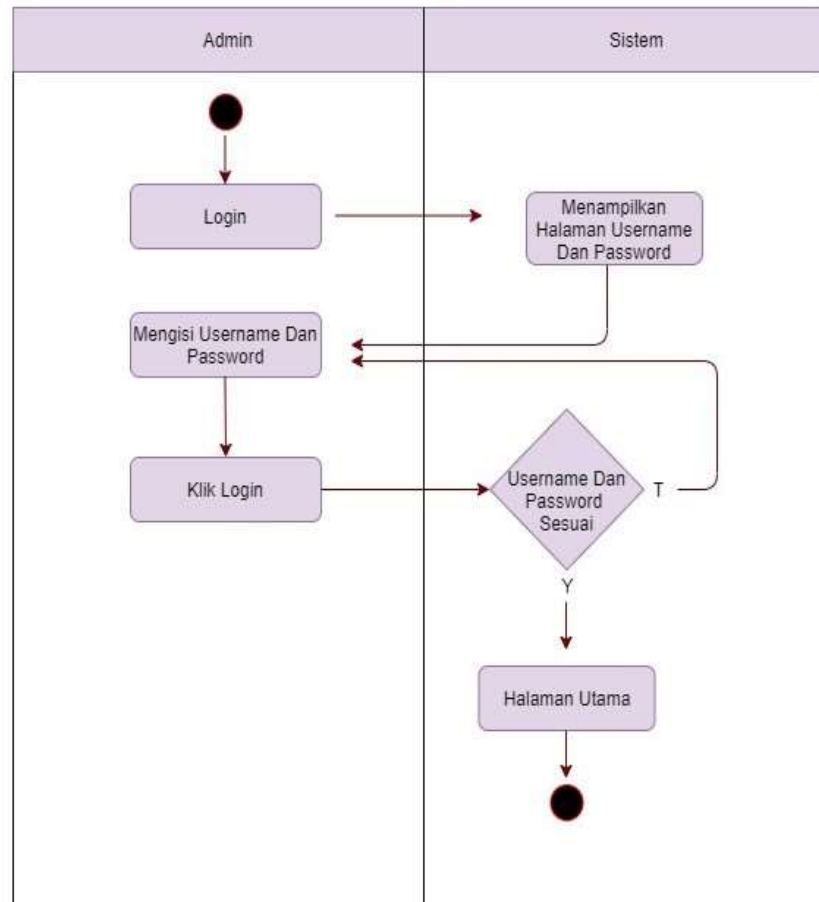
4.3.1 UML (Unified Modeling Language)

- a. Functional Modelling
 - *Use Case Diagram*



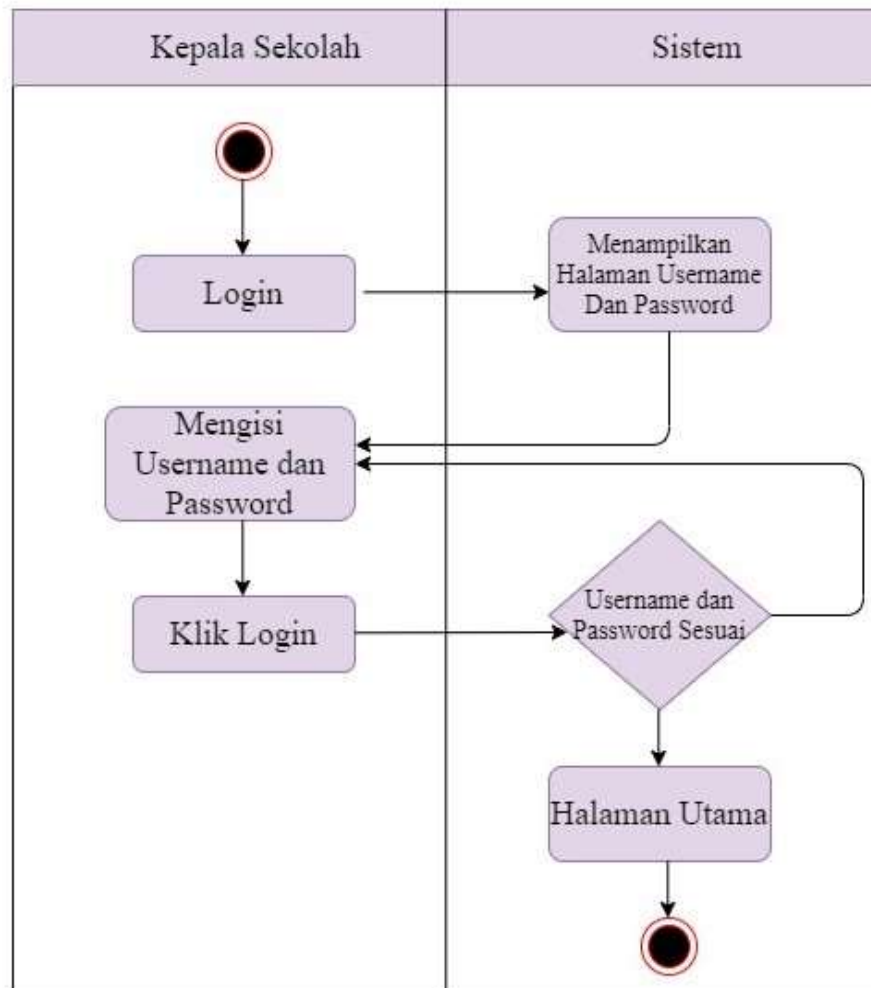
Gambar 4.1 *Use Case Diagram* Admin Clustering Siswa

- Activity Diagram Login Admin



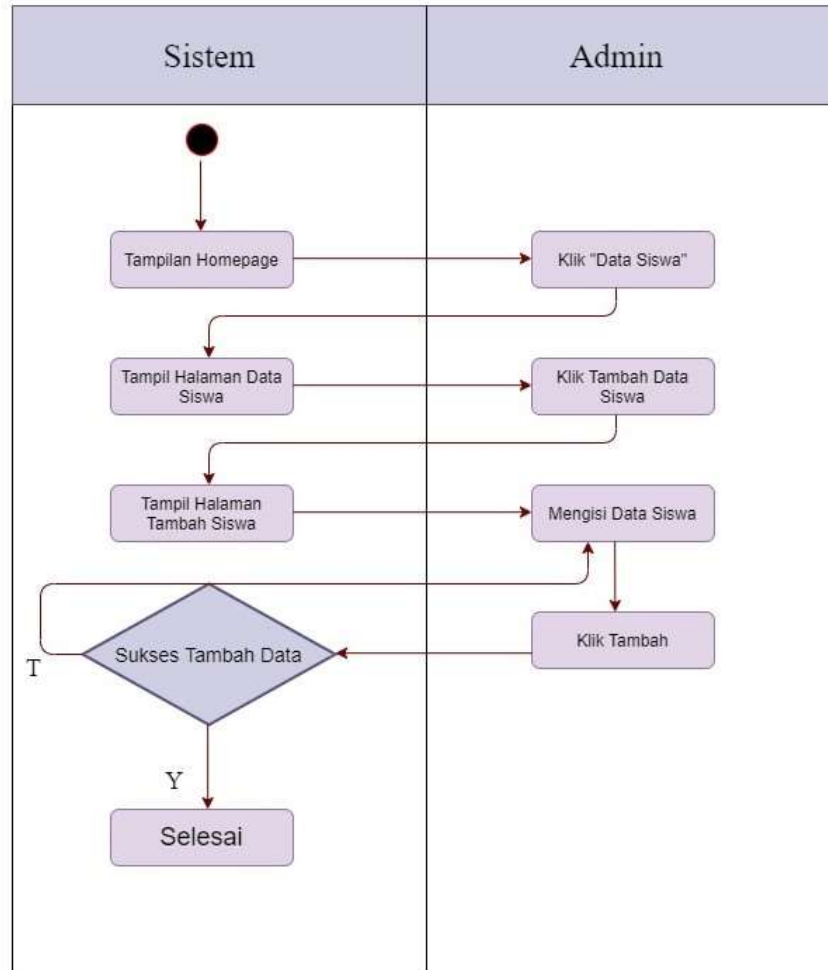
Gambar 4.2 Activity Diagram Login Admin

- Activity Diagram Login Kepala Sekolah



Gambar 4.3 Activity Diagram Login Kepala Sekolah

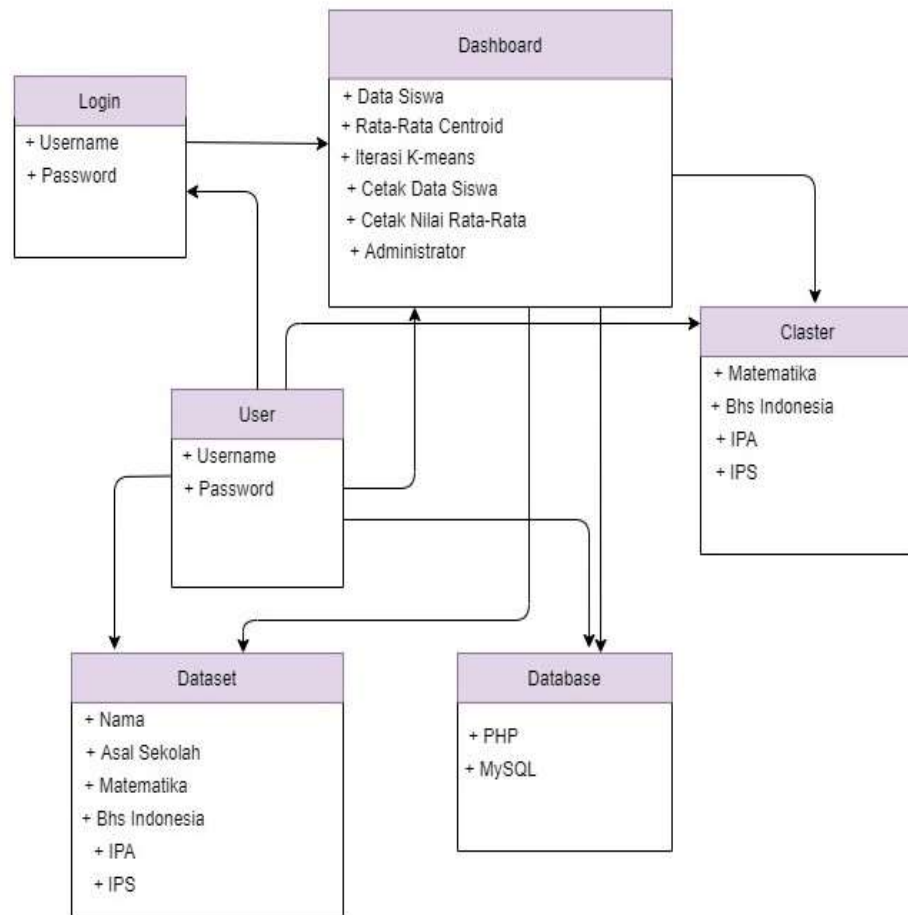
- Activity Diagram Data Siswa Baru



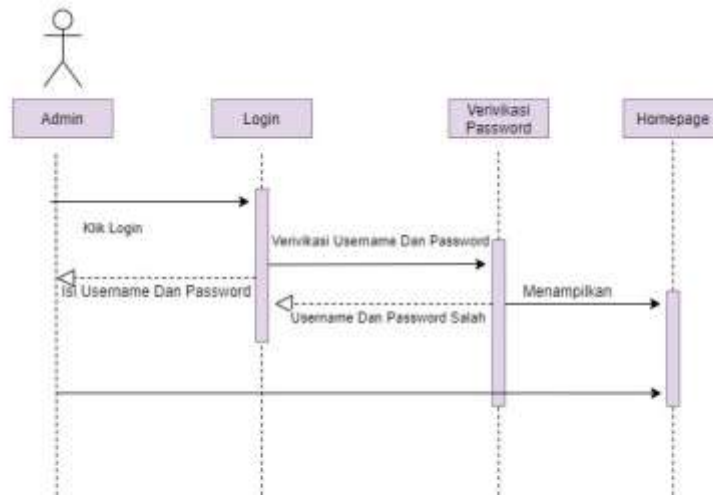
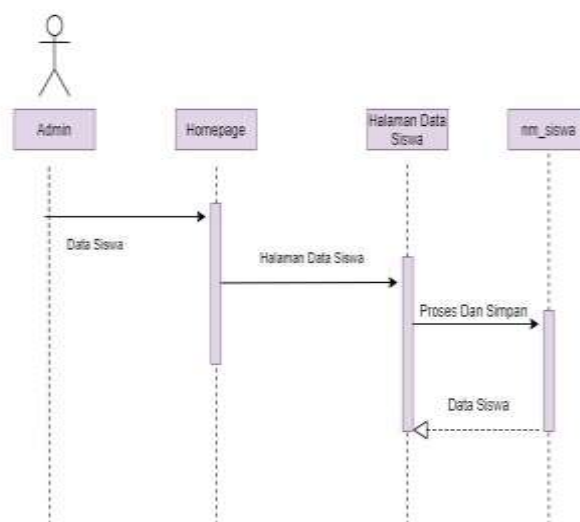
Gambar 4.4 Activity Diagram Data Siswa baru

b. Structural Modelling,

- Class Diagram

**Gambar 4.5** *Class Diagram*

c. Behavioral Modelling

- *Sequence Diagram Login***Gambar 4.6** *Sequence Diagram Login*- *Sequence Diagram Data siswa***Gambar 4.7** *Sequence Diagram Data Siswa*

4.4. Arsitektur Sistem

Spesifikasi hardware dan software yang direkomendasikan, yaitu

1. Processor : Intel(R) Celeron(R) N4000 CPU @ 1.10GHz
2. RAM : 4GB
3. VGA : Intel(R) UHD Graphics 600
4. Hardisk : 500GB
5. Operating System : Windows 10

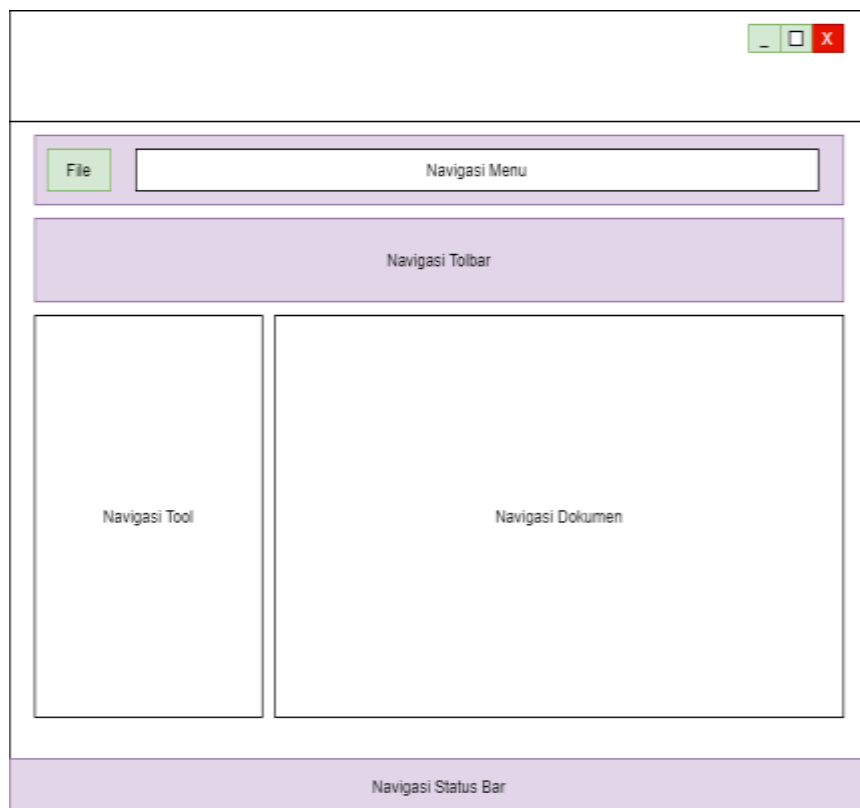
4.5. Interface Design

4.5.1 Mekanisme User

Tabel 4.8 Mekanisme User

User	Kategori	Akses Input	Akses Output
Admin	Administrator	All	All
Kepala Sekolah	Administrator		

4.5.2 Mekanisme Navigasi



Gambar 4.8 Mekanisme Navigasi Home Admin

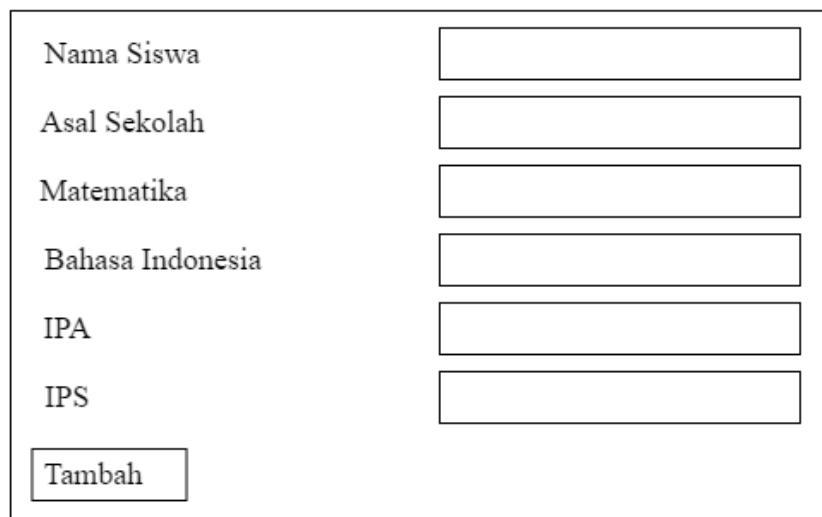
4.5.3 Mekanisme Login



A login form titled "Masuk" (Login) is shown. It contains two input fields: "Username" and "Password". Below these fields is a green button labeled "Masuk".

Gambar 4.9 Mekanisme Login

4.5.4 Mekanisme Input Data Siswa



A form for inputting student data is shown. It contains six input fields with labels: "Nama Siswa", "Asal Sekolah", "Matematika", "Bahasa Indonesia", "IPA", and "IPS". Below these fields is a button labeled "Tambah" (Add).

Gambar 4.10 Mekanisme Input Data Siswa

4.5.5 Mekanisme Output

DATA HASIL ITERASI			
MULAI AWAL			
ITERAASI KE-1			
C1	C2	C3	
0	1	0	
0	1	0	
ITERAASI KE-2			
C1	C2	C3	
1	0	0	
1	0	0	
ITERAASI KE-3			
C1	C2	C3	
1	0	0	
1	0	0	
ITERAASI KE-4			
C1	C2	C3	
1	0	0	
1	0	0	
ITERAASI KE-5			
C1	C2	C3	
1	0	0	
1	0	0	

Gambar 4.11 Mekanisme Output

4. 6. Data Design

4.6.1 Struktur Data

Tabel 4.9 Data Nilai Matematika

Nama : matematika					
Primery Key : -					
Forigen Key : -					
Media : Hardisk					
Fungsi : Menyimpan Data Matematika					
Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1.	Nama Siswa	Varchar	50	100	Nama Siswa

2.	Asal Sekolah	Text	50	50	Asal Ssekolah
3.	Matematika	Varchar	5	10	Matematika

Tabel 4.10 Data Nilai Bahasa Indonesia

Nama : bahasa_indonesia Primery Key : - Forigen Key : - Media : Hardisk Fungsi : Menyimpan Data Bahasa Indonesia Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1.	Nama Siswa	Varchar	50	100	Nama Siswa
2.	Asal Sekolah	Text	50	50	Asal Ssekolah
4.	Bahasa Indonesia	Varchar	10	20	Bahasa Indonesia

Tabel 4.11 Data Nilai IPA

Nama : ipa Primery Key : - Forigen Key : - Media : Hardisk Fungsi : Menyimpan Data IPA Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1.	Nama Siswa	Varchar	50	100	Nama Siswa
2.	Asal Sekolah	Text	50	50	Asal Ssekolah
4.	IPA	Varchar	10	20	IPA

Tabel 4.12 Data Nilai IPS

Nama : ips Primery Key : - Forigen Key : - Media : Hardisk Fungsi : Menyimpan Data IPS Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1.	Nama Siswa	Varchar	50	100	Nama Siswa
2.	Asal Sekolah	Text	50	50	Asal Ssekolah
3.	IPS	Varchar	10	20	IPS

Tabel 4.13 Data Siswa

Nama : nm_siswa Primery Key : - Forigen Key : - Media : Hardisk Fungsi : Menyimpan Data Siswa Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1.	Nama Siswa	Varchar	50	100	Nama Siswa
2.	Asal Sekolah	Text	50	50	Asal Ssekolah
3.	Matematika	Varchar	5	10	Matematika
4.	Bahasa Indonesia	Varchar	10	20	Bahasa Indonesia
.	IPA	Varchar	10	20	IPA
6.	IPS	Varchar	10	20	IPS

Tabel 4.14 Asal Sekolah

Nama : asal_sekolah Primery Key : - Forigen Key : - Media : Hardisk Fungsi : Menyimpan Data Asal Sekolah Struktur Data :					
NO	FIELD	TYPE	SIZE	RANGE	KETERANGAN
1.	Nama Siswa	Varchar	50	100	Nama Siswa
2.	Asal Sekolah	Text	50	50	Asal Ssekolah
3.	Matematika	Varchar	5	10	Matematika
4.	Bahasa Indonesia	Varchar	10	20	Bahasa Indonesia
.	IPA	Varchar	10	20	IPA
6.	IPS	Varchar	10	20	IPS

Tabel 4.15 Tabel *Program Design*

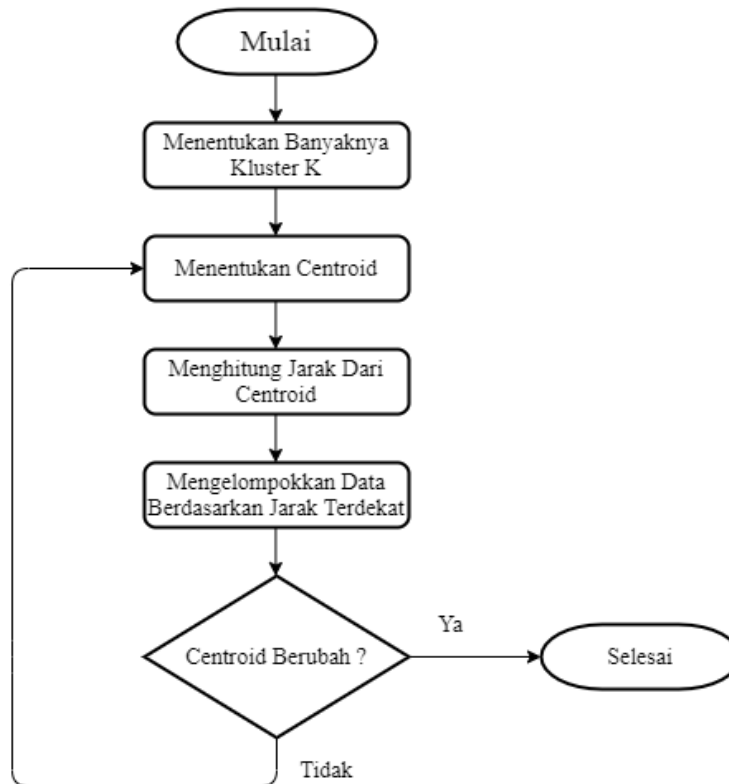
CLASS/TYPE	ATRIBUTES[TYPE]	METHODS[EVENTS or TYPE]
From Home	Home [Menu] Hasil Clastering [Menu] Logout [Menu]	From Main [Load] Home [Click] Hasil Clustering [Click] Logout [Click]
From Hasil Clastering	Tekan Tombol [button]	From Main [Load] Tekan [Click]
From Profil	Kembali [button]	From Main [Load] Kembali [click]
From Login	Username [button] Password [button]	From Main [Load] From Login [Close] Login [Click]

4. 7. Hasil Pengujian Sistem

4.7.1 Pengujian White Box

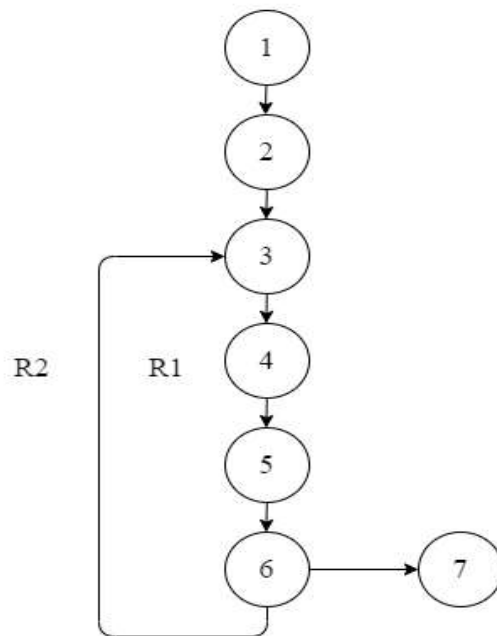
<?php	1
if(\$hc1<=\$hc2)	2
{	3
if(\$hc1<=\$hc3)	4
{	5
if(\$hc1<=\$hc4)	6
{	7
\$arr_c1[\$no] = 1;	8
}	9
Else	10
{	11
\$arr_c1[\$no] = '0';	12
}	13
}	14
Else	15
{	16
\$arr_c1[\$no] = '0';	17
}	18
}	19
Else	20
{	21
\$arr_c1[\$no] = '0';	22
}	23

4.7.2 Flowchart



Gambar 4.12 Gambar *Flowchart*

4.7.3 Flowgraph



Gambar 4.13 Gambar *Flowgraph*

Dari Flowgraph tersebut, didapatkan

Diketahui Region(R) = 2

Node(N) = 7

Edge(E) = 7

Predikat Node P = 1

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

Atau $V(G) = P + 1$

Penyelesaian : $V(G) = (7 - 7) + 2 = 2$

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

(R1, R2)

Tabel 4.16 *Basis Path*

NO	PATH	KET
1.	1-2-7	OK
2.	1-2-3-4-5-6-7	OK

4.7.6 Hasil Pengujian Black Box

Tabel 4.17 *Pengujian Black Box Admin*

Input/Event	Fungsi	Hasil Yang Diharapkan	Hasil Uji
Masukkan Username Dan Password Salah, Klik Login	Validasi Username Dan Pasword	Tampi Pesan “Maaf, kombinasi username dan password anda tidak sesuai	Sesuai
Masukkan Username Dan Password Benar, Klik Login	Validasi Username Dan Pasword	Tampil Halaman Home Admin	Sesuai
Klik Menu Data Siswa	Menampilkan Data Siswa	Tampil Data Siswa	sesuai
Klik Tambah Data Siswa	Menambahkan Data Siswa	Tampil Form Tambah Data Siswa	Sesuai
Masukkan Data Siswa Klik Tambah	Menyimpan Data Siswa	Tampil Pesan Data Ditambahkan	Sesuai
Klik Aksi Centang Pada Data Siswa	Mengubah Data Siswa	Tampil Form Edit Data Siswa	Sesuai
Masukkan Perubahan Data Siswa Klik Simpan	Mengubah Data Siswa	Tampil Pesan Data Telah Diubah	Ssesuai
Klik Aksi Hapus	Menhapus Data Siswa	Tampil Pesan “Anda Yakin Ingin Menghapus?”	Sesuai

Tabel 4.18 Tabel *Pengujian Black Box* Kepala Sekolah

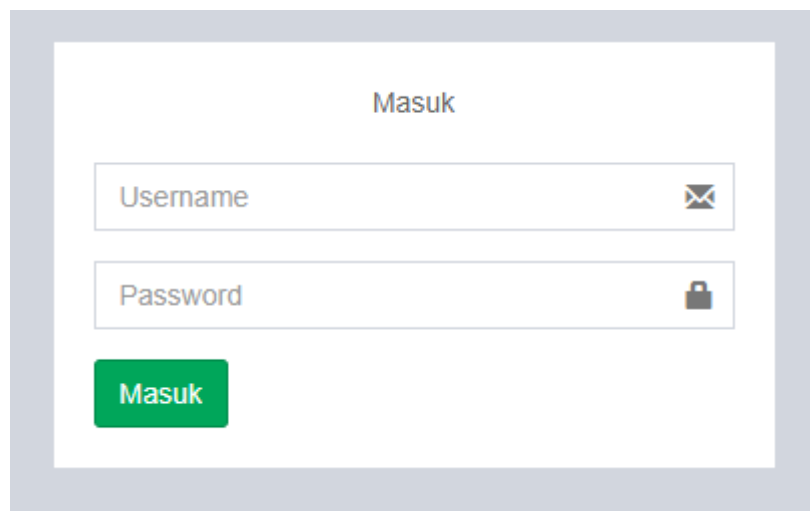
Input/Event	Fungsi	Hasil Yang Diharapkan	Hasil Uji
Masukkan Username Dan Password Salah, Klik Login	Validasi Username Dan Password	Tampi Pesan “Maaf, kombinasi username dan password anda tidak sesuai	Sesuai
Masukkan Username Dan Password Benar, Klik Login	Validasi Username Dan Password	Tampil Halaman Home Admin	Sesuai
Klik Menu Data Siswa	Menampilkan Data Siswa	Tampil Data Siswa	sesuai

BAB V

HASIL DAN PEMBAHASAN

5.1 Pembahasan Sistem

5.1.1 Tampilan Halaman Login

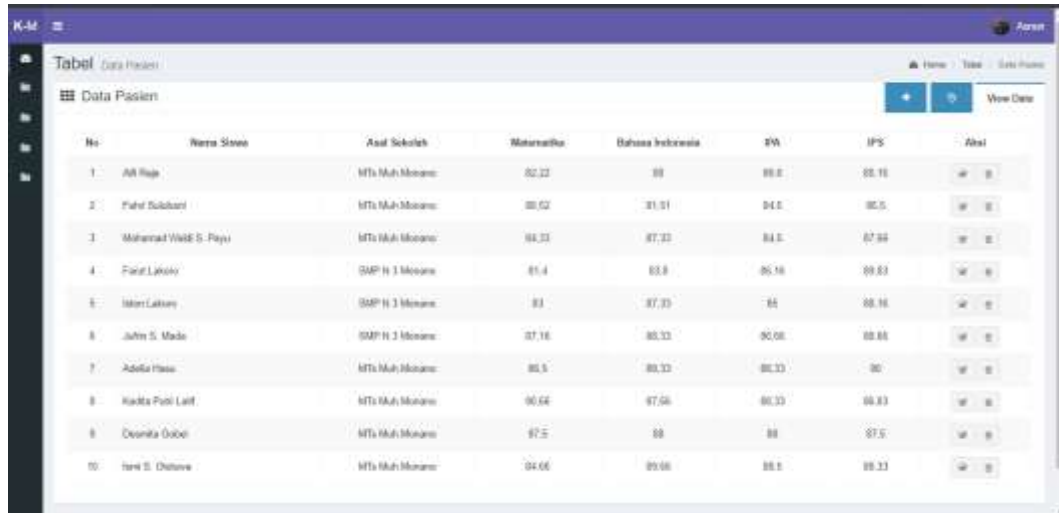






















The image shows a login interface within a light gray border. At the top center is the title 'Masuk'. Below it are two input fields: 'Username' with an envelope icon on the right, and 'Password' with a lock icon on the right. At the bottom left is a green button with the text 'Masuk' in white.

Gambar 5.1 Tampilan Halaman Login

Halaman ini digunakan untuk masuk ke halaman admin. Dimulai dengan memasukkan username, password untuk melanjutkan proses login silahkan klik tombol Masuk.

5.1.2 Tampilan Halaman Data Siswa



No	Nama Siswa	Asal Sekolah	Matematika	Bahasa Indonesia	IPA	IPS	Aksi
1	Ali Raja	MTs Muli-Mulani	82.22	88	88.8	88.18	 
2	Fahri Sukardi	MTs Muli-Mulani	88.52	81.51	84.8	85.5	 
3	Muhammad Wafid S. Pajus	MTs Muli-Mulani	88.33	87.33	84.8	87.84	 
4	Fahri Laksono	SDP H 3 Muli-Mulani	81.4	83.8	85.18	89.83	 
5	Iskren Laksono	SDP H 3 Muli-Mulani	83	87.33	84	88.18	 
6	Julien S. Mada	SDP H 3 Muli-Mulani	87.18	88.33	90.08	88.88	 
7	Adelia Hana	MTs Muli-Mulani	88.5	88.33	88.33	90	 
8	Kadita Putri Laili	MTs Muli-Mulani	90.66	87.66	88.33	88.83	 
9	Desmita Gubel	MTs Muli-Mulani	87.5	88	88	87.5	 
10	Irene D. Dhotave	MTs Muli-Mulani	84.66	85.66	88.8	88.33	 

Gambar 5.2 Halaman Data Siswa

Halaman ini merupakan tampilan dari Data Siswa yang terdiri dari Nama Siswa, Asal Sekolah, Nilai Rata-Rata IPA, Nilai Rata-Rata IPS, Dan Nilai UN. Untuk menambahkan data siswa klik tombol tambah, untuk mengubah data siswa klik aksi, dan untuk menghapus data klik hapus pada aksi.

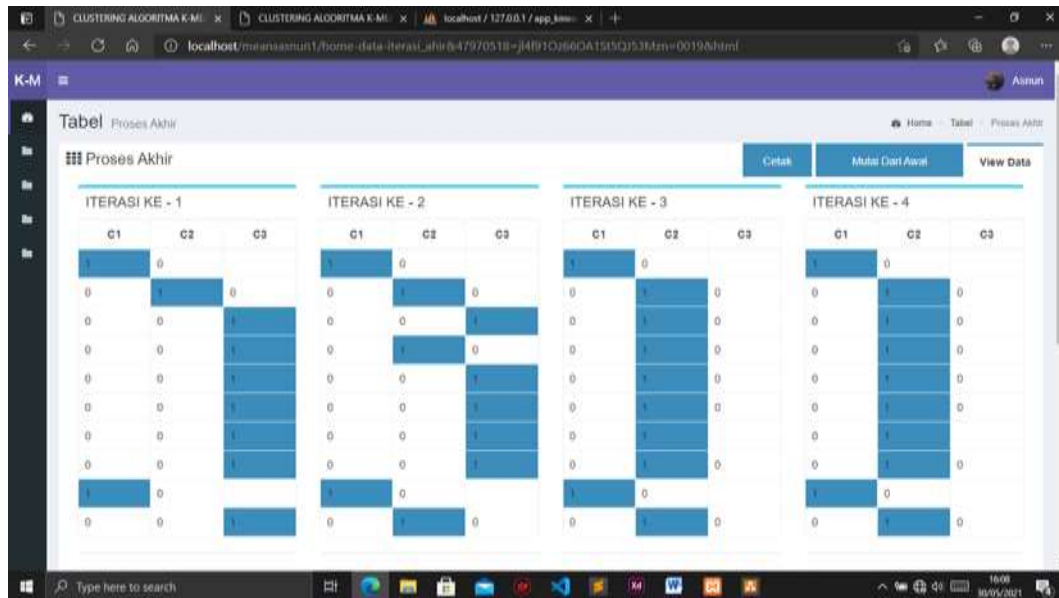
5.1.3 Tampilan Halaman Iterasi K-means

K-4

Gambar 5.3 Halaman Iterasi K-means

Halaman ini menampilkan hasil iterasi K-means yang terdiri dari Nama Siswa, Asal Sekolah, C1,C2, dan C3. Untuk melanjutkan Ke iterasi selanjutnya klik iterasi selanjutnya sampai muncul iterasi terakhir.

5.1.4 Tampilan Halaman Hasil Cluster



Gambar 5.4 Halaman Hasil Clustering

Halaman ini menampilkan hasil clustering, untuk kembali ke perhitungan awal klik data mulai dari awal

BAB VI

PENUTUP

6.1 Kesimpulan

Adapun hasil dari penelitian ini dapat disimpulkan bahwa :

- c. Bahwa perancangan sistem dalam mengelompokkan data penerimaan siswa baru menggunakan metode *K-mean* di Madrasah Aliyah Anggrek. Menghasilkan Clustering Siswa berdasarkan Nilai Mata Pelajaran Matematika, Bahasa Indonesia, Ipa dan Ips.
- d. Untuk mengetahui penerapan hasil dari pengelompokan data penerimaan siswa baru menggunakan metode *K-means*.

6.2 Saran

Setelah melakukan penelitian ini dengan kasus Pengelompokkan data penerimaan siswa baru. Maka ada beberapa hal yang perlu diperhatikan :

- a. Penelitian ini menggunakan Algoritma K-means dan penelitian ini juga bisa digunakan pada algoritma lain selain K-means.
- b. Penulis juga berharap agar penelitian ini dapat dijadikan sebagai acuan untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- Fina Nasari, Surya Darma, 2015, Penerapan K-means Clustering Pada Data Penerimaan Mahasiswa Baru (Studi kasus : UNIVERSITAS POTENSI UTAMA). Sistem Informasi Universitas Potensi Utama. JL.KL. Yos Sudarso Km 6,5 No 3A Tanjung Mulia-Medan.
Alamat : <http://ojs.amikom.ac.id>
- Fitri Yunita, 2018, Penerapan Data Mining Menggunakan K-means Clustering Pada Penerimaan Mahasiswa Baru (Studi kasus : UNIVERSITAS ISLAM IDRAGIRI). Program Studi Sistem Informasi, Fakultas Teknik Dan Ilmu Komputer Universitas Islam Indragiri (UNISI) JL. Propinsi, Parit 1 Tambilahan, Riau, Indonesia. Volume 7 Nomer 3 September 2018.
Alamat : <http://sistemasi.ftik.unisi.ac.id>
- Juliyana Tandy, Setiawan Assegaff, Analisis Dan Perancangan Clustering Siswa Baru Menggunakan Metode K-means (Studi kasus : SMK NEGERI 1 KOTA JAMBI). Program Studi Magister Sistem Informasi, STIKOM Dinamika Bangsa, Jambi. JL. Jendral Sudirman, Kec. Thehok, (0741) 305096. Vol. 4 No 3 Desember 2019.
Alamat : <http://ejournal.stikom.db.ac.id>
- Khotimah, T., Teknik, D. F., Studi, P., Informatika, T., & Kudus, U. M. (2014). Pengelompokan Surat Dalam AL-QUR'AN Menggunakan Algoritma K-means, 5(1)
- Langgeni, D. P., Baizal, Z. K. A., & W, Y. F.A. (2010). Clustering Artikel Berita Bahasa Indonesia, 2010(semnasIF).
- Mabrur A.G.R.L. (2012). Penerapan Data Mining Untuk Memprediksi Program Studi Teknik Informatika Jurnal Komputer Dan Informatika (KOMPUTA).

Nasari, F., Darma, S., & Informasi, S. (2015). Penerapan K-means Clustering Pada Data Penerimaan Mahasiswa Baru.

Jogianto HM, 2005 : 52. Perencanaan Sistem(Sistem Planning) Merupakan Pedoman Untuk Melakukan Pengembangan Sistem.

KODE PROGRAM

```

<div class="row">
  <div class="col-md-12">
    <!-- Custom Tabs -->
    <div class="nav-tabs-custom">
      <ul class="nav nav-tabs pull-right">
        <li class="active"><a href="#admin" data-
toggle="tab"><b>View Data</b></a></li>

        <li style="background-color:#3c8dbc; color:#FFF; width:15%; text-
align:center;" >
          <a href="home-data-
iterasi_lan&<?php echo $gett; ?>=<?php echo $result; ?>=0016&html" style="ba
ckground-color:#3c8dbc; color:#FFF; width:100%; text-align:center; border-
style:none; height:40px;">Proses Iterasi Selanjutnya</a></li>
        <li class="pull-left header"><i class="fa fa-th"></i>Data Awal</li>
      </ul>
      <div class="tab-content">
        <!-- /.admin -->
        <div class="tab-pane active" id="admin">
          <div class="box-body">
            <table id="tab22" class="table table-bordered table-striped">
              <thead>
                <tr>
                  <th rowspan="2" style="width:15px; vertical-align:middle; text-
align:center;">No</th>
                  <th rowspan="2" style="width:120px; vertical-align:middle; text-
align:center;">Nama Siswa</th>
                  <th rowspan="2" style="width:120px; vertical-align:middle; text-
align:center;">Matematika</th>
                  <th rowspan="2" style="width:50px; vertical-align:middle; text-
align:center;">Bahasa Indonesia</th>
                  <th rowspan="2" style="width:80px; vertical-align:middle; text-
align:center;">IPA</th>
                  <th rowspan="2" style="width:80px; vertical-align:middle; text-
align:center;">IPS</th>

                  <th colspan="4" style="width:50px; vertical-align:middle; text-
align:center;">Centorid 1</th>
                  <th colspan="4" style="width:50px; vertical-align:middle; text-
align:center;">Centorid 2</th>

```

```

        <th colspan="4" style="width:50px; vertical-align:middle; text-align:center;">Centorid 3</th>
        <th colspan="4" style="width:50px; vertical-align:middle; text-align:center;">Centorid 4</th>

        <th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C1</th>
        <th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C2</th>
        <th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C3</th>
        <th rowspan="2" style="width:40px; vertical-align:middle; text-align:center;">C4</th>
    </tr>
    <tr>
        <th>70</th><th>1</th><th>1</th><th>2</th>
        <th>27</th><th>2</th><th>2</th><th>2</th>
        <th>22</th><th>3</th><th>3</th><th>1</th>
        <th>18</th><th>4</th><th>4</th><th>2</th>
    </tr>
</thead>
<tbody>
<?php
    $c1a = 70;
    $c1b = 1;
    $c1c = 1;
    $c1d = 2;

    $c2a = 18;
    $c2b = 2;
    $c2c = 2;
    $c2d = 2;

    $c3a = 22;
    $c3b = 3;
    $c3c = 3;
    $c3d = 1;

    $c4a = 27;
    $c4b = 4;
    $c4c = 4;
    $c4d = 2;

    $c1a_b = "";
    $c1b_b = "";

```

```
$c1c_b = "";  
$c1d_b = "";
```

```
$c2a_b = "";  
$c2b_b = "";  
$c2c_b = "";  
$c2d_b = "";
```

```
$c3a_b = "";  
$c3b_b = "";  
$c3c_b = "";  
$c3d_b = "";
```

```
$c4a_b = "";  
$c4b_b = "";  
$c4c_b = "";  
$c4d_b = "";
```

```
$hc1=0;  
$hc2=0;  
$hc3=0;  
$hc4=0;
```

```
$no=0;  
$arr_c1 = array();  
$arr_c2 = array();  
$arr_c3 = array();  
$arr_c4 = array();
```

```
$arr_c1_temp = array();  
$arr_c2_temp = array();  
$arr_c3_temp = array();  
$arr_c4_temp = array();
```


DAFTAR RIWAYAT HIDUP

Nama : Asnun S Lagarusu

NIM : T3117236

TTL : Monano, 17 Juli 1999

Agama : Islam

Email : ayinunasnun@gmail.com

Riwayat Pendidikan :

1. Menyelesaikan Pendidikan Di Sekolah Dasar Negri 2 Monano Pada Tahun 2011
2. Menyelesaikan Pendidikan Di Madrasah Tsanawiyah Monano Pada Tahun 2014
3. Menyelesaikan Pendidikan Di Sekolah Menengah Kejuruan Pada Tahun 2017
4. Menyelesaikan Pendidikan Strata 1(Satu) Di Universitas Ichsan Gorontalo Pada Tahun 2021



KEMENTERIAN AGAMA REPUBLIK INDONESIA
MADRASAH ALIYAH ANGGREK
KABUPATEN GORONTALO UTARA
STATUS TERAKREDITAS "A" (AMAT BAIK)
NSM : 131275050001 NPSN : 60728037
Jln Trans. Buol – Gorontalo Kode Pos 96256 Tlf.....

SURAT KETERANGAN

No : 051 /MA-Angg/05/2021

Yang bertanda tangan di bawah ini :

Nama : Since G. Hajamati, S.Pd.I
NIP : 196406072014082001
Pangkat/Gol : Penata Muda Tingkat I/III.b
Jabatan : Kepala Madrasah
Unit Kerja : Madrasah Aliyah Anggrek

Menerangkan dengan sesungguhnya bahwa :

Nama : Asnun S. Lagarusu
N I M : T3117236
Fakultas : Ilmu Komputer
Program Studi : Teknik Informatika

Telah melaksanakan penelitian di Madrasah Aliyah Anggrek Mulai September sampai dengan Desember 2020 untuk memperoleh data guna penyusunan Tugas akhir Skripsi dengan judul : Implementasi Data Mining untuk Pengelompokan Data Penerimaan Siswa Baru menggunakan K-Means di Madrasah Aliyah Anggrek.

Demikian Surat Keterangan ini di buat untuk dapat digunakan sebagaimana mestinya.

Monsop, 31 Mei 2021
Kepala Madrasah

Since G. Hajamati, S.Pd.I
NIP. 19640607 201408 2 001



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS ICHSAN
(UNISAN) GORONTALO**

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001
Jl. Achmad Nadjamuddin No. 17 Telp (0435) 829675 Fax (0435) 829676 Gorontalo

SURAT REKOMENDASI BEBAS PLAGIASI

No. 0723/UNISAN-G/S-BP/VI/2021

Yang bertanda tangan di bawah ini:

Nama : Sunarto Taliki, M.Kom
NIDN : 0906058301
Unit Kerja : Pustikom, Universitas Ichsan Gorontalo

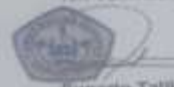
Dengan ini Menyatakan bahwa :

Nama Mahasiswa : AJSNUN S. LAGARUSU
NIM : T3117236
Program Studi : Teknik Informatika (S1)
Fakultas : Fakultas Ilmu Komputer
Judul Skripsi : Implementasi Data Mining Untuk Pengelompokan Data Penerimaan Siswa Baru Menggunakan K-means

Sesuai dengan hasil pengecekan tingkat kemiripan skripsi melalui aplikasi Turnitin untuk judul skripsi di atas diperoleh hasil Similarity sebesar 23%, berdasarkan SK Rektor No. 237/UNISAN-G/RSK/IX/2019 tentang Panduan Pencegahan dan Penanggulangan Plagiarisme, bahwa batas kemiripan skripsi maksimal 35% dan sesuai dengan Surat Pernyataan dari kedua Pembimbing yang bersangkutan menyatakan bahwa isi softcopy skripsi yang diolah di Turnitin SAMA ISINYA dengan Skripsi Aslinya serta format penulisannya sudah sesuai dengan Buku Panduan Penulisan Skripsi, untuk itu skripsi tersebut di atas dinyatakan BEBAS PLAGIASI dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Gorontalo, 01 Juni 2021
Tm Verifikasi,



Sunarto Taliki, M.Kom
NIDN. 0906058301

Tembusan :

1. Dekan
2. Ketua Program Studi
3. Pembimbing I dan Pembimbing II
4. Yang bersangkutan
5. Arsip

[illegible]

THEODORE ROBERTS, JR. AND J. L. ALLEN

Implementasi Data Mining Untuk Pengelompokan Data Pener...

Source: <http://www.irs.gov>

23%

Rank	Domain	Percentage
1	www.fox.com.au	10%
2	www.abc.com.au	4%
3	digitalarts.au	2%
4	www.fox.com.au	<1%
5	www.fox.com.au	<1%
6	www.fox.com.au	<1%
7	www.fox.com.au	<1%
8	www.fox.com.au	<1%
9	www.fox.com.au	<1%
10	www.fox.com.au	<1%
11	www.fox.com.au	<1%
12	www.fox.com.au	<1%
13	www.fox.com.au	<1%
14	www.fox.com.au	<1%
15	www.fox.com.au	<1%
16	www.fox.com.au	<1%
17	www.fox.com.au	<1%
18	www.fox.com.au	<1%
19	www.fox.com.au	<1%
20	www.fox.com.au	<1%

