

**IMPLEMNTASI DATA MINING UNTUK  
PENGELOMPOKKAN PENYAKIT HYPERTENSI  
DENGAN MENGGUNAKAN METODE K-MEANS**

**(Studi Kasus di Puskesmas Atinggola)**

**Oleh**

**ARDAN DJAFAR**

**T3117241**

**SKRIPSI**

Untuk memenuhi salah satu syarat ujian  
guna memperoleh gelar sarjana



**PROGRAM SARJANA  
TEKNIK INFORMATIKA  
UNIVERSITAS ICHSAN GORONTALO  
GORONTALO  
2023**

**PERSETUJUAN SKRIPSI**  
**IMPLEMENTASI *DATA MINING***  
**UNTUK PENGELOMPOKAN PENYAKIT *HYPERTENSI***  
**DENGAN MENGGUNAKAN METODE K-MEANS**

**(Studi Kasus di Puskesmas Atinggola)**

**Oleh**

**ARDAN DJAFAR**

T3117241

**SKRIPSI**

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar sarjana program studi teknik informatika,

Ini telah disetujui oleh tim pembimbing

Gorontalo, 12 April 2023

Pembimbing Utama



**Irvan Abraham Salihi, M.Kom**  
**NIDN : 0928028101**

Pembimbing Pendamping



**Suhardi Rustam, M.Kom**  
**NIDN :0915088403**

## PENGESAHAN SKRIPSI




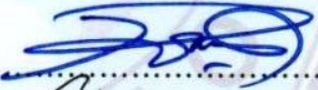

# IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN PENYAKIT HYPERTENSI DENGAN MENGGUNAKAN METODE K-MEANS

OLEH

ARDAN DJAFAR

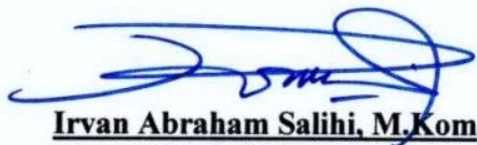
T3117241

Diperiksa Oleh Panitia Ujian Strata Satu (SI)  
Universitas Ichsan Gorontalo

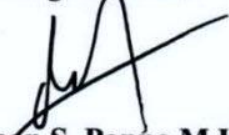
1. Ketua Penguji  
Irma Surya Kumala Idris, M.Kom ..... 
2. Anggota  
Yulianti Lasena, M.Kom ..... 
3. Anggota  
Sumarni, M.Kom ..... 
4. Anggota  
Irvan Abraham Salihi, M.Kom ..... 
5. Anggota  
Suhardi Rustam, M.Kom ..... 

Mengetahui

**Dekan Fakultas Ilmu Komputer**

  
**Irvan Abraham Salihi, M.Kom**  
NIDN 0928028101

**Ketua Program Studi**

  
**Sudirman S. Panna M.Kom**  
NIDN 0924038205

## 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 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/sitasi 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, 12 April 2023

Yang membuat pernyataan



ARDAN DJAFAR

## **ABSTRACT**

### **ARDAN DJAFAR. T3117241. THE IMPLEMENTATION OF DATA MINING FOR HYPERTENSION DISEASE CLUSTERING USING THE K-MEANS METHOD**

*This study aims to obtain (1) the performance results of the K-Means method in clustering hypertension disease data, and (2) find the results of clustering hypertension disease data using the K-Means method. This study uses a qualitative approach. The method used is data collection, done systematically. This study applies data mining techniques in grouping existing data into several clusters. Data or objects grouped into one cluster have the same characteristics. This study employs several variables used in the calculation of the K-Means clustering algorithm, namely name, gender, age, systolic blood pressure, diastolic blood pressure, weight, and heredity factor. This study uses 100 data obtained from the site by producing Cluster C1 with 35 data, C2 with 31 data, and C3 with 34 data. C1 is Stage 1 Hypertension, C2 is Pre-hypertension, and C3 is Normal. The results of accurate data grouping using the K-Means algorithm are considered very effective in grouping data. Coupled with the use of RapidMiner tools that are good at grouping data, this study can be completed properly.*

*Keywords: data mining, K-Means, hypertension disease data, Rapidminer*

## ABSTRAK

**ARDAN DJAFAR. T3117241 IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKAN PENYAKIT HYPERTENSI DENGAN MENGGUNAKAN METODE K-MEANS**

Penelitian ini bertujuan 1. Memperoleh hasil kinerja metode *k-means* dalam mengelompokkan data penyakit hipertensi. 2. Mengetahui hasil pengelompokan data penyakit hipertensi menggunakan metode *k-means*. Penelitian ini menggunakan pendekatan kualitatif dengan metode yang digunakan yakni pengumpulan data yang dikerjakan secara sistematis. Penelitian ini menerapkan teknik *data mining* dalam pengelompokan data yang ada ke dalam beberapa *cluster*. Data atau objek yang dikelompokkan ke dalam satu *cluster* memiliki karakteristik yang sama. Penelitian ini menggunakan beberapa variabel yang digunakan dalam perhitungan algoritma *K-Means clustering*, yaitu nama, jenis kelamin, umur, tekanan darah sistolik (TDS), tekanan darah diastolik (TDD) berat badan dan faktor keturunan. Penelitian ini menggunakan 100 data yang diperoleh dari tempat penelitian dengan menghasilkan *Cluster C1* sebanyak 35 data, *C2* sebanyak 31 data, *C3* sebanyak 34 data. *C1* merupakan hipertensi stage 1, *C2* merupakan pre-hipertensi, *C3* merupakan Normal. Hasil pengelompokan data yang akurat dalam menggunakan algoritma *k-means* dinilai sangat efektif dalam mengelompokkan data. Ditambah lagi dengan penggunaan tools *rapidminer* yang bagus dalam mengelompokkan data maka penelitian ini dapat di selesaikan dengan baik.

Kata Kunci : Data mining , K-Means, Data penyakit hipertensi, *Rapidminer*





## KATA PENGANTAR

Alhamdulillahirobbil'alamiin, penulis dapat menyelesaikan skripsi ini dengan judul: “IMPLEMENTASI DATA MINING UNTUK PENGELOMPOKKAN PENYAKIT HYPERTENSI DENGAN MENGGUNAKAN METODE K-MEANS”, untuk memenuhi salah satu syarat memperoleh gelar Sarjana Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

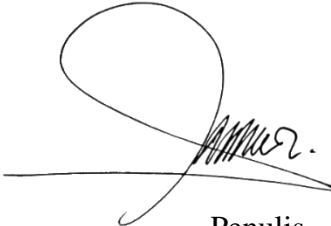
Penulis menyadari sepenuhnya bahwa penulisan skripsi ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, dengan segala keikhlasan dan kerendahan hati, penulis mengucapkan banyak terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Ibu Dr. Hj. Juriko Abdussamad, 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. Bapak Irvan Abraham Salihi, S.kom, M.kom, selaku Pembimbing utama
4. Bapak Sudirman S. Melangi, 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 Adminstrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
6. Bapak Sudirman S. Panna, S.Kom M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
7. Bapak Suhardi Rustam, S.kom M.kom, selaku Pembimbing Pendamping yang telah membimbing penulis selama menyusun usulan penelitian;
8. Tenaga Kesehatan dan Kepala Puskesmas Atinggola yang telah membantu penulis dalam pengumpulan data di lapangan;
9. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis;

10. Orang tua yang tercinta, atas segala kasih sayang, jerih payah dan doa restunya dan membersarkan serta mendidik penulis;
11. Kakak - kakak yang selalu memotivasi dan menyemangati penulis;
12. Rekan-rekan seperjuangan yang telah banyak memberikan bantuan dan dukungan moril yang sangat besar kepada penulis;
13. Kepada semua pihak yang ikut membantu penyelesaian skripsi ini yang tak sempat penulis sebutkan satu-persatu.

Penulis menyadari sepenuhnya bahawa apa yang telah dicapai ini masih jauh dari kesempurnaan dan masih banyak terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif. Akhirnya, penulis berharap semoga hasil yang telah dicapai ini dapat bermanfaat bagi kita semua, Aamiin.

Gorontalo, 12 April 2023



Penulis



## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>PERSETUJUAN SKRIPSI.....</b>	<b>ii</b>
<b>PENGESAHAN SKRIPSI.....</b>	<b>iii</b>
<b>PERNYATAAN SKRIPSI.....</b>	<b>iv</b>
<b>ABSTRACT .....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vi</b>
<b>KATA PENGANTAR.....</b>	<b>vii</b>
<b>DAFTAR ISI.....</b>	<b>ix</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR TABEL.....</b>	<b>xii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	3
1.3 Rumusan Masalah.....	4
1.4 Tujuan Penelitian .....	4
1.5 Manfaat Penelitian .....	4
<b>BAB II LANDASAN TEORI .....</b>	<b>5</b>
2.1 Tinjauan Studi.....	5
2.2 Tinjauan Pustaka.....	7
2.2.1 Hipertensi .....	7
2.2.2 <i>Clustering</i> .....	7
2.2.3 Data Mining .....	7
2.2.4 Algoritma K-Means .....	9
2.2.5 Penerapan Algoritma K-Means.....	10
2.2.6 Rapid Miner .....	11
2.2.7 Hipertensi .....	12
2.2.8 Hipertensi Stage 1 .....	12
2.2.9 Pre-Hipertensi .....	12
2.2.10 Normal .....	12
2.2.11 Tekanan Sistolik Darah.....	12
2.2.12 Tekanan Sistolik Diastolik.....	13
2.3 Kerangka Pikir .....	14

<b>BAB III METODE PENELITIAN .....</b>	<b>15</b>
3.1 Jenis, Metode, Subjek, Waktu dan Lokasi Penelitian.....	15
3.2 Pengumpulan Data.....	15
3.3 Permodelan .....	16
3.4 Pra pengolahan.....	17
3.5 Hasil Clustering .....	17
3.6 Evaluasi.....	17
<b>BAB IV PEMBAHASAN.....</b>	<b>18</b>
4.1 Hasil Pengambilan Data.....	18
4.2 Penjelasan Algoritma .....	19
4.3 Perhitungan K-means.....	19
4.4 Perhitungan Excel.....	23
4.4.1 Iterasi 1 .....	23
4.4.2 Iterasi 2.....	24
4.4.3 Iterasi 3.....	25
<b>4.4.4 Hasil Perhitungan Data Training.....</b>	<b>26</b>
<b>BAB V PEMBAHASAN PENELITIAN .....</b>	<b>28</b>
5.1 Pembahasan Model.....	28
5.2 Pembahasan Tools .....	28
5.2.1 Tampilan Data Excel Yang Akan Digunakan .....	28
5.2.2 Tampilan utama Rapidminer.....	29
5.2.3 Import data .....	29
5.2.4 Lokasi penyimpanan data.....	30
5.2.5 Data yang digunakan.....	30
5.2.6 Tampilan Proses Pemasukan Data dan K-Means.....	31
5.2.7 Result Data .....	31
5.2.8 Tampilan Cluster Data Model.....	32
5.2.9 Pembahasan Algoritma .....	32
<b>BAB VI PENUTUP .....</b>	<b>36</b>
6.1 Kesimpulan .....	36
6.2 Saran .....	36
<b>DAFTAR PUSTAKA .....</b>	<b>37</b>
LAMPIRAN	

## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Tahapan knowledge discovery in database .....	8
<b>Gambar 2. 2</b> Diagram Alur Implementasi Algoritma K-Means .....	10
<b>Gambar 2. 3</b> Kerangka Pikir .....	14
<b>Gambar 3. 1</b> Permodelan K- Menas .....	14
<b>Gambar 5. 1</b> Permodelan K-Means .....	26
<b>Gambar 5. 2</b> Tampilan Data Excel yang akan digunakan .....	26
<b>Gambar 5. 3</b> Tampilan Utama Rapid Miner .....	27
<b>Gambar 5. 4</b> Import Data.....	27
<b>Gambar 5. 5</b> Data yang digunakan .....	28
<b>Gambar 5. 6</b> Tampilan Utama Rapid Miner .....	28
<b>Gambar 5. 7</b> Tampilan proses pemasukan data dan K-Means .....	29
<b>Gambar 5. 8</b> Tampilan Cluster model .....	29
<b>Gambar 5. 9</b> Tampilan Result Data .....	30

## DAFTAR TABEL

<b>Table 1. 1</b> Data Set Pasien Penyakit Hipertensi .....	2
<b>Tabel 2. 1</b> Tinjauan Studi .....	5
<b>Tabel 2. 2</b> Cluster .....	11
<b>Tabel 3. 1</b> Variabel Data.....	16
<b>Tabel 4. 1</b> Pengumpulan Data .....	18
<b>Tabel 4. 2</b> Data Mining.....	19
<b>Tabel 4. 3</b> Data Centroid Awal.....	20
<b>Tabel 4. 4</b> Iterasi 1 .....	24
<b>Tabel 4. 5</b> Centroid Baru Iterasi 2 .....	24
<b>Tabel 4. 6</b> Iterasi 2 .....	25
<b>Tabel 4. 7</b> Centroid Baru Iterasi 3 .....	26
<b>Tabel 4. 8</b> Iterasi 3 .....	26
<b>Tabel 5. 1</b> Hasil <i>Cluster</i> .....	33

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Penyakit merupakan suatu keadaan abnormal tubuh atau pikiran yang menyebabkan ketidaknyamanan disfungsi atau kesukaran terhadap orang yang di pengaruhinya. Oleh karena itu, kita harus tahu tentang kesehatan kita, setidaknya kita tahu tentang diri kita sendiri. Apabila merasa sakit sebaiknya kita langsung memeriksakan diri ke dokter atau puskesmas terdekat untuk mencegah agar rasa sakit yang kita rasakan tidak menjadi parah.

Hipertensi adalah keadaan di mana tekanan darah mengalami peningkatan yang memberikan gejala berlanjut pada suatu organ target ditubuh. Hal ini dapat menimbulkan kerusakan yang lebih berat, misalnya *stroke* (terjadi pada otak dan menyebabkan kematian yang cukup tinggi), penyakit jantung koroner (terjadi kerusakan pembuluh darah jantung), dan hipertrofi ventrikel kiri (terjadi pada otot jantung). Hipertensi juga dapat menyebabkan penyakit gagal ginjal, penyakit pembuluh lain dan penyakit lainnya [1].

Puskesmas Atinggola merupakan suatu pusat pelayanan kesehatan masyarakat yang bertempat di Jalan Trans Sulawesi Desa Monggupo Kecamatan Atinggola Kabupaten Gorontalo Utara. Setiap harinya, Puskesmas Atinggola melayani pasien dari berbagai wilayah yang ada di Kecamatan Atinggola di antaranya adalah pasien penderita penyakit hipertensi.

Dalam dunia kesehatan, penyakit jantung merupakan penyakit yang mendorong angka kematian yang cukup tinggi, sehingga banyak penelitian yang dilakukan sebelumnya untuk memprediksi penyakit jantung, diantaranya penelitian tentang faktor risiko hipertensi [2] tentang faktor yang dapat memperbesar risiko atau kecenderungan seseorang menderita hipertensi, [3]. Berdasarkan penelitian yang telah dilakukan sebelumnya, perlu adanya pengelompokan tingkat penyakit Hipertensi berdasarkan dataset yang diperoleh menggunakan salah satu teknik *data mining* yaitu metode *clustering*.

Data mining merupakan metode pengolahan data dalam jumlah besar. Data mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting [4]. Sesingga metode data mining sangat cocok untuk pengolahan data hipertensi. Berdasarkan fungsionalitas, data mining terbagi dalam 5 metode yaitu estimasi, prediksi, klasifikasi, clustering, dan asosiasi.

Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database (KDD)* secara keseluruhan. Sebab tujuan penelitian untuk mengelompokan data hipertensi, maka algoritma yang cocok adalah clustering. Clustering yang merupakan metode pengelompokan data. Tujuan dari pengelompokan cluster ini adalah untuk menemukan pengelompokan dari serangkaian pola, titik, objek maupun dokumen. Objek yang berada didalam pengelompokan cluster yang sama memiliki kemiripan antar satu kelompok dan memiliki perbedaan dengan objek oleh kelompok cluster lain [5].

Pada penelitian ini penulis menemukan masalah yang ada di tempat penelitian yaitu data pasien penyakit hipertensi belum dikelompokan. Data pasien ditulis secara manual dalam satu buku besar yakni buku registrasi data pasien penderita penyakit hipertensi. Dengan begitu sangat besar kemungkinan data tersebut akan rusak atau hilang jika data tersebut tidak diolah menggunakan metode *data mining* dengan algoritma *clustering*. Berikut adalah data pasien penyakit hipertensi dapat dilihat dalam table 1.1

**Table 1.1** Data Set Pasien Penyakit Hipertensi

No	Nama	L/ P	Umur	TDS (mmHg)	TDD (mmHg)	Berat Badan	Faktor Keturunan
	Arman Nusi	L	58	170	80	56	Ada
	Mei Hasan	P	52	160	90	48	Tidak Ada

	Makmur Panigoro	L	72	160	100	68	Tidak Ada
	Gustin Mahmud	P	51	160	90	62	Tidak Ada
	Abd. Najib Blongkod	L	49	150	90	69	Tidak Ada
	Yahya Gobel	L	45	180	90	54	Tidak Ada
	Rostin Humagi	P	50	150	90	76	Ada
	Dahlan Buheli	L	51	220	130	68	Tidak Ada
	Nino Buheli	P	53	200	100	59	Tidak Ada
	Arisna Dunggio	P	55	190	60	60	Tidak Ada

**Tabel 1.1** Data Set Pasien Penyakit Hypertensi

(Sumber : Puskesmas Atinggola, 2019-2020).

Attribut yang digunakan pada penelitian ini adalah nama, jenis kelamin, umur, tekanan darah sistolik (TDS), tekanan darah diastolic (TDD), berat badan dan factor keturunan.

Berdasarkan uraian diatas, maka di anggap perlu untuk melakukan penelitian mengenai proses yang berjalan diatas, dengan judul ***“IMPLEMENTASI DATA MINING UNTUK MENGELOMPOKAN PENYAKIT HIPERTENSI DENGAN MENGGUNAKAN METODE K-MEANS CLUSTERING ”***

## **1.2 Identifikasi Masalah**

Berdasarkan uraian di atas, dapat dirumuskan masalah pokok yang berkaitan yaitu belum adanya pengelompokan penyakit hipertensi di Puskesmas Atinggola.



### 1.3 Rumusan Masalah

1. Bagaimana kinerja metode *k-means clustering* dalam pengelempokan data penyakit hipertensi menggunakan metode *k-means clustering* ?
2. Bagaimana hasil pengelempokan data penyakit hipertensi menggunakan menggunakan metode *k-means clustering*?

### 1.4 Tujuan Penelitian

1. Memperoleh hasil kinerja metode *k-means clustering* dalam pengelempokan data penyakit hipertensi menggunakan?
2. Mengetahui hasil pengelempokan data penyakit hipertensi menggunakan menggunakan metode *k-means clustering*?

### 1.5 Manfaat Penelitian

Adapun manfaat penelitian sebagai berikut :

#### 1. Pengembangan Ilmu

Dari hasil penelitian ini diharapkan dapat menambah pengetahuan dibidang teknoogi pada umumnya khususnya pada bagaimana cara menerapkan algoritma *K-Means* untuk mengelopokan penyakit Hipertensi

#### 2. Praktis

Sebagai bahan masukan untuk mendukung dalam melaksanakan penelitian tentang Implementasi Data Mining Dengan Algoritma K-Means Clustering Penyakit Hipertensi di Puskesmas Atinggola

#### 3. Peneliti

Dapat dilakukan pengembangan lebih lanjut dengan mengcluster penyakit Hipertensi berdasarkan jenis penyakitnya.

## 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** Tinjauan Studi

No	Peneliti	Judul	Tahun	Metode	Hasil
1.	Anindya Khrisna Wardani	Implementasi algoritma k-means untuk mengelompokkan penyakit pasien pada puskesmas kaje pekalongan	2016	K-Means Clustering	Penelitian ini menghasilkan sebuah kesimpulan bahwa dalam perangkaian kesehatan pada suatu daerah masih menggunakan cara manual untuk perhitungan data. Untuk menentukan konsistensi data kesehatan dapat digunakan teknik data mining yang mampu mengenali informasi tersembunyi dari kumpulan data multidimensi yang telah di peroleh.
2	Ahmad chusayairi, pelsri ramadar	Pengelompokan data puskesmas banyuwangi dalam	2019	K-Means Clustering	Penormalisasian data dilakukan dengan menggunakan metode min-max di mana akan dicari nilai terkecil dan

	noor saputra	Imunisasi menggunakan metode k- means clustering			terbesar. Dari perhitungan jarak tersebut maka akan dihasilkan data-data imunisasi dengan jarak yang terdekat terhadap titik dari ketiga pusat <i>centroid</i> tersebut.
3	Diajeng tyas purwa hapsari, 2017	Penerapan algoritma k- means pada kualitas gizi bayi di Indonesia	2017	Menggunakan metode k-means	Pada <i>cluster</i> 1 terdapat 15 responden dan cluster2 terdapat 5 responden pada hasil clustering k-means tersebut tingkat gizi pada 5 provinsi berada dalam kategori baik. Ditinjukan dengan inisial cluster yang Menunjukkan kelompok-kelompok cluster. Final cluster yang menunjukkan hasil cluster berada pada rata-rata total. Dan sumber of case in each cluster menunjukkan terdapat 15 responden pada cluster 1 serta 5 responden pada cluster 2.

## 2.2 Tinjauan Pustaka

### 2.2.1 Hipertensi

Hipertensi atau yang biasa disebut tekanan darah tinggi merupakan peningkatan tekanan darah sistolik di atas batas normal yaitu lebih dari 140 mmHg dan tekanan darah diastolik lebih dari 90 mmHg [6].

Penyakit hipertensi atau tekanan darah tinggi adalah salah satu jenis penyakit yang mematikan di dunia dan faktor risiko paling utama terjadinya hipertensi yaitu faktor usia sehingga tidak heran penyakit hipertensi sering dijumpai pada usia senja/ usia lanjut [7]. Hipertensi merupakan tanda klinis ketidakseimbangan hemodinamik suatu sistem kardiovaskular, di mana penyebab terjadinya disebabkan oleh beberapa faktor/ multi faktor sehingga tidak bisa terdiagnosis dengan hanya satu faktor tunggal [8].

### 2.2.2 Clustering

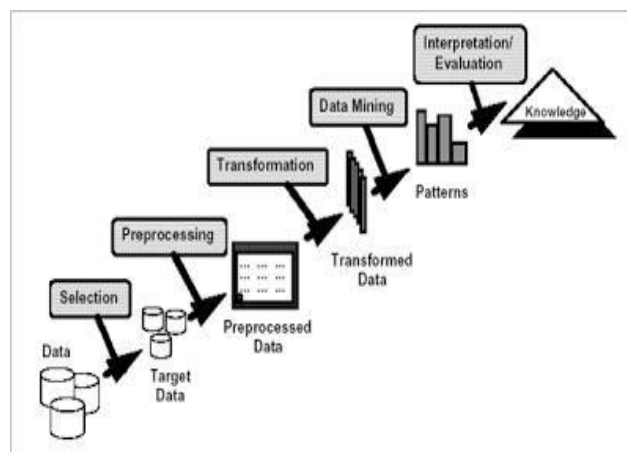
*Clustering* adalah suatu metode pengelompokan berdasarkan kedekatan atau kemiripan. *Clustering* berbeda dengan grup, kalau grup berarti kelompok yang sama kondisinya kalau tidak ya pasti bukan kelompoknya[10]. Tetapi kalau *cluster* tidak harus sama akan tetapi pengelompokannya berdasarkan pada kedekatan dari suatu karakteristik simpel yang ada. Salah satunya dengan menggunakan rumus jarak *euclidean*. Aplikasi *cluster* ini sangat banyak, karena hampir dalam mengidentifikasi permasalahan atau pengambilan keputusan selalu tidak sama persis akan tetapi cenderung memiliki kemiripan saja.

### 2.2.3 Data Mining

Data mining didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan kecenderungan baru yang bermakna dengan menyaring data yang sangat besar, yang tersimpan dalam penyimpanan, menggunakan teknik pengenalan pola seperti teknik statistik dan matematika. Data Mining adalah proses mengekstraksi informasi atau sesuatu yang penting atau menarik dari data yang ada didalam database sehingga menghasilkan informasi. Suatu istilah yang digunakan untuk menguraikan penemuan

pengetahuan didalam *database* adalah data mining. Data mining merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning*, untuk mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari berbagai database besar [9]. Data mining merupakan suatu tahapan dalam *knowledge discovery in database* (KDD).

KDD (*Knowledge discovery in database*) merupakan keseluruhan proses *non-trivial* untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru dapat bermanfaat dan dapat dimengerti [10]. Berikut merupakan tahapan proses KDD dapat dilihat pada gambar 2.1 berikut ini:



**Gambar 2. 1** Tahapan *knowledge discovery in database*

Tahapan proses KDD terdiri dari :

1. Data selection

Perlu dilakukan pemilihan atau seleksi data dari sekumpulan data operasional sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil pemilihan disimpan dulu dalam satu berkas terpisah dari database operasional sebelum digunakan.

2. *Pre-processing / cleaning*

Perlu dilakukan pembersihan data yang menjadi fokus KDD sebelum proses data mining dilaksanakan.

### 3. *Transformation*

Merupakan proses transformasi pada data yang dipilih, sehingga sesuai untuk proses data mining. Proses ini merupakan proses yang sangat tergantung pada jenis/pola informasi yang akan dicari dalam *database*.

### 4. Data mining

Data mining merupakan proses pencarian pola atau informasi menarik dalam data terpilih dengan menggunakan algoritma atau metode tertentu. Algoritma atau metode tertentu dalam data mining sangat bervariasi. Pemilihan algoritma atau metode yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

### 5. *Interpretation / evaluation*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dipahami oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*.

Tahap ini meliputi pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

#### 2.2.4 **Algoritma K-Means**

K-means adalah salah satu metode pengelompokan data nonhirarki yang berusaha mempartisi data yang ada kedalam bentuk dua atau lebih kelompok. Metode ini mempartisi data kedalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam satu kelompok yang sama. Dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain [2]. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok.

Hitung jarak setiap data yang ada terhadap masing-masing *centroid* menggunakan rumus *Euclidean* hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Klasifikasikan setiap data berdasarkan kedekatannya dengan *centroid*. Lakukan langkah tersebut hingga nilai *centroid* tidak berubah. Jarak *Euclidean* yang dirumuskan sebagai berikut:

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2}$$

dimana:

De adalah *Euclidean distance*

i adalah banyaknya objek,

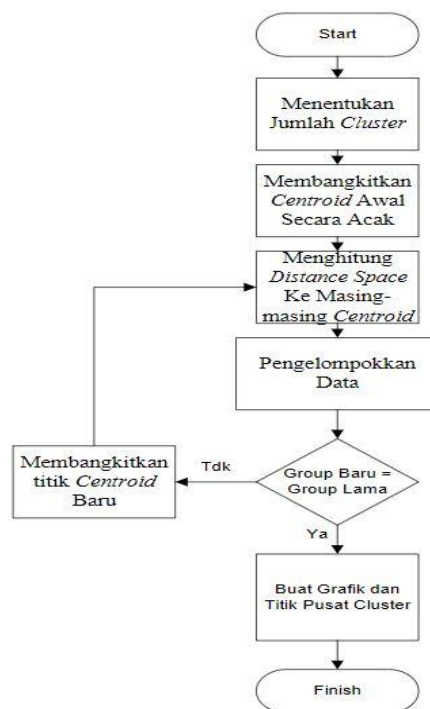
(x,y) merupakan koordinat objek dan

(s,t) merupakan koordinat *centroid*.

### 2.2.5 Penerapan Algoritma K-Means

#### a. Menentukan Jumlah *Cluster*

Dapat dilihat pada gambar dibawah ini merupakan diagram alur dari metode k-means yang digunakan dalam pengelompokan data penyakit hipertensi di Puskesmas Atinggola, pada umumnya kinerja metode k-means secara berurutan adalah sebagai berikut :



**Gambar 2. 2** Diagram Alur Implementasi Algoritma K-Means

Pada pengelompokan penyakit Hipertensi di Puskesmas Atinggola di buat menjadi 3 *Cluster*, penentuan *Cluster* tersebut terdiri dari 7 variabel nama, jenis kelamin, umur, tekanan darah sistolik (TDS), tekanan darah diastolic (TDD), berat dadan dan factor keturunan, dengan *Cluster* sebagai berikut :



**Tabel 2. 2 Cluster**

NO	Cluster	Nama Cluster
1	C1	Hipertensi Stage 1
2	C2	Pre-hipertensi
3	C3	Normal

b. Menentukan Titik Pusat Awal *Cluster (Centroid)*

Dalam penerapan algoritma k-means dihasilkan nilai titik tengah dari data yang ditemukan dengan ketentuan bahwa pengelompokan yang diinginkan adalah 3, yaitu (C1) Normal, (C2) Pre-Hipertensi, dan (C3) Hipertensi stage 1. Maka nilai titik tengah juga terdapat 3 titik. Penentuan titik *cluster* ini dilakukan dengan mengambil nilai terbesar untuk cluster Normal (C1), nilai rata-rata untuk cluster Pre-hipertensi (C2) dan nilai terkecil untuk *cluster* Hipertensi stage 1 (C3).

### 2.2.6 Rapid Miner

RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner merupakan *software* yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri.

Rapid Miner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik [11].

### **2.2.7 Hipertensi**

Menurut World Health Organization [12], hipertensi merupakan suatu kondisi dimana pembuluh darah terus-menerus mengalami peningkatan tekanan. Adanya peningkatan tekanan pada pembuluh darah mengakibatkan kerja jantung untuk memompa darah semakin keras/cepat. Hipertensi juga dapat didefinisikan sebagai peningkatan tekanan sistolik lebih besar atau sama dengan 140 mmHg dan atau tekanan diastolik sama atau lebih besar 90 mmHg.

Menurut Mansjoer, berdasarkan penyebabnya hipertensi dibagi menjadi dua golongan, yaitu hipertensi primer adalah hipertensi yang tidak diketahui penyebabnya secara pasti dan hipertensi sekunder adalah hipertensi yang diketahui secara spesifik penyebabnya [13].

### **2.2.8 Hipertensi Stage 1**

Hipertensi Stage 1 merupakan kondisi dimana tekanan darah sistolik 140-159 mmHg atau tekanan darah diastolik 90-99 mmHg dan memungkinkan penderita memerlukan pengobatan karena resiko terjadinya kerusakan pada organ menjadi lebih tinggi

### **2.2.9 Pre-Hipertensi**

Pre-Hipertensi merupakan kondisi dimana tekanan darah sistolik 120-139 mmHg atau tekanan darah diastolik 80-89 mmHg. Individu dengan Pre-Hipertensi tergolong beresiko lebih tinggi terkena hipertensi.

### **2.2.10 Normal**

Normal merupakan merupakan kondisi apabila tekanan darah sistolik kurang dari 120 mmHg dan Tekanan Diastolik kurang dari 80 mmHg atau bisa ditulis dengan 120/80 mmHg.

### **2.2.11 Tekanan Sistolik Darah**

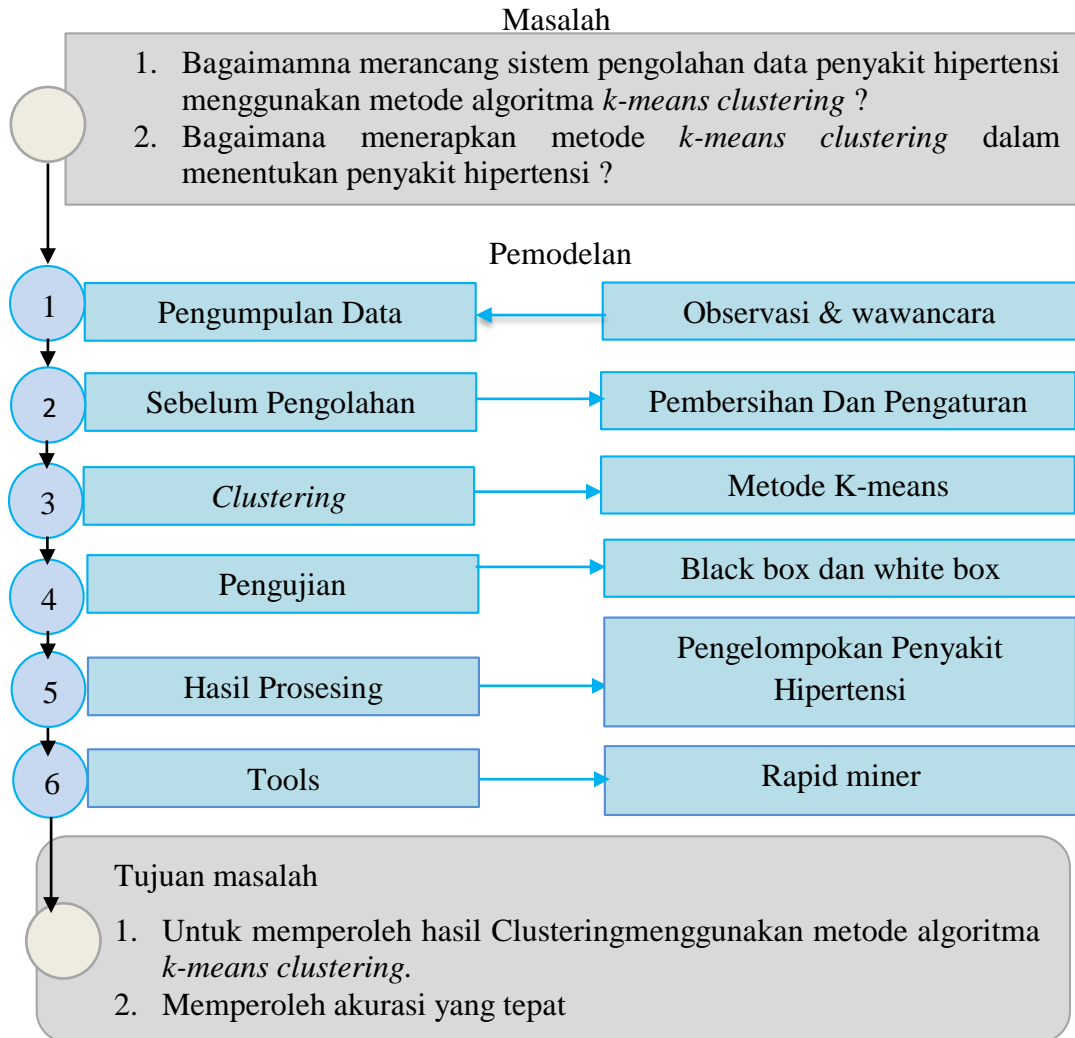
Menurut Ronny, dkk [14] tekanan darah sistole merupakan tekanan darah yang terukur pada saat ventrikel kiri jantung berkontraksi (sistole). Darah mengalir dari jantung ke pembuluh darah sehingga pembuluh dasar sehingga pembuluh darah teregang maksimal. Pada pemeriksaan fisik, bunyi “lup” pertama yang terdengar adalah tekanan darah sistolik. Tekanan darah sistolik orang normal 120 mmHg.

Seseorang bisa dikatakan mengalami hipertensi bila pembacaan tekanan darah sistolik pada pengukuran selama dua hari berturut-turut menunjukkan hasil yang lebih besar dari 140 mmHg.

#### **2.2.12 Tekanan Sistolik Diastolik**

Menurut Ronny dkk [15] tekanan diastole merupakan tekanan darah yang terukur yang terjadi pada saat jantung berelaksasi (diastol). Pada saat diastol, tidak ada darah mengalir dari jantung ke pembuluh sehingga pembuluh darah dapat kembali ke ukuran normalnya sementara darah didorong ke bagian arteri yang lebih distal. Pada pemeriksaan fisik, tekanan darah diastol dapat ditentukan melalui bunyi “dup” terakhir yang terdengar. Pada orang normal, diastol adalah 80 mmHg

### 2.3 Kerangka Pikir



**Gambar 2. 3** Kerangka Pikir

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis, Metode, Subjek, Waktu dan Lokasi Penelitian**

Dipandang dari jenis informasi yang diolah, maka penelitian ini adalah penelitian kuantitatif. Penelitian ini menggunakan metode study kasus. Data penelitian ini diperoleh dengan survey lokasi dan wawancara sehingga Metode yang digunakan dalam penelitian untuk masalah *Clustering* Data Penyakit Hypertensi yaitu metode *Deskriptif* dimana metode ini bertujuan untuk mengungkapkan kebenaran yang *Objektif*.

Berdasarkan latar belakang dan kerangka pemikiran seperti yang telah diuraikan dalam Bab I dan Bab II, maka menjadi objek penelitian adalah Clustering Data Penyakit Hypertensi. Penelitian ini dimulai tanggal 14 bulan November tahun 2021 yang berlokasi di Puskesmas Atinggola Desa Monggupo Kecamatan Atinggola Kabupaten Gorontalo Utara.

#### **3.2 Pengumpulan Data**

Metode pengumpulan data yang digunakan untuk mendapatkan dan informasi digunakan 2 (dua) jenis data, yaitu data primer dan data sekunder sebagai berikut:

a) Data Primer (Observasi, Wawancara)

Penelitian primer merupakan proses pengambilan data dengan melihat dan mempelajari masalah yang ada di lokasi penelitian yang berkaitan dengan objek yang diteliti. Selain itu wawancara dilakukan dengan mengajukan pertanyaan ke pihak Puskesmas Atinggola.

b) Data Sekunder (Kepustakaan)

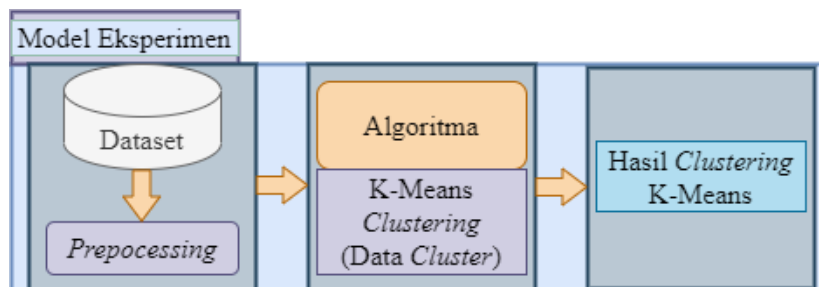
Penelitian sekunder merupakan pengambilan informasi dengan melakukan pengkajian kepustakaan yang berisi dasar-dasar teori. metode kepustakaan digunakan untuk mengambil contoh dokumen yang berhubungan dengan objek penelitian.

Adapun variable dengan tipe datanya masing-masing ditunjukkan pada tabel berikut ini:

**Tabel 3. 1** Variabel Data

No	Nama	Type	Keterangan
1	Nama Pasien	Binominal	Parameter Input
2	Umur	Nominal	Parameter Input
3	Jenis Kelamin	Binominal	Parameter Input
4	Tekanan Darah Sistolik(TDS)	Nominal	Parameter Input
5	Tekanan Darah Diastolik(TDD)	Nominal	Parameter Input
6	Berat Badan	Nominal	Parameter Input
7	Faktor Keturunan	Binominal	Parameter Input
8	Normal	Nominal	C1(Output)
9	Pre-hipertensi	Nominal	C2 (Output)
10	Hipertensi Stage 1	Nominal	C3 (Output)

### 3.3 Permodelan

**Gambar 3. 1** Permodelan K-Means

Pemodelan K-Means sendiri dimulai dari tahap pengumpulan dataset oleh peneliti di lokasi penelitian, setelah pengumpulan data, dilakukan pemilihan data (seleksi). Setelah diperoleh data seleksi dilakukan pembersihan data yang tidak perlu kemudian dilakukan proses algoritma K-means Clustering. Proses algoritma K-means digunakan 100 dataset dengan jumlah  $k=3$ . Dari proses tersebut mendapatkan hasil cluster.

### **3.4 Pra pengolahan**

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak.

### **3.5 Hasil Clustering**

Proses algoritma K-means digunakan 100 dataset dengan jumlah  $k=3$ . Dari proses tersebut mendapatkan hasil cluster. Cluster 1 berjumlah 35 data, Cluster 2 berjumlah 31 data, Cluster 3 berjumlah 34 data.

### **3.6 Evaluasi**

Evaluasi pengelompokan penyakit hipertensi menggunakan algoritma K-means Clustering dilakukan untuk mengetahui kinerja algoritma k-means dan hasil yang diperoleh dari hasil cluster tersebut.



## BAB IV PEMBAHASAN

### 4.1 Hasil Pengambilan Data

Data diperoleh dengan observasi lokasi penelitian dan wawancara serta observasi yang dilakukan di Puskesmas Atinggola Desa Monggupo Kecamatan Atinggola Kabupaten Gorontalo Utara. Setelah dilakukan wawancara dan observasi diperoleh masalah yang ada di lokasi penelitian yakni data pasien penyakit hipertensi belum dikelompokkan. Berikut data yang diperoleh dari lokasi penelitian :

**Tabel 4. 1** Pengumpulan Data

No	Nama	L/P	Umur	TDS (mmHg)	TDD (mmHg)	Berat Badan
1	Arman Nusi	L	58	170	80	56
2	Mei Hasan	P	52	160	90	48
3	Makmur Panigoro	L	72	160	100	68
4	Gustin Mahmud	P	51	160	90	62
5	Abd. Najib Blongkod	L	49	150	90	69
6	Yahya Gobel	L	45	180	90	54
7	Rostin Humagi	P	50	150	90	76
8	Dahlan Buheli	L	51	220	130	68
9	Nino Buheli	P	53	200	100	59
10	Arisna Dunggio	P	55	190	60	60
...	...	..	...	...	...	...
100	Jumria Opaladu	P	72	140	70	54

(Sumber : Puskesmas Atinggola, 2022)

Data yang digunakan pada penelitian ini adalah data Puskesmas Atinggola tahun 2022. Data tersebut meliputi nama, jenis kelamin, umur, tekanan darah sistolik (TDS), tekanan darah diastolic (TDD), berat badan dan factor keturunan. Dataset berjumlah 100 Data dan Data Training berjumlah 20 Data.

#### 4.2 Penjelasan Algoritma

K-means clustering merupakan metode data clustering non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih cluster. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster yanglain sehingga data yang berada dalam satu cluster/kelompok memiliki tingkat variasi yang kecil.

Clustering adalah proses pengelompokan kumpulan data menjadi beberapa kelompok sehingga objek di dalam satu kelompok memiliki banyak kesamaan dan memiliki banyak perbedaan dengan objek dikelompok lain. Perbedaan dan persamaannya biasanya berdasarkan nilai atribut dari objek tersebut dan dapat juga berupa perhitungan jarak.

#### 4.3 Perhitungan K-means

**Tabel 4. 2** Data Mining

No	Nama	L/P	Umur	TDS (mmHg)	TDD (mmHg)	Berat Badan
1	Arman Nusi	L	58	170	80	56
2	Mei Hasan	P	52	160	90	48
3	Makmur Panigoro	L	72	160	100	68
4	Gustin Mahmud	P	51	160	90	62
5	Abd. Najib Blongkod	L	49	150	90	69
6	Yahya Gobel	L	45	180	90	54

7	Rostin Humagi	P	50	150	90	76
8	Dahlan Buheli	L	51	220	130	68
9	Nino Buheli	P	53	200	100	59
10	Arisna Dunggjo	P	55	190	60	60
11	Makmur Panigoro	L	72	160	100	68
12	Fatma Ahaya	P	46	140	90	63
13	Rusdin Manti	L	70	150	90	50
14	Gustin Mahmud	P	51	160	90	62
15	Absan Kamaru	L	77	160	100	72
16	Abd. Najib Blongkod	L	49	150	90	69
17	Dasma Gani	P	52	190	90	82
18	Lula Moha	P	44	140	90	56
19	Miman Desei	L	25	160	90	60
20	Elma Kamaru	P	49	140	90	63

Untuk melakukan perhitungan Algoritma Kmeans, harus menentukan Centroid Awal. Proses penentuan Centroid Awal dilakukan dengan mencari 3 Data secara acak. Jumlah data centroid disesuaikan dengan jumlah  $k=3$ . Maka dari 20 Data Training dipilih 3 Data secara acak.

Penentuan Centroid  $y_1$  diambil dari variabel "Umur",  $y_2$  dari variabel "TDS",  $y_3$  dari variabel "TDD",  $y_4$  dari variabel "Berat Badan" seperti pada Tabel 4.3 berikut.

**Tabel 4. 3** Data Centroid Awal

Cluster	$y_1$	$y_2$	$y_3$	$y_4$
C1	35	180	90	80
C2	77	160	100	72
C3	49	140	90	63

Diketahui Rumus

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

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

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

Keterangan :

$x$  : Data Record (Data Training)

$y$  : Data Centroid

Cara perhitungan manual sebagai berikut:

$$\begin{aligned} C1 \quad (1) &= \sqrt{(48 - 35)^2 + (170 - 180)^2 + (100 - 90)^2 + (63 - 80)^2} \\ &= \sqrt{168 + 100 + 100 + 289} \\ &= \sqrt{657} \\ &= 25,632 \end{aligned}$$

$$\begin{aligned} (2) &= \sqrt{(57 - 35)^2 + (160 - 180)^2 + (90 - 90)^2 + (52 - 80)^2} \\ &= \sqrt{484 + 400 + 0 + 784} \\ &= \sqrt{1.668} \\ &= 40,841 \end{aligned}$$

$$\begin{aligned} (3) &= \sqrt{(47 - 35)^2 + (170 - 180)^2 + (110 - 90)^2 + (68 - 80)^2} \\ &= \sqrt{144 + 100 + 400 + 144} \\ &= \sqrt{788} \\ &= 28,071 \end{aligned}$$

$$\begin{aligned} (4) &= \sqrt{(68 - 35)^2 + (160 - 180)^2 + (90 - 90)^2 + (54 - 80)^2} \\ &= \sqrt{1089 + 400 + 0 + 676} \\ &= \sqrt{2165} \\ &= 46,529 \end{aligned}$$

$$\begin{aligned}
 (5) &= \sqrt{(58 - 35)^2 + (160 - 180)^2 + (80 - 90)^2 + (56 - 80)^2} \\
 &= \sqrt{529 + 400 + 100 + 576} \\
 &= \sqrt{1.605} \\
 &= 40,062
 \end{aligned}$$

$$\begin{aligned}
 \text{C2 (1)} &= \sqrt{(48 - 77)^2 + (170 - 160)^2 + (100 - 100)^2 + (63 - 72)^2} \\
 &= \sqrt{841 + 100 + 0 + 81} \\
 &= \sqrt{1022} \\
 &= 31,968
 \end{aligned}$$

$$\begin{aligned}
 (2) &= \sqrt{(57 - 77)^2 + (160 - 160)^2 + (90 - 100)^2 + (52 - 72)^2} \\
 &= \sqrt{400 + 0 + 100 + 400} \\
 &= \sqrt{900} \\
 &= 30
 \end{aligned}$$

$$\begin{aligned}
 (3) &= \sqrt{(47 - 77)^2 + (170 - 160)^2 + (110 - 100)^2 + (68 - 72)^2} \\
 &= \sqrt{900 + 100 + 100 + 16} \\
 &= \sqrt{1.116} \\
 &= 33,406
 \end{aligned}$$

$$\begin{aligned}
 (4) &= \sqrt{(68 - 77)^2 + (160 - 160)^2 + (90 - 100)^2 + (54 - 72)^2} \\
 &= \sqrt{81 + 0 + 100 + 324} \\
 &= \sqrt{475} \\
 &= 21,794
 \end{aligned}$$

$$\begin{aligned}
 (5) &= \sqrt{(58 - 77)^2 + (160 - 160)^2 + (80 - 100)^2 + (56 - 72)^2} \\
 &= \sqrt{361 + 0 + 400 + 256} \\
 &= \sqrt{1017} \\
 &= 31,890
 \end{aligned}$$

$$\begin{aligned}
 \text{C3 (1)} &= \sqrt{(48 - 49)^2 + (170 - 140)^2 + (100 - 90)^2 + (63 - 63)^2} \\
 &= \sqrt{1 + 900 + 100 + 0} \\
 &= \sqrt{1001} \\
 &= 31,638
 \end{aligned}$$

$$\begin{aligned}
 (2) &= \sqrt{(57 - 49)^2 + (170 - 140)^2 + (90 - 90)^2 + (52 - 63)^2} \\
 &= \sqrt{64 + 400 + 0 + 121} \\
 &= \sqrt{585} \\
 &= 24,186
 \end{aligned}$$

$$\begin{aligned}
 (3) &= \sqrt{(47 - 49)^2 + (170 - 140)^2 + (110 - 90)^2 + (68 - 63)^2} \\
 &= \sqrt{4 + 400 + 400 + 25} \\
 &= \sqrt{842} \\
 &= 29,017
 \end{aligned}$$

$$\begin{aligned}
 (4) &= \sqrt{(68 - 49)^2 + (160 - 140)^2 + (80 - 90)^2 + (56 - 63)^2} \\
 &= \sqrt{361 + 400 + 0 + 1} \\
 &= \sqrt{762} \\
 &= 27,604
 \end{aligned}$$

$$\begin{aligned}
 (5) &= \sqrt{(58 - 49)^2 + (160 - 140)^2 + (80 - 90)^2 + (56 - 63)^2} \\
 &= \sqrt{81 + 400 + 100 + 49} \\
 &= \sqrt{630} \\
 &= 25,099
 \end{aligned}$$

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

#### 4.4 Perhitungan Excel

##### 4.4.1 Iterasi 1

Selain perhitungan manual menggunakan rumus Eucliden Distance, perhitungan algoritma kmeans dilakukan di Ms.Excel 2016 juga menggunakan rumus Eucliden Distance.

Pada Tabel 4.4 berisi hasil perhitungan data training Iterasi Pertama. Pada tabel dapat dilihat dari 20 data training diperoleh hasil Cluster 1 ada 6 data yakni data ke- 1, 3, 7, 10, 17, 19. Cluster 2 ada 11 data yakni data ke- 2, 4, 5, 6, 8, 9, 11,

13, 15, 16, 18. Cluster 3 terdapat 3 data yakni data ke- 12, 14, 20 seperti pada tabel berikut.

**Tabel 4. 4 Iterasi 1**

Data ke- i	C1	C2	C3	CLUSTER
1	25,65151068	31,96873473	31,63858404	1
2	40,84115571	30	2,85312E+11	2
3	28,0713377	33,40658618	48828125	1
4	46,5295605	22,47220505	31381059609	2
5	40,06245125	31,89043744	1977326743	2
6	41,38840417	36,06937759	8,64976E+12	2
7	18,97366596	31,36877428	31381059609	1
8	56,78908346	28,86173938	6,42684E+13	2
9	37,22902094	15,16575089	7,43008E+11	2
10	0	48,24935233	3,42719E+13	1
11	44,86646855	6,403124237	48828125	2
12	44,83302354	39,26830783	3	3
13	55	27,07397274	1,79216E+12	2
14	31,30495168	29,59729717	20,1246118	3
15	48,24935233	0	31381059609	2
16	34,88552709	31,51190251	362797056	2
17	19,8242276	41,53311931	1,1649E+14	1
18	47,50789408	42,95346319	1977326743	2
19	30	54,29548784	177147,0028	1
20	45,66180023	36,9459064	0	3

#### 4.4.2 Iterasi 2

Setelah diketahui hasil cluster pada iterasi 1 kemudian ditentukan Centroid Baru berdasarkan cluster seperti pada tabel 4.5 berikut

**Tabel 4. 5 Centroid Baru Iterasi 2**

CENTROID BARU ITERASI 2				
<b>C1</b>	42,66666667	173,3333333	95	70,83333
<b>C2</b>	62,54545455	156,3636364	90,90909091	58,63636
<b>C3</b>	48,66666667	146,6666667	90	62,66667

Setelah mendapatkan Centroid baru, perhitungan diulangi menggunakan rumus yang sama tetapi menggunakan centroid yang baru. Hasil perhitungan Iterasi Ke-2 terdapat pada tabel 4.6.

**Tabel 4. 6** Iterasi 2

Data ke- i	C1	C2	C3	CLUSTER
1	11,22125958	22,34293203	25,3968502	1
2	27,62094616	9,425655161	19	2
3	16,21470526	29,66061473	31,23566338	1
4	33,58447062	8,08068813	25,03331114	2
5	29,29021452	12,64290217	20,23198787	2
6	28,48245074	15,43983257	20,09975124	2
7	8,808140174	23,42739796	25,13297966	1
8	43,7673775	18,84451386	34,3996124	2
9	27,65110004	17,13919678	25,10644008	2
10	14,56880228	42,12672735	39,97916124	1
11	32,7299557	16,52071077	29,16619047	2
12	34,76468515	23,6936369	7,187952884	3
13	41,83997291	13,09501198	25,03331114	2
14	18,71496727	12,59607705	13,55236757	2
15	37,18758753	21,98609628	34,17113792	2
16	24,75715385	18,22653592	7,16472842	3
17	22,68442931	42,3000889	47,5674959	1
18	36,84992085	24,88934187	10,51982256	3
19	25,14458192	37,75672454	27,29468813	1
20	35,17930263	21,70519989	6,683312552	3

Dari perhitungan Iterasi Ke-2 terdapat perubahan Cluster dimana Cluster I ada 6(enam) data yakni 1, 3, 7, 10, 17, 19. Cluster II ada 10(sepuluh) data yakni 2, 4, 5, 6, 8, 9, 11, 13, 14, 15. Cluster III ada 4(empat) data yakni 12, 16, 18, 20.

#### 4.4.3 Iterasi 3

Pada perhitungan Iterasi ke-2 mengalami perubahan data cluster maka perhitungan dilakukan kembali menggunakan rumus yang sama tetapi centroid baru berdasarkan hasil custer iterasi ke-2. Data centroid baru dapat dilihat pada tabel 4.7 berikut.



**Tabel 4. 7** Centroid Baru Iterasi 3

C1	42,66666667	173,3333333	95	70,83333
C2	64,6	159	91	58,2
C3	47	142,5	90	62,75

**Tabel 4. 8** Iterasi 3

Data ke- i	C1	C2	C3	CLUSTER
1	11,22126	22,37409216	29,27989925	1
2	27,62095	9,909591313	22,84321562	2
3	16,21471	29,79597288	34,40657641	1
4	33,58447	5,585696018	28,70213407	2
5	29,29021	13,0537351	23,93350162	2
6	28,48245	16,27267649	23,42674753	2
7	8,80814	23,57541092	29,08285577	1
8	43,76738	16,86416319	37,48749792	2
9	27,6511	16,91744662	28,58867783	2
10	14,5688	42,34855369	42,98618964	1
11	32,72996	15,2577849	32,5393992	2
12	34,76469	27,03701167	2,704163457	3
13	41,83997	13,35664629	27,34616061	2
14	18,71497	14,19154678	17,96698361	2
15	37,18759	20,64461189	37,30700336	2
16	24,75715	21,02379604	9,965565714	3
17	22,68443	41,07554017	51,49575225	1
18	36,84992	28,12827759	7,79823698	3
19	25,14458	39,66610644	28,24557487	1
20	35,1793	25,06790777	3,211308145	3

Hasil perhitungan Iterasi ke-3 memperoleh hasil cluster sama dengan hasil cluster pada iterasi ke-2. Dimana Cluster I ada 6(enam) data yakni 1, 3, 7, 10, 17, 19. Cluster II ada 10(sepuluh) data yakni 2, 4, 5, 6, 8, 9, 11, 13, 14, 15. Cluster III ada 4(empat) data yakni 12, 16, 18, 20.

#### 4.4.4 Hasil Perhitungan Data Training

Pada perhitungan 20 data training posisi cluster tidak berubah atau sama dengan iterasi ke-2, maka didapat hasil sebagai berikut:

Anggota Cluster I (C1) : {1, 3, 7, 10, 17, 19}= 6 Orang

Anggota Cluster II (C2) : {2, 4, 5, 6, 8, 9, 11, 13, 14, 15}= 10 Orang

Anggota Cluster III(C3) : {12, 16, 18, 20}= 4 Orang

Berdasarkan perhitungan Excel 20 Data Training, Kelompok Hipertensi Stage 1 terdapat pada cluster I dan terdiri dari 6 orang. Kelompok Pre-Hipertensi terdapat pada cluster II dan terdiri dari 10 orang. Kelompok Normal terdapat pada cluster III dan terdiri dari 4 orang.

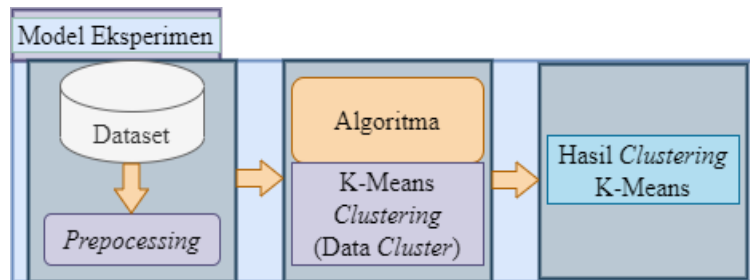
Data ke- i	C1	C2	C3	CLUSTER
1	11,22126	22,37409216	29,27989925	1
2	27,62095	9,909591313	22,84321562	2
3	16,21471	29,79597288	34,40657641	1
4	33,58447	5,585696018	28,70213407	2
5	29,29021	13,0537351	23,93350162	2
6	28,48245	16,27267649	23,42674753	2
7	8,80814	23,57541092	29,08285577	1
8	43,76738	16,86416319	37,48749792	2
9	27,6511	16,91744662	28,58867783	2
10	14,5688	42,34855369	42,98618964	1
11	32,72996	15,2577849	32,5393992	2
12	34,76469	27,03701167	2,704163457	3
13	41,83997	13,35664629	27,34616061	2
14	18,71497	14,19154678	17,96698361	2
15	37,18759	20,64461189	37,30700336	2
16	24,75715	21,02379604	9,965565714	3
17	22,68443	41,07554017	51,49575225	1
18	36,84992	28,12827759	7,79823698	3
19	25,14458	39,66610644	28,24557487	1
20	35,1793	25,06790777	3,211308145	3

**Tabel 4.9** Hasil Perhitungan Data Training

## BAB V

### PEMBAHASAN PENELITIAN

#### 5.1 Pembahasan Model



Gambar 5.1 Permodelan K-Means

Pemodelan K-Means sendiri dimulai dari tahap pengumpulan dataset oleh peneliti di lokasi penelitian, setelah pengumpulan data, dilakukan pemilihan data (seleksi). Setelah diperoleh data seleksi dilakukan pembersihan data yang tidak perlu kemudian dilakukan proses algoritma K-means Clustering. Proses algoritma K-means digunakan 100 dataset dengan jumlah  $k=3$ . Dari proses tersebut mendapatkan hasil cluster.

#### 5.2 Pembahasan Tools

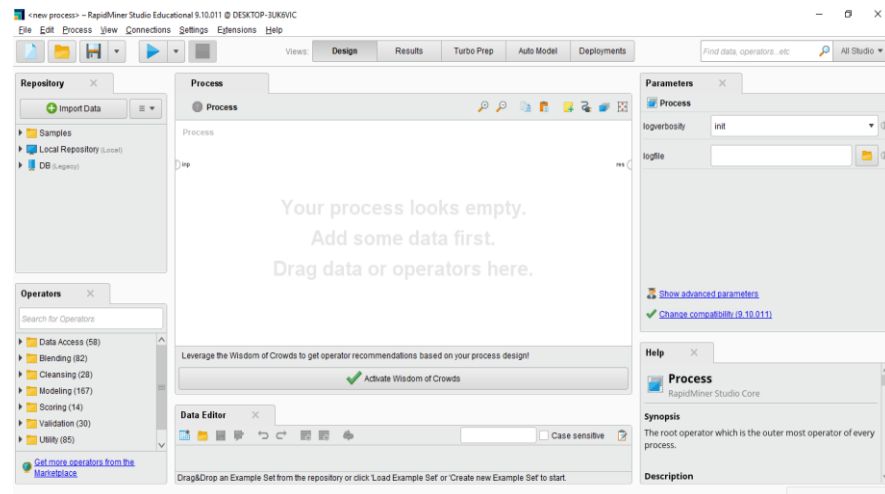
##### 5.2.1 Tampilan Data Excel Yang Akan Digunakan

NO	NAMA	UMUR (Tahun)	TEKANAN DARAH SISTOLIK	TEKANAN DARAH DIASTOLIK(TDD)	BERAT BADAN(KG)
1	AMIR MAHMUD	48	170	100	63
2	HADIA BLONGKOD	57	160	90	52
3	HARTONO DATAU	47	170	110	68
4	DAUD ABUNYE	68	160	90	54
5	ARMAN NUSI	58	160	80	56
6	MEI HASAN	52	160	90	48
7	TAJRI SUNGE	49	170	90	72
8	NORMA ISAI	75	160	90	45
9	RUSNI TANGAHU	66	160	90	75
10	YUNUS DESEI	35	180	90	80
11	MAKMUR PANIGORO	72	160	100	68
12	FATMA AHAYA	46	140	90	63
13	RUSDIN MANTI	70	150	90	50
14	GUSTIN MAHMUD	51	160	90	62
15	ABSAN KAMARU	77	160	100	72
16	ABD. NAJIB BLONGKOD	49	150	90	69
17	DASMA GANI	52	190	90	82
18	LULA MOHA	44	140	90	56
19	MIMAN DESEI	25	160	90	60
20	ELMA KAMARU	49	140	90	63

Gambar 5.2 Tampilan data Excel yang akan digunakan

Pada tampilan ini sediakan data yang akan digunakan dengan format Excel.

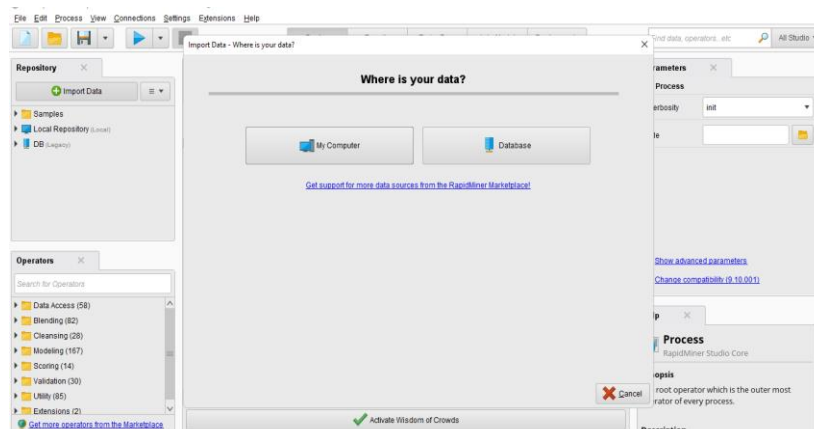
## 5.2.2 Tampilan utama Rapidminer



**Gambar 5.3** Tampilan Utama Rapidminer

Tampilan utama Rapidminer berisi berbagai view untuk mengoperasikan data diantaranya view Repository digunakan untuk penyimpanan data. Selain itu ada menu Operators yang berisi alat kerja Rapidminer. Kemudian ada view Parameter serta view Proses yang menunjukkan langkah-langkah tertentu dalam proses analisis. Berbagai view tersebut dapat dilihat pada Gambar 5.3 diatas.

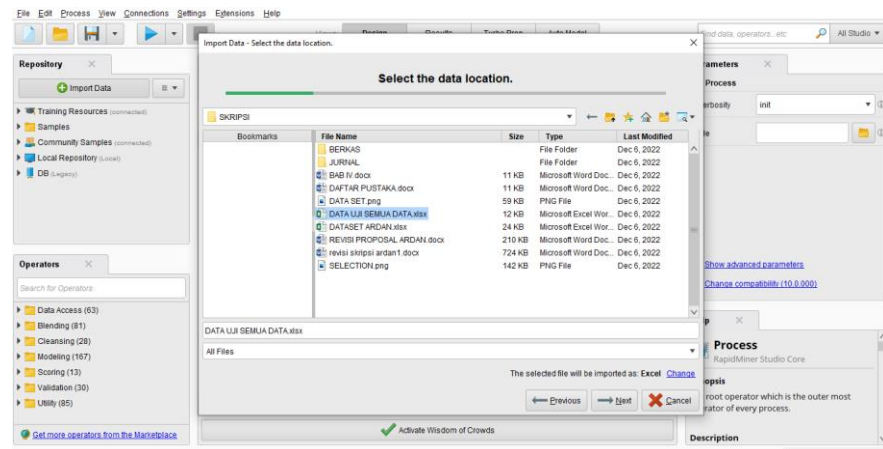
## 5.2.3 Import data



**Gambar 5.4 I Gambar 5.5** import Data

Pada gambar import data, klik my computer untuk menemukan di mana data sebelumnya sudah di siapkan untuk digunakan.

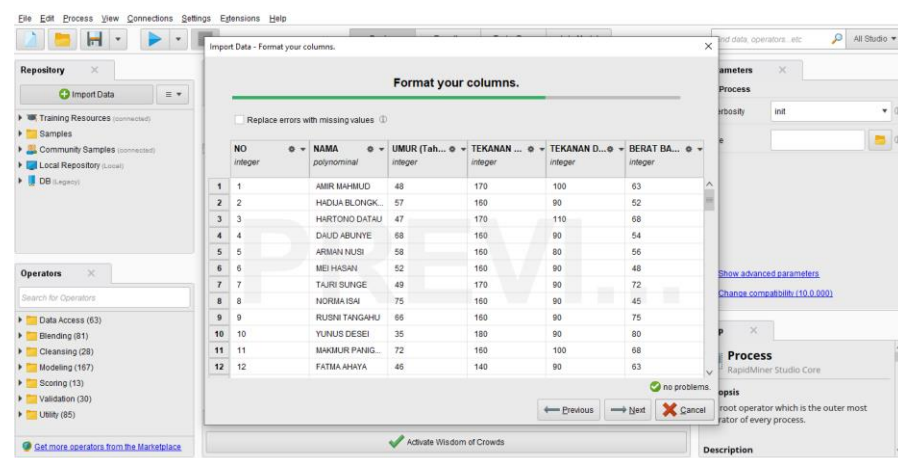
## 5.2.4 Lokasi penyimpanan data



Gambar 5.5 Lokasi penyimpanan data

Tampilan ini merupakan tampilan lokasi penyimpanan data yang akan digunakan.

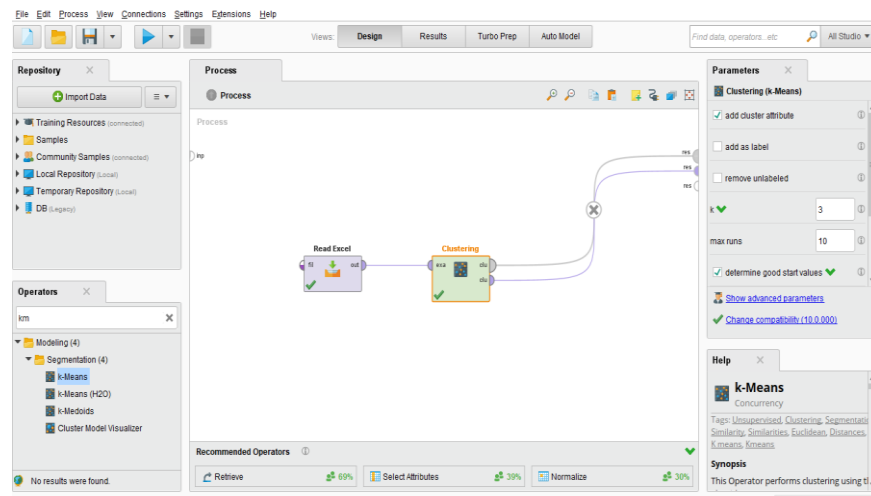
## 5.2.5 Data yang digunakan



Gambar 5.6 Data yang akan digunakan

Tampilan ini merupakan tampilan dari data yang akan digunakan, setelah selesai mengimport data yang akan digunakan kemudian tekan tombol finish untuk mengakhiri pengimporan data.

## 5.2.6 Tampilan Proses Pemasukan Data dan K-Means



**Gambar 5.7** Tampilan proses pemasukan data dan k-means

Tampilan ini merupakan tampilan proses pemasukan dan algoritma k-means yang akan digunakan, setelah memasukan data dan algoritma k-means, kemudian tentukan berapa jumlah cluster yang terletak di samping kiri tampilan. Jika kedua proses tersebut sudah dilakukan makan klik Run (segitiga biru) yang berada di bagian kiri atas tampilan.

## 5.2.7 Result Data

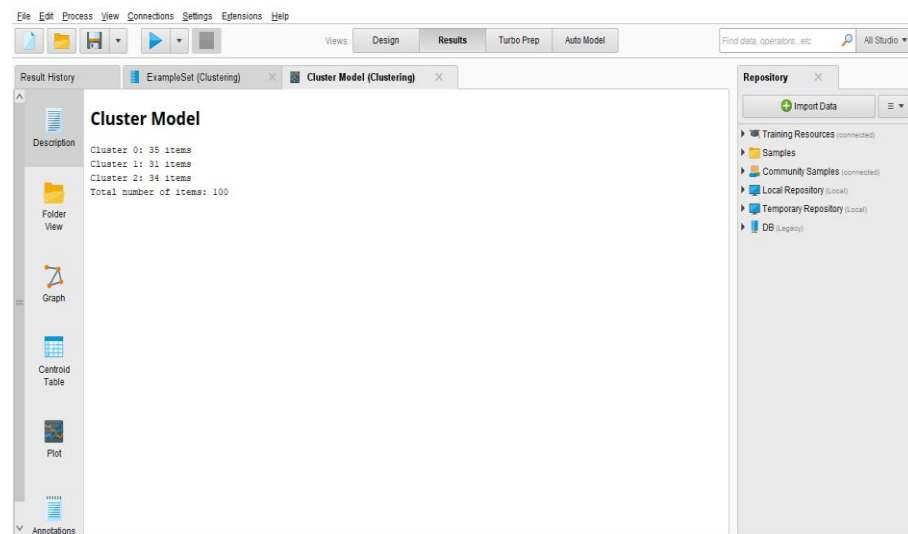
The screenshot shows the 'Result History' window in Orange3, displaying the output of the clustering process. The table below shows the data for 15 rows, with each row assigned to a specific cluster (cluster\_0 through cluster\_15).

Row No.	id	cluster	NO	NAMA	UMUR (Tahun)	TEKANAN D.	BERAT BAD...
1	1	cluster_0	1	AMIR MAHJUD	48	170	100
2	2	cluster_0	2	HADJIA BLO.	57	160	90
3	3	cluster_0	3	HARTONO D.	47	170	118
4	4	cluster_0	4	DAUD ABUNJ	68	160	90
5	5	cluster_0	5	ARMAN NUSI	58	160	80
6	6	cluster_0	6	MEI HASAN	52	160	90
7	7	cluster_0	7	TAJRI SUNJG	49	170	90
8	8	cluster_0	8	NCORMA ISAI	75	160	90
9	9	cluster_0	9	RUSDI TANG.	66	160	90
10	10	cluster_0	10	YURUS DESEI	35	180	90
11	11	cluster_0	11	MAKJUR PA.	72	160	100
12	12	cluster_0	12	FATMA AHAYA	46	140	90
13	13	cluster_0	13	RUSDIN MAH.	70	150	90
14	14	cluster_0	14	GUSTIN MAH.	51	160	90
15	15	cluster_0	15	ABSAH KAMA.	77	160	100

**Gambar 5.8** Tampilan Result Data

Tampilan ini merupakan tampilan result data dimana pada tampilan ini data yang telah di import sebelumnya sudah di *cluster*.

## 5.2.8 Tampilan Cluster Data Model



**Gambar 5.9** Tampilan *cluster* model

Pada tampilan ini data yang telah di *cluster* sebelumnya telah didapatkan hasil *cluster* berdasarkan jumlah yang telah di tentukan sebelumnya.

## 5.2.9 Pembahasan Algoritma

Pada penelitian ini algoritma yang digunakan ialah Algoritma K-Means, algoritma k-means dinilai cukup efektif dalam mengelompokkan data yang banyak. Pada penelitian ini data yang digunakan berjumlah 100 data yang di ambil dari lokasi penelitian, hasil dari pengelompokan menggunakan algoritma k-means adalah *cluster* 0 yang merupakan pasien Hipertensi Stage 1 terdiri dari 35 data, *cluster* 1 yang merupakan pasien Pre-Hipertensi terdiri dari 31, dan *cluster* 2 yang merupakan pasien normal terdiri dari 34 data. dengan hasil pengelompokan tersebut dapat dikatakan bahwa penggunaan algoritma k-means mampu mengelompokkan data secara efektif.

**Tabel 5. 1** Hasil *Cluster*

No	Nama	Umur (Tahun)	TDS	TDD	Berat Badan	Cluster
1	AMIR MAHMUD	48	170	100	63	cluster_0
2	HADIJA BLONGKOD	57	160	90	52	cluster_0
3	HARTONO DATAU	47	170	110	68	cluster_0
4	DAUD ABUNYE	68	160	90	54	cluster_0
5	ARMAN NUSI	58	160	80	56	cluster_0
6	MEI HASAN	52	160	90	48	cluster_0
7	TAJRI SUNGE	49	170	90	72	cluster_0
8	NORMA ISAI	75	160	90	45	cluster_0
9	RUSNI TANGAHU	66	160	90	75	cluster_0
10	YUNUS DESEI	35	180	90	80	cluster_0
11	MAKMUR PANIGORO	72	160	100	68	cluster_0
12	FATMA AHAYA	46	140	90	63	cluster_0
13	RUSDIN MANTI	70	150	90	50	cluster_0
14	GUSTIN MAHMUD	51	160	90	62	cluster_0
15	ABSAN KAMARU	77	160	100	72	cluster_0
16	ABD. NAJIB BLONGKOD	49	150	90	69	cluster_0
17	DASMA GANI	52	190	90	82	cluster_0
18	LULA MOHA	44	140	90	56	cluster_0
19	MIMAN DESEI	25	160	90	60	cluster_0
20	ELMA KAMARU	49	140	90	63	cluster_0
21	ROHANI ALHASNI	44	180	100	54	cluster_0
22	HASMIN HULALATA	59	140	90	64	cluster_0
23	DIANA MANTI	64	140	90	47	cluster_0
24	SARTIN TETEDULO	77	160	90	65	cluster_0
25	HAMIDA SALOTE	70	180	90	70	cluster_0
26	SULA ARTADI	73	160	100	58	cluster_0
27	YAHYA GOBEL	45	180	90	54	cluster_0
28	ZULAEHA KARIM	72	160	90	53	cluster_0
29	ANISA BUTOLO	45	140	90	63	cluster_0
30	ROSTIN HUMAGI	50	150	90	76	cluster_0
31	SARIA BAID	50	170	90	65	cluster_0
32	SARCI MOHA	45	180	90	53	cluster_0
33	KARSUM GOBEL	76	170	90	58	cluster_0
34	AHMAD V. GOBEL	65	160	90	69	cluster_0
35	MARNI SUNGE	64	170	90	57	cluster_0
36	YURNI BIMBING	68	140	90	59	cluster_2



37	ASNI GOBEL	51	150	90	74	cluster_2
38	RABI BLONGKOD	48	140	90	58	cluster_2
39	MINCE IBRAHIM	50	140	90	52	cluster_2
40	MIRNA SUNGE	40	140	90	63	cluster_2
41	SALHA DJAMALU	56	140	90	73	cluster_2
42	NURLAILA BAID	53	140	90	71	cluster_2
43	SARTIN BIMBING	44	180	90	65	cluster_2
44	SAFRUDIN LAMADI	54	160	100	59	cluster_2
45	IDHAM DAI	47	140	90	52	cluster_2
46	HERLINA ABAS	51	180	90	68	cluster_2
47	YAKUB BLONGKOD	67	150	80	84	cluster_2
48	UPIK LAMADI	60	140	90	86	cluster_2
49	SURYANI PUABENGGGA	39	150	90	80	cluster_2
50	SANUBANA POLAPA	75	160	90	48	cluster_2
51	DISMAN SALOTE	51	140	90	57	cluster_2
52	ASRIN HIKAYA	74	150	90	47	cluster_2
53	NETI DANGKUA	78	180	100	45	cluster_2
54	YANTI V. GOBEL	45	160	100	71	cluster_2
55	SATRIA ILOPONU	40	180	90	63	cluster_2
56	DAHLAN BUHELI	51	220	130	68	cluster_2
57	NINO BUHELI	53	200	100	59	cluster_2
58	TAHIR SUPU	53	160	90	57	cluster_2
59	ARISNA DUNGGIO	55	190	60	60	cluster_2
60	AHMAD GUGE	65	140	90	49	cluster_2
61	MINI PAKAYA	53	160	90	53	cluster_2
62	OLAN ILOPONU	41	160	100	58	cluster_2
63	HARDI LANTER	43	150	90	68	cluster_2
64	HENGKI MAMONTO	75	150	90	49	cluster_2
65	RAMNA PASILIA	50	150	90	56	cluster_2
66	KASMAD DUNGGIO	55	140	90	60	cluster_2
67	HAJIRA PATINGKI	64	170	100	52	cluster_2
68	ASTIN MALOHO	42	150	100	76	cluster_1
69	ULIN LANTER	73	150	90	52	cluster_2
70	SARCE TOMAYAHU	56	160	80	48	cluster_2
71	SARIPA YUNUS	51	140	90	62	cluster_1
72	KARTIN HADIA	58	140	90	57	cluster_1
73	NITA MADI	52	160	90	51	cluster_1
74	IBRAHIM HILOMALO	68	150	90	59	cluster_1
75	ZUN Aidin MOODUTO	27	170	100	79	cluster_1
76	MARNI HUSAIN	27	140	90	58	cluster_1
77	KONA ALI	67	190	100	61	cluster_1

78	ANGGRAINI MOODUTO	24	140	100	60	cluster_1
79	ESLA DAMA	47	170	90	55	cluster_1
80	MARWAN BOEYA	41	150	90	64	cluster_1
81	HASMIN KUN	49	150	90	59	cluster_1
82	DAKO PALIA	57	160	90	49	cluster_1
83	ELMI ILOPONU	49	170	90	63	cluster_1
84	SALEH PADJIRI	63	150	90	58	cluster_1
85	LISNA IYOHUN	36	160	90	70	cluster_1
86	ASRAWATI ILOPONU	36	140	90	57	cluster_1
87	HELDI BUHANG	53	160	90	59	cluster_1
88	MISRATI LAMADI	56	140	90	61	cluster_1
89	RASUNA YASIN	50	150	90	64	cluster_1
90	MURTIN ANTU	28	140	100	68	cluster_1
91	TENA BOEYA	61	170	100	57	cluster_1
92	YURNI BUKULU	52	150	90	58	cluster_1
93	SUMARNI NASARU	58	140	90	63	cluster_1
94	SEU NOI	59	140	90	62	cluster_1
95	MISDA DARISE	49	150	100	73	cluster_1
96	HASISA PARAMATA	58	140	90	65	cluster_1
97	SAPIA LAHABILA	62	150	90	72	cluster_1
98	ARIYATI GUGE	40	190	100	68	cluster_1
99	RAHIMUDIN	38	160	90	71	cluster_1
100	JUMRIA OPALADU	72	140	70	54	cluster_1

## **BAB VI**

### **PENUTUP**

#### **6.1 Kesimpulan**

Setelah dilakukan penelitian di Puskesmas Atinggola Kabupaten Gorontalo Utara dan pembahasan yang telah diuraikan sebelumnya, maka dapat diperoleh kesimpulan bahwa:

1. Hasil kinerja metode *k-means clustering* dalam pengelempokan data penyakit hipertensi menggunakan metode *k-means clustering* berjalan baik.
2. Hasil perhitungan 100 Data Training memperoleh hasil 3 Cluster. adalah *cluster 0* yang merupakan pasien Hipertensi Stage 1 terdiri dari 35 data, *cluster 1* yang merupakan pasien Pre-Hipertensi terdiri dari 31, dan *cluster 2* yang merupakan pasien Normal terdiri dari 34 data

#### **6.2 Saran**

Setelah dilakukan penelitian dan pengujian dengan aplikasi rapidminer, terdapat beberapa saran yang harus diperhatikan untuk memperoleh tujuan yang diharapkan sebagai berikut:

1. Untuk penelitian selanjutnya perlu ditambahkan data training serta variabel lain untuk mendapatkan hasil clustering yang lebih baik lagi.
2. Penulis mengharapkan agar dilakukan penambahan dataset agar hasil cluster lebih banyak.

## DAFTAR PUSTAKA

- [1] Y. Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [2] Y. Lasena, “CLUSTERING KOMODITI UNGGULAN MENGGUNAKAN ALGORITMA K-MEANS,” vol. 2, pp. 14–18, 2020.
- [3] Z. Nabila, A. Rahman Isnain, and Z. Abidin, “Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>.
- [4] A. Rahmat Dian Nugraha, K. Auliasari, and Y. Agus Pranoto, “IMPLEMENTASI METODE K-NEAREST NEIGHBOR (KNN) UNTUK SELEKSI CALON KARYAWAN BARU (Studi Kasus : BFI Finance Surabaya),” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 4, no. 2, pp. 14–20, 2020, doi: 10.36040/jati.v4i2.2656.
- [5] R. T. Wulandari, “Pengertian Data Mining,” *Data Min.*, vol. 7, no. 3, pp. 3–9, 2010.
- [6] Y. R. Kaesman, “Penentuan Penerima Beras Raskin di Kelurahan Oesapa Barat Menggunakan Metode K-Nearest Neighbor,” *Teknol. Terpadu*, vol. 2, no. 2, 2016.
- [7] Z. Arifin, “Penerapan Metode Knn (K-Nearest Neighbor) Dalam Sistem Pendukung Keputusan Penerimaan Kip (Kartu Indonesia Pintar) Di Desa Pandean Berbasis Web Dan Mysql,” *NJCA (Nusantara J. Comput. Its Appl.*, vol. 4, no. 1, 2019, doi: 10.36564/njca.v4i1.101.
- [8] Agus Gunawan, “Bab Ii Landasan Teori,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 8–24, 2019.
- [9] J. De Leeuw, “Journal of statistical software,” *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 1, no. 1, pp. 128–129, 2009, doi: 10.1002/wics.10.

- [10] W. Pratama and S. Kom, "KONSEP DATA MINING PERTEMUAN 2 - KDM Data Mining dan KDD."
- [11] W. Maharani, "Klasifikasi Data Menggunakan JST Backpropagation Momentum Dengan Adaptive Learning Rate," *Semin. Nas. Inform.*, vol. 1, no. semnasIF, pp. 25–31, 2009.
- [12] I. Irawan, "Pengembangan Sistem Informasi Tagihan (Billing System) Pasien Rawat Inap Pada Rumah Sakit Umum Daerah Bangkinang Kampar," *Riau J. Comput. Sci.*, vol. 4, no. 2, pp. 31–40, 2018.
- [13] N. Sastrawati, U. Islam, and N. Alauddin, "KONSUMTIVISME DAN STATUS SOSIAL," vol. 2, pp. 17–26, 2020.
- [14] S. D. A. Ambarwati, "Managing Productive Performance Appraisal: Sebuah Upaya Menjawab Kebutuhan Penilaian Kinerja Karyawan yang Bebas KKN," *J. Siasat Bisnis*, vol. 1, no. 7, pp. 93–111, 2002, doi: 10.20885/jsb.vol1.iss7.art6.
- [15] Teguh Hariyadi, Penerapan Algoritma *K-Means* Untuk Pengelompokan Data Nilai Siswa, 2013. Tinus Septiko, *Aplikasi K-Means Clustering Pada Data rumah tangga* 2013.

PAPER NAME

**SKRIPSI\_T3117241\_ARDAN DJAFAR.doc  
x**

AUTHOR

**T3117241 - ARDAN DJAFAR ardhandjafa  
r94@gmail.com**

WORD COUNT

**6336 Words**

CHARACTER COUNT

**36641 Characters**

PAGE COUNT

**40 Pages**

FILE SIZE

**919.6KB**

SUBMISSION DATE

**Dec 13, 2022 12:01 PM GMT+8**

REPORT DATE

**Dec 13, 2022 12:02 PM GMT+8**

### ● 21% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 21% Internet database
- 11% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

### ● Excluded from Similarity Report

- Bibliographic material
- Cited material
- Small Matches (Less than 20 words)

## ● 21% Overall Similarity

Top sources found in the following databases:

- 21% Internet database
- 11% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

### TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	<b>ejournal.akademitelkom.ac.id</b>	2%
	Internet	
2	<b>123dok.com</b>	2%
	Internet	
3	<b>repository.uin-suska.ac.id</b>	1%
	Internet	
4	<b>journal.universitaspahlawan.ac.id</b>	1%
	Internet	
5	<b>repository.uinsu.ac.id</b>	1%
	Internet	
6	<b>ejournal-s1.undip.ac.id</b>	1%
	Internet	
7	<b>ojs.trigunadharma.ac.id</b>	1%
	Internet	
8	<b>jim.teknokrat.ac.id</b>	1%
	Internet	

9	<b>prosiding.seminar-id.com</b> Internet	1%
10	<b>repository.unmuhjember.ac.id</b> Internet	1%
11	<b>repository.poltekkes-kdi.ac.id</b> Internet	<1%
12	<b>docplayer.info</b> Internet	<1%
13	<b>ojs.amikom.ac.id</b> Internet	<1%
14	<b>id.scribd.com</b> Internet	<1%
15	<b>jurnal.stmikroyal.ac.id</b> Internet	<1%
16	<b>conference.binadarma.ac.id</b> Internet	<1%
17	<b>jurnal.uisu.ac.id</b> Internet	<1%
18	<b>media.neliti.com</b> Internet	<1%
19	<b>repository.pelitabangsa.ac.id</b> Internet	<1%
20	<b>dcckotabumi.ac.id</b> Internet	<1%



21	<b>id.berita.yahoo.com</b> Internet	<1%
22	<b>coursehero.com</b> Internet	<1%
23	<b>e-journal.sari-mutiara.ac.id</b> Internet	<1%
24	<b>scribd.com</b> Internet	<1%



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI

**UNIVERSITAS ICHSAN GORONTALO**

**FAKULTAS ILMU KOMPUTER**

**SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001**

**Jl. Achmad Najamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo**

**SURAT REKOMENDASI BEBAS PLAGIASI**

**No. 168/FIKOM-UIG/R/V/2023**

Yang bertanda tangan di bawah ini :

Nama : Irvan Abraham Salihi, M.Kom  
NIDN : 0928028101  
Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini menerangkan bahwa :

Nama Mahasiswa : Ardan Djafar  
NIM : T3117241  
Program Studi : Teknik Informatika (S1)  
Fakultas : Fakultas Ilmu Komputer  
Judul Skripsi : Implementasi Data Mining Untuk Pengelompokan Penyakit Hipertensi Dengan Menggunakan Metode K-Means

Sesuai hasil pengecekan tingkat kemiripan skripsi melalui aplikasi **Turnitin** untuk judul skripsi di atas diperoleh hasil *Similarity* sebesar **21%**, berdasarkan Peraturan Rektor No. 32 Tahun 2019 tentang Pendeteksian Plagiat pada Setiap Karya Ilmiah di Lingkungan Universitas Ichsan Gorontalo dan persyaratan pemberian surat rekomendasi verifikasi calon wisudawan dari LLDIKTI Wil. XVI, bahwa batas kemiripan skripsi maksimal 30%, untuk itu skripsi tersebut di atas dinyatakan **BEBAS PLAGIASI** dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Mengetahui  
Dekan,  
  
**Irvan Abraham Salihi, M.Kom**  
NIDN. 0928028101

Gorontalo, 19 Mei 2023  
Tim Verifikasi,

  
**Zulfrianto Y. Lamasiqi, M.Kom**  
NIDN. 0914089101

Terlampir :  
Hasil Pengecekan Turnitin.



**DINAS KESEHATAN KABUPATEN GORONTALO UTARA**  
**PUSKESMAS ATINGGOLA**

*Desa Monggupo Kec. Atinggola Kab. Gorontalo Utara*



**SURAT KETERANGAN SELESAI MELAKSANAKAN PENELITIAN**

No. 094/PKM-ATG/2916/XII/2022

Yang bertanda tangan dibawah ini :

Nama : Meike Ibrahim, S.ST, M.Kes  
Nip : 19760528 200604 2 021  
Jabatan : Kepala Puskesmas

Dengan ini menerangkan bahwa mahasiswa yang beridentitas :

Nama : Ardan Djafar  
Nim : T3117241  
Fakultas : Ilmu Komputer  
Jurusan : Teknik Informatika  
Universitas : Universitas Ichsan Gorontalo

Bahwa benar telah melakukan penelitian diwilayah kerja Puskesmas Atinggola untuk menyusun Skripsi dengan Judul ***“Implementasi Data Mining Untuk Pengelompokan Penyakit Hypertensi Dengan Menggunakan Metode K-Means”***

Demikian surat keterangan ini dibuat, agar dapat dipergunakan sebagai mana mestinya

Atinggola, 08 Desember 2022  
Kepala Puskesmas Atinggola



Meike Ibrahim, S.ST, M.Kes  
Nip. 19760528 200604 2 021

## DAFTAR RIWAYAT HIDUP



Nama : ARDAN DJAFAR  
Nim : T3117214  
Tempat & Tgl Lahir : Atinggola, 09 Oktober 1993  
Jenis Kelamin : Laki-laki  
Agama : Islam  
Email : [ardhandjafar94@gmail.com](mailto:ardhandjafar94@gmail.com)

### Riwayat Pendidikan :

1. Tahun 2006 Telah Menyelesaikan Pendidikan Di Sekolah Dasar Negeri 1 Ilomata,
2. Tahun 2010 Telah menyelesaikan Pendidikan Di SMP Negeri 4 Atinggola
3. Tahun 2013 Telah Menyelesaiakan Pendidikan Di SMKN 1 Atinggola
4. Tahun 2017 Telah Diterima Menjadi Mahasiswa Di Perguruan Tinggi Universitas Ichsan Gorontalo.