



**PENGELOMPOKAN PENYEBARAN PENYAKIT TUBERCULOSIS
DI KABUPATEN ACEH TAMIANG MENGGUNAKAN
METODE K-MEDOIDS**

PROPOSAL SKRIPSI

**Disusun Sebagai Syarat Memperoleh Gelar Sarjana Komputer
Prodi Teknik Informatika Fakultas Teknik
Universitas Malikussaleh**

DISUSUN OLEH :

**NAMA : RANIA SOFIA SALSABILA HARAHAP
NIM : 200170252
PRODI : TEKNIK INFORMATIKA**

**JURUSAN INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MALIKUSSALEH
LHOKSEUMAWE**

2024

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan kesehatan, rahmat dan hidayahNya, sehingga penulis dapat menyelesaikan proposal skripsi ini. Sholawat dan salam penulis sanjung sajian kepada Baginda Nabi Besar Rasulullah Muhammad SAW yang telah mengantarkan kita kehidupan dari jaman jahiliyah menuju jaman yang kaya akan ilmu pengetahuan. Adapun judul proposal skripsi yang dibahas yaitu **“Pengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang Menggunakan Metode K-Medoids”**.

Proposal skripsi ini disusun sebagai salah satu syarat dalam menyelesaikan dan memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika Universitas Malikussaleh. Tak lupa ucapan terima kasih kepada kedua orang tua yang selalu memberikan dukungan dan mendoakan penulis dalam menyelesaikan penulisan proposal skripsi ini agar segera cepat terselesaikan dan segera maju naik seminar proposal.

Dalam menyelesaikan proposal skripsi ini, banyak pihak yang telah memberikan bantuan secara nasihat, semangat, dan materil. Dan kesempatan kali ini penulis ingin mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Herman Fithra, S.T., M.T., IPM., ASEAN.Eng selaku Rektor Universitas Malikussaleh.
2. Bapak Dr. Muhammad Daud, S.T., M.T selaku Dekan Teknik Universitas Malikussaleh.
3. Bapak Munirul Ula, S.T., M.Eng.,Ph.D selaku Ketua Jurusan Teknik Informatika.
4. Ibu Zara Yunizar, S.T., M.Kom selaku Ketua Prodi Teknik Informatika.
5. Ibu Eva Darnilla, S.T., M.T selaku Dosen Pembimbing Utama yang banyak membantu dalam memberikan arahan dan nasihat yang berguna dalam penyusunan proposal skripsi ini.
6. Ibu Yesy Afrillia, S.T., M.Kom selaku Dosen Pembimbing Pendamping yang banyak membantu dalam memberikan arahan dan

nasihat yang berguna dalam penyusunan proposal skripsi ini.

7. Ibu Cut Agusniar, S.T., M.Cs selaku dosen penguji I dan bapak Rizal, S.Si., M.IT selaku dosen penguji II.
8. Seluruh dosen jurusan Teknik Informatika Universitas Malukuusaleh yang telah memberikan ilmu yang sangat bermanfaat bagi penulis.
9. Terimakasih kepada kedua orang tua penulis, keluarga yang memberikan semangat dan doa sehingga penulis dapat menyelesaikan proposal skripsi ini.
10. Terimakasih juga kepada teman-teman kos penulis Rahmita Ayu, Cheltra, Nurul Khairunnisa, Nurul Hidayah, Natasya, Mahzura, Syafrida, yang setia memberikan semangat dan saling mendukung di perantauan.

Akhir kata, terimakasih penulis kepada semua pihak yang telah membantu dalam penyusunan skripsi ini dan mohon maaf jika ada kekurangan. Kritik dan saran dibutuhkan dalam penyempurnaan penyusunan ini dan semoga proposal skripsi ini dapat memberikan manfaat untuk pengembangan wawasan dan peningkatan ilmu pengetahuan bagi pembaca.

Lhokseumawe, 26 September 2023

Penulis,

RANIA SOFIA SALSABILA

NIM. 200170252

DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR TABEL	v
DAFTAR GAMBAR.....	viz
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA	8
2.1 Tuberculosis (TBC)	8
2.1.1 Definisi Tuberculosis (TBC).....	8
2.1.2 Penyebab Penyakit Tuberculosis (TBC)	9
2.1.3 Gejala Penyakit Tuberculosis (TBC)	9
2.2 Data Mining.....	10
2.2.1 Definisi Data mining.....	10
2.2.2 Metode Data Mining	10
2.3 Clustering	12
2.4 Algoritma K-Medoids	12
2.4.1 Definisi Algoritma K-Medoids	12
2.5 Flowchart.....	13
2.6 PHP	15
2.7 MySQL	15
2.8 XAMPP	15
2.9 Penelitian Terdahulu	15
BAB III METODE PENELITIAN	21
3.1 Tempat dan Waktu Pelaksanaan Penelitian	21
3.2 Alat dan Bahan Penelitian	21
3.2.1 Perangkat Keras (Hardware).....	21
3.2.2 Perangkat Lunak (Software)	21

3.3 Metode Pengumpulan Data	21
3.3.1 Data Primer	22
3.3.2 Data Sekunder	22
3.4 Analisa Kebutuhan Sistem	22
3.5 Alur Penelitian.....	22
3.6 Skema dan Perancangan Sistem	25
DAFTAR PUSTAKA	45

DAFTAR TABEL

Tabel 2.1 Simbol Flowchart	14
Tabel 2.2 Penelitian Terdahulu	14

DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	19
Gambar 3.2 Skema Sistem	21

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini penyakit Tuberculosis menjadi salah satu problematika kesehatan masyarakat yang sangat serius yang harus diperhatikan, karena penyakit tersebut dikategorikan dapat menyebar dan menular. Penyakit ini ditularkan melalui udara yang tercemar oleh bakteri *Mycrobacterium Tuberculosis* yang bercampur dengan pencemaran udara akibat polusi yang dapat menyerang paru-paru. Bakteri ini dikategorikan sebagai bakteri basil yang sangat kuat sehingga memerlukan waktu yang lama untuk mengobatinya atau berpotensi besar membunuh dan mematikan pengidap TBC. Penyakit TBC masih terus menjadi penyebab dari 10 daftar kematian tertinggi di dunia, angka kematian akibat penyakit ini secara keseluruhan mencapai 1,3 jiwa. Sudah ditemukannya Obat Anti Tuberculosis (OAT) dan *Vaksinasi Bacillus Calmette Guerin (BCG)* dan telah dilakukan, tetapi TBC masih belum bisa di berantas.

Secara global negara Indonesia berada pada posisi ketiga dengan jumlah penyakit Tuberculosis terbanyak di dunia. TBC di Indonesia menurut WHO (World Health Organization) pada tahun 2020 terdapat 9,9 juta orang mengidap penyakit ini dan 1,5 juta meninggal dunia. Kasus ini di Indonesia hampir tersebar disemua daerah, salah satunya yaitu Kabupaten Aceh Tamiang. Dinas Kesehatan Kabupaten Aceh Tamiang mencatat jumlah Tuberculosis tahun 2019 sampai 2020 kasus penyakit TBC di Kabupaten Aceh Tamiang sebanyak 610 kasus. Data itu belum termasuk jumlah data yang masuk ke Dinas Kesehatan dari Faskes (Fasilitas Kesehatan) swasta yang ada di Kabupaten Aceh Tamiang. Jika semua faskes yang ada melaporkan kasus itu ke dinas, maka data TBC ini bisa lebih (Fahzian Aldevan., 2020). Dalam beberapa tahun terakhir di Provinsi Aceh, Indonesia, layanan kesehatan telah menarik banyak perhatian mencari lebih banyak analisis data dalam layanan kesehatan untuk meringankan masalah medis pada staf medis, populasi, orang yang hidup sendiri, dan kualitas hidup. Penambangan dan peramalan data memainkan peran penting dalam bidang sosial dan medis modern (Darnila et al.,

2019). Sementara itu, pengolahan data di Dinas Kabupaten Aceh Tamiang belum bisa mengelompokkan penyebaran penyakit TBC, karena masih menggunakan cara manual dengan menggunakan excel.

Dengan adanya permasalahan tersebut, perlu adanya suatu upaya penanganan yang efektif yaitu dengan mengelompokkan daerah penyebaran penyakit Tuberculosis di Kabupaten Aceh tamiang menggunakan data mining dengan metode K-Medoids Clustering. Cara kerja dari teknik clustering yaitu populasi atau beberapa data yang kita punya dikelompokkan di dalam beberapa cluster sehingga titik data dalam kelompok yang sama lebih mirip dengan titik data lain dalam kelompok yang sama daripada yang ada di kelompok lain, dengan memanfaatkan kemajuan pengetahuan dan teknologi di era milenial ini yang berkembang sangat pesat.

Kemajuan pengetahuan teknologi beserta aplikasinya dalam segala bidang tidak bisa lepas dari perangkat komputer. Kebutuhan akan layanan informasi sangatlah penting khususnya pada bidang kesehatan. Salah satunya adalah aplikasi yang menggunakan data mining untuk mempermudah dalam proses analisis dalam penentuan pola penyebaran penyakit Tuberculosis . Pengelompokan ini bertujuan untuk mengetahui daerah mana saja yang memiliki tingkat penyakit Tuberculosis rendah, sedang hingga daerah yang paling tinggi. Hasil penelitian ini diharapkan dapat menjadi bahan rekomendasi untuk pemerintah Kabupaten Aceh Tamiang dalam melakukan penanganan yang lebih efektif dengan memberikan obat-obatan terhadap penyakit Tuberculosis. Pemerintah nantinya dapat lebih fokus terlebih dahulu ke daerah-daerah yang memiliki kasus penyakit Tuberculosis tingkat tinggi, sehingga kasus Tuberculosis di Kabupaten Aceh Tamiang semakin berkurang.

Berdasarkan penelitian sebelumnya terkait tentang pengelompokan penyebaran penyakit tuberculosis, yaitu : Penelitian oleh Dini Marlina, Nurelina Fauzer Putri, Andri Fernando, dan Aditya Ramadhan "Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokan Wilayah Sebaran Cacat pada Anak" Metode yang di gunakan untuk penelitian ini adalah metode K-Means dan K-Medoids. Penelitian ini bertujuan untuk pengelompokan pada data sebaran anak cacat yang ada pada Provinsi Riau yang dapat memberikan solusi bagi pemerintah

untuk memudahkan dalam mengetahui wilayah-wilayah yang memiliki tingkat anak cacat yang tinggi sehingga dapat dilakukan penanggulangan secara dini. Klaster yang dihasilkan pada penelitian ini menggunakan metode K-Means dan K-Medoids adalah berjumlah tiga klaster. Validitas yang digunakan pada penelitian ini adalah validitas Silhoutte Coefficient Adapun nilai validitas yang dihasilkan pada algoritma K-Medoids adalah sebesar 0.5009. Sedangkan nilai validitas yang dihasilkan pada algoritma K-Means adalah 0.1443. Hal ini menunjukkan bahwa algoritma K-Medoids lebih baik dalam melakukan pengelompokan pada data sebaran Anak Cacat dibandingkan dengan algoritma K-Means (Marlina et al., 2018).

Penelitian oleh Tri Juninda, Mustakim, dan Elvia Andri "Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau" Metode yang di gunakan adalah algoritma K-Medoids. Penelitian ini untuk pengelompokan penyakit di Pekanbaru Riau, dengan memiliki empat atribut yaitu kecamatan, sakit yang sering diderita, usia dan jenis kelamin. Dari data tersebut akan ditentukan cluster terbaik menggunakan validitas Devies Bouildien Indeks (DBI). Penelitian ini diharapkan mampu membantu pemerintah serta instansi terkait tentang pola penyakit yang sering diderita masyarakat Pekanbaru berdasarkan cluster terbaik menggunakan algoritma K-Medoids. Didapatkan output berupa sistem berbasis web dalam pengaplikasiannya. Adapun hasil analisa klastering yang didapatkan adalah 4 cluster sebagai pengelompokan terbaik dengan nilai Devies Bouildien Indeks sebesar 0,043. Pada cluster 1 didapatkan 420 record dengan penyakit dominan adalah Maag sebesar 44,39%, cluster 2 didapatkan 349 record dengan penyakit dominan adalah Diare dan Sakit Perut sebesar 16,98%, pada cluster 3 didapatkan 794 record dengan penyakit dominan adalah Batuk dan Pilek sebesar 65,21% dan pada cluster 4 didapatkan 1248 record dengan penyakit dominan adalah Batuk dan Pilek sebesar 54,10%. Hasil dari penelitian ini menunjukkan bahawa algoritma K-Medoids mampu melakukan pengelompokan terhadap penyakit di Pekanbaru Riau (Juninda et al., 2019).

Penelitian selanjutnya oleh Astrika Praokta Fialine, Daffa Alya Alodia, Desy Endriani, dan Edy Widodo "Implementasi Metode K-Medoids Clustering

untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan" Metode yang digunakan adalah metode K-Medoids. Penelitian ini memiliki tujuan untuk Pemerataan program wajib belajar. Data yang digunakan diperoleh dari publikasi Statistik Pendidikan tahun ajaran 2019/2020 oleh Badan Pusat Statistik dengan 8 variabel indikator pendidikan SMA/MA sederajat menurut provinsi di Indonesia, yaitu Angka Partisipasi Kasar (APK), Angka Partisipasi Murni (APM), Angka Partisipasi Sekolah (APS), persentase guru layak, Rata-Rata Lama Sekolah (RLS), Angka Melanjutkan (AM), tingkat menyelesaikan sekolah (SMA), dan angka putus sekolah. Hasil penelitian ini diperoleh tiga cluster. Cluster pertama terdapat 6 provinsi dengan indikator pendidikan tinggi, cluster kedua terdapat 13 provinsi dengan indikator pendidikan sedang, dan cluster ketiga terdapat 15 provinsi dengan indikator pendidikan rendah. Berdasarkan hasil analisis data indikator Pendidikan di Indonesia tahun ajaran 2019/2020 diperoleh kesimpulan bahwa gambaran umum provinsi dengan nilai persentase indikator pendidikan paling tinggi adalah AM dengan rata-rata 90,09%. Namun jika dilihat dari nilai maksimum yang paling tinggi adalah indikator APK dengan nilai persentase sebesar 98,31% berada pada provinsi Kalimantan Utara. Sedangkan untuk indikator pendidikan paling rendah adalah APtS dengan persentase 0.18% berada pada provinsi Yogyakarta (Alodia et al., 2021).

Penelitian oleh Fina Nasari, Dahriani Hakim Tanjung, dan Fitri Handayanic "Optimasi Metode K-Means dan K-Medoids Berdasarkan Jumlah Cluster dan Nilai DBI Dalam Pengelompokkan Produksi Kelapa Sawit Di Provinsi Riau" dengan menggunakan metode K-Means dan K-Medoids. Pada penelitian ini Provinsi riau menjadi salah satu provinsi dengan produksi kelapa sawit tertinggi di Indonesia dengan luas kebun 2.8 juta Ha produksi 8.8 juta ton per tahun. Penyebaran kelapa sawit di provinsi riau hampir diseluruh kabupaten, sehingga perlu adanya pengelompokkan daerah produksi kelapa sawit. Clustering menjadi salah satu metode yang dapat mengelompokkan data pada data yang sejenis. Berdasarkan data Direktorat Perusahaan Perkebunan Kelapa Sawit Provinsi Riau tahun 2020 terdapat 219 Perusahaan Besar Swasta, 19 Perusahaan Besar Negara yang tersebar. Hasil penelitian ini menjelaskan bahwa kombinasi kedua algoritma

ini berfungsi dengan sangat baik, terdapat 72% produksi kelapa sawit dalam Cluster rendah, 20% dalam cluster sedang dan 2% dalam cluster tinggi. Jumlah Cluster yang digunakan dalam pengujian ini adalah 2, 3 dan 5. Hasil penelitian ini menunjukkan Jumlah Cluster 2 menjadi cluster terbaik dengan nilai DBI untuk metode K-Medoid - 159796492242,667 dan metode K-Means -82338884292,014. Metode K-Medoid menjadi metode cluster terbaik dengan cluster yang dihasilkan berupa 7 kabupaten pada kelompok jumlah produksi Tinggi dan 5 kabupaten masuk pada kelompok jumlah produksi rendah (Nasari et al., 2023) .

Atas dasar uraian diatas, sehingga penulis tertarik untuk mengangkatnya sebagai tugas akhir dengan membuat judul “ **Pengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang Menggunakan Metode K-Medoids**”

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka permasalahan yang penulis rumuskan adalah:

1. Bagaimana membangun sebuah sistem yang dapat digunakan untuk Pengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang menggunakan metode K-Medoids?
2. Bagaimana cara kerja algoritma K-Medoids untuk mengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang ?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut :

1. Data set yang digunakan pada penelitian ini diperoleh dari Dinas Kesehatan Kabupaten Aceh Tamiang.
2. Data set yang digunakan pada penelitian ini adalah data pasien penyakit Tuberculosis pada tahun 2019,2020,2021,2022 dan 2023.
3. Atribut yang digunakan pada penelitian ini yaitu jenis kelamin, umur, alamat, Penyakit Tuberculosis, Tipe Diagnosis TBC, dan Klasifikasi Berdsarkan Lokasi Anatomi.

4. Metode yang digunakan dalam penelitian ini untuk Pengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang adalah menggunakan metode K-Medoids.
5. *Output* yang dihasilkan adalah Pengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang dalam bentuk website yang menampilkan hasil berbentuk peta.
6. Daerah Penyebaran Penyakit Tuberculosis terdapat di Kabupaten Aceh Tamiang yang terdiri dari 12 kecamatan yaitu Banda Mulia, Banda Pusaka, Bendahara, Karang Baru, Kejuruan Muda, Kota Kuala Simpang, Manyak Payed, Rantau, Sekerak, Seruway, Tamiang Hulu dan Teunggulun.
7. Data yang dikelompokkan menjadi 3 cluster, yaitu cluster rendah, cluster sedang, dan cluster tinggi.
8. Sistem yang dibangun berbasis web dengan menggunakan bahasa pemrograman *Hypertext Preprocessor* (PHP) dengan database MYSQL.

1.4 Tujuan Penelitian

Adapun tujuan penelitian yang ingin dicapai adalah :

1. Membangun sebuah sistem untuk mengelompokkan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang menggunakan metode K-Medoids.
2. Sistem ini diharapkan dapat membantu pemerintah Kabupaten Aceh Tamiang dalam mengatasi penyakit Tuberculosis berdasarkan wilayah penyebarannya.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberi manfaat. Adapun manfaat dari penelitian ini adalah :

1. Sebagai referensi bagi pembaca dan sebagai tambahan pengetahuan tentang Pengelompokkan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang menggunakan metode K-Medoids.
2. Memperluas pengetahuan tentang metode K-Medoids.
3. Sebagai sarana untuk penerapan ilmu yang diperoleh selama studi,

khususnya dalam perancangan dan pembuatan sistem berbasis web.

4. Membantu Pemerintah Kabupaten Aceh Tamiang dalam memberikan obat-obatan untuk masyarakat yang terkena penyakit Tuberculosis.

BAB II

TINJAUAN PUSTAKA

2.1 Tuberculosis (TBC)

2.1.1 Definisi Tuberculosis (TBC)

Tuberculosis atau yang sering dikenal dengan TBC merupakan salah satu penyakit bisa menular yang menyerang paru-paru, disebabkan oleh bakteri *Mycobacterium Tuberculosis* yang ditandai dengan jaringan Granulasi nekrotik (Perkijauan) sebagai respon terhadap kuman tersebut (Sejati & Sofiana, 2015). Paru-paru adalah organ penting dalam tubuh yang peran utamanya yaitu untuk mendukung sistem pernafasan. Ketika kita menarik napas, kita akan memasukkan oksigen (O_2 = zat asam) ke dalam paru-paru, dan pada waktu mengeluarkan napas kita akan mengeluarkan karbon dioksida (CO_2 = zat asam arang) dari paru-paru. Tingkat polusi udara yang semakin tinggi memicu timbulnya berbagai penyakit paru. Asap industri pabrik ataupun berbagai asap lainnya, apabila banyak terhirup oleh manusia dapat mengganggu fungsi paru serta kebiasaan merokok juga memicu timbulnya penyakit kanker paru (Yunus & Setyowibowo, 2011).

Tuberculosis (TBC) merupakan infeksi yang sangat umum dan berpotensi besar untuk membunuh atau mematikan dalam sebagian kasus nya. Diperkirakan seorang penderita Tuberculosis pada 1 dari 10 orang penderitanya. Tuberkulosis merupakan penyakit menular langsung dan kronis yang ditularkan oleh penderita TBC, sehingga penderita TBC harus segera diobati hingga sembuh (Purba et al., 2019). Pengidap Tuberculosis mengalami berbagai macam gejala yaitu demam, kurus dan batuk. Meskipun pengobatan yang berhasil mencegah kematian, banyak penderita Tuberculosis yang mengalami masalah kesehatan yang berkelanjutan setelah episode penyakit, dan ada peningkatan bukti kecacatan jangka panjang dan peningkatan resiko kematian pada populasi ini. TBC penyakit yang mengganggu sumber daya manusia dan umumnya menyerang kelompok masyarakat dengan golongan sosial ekonomi rendah.

2.1.2 Penyebab Penyakit Tuberculosis (TBC)

Terdapat 5 bakteri yang berkaitan erat dengan infeksi *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Mycobacterium africanum*, *Mycobacterium microti* dan *Mycobacterium canettii.p. Mycobacterium tuberculosis* hingga saat ini merupakan bakteri yang paling sering ditemukan, dan menular antar manusia ke manusia melalui rute udara. Penyakit ini menular dengan cepat pada orang yang rentan dan daya tahan tubuh yang lemah. Penularan Tuberculosis cukup mudah, yaitu saat orang yang terpapar dan tidak melakukan pengobatan menyebarkan ke udara melalui batuk, bersin berbicara, tertawa hingga meludah (Yunus & Setyowibowo, 2011).

2.1.3 Gejala Penyakit Tuberculosis (TBC)

Gejala Tuberculosis sangat bervariasi tergantung bagian tubuh mana yang terkena. Penyakit TBC biasanya berkembang perlahan dan perlu waktu beberapa minggu untuk menyadari ada infeksi tersebut. Infeksi TBC yang tidak menimbulkan gejala apapun disebut TBC Laten. Sedangkan infeksi yang menimbulkan gejala disebut TBC aktif. Gejala biasanya berhubungan dengan sistem pernafasan yang dapat memengaruhi bagian tubuh lainnya, tergantung dimana bakteri TBC tumbuh. Gejala umum bagi penderita penyakit TBC diantaranya adalah (Suherni & Maduratna, 2013) :

1. Batuk berdarah
2. Nyeri di dada
3. Sesak nafas
4. Batuk yang berlangsung lama sekitar 3 minggu lebih
5. Berat badan turun drastis
6. Kehilangan nafsu makan
7. Demam dan menggigil
8. Berkeringat pada malam hari

2.2 Data Mining

2.2.1. Definisi Data mining

Data mining merupakan teknik statistik, matematika, kecerdasan buatan, dan machine learning yang mengekstraksi serta mengidentifikasi informasi yang bermanfaat terakut dari sebuah database yang besar (Ali, 2019). Informasi dari database yang digunakan untuk membuat keputusan bisnis yang sangat penting dimana data mining merupakan proses menemukan pola pada data. Pada dasarnya data mining adalah ilmu disiplin yang berusaha menemukan, menyelidiki serta mengekstraksi informasi dari data yang sudah ada.

2.2.2. Metode Data Mining

Data mining memiliki tujuh metode yaitu (Jolly et al, 2020) :

1. *Association* merupakan aturan yang digunakan untuk mengidentifikasi serta mengkorelasi satu sumber data. Asosiasi mengkorelasi produk didalam keranjang ketika menganalisis pola pembelian klien. Sehingga, bisnis dapat membuat strategi penjualan yang lebih baik.
2. *Classification* merupakan proses dalam menemukan kesamaan karakteristik suatu kelompok atau kategori (kelas). Salah satu metode yang paling umum adalah klasifikasi. Mengevaluasi kelas dari suatu objek yang identifiernya tidak diketahui merupakan tujuan dari metode klasifikasi.
3. *Regression* merupakan suatu teknik yang mencari pola numerik daripada kelas. Fungsi yang berfungsi untuk memprediksi hasil nilai input merupakan output dari metode regresi.
4. *Clustering* merupakan metode yang digunakan untuk membagi suatu kumpulan data menjadi banyak kelompok berdasarkan mirip karakteristiknya.
5. *Forecasting* merupakan teknik dalam memperkirakan suatu nilai yang akan dicapai dari waktu ke waktu. Dalam menggunakan metode ini, data dan nilai dari informasi sebelumnya didasarkan pada prakiraan.
6. *Sequence* merupakan analisis yang digunakan dalam mencari pola untuk merangkai urutan atau peristiwa. Dalam kehidupan sehari dapat kita ambil

contoh adalah makan. Urutan dalam makan yaitu mengambil piring, garpu dan sendok. Setelah itu baru mengambil lauknya.

7. *Descriptive* bertujuan untuk memahami data yang mengarah pada pengamatan hasil perilaku data itu sendiri.

2.3. Sistem Informasi Geografis

Istilah “spasial” memiliki makna berkaitan dengan ruang atau tempat yang Melibatkan koordinat geografis. Informasi spasial dapat diketahui dari sistem Informasi Geografis (SIG) telah diacu oleh konteks ini. Dalam hal ini, informasi yang berhubungan dengan lokasi atau ruang pada bumi, kami dapat memperoleh wawasan tentang koordinat sebuah objek di permukaanya serta informasi terkait atribut maupun karakter Kawasan. Sistem Informasi Geografis atau bisa di singkat dengan (SIG) adalah sebuah sistem yang menyampaikan informasi berbasis *computer* dengan menggunakan data dengan komponen iformasi spasial. Sistem ini melakukan pengambilan , analisis, pemeriksaan, integrasi, manajemen, kondisi di bumi. Dengan memperhatikan unsur – unsur utamanya, dapat di rangkum bahwa Gis “Informasi Geografis”. SIG ialah satu buah alat lunak yang mengizinkan konsumen untuk memasukkan, menyimpan, memanipulasi, menampilkan serta membuahkan informasi geografis Bersama atributnya (Jannah et al., 2022).

Implementasi teknologi Sistem Informasi Geografis (SIG) menyatukan bermacam cara pengerjaan informasi yang biasa dipakai masa ini, termasuk pengumpulan informasi yang relevan serta Analisa statistik. Kelebihan teknologi ini terdapat pada penggambaran yang khas serta memberikan banyak manfaat dengan Analisa geografis dengan dukungan ilustrasi yang ada (Riswandi & Afrilia, 2022).

2.4. Clustering

Clustering merupakan sebuah proses pengelompokan semua anggota dari setiap partisi yang mempunyai kesamaan berdasarkan matriks tertentu (Harahap et al., 2022). Clustering dilakukan dengan pengamatan serta menggambarkan proses sekelompok objek data dari suatu set ke banyak kelas dalam melakukan partisi. Akan tetapi, proses ini dilakukan dengan mengimplementasi Euclidean Distance sebagai langkah-langkah dan persamaan dalam jarak algoritma (Venkateswaslu & Raju. 2013).

Dalam pengelompokan nya teknik cluster mempunyai dua metode yaitu hierarchical clustering dan non-hierarchical clustering. Hierarchical clustering adalah metode pengelompokan data dengan mengelompokkan dua data atau lebih yang mempunyai kesamaan yang kemudian proses dilanjutkan ke objek lain yang memiliki kedekatan dua sehingga proses ini terus berlangsung hingga cluster membentuk seperti tree dimana ada hirarki atau tingkatan yang jelas antar objek yang paling mirip hingga yang tidak mirip.

Sedangkan non-hierarchical adalah menentukan jumlah cluster yang diinginkan (dua cluster, tiga cluster atau lebih). Setelah jumlah cluster yang diinginkan maka proses cluster dimulai tanpa mengikuti proses hirarki (Anindya Khrisna Wardhani, 2016).

2.5. Algoritma K-Medoids

2.5.1. Definisi Algoritma K-Medoids

K-Medoids atau algoritma *Partition Around Medoids* (PAM) pada tahun 1987 dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw. Metode *Partitioning clustering* termasuk metode PAM yang digunakan untuk mengelompokkan sekumpulan objek menjadi cluster. Medoid adalah representasi cluster pada PAM dari sekumpulan objek yang mewakili cluster. K-Medoids adalah algoritma clustering yang hampir sama dengan algoritma k-means. Perbedaannya adalah K-Medoids menggunakan objek sebagai pusat cluster untuk setiap cluster sedangkan K-Means menggunakan nilai rata-rata sebagai pusat cluster untuk setiap cluster. Algoritma K-Medoids digunakan untuk mengatasi kelemahan dari algoritma K-

Means yang sangat sensitif terhadap noise dan outlier, dimana objek yang bernilai besar memungkinkan menyimpang pada distribusi data.

Adapun langkah penyelesaian dalam algoritma K-Medoids diantaranya :

1. Inisialisasi pusat cluster sebanyak k (jumlah cluster).
2. Hitung setiap objek ke cluster terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* :




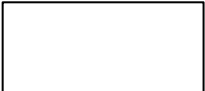

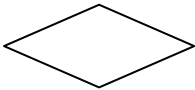

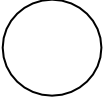
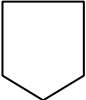
$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots \dots \dots (2.1)$$

3. Pilih secara acak objek pada masing-masing cluster sebagai medoid baru.
4. Hitung jarak objek pada masing-masing cluster dengan kandidat medoid
5. baru.
6. Hitung total simpangan (S), dengan menghitung nilai total distance baru-total distance lama. Jika $S < 0$, maka tukar objek dengan data cluster untuk membentuk sekumpulan k objek baru sebagai medoid.
7. Ulangi langkah 3-5 hingga tidak terjadi perubahan medoid, sehingga didapatkan cluster beserta anggota cluster masing-masing.

2.6. Flowchart

Flowchart adalah penggambaran dalam membuat langkah-langkah prosedur dari suatu program. Dalam membuat flowchart menggunakan anotasi bidang-bidang geometri seperti lingkaran, belah ketupat, oval, persegi empat dan sebagainya untuk mempresentasikan alur program (Sutanti et al., 2020). Berikut simbol pada flowchart beserta artinya :

Tabel 2.1
Simbol Flowchart

Atribut	Nama	Fungsi
	Terminator	Menyatakan permulaan/akhir suatu program
	Garis Alur (<i>Flow Line</i>)	Menyatakan arah aliran program yang menggabungkan antara simbol satu dengan simbol lainnya
	<i>Preparation</i>	Proses inisialisasi atau pemberian harga awal
	<i>Process</i>	Proses perhitungan atau proses pengolahan data
	<i>Input/Output</i>	Proses input/output data tanpa tergantung peralatan, parameter, informasi
	Logika (<i>Decision</i>)	Menunjukkan suatu kondisi tertentu dengan kemungkinan YA /TIDAK
	Sub Program (<i>Predefined Process</i>)	Pelaksanaan suatu bagian program (sub program)
	<i>On Page Conector</i>	Penghubung bagian-bagian flowchart yang berada pada satu halaman
	<i>Off Page Conector</i>	Penghubung bagian-bagian flowchart yang berada pada halaman berbeda

Sumber: (Sutanti et al., 2020)

2.7. PHP

PHP adalah singkatan dari HyperTextbPreprocessor. PHP merupakan bahasa pemograman yang digunakan untuk membuat sebuah aplikasi berbasis website menurut Enterprise. PHP adalah bahasa pemograman yang berjenis server-side yang artinya adalah php akan diproses oleh server yang hasilnya akan dikirim kembali ke browser. Oleh karena itu tool yang harus tersedia sebelum memulai pemograman php adalah server (Kadarsih & Andrianto, 2022).

2.8. MySQL

MySQL merupakan salah satu jenis database yang banyak digunakan untuk membuat sebuah aplikasi berbasis web. MySQL mempunya query atau bahasa SQL (Structured Query Language) yang simpel dan menggunakan escape character. Menurut Enterprise, MySQL adalah Relational Database Management System (RDBMS) yang mudah dan cepat digunakan berbagai kebutuhan (Kadarsih & Andrianto, 2022).

2.9. XAMPP

Menurut Enterprise, XAMPP adalah sebuah aplikasi web server instan dan lengkap karena untuk membuat sebuah web dengan Content Management System (Joomla) bisa digunakan dalam aplikasi ini. XAMPP memiliki sebuah paket installer AMP (Apache, MySQL dan PHP) yang sangat mudah diaplikasikan dalam komputer yang belum memiliki server (Kadarsih & Andrianto, 2022).

2.10. Penelitian Terdahulu

Penelitian terdahulu menjadi salah satu rujukan untuk menemukan perbandingan ketika melakukan sebuah penelitian sehingga dapat memberikan wawasan kepada penulis untuk mengevaluasi dan mencari inspirasi baru untuk penelitiannya. Berikut beberapa hasil penelitian terdahulu yang bertujuan memberikan bahan acuan diuraikan sebagai berikut :

Tabel 2.2
Penelitian Terdahulu

No.	Peneliti	Judul Penelitian
1.	Ruly Adi Permana, Dian Puspita Hapsari, Rotul Rotul Muhima (2022)	Penerapan Metode <i>CLUSTERING</i> K-MEDOIDS Untuk Pengelompokan Abstrak Skripsi Berbasis Web
	<p>Hasil Penelitian : Tujuan penelitian ini adalah untuk mengembangkan aplikasi pencarian yang dapat mengelompokkan skripsi. Data terdiri dari 105 data teks ringkasan abstrak skripsi jurusan Teknik Informatika ITATS dengan rentang waktu tiga tahun. Praproses menghitung total deviasinya. Hasil simulasi clustering K=4 maka abstrak makalah nilai K yang lebih optimal. Silhouette Coefficient sebesar -0,028975 (Permana et al., 2022).</p>	
2.	Tri Juninda Mustakim, Elvia Andri (2019)	Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau
	<p>Hasil Penelitian : Berdasarkan penelitian ini penyakit yang sering diderita di Pekanbaru Riau didapatkan hasil 4 cluster . Pada cluster 1 didapatkan 420 record penyakit dominan adalah Maag sebesar 44,39%, cluster 2 didapatkan 349 record penyakit dominan adalah Diare dan Sakit Perut sebesar 16,98%, pada cluster 3 didapatkan 794 record dengan penyakit dominan adalah Batuk dan Pilek sebesar 65,21% dan pada cluster 4 didapatkan 1248 record dengan penyakit dominan adalah Batuk dan Pilek sebesar 54,10% (Juninda et al., 2019)</p>	
3.	Astrika Praokta Fialine, Daffa Alya Alodia, Desy Endriani, Edy Widodo (2021)	Implementasi Metode K-Medoids <i>Clustering</i> untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan

Sambungan tabel di belakang

Tabel Sambungan

	<p>Hasil Penelitian : Dari hasil penelitian ini, diperoleh tiga cluster. Cluster pertama terdapat 6 provinsi dengan indikator pendidikan tinggi, cluster kedua terdapat 13 provinsi dengan indikator pendidikan sedang, dan cluster ketiga terdapat 15 provinsi dengan indikator pendidikan rendah (Alodia et al., 2021)</p>	
4.	<p>Eva Darnila, Mutamminul Ula, Mauliza, Iwan Pahendra, Ermatita (2020)</p>	<p><i>Classification Of Treatment Tuberculosis History Based On Machine Learning Techniques</i></p>
	<p>Hasil Penelitian : Pengobatan penyakit Tuberculosis sangat penting untuk melindungi pasien dan dapat menyebabkan kematian jika tidak diobati. Penelitian ini menggunakan teknik pembelajaran mesin dengan menggunakan Decision Tree dan Random Forest untuk mengklasifikasi Tuberculosis untuk dianalisis dan direpresentasikan berdasarkan riwayat pengobatan. Hasilnya menunjukkan kinerja yang dirancang berhasil dan dapat digunakan dalam menganalisis pengobatan Tuberculosis berdasarkan sejarah di Aceh Utara dan Lhokseumawe. Tidak hanya itu, hasilnya juga memudahkan untuk memahami data TBC dan membuat beberapa interpretasi yang sangat baik (Darnila et al., 2020).</p>	
5.	<p>Dedy Sutris Martua Simanjuntak, Indra Gunawan, Sumarno, Poningsih (2023)</p>	<p>Penerapan Algoritma K-Medoids Untuk Pengelompokan Pengangguran Umur 25 Tahun Keatas Di Sumatera Utara</p>
	<p>Hasil Penelitian : Tujuan penelitian ini adalah untuk Pengelompokan tingkat pengangguran di Sumatera Utara. Jumlah sampel yang digunakan adalah 34 kabupaten, khususnya yang berusia 25 tahun ke atas di Sumatera Utara, Berdasarkan hasil pengelompokan terdapat tiga cluster, yaitu 14 kabupaten di cluster nol, 2 Kabupaten di cluster satu, dan 17 Kabupaten di Cluster dua (Simanjuntak et al., 2023).</p>	

Sambungan tabel di belakang

6	Fadhlan Sulistiyo Hidayat, Rizma Berliana Putri Affandi, Virgaria Zuliana, Tesa Nur Padilah (2022)	Penerapan K-Means <i>Clustering</i> dalam Pengelompokan Kasus Tuberkulosis di Provinsi Jawa Barat
	<p>Hasil Penelitian : Jenis kelamin atau gender tidak mempengaruhi jumlah kasus tuberkulosis, karena hasil cluster menunjukkan tidak ada perbedaan signifikan dalam total kasus antara jenis kelamin laki laki dan jenis kelamin perempuan. Hasil penelitian juga menghasilkan 2 cluster yaitu cluster dengan jumlah kasus tinggi dan rendah. Jumlah kasus dengan cluster tinggi terdapat 6 daerah, sedangkan cluster terendah terdapat 21 daerah. Evaluasi menggunakan silhouette coefficient mendapatkan nilai sebesar 0.59 yang berarti hasil cluster dari penelitian ini cukup baik (Sulistiyo Hidayat et al., 2022).</p>	
7	Hendra Kusumah, Muhammad Ridwan Lubis, Heru Satria Tambunan (2021)	Penerapan Algoritma <i>K-Medoids</i> Dalam Pengelompokan Imunisasi Lanjutan Pada Anak 2 Tahun
	<p>Hasil Penelitian : Data pada penelitian ini bersumber dari Departemen Kesehatan Tahun 2017-2019. Pengelompokan dilakukan berdasarkan jumlah penerima vaksin DPT-HB-HiB dan Campak/MR dari 34 Provinsi yang ada di Indonesia. Hasil penelitian didapatkan 3 cluster dengan anggota yaitu: cluster 1 di isi 1 Provinsi, cluster 2 di isi 18 Provinsi, cluster 3 di isi 15 Provinsi (Kusumah et al., 2021).</p>	

Tabel Sambungan

8	Astrika Praokta Fialine, Daffa Alya Alodia, Desy Endriani, Edy Widodo (2021)	Implementasi Metode <i>K-Medoids Clustering</i> untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan
---	--	--

Sambungan tabel di belakang

Tabel Sambungan

	Data yang digunakan diperoleh dari publikasi Statistik Pendidikan tahun ajaran 2019/2020 oleh Badan Pusat Statistik dengan 8 variabel indikator Pendidikan SMA/MA sederajat menurut provinsi di Indonesia, yaitu Angka Partisipasi Kasar (APK), Angka Partisipasi Murni (APM), Angka Partisipasi Sekolah (APS), persentase guru layak, Rata-Rata Lama Sekolah (RLS), Angka Melanjutkan (AM), tingkat menyelesaikan sekolah (SMA), dan angka putus sekolah. Dari hasil analisis, diperoleh tiga cluster. Cluster pertama terdapat 6 provinsi dengan indikator pendidikan tinggi, cluster kedua terdapat provinsi dengan indicator pendidikan sedang, dan cluster ketiga terdapat 15 provinsi dengan indicator Pendidikan rendah (Alodia et al., 2021).	
9	Eva Darnila, Muhammada Ula, Mauliza, Ermatita, Iwan Pahendra (2018)	Implementasi Sistem Pakar Pada Pasien Penderita Tuberkulosis Potential Drop Out di Rumah Sakit Cut Meutia Aceh Utara
	Berdasarkan hasil dari aplikasi Sistem Pakar pada Pasien Penderita TBC Potential Drop Out di Rumah Sakit Cut Meutia Aceh Utara berdasarkan kode Kasus 31 dengan system pendeteksian Sistem Pasien Paru BTA (+) dengan gejala pendeteksian nya adalah Pasien batuk berdahak selama 2-3 minggu Pasien tersebut pada hasil pemeriksaan sputum, Pasien yang sudah	

Tabel Sambungan

<p>pernah diobati dengan obat TBC kurang dari 1 bulan. pada hasil pemeriksaan sputum, Pasien yang sudah pernah diobati dengan obat TBC kurang dari 1 bulan, Pasien TB menghentikan pengobatannya dan Pasien TB kembali ke sarana fasilitas pelayanan kesehatan dengan nilai kasus tertinggi 0.6111 dari semua system pendeteksian yang telah di ujikan (Darnila et al., 2018).</p>		
10	Yanti Puspita Sari, Aji Primajaya, Agung Susilo Yuda Irawan (2020)	Implementasi Algoritma <i>K-Means</i> untuk <i>Clustering</i> Penyebaran Tuberkulosis di Kabupaten Karawang
<p>Metode yang dapat digunakan untuk mengelompokkan daerah penyebaran Penyakit tuberkulosis adalah data mining. Metodologi data mining yang diterapkan yaitu CRISP-DM. Dataset yang digunakan bersumber dari Dinas Kesehatan Kabupaten Karawang tahun 2018. Kemudian data tersebut diproses menggunakan algoritma k-means clustering dengan bantuan tools WEKA. Penelitian diawali dengan mencari jumlah cluster terbaik menggunakan metode elbow, dimana K=3 adalah jumlah cluster terbaik. K=3 dikombinasikan dengan S (random seed) 5, 10, 15 untuk diuji coba. Hasil dari uji coba tersebut kemudian dievaluasi menggunakan SSE (Sum of Square Error) dan Silhouette. Clustering terbaik dihasilkan oleh K=3 dengan kombinasi S=10 yang terdiri dari cluster 0 (7 anggota), cluster 1 (9 anggota), dan cluster 2 (14 anggota). Nilai evaluasi yang dihasilkan SSE 2,4402 dan Silhouette 0,5629 (Sari et al., 2020).</p>		

BAB III

METODE PENELITIAN

3.1 Tempat dan Waktu Pelaksanaan Penelitian

Penelitian ini dilakukan di Dinas Kesehatan Kabupaten Aceh Tamiang yang beralamat di Jl.Ir. H Juanda, Karang Baru, Komplek Perkantoran Pemda Aceh Tamiang, Aceh. Data set yang digunakan pada penelitian ini adalah data pasien Tuberculosis pada tahun 2019,2020,2021,2022 dan 2023 dengan waktu penelitian dimulai pada bulan februari.

3.2 Alat dan Bahan Penelitian

3.2.1. Perangkat Keras (Hardware)

Pada penelitian ini spesifikasi perangkat keras yang digunakan adalah sebagai berikut :

1. Laptop Lenovo IdeaPad 3 14IIL05
2. Ram 8.00 GB
3. SSD 182 GB

3.2.2. Perangkat Lunak (Software)

1. Windows 10 64 bit
2. Microsoft Office 2019
3. Microsoft Excel 2019
4. Google Chrome
5. Visual Studio Code
6. Apache 2.0
7. Database MySQL

3.3 Metode Pengumpulan Data

Pada tahap penelitian ini penulis melakukan pengumpulan data yang memiliki beberapa proses yang dibutuhkan dalam menentukan pengelompokan penyakit Tuberculosis di Kabupaten Aceh Tamiang. Pada tahap pertama, pembuatan berkas surat pengantar penelitian skripsi dari pihak program studi (prodi) yang kemudian surat tersebut dikirimkan ke pihak jurusan sehingga keluar

surat pengantar penelitian skripsi dari jurusan yang dapat digunakan dalam meneliti suatu tempat tujuan yang dituju untuk diteliti. Setelah itu, didapatkan data penelitian yang dibutuhkan dari administrator untuk penulisan skripsi sehingga dari data yang didapat, penulis dapat merancang dan membangun sistem. Data yang dibutuhkan dalam penelitian terdiri dari:

3.3.1. Data Primer

Data primer yang digunakan dalam mengelompokkan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang yang didapatkan dari administrator Dinas Kesehatan Kabupaten Aceh Tamiang.

3.3.2. Data Sekunder

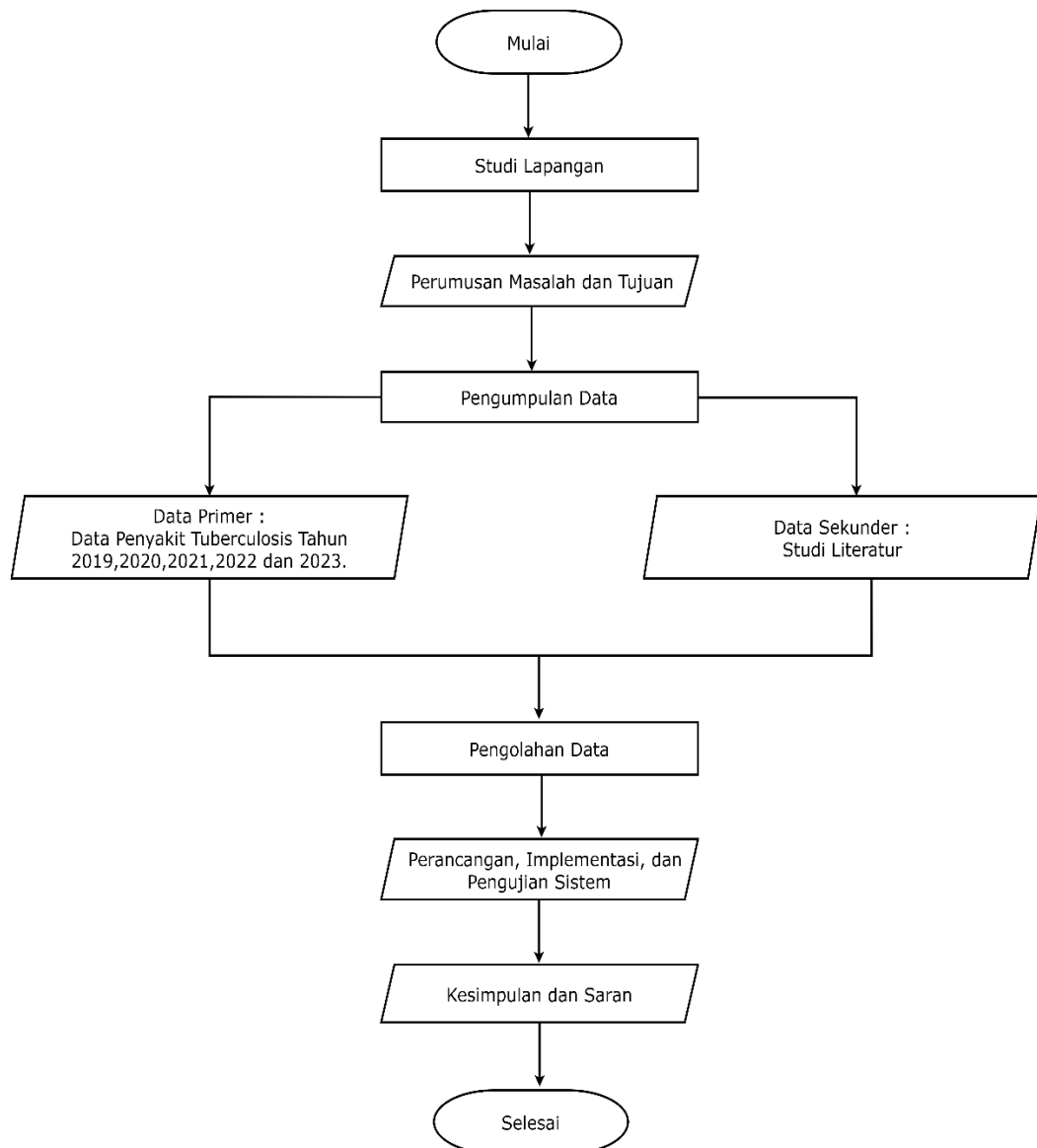
Data sekunder yang digunakan penulis sebagai pelengkap dalam penyusunan tugas akhir. Penulis mengambil referensi lain sebagai pendukung, diantaranya jurnal, buku yang sesuai dengan judul penulis yaitu tentang Clustering, Penyakit Tuberculosis dan algoritma K-Medoids.

3.4 Analisa Kebutuhan Sistem

Pada tahap ini peneliti melakukan analisa sistem dalam mengolah data yang sudah didapatkan dan yang digunakan dalam merancang sebuah sistem clustering untuk Pengelompokan Penyebaran penyakit Tuberculosis di Kabupaten Aceh Tamiang.

3.5 Alur Penelitian

Alur penelitian pada sistem Pengelompokan Penyebaran Penyakit Tuberculosis di Kabupaten Aceh Tamiang Menggunakan Metode K-Medoid dijelaskan pada gambar berikut :



Gambar 3.1 Alur Penelitian

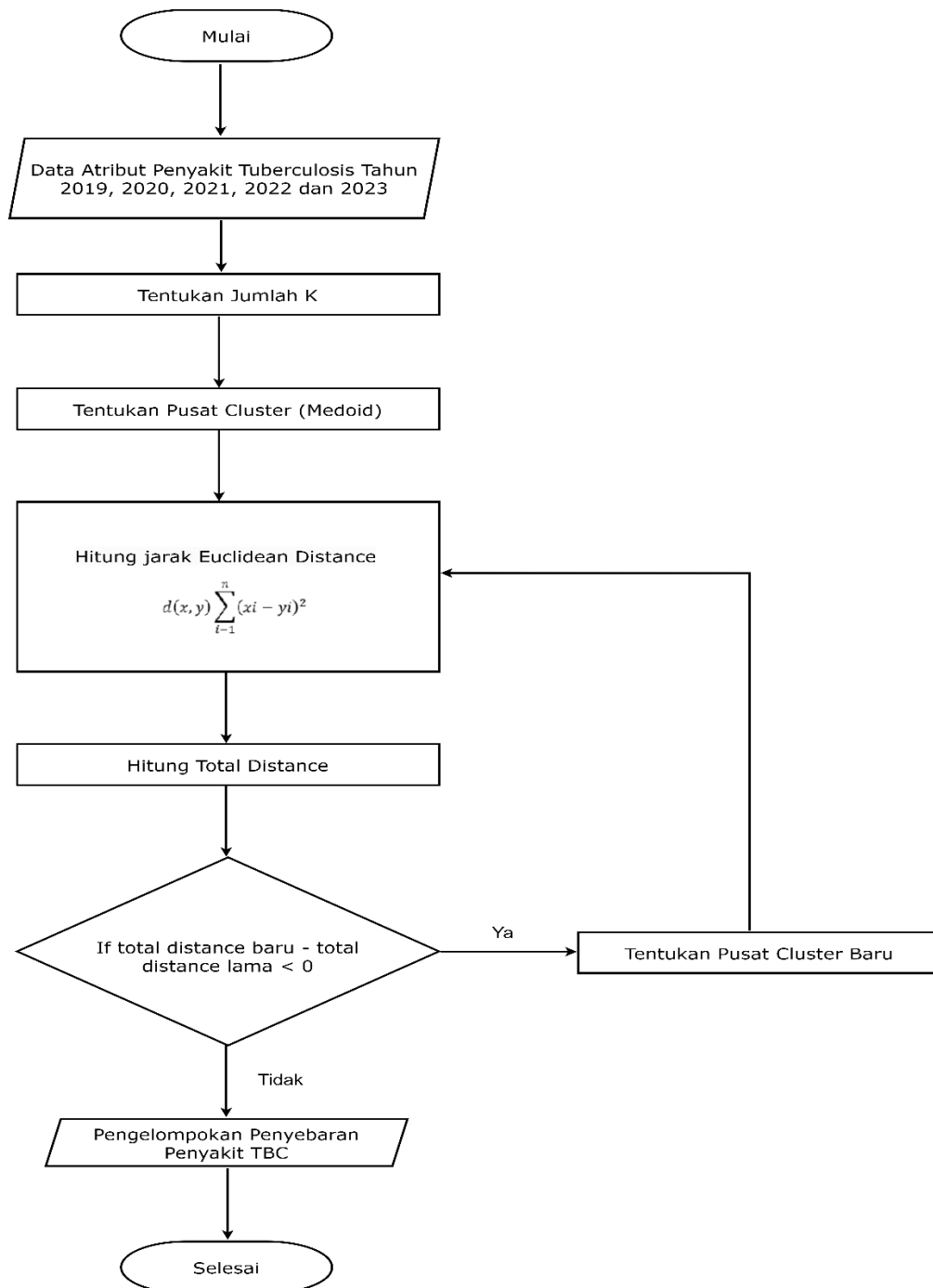
Alur penelitian diuraikan sebagai berikut:

1. Dimulai dari studi lapangan melihat dan mencari masalah yang akan diteliti agar menemukan sebuah solusi yang dapat memecahkan masalah tersebut.
2. Setelah menemukan topik permasalahan, maka langkah selanjutnya menentukan perumusan masalah yang akan dikupas secara tuntas demi mendapatkan sebuah solusi yang tepat dan menemukan tujuan penelitian dari rumusan masalah yang akan dibahas.

9. Kemudian pada tahap pengumpulan data terbagi dua bagian diantaranya data primer yang merupakan data penyakit Tuberculosis tahun 2019,2020,2021dan 2023 dengan memiliki atribut yaitu jenis kelamin, umur, alamat, Penyakit Tuberculosis, Tipe Diagnosis TBC, dan Klasifikasi Berdsarkan Lokasi Anatomi serta data sekunder merupakan data pelengkap dalam menyusun penelitian.
3. Setelah mendapatkan data yang dibutuhkan kemudian, data tersebut diolah menggunakan *Microsoft Excel* sehingga menghasilkan nilai yang sesuai dengan algoritma yang diterapkan.
4. Setelah diolah dengan perhitungan manual, maka selanjutnya tahap perancangan yang akan dirancang menggunakan *flowchart*, dibangun menggunakan bahasa pemograman *php*, dan setelah itu, diuji kembali program yang telah dibuat apakah berhasil ataupun tidak.
5. Setelah mendapatkan hasil pengujian tersebut maka dapat menghasilkan kesimpulan dan saran yang membangun guna memperbaiki penelitian ini kedepannya. Kemudian program berhenti.

3.6 Skema dan Perancangan Sistem

Berikut merupakan susunan model yang akan diterapkan kedalam bahasa pemograman :



Gambar 3.2 Skema Sistem

Alur skema sistem dari Algoritma *K-Medoids* diuraikan sebagai berikut:

1. Sistem dimulai dari menginput data atribut pengelompokkan penyakit Tuberculosis tahun 2019, 2020, 2021, 2022 dan 2023 diantaranya jenis kelamin, umur, alamat, penyakit Tuberculosis, tipe diagnosis TBC, dan klasifikasi berdasarkan lokasi anatomi.
2. Setelah diinput, maka langkah selanjutnya tentukan nilai jumlah k .
3. Menentukan 3 pusat cluster awal secara acak.
4. Kemudian menghitung jarak dari obyek ke medoid sementara, setelah itu beri tanda pada nilai jarak terkecil obyek tersebut ke medoid kemudian hitung jumlahnya dengan menggunakan rumus *Euclidean Distance*.
5. Setelah mendapatkan hasil jarak dari rumus *Euclidean Distance*, selanjutnya hitung total distance.
6. Kemudian hitung total distance dengan menghitung total cost baru – total cost lama.
7. Jika total simpangan < 0 , maka perhitungan dilanjutkan sampai dengan total simpangan > 0 .
8. Tahap selanjutnya tentukan kembali medoids baru sebagai titik pusat cluster pada iterasi yang dihitung.
9. Hitung kembali jarak setiap data ke masing-masing non medoids menggunakan rumus *Euclidean Distance* sehingga diperoleh hasil penghitungan jarak setiap data untuk masing-masing cluster sampai dengan total simpangan > 0 .
10. Setelah dilakukan perhitungan pada semua data, maka proses selanjutnya adalah memilah cluster mana dari tiap data dimana nilai terkecil dari tiap perhitungan cluster akan menjadi penentu cluster yang akan diikuti dari data tersebut hitung kembali total simpangan baru dengan mencari selisih total cost baru dengan total cost lama. $S = \text{total cost baru} - \text{total cost lama}$.
11. Jika total simpangan > 0 , maka perhitungan dihentikan sehingga didapatkan hasil cluster akhir.
12. Kemudian sistem menampilkan output hasil pengelompokan dari penyakit

Tuberculosis yang dikelompokkan sesuai dengan cluster masing-masing yaitu cluster rendah, cluster sedang dan cluster tinggi.

13. Tahap terakhir adalah selesai dimana ini adalah tanda perhentian dari sistem.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Pada penelitian ini penulis mengimplementasikan metode K-Medoids untuk mengelompokkan wilayah penyebaran penyakit Tuberkulosis berdasarkan kecamatan di Kabupaten Aceh Tamiang, yang memiliki 12 kecamatan yaitu Banda Mulia, Bandar Pusaka, Bendahara, Karang Baru, Kejuruan Muda, Kota Kuala Simpang, Manyak Payed, Rantau, Sekerak, Seruway, Tamiang Hulu dan Tenggulun. Hal ini dilakukan untuk mempermudah pemerintah untuk mengetahui daerah penyebaran penyakit TBC dengan Tingkat penyebaran rendah, sedang dan tinggi. Data yang digunakan pada penelitian ini adalah data tahun 2019, 2020, 2021, 2022 dan tahun 2023.

4.1.1. Analisis Sistem

Analisis sistem adalah tahapan paling awal dalam pengembangan sistem yang nantinya menjadi fondasi menentukan keberhasilan sistem yang dihasilkan. Analisa sistem memiliki 3 tahapan yaitu Analisa masalah, Analisa kebutuhan sistem, dan analisa proses.

4.1.1.1 Analisa Masalah

Kabupaten Aceh Tamiang terletak di perbatasan Aceh-Sumatera Utara. Kondisi ini dapat menyebabkan penularan penyakit tuberculsis semakin tinggi. Dengan adanya permasalahan tersebut, perlu adanya suatu upaya penanganan yang efektif yaitu dengan mengelompokkan daerah penyebaran penyakit Tuberculosis di Kabupaten Aceh tamiang menggunakan data mining dengan metode K-Medoids Clustering.

4.1.1.2 Analisa Kebutuhan Sistem

Agar sistem berjalan sesuai yang diharapkan, penulis melakukan pengumpulan data pasien tuberculosis paru yang diperoleh dari RSUD Aceh Tamiang. Didapatkan data pasien tuberculosis paru dan kesimpulan sebagai berikut:

1. Data pasien Tuberkulosis paru diperoleh dari tahun 2019 sampai tahun 2023 yaitu umur, jenis kelamin, alamat, penyakit tuberkulosis paru, dan penunjang.
2. Penentuan parameter data yang di proses yaitu kecamatan, jumlah kasus tb paru, jumlah laki-laki, dan jumlah perempuan .
3. Dalam menentukan klistor dibagi menjadi 3 kluster yaitu kluster rendah, sedang dan tinggi.
4. Penentuan centroid awal di tentukan secara acak atau random.
5. Setelah jumlah K ditetapkan maka Langkah selanjutnya tiap nilai data dihitung menggunakan jarak Euclidean distance.

4.1.1.3 Analisa Proses

Membangun sistem yang dapat mengkluster daerah penyebaran penyakit tuberkulosis di Kabupaten Aceh Tamiang. Output dari sistem ini dapat menjadi bahan rekomendasi untuk pemerintah Kabupaten Aceh Tamiang dalam melakukan penanganan yang lebih efektif terlebih dahulu ke daerah yang memiliki kasus penyakit Tuberculosis tingkat tinggi, sehingga kasus Tuberculosis di Kabupaten Aceh Tamiang semakin berkurang.

4.1.2. Perhitungan Data tahun 2019

Berikut adalah perhitungan yang dilakukan pada tahun 2019.

4.1.2.1. Dataset

Penentuan parameter data yang di proses adalah kecamatan, jumlah kasus tb paru, jumlah laki-laki, dan jumlah Perempuan. Dataset dapat dilihat pada table di bawah ini :

Tabel 4.1
Data Tahun 2019

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	5	1	4
2	BANDAR PUSAKA	4	2	2
3	BENDAHARA	9	5	4
4	KARANG BARU	19	10	9
5	KEJURUAN MUDA	21	8	13
6	KOTA KUALA SIMPANG	15	4	11
7	MANYAK PAYED	10	6	4
8	RANTAU	25	9	16
9	SEKERAK	3	1	2
10	SERUWAY	13	6	7
11	TAMIANH HULU	9	3	6
12	TENGGULUN	7	3	4

4.1.2.2 Normalisasi Data

Tujuan normalisasi data adalah mengubah nilai kolom numerik dalam himpunan data untuk menggunakan skala umum, tanpa mendistorsi perbedaan dalam rentang nilai. Normalisasi juga diperlukan untuk beberapa algoritma dalam membuat model data dengan benar. Normalisasi data dilakukan dengan cara metode Min-Max. Adapun normalisasi data dapat dilihat pada table di bawah ini :

Tabel 4.2
Normalisasi Data Tahun 2019

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	0.090909091	0	0.142857143
2	BANDAR PUSAKA	0.045454545	0.111111111	0
3	BENDAHARA	0.272727273	0.444444444	0.142857143
4	KARANG BARU	0.727272727	1	0.5
5	KEJURUAN MUDA	0.818181818	0.777777778	0.785714286
6	KOTA KUALA SIMPANG	0.545454545	0.333333333	0.642857143
7	MANYAK PAYED	0.318181818	0.555555556	0.142857143
8	RANTAU	1	0.888888889	1
9	SEKERAK	0	0	0
10	SERUWAY	0.454545455	0.555555556	0.357142857
11	TAMIANG HULU	0.272727273	0.222222222	0.285714286
12	TENGGULUN	0.181818182	0.222222222	0.142857143

4.1.2.3 Pengujian Data

Langkah - Langkah yang dilakukan sebagai berikut :

1. Inisialisasi pusat kluster sebanyak 3 kluster dari data sampel untuk pemilihan setiap medoids yang dipilih secara acak atau random. Berikut ini adalah nilai pusat kluster awal pada pengujian pertama yaitu penjabaran pada table data testing.

Tabel 4.3
Nilai Medoid Awal

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	0.090909091	0	0.142857143
2	BANDAR PUSAKA	0.045454545	0.111111111	0
3	BENDAHARA	0.272727273	0.444444444	0.142857143

2. Menghitung jarak terdekat dengan persamaan Euclidean Distance

Untuk melakukan klustering pada setiap data yang telah diperoleh, maka langkah selanjutnya adalah menghitung persamaan jarak menggunakan metode Euclidean Distance .

Menghitung Distance data dengan pusat kluter pertama :

$$d_{(1)} = \sqrt{(0,090909091-0,090909091)^2 + (0-0)^2 + (0,142857143-0,142857143)^2}$$

$$= 0$$

Menghitung Distance data dengan pusat kluter kedua :

$$d_{(2)} = \sqrt{(0,090909091-0,045454545)^2 + (0-0,011111111)^2 + (0,142857143-0)^2}$$

$$= 0,186601066$$

Menghitung Distance data dengan pusat kluter ketiga :

$$d_{(3)} = \sqrt{(0,090909091-0,272727273)^2 + (0-0,444444444)^2 + (0,142857143-0,142857143)^2}$$

$$= 0,480196538$$

Perhitungan yang sama tetap dilakukan kepada seluruh data. Setelah dilakukan perhitungan ke seluruh data dan atribut maka akan mendapatkan jarak terdekat. Setiap data pada masing-masing kluster seperti table dibawah ini :

Tabel 4.4
Hasil Perhitungan Algoritma K-Medoids iterasi ke-1

NO	C1	C2	C3	KEDEKATAN	KLUSTER
1	0	0.186601066	0.480196538	0	1
2	0,186601066	0	0.427986176	0	2
3	0,480196538	0.427986176	0	0	3
4	1,193887282	1.156128082	0.731889137	0.731889137	3
5	1,243836483	1.287982462	0.906585395	0.906585395	3
6	0,753473743	0.844184827	0.580280832	0.580280832	3
7	0.600245673	0.540665509	0.120049135	0.120049135	3
8	1.533383062	1.586220444	1.208780527	1.208780527	3
9	0.169329933	0.120049135	0.540665509	0.120049135	2
10	0.697704628	0.701738738	0.302195132	0.302195132	3
11	0.320700375	0.381616594	0.264179635	0.264179635	3
12	0.250497943	0.189849312	0.250497943	0.189849312	2
JUMLAH KEDEKATAN				4.423858238	

3. Menentukan nilai *Medoids* Baru (non *Medoids*)

Selanjutnya menentukan nilai *Medoids* baru dengan cara memilih secara acak atau random, setiap *medoids* yang sudah terpilih maka tidak dapat di jadikan lagi sebagai *medoids* baru. Nilai *medoids* baru dapat dilihat pada table di bawah ini:

Tabel 4.5
Nilai *Medoids* Baru

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
10	SERUWAY	0.454545455	0.555555556	0.357142857
11	TAMANG HULU	0.272727273	0.222222222	0.285714286
12	TENGULUN	0.181818182	0.222222222	0.071428571

4. Hitung Kembali jarak dari setiap objek pada iterasi kedua dengan menggunakan Medoids baru pada tabel 4.4.

Menghitung Distance data dengan pusat kluster pertama :

$$d_{(1)} = \sqrt{(0,090909091-0,454545455)^2 + (0-0,555555555)^2 + (0,142857143-0,357142857)^2}$$

$$= 0,697704628$$

Menghitung Distance data dengan pusat kluster kedua :

$$d_{(2)} = \sqrt{(0,090909091-0,272727273)^2 + (0-0,222222222)^2 + (0,142857143-0,285714286)^2}$$

$$= 0,320700375$$

Menghitung Distance data dengan pusat kluster ketiga :

$$d_{(3)} = \sqrt{(0,090909091-0,181818182)^2 + (0-0,222222222)^2 + (0,142857143-0,071428571)^2}$$

$$= 0,250497943$$

Perhitungan yang sama tetap dilakukan kepada seluruh data. Setelah dilakukan perhitungan ke seluruh data dan atribut maka akan mendapatkan jarak terdekat. Setiap data pada masing-masing kluster seperti tabel dibawah ini :

Tabel 4.6

Hasil Perhitungan Algoritma K-Medoids iterasi ke-2

NO	C1	C2	C3	KEDEKATAN	KLUSTER
1	0.697704628	0.320700375	0.250497943	0.250497943	3
2	0.701738738	0.381616594	0.189849312	0.189849312	3
3	0.302195132	0.264179635	0.250497943	0.250497943	3
4	0.526320312	0.900860612	0.973846651	0.526320312	1
5	0.604390263	0.925290569	1.106257083	0.604390263	1
6	0.373202133	0.462900491	0.686372855	0.373202133	1
7	0.253994899	0.365493352	0.367162353	0.253994899	1
8	0.906585395	1.218020585	1.405742092	0.906585395	1
9	0.801750938	0.453205841	0.295876001	0.295876001	3

Lanjutan tabel dibelakang

Tabel lanjutan

10	0	0.386356058	0.516840333	0	1
11	0.386356058	0	0.23277205	0	2
12	0.516840333	0.232772056	0	0	3
JUMLAH KEDEKATAN				3.651214201	

5. Menghitung total simpangan (S)

Setelah didapatkan nilai jarak antara iterasi ke-1 dan iterasi ke-2, hitung total simpangan (S) dengan mencari selisih dari nilai total cost baru dikurang (-) nilai total cost lama. Dengan ketentuan $S < 0$, maka tukar nilai objek dengan menentukan medoid baru.

$$\begin{aligned}
 S &= \text{total cost baru} - \text{total cost lama} \\
 &= 3.651214201 - 4.423858238 \\
 &= -0.7726
 \end{aligned}$$

Karena hasil total simpangan kurang dari nol $S < 0$, maka pengujian ini dilanjutkan pada iterasi ke-3. Ulangi lagi langkah ke 3 yaitu menentukan nilai medoid baru .

Tabel 4.7

Nilai Medoids Baru

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
6	KOTA KUALA SIMPANG	0.545454545	0.333333333	0.642857143
7	MANYAK PAYED	0.318181818	0.555555556	0.142857143
8	RANTAU	1	0.888888889	1

Kemudian hitung Kembali jarak dari setiap objek pada iterasi ketiga menggunakan medoids baru pada tabel 4.7.

Menghitung Distance data dengan pusat kluster pertama :

$$d_{(1)} = \sqrt{(0,090909091-0,545454545)^2 + (0-0,333333333)^2 + (0,142857143-0,642857143)^2}$$

$$= 0,753473743$$

Menghitung Distance data dengan pusat kluster kedua :

$$d_{(2)} = \sqrt{(0,090909091-0,318181818)^2 + (0-0,555555556)^2 + (0,142857143-0,142857143)^2}$$

$$= 0,600245673$$

Menghitung Distance data dengan pusat kluster ketiga :

$$d_{(3)} = \sqrt{(0,090909091-1)^2 + (0-0,888888889)^2 + (0,142857143-1)^2}$$

$$= 1,533383062$$

Perhitungan yang sama tetap dilakukan kepada seluruh data dan atribut untuk mendapatkan jarak terdekat seperti table dibawah ini :

Tabel 4.8

Hasil Perhitungan Algoritma K-Medoids iterasi ke-3

NO	C1	C2	C3	KEDEKATAN	KLUSTER
1	0.753473743	0.600245673	1.533383062	0.600245673	2
2	0.844184827	0.540665509	1.586220444	0.540665509	2
3	0.580280832	0.120049135	1.208780527	0.120049135	2
4	0.777851731	0.620720871	0.772612403	0.620720871	2
5	0.540665509	0.844184827	0.302195132	0.302195132	3
6	0	0.592482581	0.801750938	0	1
7	0.592482581	0	1.144849781	0	2
8	0.801750938	1.144849781	0	0	3
9	0.906585395	0.655964792	1.670366264	0.655964792	2
10	0.373202133	0.253994899	0.906585395	0.253994899	2
11	0.462900491	0.365493352	1.218020585	0.365493352	2
12	0.686372855	0.367162353	1.405742092	0.367162353	2
JUMLAH KEDEKATAN				3.826491714	

Setelah didapatkan nilai jarak antara iterasi ke-2 dan iterasi ke-3, hitung total simpangan (S) dengan mencari selisih dari nilai total cost baru dikurang (-) nilai total cost lama. Dengan ketentuan $S < 0$, maka tukar nilai objek dengan menentukan medoid baru.

$S = \text{total cost baru} - \text{total cost lama}$

$= 3.826491714 - 3.651214201$

$= 0.17528$

Karena hasil total simpangan lebih besar dari nol $S > 0$, maka pengujian ini dihentikan pada iterasi ke-3. Hasil dari iterasi terakhir akan menjadi parameter pengklasteran. Penulis menentukan anggota kluster mana yang masuk kedalam anggota kluster penyebaran penyakit tuberculosis rendah, sedang dan tinggi berdasarkan centroid yaitu $C1 = \text{rendah}$, $C2 = \text{sedang}$, $C3 = \text{tinggi}$, maka anggota dari $C1$ adalah kluster rendah, $C2$ adalah kluster sedang, dan anggota $C3$ adalah kluster tinggi.

4.1.2.4 Hasil Akhir

Tabel 4.9

Hasil akhir perhitungan tahun 2019

NO	KECAMATAN	CLUSTER
1	BANDA MULIA	SEDANG
2	BANDAR PUSAKA	SEDANG
3	BENDAHARA	SEDANG
4	KARANG BARU	SEDANG
5	KEJURUAN MUDA	TINGGI
6	KOTA KUALA SIMPANG	RENDAH
7	MANYAK PAYED	SEDANG
8	RANTAU	TINGGI
9	SEKERAK	SEDANG
10	SERUWAY	SEDANG
11	TAMIANG HULU	SEDANG
12	TENGGULUN	SEDANG

4.1.3 Perhitungan Data tahun 2020

4.1.3.1 Data Set

Tabel 4.10
Data Tahun 2020

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	6	5	1
2	BANDAR PUSAKA	2	1	1
63	BENDAHARA	3	2	1
4	KARANG BARU	13	8	5
5	KEJURUAN MUDA	13	10	3
6	KOTA KUALA SIMPANG	16	8	8
7	MANYAK PAYED	2	2	0
8	RANTAU	15	12	3
9	SEKERAK	6	4	2
10	SERUWAY	10	7	3
11	TAMIANG HULU	2	1	1
12	TENGGULUN	3	2	1

4.1.3.2 Hasi Akhir

Tabel 4.11
Data Tahun 2020

NO	KECAMATAN	CLUSTER
1	BANDA MULIA	SEDANG
2	BANDAR PUSAKA	RENDAH
3	BENDAHARA	RENDAH

Sambungan Tabel Dibelakang

Tabel Sambungan

4	KARANG BARU	RENDAH
5	KEJURUAN MUDA	SEDANG
6	KOTA KUALA SIMPANG	TINGGI
7	MANYAK PAYED	SEDANG
8	RANTAU	SEDANG
9	SEKERAK	RENDAH
10	SERUWAY	SEDANG
11	TAMIANG HULU	RENDAH
12	TENGGULUN	RENDAH

4.1.4 Perhitungan Data tahun 2021

4.1.4.1 Data Set

Tabel 4.12

Data Tahun 2021

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	5	2	3
2	BANDAR PUSAKA	6	4	2
63	BENDAHARA	6	3	3
4	KARANG BARU	25	18	7
5	KEJURUAN MUDA	14	10	4
6	KOTA KUALA SIMPANG	9	4	5
7	MANYAK PAYED	5	2	3
8	RANTAU	9	6	3
9	SEKERAK	7	3	4

Lanjutan Tabel Dibelakang

Tabel Lanjutan

10	SERUWAY	7	3	4
11	TAMANG HULU	4	4	0
12	TENGKULUN	7	6	1

4.1.4.5 Hasil Akhir

Tabel 4.13
Data Tahun 2021

NO	KECAMATAN	CLUSTER
1	BANDA MULIA	RENDAH
2	BANDAR PUSAKA	SEDANG
3	BENDAHARA	TINGGI
4	KARANG BARU	TINGGI
5	KEJURUAN MUDA	TINGGI
6	KOTA KUALA SIMPANG	TINGGI
7	MANYAK PAYED	RENDAH
8	RANTAU	TINGGI
9	SEKERAK	TINGGI
10	SERUWAY	TINGGI
11	TAMANG HULU	SEDANG
12	TENGKULUN	SEDANG

4.1.5 Perhitungan Data tahun 2022

4.1.5.1 Data Set

Tabel 4.14
Data Tahun 2022

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	3	2	1
2	BANDAR PUSAKA	5	3	2
3	BENDAHARA	7	3	4
4	KARANG BARU	12	9	3
5	KEJURUAN MUDA	11	9	2
6	KOTA KUALA SIMPANG	12	10	2
7	MANYAK PAYED	4	4	0
8	RANTAU	16	10	6
9	SEKERAK	3	2	1
10	SERUWAY	6	3	3
11	TAMIANG HULU	2	1	1
12	TENGGULUN	4	2	2

4.1.5.2 Hasil Akhir

Tabel 4.15
Hasil Akhir Perhitungan Tahun 2022

NO	KECAMATAN	CLUSTER
1	BANDA MULIA	RENDAH
2	BANDAR PUSAKA	SEDANG
3	BENDAHARA	TINGGI
4	KARANG BARU	TINGGI

Lanjutan Tabel Dibelakang

Tabel Lanjutan

5	KEJURUAN MUDA	SEDANG
6	KOTA KUALA SIMPANG	TINGGI
7	MANYAK PAYED	RENDAH
8	RANTAU	TINGGI
9	SEKERAK	RENDAH
10	SERUWAY	TINGGI
11	TAMIAN HULU	RENDAH
12	TENGGULUN	SEDANG

4.1.6 Perhitungan Data tahun 2023**4.1.6.1 Data Set****Tabel 4.16****Data Tahun 2023**

NO	KECAMATAN	JUMLAH KASUS TB PARU	JUMLAH LAKI-LAKI	JUMLAH PEREMPUAN
1	BANDA MULIA	0	0	0
2	BANDAR PUSAKA	2	1	1
3	BENDAHARA	1	1	1
4	KARANG BARU	9	8	1
5	KEJURUAN MUDA	7	5	2
6	KOTA KUALA SIMPANG	8	7	1
7	MANYAK PAYED	3	3	0
8	RANTAU	10	9	1
9	SEKERAK	2	1	1
10	SERUWAY	4	2	2
11	TAMIAN HULU	2	0	2
12	TENGGULUN	0	0	0

4.1.6.2 Hasil Akhir

Tabel 4.17
Data Tahun 2023

NO	KECAMATAN	CLUSTER
1	BANDA MULIA	SEDANG
2	BANDAR PUSAKA	SEDANG
3	BENDAHARA	SEDANG
4	KARANG BARU	TINGGI
5	KEJURUAN MUDA	RENDAH
6	KOTA KUALA SIMPANG	RENDAH
7	MANYAK PAYED	SEDANG
8	RANTAU	TINGGI
9	SEKERAK	SEDANG
10	SERUWAY	RENDAH
11	TAMIANH HULU	SEDANG
12	TENGGULUN	SEDANG

KESIMPULAN DAN SARAN

5.1 Kesimpulan

5.2 Saran

Berdasarkan hasil analisis yang telah dilakukan, maka diperoleh beberapa saran sebagai berikut :

1. Penelitian selanjutnya dapat melakukan pengembangan terhadap sistem misal dengan menggunakan algoritma lain, untuk dapat membandingkan algoritma mana yang paling baik digunakan.
2. Untuk penelitian selanjutnya dapat melakukan penambahan jumlah data dan atribut.

DAFTAR PUSTAKA

- Alodia, D. A., Fialine, A. P., Endriani, D., & Widodo, E. (2021). Implementasi Metode K-Medoids Clustering untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Pendidikan. *Sepren*, 2(2), 1–13. <https://doi.org/10.36655/sepren.v2i2.606>
- Darnila, E., Ula, M., Mauliza, M., Ermatita, E., & Pahendra, I. (2018). Implementasi Sistem Pakar Pada Pasien Penderita Tuberkulosis Potential Drop Out Di Rumah Sakit Cut Meutia Aceh Utara. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 2(1), 415–419. <https://doi.org/10.30865/komik.v2i1.968>
- Darnila, E., Ula, M., Mauliza, Pahendra, I., & Ermatita. (2019). Machine Learning for Tuberculosis Classification Based on Treatment History. *Journal of Physics: Conference Series*, 1361(1). <https://doi.org/10.1088/1742-6596/1361/1/012064>
- Darnila, E., Ula, M., Mauliza, Pahendra, I., & Ermatita. (2020). Classification of treatment tuberculosis history based on machine learning techniques. *IOP Conference Series: Materials Science and Engineering*, 725(1), 1–6. <https://doi.org/10.1088/1757-899X/725/1/012102>
- Harahap, L. M., Fuadi, W., Rosnita, L., Darnila, E., & Meiyanti, R. (2022). Klastering Sayuran Unggulan Menggunakan Algoritma K-Means. *Jurnal Teknik Informatika Dan Sistem Informasi*, 8(3), 567–579. <https://doi.org/10.28932/jutisi.v8i3.5277>
- Juninda, T., Mustasim, & Andri, E. (2019). Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau. *Seminar Nasional Teknologi Informasi, Komunikasi Dan Industri*, 11(1), 42–49.
- Kadarsih, K., & Andrianto, S. (2022). Membangun Website SMA PGRI Gunung Raya Ranau Menggunakan PHP dan MYSQL. *JTIM: Jurnal Teknik Informatika Mahakarya*, 03(2), 37–44.
- Kusumah, H., Lubis, M. R., & Tambunan, H. S. (2021). Penerapan Algoritma K-Medoids Dalam Pengelompokan Imunisasi Lanjutan Pada Anak Usia 2 Tahun. *Rekayasa Teknik Informatika Dan Informasi*, 1(4), 265–273.
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 4(2), 64. <https://doi.org/10.24014/coreit.v4i2.4498>
- Nasari, F., Tanjung, D. H., & Handayani, F. (2023). Optimasi Metode K-Means dan K-Medoids Berdasarkan Jumlah Cluster dan Nilai DBI Dalam Pengelompokkan

- Produksi Kelapa Sawit Di Provinsi Riau. *InfoSys Journal*, 7(2), 129–141.
<https://www.doi.org/10.22303/infosys.7.2.2023.129-141>
- Permana, R. A., Hapsari, D. P., & Muhima, R. R. (2022). Penerapan Metode Clustering K-Medoids Untuk Pengelompokan Abstrak Skripsi Berbasis Web. *SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, Dan Teknik Informatika*, 277. <https://ejurnal.itats.ac.id/snestikdanhttps://snestik.itats.ac.id>
- Purba, J., Bahri, S., Khair, H., & Darnila, E. (2019). *The Effect of Compensation , Supervision and Competence on Performance of Tuberculosis Health Center Officers using Multiple Linear Regression*. 13(2), 6–11.
- Sari, Y. P., Primajaya, A., & Irawan, A. S. Y. (2020). Implementasi Algoritma K-Means untuk Clustering Penyebaran Tuberkulosis di Kabupaten Karawang. *INOVTEK Polbeng - Seri Informatika*, 5(2), 229.
<https://doi.org/10.35314/isi.v5i2.1457>
- Sejati, A., & Sofiana, L. (2015). Faktor-Faktor Terjadinya Tuberkulosis. *Jurnal Kesehatan Masyarakat*, 10(2), 122. <https://doi.org/10.15294/kemas.v10i2.3372>
- Simanjuntak, D. S. M., Gunawan, I., Sumarno, S., Poningsih, P., & Sari, I. P. (2023). Penerapan Algoritma K-Medoids Untuk Pengelompokkan Pengangguran Umur 25 tahun Keatas Di Sumatera Utara. *Jurnal Krisnadana*, 2(2).
<https://doi.org/10.58982/krisnadana.v2i2.264>
- Suherni, N., & Maduratna. (2013). Analisis Pengelompokan Kecamatan di Kota Surabaya Berdasarkan Faktor Penyebab Terjadinya Penyakit Tuberkulosis. *Jurnal Sains Dan Seni ITS*, 2(1), 2337–3520.
http://ejurnal2.its.ac.id/index.php/sains_seni/article/view/3083
- Sulistiyo Hidayat, F., Berliana, R., Affandi, P., Zuliana, V., & Padilah, T. N. (2022). Penerapan K-Means Clustering dalam Pengelompokan Kasus Tuberkulosis di Provinsi Jawa Barat. *Jurnal Ilmiah Wahana Pendidikan*, 8(15), 213–227.
<https://doi.org/10.5281/zenodo.7049113>
- Sutanti, A., MZ, M. K., Mustika, M., & Damayanti, P. (2020). Rancang Bangun Aplikasi Perpustakaan Keliling Menggunakan Pendekatan Terstruktur. *Komputa : Jurnal Ilmiah Komputer Dan Informatika*, 9(1), 1–8.
<https://doi.org/10.34010/komputa.v9i1.3718>
- Yunus, M., & Setyowibowo, S. (2011). Aplikasi Sistem Pendukung Keputusan Diagnosa Penyakit Paru-Paru Dengan Metode Forward Chaining. *Jurnal Teknologi Informasi*, 2(2), 95–114. <https://doi.org/10.36382/jti-tki.v2i2.130>

