

**ANALISIS SPASIAL PENYEBARAN KASUS TUBERKULOSIS  
BERDASARKAN FAKTOR SOSIAL-EKONOMI  
MENGUNAKAN DISTANCE-BASED ALGORITHM (STUDI  
KASUS KOTAMADYA MAKASSAR)**

*SPATIAL ANALYSIS OF THE SPREAD OF TUBERCULOSIS CASES  
BASED ON SOCIO-ECONOMIC FACTORS USING DISTANCE-  
BASED ALGORITHM (CASE STUDY OF MAKASSAR CITY)*

**AYU AZIZAH  
D032182008**



**PROGRAM STUDI MAGISTER TEKNIK ELEKTRO  
FAKULTAS TEKNIK  
UNIVERSITAS HASANUDDIN  
GOWA  
2023**

**ANALISIS SPASIAL PENYEBARAN KASUS TUBERKULOSIS  
BERDASARKAN FAKTOR SOSIAL-EKONOMI  
MENGUNAKAN DISTANCE-BASED ALGORITHM (STUDI  
KASUS KOTAMADYA MAKASSAR)**

Tesis

Sebagai Salah Satu Syarat untuk Mencapai Gelar Magister

Program Studi Teknik Elektro

Disusun dan diajukan oleh:

**AYU AZIZAH**

**D032182008**

Kepada

**PROGRAM STUDI MAGISTER TEKNIK ELEKTRO**

**FAKULTAS TEKNIK**

**UNIVERSITAS HASANUDDIN**

**GOWA**

**2023**

**TESIS**  
**ANALISIS SPASIAL PENYEBARAN KASUS TUBERKULOSIS**  
**BERDASARKAN FAKTOR SOSIAL-EKONOMI**  
**MENGGUNAKAN DISTANCE-BASED ALGORITHM (STUDI**  
**KASUS KOTAMADYA MAKASSAR)**

**AYU AZIZAH**  
**D032 182 008**

Telah dipertahankan di hadapan Panitia Ujian Tesis yang dibentuk dalam rangka penyelesaian studi pada Program Magister Teknik Elektro, Fakultas Teknik

Universitas Hasanuddin  
pada tanggal 19 Januari 2023  
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama



**Dr. Amil Ahmad Ilham, S.T., M.IT**  
NIP. 19731010 199802 1 001

Pembimbing Pendamping



**Prof. Dr. Eng. Ir. Syafaruddin, S.T., M.Eng. IPU**  
NIP. 19740530 199903 1 003

Dekan Fakultas Teknik  
Universitas Hasanuddin



**Prof. Dr. Eng. Ir. Muhammad Isran Ramli, S.T., M.T**  
NIP. 19730926 2000 121 002

Ketua Program Studi  
S2 Teknik Elektro



**Dr. Eng. Ir. Wardi, S.T., M.Eng**  
NIP. 19720828 199903 1 003

## **PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA**

Yang bertanda tangan di bawah ini:

Nama : Ayu Azizah  
Nomor Mahasiswa : D032182008  
Program Studi : S2 Teknik Elektro

Dengan ini menyatakan bahwa, tesis yang saya tulis berjudul “Analisis Spasial Penyebaran Kasus Tuberkulosis Berdasarkan Faktor Sosial-Ekonomi Menggunakan *Distance-Based Algorithm* (Studi Kasus Kotamadya Makassar)” adalah benar-benar merupakan hasil karya tulisan saya dengan arahan dari komisi pembimbing (Dr. Amil Ahmad Ilham, S.T., M.IT sebagai Pembimbing Utama dan Prof. Dr. Eng. Syafaruddin, S.T., M.Eng., IPU sebagai Pembimbing Pendamping). Karya ilmiah ini belum diajukan dan sedang tidak diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Jurnal/Prosiding (*2022 International Conference on Electrical Engineering and Informatics (ICELTICs)*) sebagai artikel dengan judul “*Spatial Analysis of the Spread of Tuberculosis Cases Based on Socio-Economic Factors Using Distance-Based Algorithm*”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Gowa, 19 Januari 2023

Yang menyatakan

  
D908AAKX349162896

Ayu Azizah

## KATA PENGANTAR

Alhamdulillah rabbil'alamin, segala puji bagi Allah Subhanahu Wa Ta'ala Yang Maha Sempurna, yang telah memberikan rahmat, hidayah dan pertolongannya sehingga penulis dapat menyelesaikan tesis dengan judul “**Analisis Spasial Penyebaran Kasus Tuberkulosis Berdasarkan Faktor Sosial-Ekonomi Menggunakan *Distance-Based Algorithm* (Studi Kasus Kotamadya Makassar)**”. Tak lupa pula shalawat dan salam kepada Nabi Muhammad SAW yang telah menyinari dunia ini dengan keindahan ilmu dan akhlak yang diajarkan kepada seluruh umatnya.

Tesis ini disusun untuk memenuhi persyaratan untuk memperoleh gelar Magister Teknik (M.T.) pada Program Pascasarjana Departemen Teknik Elektro, Konsentrasi Teknik Informatika Universitas Hasanuddin Makassar. Tentunya penyelesaian tesis ini tidak terlepas dari dukungan dan bantuan dari semua pihak. Untuk itu, dengan penuh kerendahan hati penulis menyampaikan terima kasih setulus-tulusnya dan setinggi-tingginya kepada yang tercinta Ibu **Andi Asma Raju** dan Ayah **Yunet Elvis, S.E** beserta Suami tersayang **Muhammad Ikramullah, S.Kom** atas segala doa, kasih sayang, serta dukungan moril maupun materil kepada penulis dalam menyelesaikan penelitian ini. Penulis juga menghaturkan terima kasih yang sebar-besarnya kepada Bapak **Dr. Amil Ahmad Ilham, S.T., M.IT** sebagai pembimbing utama dan Bapak **Prof. Dr. Eng. Ir. Syafaruddin, S.T., M.Eng. IPU** selaku dosen pembimbing pendamping yang telah meluangkan waktunya kepada penulis untuk membimbing, memberikan masukan, memotivasi tiada henti-hentinya hingga tahap penyelesaian tesis ini. Serta tim penguji yang banyak memberikan masukan dan arahan untuk kesempurnaan tesis ini. Adapun penulis sampaikan ucapan terima kasih kepada rekan-rekan Mahasiswa S2 Departemen Teknik Elektro angkatan 2018 dan 2019 yang selalu mendukung dalam proses penyelesaian tesis ini. Serta semua pihak manapun yang telah membantu dalam penyusunan tesis ini, yang tidak bisa saya sebutkan satu persatu.

Penulis menyadari bahwa penulisan tesis ini masih jauh dari kata sempurna dan di dalam penyelesaiannya masih menemui kesulitan dan hambatan, sehingga penulis tetap mengharapkan saran dan kritik untuk pengembangan lebih lanjut. Harapan penulis, semoga tesis ini dapat memberikan manfaat yang banyak bagi semua pihak.

Gowa, 19 Januari 2023

Penulis

## ABSTRAK

**Ayu Azizah.** Analisis Spasial Penyebaran Kasus Tuberkulosis Berdasarkan Faktor Sosial-Ekonomi Menggunakan *Distance-Based Algorithm* (Studi Kasus Kotamadya Makassar). (dibimbing oleh **Amil Ahmad Ilham** dan **Syafaruddin**).

Tuberkulosis merupakan penyakit infeksi menular yang merupakan salah satu dari sepuluh penyakit mematikan di dunia. Indonesia sendiri merupakan negara tertinggi ketiga dalam kasus penyebaran TBC. Oleh karena itu, diperlukan pengawasan yang komprehensif dengan mengelompokkan kasus tuberkulosis berdasarkan wilayah. Tujuannya untuk mendapatkan hasil analisis spasial pemetaan tuberkulosis berdasarkan faktor usia, jenis kelamin, dan sosial ekonomi dengan memanfaatkan data mining dan algoritma *K-Means clustering*. Kemudian, pola sebaran spasial tersebut divisualisasikan menggunakan sistem informasi geografis (SIG). Data kasus tuberkulosis yang digunakan adalah data sekunder yang diperoleh dari Balai Kesehatan Paru Masyarakat (BBKPM) Makassar selama tiga tahun terakhir (2018 - 2020). Dari hasil analisis diketahui bahwa kasus tuberkulosis tertinggi didominasi oleh laki-laki lanjut usia. Sedangkan penyebaran tuberkulosis tertinggi berada di Kecamatan Panakukang yang didominasi oleh laki-laki dewasa. Hasil penelitian ini dapat digunakan untuk membantu dalam pengambilan keputusan tentang perencanaan kasus kontrol tuberkulosis di Kota Makassar.

**Kata Kunci:** Analisis Spasial, Tuberkulosis, *K-Means Clustering*, *Geographic Information System*

## ABSTRACT

**Ayu Azizah.** Spatial Analysis Of The Spread Of Tuberculosis Cases Based on Socio-Economic Factors Using Distance-Based Algorithm (Case Study Of Makassar City). (supervised by **Amil Ahmad Ilham** and **Syafaruddin**).

Tuberculosis is a contagious infectious disease that is one of the ten deadly diseases in the world. Indonesia itself is the third highest country in the case of the spread of tuberculosis. Therefore, comprehensive supervision is needed by grouping tuberculosis cases by region. The aim is to obtain the results of the spatial analysis of tuberculosis mapping based on age, gender, and socio-economic factors by utilizing data mining and the K-Means *clustering* algorithm. Then, the spatial distribution pattern is visualized using a geographic information system (GIS). The tuberculosis case data used is secondary data obtained from the Makassar Community Lung Health Center (BBKPM) for the last three years (2018 - 2020). From the results of the analysis, it was found that the highest tuberculosis cases were dominated by elderly men. Meanwhile, the highest spread of tuberculosis was in the Panakukang sub-district which was dominated by adult males. The results of this study can be used to assist in making decisions about tuberculosis case-control planning in Makassar City.

**Keywords:** Spatial Analysis, Tuberculosis, K-Means *Clustering*, Geographic Information System



## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>PENGAJUAN TESIS.....</b>	<b>ii</b>
<b>PERSETUJUAN TESIS .....</b>	<b>iii</b>
<b>PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA ....</b>	<b>iv</b>
<b>KATA PENGANTAR.....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vii</b>
<b>ABSTRACT .....</b>	<b>viii</b>
<b>DAFTAR ISI.....</b>	<b>ix</b>
<b>DAFTAR TABEL .....</b>	<b>xi</b>
<b>DAFTAR GAMBAR.....</b>	<b>xii</b>
<b>DAFTAR LAMPIRAN.....</b>	<b>xiii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
A. Latar Belakang .....	1
B. Rumusan Masalah.....	3
C. Tujuan Penelitian .....	3
D. Manfaat Penelitian .....	4
E. Batasan Masalah .....	4
F. Sistematika Penulisan .....	4
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>7</b>
A. Landasan Teori.....	7
1. Tuberkulosis.....	7

2. Data Mining .....	9
3. Algoritma <i>K-Means</i> .....	11
4. <i>Clustering</i> .....	12
5. Sistem Informasi Geografis .....	13
B. Penelitian Terkait .....	15
C. Kerangka Pikir .....	20
<b>BAB III METODE PENELITIAN .....</b>	<b>22</b>
A. Tahapan Penelitian.....	22
B. Waktu Dan Lokasi Penelitian .....	23
C. Jenis Penelitian.....	24
D. Instrumen Penelitian .....	24
E. Sumber Data.....	25
F. Sampel Data .....	25
G. Rancangan Sistem.....	28
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>32</b>
A. Hasil <i>Clustering</i> Menggunakan Algoritma <i>K-Means</i> .....	32
B. Hasil Analisis Pola Pesebaran.....	43
C. Hasil Analisis Spasial .....	45
<b>BAB V PENUTUP.....</b>	<b>48</b>
A. Kesimpulan .....	48
B. Saran .....	49
<b>DAFTAR PUSTAKA .....</b>	<b>50</b>
<b>LAMPIRAN.....</b>	<b>53</b>

## DAFTAR TABEL

Nomor	Halaman
<b>Tabel 1.</b> <i>State Of The Art</i>	18
<b>Tabel 2.</b> Rincian Jumlah Data Pada Kategori Jenis Kelamin Dan Jenis Pembayaran Pelayanan	26
<b>Tabel 3.</b> Data Sampel	27
<b>Tabel 4.</b> Dataset Awal	33
<b>Tabel 5.</b> Hasil Normalisasi Data	34
<b>Tabel 6.</b> Keanggotaan <i>Centroid</i>	38
<b>Tabel 7.</b> Keanggotaan Tiap Klaster	41
<b>Tabel 8.</b> Penderita Tuberkulosis Disetiap Kecamatan	47

## DAFTAR GAMBAR

Nomor	Halaman
<b>Gambar 1.</b> Faktor Risiko Penyebaran Tuberkulosis	8
<b>Gambar 2.</b> Proses Data Mining	10
<b>Gambar 3.</b> Subsistem-subsistem SIG	15
<b>Gambar 4.</b> Kerangka Pikir Penelitian	21
<b>Gambar 5.</b> Tahapan Penelitian	22
<b>Gambar 6.</b> Batas Wilayah Pengambilan Data	24
<b>Gambar 7.</b> Blok Diagram	28
<b>Gambar 8.</b> Flowchart Algoritma <i>K-Means Clustering</i>	29
<b>Gambar 9.</b> Proses Normalisasi Data	35
<b>Gambar 10.</b> Menghitung Nilai SSE	36
<b>Gambar 11.</b> Garifik Hubungan Nilai SSE ddengan $K = 1$ s.d. $K = 10$	36
<b>Gambar 12.</b> Proses Pengolompokan dengan <i>K-Means</i>	40
<b>Gambar 13.</b> Hasil Pengelompokan <i>K-Means</i> ( $K = 4$ )	40
<b>Gambar 14.</b> Presentase Jumlah Anggota Setiap Klaster	41
<b>Gambar 15.</b> Peta Pola Penyebaran Penderita Tuberkulosis	43
<b>Gambar 16.</b> Peta Pola Penyebaran Penderita Tuberkulosis Di Kecamatan Biringkanaya	44
<b>Gambar 17.</b> Menampilkan Informasi Penderita Tuberkulosis	45
<b>Gambar 18.</b> Peta Spasial Penyebaran Tuberkulosis	45
<b>Gambar 19.</b> Peta Spasial Penyebaran Tuberkulosis Di Kecamatan Panakukang	46
<b>Gambar 20.</b> Statistik Penderita Tuberkulosis Setiap Kecamatan	46

## DAFTAR LAMPIRAN

Nomor	Halaman
<b>Lampiran 1.</b> Tabel Data Penderita Tuberkulosis	53
<b>Lampiran 2.</b> <i>Source Code K-Means Clustering Algorithm</i>	97
<b>Lampiran 3.</b> <i>Source Code Mapping Penyebaran Tuberkulosis</i>	101

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang**

Tuberkulosis merupakan penyakit infeksi menular yang disebabkan oleh *Mycobacterium Tuberculosis* yang sebagian besar menyerang paru, tetapi dapat juga mengenai berbagai organ atau jaringan tubuh lainnya. Mikobakterium ini ditransmisikan melalui droplet di udara, sehingga seorang penderita tuberkulosis merupakan sumber penyebab penularan tuberkulosis pada populasi di sekitarnya. Sampai saat ini penyakit tuberkulosis masih menjadi masalah kesehatan yang utama, baik di dunia maupun di Indonesia (Departemen Kesehatan RI, 2008).

Berdasarkan data dari WHO (2020) Indonesia menjadi negara dengan kasus tuberkulosis tertinggi ketiga di Asia setelah China dan India. Kebanyakan kasus tuberkulosis terjadi pada negara-negara yang berkembang, serta negara yang memiliki kepadatan penduduk yang tinggi. Setiap tahun estimasi kasus tuberkulosis di Indonesia mencapai 842.000 penderita (data per 27 Maret 2019).

Makassar, salah satu kota terbesar keempat di Indonesia yang termasuk daerah penyebaran tuberkulosis tertinggi. Dengan jumlah penduduk 1.536.600 orang, tercatat pada tahun 2020 ditemukan 3.255 kasus tuberkulosis yang merupakan daerah kasus tertinggi pertama di Provinsi Sulawesi Selatan (Aswi, 2021).

Menurut H.L Blum, ada 4 faktor yang dapat mempengaruhi derajat kesehatan baik individu, kelompok, maupun masyarakat yaitu lingkungan, perilaku, pelayanan kesehatan dan keturunan. Keempat faktor tersebut saling berkaitan dalam mempengaruhi kesehatan (Notoatmodjo, 2010).

Pada penelitian Hiswani 2009, mengatakan bahwa penyebaran tuberkulosis pada seseorang dipengaruhi oleh beberapa faktor seperti: status sosial ekonomi (kondisi rumah, kepadatan hunian, lingkungan perumahan, lingkungan dan sanitasi tempat kerja, dan pendapatan keluarga), status gizi (keadaan malnutrisi), umur, jenis kelamin dan faktor sosial lainnya.

Faktor pertama yang mempengaruhi penyebaran tuberkulosis adalah faktor umur, karena insiden tertinggi penyakit tuberkulosis terdapat pada usia dewasa muda di Indonesia, yang diperkirakan sekitar 75% penderita tuberkulosis adalah kelompok usia produktif. Faktor kedua adalah jenis kelamin yang lebih banyak menyerang laki-laki dari pada wanita, hal ini disebabkan karena sebagian besar dari laki-laki mempunyai kebiasaan merokok (Aswi, 2021).

Dalam penelitian Wulandari 2012, diperoleh hasil bahwa terdapat hubungan yang signifikan antara kepadatan penduduk dengan jumlah kasus baru tuberkulosis paru BTA+ (Basil Tahan Asam). Hal tersebut dapat terjadi karena jika kepadatan penduduknya tinggi maka peluang kontak terhadap penderita tuberkulosis lebih besar. Demikian pula dengan faktor rumah sehat, dapat mempengaruhi terjadinya tuberkulosis paru.

Hal tersebut sesuai dengan penelitian yang dilakukan oleh Mahmuda et al (2014), diperoleh hasil bahwa orang yang memiliki rumah tidak sehat pada status ekonomi rendah memiliki peluang lebih besar untuk menderita tuberkulosis paru dibandingkan dengan orang yang memiliki rumah sehat. Di Kabupaten Batang Provinsi Banten pada tahun 2010 distribusi cakupan rumah sehatpun tidak merata, terdapat kecamatan dengan cakupan rumah sehat mencapai 94,7% dengan tingkat penyebaran tuberkulosis rendah dan terdapat pula kecamatan dengan cakupan rumah sehat yang hanya 9,3% dengan tingkat penyebaran tuberkulosis tinggi.

Sekitar 75% penderita tuberkulosis adalah kelompok usia produktif secara ekonomis (15-50 tahun). Diperkirakan seorang penderita tuberkulosis dewasa akan kehilangan rata-rata waktu kerjanya 3 sampai 4 bulan, hal tersebut berakibat pada kehilangan pendapatan tahunan rumah tangganya sekitar 20-30%. Sehingga kerugian yang diakibatkan oleh penyakit tuberkulosis bukan hanya dari aspek kesehatan semata tetapi juga dari aspek sosial ekonomi. Karenanya perang terhadap penyakit tuberkulosis berarti pula perang terhadap kemiskinan, ketidak produktifan dan kelemahan akibat tuberkulosis (Ruswanto, 2010).

Selain itu, salah satu alat epidemiologi yang paling penting untuk mengidentifikasi wilayah dengan risiko tinggi tuberkulosis adalah Sistem Informasi Geografis (SIG) (Purwoko, 2020). Pengelolaan perawatan, pencegahan dan

pengendalian penyakit termasuk tuberkulosis, semuanya ditangani oleh aplikasi Sistem Informasi Geografis (SIG) ini.

Seperti pada penelitian Resha (2019) menggunakan metode data mining dan *special scan statistic* serta Sistem Informasi Geografis (SIG) untuk melakukan pemetaan pola penyebaran kasus tuberkulosis di Kota Makassar berdasarkan teknik spatio-temporal. Pemetaan yang dilakukan pada penelitian ini berdasarkan jumlah penderita tuberkulosis perwilayah tanpa memperhatikan faktor jumlah penduduk dan keadaan sosial ekonomi setiap wilayah.

Dengan mengacu pada kajian literatur tersebut, penelitian ini bertujuan untuk mengelompokkan penderita tuberkulosis di setiap kecamatan di Kota Makassar berdasarkan jumlah penderita dan faktor sosial ekonomi yang mempengaruhi penyebaran tuberkulosis dengan menggunakan metode data mining dan *K-Means clustering* dengan visualisasi spasial pola penyebaran tuberkulosis menggunakan Sistem Informasi Geografis (GIS).

## **B. Rumusan Masalah**

Berdasarkan pada latar belakang di atas, maka rumusan masalah pada penelitian adalah:

1. Bagaimana mengelompokkan penderita tuberkulosis berdasarkan faktor sosial ekonomi yang mempengaruhi penyebaran tuberkulosis dengan menggunakan parameter *clustering* yang tepat?
2. Bagaimana memvisualisasikan pola penyebaran tuberkulosis yang mudah dipahami dalam bentuk spasial?

## **C. Tujuan Penelitian**

Berdasarkan rumusan masalah maka tujuan yang akan dicapai dari penelitian ini adalah:

1. Mengelompokkan penderita tuberkulosis berdasarkan faktor sosial ekonomi yang mempengaruhi penyebaran tuberkulosis dengan menggunakan parameter *clustering* yang tepat.



2. Menvisualisasikan pola penyebaran tuberkulosis yang mudah dipahami dalam bentuk spasial.

#### **D. Manfaat Penelitian**

Adapun manfaat yang dapat diperoleh dari penelitian ini adalah menambah pengetahuan ataupun kemampuan mengenai proses clustering dan analisis spasial pola penyebaran tuberkulosis. Dan bagi Dinas Kesehatan dapat digunakan sebagai referensi dalam hal penanggulangan tuberkulosis di setiap daerah, termasuk mempromosikan kebiasaan hidup sehat, program imunisasi, kunjungan rumah dan mengoptimalkan kegiatan skrining tuberkulosis setiap daerah terutama yang memiliki tingkat penyebaran penyakit tuberkulosis yang tinggi.

#### **E. Batasan Masalah**

Adapun batasan masalah dalam penelitian ini adalah sistem yang dibuat hanya untuk mengelompokkan dan memvisualisasikan pola penyebaran tuberkulosis berdasarkan faktor sosial ekonomi yaitu umur, jenis kelamin dan alamat domisili penderita tuberkulosis.

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Balai Besar Kesehatan Paru Masyarakat (BBKPM) Makassar mulai dari tahun 2018 sampai dengan tahun 2020.

#### **F. Sistematika Penulisan**

Adapun sistematika penulisan pada penelitian ini adalah:

##### **Bab I Pendahuluan**

Pada bab I ini berisi penjelasan tentang latar belakang yang menjabarkan alasan dilakukannya penelitian terkait *clustering* dan analisis spasial penyebaran tuberkulosis berdasarkan peluang penelitian dan uraian penelitian awal, terkait rumusan masalah, tujuan, manfaat, ruang lingkup serta sistematika penulisan penelitian dibahas pada bagian ini.

## **Bab II Tinjauan Pustaka**

Pada bab II ini berisi penjelasan tentang landasan teori yang digunakan dalam penelitian meliputi tuberkulosis, *K-Means clustering*, Sistem Informasi Geografis (GIS) dan beberapa landasan teori lainnya. Diuraikan pula tentang tinjauan pustaka yang merupakan penjelasan tentang hasil-hasil penelitian sebelumnya yang berkaitan dengan penelitian yang dilakukan. Landasan teori merupakan suatu penjelasan tentang sumber acuan terbaru dari pustaka *primer* seperti buku, artikel, jurnal, prosiding dan tulisan asli lainnya untuk mengetahui perkembangan penelitian yang relevan dengan judul atau tema penelitian yang dilakukan dan juga sebagai arahan dalam memecahkan masalah yang diteliti. Dalam bab ini juga diuraikan tentang kerangka pemikiran yang merupakan penjelasan tentang kerangka berpikir untuk memecahkan masalah yang sedang diteliti, termasuk menguraikan objek penelitian serta *state of the art* dari beberapa penelitian terkait.

## **Bab III Metodologi Penelitian**

Pada bab III ini merupakan penjelasan tentang tahapan-tahapan yang dilakukan di dalam penelitian ini berdasarkan *Research & Development (R&D)*, serta waktu dan lokasi penelitian yang dilakukan selama  $\pm 2$  tahun. Adapun penjelasan tentang jenis penelitian ini yang bersifat analisis dengan *studi literature* yang ada (*Library Research*) dengan didukung oleh beberapa *instrument* penelitian. Pembahasan tentang sumber pengambilan data dijelaskan pada bab ini, serta sampel data yang digunakan juga dijelaskan pada bagian ini berupa jumlah data maupun jenis data yang digunakan sebagai sampel dalam penelitian ini. Pada bab ini juga terdapat rancangan sistem yang dirangkum pada setiap proses dalam bentuk *flowchart* pada setiap tahapan yang ada.

## **Bab IV Hasil dan Pembahasan**

Pada bab IV berisi penjabaran dari hasil penelitian yang telah diperoleh. Hasil dan pembahasan pada penelitian terbagi atas dua yaitu *clustering* dan analisis spasial pola penyebaran tuberkulosis yang dirangkum dalam bentuk grafik, gambar dan tabel yang tersusun dalam poin-poin pembahasan.

## **Bab V Kesimpulan dan Saran**

Pada bab V berisi kesimpulan terhadap hasil yang didapatkan dalam penelitian ini, yang merujuk pada rumusan masalah. Saran pengembangan dari penelitian ini untuk menyempurnakan kekurangan-kekurangan atau capaian-capaian yang belum tercapai pada penelitian ini, sehingga kedepannya penelitian yang dilakukan dapat dikembangkan dan bisa diperoleh hasil yang jauh lebih baik.

## BAB II

### TINJAUAN PUSTAKA

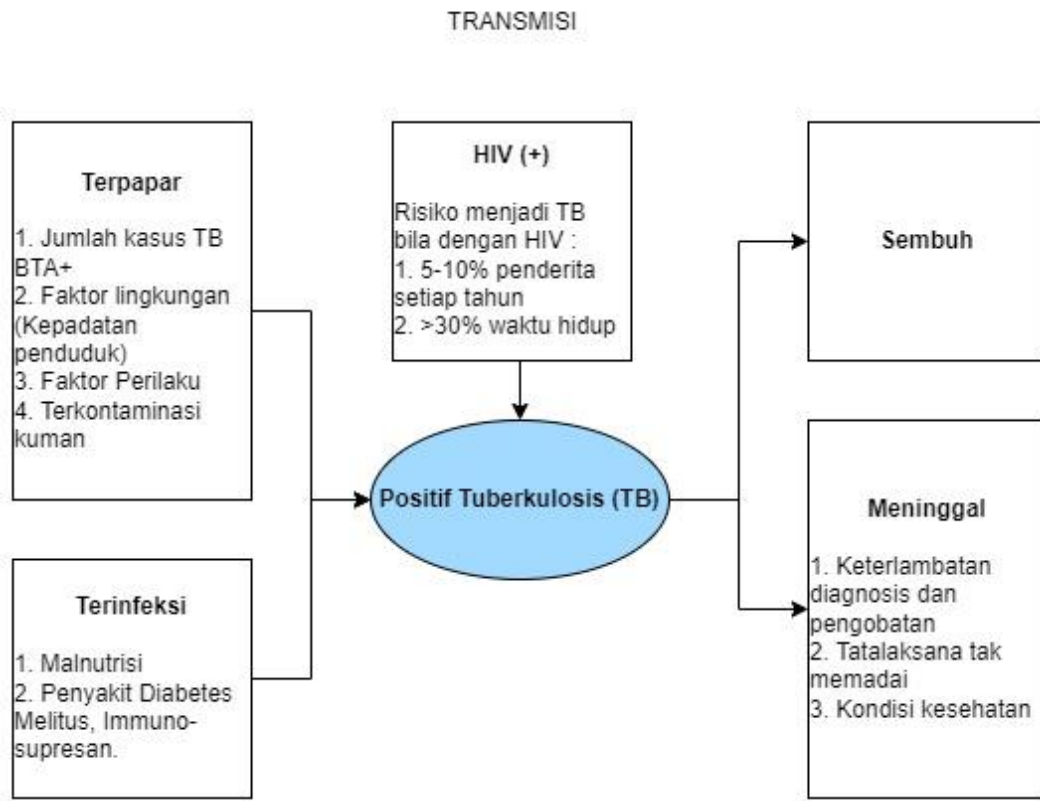
#### A. Landasan Teori

##### 1. Tuberkulosis

Tuberkulosis (TB) merupakan penyakit infeksi menular yang disebabkan oleh *Mycobacterium Tuberculosis* yang sebagian besar menyerang paru, tetapi dapat juga mengenai berbagai organ atau jaringan tubuh lainnya. Mikobakterium ini ditransmisikan melalui droplet di udara, sehingga seorang penderita tuberkulosis merupakan sumber penyebab penularan tuberkulosis pada populasi di sekitarnya. (Departemen Kesehatan RI, 2008).

Terdapat beberapa faktor risiko penyebaran tuberkulosis:

- a. Sekitar 10% yang terinfeksi tuberkulosis akan menjadi sakit tuberkulosis.
- b. Berdasarkan *Annual Risk of Tuberculosis Infection* (ARTI) diperkirakan 1% diantara 100.000 penduduk (1.000 orang) rata-rata terinfeksi tuberkulosis. Sekitar 10% diantaranya (100 orang) akan terinfeksi tuberkulosis setiap tahunnya dan 50 orang diantaranya adalah pasien tuberkulosis BTA+.
- c. Salah satu faktor yang kemungkinan mempengaruhi seseorang menjadi pasien tuberkulosis adalah daya tahan tubuh yang rendah, diantaranya infeksi HIV/AIDS dan malnutrisi (gizi buruk).
- d. Infeksi HIV mengakibatkan kerusakan luas sistem daya tahan tubuh seluler (*cellular immunity*) dan merupakan faktor risiko seseorang dapat terinfeksi tuberkulosis untuk menjadi sakit tuberkulosis (TB Aktif). Bila jumlah orang terinfeksi HIV meningkat, maka jumlah pasien tuberkulosis akan meningkat, dengan demikian penularan tuberkulosis di masyarakat akan meningkat pula (Kemenkes RI, 2011).



**Gambar 1.** Faktor Risiko Penyebaran Tuberkulosis (Kemenkes RI, 2011)

Penyebab utama meningkatnya penyebaran tuberkulosis antara lain adalah:

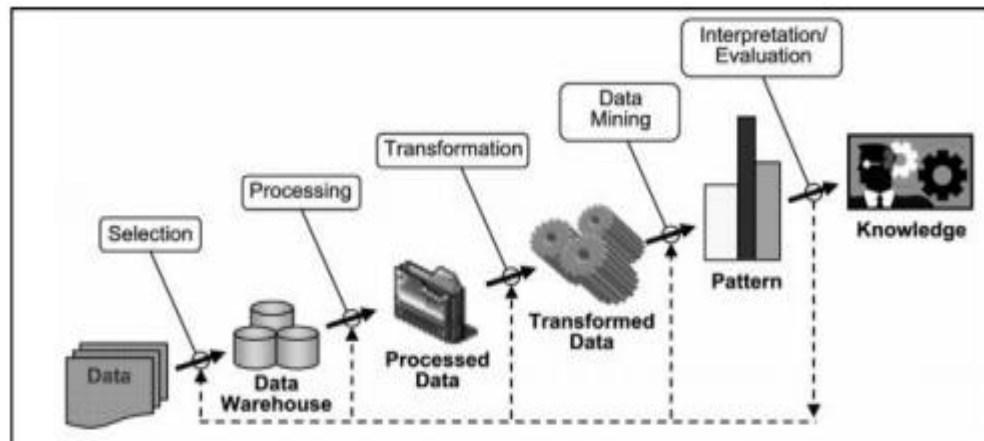
- a. Kemiskinan pada berbagai kelompok masyarakat, seperti pada negara-negara yang sedang berkembang.
- b. Kegagalan program tuberkulosis selama ini. Hal ini diakibatkan oleh:
  - Tidak memadainya komitmen politik dan pendanaan.
  - Tidak memadainya organisasi pelayanan tuberkulosis (kurang terakses oleh masyarakat, penemuan kasus /diagnosis yang tidak standar, obat tidak terjamin penyediaannya, tidak dilakukan pemantauan, pencatatan dan pelaporan yang standar, dan sebagainya).
  - Tidak memadainya tatalaksana kasus (diagnosis dan paduan obat yang tidak standar, gagal menyembuhkan kasus yang telah didiagnosis).

- Salah persepsi terhadap manfaat dan efektifitas BCG.
  - Infrastruktur kesehatan yang buruk pada negara-negara yang mengalami krisis ekonomi atau pergolakan masyarakat.
- c. Perubahan demografik karena meningkatnya penduduk dunia dan perubahan struktur umur kependudukan.
  - d. Dampak pandemi HIV (Kemenkes RI, 2011).

## 2. Data Mining

Data Mining adalah salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database dengan skala besar. Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. Data Mining memiliki hubungan dari bidang ilmu seperti artificial intelligent, machine learning, statistik dan database. Beberapa teknik Data Mining antara lain: *clustering, classification, association rule mining, neural network, genetic algorithm* dan lain-lain (Aradea, 2011).

Teknik data mining digunakan untuk memeriksa basis data berukuran besar sebagai cara untuk menemukan pola yang baru dan berguna. Pada *data mining* proses untuk menggali pengetahuan baru dan berguna dari sekumpulan data yang besar sering juga dikenal sebagai penemuan pengetahuan dari pangkalan data (*Knowledge Discovery in Databases*) / KDD yaitu tahap-tahap yang dilakukan dalam menggali pengetahuan dari sekumpulan data. Tahap-tahap yang dimaksud digambarkan seperti Gambar 2.



**Gambar 2.** Proses Data Mining (Aradea, 2011)

a. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining. Disimpan dalam suatu berkas, terpisah dari basis data operasional.

- b. *Pre-processing* Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD. *Pre-Processing* dilakukan untuk membuang data yang tidak konsisten dan noise, duplikasi data, memperbaiki kesalahan data dan boleh juga diperkaya dengan data eksternal yang relevan. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

c. *Transformation*

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

d. Data Mining

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

e. *Interpretation/Evaluation*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation* (Aradea, 2011)

### 3. Algoritma *K-Means*

Algoritma *K-Means* merupakan salah satu metode data *clustering* non hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*/kelompok dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster*/kelompok yang lain sehingga data yang berada dalam satu *cluster*/kelompok memiliki tingkat variasi yang kecil (Agusta, 2007).

Menurut Santosa (2007), langkah-langkah melakukan *clustering* dengan metode *K-Means* adalah sebagai berikut:

- a. Pilih jumlah *cluster*  $k$ .
- b. Inisialisasi  $k$  pusat *cluster* ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat *cluster* diberi nilai awal dengan angka-angka random
- c. Alokasikan semua data/ objek ke *cluster* terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster*. Jarak terdekat antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster*



mana. Untuk menghitung jarak semua data ke setiap titik pusat *cluster* dapat menggunakan teori jarak Euclidean.

- d. Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari semua data/ objek dalam *cluster* tertentu. Jika dikehendaki bisa juga menggunakan median dari *cluster* tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa digunakan.
- e. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru. Jika pusat *cluster* tidak berubah lagi maka proses *clustering* selesai. Atau, kembali ke langkah nomor 3 sampai pusat *cluster* tidak berubah lagi.

#### 4. *Clustering*

Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (similarity) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode data mining yang bersifat tanpa arahan (unsupervised), maksudnya metode ini diterapkan tanpa adanya latihan (training) dan tanpa ada guru (teacher) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode *clustering* yang digunakan dalam pengelompokan data, yaitu hierarchical *clustering* dan non-hierarchical *clustering* (Santosa, 2007).

*Hierarchical clustering* adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga *cluster* akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut (Santoso, 2010).

Berbeda dengan metode hierarchical *clustering*, metode non-hierarchical *clustering* justru dimulai dengan menentukan terlebih dahulu jumlah *cluster*

yang diinginkan (dua *cluster*, tiga *cluster*, atau lain sebagainya). Setelah jumlah *cluster* diketahui, baru proses *clustering* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means clustering* (Santoso, 2010).

Analisis *Clustering* memiliki beberapa kelebihan dan juga kekurangan sebagai berikut:

a. Kelebihan

- 1) Dapat mengelompokkan data observasi dalam jumlah besar dan variabel yang relatif banyak. Data yang di reduksi dengan kelompok akan mudah dianalisis.
- 2) Dapat dipakai dalam skala data ordinal, interval dan rasio.

b. Kelemahan

- 1) Pengelompokan bersifat subjektifitas peneliti karena hanya melihat dari gambar dendogram.
- 2) Untuk data heterogen antara objek penelitian yang satu dengan yang lain akan sulit bagi peneliti untuk menentukan jumlah kelompok yang di bentuk.
- 3) Metode-metode yang dipakai memberikan perbedaan yang signifikan, sehingga dalam perhitungan biasanya masing-masing metode dibandingkan.
- 4) Semakin besar observasi, biasanya tingkat kesalahan akan semakin besar.

## 5. Sistem Informasi Geografis (SIG)

Istilah geografi digunakan karena SIG dibangun berdasarkan pada geografi atau spasial. Objek ini mengarah pada spesifikasi lokasi dalam suatu space. Sistem informasi geografis (SIG) merupakan sistem komputer yang berbasis pada sistem informasi yang digunakan untuk memberikan bentuk digital dan analisis terhadap permukaan geografi bumi.

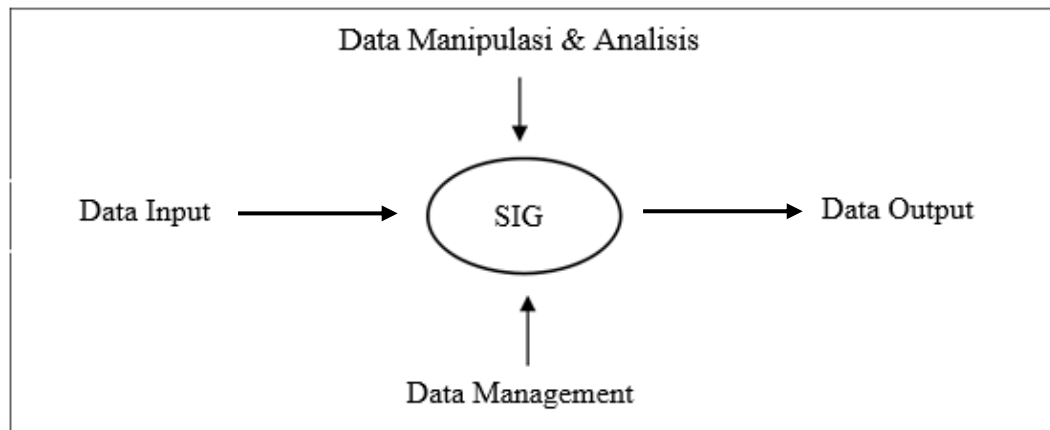
Geografi adalah informasi mengenal permukaan bumi dan semua obyek yang berada di atasnya, sedangkan sistem informasi geografis (SIG) adalah sistem informasi khusus yang mengelola data yang memiliki informasi spasial

(berefrensi keruangan). Sistem informasi geografis adalah bentuk sistem informasi yang menyajikan informasi dalam bentuk grafis dengan menggunakan peta sebagai antar muka. SIG tersusun atas konsep beberapa lapisan (layer) dan relasi (Eka, 2011)

Berikut adalah beberapa definisi SIG:

- a. Kang-Tsung Chang (2002), mendefinisikan SIG sebagai: is an a computer system for capturing, storing, querying, analyzing, and displaying geographic data.
- b. Arronoff (1989), mendefinisikan SIG sebagai suatu sistem berbasis komputer yang memiliki kemampuan dalam menangani data bereferensi geografi yaitu pemasukan data, manajemen data (penyimpanan dan pemanggilan kembali), manipulasi dan analisis data, serta keluaran sebagai hasil akhir (output). Hasil akhir (output) dapat dijadikan acuan dalam pengambilan keputusan pada masalah yang berhubungan dengan geografi Arronoff (1989).
- c. Menurut Gistut (1994), SIG adalah sistem yang dapat mendukung pengambilan keputusan spasial dan mampu mengintegrasikan deskripsi lokasi dengan karakteristik-karakteristik fenomena yang ditemukan di lokasi tersebut. SIG yang lengkap mencakup metodologi dan teknologi yang diperlukan yaitu data spasial perangkat keras, perangkat lunak dan struktur organisasi Gistut .
- d. (Burrough,1986) mendefinisikan SIG adalah sistem berbasis komputer yang digunakan untuk memasukan, menyimpan, mengelola, menganalisis dan mengaktifkan kembali data yang mempunyai referensi keruangan untuk berbagai tujuan yang berkaitan dengan pemetaan dan perencanaan.

Dari defenisi-definisi diatas dapat diambil kesimpulan bahwa SIG terdiri atas beberapa subsistem yaitu: data input, data output, data management, data manipulasi dan analisis (Prahasta, 2005).



**Gambar 3.** Subsistem-subsistem SIG (Prahasta, 2005)

## B. Penelitian Terkait

Beberapa penelitian terkait pengelompokan penderita tuberkulosis berdasarkan faktor - faktor yang mempengaruhi penyebaran tuberkulosis dengan beberapa metode berbeda seperti pada beberapa penelitian:

1. Shaweno et al, pada tahun 2018 melakukan *systematic review* dari 168 penelitian mengenai metode yang digunakan dalam analisis spasial epidemiologi tuberkulosis. Penelitian ini bertujuan untuk mengetahui pendekatan metode yang dapat digunakan dalam analisis spasial terkhusus pada kasus tuberkulosis. Hasil dari penelitian ini, yang berkaitan dengan penelitian yang akan dilakukan ialah, bahwa dari 131 penelitian dengan kategori metode spatial *cluster analysis* terdapat 48 penelitian yang menggunakan metode kulldorff's spatial scan statistic sebagai metode yang akurat dalam melakukan analisis spasial penyebaran kasus tuberkulosis (Shaweno, 2018).
2. Resha et al, pada tahun 2019 melakukan penelitian pemetaan pola penyebaran kasus tuberkulosis dengan menggunakan pendekatan spatio-temporal. Tujuan penelitian ini untuk mengetahui tingkat daerah resiko tertinggi dan terendah penyebaran tuberkulosis pada kurun waktu 3 tahun (2015-2017) dengan jumlah 15 kecamatan dan jumlah penderita 5.896 pasien. Dari hasil distribusi terbesar penyebaran tuberkulosis

berada pada kecamatan Panakukang dan terkecil berada pada kabupaten Wajo. Dengan rata-rata usia penderita 42 tahun, sementara usia penderita 50 tahun keatas berjumlah 816 pasien.

3. Wardani et al, pada tahun 2019 melakukan peneliti tentang pengklasifikasian kasus penderita tuberkulosis pada anak-anak di kota Semarang pada tahun 2014-2017 dengan jumlah penderita sebanyak 1.965 kasus. Metode yang digunakan untuk mengelompokkan pola spasial distribusi tuberkulosis pada anak-anak menjadi tiga kategori yaitu: daerah dengan prevalensi tinggi, daerah dengan prevalensi sedang, dan daerah dengan prevalensi rendah. Metode pengelompokan yang digunakan adalah metode *K-Means clustering* dan mengidentifikasi pola distribusi menggunakan GIS.
4. Harling et al, pada tahun 2013 penelitian ini menggunakan metode berbeda pada objek penelitian yang sama yaitu *metode Conditional Autoregressive Models (CAR)* dalam analisis spasial pendistribusian penyebaran tuberkulosis dan hubungannya terhadap faktor sosial-ekonomi. Hasil dari metode *CAR* ialah hubungan sosial-ekonomi dan penyebaran tuberkulosis sangat erat, terutama di beberapa daerah perkotaan. Namun, berdasarkan model spasial autokorelasi penyebaran tuberkulosis tidak berhubungan dengan pendapatan, urbanitas, dan kepadatan penduduk. Terdapat faktor-faktor lain yang menjadi penyebab penyebaran tuberkulosis yang tinggi. Ini menunjukkan bahwa masih ada variabel penting yang berkorelasi dan tidak ditemukan oleh metode yang digunakan.
5. Li et al, pada tahun 2019 melakukan penelitian dengan menggunakan pendekatan Bayesian dalam penelitian untuk menguji distribusi spatio-temporal tuberkulosis dan hubungan antara variabel sosial ekonomi, gaya hidup sehat, demografi, dan meteorologi menggunakan data pasien tuberkulosis. Tingkat kematian tuberkulosis yang berkorelasi dengan ketimpangan sosial ekonomi dihitung menggunakan metodologi Bayesian berdasarkan pendekatan bersarang Laplace. Hubungan antara

faktor sosial ekonomi, pola hidup sehat, demografi, dan meteorologi dapat mempengaruhi jumlah penderita tuberkulosis, ditemukan berdasarkan teknik Bayesian Spatio-temporal. Untuk jumlah tingkat populasi tuberkulosis, diperkirakan parameter seperti rasio laki-laki terhadap perempuan, layanan medis, kepadatan penduduk, jumlah curah hujan, kecepatan angin dan durasi paparan sinar matahari, menunjukkan bahwa hubungannya positif.

State of The Art merupakan sebuah ringkasan tabel yang memuat penelitian-penelitian sebelumnya yang memiliki keterkaitan dengan penelitian yang dilakukan seperti judul, penulis, penerbit, tahun, metode dan hasil yang dicapai.

**Tabel 1.** State of The Art

No	Judul	Penulis	Penerbit	Tahun	Metode	Hasil
1.	<i>Methods used in the spatial analysis of epidemiology a systematic review</i>	Debebe S. Malancha K. Kefyalew A. A. Romain R. Archie CA C. James M. T. Justin T. D. Emma S. M.	<i>BMC Medice</i>	2018	Studi Literature Sistematis	Dari 131 penelitian dengan kategori metode spatial <i>cluster</i> analysis terdapat tiga metode yang paling seringdigunakan yaitu: Kulldorff's Spatial Scan Statistic dengan 48 penelitian, Global Moran's dengan 28 penelitian dan Local Moran's dengan 14 penelitian.
2.	<i>Mapping the Pattern of Spread of Tuberculosis Cases using Spatio-Temporal Approach in Makassar – South Sulawesi</i>	Muhammad Resha, Ansar Suyuti, Rhiza S Sadjad, Muhammad Niswar	International Conference on Computer, Control, Informatics and its Applications (IC3INA)	2019	Metode <i>K-Means clustering</i> dan <i>Quantum GIS</i>	Dengan metode Spatio-Temporal ditemukan bahwa distribusi terbesar penyebaran tuberkulosis berada pada kabupaten Panakukang dan terkecil berada pada kabupaten Wajo. Dengan rata-rata usia penderita 42 tahun. Dalam kurun waktu 3 tahun (2015 - 2017) dengan jumlah 15 kecamatan dan jumlah penderita 5896 pasien.

No	Judul	Penulis	Penerbit	Tahun	Metode	Hasil
3.	<i>Clustering tuberculosis in children using KMeans based on geographic information</i>	Ratih Sari Wardani, Purwanto Sayono, Aditya Paramananda	AIP Conference Proceedings	2019	Metode <i>K-Means</i> dan GIS	Metode <i>K-means</i> meng <i>clustering</i> data menjadi tiga kategori: distribusi terbesar, distribusi sedang, distribusi terkecil pada kasus tuberkulosis pada anak di kota Semarang. Dengan metode ini tidak terdapat <i>everlab</i> pada hasil Analisa yang dilakukan.
4.	<i>A spatial analysis of social and economic determinants of tuberculosis in Brazil</i>	Guy Harling, Marcia C.Castro	Health & Place 25 (2014) 56–67	2013	<i>Conditional Autoregressiv Models</i> (CAR)	Hasil dari analisa CAR, hubungan sosial-ekonomi dan penyebaran tuberkulosis sangat erat, terutama di beberapa daerah perkotaan. Namun, berdasarkan model spasial autokorelasi penyebaran tuberkulosis tidak berhubungan dengan pendapatan, urbanitas, dan kepadatan penduduk.
5.	<i>The spatio-temporal analysis of the incidence of tuberculosis and the associated factors in mainland China, 2009-2015</i>	Qihuan Li, Mengyang Liu, Yingjie Zhang, Shangwu Wu, Yang Y, Yue Liua, Xiangtong Liu, Feng Z, Yanxia Luo, Xinghua Y, Weimin Li, Xia Li, Wei Wang, Xiaonan Wang, Xiuhua Guo	Infection, Genetics and Evolution 75 (2019) 103949	2019	Metode <i>Spatio-Temporal Bayesian</i>	Berdasarkan metode spatio-temporal Bayesian di temukan bahwa faktor sosial-ekonomi, gaya hidup sehat, demografi dan meteorologi dapat mempengaruhi jumlah kasus penderita tuberkulosis. Untuk jumlah tingkat populasi tuberkulosis, diperkirakan parameter seperti rasio laki-laki terhadap perempuan, layanan medis, kepadatan penduduk, jumlah curah hujan, kecepatan angin dan durasi peparan sinar matahari, menunjukkan bahwa hubungannya positif.

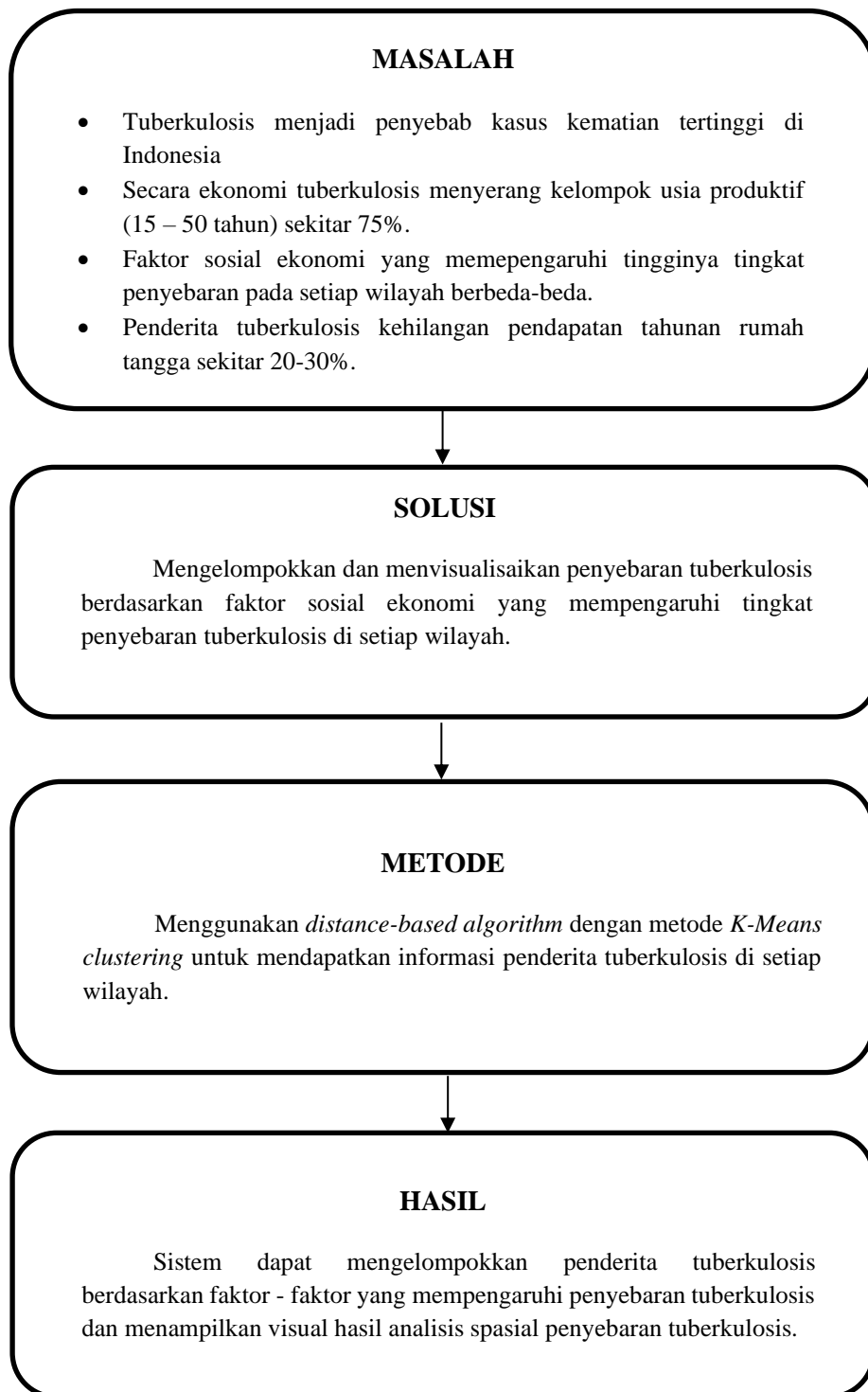


### C. Kerangka Pikir

Kerangka pikir menjelaskan alur penelitian yang akan dilakukan. Pada tahap pertama menjelaskan permasalahan yang ada dalam hal survei peningkatan penyebaran tuberkulosis. Penyebaran tuberkulosis menyerang semua kelompok umur, baik anak-anak, dewasa bahkan lansia. Hal ini dapat membahayakan tidak hanya dari aspek kesehatan semata tetapi juga dari aspek sosial ekonomi.

Oleh karena itu setelah melalui proses *literature review*, penelitian ini menawarkan solusi untuk pemanfaatan suatu sistem data mining yang di harapkan dapat mampu memberikan suatu gambaran atau tampilan spasial penyebaran kasus tuberkulosis di wilayah Kota Makassar berdasarkan jumlah penderita dan faktor sosial ekonomi yang mempengaruhi penyebaran tuberkulosis dengan mengacu dari beberapa penelitian sebelumnya.

Selanjutnya penggunaan algoritma *K-Means clustering* dapat mengelompokkan penderita tuberkulosis tertinggi suatu wilayah berdasarkan faktor sosial ekonomi yang mempengaruhi peningkatan penyebaran tuberkulosis. Hasil analisis di tampilkan secara visual dalam bentuk spasial yang akan lebih mudah untuk dipahami. Adapun kerangka pikir penelitian dapat dilihat pada Gambar 4.



**Gambar 4.** Kerangka Pikir Penelitian