

**SKRIPSI**

**ANALISIS TUBERKULOSIS PADA ANAK BERBASIS DATA  
MINING MENGGUNAKAN ALGORITMA FP-GROWTH**

**Disusun dan diajukan oleh:**

**MUH. ALFARABI ALIF PUTRA**

**D121171510**



**DEPARTEMEN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS HASANUDDIN**

**MAKASSAR**

**2022**

**LEMBAR PENGESAHAN SKRIPSI**

**ANALISIS TUBERKULOSIS PADA ANAK BERBASIS DATA MINING  
MENGUNAKAN ALGORITMA FP-GROWTH**

**Disusun dan diajukan oleh**

**MUH. ALFARABI ALIF PUTRA**

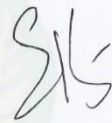
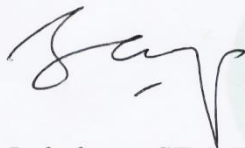
**D121171510**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana Program Studi Teknik Informatika Fakultas Teknik Universitas Hasanuddin pada tanggal 08 September 2022 dan dinyatakan telah memenuhi syarat kelulusan.

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,



Dr. Indrabayu, ST., MT., M. Bus. Sys  
Nip. 197507162002121004

Elly Warni, ST., MT.  
Nip. 198202162008122001

Ketua Program Studi,



Dr. Indrabayu, S.T., M.T., M. Bus. Sys  
Nip. 19750716 200212 1 004

## PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Muh. Alfarabi Alif Putra

NIM : D121171510

Departemen : Teknik Informatika

Jenjang : S1

Menyatakan dengan ini karya tulisan saya berjudul:

ANALISIS TUBERKULOSIS PADA ANAK BERBASIS DATA MINING  
MENGUNAKAN ALGORITMA FP-GROWTH

Adalah karya tulisan saya sendiri dan bukan merupakan pengambilalihan tulisan orang lain bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri.

Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan skripsi ini hasil karya orang lain, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Makassar, 08 September 2022

Yang menyatakan,



Muh. Alfarabi Alif Putra

## ABSTRAK

Tuberkulosis (TB) adalah penyakit menular yang merupakan penyebab utama kesehatan yang buruk, salah satu dari 10 penyebab kematian teratas di dunia dan penyebab utama kematian dari agen infeksi tunggal (peringkat di atas HIV/AIDS). Ketua UKK Respiralogi IDAI pada tahun 2016 mengatakan bahwa masalah yang dihadapi pada TB anak adalah kesulitan diagnosis, pengobatan dan belum optimalnya program pencegahan. Berdasarkan data yang diperoleh melalui pengisian kuesioner oleh orang tua pasien tuberkulosis anak Klinik Pediatria Husada, terdapat beberapa variabel yang berhubungan dengan tuberkulosis anak. Dengan variabel tersebut, maka akan dilakukan asosiasi untuk mengetahui keterkaitan antar variabel tuberkulosis anak. Salah satu teknologi yang dapat digunakan adalah *Data Mining* dengan metode *Association Rules*. Hasil dari penelitian ini menunjukkan bahwa penggunaan algoritma FP-Growth yang diujicobakan pada *dataset* tuberkulosis anak yang diperoleh dari Klinik Pediatria Husada dengan jumlah data yang diolah sebanyak 211 kejadian, dengan menggunakan *minimum support* = 0.35 diperoleh sebanyak 253 *frequent itemset*. Dari *frequent itemset*, proses penemuan aturan asosiasi dilakukan dengan menerapkan *minimum confidence* = 0.9, maka didapatkan aturan asosiasi setiap kecamatan yang terdapat dalam *dataset*. Aturan asosiasi yang didapatkan pada setiap kecamatan dilakukan uji lift ratio untuk menentukan kekuatan dari aturan sebagai hasil akhir aturan asosiasi. Jika nilai lift ratio kurang dari atau sama dengan ( $\leq$ ) 1, maka hubungan sebab-akibat yang terjadi bersifat saling lepas satu sama lain. Sedangkan, jika nilai lift ratio lebih dari ( $>$ ) 1, maka hubungan sebab-akibat yang terjadi bersifat saling berhubungan satu sama lain dan dapat dikatakan kejadian tersebut bukan kebetulan dan akan berulang.

**Kata kunci:** *tuberkulosis, data mining, association rules, algoritma fp-growth*

## KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, yang telah memberikan rahmat dan karunia-Nya, sehingga dapat menyelesaikan tugas akhir dengan judul “Analisis Tuberkulosis pada Anak Berbasis Data Mining menggunakan Algoritma FP-Growth” sebagai salah satu syarat dalam menyelesaikan jenjang Strata-1 di Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin.

Penulis menyadari bahwa penyusunan dan penulisan tugas akhir ini tidak dapat terselesaikan dengan baik tanpa adanya bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih banyak kepada:

1. Kedua orang tua penulis, Bapak Ridwan Dian dan Ibu Murniaty Hamid yang selalu memberikan dukungan, doa, dan semangat yang tiada hentinya, serta selalu sabar dalam mendidik penulis sejak kecil.
2. Bapak Dr. Indrabayu, S.T., M.T., M.Bus.sys., selaku pembimbing I dan Ibu Elly Warni, S.T., M.T. selaku pembimbing II, yang senantiasa menyediakan waktu, tenaga, pikiran, dan perhatian yang luar biasa dalam mengarahkan penulis untuk menyelesaikan tugas akhir.
3. Klinik Pediatrica, Dr. dr. Bob Wahyudin, Sp. A(K), CIMI yang senantiasa memberikan nasihat dan masukan, seluruh staf klinik Pediatrica yang membantu penulis selama pengambilan data, serta Muhammad Irzam Kasyfillah sebagai tim Tuberkulosis Anak.
4. Segenap keluarga *AIMP Research Group* Universitas Hasanuddin yang telah memberikan begitu banyak bantuan selama penelitian, pengambilan data dan diskusi *progress* penyusunan tugas akhir serta memberikan semangat di masa-masa sulit.
5. Muhammad Zulfahmi Sadrah, Muhammad Ilham Askari, Fitriani Nasir, Khaerul Hidayat dan Muhammad Fadhil yang telah membantu penulis sejak awal perkuliahan dan selalu membantu dalam penyelesaian tugas akhir serta Fauzan Alif Anwar yang bersedia menampung penulis di Ruko selama proses penyelesaian tugas akhir.

6. Ghalib, Yusril, Nublan, Sanji, Irsyad, Famzer, Ilmi, Taslinda, Devy, Iju, Giyani, Andar, Aries, Wahyu, Bishram, Asjar, dan seluruh anggota Recognizer yang telah memberi dukungan kepada penulis.
7. Syukur, Budi, dan Sandi yang sejak SMP hingga saat ini selalu mendukung dan menghibur penulis.
8. Seluruh warga Charchar Reza, Dhanu, Aisyah, Filda, Ikki(Kibo), Lastry, Muna, Namira, dan Rinaldy yang sejak SMA hingga saat ini selalu mendukung penulis serta Ibu Irianti yang bersedia menampung penulis selama pengerjaan tugas akhir.
9. Segenap staf dan dosen Departemen Teknik Informatika, Fakultas Teknik, Universitas Hasanuddin yang telah membantu kelancaran penyelesaian tugas akhir.
10. Serta pihak-pihak lain yang tidak disebutkan dan tanpa sadar telah menjadi inspirasi dan membantu penulis dalam menyelesaikan tugas akhir.

Penulis berharap semoga Tuhan Yang Maha Esa berkenan membalas segala kebaikan serta jasa dari semua pihak yang telah banyak membantu penulis dalam menyelesaikan tugas akhir ini. Penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna, oleh karena itu penulis mengharapkan segala bentuk saran serta masukan yang membangun dari berbagai pihak. Semoga tugas akhir ini dapat memberi manfaat bagi para pembaca dan semua pihak. Aamiin.

Makassar, Juli 2022

Penulis

## DAFTAR ISI

ABSTRAK .....	ii
KATA PENGANTAR .....	iii
DAFTAR ISI.....	v
DAFTAR GAMBAR .....	vii
DAFTAR TABEL.....	ix
BAB I PENDAHULUAN .....	1
I.1    Latar Belakang.....	1
I.2    Rumusan Masalah.....	3
I.3    Tujuan Penelitian .....	3
I.4    Manfaat Penelitian .....	3
I.5    Batasan Masalah .....	4
I.6    Sistematika Penulisan .....	5
BAB II TINJAUAN PUSTAKA.....	6
II.1    Tuberkulosis.....	6
II.2    Anak.....	7
II.3    Data Mining .....	9
II.4    Metode <i>Association Rules</i> .....	12
II.5    Algoritma FP-Growth.....	14
II.5.1    Pembangunan FP-Tree .....	15
II.5.2    Penerapan Algoritma FP-Growth.....	21
II.6    GeoCoding .....	23
II.7    WebGis .....	24
II.8    System Usability Scale .....	25
BAB III METODOLOGI PENELITIAN.....	28

III.1	Tahapan Penelitian.....	28
III.2	Waktu dan Lokasi Penelitian .....	30
III.3	Instrumen Penelitian .....	30
III.4	Teknik Pengambilan Data.....	31
III.5	Perancangan Sistem .....	32
III.6	Asosiasi Algoritma FP-Growth .....	38
III.7	Evaluasi Sistem.....	39
III.8	Visualisasi WebGIS.....	39
III.9	Evaluasi Usability.....	39
BAB IV HASIL DAN PEMBAHASAN .....		40
IV.1	Penerapan Algoritma FP-Growth dengan Menggunakan Python.....	40
IV.2	Penerapan Algoritma FP-Growth dengan Menggunakan <i>Library</i> Python .....	65
IV.3	Pembahasan .....	68
IV.4	Visualisasi WebGIS.....	73
IV.5	Hasil Evaluasi Usability.....	76
BAB V KESIMPULAN DAN SARAN.....		78
V.1	Kesimpulan .....	78
V.2	Saran .....	78
DAFTAR PUSTAKA .....		80
LAMPIRAN.....		82



## DAFTAR GAMBAR

<b>Gambar 2. 1</b> Tahap Data Mining (Han et al., 2011).....	10
<b>Gambar 2. 2</b> Hasil pembentukan FP-tree TID 1 (Samuel, n.d.) .....	19
<b>Gambar 2. 3</b> Hasil pembentukan FP-Tree TID 2 (Samuel, n.d.).....	19
<b>Gambar 2. 4</b> Hasil pembentukan FP-Tree TID 3 (Samuel, n.d.).....	19
<b>Gambar 2. 5</b> Hasil pembentukan FP-Tree TID 10 (Samuel, n.d.).....	20
<b>Gambar 2. 6</b> Interpretasi skor SUS (Vaziri et al., 2016) .....	27
<b>Gambar 3. 1</b> Tahapan Penelitian.....	28
<b>Gambar 3. 2</b> Rancangan Sistem.....	32
<b>Gambar 3. 3</b> Sampel dataset Tuberkulosis di excel.....	33
<b>Gambar 3. 4</b> Sampel dataframe pada data selection.....	34
<b>Gambar 3. 5</b> Hasil data clening .....	36
<b>Gambar 3. 6</b> Dataset dalam bentuk array .....	37
<b>Gambar 3. 7</b> Sampel hasil transformasi dataset.....	37
<b>Gambar 4. 1</b> Perbandingan jenis kelamin.....	41
<b>Gambar 4. 2</b> Perbandingan pendapatan orang tua .....	41
<b>Gambar 4. 3</b> Perbandingan diabetes anak.....	42
<b>Gambar 4. 4</b> Perbandingan vaksin BCG.....	42
<b>Gambar 4. 5</b> Perbandingan ASI eksklusif .....	43
<b>Gambar 4. 6</b> Perbandingan riwayat tuberkulosis orang serumah.....	43
<b>Gambar 4. 7</b> Contoh Dataset Tuberkulosis.....	44
<b>Gambar 4. 8</b> Dataset setelah dilakukan cleaning data .....	44
<b>Gambar 4. 9</b> FP-tree.....	45
<b>Gambar 4. 10</b> Contoh Conditional Pattern Base .....	46
<b>Gambar 4. 11</b> Sampel Frekuensi Item .....	46
<b>Gambar 4. 12</b> Nilai Support itemset .....	48
<b>Gambar 4. 13</b> FP-Growth rekursif.....	48
<b>Gambar 4. 14</b> Frequent Itemset .....	49
<b>Gambar 4. 15</b> Aturan Asosiasi kecamatan Bajeng .....	63
<b>Gambar 4. 16</b> Aturan Asosiasi kecamatan Manggala.....	64
<b>Gambar 4. 17</b> Aturan Asosiasi kecamatan Soreang .....	64

<b>Gambar 4. 18</b> FP-Growth dengan Library.....	66
<b>Gambar 4. 19</b> Aturan asosiasi dengan Library .....	66
<b>Gambar 4. 20</b> Frequent Itemset dengan Library.....	66
<b>Gambar 4. 21</b> Aturan asosiasi kecamatan Bajeng menggunakan library .....	67
<b>Gambar 4. 22</b> Aturan asosiasi kecamatan Manggala dengan library .....	67
<b>Gambar 4. 23</b> Aturan asosiasi kecamatan Soreang dengan library .....	68
<b>Gambar 4. 24</b> Tampilan Dashboard WebGIS.....	73
<b>Gambar 4. 25</b> Menu pada sidebar .....	74
<b>Gambar 4. 26</b> Halaman hasil aturan asosiasi WebGIS.....	75
<b>Gambar 4. 27</b> Hasil aturan asosiasi setiap kecamatan .....	75

## DAFTAR TABEL

<b>Tabel 2. 1</b> Contoh market basket transaction (Tan & Vipin Kumar, 2005) .....	12
<b>Tabel 2. 2</b> Tabel data transaksi mentah (Samuel, n.d.).....	17
<b>Tabel 2. 3</b> Frekuensi kemunculan tiap karakter(Samuel, n.d.).....	17
<b>Tabel 2. 4</b> Tabel Data Transaksi (Samuel, n.d.) .....	18
<b>Tabel 2. 5</b> Pseudocode Algoritma FP-Growth .....	22
<b>Tabel 2. 6</b> Daftar pertanyaan SUS .....	26
<b>Tabel 4. 1</b> Aturan Asosiasi Kota Makassar .....	50
<b>Tabel 4. 2</b> Aturan Asosiasi Kab. Maros .....	53
<b>Tabel 4. 3</b> Aturan Asosiasi Kab. Gowa .....	54
<b>Tabel 4. 4</b> Aturan Asosiasi Kab. Barru.....	56
<b>Tabel 4. 5</b> Aturan Asosiasi Kab. Takalar .....	56
<b>Tabel 4. 6</b> Aturan Asosiasi Kab. Bantaeng.....	57
<b>Tabel 4. 7</b> Aturan Asosiasi Kab. Sinjai .....	58
<b>Tabel 4. 8</b> Aturan Asosiasi Kab. Bone .....	58
<b>Tabel 4. 9</b> Aturan Asosiasi Kab. Luwu Utara .....	59
<b>Tabel 4. 10</b> Aturan Asosiasi Kota Parepare .....	59
<b>Tabel 4. 11</b> Aturan Asosiasi Kab. Wajo .....	60
<b>Tabel 4. 12</b> Aturan Asosiasi Kab. Pangkajene dan Kepulauan .....	61
<b>Tabel 4. 13</b> Aturan Asosiasi Kab. Bekasi.....	61
<b>Tabel 4. 14</b> Aturan Asosiasi Kab. Polewali Mandar .....	62
<b>Tabel 4. 15</b> Aturan Asosiasi Kab. Maluku Utara .....	62
<b>Tabel 4. 16</b> Aturan Asosiasi dari program tanpa library dan library mlxtend.....	71
<b>Tabel 4. 17</b> Data responden SUS.....	76
<b>Tabel 4. 18</b> Hasil perhitungan skor SUS .....	77

# **BAB I**

## **PENDAHULUAN**

### **I.1 Latar Belakang**

Tuberkulosis (TB) adalah penyakit menular yang merupakan penyebab utama kesehatan yang buruk, salah satu dari 10 penyebab kematian teratas di dunia dan penyebab utama kematian dari agen infeksi tunggal (peringkat di atas HIV/AIDS). Pada 2019, sekitar 10 juta orang menderita TB dan 1,4 juta meninggal. Tuberkulosis disebabkan oleh basil *Mycobacterium tuberculosis*, yang menyebar ketika orang yang sakit tuberkulosis mengeluarkan bakteri ke udara misalnya dengan batuk. Penyakit ini biasanya mempengaruhi paru-paru yang biasa di sebut dengan TB paru tetapi juga dapat mempengaruhi organ lain atau biasa disebut dengan TB ekstraparu. Indonesia berada pada urutan ketiga setelah Cina dan India dengan jumlah kasus Tuberkulosis sebanyak 842 ribu berdasarkan data dari *World Health Organization* (WHO). Tuberkulosis dapat menyerang siapa saja, baik anak-anak, remaja, dewasa, maupun lansia (World Health Organization, 2020).

Tuberkulosis pada anak mempunyai permasalahan khusus yang berbeda dengan TB dewasa. Ketua UKK Respiralogi IDAI pada tahun 2016 (Kementerian Kesehatan Republik Indonesia, n.d.) mengatakan bahwa masalah yang dihadapi pada TB anak adalah kesulitan diagnosis, pengobatan, dan belum optimalnya program pencegahan. Melakukan diagnosa Tuberkulosis pada anak lebih sulit karena pada kasus anak jarang menunjukkan gejala tuberkulosis, seperti batuk dan masalah saluran pernafasan. Kesulitan

tersebut menyebabkan risiko terjadinya *under* diagnosis dan *over* diagnosis tuberkulosis pada anak. Kedua hal tersebut mengakibatkan penanganan tuberkulosis yang tidak tepat pada anak (Kementerian Kesehatan Republik Indonesia, n.d.).

Berdasarkan Peraturan Presiden Republik Indonesia Nomor 67 Tahun 2021 tentang penanggulangan tuberkulosis, tuberkulosis masih menjadi masalah kesehatan di Indonesia dan menimbulkan masalah yang sangat kompleks baik dari segi medis maupun sosial, ekonomi, dan budaya. Penanggulangan tuberkulosis adalah segala upaya kesehatan yang mengutamakan aspek promotif dan preventif tanpa mengabaikan aspek kuratif dan rehabilitatif untuk melindungi kesehatan masyarakat, menurunkan angka kesakitan, kecacatan atau kematian, memutuskan penularan, mencegah resistensi obat tuberkulosis, dan mengurangi dampak yang ditimbulkan akibat tuberkulosis. Oleh sebab itu, dengan sulitnya diagnosa tuberkulosis pada anak dan berdasarkan Peraturan Presiden Republik Indonesia Nomor 67 Tahun 2021, maka penelitian ini dilakukan untuk membantu dalam melakukan analisis terhadap penyakit tuberkulosis pada anak menggunakan *data mining*, sehingga dapat diharapkan memberikan informasi sebagai dasar untuk melakukan tindakan yang diperlukan dalam pengobatan, pengendalian, dan pencegahan.

## **I.2 Rumusan Masalah**

Adapun rumusan masalah dari penelitian ini ialah sebagai berikut:

- a. Bagaimana mengidentifikasi keterkaitan antar variabel menggunakan *data mining* dengan teknik asosiasi dalam analisis penyakit tuberkulosis pada anak?
- b. Bagaimana penerapan algoritma FP-Growth dalam menemukan pola keterkaitan antar variabel pada penyakit tuberkulosis anak?

## **I.3 Tujuan Penelitian**

Adapun tujuan dari penelitian ini ialah sebagai berikut:

- a. Untuk menerapkan teknologi *data mining* dengan teknik asosiasi dalam mengidentifikasi keterkaitan antar variabel dalam analisis penyakit tuberkulosis pada anak.
- b. Untuk mengetahui penerapan algoritma FP-Growth dalam menemukan pola keterkaitan antar variabel pada penyakit tuberkulosis anak.

## **I.4 Manfaat Penelitian**

Adapun manfaat dari penelitian ini ialah sebagai berikut:

- a. Memberikan dorongan penggunaan teknologi *data mining* dalam mendapatkan hasil aturan asosiasi pada analisis terhadap penyakit tuberkulosis pada anak.

- b. Memberikan informasi yang dapat digunakan sebagai dasar untuk melakukan tindakan yang diperlukan dalam pengendalian dan pencegahan tuberkulosis pada anak di masa yang akan datang.

### **I.5 Batasan Masalah**

Batasan masalah dalam penelitian ini ialah sebagai berikut:

- a. Pengambilan data dengan kuesioner yang dilakukan di Klinik Pediatrica Husada melalui *Google forms* yang disebarakan menggunakan WhatsApp, SMS, dan *Broadcast* menggunakan sistem informasi Klinik Pediatrica Husada.
- b. Variabel – variabel penelitian terdiri dari:
  - Umur anak pada saat tuberkulosis
  - Alamat tempat tinggal anak
  - Status gizi dihitung berdasarkan tinggi badan dan berat badan anak menggunakan rumus IMT
  - Pekerjaan orang tua
  - Riwayat penyakit pada keluarga
  - Riwayat opname anak
  - Riwayat tuberkulosis orang serumah tempat tinggal anak
  - Kondisi rumah tempat tinggal anak
  - Riwayat konsumsi ASI eksklusif yaitu konsumsi ASI selama 6 bulan tanpa menambahkan dan/atau mengganti dengan makanan atau minuman lain

## **I.6 Sistematika Penulisan**

Berikut adalah gambaran singkat mengenai isi tulisan secara keseluruhan:

### **BAB 1 PENDAHULUAN**

Pada bab ini diuraikan mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah penelitian, metode penelitian, serta sistematika penulisan.

### **BAB 2 TINJAUAN PUSTAKA**

Pada bab ini membahas landasan teori yang digunakan untuk menganalisis masalah yang akan diteliti serta hal-hal lain yang berhubungan dengan variabel-variabel data yang akan digunakan, *data mining*, *association rules*, dan algoritma FP-Growth.

### **BAB 3 METODOLOGI PENELITIAN**

Pada bab ini berisi mengenai tahap penelitian, instrumen penelitian, pengumpulan data, penerapan algoritma, teknik pengolahan data, serta hasil asosiasi.

### **BAB 4 HASIL DAN PEMBAHASAN**

Bab ini berisi tentang sistem yang telah berhasil dibangun serta pembahasannya.

### **BAB 5 PENUTUP**

Bab ini berisi tentang kesimpulan yang di dapatkan berdasarkan hasil penelitian yang telah dilakukan serta saran-saran untuk pengembangan lebih lanjut.



## **BAB II**

### **TINJAUAN PUSTAKA**

#### **II.1 Tuberkulosis**

Tuberkulosis (TB) adalah penyakit menular yang merupakan penyebab utama kesehatan yang buruk dan salah satu penyebab utama kematian di seluruh dunia. Hingga pandemi virus corona (COVID-19), Tuberkulosis penyebab utama kematian dari agen infeksi tunggal, peringkat di atas HIV/AIDS. Tuberkulosis di sebabkan oleh basil *Mycobacterium Tuberculosis* yang menyebar ketika orang yang sakit tuberkulosis mengeluarkan bakteri ke udara (misalnya melalui batuk). Sekitar 90% orang yang mengidap penyakit ini adalah orang dewasa dan lebih banyak kasus di antaranya adalah pria daripada wanita. Sekitar seperempat populasi dunia terinfeksi *Mycobacterium Tuberculosis* (World Health Organization, 2021).

Indonesia menghadapi beban ganda penyakit karena meningkatnya penyakit tidak menular dan masih tingginya insiden penyakit menular. Malnutrisi masih menjadi masalah kesehatan utama. Terdapat sekitar 30,8% (7 juta) anak di bawah usia lima tahun mengalami kerdil (stunting). Sementara itu, masalah kelebihan berat badan dan obesitas pada orang dewasa meningkat hingga lima kali lipat lebih tinggi daripada target RPJMN 2019 dalam kurun waktu 3 tahun (2016-2019). Angka kematian ibu di Indonesia sebesar 305 per 100.000 kelahiran hidup merupakan angka kematian ibu yang tertinggi di Asia Tenggara. Tuberkulosis menjadi salah satu dari lima penyebab utama beban penyakit pada tahun 2017. Selain itu, diabetes melitus (DM) yang juga

merupakan faktor risiko Tuberkulosis menjadi penyumbang beban penyakit ke-3 terbesar (Kesehatan & Indonesia, n.d.).

Tuberkulosis anak mempunyai permasalahan khusus yang berbeda dengan tuberkulosis paru dewasa. Salah satu permasalahan tuberkulosis anak di Indonesia adalah penegakan diagnosis. Permasalahan lain dalam program penanggulangan tuberkulosis anak adalah semakin meningkatnya jumlah kasus tuberkulosis resisten obat (TB RO) pada dewasa, yang bisa menjadi sumber penularan bagi anak. Jumlah pasti kasus tuberkulosis resisten obat pada anak di Indonesia saat ini belum diketahui, tetapi semakin meningkat. Faktor risiko penularan tuberkulosis anak sama halnya dengan tuberkulosis pada umumnya, tergantung dari tingkat penularan, lama pajanan, dan daya tahan tubuh. Salah satu cara penemuan pasien tuberkulosis anak dapat ditemukan melalui upaya Penemuan Secara Aktif yaitu upaya yang dilakukan berbasis keluarga dan masyarakat melalui kegiatan investigasi kontak pada anak yang kontak erat dengan pasien tuberkulosis yang menular. Yang dimaksud dengan kontak erat adalah anak yang tinggal serumah atau sering bertemu dengan pasien tuberkulosis menular (Kementerian Kesehatan Republik Indonesia, n.d.).

## **II.2 Anak**

Merujuk dari Kamus Umum bahasa Indonesia mengenai pengertian anak secara etimologis, diartikan dengan manusia yang belum dewasa. Menurut The Minimum Age Convention Nomor 138 tahun 1973, pengertian anak adalah seseorang yang berusia 15 tahun ke bawah. Sebaliknya dalam Convention The

Right Of the Child tahun 1989 yang telah diratifikasi pemerintah Indonesia melalui Keppres Nomor 39 tahun 1990 disebutkan bahwa anak adalah mereka yang berusia 18 tahun ke bawah. Pengertian anak menurut peraturan perundang-undangan dapat dilihat sebagai berikut (Latifa & Dhita Novika, n.d.):

a. Anak Menurut UU No. 35 Tahun 2014 Tentang Perlindungan Anak.

Pengertian anak berdasarkan Pasal 1 ayat (1) UU No. 35 Tahun 2014 tentang Perlindungan Anak adalah seseorang yang belum berusia 18 (delapan belas) tahun, termasuk anak yang masih dalam kandungan.

b. Anak menurut Kitab Undang-Undang Hukum perdata

Di jelaskan dalam Pasal 330 Kitab Undang-undang Hukum Perdata, mengatakan orang belum dewasa adalah mereka yang belum mencapai umur 21 tahun dan tidak lebih dahulu telah kawin. Jadi anak adalah setiap orang yang belum berusia 21 tahun dan belum menikah. Seandainya seorang anak telah menikah sebelum umur 21 tahun kemudian bercerai atau ditinggal mati oleh suaminya sebelum genap umur 21 tahun, maka ia tetap dianggap sebagai orang yang telah dewasa bukan anak-anak.

c. Menurut Kitab Undang-undang Hukum Pidana

Anak dalam Pasal 45 KUHPidana adalah anak yang umurnya belum mencapai 16 (enam belas) tahun.

### II.3 Data Mining

*Data mining* merupakan ekstraksi informasi berguna dari sejumlah data yang besar (Bhatia, 2019). *Data mining* juga biasa disebut dengan *Data Discovery* atau *Knowledge Discovery from Databases* (KDD), yaitu proses menganalisis data dari berbagai sudut pandang sehingga menjadi informasi yang berguna (Hussain et al., 2019). *Data mining* telah ada sejak tahun 1930-an, ditemukan oleh seorang ahli matematika, logika, dan seorang kriptografer, bernama Alan Turing. Menurut Jiawei Han (Han et al., 2011), *data mining* yang merupakan bagian dari *machine learning* menyelidiki bagaimana komputer dapat belajar atau meningkatkan kinerjanya berdasarkan data. Berikut merupakan beberapa pendekatannya:

a. *Supervised Learning*

*Supervised learning* adalah sebuah pendekatan untuk data yang memiliki informasi kelas atau label. Label merupakan variabel yang menjadi identitas dari tiap data dalam sekumpulan data, sehingga tujuan dari pendekatan ini adalah mengklasifikasi suatu data ke data yang sudah ada.

b. *Unsupervised Learning*

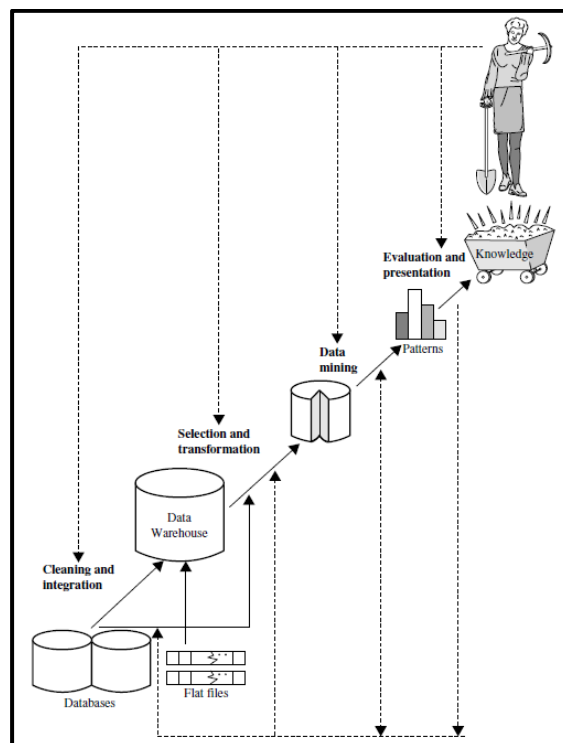
*Unsupervised learning* adalah proses belajar yang tidak diawasi (*unsupervised*) sebab set data yang diberikan tidak memiliki label, sehingga keluaran dari *unsupervised learning* tidak dapat ditebak karena model tersebut belajar sendiri untuk menemukan informasi dari set data yang telah diberikan. Salah satu penerapan dari *unsupervised learning* adalah *association rules*.

c. *Semi-supervised Learning*

*Semi-supervised learning* adalah kombinasi antara *supervised* dan *unsupervised learning*. Data yang diberikan pada pendekatan ini memiliki kombinasi antara data yang berlabel dan tidak berlabel.

d. *Active Learning*

*Active learning* adalah kasus yang khusus atau spesial pada *machine learning*, yang membuat mesin meminta peneliti atau pengguna secara interaktif memberi label data dengan keluaran yang diinginkan.



**Gambar 2. 1** Tahap Data Mining (Han et al., 2011)

Jiawei Han (Han et al., 2011) juga menyebutkan proses dari *Data Mining* atau *Knowledge Discovery from Database* (KDD) yang terbagi ke dalam 7 tahap, yaitu sebagai berikut:

- a. *Data cleaning*: tahap untuk membersihkan data yang hilang, *noise* dan tidak konsisten.
- b. *Data integration*: tahap di mana beberapa sumber data dapat digabungkan.
- c. *Data selection*: tahap untuk data yang relevan dengan analisis diambil dari *database*.
- d. *Data transformation*: tahap untuk data ditransformasikan dan dikonsolidasikan ke dalam bentuk yang sesuai untuk penambangan dengan melakukan operasi *summary* atau *aggregation*.
- e. *Data mining*: tahap penting di mana metode cerdas diterapkan untuk mengekstrak pola data.
- f. *Pattern evaluation*: tahap untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan berdasarkan ukuran ketertarikannya (*distance/interestingness measure*).
- g. *Knowledge presentation*: tahap di mana teknik gambaran visualisasi dan pengetahuan digunakan untuk menyajikan pengetahuan kepada pengguna.

## II.4 Metode *Association Rules*

Ide dari aturan asosiasi adalah untuk memeriksa semua kemungkinan hubungan *if-then* antar item dan memilih hanya yang paling mungkin sebagai indikator dari hubungan ketergantungan antar item. Biasanya digunakan istilah *antedecent* untuk mewakili bagian “jika” dan *consequent* untuk mewakili bagian “maka”. Dalam analisis ini, *antedecent* dan *consequent* adalah sekelompok item yang tidak punya hubungan secara bersama (Listriani et al., 2016).

Kebanyakan perusahaan bisnis mengakumulasikan data dalam jumlah besar dari operasi sehari-hari mereka. Misalnya, sejumlah besar data pembelian pelanggan dikumpulkan setiap hari di kasir toko grosir.

**Tabel 2. 1** Contoh market basket transaction (Tan & Vipin Kumar, 2005)

TID	BARANG
1	{ Roti, Susu }
2	{ Roti, Popok, Bir, Telur }
3	{ Susu, Popok, Bir, Cola }
4	{ Roti, Susu, Popok, Bir }
5	{ Roti, Susu, Popok, Cola }

Tabel 2.1 mengilustrasikan contoh dari data tersebut yang umumnya dikenal sebagai transaksi keranjang pasar (*market basket transaction*). Setiap baris dalam Tabel 2.1 sesuai dengan transaksi yang terjadi disertai atribut tanda pengenal transaksi (TID) dan sekumpulan item yang dibeli oleh pelanggan

tertentu. Penjual menganalisis data transaksi yang terjadi untuk mempelajari perilaku pembelian pelanggan mereka. Informasi berharga seperti ini dapat digunakan untuk mendukung berbagai aplikasi terkait bisnis seperti promosi pemasaran, manajemen inventaris, dan manajemen hubungan pelanggan (Tan, 2005).

Metodologi dasar aturan asosiasi dijelaskan sebagai berikut(Listriani et al., 2016):

#### 1. Pembentukan Pola Frekuensi Tinggi

Tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam suatu *database*. Nilai *support* adalah nilai penunjang atau persentase kombinasi sebuah item bersamaan dalam suatu *database*. Semakin besar nilai *support* menandakan semakin banyak data pendukung yang ditemukan dalam *database*. Nilai *support* sebuah item diperoleh dari persamaan:

$$Support(A) = \frac{jumlah\ kejadian\ mengandung\ A}{total\ kejadian} \quad (2.1)$$

Nilai *support* dari 2 item diperoleh dari persamaan:

$$Support(A, B) = \frac{jumlah\ kejadian\ mengandung\ A\ dan\ B}{total\ kejadian} \quad (2.2)$$

#### 2. Pembentukan Aturan Asosiasi

Setelah seluruh pola frekuensi tinggi ditemukan, maka tahap selanjutnya adalah membentuk aturan asosiasi dengan melihat kombinasi item yang memenuhi syarat minimum dari nilai *confidence*. Nilai *confidence* adalah nilai keyakinan berupa kuatnya hubungan antar item yang



didapatkan. Semakin besar nilai *confidence* menandakan semakin besar kemungkinan kombinasi item muncul secara bersamaan. Persamaan dari nilai *confidence* dengan menggunakan kondisi “jika A maka B” adalah sebagai berikut.

$$Confidence(A \rightarrow B) = \frac{\text{jumlah kejadian mengandung A dan B}}{\text{jumlah kejadian mengandung A}} \quad (2.3)$$

### 3. Rasio Peningkatan (*Lift Ratio*)

*Lift ratio* merupakan nilai yang menunjukkan keabsahan aturan yang terbentuk dalam proses transaksi dan memberikan informasi apakah benar produk A dibeli bersamaan dengan produk B. *Lift ratio* mengukur seberapa penting aturan yang telah terbentuk berdasarkan nilai *support* dan *confidence* yang telah didapatkan sebelumnya. Jika nilai *lift ratio* kurang dari atau sama dengan ( $\leq$ ) 1, maka hubungan sebab-akibat yang terjadi bersifat saling lepas satu sama lain. Sedangkan, jika nilai *lift ratio* lebih dari ( $>$ ) 1, maka hubungan sebab-akibat yang terjadi bersifat saling berhubungan satu sama lain dan dapat dikatakan kejadian tersebut bukan kebetulan dan akan berulang. Nilai *lift ratio* diperoleh dari persamaan:

$$Lift\ ratio = \frac{Confidence\ Antecedent}{Support\ Consequent} \quad (2.4)$$

## II.5 Algoritma FP-Growth

FP-Growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. FP-growth menggunakan pendekatan yang berbeda dari paradigma yang selama ini sering digunakan, yaitu paradigma

apriori. Paradigma apriori yang dikembangkan oleh Agrawal dan Srikan (1994), yaitu *anti-monotone Apriori Heuristic*: Setiap pola dengan panjang pola  $k$  yang tidak sering muncul (tidak frequent) dalam sebuah kumpulan data, maka pola dengan panjang  $(k+1)$  yang mengandung sub pola  $k$  tersebut tidak akan sering muncul pula (tidak frequent). Ide dasar paradigma apriori ini adalah dengan mencari himpunan kandidat dengan panjang  $(k+1)$  dari sekumpulan pola frequent dengan panjang  $k$ , lalu mencocokkan jumlah kemunculan pola tersebut dengan informasi yang terdapat dalam *database*. Adapun hal ini akan mengakibatkan algoritma apriori akan melakukan *scanning database* yang berulang-ulang, apalagi jika jumlah data cukup besar. Berbeda dengan Algoritma FP-growth yang hanya memerlukan dua kali *scanning database* untuk menentukan *frequent itemset* (Samuel, n.d.).

Struktur data yang digunakan untuk mencari *frequent itemset* dengan algoritma FP-growth adalah perluasan dari penggunaan sebuah pohon *prefix*, yang biasa disebut adalah FP-tree. Dengan menggunakan FP-tree, algoritma FP-growth dapat langsung mengekstrak *frequent Itemset* dari FP tree yang telah terbentuk dengan menggunakan prinsip *divide and conquer*.

### **II.5.1 Pembangunan FP-Tree**

FP-tree merupakan struktur penyimpanan data yang dimampatkan. FP-tree dibangun dengan memetakan setiap data transaksi ke dalam setiap lintasan tertentu dalam FP-tree. Karena dalam setiap transaksi yang dipetakan, mungkin ada transaksi yang memiliki item yang sama, maka

lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data FP-tree semakin efektif. Kelebihan dari FP-tree adalah hanya memerlukan dua kali pemindaian data transaksi yang terbukti sangat efisien (Samuel, n.d.).

Misal  $I = \{a_1, a_2, \dots, a_n\}$  adalah kumpulan dari item. Dan basis data transaksi  $DB = \{T_1, T_2, \dots, T_n\}$ , di mana  $T_i$  ( $i \in [1..n]$ ) adalah sekumpulan transaksi yang mengandung item di  $I$ . Sedangkan *support* adalah penghitung (*counter*) frekuensi kemunculan transaksi yang mengandung suatu pola. Suatu pola dikatakan sering muncul (*frequent pattern*) apabila *support* dari pola tersebut tidak kurang dari suatu konstanta  $\xi$  (batas ambang *minimum support*) yang telah didefinisikan sebelumnya. Permasalahan mencari pola frequent dengan batas ambang *minimum support count*  $\xi$  inilah yang dicoba untuk dipecahkan oleh FP-Growth dengan bantuan Struktur FP-tree. Adapun FP- tree adalah sebuah pohon dengan definisi sebagai berikut:

- FP-Tree dibentuk oleh sebuah akar yang diberi label null, sekumpulan upapohon yang beranggotakan item-item tertentu, dan sebuah tabel *frequent header*.
- Setiap simpul dalam FP-tree mengandung dua informasi penting, yaitu label item, menginformasikan jenis item yang direpresentasikan simpul tersebut dan *support count* merepresentasikan jumlah lintasan transaksi yang melalui simpul tersebut (Samuel, n.d.).

Misalkan diberikan tabel data transaksi sebagai berikut, dengan minimum *support count*  $\zeta = 2$

**Tabel 2. 2** Tabel data transaksi mentah (Samuel, n.d.)

No.	Transaksi
1	a,b
2	b,c,d,g,h
3	a,c,d,e,f
4	a,d,e
5	a,b,z,c
6	a,b,c,d
7	a,r

Frekuensi kemunculan tiap item dapat dilihat pada tabel berikut:

**Tabel 2. 3** Frekuensi kemunculan tiap karakter(Samuel, n.d.)

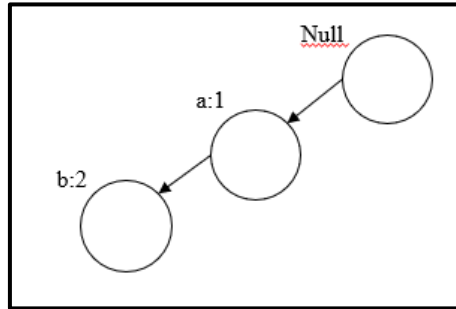
Item	Frekuensi
a	8
b	7
c	6
d	5
e	4
f	1
g	1
h	1

Setelah dilakukan pemindaian pertama didapat item yang memiliki frekuensi di atas *support count*  $\xi = 2$  adalah a,b,c,d, dan e. Kelima item inilah yang akan berpengaruh dan akan dimasukkan ke dalam FP-tree, selebihnya (r, z, g, dan h) dapat dibuang karena tidak berpengaruh signifikan. Tabel berikut mendata kemunculan item yang *frequent* dalam setiap transaksi, diurut berdasarkan yang frekuensinya paling tinggi.

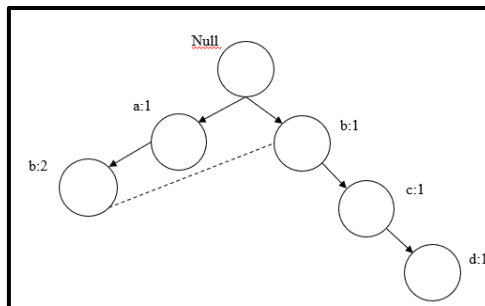
**Tabel 2. 4** Tabel Data Transaksi (Samuel, n.d.)

TID	Item
1	{a,b}
2	{b,c,d}
3	{a,c,d,e}
4	{a,d,e}
5	{a,b,c}
6	{a,b,c,d}
7	{a}
8	{a,b,c}
9	{a,b,d}
10	{b,c,e}

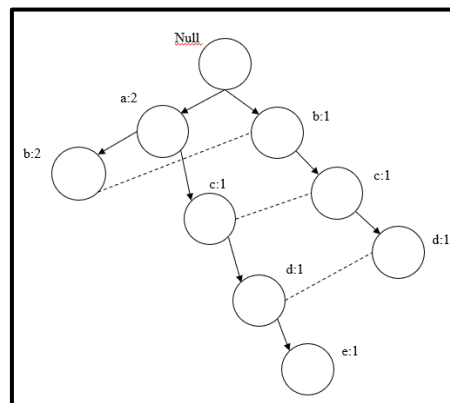
Gambar di bawah ini memberikan ilustrasi mengenai pembentukan FP-tree setelah pembacaan TID 1.



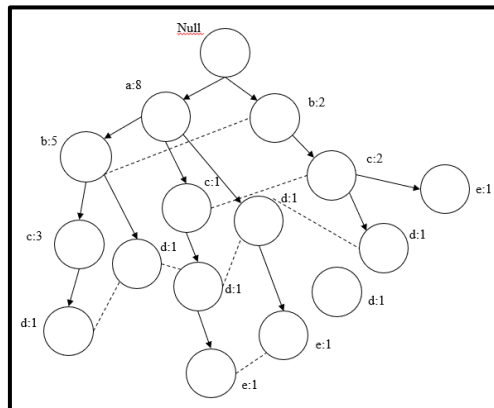
**Gambar 2. 2** Hasil pembentukan FP-tree TID 1 (Samuel, n.d.)



**Gambar 2. 3** Hasil pembentukan FP-Tree TID 2 (Samuel, n.d.)



**Gambar 2. 4** Hasil pembentukan FP-Tree TID 3 (Samuel, n.d.)



**Gambar 2. 5** Hasil pembentukan FP-Tree TID 10 (Samuel, n.d.)

Diberikan 10 data transaksi dengan 5 jenis item seperti pada Tabel 2.4. Gambar 2.2, Gambar 2.3, Gambar 2.4, dan Gambar 2.5 menunjukkan proses terbentuknya FP-Tree setiap TID dibaca. Setiap simpul pada FP-Tree mengandung nama sebuah item dan *counter support* yang berfungsi untuk menghitung frekuensi kemunculan item tersebut dalam tiap lintasan transaksi (Samuel, n.d.). FP-tree yang merepresentasikan data transaksi pada Tabel 2.1 dibentuk dengan cara sebagai berikut:

1. Kumpulan data dipindai pertama kali untuk menentukan *support count* dari setiap item. Item yang tidak *frequent* dibuang, sedangkan *frequent item* dimasukkan dan disusun dengan urutan menurun, seperti yang terlihat pada Tabel 2.3.
2. Pemindaian kedua, yaitu pembacaan TID pertama {a,b} akan membuat simpul a dan b, sehingga terbentuk lintasan transaksi Null  $\rightarrow$  a  $\rightarrow$  b. *Support count* dari setiap simpul bernilai awal 1.
3. Setelah pembacaan transaksi kedua {b,c,d}, terbentuk lintasan kedua yaitu Null  $\rightarrow$  b  $\rightarrow$  c  $\rightarrow$  d. *Support count* masing-masing *count* juga

bernilai awal 1. Walaupun b ada pada transaksi pertama, namun karena *prefix* transaksinya tidak sama, maka transaksi kedua ini tidak bisa dimampatkan dalam satu lintasan.

4. Transaksi keempat memiliki *prefix* transaksi yang sama dengan transaksi pertama, yaitu a, maka lintasan transaksi ketiga dapat ditimpakan di a, dan selanjutnya membuat lintasan baru sesuai dengan transaksi ketiga. (Gambar 2.3)
5. Proses ini dilanjutkan sampai FP-tree berhasil dibangun berdasarkan tabel data transaksi yang diberikan.

### **II.5.2 Penerapan Algoritma FP-Growth**

Setelah tahap pembangunan FP-tree dari sekumpulan data transaksi, akan diterapkan algoritma FP-Growth untuk mencari *frequent itemset* yang signifikan (Samuel, n.d.). Algoritma FP-Growth dibagi menjadi tiga langkah utama, yaitu:

1. Tahap Pembangkitan *Conditional Pattern Base*

*Conditional Pattern Base* merupakan subdatabase yang berisi *prefix path* (lintasan prefix) dan *suffix pattern* (pola akhiran). Pembangkitan *conditional pattern base* didapatkan melalui FP-tree yang telah dibangun sebelumnya.

2. Tahap Pembangkitan *Conditional FP-tree*

Pada tahap ini, *support count* dari setiap item pada setiap *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah



*support count* lebih besar sama dengan minimum *support count*  $\xi$  akan dibangkitkan dengan *conditional* FP-tree.

3. Tahap Pencarian *frequent itemset*

Apabila *conditional* FP-tree merupakan lintasan tunggal (*single path*), maka didapatkan *frequent itemset* dengan melakukan kombinasi item untuk setiap *conditional* FP-tree. Jika bukan lintasan tunggal, maka dilakukan pembangkitan FP-Growth secara rekursif.

Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*, yang dapat dilihat pada algoritma berikut:

**Tabel 2. 5** Pseudocode Algoritma FP-Growth

```
Input: FP-tree Tree
Output :  $R_t$  sekumpulan lengkap pola frequent
Method : FP-Grwoth(Tree, null)
    Procedure : FP-Growth(Tree,  $\alpha$ )
    {
    1: IF Tree mengandung single path P;
    2: THEN untuk tiap kombinasi (dinotasikan  $\beta$ ) dari
        node-node dalam path P DO
    3: bangkitkan pola  $\beta$   $\alpha$  dengan support = minimum
        support dari node-node dalam  $\beta$ 
    4: ELSE untuk tiap  $a_1$  dalam header dari Tree DO {
    5: bangkitkan pola
```

```
6: bangun  $\beta = a_1 \alpha$  dengan  
Support =  $a_1.support$   
7: IF Tree  $\beta = \emptyset$   
8: THEN panggil FP-Growth(Tree,  $\beta$ )  
}
```

## II.6 GeoCoding

*Geocoding* merupakan proses mengubah alamat berbasis teks ke dalam kode geospasial. Alamat dapat dilihat sebagai deskripsi tekstual di kehidupan nyata, sedemikian rupa sehingga mudah dipahami oleh manusia. Namun, informasi geospasial yang diberikan dalam bentuk tekstual tidak mudah untuk dianalisis oleh komputer sehingga alamat harus dikonversi ke referensi geografis (Rooth & Lambrix, n.d.).

Sistem yang mengubah alamat berbasis teks ke dalam koordinat geospasial disebut sebagai *Geocoders*. Saat ini sudah banyak *geocoder* komersial yang disediakan, beberapa di antaranya berbentuk perangkat lunak, dimana pengguna bertugas menyediakan referensi data geospasial, memelihara *database* yang akan menampung data, dan memberikan parameter untuk proses geocoding, seperti ArcView, ArcGis, dan Automatch. Selain dalam bentuk perangkat lunak, terdapat juga *geocoders* yang disediakan berbasis *online*, seperti Google Maps API, ArcGis server, dan Nominatim.

## II.7 WebGis

WebGis adalah salah satu teknologi geografis yang menggunakan *website* untuk menampilkan informasi peta wilayah sehingga dapat diakses dengan fleksibel (Kanthi & Purwanto, 2016). Untuk membangun webGis, diperlukan bahasa markah dan bahasa-bahasa pemrograman lainnya. Komponen website pada umumnya dibagi ke dalam dua bagian, yaitu sisi *FrontEnd* dan sisi *BackEnd*.

*FrontEnd* merupakan sisi atau bagian *website* yang dapat membuat pengguna dapat berinteraksi secara langsung dengan sistem. Bahasa markah yang menjadi komponen utama untuk membangun sebuah *website* ialah HTML (*HyperText Markup Language*). HTML adalah bahasa untuk menggambarkan bagaimana teks, grafik, dan file yang berisi informasi lain diatur dan dihubungkan bersama dalam sebuah halaman *website*. Untuk menampilkan struktur HTML dengan baik, diperlukan bahasa yang dapat mendefinisikan konstruksi gaya seperti *font*, warna, dan posisi, yang digunakan untuk menggambarkan bagaimana informasi di halaman web diformat dan ditampilkan, dua diantaranya adalah CSS (*Cascading Styles Sheets*) atau SCSS (*Syntactically Awesome Styles Sheets*). Untuk menampilkan informasi dari *database*, melakukan proses kalkulasi atau validasi, membuat animasi, menjalankan fitur-fitur dari *website* dengan baik, Javascript merupakan bahasa pemrograman yang tepat dan sangat populer (Meloni, 2012).

*BackEnd* merupakan bagian *website* yang berada di *server website* (Meloni, 2012). Untuk membangun sisi *BackEnd* diperlukan bahasa pemrograman yang dapat berkomunikasi dengan *database* secara langsung, beberapa bahasa pemrograman yang dapat digunakan ialah Python, PHP, dan JavaScript. Adapun produk *database warehouse* yang paling umum digunakan di antaranya *mySQL*, PostgreSQL, MariaDB, MongoDB, SQLAlchemy, dan masih banyak lagi.

Untuk menampilkan informasi geografis pada WebGis, diperlukan sebuah alat yang dapat mentranslasikan koordinat lokasi ke dalam map atau peta wilayah geografis digital. Salah satu alat yang palung umum digunakan pada perkembangan *website* ialah Leaflet. Leaflet adalah pustaka *open source* yang sangat populer dan memiliki fitur peta yang interaktif. Leaflet dirancang dengan mempertimbangkan kesederhanaan, kinerja, dan kegunaan sehingga dapat bekerja secara efisien di semua platform *dekstop* dan seluler utama, dapat diperluas dengan banyak *plugin*, memiliki layanan API, mudah digunakan, dan terdokumentasi dengan baik, serta kode sumber yang sederhana dan mudah dibaca (leafletjs.com).

## **II.8 System Usability Scale**

*System Usability Scale* (SUS) merupakan salah satu alat pengujian usability yang paling populer. SUS dikembangkan oleh John Brooke pada tahun 1986. SUS ini merupakan skala usability yang handal, populer, efektif dan murah. Zahra Sharfina pada tahun 2016 (Sharfina & Santoso, n.d.) menyebutkan daftar pertanyaan SUS dalam Tabel 2.6.

**Tabel 2. 6** Daftar pertanyaan SUS

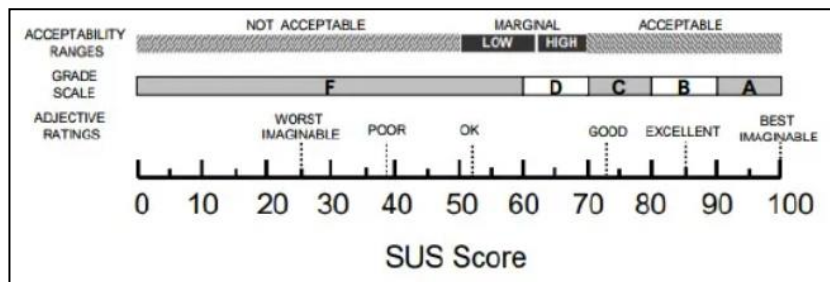
No	Pertanyaan
1	Saya berpikir akan menggunakan sistem ini lagi.
2	Saya merasa sistem ini rumit untuk digunakan.
3	Saya merasa sistem ini mudah untuk digunakan
4	Saya membutuhkan bantuan dari orang lain atau teknisi dalam menggunakan sistem ini.
5	Saya merasa fitur-fitur sistem ini berjalan dengan semestinya.
6	Saya merasa ada banyak hal yang tidak konsisten (tidak serasi) pada sistem ini.
7	Saya merasa orang lain akan memahami cara menggunakan sistem ini dengan cepat.
8	Saya merasa sistem ini membingungkan.
9	Saya merasa tidak ada hambatan dalam menggunakan sistem ini.
10	Saya perlu membiasakan diri terlebih dahulu sebelum menggunakan sistem ini.

SUS sendiri terdiri dari 10 nomor, angka ganjil adalah untuk pernyataan positif dan angka genap untuk sebaliknya. Responden SUS diminta untuk menilai kegunaan dari suatu produk menggunakan 5 skala, yaitu skala nomor 1 (sangat tidak setuju) sampai dengan 5 (sangat setuju). Berikut ini aturan-aturan saat perhitungan skor dengan menggunakan SUS:

1. Setiap pertanyaan bernomor ganjil, skor setiap pertanyaan yang didapat dari skor pengguna akan dikurangi 1.
2. Setiap pertanyaan bernomor genap, skor akhir didapat dari nilai 5 dikurangi skor pertanyaan yang didapat dari pengguna.

3. Skor SUS didapat dari hasil penjumlahan skor setiap pertanyaan yang kemudian dikali 2,5.

Selanjutnya adalah menghitung nilai SUS rata-rata dari tiap responden. Sebuah produk dapat dikatakan memiliki tingkat *usability* yang baik apabila memiliki skor SUS lebih dari atau sama dengan 68. Vaziri pada tahun 2016 menjelaskan interpretasi dari skor SUS pada Gambar 2.5 (Vaziri et al., 2016).



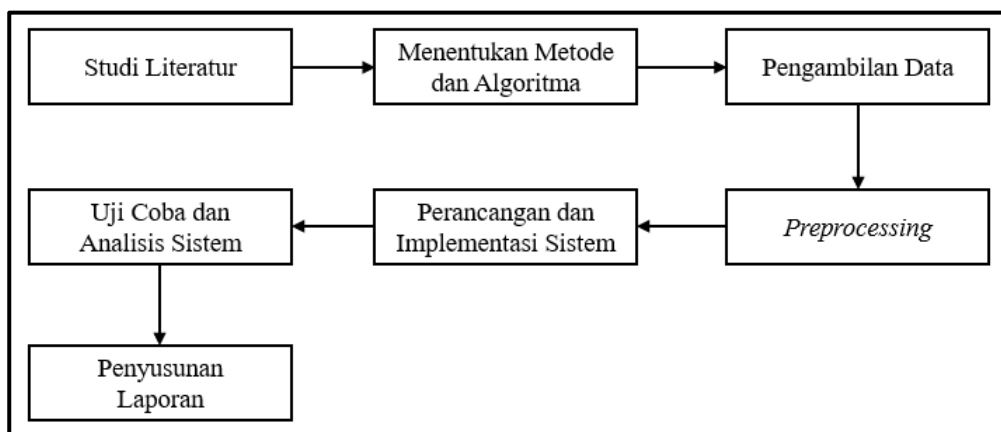
**Gambar 2. 6** Interpretasi skor SUS (Vaziri et al., 2016)

## BAB III

### METODOLOGI PENELITIAN

#### III.1 Tahapan Penelitian

Penelitian ini menggunakan algoritma asosiasi yaitu FP-Growth dan hasil asosiasi akan ditampilkan pada sistem WebGIS. Set data yang digunakan pada penelitian ini adalah data pasien anak klinik *Pediatrica* dari tahun 2018 hingga tahun 2022. Pasien yang terindikasi tuberkulosis akan dikirimkan *link* kuesioner tuberkulosis anak yang dapat dilihat pada Lampiran 1, melalui SMS, WhatsApp, dan *Broadcasat* melalui sistem informasi klinik *Pediatrica*. Sistem WebGIS dengan *data mining* akan menghasilkan *output* berupa asosiasi tiap kecamatan yang terdaftar dalam set data. Gambar 3.1 menunjukkan tahapan penelitian yang dilakukan:



Gambar 3. 1 Tahapan Penelitian

Studi literatur merupakan tahap awal penelitian yang dilakukan. Pada tahap ini, penulis menganalisis permasalahan yang akan diangkat dan mencari literatur atau penelitian lain yang berkaitan dengan permasalahan yang ingin diselesaikan, yaitu terkait penggunaan teknologi *data mining* untuk