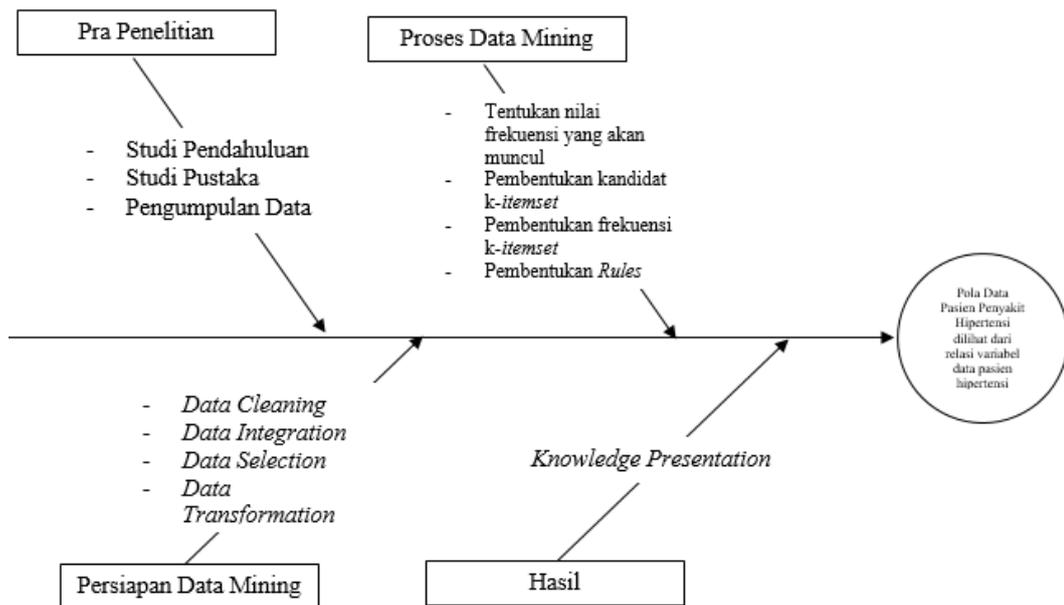


BAB III

METODOLOGI

3.1. Metode Penelitian



Gambar 3.1 Diagram *Fishbone* Tahapan Penelitian

Metodologi penelitian yang digunakan yaitu *fishbone diagram* atau sering juga disebut *Cause-and-Effect Diagram* atau Ishikawa Diagram diperkenalkan oleh Dr. Kaoru Ishikawa, seorang ahli pengendalian kualitas dari Jepang, sebagai alat kualitas dasar (*7 basic quality tools*).

Penelitian ini dibagi menjadi tiga tahapan utama yaitu:

1. Pra penelitian dimulai dengan melakukan studi pendahuluan untuk mengidentifikasi masalah, *literature review* dan pengumpulan data.

2. Persiapan proses data mining terdiri dari, *Data Cleaning*, *Data Authentication*, *Data Integration*, *Data Selection* dan *Data Transformation*.

Proses *Data Mining* dengan melakukan pola relasi data dari pasien dengan diagnosa Hipertensi.

3.2. Pra Penelitian

a) Studi Pendahuluan

Studi pendahuluan dilakukan untuk memperoleh informasi tentang penelitian yang akan dikerjakan. Langkah awal yang dilakukan yaitu menentukan masalah penelitian serta menjajaki dapat tidaknya suatu penelitian dilakukan pada bidang tersebut. Studi pendahuluan dimaksudkan untuk memperjelas kedudukan masalah peneliti dengan keadaan lingkungan atau bidang yang sesungguhnya. Studi pendahuluan dalam memenuhi penelitian ini yaitu dengan mendatangi pihak Dinas Kesehatan Kota Tasikmalaya untuk mendapatkan informasi lengkap mengenai data dan keterangan kasus penyakit Hipertensi.

b) Studi Pustaka

Studi pustaka merupakan sebuah usaha yang dilakukan untuk mendapatkan bahan pengetahuan dan informasi yang relevan, informasi tersebut dapat diperoleh dari buku, laporan penelitian, peraturan-peraturan, ketetapan-ketetapan, jurnal dan sumber tertulis lainnya baik tercetak maupun elektronik.

c) Pengumpulan Data

Di dalam memperoleh data kasus penyakit Hipertensi yang akan digunakan pada proses data mining, diawali dengan melalui proses perijinan untuk memperoleh data, setelah ijin didapatkan kemudian data dapat diperoleh dari instansi terkait sehingga didapatkan data penyakit Hipertensi Kota Tasikmalaya tahun 2018.

3.3. Metode Persiapan Proses *Data Mining*

Persiapan proses data mining perlu dilakukan untuk meningkatkan kualitas data, karena dengan mengetahui data-data yang anomali, memisahkan dan membuangnya dari data training yang diambil bisa menjadikan hasil analisis yang didapatkan jauh lebih baik dari data training yang tidak melalui proses *processing* data mining. (Han, Kamber & Pei, 2012).

Tahapan *Knowledge Discoveries* terdapat beberapa proses yang harus dilakukan pada *data mining* sehingga dapat digunakan untuk proses data mining. *Preprocessing* data mining atau proses mempersiapkan data dari data mentah menjadi data yang siap digunakan, terdapat tahapan terstruktur yang terdiri dari *Data Cleaning*, *Data Integration*, *Data Selection* dan *Data Transformation*.

3.3.1 *Data Cleaning*

Han, Kamber & Pei menjelaskan *Data Cleaning* merupakan proses untuk dapat mengatasi nilai yang hilang, noise dan data yang tidak konsisten. (Han, Kamber & Pei, 2012). Data kasus penyakit Hipertensi diperoleh dari Dinas

Kesehatan Kota Tasikmalaya dengan mencakup 21 UPTD Puskesmas di Kota Tasikmalaya tahun 2018.

Data hipertensi di Dinas Kesehatan Kota Tasikmalaya pada tahun 2018 setelah dilakukan *Data Cleaning* terkumpul sebanyak 11.407 data pasien penyakit hipertensi dengan variabel Jenis Kelamin, Umur, Pekerjaan, dan Klasifikasi Hipertensi. Data tersebut yang nantinya akan dicari kombinasi itemset dengan aturan *Association Rule* menggunakan algoritma Apriori. Variabel yang digunakan untuk kombinasi *itemset* pada aturan Asosiasi adalah :

- 1) Jenis Kelamin, keterangan *gender* pasien dengan diagnose hipertensi dapat digunakan dalam menemukan pola kecelakaan dalam pencarian relasi antar variabel pasien hipertensi.
- 2) Usia, data hipertensi berdasarkan usia digunakan sebagai salah satu variabel yang berhubungan dengan identitas pasien. Keterangan usia dapat menspesifikasikan pasien dalam pendeskripsian. Pengelompokkan usia pasien:
 - 1) 16 – 30
 - 2) 31 – 40
 - 3) 41 – 50
 - 4) 51 – 60
 - 5) 61 – 70
 - 6) 71 – 80
 - 7) >80

3) Profesi, data hipertensi berdasarkan profesi/pekerjaan. Data profesi sebagai berikut:

- a) PNS
- b) TNI/POLRI
- c) Karyawan Swasta
- d) Pelajar/Mahasiswa
- e) Ibu rumah tangga
- f) Sopir
- g) Pedagang
- h) Buruh Harian Lepas
- i) Wiraswasta
- j) Tidak Bekerja
- k) Petani
- l) Mekanik
- m) Guru
- n) Karyawan Honorer
- o) Buruh Nelayan/Perikanan
- p) Pensiunan
- q) Tukang Jahit
- r) Pekerjaan Lainnya
- s) Karyawan BUMN
- t) Dosen

- 4) Klasifikasi data, data hipertensi berdasarkan klasifikasi hipertensi. Data klasifikasi sebagai berikut:
- a) Normal
 - b) Prahipertensi
 - c) Hipertensi1
 - d) Hipertensi2

3.3.2 Data Integration

Menggabungkan data dari banyak sumber data dapat membantu mengurangi redudansi data dan data yang tidak konsisten yang disebabkan pengambilan data dari banyak sumber data, hal ini tentu saja akan berpengaruh terhadap kecepatan dan akurasi saat melakukan data mining. (Han, Kamber & Pei, 2012).

Proses penggabungan pada penelitian ini dilakukan pada saat menggabungkan seluruh data dari 21 UPTD Puskesmas terkait yakni UPTD Puskesmas Bantar, UPTD Puskesmas Bungursari, UPTD Puskesmas Cibeureum, UPTD Puskesmas Cigeureung, UPTD Puskesmas Cihideung, UPTD Puskesmas Cilembang, UPTD Puskesmas Cipedes, UPTD Puskesmas Indihiang, UPTD Puskesmas Kahuripan, UPTD Puskesmas Karanganyar, UPTD Puskesmas Kawalu, UPTD Puskesmas Mangkubumi, UPTD Puskesmas Panglayungan, UPTD Puskesmas Parakanyasag, UPTD Puskesmas Purbaratu, UPTD Puskesmas Sambongpari, UPTD Puskesmas Sangkali, UPTD Puskesmas Sukalaksana, UPTD Puskesmas Tamansari, UPTD Puskesmas Tawang, UPTD Puskesmas Urug.

3.3.3 Data Selection

Meminimalkan jumlah data yang digunakan untuk proses *data mining* dengan tetap merepresentasikan data aslinya. Mengurangi jumlah data yang digunakan untuk proses mining akan lebih efisien mengingat hasil yang didapatkan sama (atau hampir sama) secara analitikal. (Han, Kamber & Pei, 2012).

Data selection pada data hipertensi, semua data hasil dari *cleaning* dan *integration* digunakan seluruhnya agar data yang diolah mencakup keseluruhan data, pada penelitian ini *data selection* dilakukan pada saat memilih jenis hipertensi yang bersifat konsisten.

3.3.4 Data Transformation

Mengubah bentuk dan format data tentunya akan sangat membantu memudahkan pengguna dalam proses data mining ataupun memahami hasil yang didapat. (Han, Kamber & Pei, 2012).

Proses mempermudah data mining data hipertensi pada penelitian ini dilakukan inialisasi data pada kolom tekanan darah diubah menjadi data klasifikasi.

3.4. Proses Data Mining

3.4.1 Association Rule

Tahapan penerapan model *Association Rule* adalah sebagai berikut:



Gambar 3.2 *Flowchart* Analisa Asosiatif

Flowchart di atas terdapat beberapa tahapan yang harus dilakukan ketika menggunakan pemodelan *Association Rule* dengan menggunakan perhitungan algoritma Apriori, sebagai berikut:

1) *Selection Minimum Confidence*

Nilai perhitungan semua proses pada setiap iterasi harus sesuai dengan nilai minimum yang telah ditentukan. Nilai frekuensi yang dihasilkan tidak lebih dari *minimum support* yang telah ditentukan yaitu 0,15 sedangkan untuk nilai

minimum confidence harus diatas nilai 70%. Jika masih ada nilai yang tidak sesuai dengan nilai minimum yang telah ditentukan maka perhitungan akan terus berlanjut dengan kandidat kombinasi *itemset* yang berbeda.

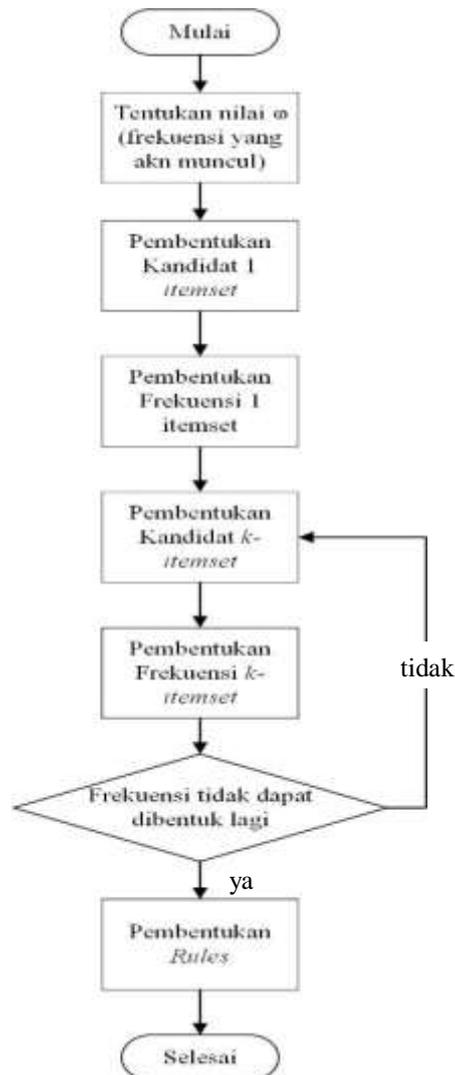
2) Aturan Asosiasi Akhir

Nilai aturan asosiasi akhir adalah nilai yang di dapatkan dari perhitungan nilai *support*. *Itemset* yang memiliki nilai lebih atau sama dengan *minimum support* 0,15 akan di kategorikan pada deretan *itemset* Frekuensi Tinggi, itemset tersebut yang nantinya akan dihitung kembali nilai *confidence* yang dimilikinya dengan rumus *confidence* dengan *minimum confidence* 70%, kemudian akan dihitung nilai untuk membuktikan kuat tidaknya aturan asosiasi yang terbentuk dengan menghitung nilai *lift ratio* dengan batas 1,1.

3) *Rules* yang diperoleh

Hasil akhir dari aturan analisa asosiasi menghasilkan sebuah deklarasi hasil kombinasi *itemset* variabel pasien hipertensi “jika ... maka ...” merupakan deklarasi yang dihasilkan dari fungsi aturan Asosiasi.

Tahapan algoritma Apriori pada aturan *Association Rule*, perhitungan rumus dan hasil kombinasi item didapat melalui perhitungan dan langkah-langkah sebagai berikut:



Gambar 3.3 *Flowchart* Algoritma Apriori

- 1) Tentukan nilai ϕ (frekuensi yang akan muncul)

Proses perhitungan data hipertensi dengan nilai *minimum support* yang digunakan 0,15, nilai *confidence minimum* 70% dan nilai *Lift Ratio* 1,1 untuk menghasilkan pola keterhubungan yang kuat antar variable.

- 2) Pembentukan kandidat I *itemset*

Data hipertensi yang dikumpulkan memiliki jumlah keterangan yang cukup banyak dan menghasilkan berbagai informasi mengenai identitas diri pasien

hipertensi. Variabel yang digunakan dalam pengolahan data hipertensi pada penelitian ini adalah relasi antar variable pasien hipertensi yaitu jenis kelamin, usia, profesi dan klasifikasi. Pembentukan kandidat I *itemset* menghitung jumlah item yang muncul pada setiap kasus hipertensi dan di hitung untuk mencari nilai *Support* (%) atau persentase kemunculan item tersebut.

3) Pembentukan frekuensi I *itemset*

Hasil *itemset* akhir yang didapat dari perhitungan kandidat I *itemset* idengan hasil nilai *Support* (%) lebih atau sama dengan 0,15.

4) Pembentukan kandidat K-*itemset*

Kombinasi antar item dihitung sampai nilai *support minimum* yang diperoleh lebih atau sama dengan 0,15. Item yang dihitung lebih dari 1, maka dilakukan kombinasi *itemset* dari mulai kandidat 2 *itemset*, 3 *itemset* dan seterusnya sampai iterasi berakhir.

5) Pembentukan frekuensi K-*itemset*

Kandidat k-*itemset* merupakan hasil *itemset* akhir yang diperoleh dari perhitungan kandidat k-*itemset* yang menghasilkan nilai *Support* (%) lebih atau sama dengan 0,15.

6) Perhitungan Nilai *Lift Ratio*

Lift ratio dihitung untuk menghitung kekuatan *rules* kejadian acak yang terjadi pada masing-masing kombinasi, sebelum proses perhitungan nilai *lift ratio*, harus dihitung nilai *expected confidence* terlebih dahulu.

3.5. *Knowledge Presentation*

Setelah dilakukan proses data mining, aturan asosiasi menggunakan algoritma Apriori kemudian dilanjutkan *knowledge presentation* yaitu proses yang dapat merepresentasikan informasi yang nantinya informasi tersebut dapat digunakan.