

BAB II

LANDASAN TEORI

2.1 Hipertensi

Hipertensi adalah peningkatan tekanan darah sistolik lebih dari 140 mmHg dan tekanan darah diastolik lebih dari 90 mmHg pada dua kali pengukuran dengan selang waktu lima menit dalam keadaan cukup istirahat/tenang (Kemenkes RI, 2013). Hipertensi didefinisikan sebagai peningkatan tekanan darah sistolik sedikitnya 140 mmHg atau tekanan darah diastolik sedikitnya 90 mmHg (Price & Wilson, 2006).

Peningkatan tekanan darah yang berlangsung dalam jangka waktu lama dapat menyebabkan kerusakan pada ginjal, jantung, dan otak bila tidak dideteksi secara dini dan mendapat pengobatan yang memadai (Kemenkes RI, 2013).

Klasifikasi tekanan darah pada orang dewasa menurut JNC 7 terbagi menjadi kelompok normal, prahipertensi, hipertensi derajat 1, dan hipertensi derajat 2 (Yogiantoro, 2009).

Tabel 2.1 Klasifikasi Tekanan Darah

Klasifikasi Tekanan Darah	Tekanan Darah Sistolik (mmHg)	Tekanan Darah Diastolik (mmHg)
Normal	<120	<80
Prahipertensi	120-139	80-89
Hipertensi derajat 1	140-159	90-99
Hipertensi derajat 2	>160	>100

Adapun klasifikasi hipertensi terbagi menjadi :

1. Berdasarkan Penyebab

a. Hipertensi Primer atau Hipertensi Esensial

Hipertensi yang penyebabnya tidak diketahui (idiopatik), walaupun dikaitkan dengan kombinasi faktor gaya hidup seperti kurang bergerak (anaktivas) dan pola makan. Hipertensi jenis ini terjadi pada sekitar 90% pada semua kasus hipertensi.

b. Hipertensi Sekunder atau Hipertensi Non Esensial

Hipertensi yang diketahui penyebabnya. Pada sekitar 5-10% penderita hipertensi, penyebabnya adalah penyakit ginjal, sekitar 1-2% penyebabnya adalah kelainan hormonal atau pemakaian obat tertentu.

2. Berdasarkan bentuk hipertensi

a. Hipertensi Pulmonal

Suatu penyakit yang ditandai dengan peningkatan tekanan darah pada pembuluh darah arteri paru-paru yang menyebabkan sesak nafas, pusing dan pingsan pada saat melakukan aktifitas. Berdasar penyebabnya hipertensi pulmonal dapat menjadi penyakit berat yang ditandai dengan penurunan toleransi dalam melakukan aktifitas dan gagal jantung kanan. Hipertensi pulmonal primer sering didapatkan pada usia muda dan usia pertengahan, lebih sering didapatkan pada perempuan dengan perbandingan 2:1, angka kejadian pertahun sekitar 2-3 kasus per 1 juta penduduk.

b. Hipertensi Pra Kehamilan

Terdapat 4 jenis hipertensi yang umumnya terdapat pada saat kehamilan, yaitu :

- a) Preeklampsia-eklampsia
- b) Hipertensi kronik
- c) Preeklampsia
- d) Hipertensi gestasional

2.2 *Data Mining*

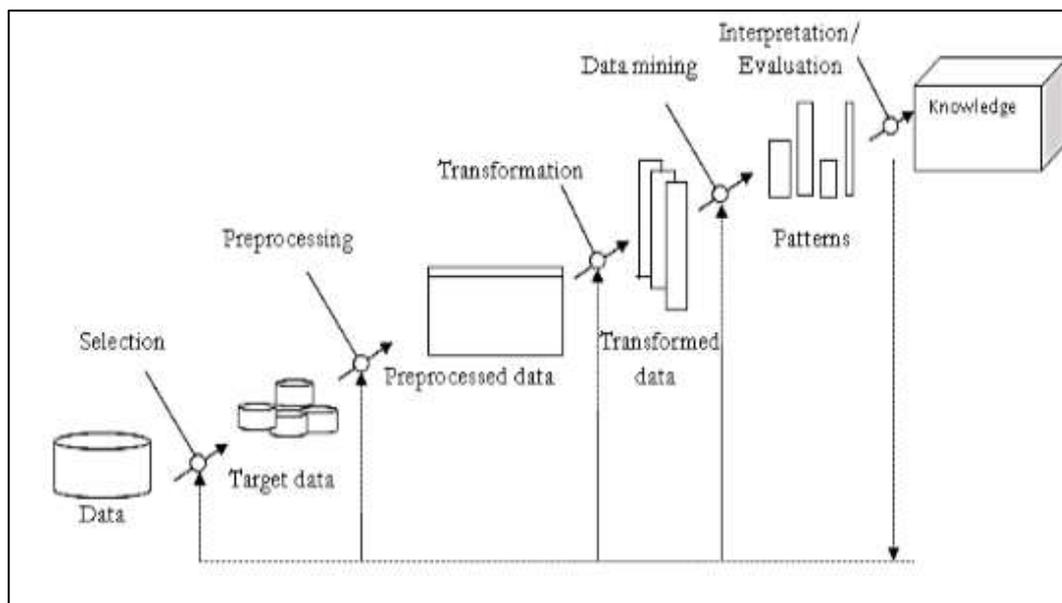
Beberapa tahun terakhir data semakin heterogen dan kompleks dengan volume yang meningkat cepat secara eksponensial. Menurut John Gantz dan David Reinsel dalam investigasi yang dilakukan di IDC, volume data pada tahun 2011 mencapai 1,8 zettabyte atau 1,8 trilyun *gigabyte*, pada tahun 2012 meningkat lebih dari 50% menjadi 2,8 *zettabytes* (Gantz et al. 2012). Tahun 2013 volume data sudah menjadi 4,4 *zettabytes* dan akan terus meningkat dengan cepat hingga mencapai 44 *zetabytes* di tahun 2020 (Turner, 2014). Saat ini dikenal istilah *big data* yang menggambarkan volume data sangat besar, terstruktur maupun tidak terstruktur, yang membanjiri dunia bisnis. *Big data* dapat dianalisis sehingga perusahaan dapat mengambil keputusan-keputusan strategi bisnis dengan lebih baik.

Data mining atau penambangan data dapat didefinisikan sebagai proses seleksi, eksplorasi dan pemodelan dari data dengan jumlah besar untuk menemukan pola atau kecenderungan yang biasanya tidak disadari keberadaannya. Data mining dapat dikatakan sebagai proses mengekstrak

pengetahuan dari data dengan jumlah besar yang tersedia. Pengetahuan yang dihasilkan dari proses data mining harus baru, mudah dimengerti dan bermanfaat. Data disimpan secara elektronik dan di proses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu (Gambbeta, 2012).

2.2.1 Tahapan Proses *Data Mining*

Data mining merupakan salah satu dari rangkaian *Knowledge Discovery in Database (KDD)*. *Knowledge Discovery in Database* merupakan proses yang terstruktur, berikut tahapannya (Han, Kamber & Pei, 2012):



Gambar 2.1 Diagram Proses KDD

1. *Data Cleaning*: Proses membersihkan data dari data noise dan tidak konsisten.
2. *Data Integration*: Proses untuk menggabungkan data dari beberapa sumber yang berbeda.
3. *Data Selection*: Proses untuk memilih data dari database yang sesuai dengan tujuan analisis.

4. *Data Transformation*: Proses mengubah bentuk data menjadi data yang sesuai dengan *data mining*.
5. *Data Mining*: Proses inti yang menggunakan sebuah metode tertentu untuk memperoleh sebuah pola dari data.
6. *Pattern Evaluation*: Proses mengidentifikasi pola.
7. *Knowledge presentation*: Proses yang dapat mempresentasikan informasi yang dibutuhkan dimana informasi yang telah didapatkan kemudian digunakan oleh pemilik data.

2.2.2 Kategori *Data Mining*

Data mining dibagi menjadi dua kategori utama (Han dan Kamber, 2012), pembagian kategori tersebut berkaitan dengan hasil akhir dari pemrosesan data, kategori tersebut yaitu:

1) Prediktif

Prediktif adalah untuk memprediksi nilai dari atribut tertentu berdasarkan pada nilai atribut-atribut lain. Atribut yang diprediksi umumnya dikenal sebagai target atau variabel tak bebas, sedangkan atribut-atribut yang digunakan untuk membuat prediksi dikenal sebagai *explanatory* atau variabel bebas.

2) Deskriptif

Deskriptif adalah untuk menurunkan pola-pola (*corelation, trend, cluster, teritory dan anomaly*) yang meringkas hubungan yang pokok dalam data. Tugas data mining deskriptif sering merupakan penyelidikan dan seringkali memerlukan Teknik *post-processing* untuk validasi dan penjelasan hasil.

2.2.3 Kegunaan *Data Mining*

Fungsi-fungsi yang umum diterapkan dalam *data mining* (Han dan Kamber, 2012), yaitu diantaranya:

- 1) *Assosiation*, adalah proses untuk menemukan aturan asosiasi antara suatu kombinasi item dalam suatu waktu,
- 2) *Sequence*, proses untuk menemukan aturan asosiasi antara suatu kombinasi item dalam suatu waktu dan diterapkan lebih dari satu periode,
- 3) *Clustering*, adalah proses pengelompokkan sejumlah data/*object* ke dalam kelompok data sehingga setiap kelompok berisi data yang mirip,
- 4) *Regression*, adalah proses pemetaan data dalam suatu nilai prediksi,
- 5) *Forecasting*, adalah proses pengestimasian nilai prediksi berdasarkan pola-pola di dalam sekumpulan data,
- 6) *Solution*, adalah proses penemuan akar masalah dan *problem solving* dari persoalan bisnis yang dipakai atau paling tidak sebagai informasi dalam pengambilan keputusan.

2.3 Algoritma Apriori

Algoritma Apriori termasuk jenis aturan asosiasi pada data mining. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut *affinity analysis* atau *market basket analysis*. Analisa Asosiasi atau *Associatin Rule Mining* adalah teknik data mining untuk menemukan aturan suatu kombinasi item. Salah satu tahap analisis Asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis (Pane, 2013).

Algoritma Apriori dapat dijabarkan sebagai berikut, jika C_k menyatakan kandidat *k-itemset* dan F_k menyatakan set *k-itemset* yang frekuen:

- a. Algoritma membaca semua transaksi satu kali untuk menentukan *support* setiap *item*. Selanjutnya pada langkah 2 akan dipilih 1-*itemset* yang frekuen.
- b. Secara *iterative* algoritma akan membangkitkan kandidat *k-itemset* yang baru ditemukan pada iterasi sebelumnya.
- c. Menghitung *support count* dari setiap kandidat, algoritma perlu melakukan pembacaan tertentu pada set data.
- d. Setelah menghitung *support*, algoritma akan membuang semua kandidat *itemset* yang nilai *support count*-nya kurang dari minimum yang ditentukan.

2.4. Analisa Asosiatif (*Association Rule*)

Association rule atau dikenal juga dengan *Market basket analysis* (analisis keranjang belanja) adalah salah satu konsep menarik di dalam *data mining* yang berusaha menemukan asosiasi atau keterkaitan data. Analisa asosiasi juga dikenal sebagai salah satu teknik *data mining* yang menjadi dasar dari berbagai teknik *data mining* lainnya. Alasan konsep asosiasi ini disebut *Market basket analysis* karena pada awal ditemukannya konsep ini berkaitan dengan barang-barang yang berada di dalam keranjang belanjaan yang secara langsung menunjukkan tingkah laku konsumen data saat berbelanja (Suyanto, 2017).

Algoritma Apriori termasuk jenis aturan asosiasi pada *data mining*, selain algoritma tersebut yang termasuk dalam golongan ini adalah metode *generalized rule induction* dan algoritma *hash based* (Tri Vlandari, R, 2017). Contoh aturan asosiasi dari analisa pembeli di suatu pasar swalayan adalah dapat diketahui berdasarkan besarkemungkinan seorang pelanggan membeli margarin bersamaan dengan tepung. Misalkan 90 dari 100 konsumen yang membeli margarin, ternyata

mereka juga membeli tepung. Hal ini berarti asosiasi antara margarin dan tepung sangatlah kuat, misalnya dinyatakan dalam ukuran prosentase sebesar 90% (Suyanto, 2017).

Analisa Asosiatif adalah teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item*. *Interesringness measure* yang dapat digunakan dalam *data mining* adalah (Kusumo, Bijaksana, & Darmantoro, 2016):

- a. *Support*: *Support* dari suatu *Association Rule* adalah presentasi kombinasi *item* tersebut dalam *database* yang mengandung A dan B.
- b. *Confidence*: *Confidence* dari *Association Rule* adalah ukuran ketepatan suatu *rule*, yaitu presentasi kuatnya hubungan antar *item* dalam aturan asosiasi A dan B.
- c. *Minimum support*: parameter yang digunakan sebagai batas frekuensi kejadian atau *support count* yang harus dipenuhi oleh *itemset* data.
- d. *Minimum confidence*: parameter yang mendefinisikan *minimum level* dari *confidence* yang harus dipenuhi oleh *itemset* data.
- e. *Kandidat itemset*: *itemset* yang akan dihitung nilai *support count*.
- f. *Large itemset*: *itemset* yang sering terjadi atau *itemset* yang melebihi batas *minimum support* yang sudah di tentukan.
- g. *Antecedent*: *item* yang merupakan sebab menjadikan *item Consequent*.
- h. *Consequent*: merupakan sebuah akibat setelah kejadian *Ancetudent* dilakukan.

Metode dasar analisis asosiasi terbagi menjadi dua tahap dengan aturan perhitungan masing-masing seperti berikut (Tri Vlandari., R, 2017):

2.4.1 Analisa Pola Frekuensi Tinggi

Tahapan ini yaitu kombinasi item yang akan dicari dengan syarat memenuhi minimum dari nilai *support* dalam *database*.

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ transaksi} \dots\dots\dots (1)$$

Sedangkan nilai *support* dari 2 item diperoleh dari rumus berikut :

$$Support(A \cap B) = \frac{Jumlah\ transaksi\ menggunakan\ A\ dan\ B}{Total\ transaksi} \dots\dots\dots (2)$$

2.4.2 Pembentukan aturan Asosiatif

Tahap berikutnya yaitu mencari aturan asosiatif yang memenuhi syarat minimum untuk *condifence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus berikut:

$$Confidence = P(B/A) = \frac{Jumlah\ transaksi\ mengandung\ A\ dan\ B}{Jumlah\ transaksi\ mengandung\ A} \dots\dots\dots (3)$$

Lift ratio di hitung setelah proses *confidence* selesai dilakukan untuk menghitung kekuatan *rules* kejadian acak yang terjadi pada masing-masing kombinasi. Proses perhitungan nilai *lift ratio*, harus dihitung nilai *expected confidence* terlebih dahulu. Rumusnya adalah sebagai berikut:

$$Expected\ confidence = \frac{Jumlah\ kejadian\ mengandung\ B}{Total\ kejadian} \dots\dots\dots (4)$$

Nilai *lift ratio* diperoleh dari hasil perbandingan nilai *confidence* dengan nilai *expected confidence*. Rumus yang digunakan adalah sebagai berikut:

$$Lift\ Ratio = \frac{Confidence}{Expected\ confidence} \dots\dots\dots (5)$$

Batas nilai *lift ratio* dalam proses perhitungan yaitu 1. Semakin tinggi nilai *lift ratio*, maka semakin besar kekuatan asosiasi yang terjadi pada kombinasi itemset tersebut.

2.5. *Software Tanagra*

Tanagra merupakan perangkat lunak data mining yang bersifat *open source* dengan tujuan akademik dan penelitian. Aplikasi ini mengusulkan beberapa metode penambahan data dari analisis data eksplorasi, pembelajaran statistik, *machine learning* dan *database*.

Aplikasi ini merupakan penerus SIPINA yang mengimplementasikan beberapa algoritma pembelajaran terutama konstruksi interaktif dan visual dari pohon keputusan. Tanagra memiliki kelebihan dalam beberapa pembelajaran dan paradigma seperti halnya pengelompokkan, analisis faktorial, statistik parametrik dan non-parametrik, aturan asosiasi, pemilihan fitur dan algoritma konstruksi. (Rakotomalala R, 2004)

Tujuan dari pengembangan aplikasi Tanagra adalah sebagai berikut: (Rakotomalala R, 2004)

- a. Aplikasi Tanagra memberikan kemudahan melakukan pengolahan data bagi para peneliti (*user*) khususnya proses *data mining* sesuai dengan aturan pengembangan perangkat lunak yang ada (terutama dalam desain GUI dan cara menggunakannya) serta memungkinkan aplikasi untuk menganalisis data nyata atau *sitisis*.

- b. Memberikan usulan kepada pengguna dalam menambah metode pengembangan data sesuai dengan keadaan data yang dimiliki. Tanagra bertindak sebagai *platform* eksperimental dalam proses pengolahan data.
- c. Pengembangan kemampuan bagi para pemula dalam memahami tahapan-tahapan yang harus dilalui dalam proses pembangunan perangkat lunak seperti halnya aplikasi Tanagra, dengan cara tersebut Tanagra dapat dianggap sebagai pedoman untuk mempelajari teknik pemrograman.

2.6. Penelitian Terkait

Tabel 2.2 Penelitian Terkait

No.	Judul	Penulis	Tahun	Tujuan	Metode Penelitian	Hasil
1.	Implementasi <i>Data Mining</i> dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat	Robi Yanto	2015	- Penentuan pola pembelian obat di sektor kesehatan.	1. Algoritma Apriori 2. Algoritma <i>FT Growth</i>	<i>Knowledge</i> dari sistem yang dibangun untuk memudahkan ketersediaan stok.
2.	Implementasi <i>Data Mining</i> Algoritma Apriori Pada Sistem Perse-diaan Alat-alat Kesehatan	Kennedi Tampubolon	2013	- Implementasi Apriori untuk memprediksi persediaan stok barang.	1. Analisa Asosiatif dengan algoritma Apriori	Menemukan kecenderungan pola kombinasi untuk pengambilan keputusan dalam persediaan stok jenis barang.
3.	Penggunaan Al-goritma Apriori <i>Data Mining</i> untuk Mengetahui Tingkat Keseti-aan Konsumen (<i>Brand Loyalty</i>) Terhadap <i>Merk</i> Kendaraan Ber-motor (Studi Kasus <i>Dealer</i> Honda Rumbai)	Wirdah Choiriah	2016	- Sistem prediksi <i>brand loyalty</i> terhadap barang yang dibeli konsumen.	1. <i>Rrule based classification</i>	Sistem pengambil keputusan terhadap <i>brand loyalty</i> .

Tabel 2.2 Penelitian Terkait (Lanjutan)

No.	Judul	Penulis	Tahun	Tujuan	Metode Penelitian	Hasil
4.	Metode <i>Association Rule</i> untuk Analisa Citra <i>CT</i> Organ Pasien Kanker <i>Ovarium</i>	Dwina Kuswardani	2011	- Menemukan kecenderungan suatu data.	1. <i>Association Rules</i>	Analisa kecenderungan atau keterkaitan antar data.
5.	Analisa <i>Data Mining</i> untuk Menentukan Pola Transaksi Obat Menggunakan Algoritma Apriori	Devi Nurmawati	2016	- Menemukan pola kombinasi itemset dari data transaksi obat.	1. Aturan asosiasi dan perhitungan algoritma Apriori	Menemukan pola kombinasi itemset dari data transaksi obat.
6.	Implementasi Apriori pada Data Kecelakaan Lalu Lintas dalam Pencarian Relasi Antar Variabel Pelaku Laka	Rusmin Saragih	2017	- Mencari solusi penanggulangan yang tepat mengenai kecelakaan lalu lintas.	1. Metode <i>Association Rules</i>	Aplikasi identifikasi keterkaitan antara variabel penyebab kecelakaan lalu lintas.

Tabel 2.3 Penelitian Terdekat

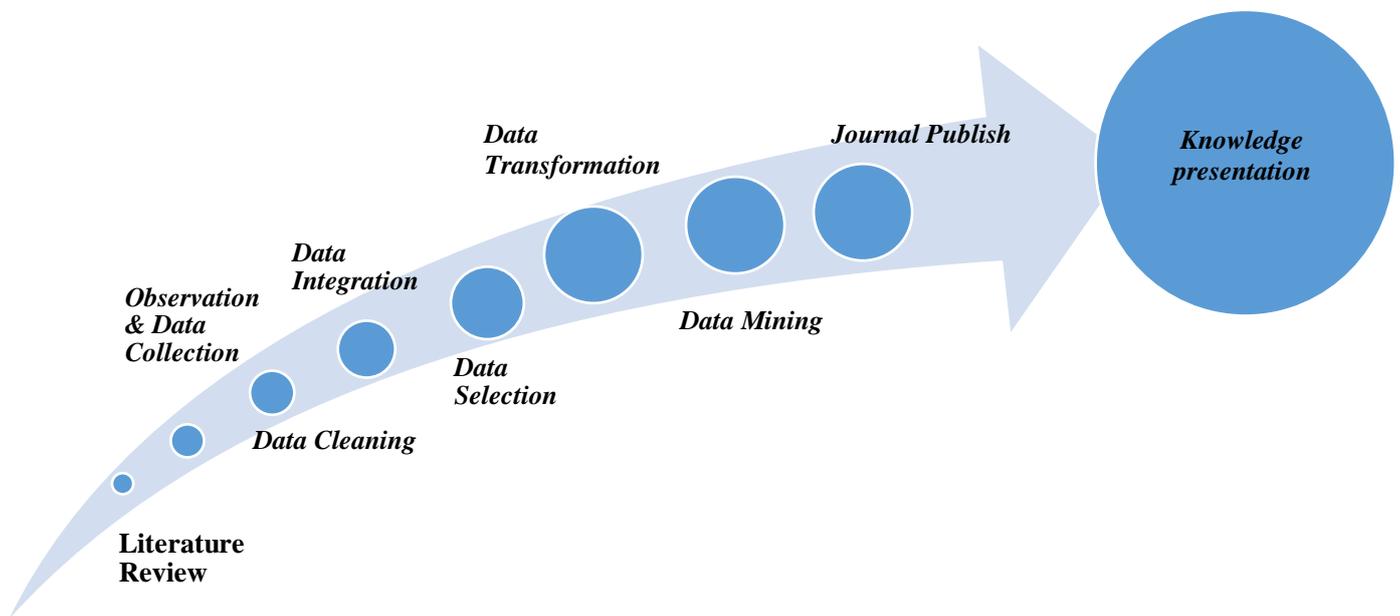
No.	Judul	Penulis	Tahun	Tujuan	Metode Penelitian	Hasil
1.	Metode <i>Association Rule</i> untuk Analisa Citra <i>CT</i> Organ Pasien Kanker <i>Ovarium</i>	Dwina Kuswardani	2011	- Menemukan kecenderungan suatu data.	1. <i>Association Rules</i>	Analisa kecenderungan atau keterkaitan antar data.

Tabel 2.4 Matriks Penelitian

No	Jurnal / Model	Lingkup Penelitian				Author References, Year
		Algoritma	Metode	Data Set	Output	
		Apriori	Association Rule	Kemenkes	Knowledge Presentation	
1	Identifikasi Penyakit Hipertensi dengan Implementasi Metode Association Rules	√	√	√	√	Nur Widiyasono Neng Ika Kurniati Rifan Aprian Hidayat 2019
2	Implementasi <i>Data Mining</i> dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat	√	√		√	Rubiyanto 2015
3	Implementasi <i>Data Mining</i> Algoritma Apriori Pada Sistem Perse-diaan Alat-alat Kesehatan	√	√		√	Kennedi Tampubolon 2013
4	Penggunaan Al-goritma Apriori <i>Data Mining</i> un-tuk Mengetahui Tingkat Keseti- aan Konsumen (<i>Brand Loyalty</i>) Terhadap <i>Merk</i> Kendaraan Ber-motor (Studi Kasus <i>Dealer</i> Honda Rumbai)	√			√	Wirdah Choiriah 2016
5	Metode <i>Association Rule</i> untuk Analisa Citra <i>CT</i> Organ Pasien Kanker <i>Ovarium</i>	√	√		√	Dwina Kuswardani 2011

Tabel 2.4 Matriks Penelitian (Lanjutan)

No	Jurnal / Model	Lingkup Penelitian				Author References, Year
		Algoritma	Metode	Data Set	Output	
		<i>Apriori</i>	<i>Association Rule</i>	<i>Kemenkes</i>	<i>Knowledge Presentation</i>	
6	Analisa <i>Data Mining</i> untuk Menentukan Pola Transaksi Obat Menggunakan Algoritma Apriori	√	√		√	Devi Nurmawati 2016
7	Implementasi Apriori pada Data Kecelakaan Lalu Lintas dalam Pencarian Relasi Antar Variabel Pelaku Laka	√	√		√	Rusmin Saragih 2017



Gambar 2.2 Roadmap Penelitian

Gambar 2.2 terdapat beberapa tahapan penelitian yang terdiri dari:

1. *Literature review*, merupakan proses kajian penelitian yang terdahulu yang berkaitan dengan penelitian ini.
2. *Observation & data collection*, merupakan tahapan untuk melakukan pengamatan masalah dan pengumpulan data untuk penelitian.
3. *Data cleaning*, merupakan proses membersihkan data dari data *noise* atau data yang tidak diperlukan dalam penelitian.
4. *Data integration*, merupakan proses untuk menggabungkan data dari beberapa sumber terkait yang berbeda.

5. *Data selection*, merupakan proses untuk memilih data dari *database* yang sesuai dengan tujuan analisis.
6. *Data transformation*, merupakan proses untuk mengubah bentuk data menjadi data yang sesuai dengan data mining.
7. *Data mining*, merupakan proses inti yang menggunakan metode tertentu untuk memperoleh *knowledge presentation* dari data.
8. *Journal publish*, merupakan proses penyusunan jurnal penelitian dan melakukan publikasi ke laman jurnal terkait.
9. *Knowledge presentation*, merupakan proses yang dapat mempresentasikan informasi yang dibutuhkan dimana informasi yang telah didapatkan kemudian digunakan oleh pemilik data.