

BAB III

LANDASAN TEORI

3.1. Rumah Sakit

Rumah Sakit adalah suatu organisasi yang melalui tenaga medis profesional yg terorganisir serta sarana kedokteran yg permanen menyelenggarakan pelayanan kedokteran, asuhan keperawatan yang berkesinambungan, diagnosis serta pengobatan penyakit yg di derita oleh pasien. (*American Hospital Association; 1974*). Menurut *World Health Organization* rumah sakit adalah bagian integral dari suatu organisasi sosial dan kesehatan dengan fungsi menyediakan pelayanan paripurna (komprehensif), penyembuhan penyakit (kuratif) dan pencegahan penyakit (preventif) kepada masyarakat. Rumah sakit juga merupakan pusat pelatihan bagi tenaga kesehatan dan pusat penelitian medic. Sedangkan di dalam Undang-Undang Republik Indonesia Nomor 44 Tahun 2009 tentang rumah sakit dijelaskan bahwa rumah sakit adalah institusi pelayanan kesehatan yang menyelenggarakan pelayanan kesehatan perorangan secara paripurna yang menyediakan pelayanan rawat inap, rawat jalan, dan gawat darurat. Rumah sakit juga merupakan tempat menyelenggarakan upaya kesehatan yaitu setiap kegiatan untuk memelihara dan meningkatkan kesehatan serta bertujuan untuk mewujudkan derajat kesehatan yang optimal bagi masyarakat. Upaya kesehatan dilakukan dengan pendekatan pemeliharaan, peningkatan kesehatan (promotif), pencegahan penyakit (preventif), penyembuhan penyakit (kuratif) dan pemulihan (rehabilitatif) yang dilaksanakan secara serasi dan terpadu serta berkesinambungan (Siregar, 2004). Berdasarkan Permenkes RI Nomor 986/Menkes/Per/11/1992 pelayanan rumah sakit umum pemerintah Departemen Kesehatan dan Pemerintah Daerah diklasifikasikan menjadi kelas/tipe A,B,C,D dan E (Azwar, 1996)

3.2. Rumah Sakit Jiwa

Rumah Sakit Jiwa termasuk kedalam Rumah Sakit Khusus (kelas E), karena melayani pasien yang menderita penyakit yang lebih dikhususkan, seperti penyakit jiwa, penyakit jantung, penyakit mata dan lainnya. Rumah sakit kelas E merupakan rumah sakit khusus (*Special Hospital*) yang menyelenggarakan hanya satu macam pelayanan kedokteran saja. Pada saat ini banyak tipe E yang didirikan pemerintah, misal Rumah Sakit Jiwa, Rumah Sakit Kusta, Rumah Sakit Paru-paru, Rumah Sakit Jantung, Rumah Sakit Ibu dan Anak. Rumah Sakit merupakan suatu kegiatan yang mempunyai potensi besar menurunkan kualitas lingkungan dan kesehatan masyarakat, terutama yang berasal dari aktivitas medis. Maksud dan tujuan rumah sakit jiwa adalah memberikan pelayanan dibidang kesehatan jiwa, bagi penderita gangguan jiwa, dengan berpegang pada prinsip : Tri Upaya Bina Jiwa, yang terdiri dari beberapa usaha sebagai berikut : Usaha profensi yaitu usaha memberikan penyuluhan dan pencegahan terjadinya gangguan jiwa kemudian usaha kuratif usaha perawatan dan penyembuhan pasien sakit jiwa dan ada dan usaha rehabilitasi yaitu memberi keterampilan untuk kembali kemasyarakat sehingga menjadi insan yang produktif. (Niko, 2012)

3.3. Rekam Medis

Menurut IFHRO (*International Federation Health Record Organization*) adalah *a health record contains all information about a patients, his illness and treatment and the end entries in it are recorded in the order in which event of care occurs* (rekam medis berisi semua informasi mengenai pasien, penyakit, pengobatan, dan rekaman yang didalamnya sesuai dengan urutan pelayanan/perawatan. Rekam medis adalah berkas yang berisikan catatan dan dokumen tentang identitas, anamnesis, diagnosa pengobatan, pemeriksaan, pengobatan, tindakan, pelayanan lain yang diberikan kepada pasien pada sarana pelayanan kesehatan meliputi pendaftaran pasien dimulai dari tempat penerimaan pasien, kemudian bertanggung jawab untuk mengumpulkan, menganalisa, mengolah, dan menjamin kelengkapan berkas rekam medis dari unit rawat jalan, unit rawat inap, unit gawat darurat, dan unit penunjang lainnya (Lubis, 2009).

Dalam PERMENKES No: 269/MENKES/PER/III/2008 yang dimaksud rekam medis adalah berkas yang berisi catatan dan dokumen antara lain identitas pasien, hasil pemeriksaan, pengobatan yang telah diberikan, serta tindakan dan pelayanan lain yang telah diberikan kepada pasien. Catatan merupakan tulisan-tulisan yang dibuat oleh dokter mengenai tindakan-tindakan yang dilakukan kepada pasien dalam rangka pelayanan kesehatan. Bentuk Rekam Medis dalam berupa manual yaitu tertulis lengkap dan jelas dan dalam bentuk elektronik sesuai ketentuan. Rekam medis terdiri dari catatan-catatan data pasien yang dilakukan dalam pelayanan kesehatan Menurut (Ramadhani, 2014) catatan-catatan tersebut sangat penting untuk pelayanan bagi pasien karena dengan data yang lengkap dapat memberikan informasi dalam menentukan keputusan baik pengobatan, penanganan, tindakan medis dan lainnya. Adapun kegunaan rekam medis secara umum adalah :

1. Sebagai alat komunikasi antara dokter dengan tenaga ahlinya yang ikut ambil bagian didalam memberikan pelayanan pengobatan, perawatan kepada pasien.
2. Sebagai dasar untuk merencanakan pengobatan/ perawatan yang harus diberikan kepada pasien.
3. Sebagai bukti tertulis atas segala tindakan pelayanan, perkembangan penyakit, dan pengobatan selama pasien berkunjung/dirawat di rumah sakit.
4. Sebagai bahan yang berguna untuk analisa, penelitian dan evaluasi terhadap kualitas pelayanan yang diberikan kepada pasien.

3.4. Penyakit

Penyakit dapat diklasifikasikan berdasarkan sifat, jenis serta tingkat keseriusannya pada suatu penyakit untuk mempermudah dalam proses penanganan kesehatan. Dalam mengklasifikasikan penyakit SIK (sistem informasi kesehatan) memiliki peran penting, karena pelayanan medik dewasa ini membutuhkan sistem yang lebih efektif dan efisien baik dalam penggunaan, waktu , tenaga ,maupun sarana (Vivi, 2012). Untuk mempermudah dalam proses mengklasifikasikan penyakit,

Indonesia menggunakan sistem informasi kesehatan yang mengacu pada *International Statistical Classification of Diseases (ICD)* yang telah ditetapkan dalam SK Menteri Kesehatan RI No.50/Menkes/SK/I/1998 atau lebih dikenal dengan nama Klasifikasi Internasional Penyakit (KIP/10).

3.5. Penyakit Jiwa

Penyakit jiwa atau biasa disebut dengan penyakit gangguan jiwa atau gangguan mental ialah sindrom atau pola perilaku, atau psikologik seseorang, yang secara klinik cukup bermakna, dan secara khas berkaitan dengan suatu gejala penderitaan (*distress*) atau hendaya (*impairment / disability*) di dalam satu atau lebih fungsi yang penting dari manusia. Sebagai tambahan bahwa disfungsi itu adalah disfungsi dari segi perilaku, psikologik, biologik, dan gangguan itu tidak semata-mata terletak dalam hubungan antara orang itu dengan masyarakat (PPDGJ-III, 2003).

Secara lebih luas gangguan mental (*mental disorder*) juga dapat didefinisikan sebagai bentuk penyakit, gangguan, dan kekacauan fungsi mental atau kesehatan mental, disebabkan oleh kegagalan mekanisme adaptasi dari fungsi-fungsi kejiwaan/mental terhadap stimuli ekstern dan ketegangan-ketegangan; sehingga muncul gangguan fungsional atau struktural dari satu bagian, satu orang, atau sistem kejiwaan/mental (Kartono, 2000). Pendapat yang sejalan juga dikemukakan Chaplin (1981) (dalam Kartono, 2000), yaitu: “Gangguan mental (*mental disorder*) ialah sebarang bentuk ketidakmampuan menyesuaikan diri yang serius sifatnya terhadap tuntutan dan kondisi lingkungan yang mengakibatkan ketidakmampuan tertentu. Sumber gangguan/kekacauannya bisa bersifat psikogenis atau organis, mencakup kasus-kasus reaksi psikopatis dan reaksi-reaksi neurotis yang gawat”

Gangguan jiwa atau gangguan mental juga dapat diartikan sebagai gangguan (*performance*) dalam peran sosial dan pekerjaan tidak digunakan sebagai komponen esensial untuk diagnosis gangguan jiwa, oleh karena ini berkaitan dengan variasi sosial-budaya yang sangat luas. Yang diartikan sebagai ‘*disability*’ keterbatasan atau kekurangan kemampuan atau kekurangan untuk melakukan suatu aktifitas pada

tingkat personal, yaitu melakukan kegiatan hidup sehari-hari yang biasa dan dilakukan untuk perawatan diri dan keberlangsungan hidup (mandi, berpakaian, makan, kebersihan diri, buang air besar dan kecil), (PPDGJIII, "The ICD-10 *classification of Mental and Behavioural Disorder* "). Konsep gangguan jiwa terdapat beberapa butir pengertian menurut (PPDGJ-III, 2003) :

1. Adanya gejala klinis yang bermakna, berupa : Sindrom atau pola perilaku dan sindrom atau pola psikologik
2. Gejala klinis tersebut menimbulkan "penderitaan" (*distress*) antara lain dapat berupa : rasa nyeri, tidak nyaman, tidak tenang, terganggu, disfungsi organ tubuh, dll.
3. Gejala klinis tersebut menimbulkan "disabilitas" (*disability*) dalam aktifitas kehidupan sehari-hari yang biasa dan diperlukan untuk perawatan diri dan kelangsungan hidup (mandi, berpakaian, makan, kebersihan diri, buang air besar dan kecil)

3.6. Klasifikasi Penyakit Jiwa

Penggolongan penyakit berdasarkan sifat, jenis serta tingkat keseriusannya pada suatu penyakit untuk mempermudah dalam proses penanganan kesehatan. Dalam mengklasifikasikan penyakit SIK (system informasi kesehatan) memiliki peran penting, karena pelayanan medik dewasa ini membutuhkan sistem yang lebih efektif dan efisien baik dalam penggunaan, waktu, tenaga, maupun sarana (Vivi, 2012).

Klasifikasi yang paling populer digunakan orang adalah klasifikasi gangguan yang dikemukakan oleh *American Psychiatric association* (APA) pada tahun 1952 yang akhirnya pada tahun 1992 telah berhasil melahirkan *Diagnostic and Statistical Manual of Mental Disorder IV* (DSM-IV), setelah mengalami tiga kali revisi sejak tahun 1979. Di Indonesia, pemerintah telah berhasil melahirkan klasifikasi gangguan kejiwaan yang memuat gangguan kejiwaan yang disebut PPDGJ atau Pedoman Penggolongan Diagnostik Gangguan Jiwa, yang saat ini telah secara resmi digunakan

adalah PPDGJ. Konsep Gangguan Jiwa dari PPDGJ II merujuk ke DSM-III, sedang PPDGJ-III merujuk pada DSM-IV, dan saat ini di Indonesia yang terbaru dan digunakan adalah PPDJ III yang merujuk pada DSM-IV.

Dalam DSM IV dijelaskan bahwa diagnosis multiaksial bertujuan untuk mencakup informasi yang komprehensif (Gangguan Jiwa, kondisi fisik umum, masalah Psikososial dan lingkungan, taraf fungsi secara global), sehingga dapat membantu dalam perencanaan terameramalkan “outcome” atau prognosis. Kemudian format yang “mudah” dan “sistematik”, sehingga dapat membantu dalam menata dan mengkomunikasikan informasi klinis, menangkap kompleksitas situasi klinis dan menggambarkan heterogenitas individual dengan diagnosis klinis yang sama. Selain itu memacu penggunaan “Model bio-psiko-sosial” dalam klinis, pendidikan dan penelitian. Berikut penggolongan diagnosis multiaksial :

1. Aksis I : Gangguan klinis, kondisi lain yang menjadi focus dan perhatian klinis
2. Aksis II : Gangguan kepribadian dan retardasi Mental
3. Aksis III : Kondisi medik umum
4. Aksis IV : Masalah Psikososial dan lingkungan
5. Aksis V : Penilaian fungsi secara global

Antara Aksis I, II, III tidak selalu harus ada hubungan *etiologik* atau *patogenese*. Hubungan antara “Aksis I-II-III” dan “Aksis IV” dapat timbal balik saling mempengaruhi. Berikut pengklasifikasiannya :

1. Aksis I
 - F00-F09 Gangguan Mental Organik & Simtomatik
 - F10-F19 Gangguan Mental & perilaku akibat zat psikoaktif
 - F20-F29 Skizophrenia, Gangguan skizotipal & gangguan waham
 - F30-F39 Gangguan suasana perasaan (afektif/mood)
 - F40-F49 Gangguan neurotik, gangguan somatoform & gangguan terkait stress
 - F50-F59 Sindrom perilaku karena gangguan fisiologis/ fisik

F62-F68	Perubahan Kepribadian karena non organic, gangguan impuls, gangguan seks
F80-F89	Gangguan Perkembangan Psikologis
F90-F98	Gangguan perilaku & emotional onset kanak –remaja
F99	Gangguan Jiwa YTT

2. Aksis II

F60	Gangguan Kepribadian khas
F60.0	Gangguan Kepribadian Paranoid
F60.1	Gangguan Kepribadian schizoid
F60.2	Gangguan Kepribadian dissosial
F60.3	Gangguan Kepribadian emosional tak stabil
F60.4	Gangguan Kepribadian histrionik
F60.5	Gangguan Kepribadian anankastik
F60.6	Gangguan Kepribadian cemas(menghindar)
F60.7	Gangguan Kepribadian dependen
F60.8	Gangguan Kepribadian khas lainnya
F60.9	Gangguan Kepribadian YTT
F61	Gangguan Kepribadian Campuran dan lainnya
F61.0	Gangguan Kepribadian Campuran
F61.1	Perubahan Kepribadian yang bermasalah
F70 -F79	Retardasi Mental

3. Aksis III

Bab I	A00-B99	Penyakit infeksi dan parasit tertentu
Bab II	C00-D48	Neoplasma
Bab IV	E00-G90	Penyakit endokrin, Nutrisi, & metabolik
Bab VI	G00-G99	Penyakit susunan syaraf
Bab VII	H00-H59	Penyakit Mata & adneksa
Bab VIII	H60-H95	Penyakit telinga & Prosesus Mastoid
Bab IX	I00-I99	Penyakit sistem sirkulasi

Bab X	J00-J99	Penyakit sistem Pernafasan
Bab XI	K00-K93	Penyakit sistem Pencernakan
Bab XII	L00-L99	Penyakit kulit & jaringan subkutan
Bab XIII	M00-M99	Penyakit sistem muskuloskeletal & Jaringan ikat
Bab XIV	N00-N99	Penyakit sistem genito-urinaria
Bab XV	O00-O99	Kehamilan, kelahiran anak & masa Nifas
Bab XVII	Q00-Q99	Malformasi congenital, deformasi, Kel.
Bab XVIII	R00-R99	Gejala, tanda & temuan klinis-lab.
Bab XIX	S00-T98	Cedera, keracunan & akibat kausa ekst
Bab XX	V01-V98	Kausa eksternal dari Morb. & mort.
Bab XXI	Z00-Z99	Faktor status kes. & Pelayanan ke

4. Aksis IV

Masalah dengan “Primary support group” (keluarga)

Masalah berkaitan dengan lingkungan sosial

Masalah Pendidikan

Masalah Pekerjaan

Masalah Perumahan

Masalah Ekonomi

Masalah Akses ke pelayanan Kesehatan

Masalah Berkaitan interaksi dengan hukum/kriminal

Masalah Psikososial & Lingkungan lain

5. Aksis V

GLOBAL ASSESSMENT OF FUNCTIONING (GAF) SCALE

100-91 Gejala tidak ada, berfungsi maksimal, tidak ada masalah yang tak tertanggulangi.

90-81 Gejala minimal, berfungsi baik, cukup puas, tidak lebih dari masalah harian yang biasa.

- 80-71 Gejala sementara dan dapat diatasi, disabilitas ringan dalam sosial, pekerjaan, sekolah dll
- 70-61 Beberapa gejala ringan dan menetap, disabilitas ringan dalam fungsi, secara umum masih baik.
- 60-51 Gejala sedang (moderate), disabilitas sedang.
- 50-41 Gejala berat (serious), disabilitas berat.
- 40-31 Beberapa disabilitas dalam hubungan dengan realita & komunikasi, disabilitas berat dalam beberapa fungsi.
- 30-21 Disabilitas berat dalam komunikasi dan daya nilai, tidak mampu berfungsi hampir semua bidang.
- 20-11 Bahaya mencederai diri/orang lain, disabilitas sangat berat dalam komunikasi dan mengurus diri.
- 10-01 Seperti diatas => persisten dan lebih serius.
- 0 Informasi tidak kuat.

3.7. Analisis Deskriptif

Dalam analisis deskriptif kebanyakan tidak dimaksudkan untuk menguji hipotesis tertentu, melainkan lebih pada menggambarkan apa adanya suatu gejala, variabel, atau keadaan. Namun demikian, tidak berarti semua analisis deskriptif tidak menggunakan hipotesis. Penggunaan hipotesis dalam penelitian deskriptif bukan dimaksudkan untuk diuji melainkan bagaimana berusaha menemukan sesuatu yang berarti sebagai alternatif dalam mengatasi masalah penelitian melalui prosedur ilmiah. (Widodo dan Mukhtar, 2000)

3.8. Basis Data

Urutan kedua setelah *database* adalah tabel, sering disebut entitas. Apabila *database* diibaratkan dengan gudang/rumah, maka tabel adalah kamarnya. Fungsi kamar dalam rumah adalah menyimpan barang yang bersifat khusus. Oleh karena itu, penggunaan basis data dalam *data mining* menjadi sebuah kewajiban dalam

mendokumentasikan sebuah data atau lebih dalam sebuah media penyimpanan. (Sari, 2008)

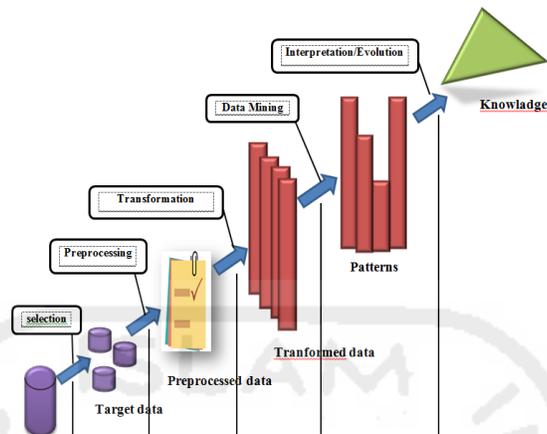
3.9. Data Mining

Secara sederhana *data mining* adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. *Data mining* juga disebut sebagai serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. *Data mining*, sering juga disebut sebagai *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Larose, 2005)

Menurut Gartner Group data mining adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan, dengan menggunakan teknik (Larose, 2005). Dalam data mining otomatisasi tidak menggantikan campur tangan manusia. Selain itu manusia harus ikut aktif dalam setiap fasa dalam proses data mining. Kehebatan kemampuan algoritma data mining yang terdapat dalam perangkat lunak analisis yang terdapat saat ini memungkinkan terjadinya kesalahan penggunaan yang berakibat fatal. Menurut Pramudiono (2006), data mining adalah analisa otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola atau kecenderungan yang penting yang biasanya tidak disadari keberadaannya.

3.10. Tahapan Data Mining

Data mining merupakan suatu rangkaian proses, *data mining* atau KDD dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar 3.1. berikut ini:



Gambar 3.1 Langkah-langkah Datamining

Adapun langkah-langkah di atas dijelaskan sebagai berikut (Sari, 2008):

1. *Data selection*

Proses menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada *subset* variabel, dimana penemuan (*discovery*) akan dilakukan. Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge discovery in database* dimulai.

2. *Preprocessing*

Mempersiapkan data, meliputi dua hal yaitu *data cleaning* (membersihkan data) dan *data reduction*. *Data cleaning* mencakup antara lain membuang duplikasi data, melakukan penghalusan data, memeriksa data yang inkosisten dan memperbaiki kesalahan pada data. *Data reduction* dilakukan untuk mengatasi ukuran data yang terlalu besar. Ukuran data yang terlalu besar dapat menimbulkan ketidakefisienan proses dan peningkatan biaya pemrosesan.

3. *Transformation*

Menggabungkan data kedalam bentuk yang sesuai untuk penggalian lewat operasi *summary*. Pencarian fitur-fitur yang berguna untuk mempresentasikan data tergantung kepada tujuan yang ingin dicapai. Transformasi data dilakukan untuk memudahkan dalam menganalisis dengan *software* pendukung teknik *data mining*.

4. *Data mining*

Proses *data mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih 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 *knowledge discovery in database* secara keseluruhan.

5. Interpretation

Dalam proses ini, *pattern* atau pola-pola yang telah diidentifikasi oleh sistem kemudian diterjemahkan/diinterpretasikan dengan menggunakan suatu program ke dalam bentuk *knowledge* (pengetahuan/informasi) yang lebih mudah dimengerti oleh pihak yang berkepentingan.

3.11. Aturan Asosiasi (*Association Rule*)

Agrawal dan Srikant menyatakan bahwa *association rules* berfungsi untuk menemukan asosiasi antar variabel, korelasi atau suatu struktur antara *item* atau objek-objek didalam *database* transaksi, *database relasional*, maupun pada penyimpanan informasi lainnya (Agrawal & Srikant, 1995).

Association rule atau aturan asosiasi adalah teknik dalam *data mining* untuk mencari hubungan antar *item*. Himpunan *item* disebut sebagai *itemset*. *Itemset* yang mengandung *k items* merupakan *k-itemset*. Kecendrungan kemunculan *itemset* dalam sejumlah transaksi disebut *frequency*. Penting atau tidaknya suatu aturan asosiasi dapat diketahui dengan dua parameter, yakni *support* dan *confidence*. *Support* adalah persentase kombinasi *item* dalam *database*. *Support* merupakan ukuran yang menunjukkan besar tingkat dominasi suatu *item* atau *itemset* keseluruhan transaksi (Zhao, 2014).

Support untuk aturan “X=>Y” yaitu probabilitas atribut atau kumpulan atribut X dan Y yang terjadi secara bersamaan. Dengan demikian, nilai *support* sebuah *item* diperoleh dengan rumus berikut (Zhao, 2014):

$$\text{Support}(X) = P(X) = \frac{n(X)}{n(S)} \dots\dots\dots \text{III}(1)$$

Sehingga,

$$\text{Support}(X) = \frac{\text{Jumlah transaksi yang mengandung } X}{\text{Jumlah transaksi keseluruhan}}$$

dengan :

$P(X)$ = probabilitas kejadian X

$n(X)$ = banyaknya anggota X

$n(S)$ = banyaknya anggota S

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

$$\text{Support}(X \Rightarrow Y) = P(X \cup Y) = \frac{n(X \cup Y)}{n(S)} \dots \dots \dots \text{III(2)}$$

$$\text{Support}(X \Rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } X \text{ dan } Y}{\text{Jumlah transaksi keseluruhan}}$$

dengan:

$P(X \cup Y)$ = probabilitas kejadian X dan Y secara bersamaan

$n(X \cup Y)$ = banyaknya kejadian X dan Y secara bersamaan

$n(S)$ = banyaknya anggota S

Support dalam penelitian ini didefinisikan sebagai probabilitas kejadian beberapa *item* (diagnosa) yang diderita atau yang dialami satu orang pasien dari keseluruhan hasil diagnosa pada data rekam medis. Dalam penelitian ini akan ditentukan pula nilai *minimum support* dengan tujuan untuk menghasilkan *item* dari suatu kumpulan data hasil diagnosa yang memberikan frekuensi paling banyak dari seluruh hasil diagnosa. *Minimum support* adalah parameter yang digunakan sebagai batasan frekuensi kejadian atau *support count* yang harus dipenuhi suatu kelompok data untuk dijadikan aturan.

Sementara itu, *confidence* yaitu kuatnya hubungan antara *item* dalam *association rules*. *Confidence* aturan “X =>Y” yaitu probabilitas terjadinya beberapa *item* secara bersamaan dimana salah satu *item* sudah pasti terjadi. Dengan demikian,

nilai *confidence* sebuah kombinasi *item* diperoleh dengan rumus berikut (Zhao, 2014):

$$\text{Confidence } (X \Rightarrow Y) = P(Y/X) = \frac{P(XUY)}{P(X)} \dots\dots\dots \text{III(3)}$$

Sehingga,

$$\text{Confidence } (X \Rightarrow Y) = \frac{\text{Jumlah transaksi mengandung } X \text{ dan } Y}{\text{Probabilitas kejadian } X}$$

$P(Y/X)$ = probabilitas bersyarat dari kejadian Y bila kejadian X telah terjadi

$P(X \cup Y)$ = probabilitas kejadian X dan Y secara bersamaan

$P(X)$ = probabilitas kejadian X

Confidence dalam penelitian ini didefinisikan sebagai probabilitas kejadian beberapa *item* (diagnosa) diderita atau dialami oleh satu pasien. Dalam penelitian ini akan ditentukan pula nilai *minimum confidence* yaitu parameter yang mendefinisikan minimum level dari *confidence* yang harus dipenuhi oleh aturan berkualitas.

Selain kedua parameter tersebut, salah satu cara yang lebih baik untuk mengetahui kekuatan suatu aturan asosiasi adalah dengan melihat nilai *lift ratio*. *Lift ratio* menunjukkan adanya tingkat kekuatan *rule* atas kejadian acak dari *antecedent* (X) dan *consequence* (Y) berdasarkan pada *supportnya* masing-masing (Zhao, 2014).

Menurut Fomby (2011) dalam (Septiani, 2015), *lift ratio* digunakan untuk mengukur seberapa kuat *rule* yang dibentuk dari algoritma *sequential pattern mining*. Nilai *lift ratio* berkisar antara 0 sampai dengan tak terhingga. Nilai *minimum* dari *lift ratio* tidak ditentukan seperti halnya *support* atau *confidence*. Jika nilai *lift ratio* kurang dari 1 dalam hal ini adalah nilai *minimum* maka *rule antecedent* berpengaruh negatif pada *rule consequent*. Jika nilai *lift ratio* sama dengan 1 maka *rule* tersebut sering muncul bersamaan tetapi *independen*. *Rule* yang *independen*

merupakan *rule* dimana untuk mendapatkan *consequent* tidak tergantung pada *antecedent*. Pada *lift ratio*, *rule* yang direkomendasikan adalah jika *lift ratio* lebih dari 1 karena *antecedent* memiliki pengaruh positif pada *consequent*

$$\text{Lift ratio } (X \Rightarrow Y) = \frac{P(X \cup Y)}{P(X)P(Y)} \dots\dots\dots \text{III(3)}$$

sehingga,

$$\begin{aligned} \text{Lift ratio} &= \frac{\text{Confidence } (X \Rightarrow Y)}{P(Y)} \\ &= \frac{\text{Confidence } (X \Rightarrow Y)}{\text{Support}(Y)} \end{aligned}$$

dengan :

$P(Y/X)$ = probabilitas kejadian X dan Y secara bersamaan

$P(X)$ = probabilitas kejadian X

$P(Y)$ = probabilitas kejadian Y

3.12. Sequential Patterns

Sequential patterns adalah pola yang menggambarkan urutan waktu terjadinya peristiwa (Agrawal, 1995). Pola tersebut bisa ditemukan jika data yang disimpan relatif besar, dan objek yang sama dalam jumlah yang relatif besar melakukan beberapa transaksi yang berulang kali. Misalnya, pelanggan yang memiliki identitas yang terekam, melakukan transaksi belanja berulang kali pada sebuah pusat perbelanjaan.

Dalam sebuah basis data yang menyimpan *market basket data*, terdapat data sejumlah transaksi dan masing-masing transaksi terdiri dari field berupa *id_pelanggan*, *waktu_transaksi*, *daftar_item*. Pada seluruh data yang tersimpan, hanya ada satu transaksi yang dilakukan oleh satu pelanggan pada satu waktu. Dalam suatu transaksi, banyaknya *item* yang dibeli tidak dipertimbangkan, namun diperlakukan sebagai sebuah variabel biner yang merepresentasikan

sebuah item dibeli atau tidak Sebagai contoh proses sequential pattern mining, terdapat tabel transaksi penjualan yang berisikan customer, tanggal dan item. Dari tabel transaksi tersebut kemudian dibentuk sequence transaksi berdasarkan customer dan diurutkan berdasar tanggal sehingga membentuk beberapa sequence (Agrawal, 1995).

3.13. Algoritma SPADE (*Sequential Patterns Discovery Using Equivalent Classes*)

Sequential Pattern Discovery using Equivalent classes adalah algoritma sequential pattern mining yang menggunakan format data vertikal pada database sequence. Dalam format data vertikal, database sequence menjadi berbentuk kumpulan urutan yang formatnya [itemset :(sequence_ID, eventID)]. Dengan kata lain, untuk setiap itemset akan disimpan sequence identifier dan event identifier yang berkoresponden. Event identifier berguna sebagai timestamp atau penanda waktu dari itemset tersebut. Sepasang (sequence_ID, eventID) untuk setiap itemset membentuk ID_list dari itemset tersebut. (Juliastio & Gunawan, 2015)

Langkah-langkah algoritma SPADE dalam mencari frequent sequence kemudian menentukan rule dari frequent sequence tersebut adalah sebagai berikut (Zaki, 2001)

1. Menghitung frequent *1-sequence*

Untuk mencari frequent 1-sequence dari sequence database yang harus dilakukan adalah dilakukan scan untuk setiap itemset dalam sequence database. Untuk masing-masing itemset, simpan id-listnya (pasangan sid dan eid). Kemudian scan id-list dari masing-masing id-list tersebut, setiap ditemui sid yang sebelumnya belum ada maka nilai supportnya ditambah. Sequence yang dimasukkan dalam frequent *1-sequence* adalah yang supportnya lebih dari min_sup.

2. Menghitung frequent 2-sequence

Dalam mencari frequent 2-sequence, data yang digunakan adalah data dari frequent 1-sequence, sehingga tidak perlu mencari dari sequence database lagi. Untuk setiap masing-masing frequent 1-sequence, gabungkan dengan semua frequent 1-sequence lainnya. Contohnya jika 1-sequence A digabungkan dengan 1-sequence B maka kemungkinan 2-sequence yang terjadi adalah A,B dimana A dan B muncul bersamaan dalam transaksi, A→B dimana item B muncul setelah item A, dan B→A dimana item B muncul setelah item A. Untuk setiap masing-masing penggabungan frequent 1-sequence ini dilakukan pengecekan apakah dalam id-listnya memiliki sid yang sama, jika sama kemudian dilakukan pengecekan apakah eid dari 1-sequence A sama dengan, kurang dari atau lebih dari eid 1-sequence B. Apabila sama maka id-listnya dimasukkan dalam 2-sequence A, B. Jika eid B lebih besar dari A maka id-listnya dimasukkan dalam 2-sequence A→B dan jika eid A lebih besar dari B maka id-listnya dimasukkan dalam 2-sequence B→A. Kemudian seperti dalam frequent 1-sequence, tambahkan supportnya untuk setiap masing – masing sid yang sebelumnya belum ditemui. Dari 2-sequence itu kemudian dilakukan pengecekan apakah supportnya lebih dari min_sup. Jika memenuhi syarat maka dimasukkan dalam frequent 2-sequence.

3. Menentukan frequent k-sequence.

Setelah mencari frequent 2-sequence, untuk mencari frequent sequence–frequent sequence berikutnya dilakukan proses yang sama, yaitu mencari frequent k-sequence. Untuk mencari frequent k-sequence ini dilakukan join pada frequent (k-1) sequence yang memiliki prefix yang sama. Contohnya untuk mencari 3-sequence, gabungkan frequent sequence dari 2-sequence yang memiliki prefix yang sama, untuk mencari 4-sequence, gabungkan frequent sequence dari 3-sequence yang memiliki prefix yang sama, dan seterusnya. Untuk mencari

prefix frequent (k-1) sequence, hilangkan item terakhir dari sequence tersebut. Contoh jika terdapat 4 -sequence $A \rightarrow B \rightarrow C \rightarrow D$, maka prefixnya adalah $A \rightarrow B \rightarrow C$. Untuk setiap penggabungan ini ada 3 kemungkinan hasil:

- a. Jika A, B digabungkan dengan A, C, maka kemungkinan hasilnya hanya A, B, C.
- b. Jika A, B digabungkan dengan $A \rightarrow C$, maka kemungkinan hasilnya hanya $A, B \rightarrow C$.
- c. Jika $A \rightarrow B$ digabungkan dengan $A \rightarrow C$, maka ada 3 kemungkinan hasil: $A \rightarrow B, C$, dan $A \rightarrow B \rightarrow C$ dan $A \rightarrow C \rightarrow B$.

Dari setiap kemungkinan ini, periksa supportnya apakah memenuhi min_sup , jika ya maka sequence itu termasuk dalam frequent k -sequence. Pencarian frequent sequence dihentikan apabila tidak ada frequent (k-1) sequence yang bisa dijoin atau sudah tidak ditemukan frequent k -sequence lagi.

4. Pembentukan Rule

Setelah ditemukan semua frequent sequence, ditentukan rule dari sequence – sequence tersebut. 1-sequence tidak digunakan untuk membentuk rule karena hanya terdiri dari 1 item. Untuk 2-sequence yang menjadi antecedent adalah item pertama dan consequentnya adalah item keduanya. Contoh untuk sequence $A \rightarrow B$ maka rule yang dibentuk adalah $A \Rightarrow B$. Sedangkan untuk sequence yang panjangnya lebih dari 2 atau k - sequence, yang dijadikan consequent adalah item terakhir, sedangkan antecedentnya adalah semua item sebelum item terakhir. Contohnya pada 4-sequence $A \rightarrow B \rightarrow C \rightarrow D$, maka rule yang dihasilkan adalah $A \rightarrow B \rightarrow C \Rightarrow D$.

Untuk masing-masing rule dihitung nilai confidencenya. Jika rule tersebut memenuhi batas min_conf , maka rule itu diterima. Kemudian dari rule yang diterima tersebut, hitung nilai lift rasionya Nilai lift rasio semakin besar semakin baik, dengan batas 1. Apabila rule memiliki nilai lift lebih dari atau sama dengan

1, maka dalam rule itu *antecedent* memiliki pengaruh positif terhadap consequent. Sehingga rule dinyatakan baik, sebaliknya apabila nilai lift kurang dari 1 maka rule dianggap kurang baik.

