



ITS
Institut
Teknologi
Sepuluh Nopember

29952/H/07



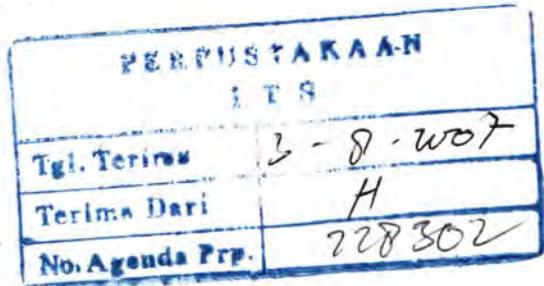
RS Ma
518.1
AKR
P-1
2007

TUGAS AKHIR - SM 1330

**PENCARIAN POLA ASOSIASI SPASIAL KASUS
DBD DI SURABAYA MENGGUNAKAN
ALGORITMA APRIORI**

HARDIAN OKTAVIANTO
NRP 1200 100 072

Dosen Pembimbing
Budi Setiyono, S.Si, MT



JURUSAN MATEMATIKA
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember
Surabaya 2007



ITS
Institut
Teknologi
Sepuluh Nopember

FINAL PROJECT - SM 1330

**FINDING SPATIAL ASSOCIATION PATTERNS
OF DHF INCIDENCE IN SURABAYA USING
APRIORI ALGORITHM**

HARDIAN OKTAVIANTO
NRP 1200 100 072

Supervisor
Budi Setiyono, S.Si, MT

MATHEMATICS DEPARTMENT
Faculty of Mathematics and Science
Sepuluh Nopember Institute of Technology
Surabaya 2007

LEMBAR PENGESAHAN
**PENCARIAN POLA ASOSIASI SPASIAL KASUS DBD
DI SURABAYA MENGGUNAKAN ALGORITMA
APRIORI**

TUGAS AKHIR
Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Bidang Studi Ilmu Komputer
Program Studi S-1 Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh:
HARDIAN OKTAVIANTO
NRP. 1200 100 072

Disetujui oleh Dosen Pembimbing Tugas Akhir :

Budi Setiyono, S.Si, MT.
NIP. 132 161 183



SURABAYA, AGUSTUS 2007

(.....)

PENCARIAN POLA ASOSIASI SPASIAL KASUS DBD DI SURABAYA MENGGUNAKAN ALGORITMA APRIORI

Nama Mahasiswa : Hardian Oktavianto
NRP : 1200 100 072
Jurusan : Matematika FMIPA-ITS
Dosen Pembimbing : Budi Setiyono, S.Si, MT

Abstrak

Spatial association rule adalah salah satu metode pada data mining spasial untuk mendapatkan suatu hubungan/relasi, asosiasi, dan korelasi dalam suatu basis data spasial. Proses rule mining terdiri dari dua buah tahap yaitu pencarian frequent itemset dan generate rule. Rule yang terbentuk merupakan kombinasi diantara item - item pada frequent itemset. Preproses data dilakukan untuk mendapatkan data yang siap digunakan pada keseluruhan proses mining.

Pada tugas akhir ini diaplikasikan metode spatial association rule mining dalam bidang kesehatan dengan fokus kasus penyakit menular Demam Berdarah Dengue (DBD) di kotamadya surabaya menggunakan algoritma apriori. Data spasial yang digunakan adalah data letak geografis rawa dan fasilitas kesehatan, sedangkan untuk data non spasial yang digunakan adalah data jumlah kasus DBD dan data kepadatan penduduk di kotamadya Surabaya. Pada preproses data dilakukan kategorisasi berdasarkan parameter yang dimasukkan oleh user, kategorisasi untuk jumlah kasus DBD dan kepadatan penduduk adalah rendah, sedang, atau tinggi, sedangkan kategorisasi untuk fasilitas kesehatan dan rawa adalah dekat atau jauh. Pada pencarian frequent itemset digunakan algoritma apriori.

Dari hasil penelitian didapatkan pola asosiasi yang sering terjadi adalah suatu kelurahan dengan tingkat kepadatan yang tinggi akan memiliki jumlah kasus DBD yang tinggi.

Kata Kunci: Data Mining, Spatial Data Mining, Association Rule, Spatial Association Rule

FINDING SPATIAL ASSOCIATION PATTERNS OF DENGUE HAEMORRHAGIC FEVER INCIDENCE IN SURABAYA USING APRIORI ALGORITHM

Name : Hardian Oktavianto
NRP : 1200 100 072
Department : Matematika FMIPA-ITS
Supervisor : Budi Setiyono, S.Si, MT

Abstract

Spatial association rule is a method in spatial data mining to discover connection/relation, association, and correlation from spatial database. The general process of rule mining consists of two main parts that is finding frequent itemset dan rule generation. Rules are formed from the combination between items in frequent itemsest. Data Preprocessing is needed to prepare data for the whole mining process.

In this research spatial association rule method was applied using apriori algorithm in healthy major focused on Dengue Haemorrhagic Fever (DHF) incidence in Surabaya. Geographic location of swamps and health facilities are used as spatial data, while the density number and coincidence number of DHF data was selected as non spatial data. Categorization in data preprocessing based on the user's parameter, coincidence of DHF and number of density are categorized to low, average, or high, and geographic location of swamps and health facilities are categorized to far away or close to. The apriori algorithm is implemented to find frequent itemsei.

From the preliminary results show that mostly the association patterns formed while the districts with high density number are also have large number of DHF coincidence.

Keywords: Data Mining, Spatial Data Mining, Association Rule, Spatial Association Rule

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir ini.

Penyusunan tugas akhir ini tidak lepas dari bantuan dan dukungan berbagai pihak selama pengerjaannya, oleh karena itu dengan segala hormat penulis mengucapkan terima kasih kepada :

1. Bapak dan Ibu tercinta, Suswanto dan Harofah atas segala cinta dan kasih sayang, semangat, dan pelajaran hidup kepada saya.
2. Dosen pembimbing, Budi Setiyono, S.Si, MT.
3. Para dosen penguji, DR. Moch. Isa Irawan, MT, Drs. Lukman Hanafi, M.Sc, Drs. Soetrisno, MIKOM.
4. Imam Mukhlas, S.Si, MT, atas segala tuntunan dan bimbingannya.
5. Seluruh staf pengajar Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Teknologi Sepuluh Nopember Surabaya, yang telah membimbing dan memberi ilmu dan pengetahuan kepada penulis selama penulis belajar di kampus tercinta.
6. Adik tersayang, Dina Prameswati Susnanda atas segala dukungan moralnya.
7. Ahmat "Cak Mat" Afif Supianto S.Si.
8. Arek – arek angkatan 2000, Teddy "Gembluk" Ramdhan Syah, Lutfi Ali M. S.Si, Zainul Arifin, S.Si, Aan Eko S., Adin Lisdiono, S.Si, Adi Poerwanto, Nur "Komeng" Qomari, Wahyudin "Brodin", S.Si, Affandy "Bothol" Abdul Ghoni.
9. Mak Dar, atas segala kemurahan hati dan pengertian serta bantuan logistik selama pengerjaan tugas akhir ini.
10. Konco-konco eks SMULABAYA : Nugroho "Koechink" Setyawan, Pras, Rahmat "Coli" Arif, Malvino "Sinyo" Yudho N., Kholid "Sipit" Firdaus, Rizal "Hostess" Kurniawan, I Gede "Yonek" Adi, Andi Prasetyawan, Dwi Nur Rochim, Reza "Kethek".

11. Arek – arek kos, M. Yahya “Ojek” M.I., Noswa S., Pak Teno, Pak Dwi, Pak Zaenal, Haniif ”kaconk”, Iwan ”Pangah”, Cak Huda, Hadi ”Mbok Lo” Susilo.
12. Arek-arek HIMATIKA ITS, Galih Pradipta (D3 Sastra Inggris UNAIR), para anggota IKASMANCA.
13. Eka Sigi Lasono, atas semua bentuk dukungan mental dan fisik, serta senyum manisnya.....aku sayang kamu.

Semoga segala bantuan yang telah diberikan kepada penulis mendapatkan pahala dari Allah SWT.

Surabaya, Agustus 2007

Penulis.

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	vi
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan.....	3
1.5 Manfaat.....	3
1.6 Sistematika Penulisan.....	3
BAB II DASAR TEORI.....	5
2.1. Data Mining.....	5
2.2. Spatial Data Mining.....	8
2.3 Metode Association Rule.....	9
2.4 Spatial Association Rule Mining.....	12
2.5 Preproses Data.....	13
2.6 Algoritma Apriori.....	14
2.6.1 Apriori Gen Step.....	17
2.6.2. Prune Step.....	18
BAB III DESAIN DAN PERANCANGAN SISTEM.....	19
3.1 Identifikasi dan Analisis Kebutuhan User.....	19
3.2 Analisis Kebutuhan Perangkat Lunak.....	19
3.3 Desain Sistem.....	20
3.3.1 PerancanganArsitektur Perangkat Lunak.....	20
3.3.2 Perancangan Data.....	21

3.3.2.1	Desain Basis Data Global.....	21
3.3.2.2	Desain Basis Data Global.....	23
3.3.2.3	Desain Basis Data Global.....	23
3.3.3	Perancangan Proses.....	30
3.3.3.1	Koneksi Basis Data.....	32
3.3.3.2	Preproses Data.....	32
3.3.3.3	Rule Mining.....	34
3.3.3.4	Algoritma apriori.....	35
3.3.4	Perancangan Antarmuka.....	36
BAB IV	IMPLEMENTASI DAN UJI COBA.....	39
4.1	Pembuatan Perangkat Lunak	39
4.1.1	Modul Modul Utama Perangkat Lunak	39
4.1.1.1	Modul Koneksi Basis Data Spasial	40
4.1.1.2	Modul Koneksi Data Non Spasial	40
4.2	Implementasi Antarmuka	41
4.3	Uji Coba Perangkat Lunak	45
4.3.1	Lingkungan Uji Coba	45
4.4	Hasil Uji Coba	46
4.4.1	Uji Coba 1.....	46
4.4.2	Uji Coba 2.....	49
4.4.3	Uji Coba 3.....	51
BAB V	PENUTUP.....	55
5.1	Kesimpulan	55
5.2	Saran	55
DAFTAR PUSTAKA	56

DAFTAR GAMBAR

Gambar 2.1.	Alur Proses Dari <i>Knowledge Discovery in Database</i>	6
Gambar 2.2.	Interaksi Proses pada <i>Spatial Data Mining</i>	9
Gambar 2.3.	Preproses Data	14
Gambar 2.4.	Diagram <i>Lattice</i> Kombinasi Itemset dari I	15
Gambar 2.5.	Penerapan Prinsip Apriori.....	16
Gambar 3.1.	Arsitektur Perangkat Lunak	21
Gambar 3.2.	Desain Basis Data Global	22
Gambar 3.3	<i>Conceptual Data Model</i> (CDM)	23
Gambar 3.4	<i>Physical Data Model</i> (PDM)	26
Gambar 3.5.	Tabel Tambahan	30
Gambar 3.6	DFD Level 0 (<i>Top Level</i>)	31
Gambar 3.7	DFD Level 1 (<i>Overview Diagram</i>)	31
Gambar 3.8.	Preproses Data	33
Gambar 3.9.	Struktur Menu Perangkat Lunak	37
Gambar 3.10.	Rancangan Desain <i>Interface</i> Perangkat Lunak	38
Gambar 4.1	Form Tampilan Awal	42
Gambar 4.2	Form Hasil Preproses	43
Gambar 4.3	Tampilan Antarmuka Hasil Pencarian <i>frequent itemset</i>	44
Gambar 4.4	Tampilan Antarmuka Rule-Rule yang Dibentuk .	44
Gambar 4.5	Tampilan Antarmuka Hasil Seleksi Rule	45
Gambar 4.6.	Grafik Waktu Proses	53
Gambar 4.7	Grafik Waktu Preproses sampai Pencarian Frequent Itemset	53
Gambar 4.8	Grafik Waktu Pencarian Frequent Itemset	54



“Halaman ini sengaja dikosongkan”

12

DAFTAR TABEL

Tabel 2.1.	Tabel Transaksi	10
Tabel 3.1	Struktur Data Layer Fasilitas	24
Tabel 3.2	Struktur Data Layer Kelurahan	24
Tabel 3.3	Struktur Data Layer Sungai Transformasi	24
Tabel 3.4	Struktur Data Fasilitas	25
Tabel 3.5	Struktur Data Kelurahan	25
Tabel 3.6	Struktur Data Sungai Transformasi	25
Tabel 3.7	Struktur Data Kepadatan	25
Tabel 3.8	Struktur Data Kasus DBD	26
Tabel 3.9	Struktur Data Preproses	27
Tabel 3.10	Struktur Data ITEMSET1	27
Tabel 3.11	Struktur Data ITEMSET2	28
Tabel 3.12	Struktur Data ITEMSET3	28
Tabel 3.13	Struktur Data ITEMSET4	29
Tabel 3.14	Struktur Data KOMBINASI	29
Tabel 4.1	Hasil Semua <i>Frequent Itemset</i>	46
Tabel 4.2	Hasil Rule	49
Tabel 4.3	Hasil Semua <i>Frequent Itemset</i>	49
Tabel 4.4	Hasil Rule	51
Tabel 4.5	Hasil Semua <i>Frequent Itemset</i>	51
Tabel 4.6	Hasil Rule	52
Tabel 4.7	Hasil Uji	52

BAB I PENDAHULUAN

BAB I

PENDAHULUAN

1.1 Latar Belakang

Munculnya sebuah teknologi baru untuk proses *knowledge discovery* pada database spasial atau biasa disebut dengan *data mining* spasial disebabkan oleh makin meluasnya aplikasi teknologi penginderaan jarak jauh yang mengakibatkan data spasial dan nonspasial dalam jumlah yang sangat besar disimpan kedalam suatu database spasial, dimana struktur pada database spasial berbeda dengan struktur pada database biasa [6]. Pada dasarnya teknologi ini adalah *data mining* yang diaplikasikan pada suatu database spasial yaitu pencarian *knowledge*, relasi-relasi spasial, atau pola-pola (*pattern*) tersembunyi pada suatu database spasial yang akan memberikan keterangan – keterangan atau informasi – informasi berguna. Beberapa metode-metode *data mining* pada database transaksional biasa secara umum yaitu asosiasi, klasifikasi, klusterisasi, juga estimasi dan summarisasi secara otomatis bisa diterapkan pada database spasial.

Spatial association rule adalah salah satu metode pada *data mining* spasial yang paling banyak mendapat perhatian saat ini. Fokus utama dari metode ini adalah untuk mendapatkan suatu hipotesa karena hasil yang didapatkan mudah dimengerti, jelas, dan dapat dipercaya [7]. *Association rule mining* pertama kali dikembangkan oleh Agrawal (1993) untuk menganalisa perilaku pembeli di sebuah swalayan dengan tujuan mendapatkan pola pembelian barang guna meningkatkan angka penjualan. Koperski dan Han (1995) mengembangkan *spatial association rule* pertama kali untuk menganalisis data geografis pegunungan Andes di Meksiko berdasarkan penelitian yang dilakukan oleh Agrawal, sehingga didapatkan pola umum dari *spatial association rule* yaitu $X \rightarrow Y$ (*support%*, *confidence%*) dimana X dan Y adalah set predikat dengan syarat paling sedikit mengandung sebuah predikat spasial, sedangkan *support%* dan *confidence%* berturut-

turut adalah presentase *support* dan *confidence* yang berfungsi sebagai *threshold* untuk mendapatkan *rule* yang diinginkan.

Dalam tugas akhir ini diaplikasikan metode *spatial association rule mining* dalam bidang kesehatan dengan fokus kasus penyakit menular Demam Berdarah *Dengue* (DBD) menggunakan algoritma apriori, sebuah algoritma pencarian *rule* yang pertama kali dipakai dan juga merupakan algoritma yang paling umum dipakai karena kemudahannya untuk diterapkan dan dioptimasi untuk mendapatkan hasil yang diinginkan. Jenis penyakit ini diambil sebagai bahan studi karena DBD adalah salah satu penyakit menular tropis yang meskipun penyebarannya tergolong tidak terlalu meningkat dalam kurun waktu beberapa tahun, tetapi penyakit ini sulit sekali untuk dihindari terutama pada musim penghujan. Didalam pandangan pihak dinas kesehatan penyakit ini sangat diperhatikan karena termasuk penyakit endemis yang berbahaya dan sudah banyak menimbulkan korban jiwa.

1.2 Perumusan Masalah

Permasalahan yang diangkat dalam tugas akhir ini adalah :

1. Bagaimana membangun sebuah sistem database spasial yang berisi obyek-obyek spasial beserta atribut-atributnya berdasarkan data yang diperoleh.
2. Bagaimana mendapatkan data yang siap untuk proses *data mining*.
3. Bagaimana menemukan *frequent pattern* di dalam tabel relasi R dengan menggunakan algoritma Apriori.
4. Bagaimana membangun suatu aplikasi untuk mendapatkan asosiasi penyebaran DBD di Surabaya.

1.3 Batasan Masalah

Permasalahan dalam tugas akhir ini dibatasi oleh :

1. Jenis penyakit yang menjadi pokok permasalahan adalah Demam Berdarah *Dengue*.

2. Sumber data spasial diambil dari peta Surabaya, yaitu lokasi fasilitas kesehatan (puskesmas) dalam wilayah Kecamatan dan rawa di Surabaya.
3. Data non spasial, yaitu jumlah kasus DBD perkelurahan (sumber Dinas Kesehatan Surabaya), data kepadatan penduduk (sumber BPS Jawa Timur).
4. Data yang diolah adalah data pada tahun 2002.

1.4 Tujuan

Tujuan dari penelitian ini adalah menemukan asosiasi penyebaran penyakit DBD di daerah Surabaya dengan hal – hal yang bisa mempengaruhinya, yaitu dengan kepadatan penduduk dan rawa di Surabaya melalui metode *spatial association rule* menggunakan algoritma apriori.

1.5 Manfaat

Manfaat dari penelitian ini adalah sebagai bahan pertimbangan selain data yang dimiliki oleh pihak-pihak yang terkait untuk mengambil keputusan seputar kasus penyakit demam berdarah.

1.6 Sistematika Penulisan

Sistematika penulisan Tugas Akhir ini adalah :

BAB I PENDAHULUAN

Bab ini berisi latar belakang dilakukannya penelitian, rumusan masalah, batasan masalah, tujuan dan manfaat, metodologi dan sistematika penulisan Tugas Akhir.

BAB II DASAR TEORI

Bab ini berisi teori penunjang yang digunakan dalam penemuan pola asosiasi spasial pada kasus demam berdarah *dengue* di Surabaya yang mencakup pengertian metode *rule* asosiasi pada data mining, data mining spasial, metode pencarian *rule* asosiasi pada data mining spasial,

pendefinisian istilah-istilah yang akan digunakan, dan pengantar mengenai penyakit demam berdarah *dengue*.

BAB III DESAIN DAN PERANCANGAN SISTEM

Bab ini berisi tentang desain dan perancangan perangkat lunak berdasarkan analisis kebutuhan.

BAB IV IMPLEMENTASI DAN UJI COBA

Berisi tentang implementasi dan uji coba perangkat lunak.

BAB V PENUTUP

Berisi kesimpulan dan saran.

BAB II DASAR TEORI

BAB II DASAR TEORI

Bidang ilmu yang mendasari tugas akhir ini adalah *Data Mining*, *Spatial Data Mining* dan *Spatial Association Rule*.

2.1 Data Mining

Data mining muncul karena adanya pertumbuhan data yang sangat cepat dan teramat banyak sehingga ukuran basis data otomatis ikut membengkak, akibatnya kita akan mengalami kesulitan jika ingin menganalisa data untuk suatu tujuan tertentu, padahal banyak sekali informasi yang bisa kita dapatkan jika bisa mengolah data dengan tepat. Jadi bisa diambil kesimpulan bahwa *data mining* adalah sebuah proses untuk menemukan pengetahuan (knowledge) yang baru, valid dan pola data yang berguna dalam suatu basis data yang besar atau bisa juga dikatakan sebagai ekstraksi informasi atau pola penting dan menarik dari sekumpulan data yang terdapat pada suatu basis data dengan algoritma tertentu [5].

Istilah *data mining* erat kaitannya dengan *Knowledge Discovery in Database* (KDD) karena *data mining* merupakan bagian terpenting pada KDD yaitu suatu proses global untuk menggali informasi dan pola yang berguna pada suatu basis data, singkatnya, KDD adalah proses pengolahan data menjadi pengetahuan (*knowledge*) [4]. Tahapan proses dari KDD cukup sederhana, meliputi :

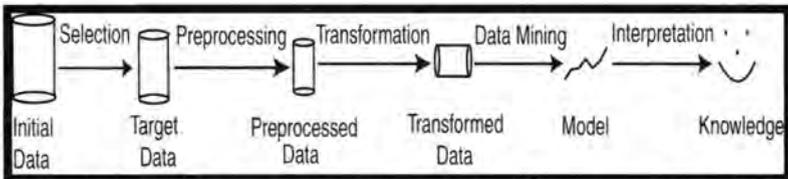
1. *Data preprocess* (termasuk di dalamnya *data cleaning*, *data integration data selection*, dan *data transformation*)
Data *preprocess* mengubah data “mentah” menjadi bentuk format yang sesuai untuk proses mining. Selain itu, pada data *preprocess* dilakukan juga “pembersihan” data untuk data-data yang termasuk noise, yang tidak valid.
2. *Data mining & Pattern Evaluation*
Pada bagian ini baru dilakukan proses mining yang meng-“ekstrak” pengetahuan dari data-data tersebut. Tidak semua

pola-pola yang dihasilkan menarik/berguna. *Pattern evaluation* mengevaluasi hasil mining berdasarkan ukuran tertentu, tergantung proses mining yang dilakukan.

3. *Knowledge Interpretation*

Knowledge Interpretation berusaha merepresentasikan pengetahuan yang diperoleh menjadi bentuk yang dapat dipahami oleh user.

Untuk lebih jelasnya bisa kita lihat bersama pada gambar dibawah ini :



Gambar 2.1 Alur proses dari *Knowledge Discovery in Database*. Perlu kita ketahui juga bahwa baik KDD maupun *Data mining* melibatkan banyak sekali disiplin ilmu yang lain. Misalnya : statistik, kecerdasan buatan (AI), database, teknik visualisasi.

Secara umum tujuan *data mining* dapat dibagi menjadi dua yaitu deskriptif *data mining* dan prediktif *data mining*. Deskriptif maksudnya bahwa *data mining* dilakukan untuk mencari pola-pola yang dapat dipahami manusia yang menjelaskan karakteristik data. Misalnya pencarian kelompok barang yang sering dibeli bersamaan pada data transaksi. Prediktif maksudnya *data mining* dilakukan untuk membentuk sebuah model pengetahuan yang akan digunakan untuk melakukan prediksi. Misalnya untuk memprediksi pengunjung *online* sebuah *website* komersial berdasarkan perilakunya apakah pengunjung ini akan membeli produk *website* tersebut.

Secara lebih spesifik tujuan *data mining* berdasarkan fungsionalitasnya adalah sebagai berikut:

- *Mining Frequent Patterns, Associations, and Correlations* : mencari pola-pola yang sering muncul dalam data. Pengetahuannya biasanya berupa *rule* yang

menunjukkan pola-pola tersebut (biasanya disebut *association rule*). Teknik yang digunakan misalnya Apriori, FP-Growth, CLOSET.

- *Classification and Prediction* : mencari sebuah model yang mampu melakukan prediksi pada suatu data baru yang belum pernah ada. *Decision tree*, *neural network*, *bayesian network*, *support vector machines*, *k-nearest neighbor* adalah contoh alat yang digunakan untuk membentuk model tersebut. *Classification* digunakan untuk prediksi *categorical data* (diskrit), sedangkan untuk *numerical data* (numerik) biasanya menggunakan analisa regresi.
- *Cluster Analysis* : mengelompokkan data dalam sebuah cluster berdasarkan kemiripannya. Prinsipnya adalah memaksimalkan kemiripan dalam sebuah *cluster*, dan meminimalisasikan kemiripan antar *cluster*. Jadi data-data yang berada pada sebuah *cluster* akan memiliki kemiripan yang tinggi, dan sebaliknya data akan memiliki nilai kemiripan yang rendah dengan data yang berada pada *cluster* yang berbeda. Beberapa teknik yang digunakan dalam *cluster analysis* ini misalnya *k-means*, *k-medoids*, SOM, CLARANS, ROCK, BIRCH, Chameleon.
- *Outlier Analysis* : mencari obyek data yang sifatnya anomali (berbeda dengan sifat umum data). Analisa ini berkaitan dengan yang namanya *fraud detection*. Justru data anomali tersebut, yang jumlahnya relatif sedikit ini menarik untuk dianalisa. Misalnya deteksi *fraud credit card*.
- *Evaluation Analysis* : mencari model atau tren untuk data-data yang sifatnya terus berubah. Analisa ini berkaitan dengan data *time-series*. Bisa meliputi *clustering*, *classification*, *association* dan *correlation analysis*.

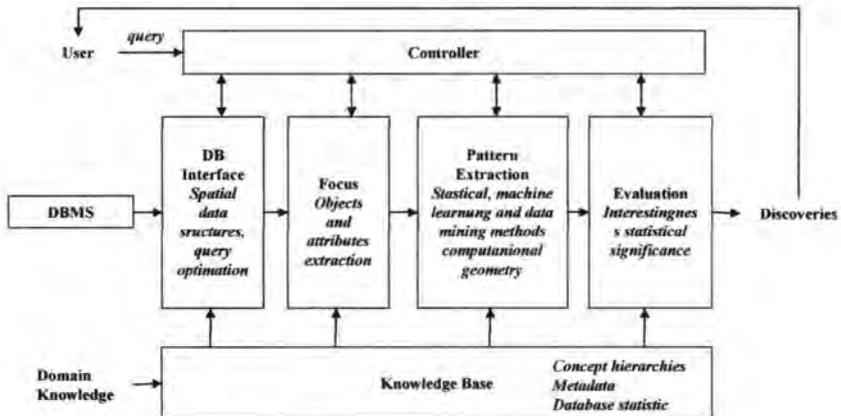


2.2 Spatial Data Mining

Spatial data mining adalah proses *data mining* pada suatu basis data spasial yang menyimpan data – data spasial dalam jumlah yang sangat besar, sedangkan data spasial adalah data yang menentukan tipe dari suatu obyek apakah berupa titik (*points*), garis (*lines*), atau luasan (*polygons*) karena pada data spasial berisi informasi geografis tentang suatu obyek tersebut [2,3,6].

Pada dasarnya konsep dari *spatial data mining* mengikuti konsep *data mining* biasa yaitu *generalization*, *clustering*, dan *mining association rules*, yang menjadikannya berbeda adalah dari basis data yang dipakai yaitu basis data spasial, sehingga muncul metode-metode yang dipakai di dalam *spatial data mining* yaitu : metode *generalization-based* untuk *spatial characteristic and discriminant rules mining*, teknik *aggregate proximity* untuk menemukan karakteristik dari *spatial clusters*, dan yang terakhir teknik *spatial computation* untuk *spatial association rules* [6].

Tahapan proses untuk *spatial data mining* pada umumnya mengadopsi alur proses dari *data mining* konvensional dimana terdapat preproses data, *pattern evaluation*, dan *knowledge interpretation*. Gambar dibawah ini menjelaskan tentang interaksi diantara proses-proses pada *spatial data mining*.



Gambar 2.2. Interaksi proses pada *Spatial Data Mining*.

Pada model ini, *user* bisa mengendalikan keseluruhan proses dalam *data mining*. *Background knowledge* (*meta knowledge*), seperti hirarki konsep spasial dan non-spasial atau informasi tentang basis data disimpan dalam *knowledge base*. Data diambil dari media penyimpanan (DBMS atau *Data Warehouse*) dengan menggunakan *DB Interface*, yang memungkinkan optimasi *query* maupun struktur data yang khusus untuk efisiensi pemrosesan. Komponen *Focus* akan menentukan bagian data mana yang akan digunakan untuk pengenalan pola terkait dengan *task* dalam *spatial data mining* yang akan digunakan. Aturan dan pola dilakukan pada modul *Pattern Extraction*, dalam hal ini menggunakan algoritma apriori untuk melakukan prosedur penemuan aturan dan relasi. Akhirnya, modul *Evaluation* akan memproses pola yang ditemukan untuk mengeliminasi informasi yang tidak digunakan.

2.3 Metode Association Rule

Association rule mining merupakan salah satu teknik di dalam data mining yang digunakan untuk mencari hubungan/relasi, asosiasi, dan korelasi dalam data. Pengetahuan yang dihasilkan juga sangat berguna untuk klasifikasi, *clustering*, dan tujuan data mining yang lainnya[1,5].

Association rule mampu menjelaskan hubungan korelasi antar item dengan lebih jelas, tidak hanya korelasi kuat atau korelasi lemah saja, hal ini dikarenakan di dalam *association rule mining* selalu digunakan parameter yaitu *support* dan *confidence* untuk evaluasi *rule* yang didapat.

Association Rule dinyatakan dalam bentuk $X \Rightarrow Y$ (*support %*, *confidence %*), di mana X dan Y merupakan *Itemset* (himpunan kelompok item), *support* (*s*) dan *confidence* (*c*) merupakan parameter yang digunakan pada *Association Rule*. *Support* menunjukkan persentasi jumlah transaksi yang berisi X dan Y. Sedangkan *confidence* menunjukkan persentasi banyaknya Y pada transaksi yang mengandung X [5]. Bentuk persamaan matematisnya dapat dituliskan seperti ini:

$$\text{Support} (X \rightarrow Y) = P (X \wedge Y)$$

$$\text{Confidence} (X \rightarrow Y) = \text{Support} (X \wedge Y) / \text{Support} (X)$$

Contoh klasik yang sering digunakan untuk menjelaskan *Association Rule Mining* adalah *market basket analysis*. Pada *market basket analysis*, kita menganalisa kebiasaan *customer* dalam membeli barang. Misalkan terdapat data transaksi seperti ini.

Tabel 2.1. Tabel Transaksi.

TID	Items
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

Dari tabel diatas diasumsikan tiap *record* adalah sebuah keranjang yang berisi pembelian oleh seorang *customer* tanpa memperhatikan jumlah barang yang dibeli sehingga dari tabel diatas bisa diketahui barang apa saja yang dibeli bersamaan

maupun barang apa saja yang jarang dibeli bersamaan oleh *customer*, hal ini bagi pemilik toko berguna sebagai pertimbangan dalam menyusun strategi penjualan, kira-kira barang-barang apa saja yang harus diletakkan berdekatan agar lebih besar lagi peluang seorang *customer* membeli produk secara bersamaan atau kira-kira persediaan produk manakah yang seharusnya dikurangi. Hal inilah yang biasanya disebut dengan *knowledge* baru.

Contoh *knowledge* yang dapat diperoleh dari tabel di atas adalah : $\{Beer\} \rightarrow \{Diaper\}$, yang artinya seseorang yang membeli *Beer* biasanya membeli *Diaper* juga. Korelasi dua buah produk ini dikatakan kuat karena dari tabel transaksi diatas apabila seorang *customer* membeli *Beer* maka selalu membeli *Diaper* juga. Korelasi dua buah produk tersebut juga dikatakan "valid" atau dapat dipercaya karena nilai *confidence*-nya adalah 1, sesuai rumus penghitungan *confidence* $A \rightarrow B$ maka *confidence* $Beer \rightarrow Diaper$ adalah $support(Beer \wedge Diaper) / support(Beer)$ yaitu $3 / 3 = 1$. Telah menjadi paradigma umum dikalangan para praktisi data mining bahwa *rule* dengan nilai *confidence* 1 adalah *rule* yang dapat dipercaya.

Secara umum proses *Association Rule Mining* dibagi menjadi dua tahap:

1. Pencarian Frequent Itemset

Pada proses ini dilakukan pencarian Frequent Itemset. Frequent Itemset yang diperoleh harus memenuhi minimum support

2. Candidate generation

Frequent Itemset yang telah dihasilkan dari proses sebelumnya digunakan untuk membentuk association rule. Association rule yang dihasilkan akan memenuhi minimum support dan minimum confidence.

Masalah utama yang muncul pada pencarian *Frequent Itemset* adalah banyaknya hasil *Frequent Itemset* yang memenuhi *threshold minimum support*. Semakin rendah *threshold minimum support*, *Frequent Itemset* yang dihasilkan akan semakin banyak. Jika terdapat d item, maka akan diperoleh $2^d - 1$ kombinasi *itemset*



yang bisa diperoleh. Contoh: 100 item akan menghasilkan kombinasi *itemset* $2^{100}-1$. Jumlah *itemset* ini sangat banyak untuk dapat disimpan atau diproses pada komputer manapun.

2.4 Spatial Association Rule Mining

Spatial association rule adalah sebuah *rule* yang terbentuk pada *spatial data mining*. Yang menjadi perhatian utama dari *spatial association rule* adalah penemuan *rule* yang merepresentasikan relasi-relasi obyek yang memuat predikat-predikat spasial.

Sebagai contoh, dibawah ini adalah *rule- rule* yang dikategorikan sebagai *spatial association rule*:

- *Consequent* non-spasial dengan *antecedent-antecedent* spasial:
 - $is_a(x,house) \wedge close_to(x,beach) \rightarrow is_expensive(x)$. (90%)
mempunyai arti bahwa jika sebuah rumah yang terletak dekat dengan pantai maka terdapat 90 % kemungkinan bahwa rumah tersebut mahal.
- *Consequent* spasial dengan *antecedent-antecedent* non-spasial / spasial:
 - $is_a(x,gas_station) \rightarrow close_to(x,highway)$. (75%)
mempunyai arti bahwa tempat pengisian bahan bakar mempunyai kemungkinan sebesar 75% akan dekat dengan jalan raya.

Penjelasan umum tentang predikat spasial adalah representasi dari relasi-relasi topologi diantara obyek-obyek spasial, seperti, *disjoint*, *intersects*, *inside/outside*, *adjacent_to*, *covers / covered_by*, *equal*, dan lain-lain. Selain itu bisa juga merepresentasikan orientasi atau urutan spasial, seperti, *left*, *right*, *north*, *east*, dan lain-lain, atau memuat informasi tentang jarak, seperti, *close_to*, *far_away*, dan lain-lain [2,6,7].

Spatial association rule adalah sebuah *rule* dengan bentuk umum seperti dibawah ini :

$$P_1 \wedge \dots \wedge P_m \rightarrow Q_1 \wedge \dots \wedge Q_n. (c\%)$$

dimana salah satu dari $P_1, \dots, P_m, Q_1, \dots, Q_n$ adalah sebuah predikat spasial, dan $c\%$ adalah *confidence* dari *rule* tersebut yang menunjukkan bahwa sebanyak $c\%$ dari suatu obyek memenuhi baik *antecedent* juga *consequent* dari *rule* yang terbentuk [6].

Dari uraian di atas bisa disimpulkan bahwa terdapat banyak sekali *rule-rule* yang terbentuk dari sebuah basis data spasial yang dipakai. Hal ini menjadikan suatu paradigma dimana hanya pola-pola yang sering terbentuk dengan parameter *support* yang besar dan *rule-rule* yang mempunyai implikasi kuat dengan parameter *confidence* yang tinggi adalah sebuah *rule* yang layak diterima.

2.5 Preproses Data

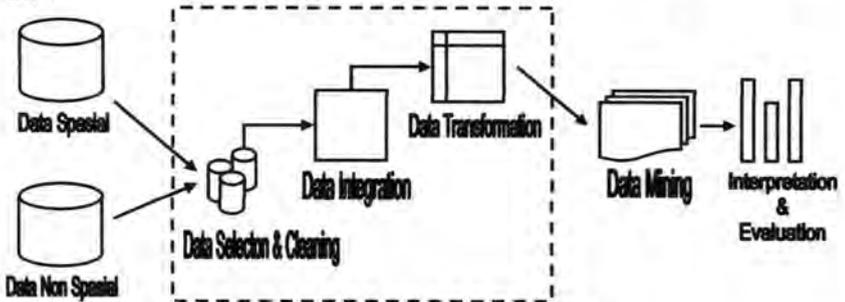
Pada uraian sebelumnya telah dijelaskan bahwa preproses data merupakan tahapan pertama dari KDD yang bertujuan untuk menyesuaikan data agar siap untuk proses mining dimana di dalam tahap ini dilakukan *data cleaning*, *data integration*, *data selection*, dan *data transformation*. Keempat langkah tersebut tidak mutlak harus dilakukan dan juga berurutan, tergantung dari data yang akan kita pakai.

Preproses data pada *spatial association rule* berbeda dengan preproses data pada *association rule*. Perbedaan ini dikarenakan pada *spatial association rule* menggunakan dua jenis data yang berbeda yaitu data spasial dan data non spasial. Data spasial dan data non spasial pada tahap preproses data kemudian akan ditransformasi menjadi sebuah tabel transaksi seperti pada umumnya.

Metode yang banyak digunakan pada preproses data adalah pertama kali mendefinisikan dan mengumpulkan semua informasi yang relevan untuk proses KDD, selain mengurangi jumlah data juga dilakukan pemilihan data, kemudian mendefinisikan relasi-relasi spasial untuk mendapatkan predikat spasial, apakah nantinya berupa relasi topologi, relasi orientasi, ataukah relasi jarak, dan yang terakhir adalah menggabungkan

data spasial dengan data non-spasial sehingga akan didapatkan sebuah tabel yang siap untuk proses [6].

Untuk lebih jelasnya bisa dilihat pada gambar di bawah ini :



Gambar 2.3. Preproses Data

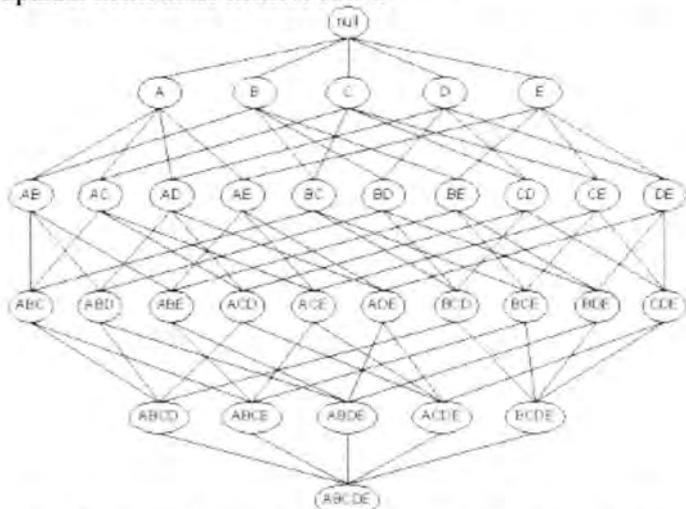
2.6 Algoritma apriori

Masalah utama di dalam pencarian *frequent itemset* adalah banyaknya jumlah kombinasi itemset yang harus diperiksa apakah memenuhi minimum support atau tidak, sehingga proses ini akan membutuhkan waktu yang cukup lama. Salah satu cara untuk mengatasinya adalah dengan cara mengurangi jumlah kandidat itemset yang harus diperiksa [8].

Algoritma apriori adalah salah satu pendekatan yang sering digunakan pada *association rule mining*. Algoritma ini menggunakan prinsip apriori, yaitu jika sebuah *itemset* merupakan *frequent* maka *itemset* yang merupakan *infrequent* tidak perlu ditelusuri supersesnya (*pruned*) sehingga jumlah kandidat yang harus diperiksa menjadi berkurang. Pembentukan *frequent itemset* dilakukan dengan mencari semua kombinasi item-item yang memiliki *support* di atas *minsup* yang telah ditentukan. Proses pada algoritma ini membangkitkan *frequent itemset* per level, dimulai dari level 1-*itemset* sampai n-*itemset*, dimana n menyatakan ukuran itemset yang paling panjang [5].

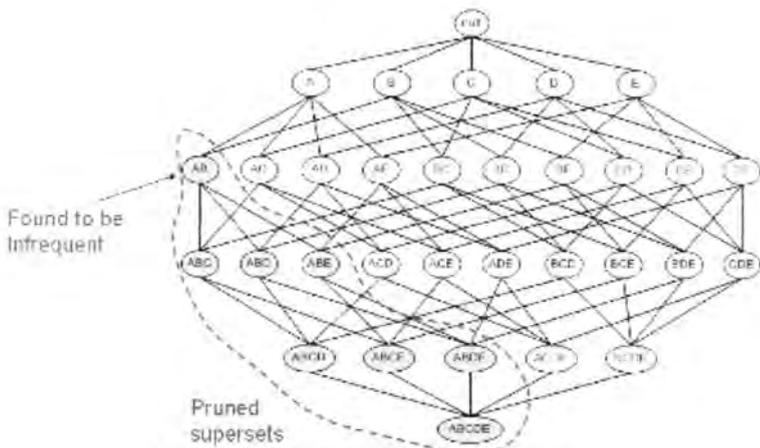
Sebagai ilustrasi, jika diberikan sebuah itemset $I = \{A, B, C, D, E\}$, maka semua kombinasi item-item yang terjadi adalah $\{AB\}$, $\{AC\}$, $\{AD\}$, $\{AE\}$, $\{BC\}$, $\{BD\}$, $\{BE\}$, $\{CD\}$, $\{CE\}$,

{DE}, {ABC}, {ABD}, {ABE}, {ACD}, {ACE}, {ADE}, {BCD}, {BCE}, {BDE}, {CDE}, {ABCD}, {ABCE}, {ABDE}, {ACDE}, {BCDE}, dan {ABCD}, gambar dibawah ini merupakan kombinasi itemset dari I.



Gambar 2.4. Diagram *Lattice* kombinasi itemset dari I.

Jika tanpa menggunakan prinsip apriori maka kita harus memeriksa tiap tiap kombinasi yang terjadi di atas apakah merupakan frequent itemset atau tidak, sehingga pencarian frequent itemset akan memakan waktu yang lama, sedangkan jika kita menggunakan prinsip apriori, apabila pada suatu level kita menemukan sebuah itemset merupakan infrequent maka dijamin bahwa superset dari itemset tersebut juga infrequent. Gambar di bawah ini menunjukkan bagaimana prinsip apriori bekerja pada kombinasi itemset gambar 2.4. :



Gambar 2.5. Penerapan prinsip apriori.

Bisa kita lihat bersama untuk 2-itemset, “AB” ternyata tidak memenuhi support yang diberikan sehingga “AB” merupakan infrequent itemset sehingga semua itemset yang mengandung “AB” akan merupakan infrequent itemset juga.

Contoh kasus di atas menjadi bukti bahwa dengan menggunakan algoritma apriori jumlah kandidat yang harus diperiksa cukup banyak berkurang, hal ini menyebabkan waktu proses pencarian frequent itemset lebih singkat.

Gambaran umum mengenai langkah – langkah di dalam algoritma apriori akan dijelaskan sebagai berikut:

1. Algoritma apriori menerima inputan yaitu data transaksi t dari tabel preproses R , parameter minimum support s , dan parameter minimum confidence c .
2. Langkah pertama yang dilakukan adalah untuk iterasi pertama dihitung jumlah kemunculan setiap item di dalam tabel preproses untuk menentukan *frequent 1-itemset*.
3. Pada iterasi selanjutnya, F_{k-1} yang merupakan *set* dari *frequent (k-1)-itemset* yang dibentuk pada iterasi ke- $(k-1)$, digunakan untuk membentuk C_k (himpunan dari *candidate k-itemset*).
4. Setelah itu, dilakukan penelusuran dalam basis data untuk menghitung *support* bagi setiap *candidate itemset* dalam C_k .

Untuk setiap transaksi t , dicari semua *candidate itemset* c dalam set C_k yang terkandung dalam transaksi tersebut. Set dari semua *candidate itemset* dalam set C_k yang terkandung dalam transaksi t disebut C_t dan ditulis dengan notasi $C_t = \text{subset}(C_k, t)$.

5. Penelusuran dilanjutkan pada transaksi berikutnya sampai semua transaksi dalam basis data ditelusuri, kemudian dicari *candidate k-itemset* yang memenuhi *minsup*. Semua *candidate k-itemset* yang memenuhi *minsup* disimpan dalam F_k yang akan digunakan untuk membentuk *frequent (k+1)-itemset*.

Detail algoritma seperti dibawah ini :

```

F1 = find_frequent_i-itemset(D);
for ( k = 2; Fk-1 ≠ ∅ ; k++ ) do begin
    Ck = apriori-gen(Fk-1, min_sup );
    for each transactions t ∈ D do begin
        Ct = subset(Ck, t);
        for each candidates c ∈ Ct do
            c.count++;
    end
    Fk = {c ∈ Ck | c.count ≥ minsup}
end
return = ∪k Fk;

```

Secara lebih detail, pembentukan *candidate k-itemset* dalam algoritma *Apriori* terdiri atas dua tahap:

2.6.1. Apriori_gen step

Pada langkah ini dilakukan *join* F_{k-1} dengan F_{k-1} untuk memperoleh C_k .

Algoritma:

```

for each itemset l1 ∈ Fk-1
    for each itemset l2 ∈ Fk-1
        if (l1[1]=l2[1]) ∧ (l1[2]=l2[2]) ∧ ... ∧ (l1[k-2]=l2[k-2]) ∧
            (l1[k-1]=l2[k-1]) then
            begin
                c = l1 ∞ l2;
            end

```

```

        if prune(c,  $F_{k-1}$ ) then
            delete c;
        else add c to  $C_k$ ;
    end
    return  $C_k$ ;

```

2.6.2. Prune step

Pada langkah ini dilakukan penelusuran dalam basis data untuk mencari superset *candidate itemset* dalam C_k .

Algoritma:

```

for each (k-1)-subset s of c
    if s  $\notin F_{k-1}$  then
        return TRUE;
    return FALSE;

```

Langkah ini bertujuan untuk menyaring kandidat mana yang mempunyai superset *infrequent*.

BAB III DESAIN DAN PERANCANGAN SISTEM

BAB III

DESAIN DAN PERANCANGAN SISTEM

Dalam bab ini akan dijelaskan mengenai desain dan perancangan dari sistem yang akan dibuat, akan tetapi akan dilakukan identifikasi dan analisis kebutuhan user terlebih dahulu agar pada desain dan implementasi sistem bisa sesuai dengan apa yang diharapkan. Pembahasan mengenai desain sistem meliputi perancangan arsitektur perangkat lunak, perancangan proses, perancangan basis data, dan perancangan antar muka perangkat lunak berdasarkan analisis kebutuhan perangkat lunak.

3.1. Identifikasi dan Analisis Kebutuhan User

Identifikasi dan analisis kebutuhan user dilakukan berdasarkan pada kemampuan perangkat lunak nantinya sehingga kebutuhan user dapat terpenuhi, diantaranya yaitu diharapkan perangkat lunak nantinya bisa mengakses basis data serta mampu melakukan *update* data non spasial yang dipakai yaitu data jumlah kasus DBD dan data kepadatan penduduk di Surabaya, mampu melakukan preproses data, menemukan pola asosiasi diantara kasus DBD, tingkat kepadatan penduduk, jarak kelurahan dengan rawa, dan jarak kelurahan dengan fasilitas kesehatan, selain itu diharapkan juga perangkat lunak nantinya dapat memberikan informasi mengenai data spasial yang tersedia yaitu *layer* kelurahan, *layer* rawa, dan *layer* fasilitas kesehatan sehingga beberapa fasilitas *tools* seperti *zoom* dan *identify* obyek-obyek spasial bisa disediakan untuk mendukung pemenuhan kebutuhan user.

3.2. Analisis Kebutuhan Perangkat Lunak

Analisis kebutuhan perangkat lunak dilakukan berdasarkan tujuan yang telah dirumuskan sebelumnya, sehingga didapatkan spesifikasi perangkat lunak sebagai berikut:

1. Perangkat lunak dapat menampilkan peta dan melakukan operasi standar pada peta, seperti *zooming*, *pan*, dan identifikasi objek.
2. Perangkat lunak dapat membaca dan menampilkan tabel dari basis data SQL Server.
3. Perangkat lunak dapat melakukan *update* data non-spasial pada basis data.
4. Perangkat lunak mempunyai fitur untuk melakukan preproses data.
5. Perangkat lunak dapat menghasilkan informasi mengenai pola asosiasi yang didapatkan menggunakan algoritma apriori berdasarkan parameter yang diberikan, yaitu *support* dan *confidence*.

Pada sub bab selanjutnya akan dijelaskan mengenai langkah – langkah perancangan sistem untuk memenuhi kebutuhan perangkat lunak yang telah disebutkan di atas.

3.3. Desain Sistem

Setelah spesifikasi perangkat lunak ditetapkan berdasarkan hasil analisis kebutuhan perangkat lunak maka selanjutnya dilakukan perancangan atau desain yang meliputi perancangan arsitektur perangkat lunak, perancangan data, perancangan proses dan perancangan antarmuka perangkat lunak.

3.3.1. Perancangan Arsitektur Perangkat Lunak

Perancangan arsitektur merupakan model fisik yang menjelaskan tentang gambaran keseluruhan proses perangkat lunak yang akan dibuat nantinya. Perangkat lunak akan melakukan koneksi dengan basis data yang dipakai terlebih dahulu, yaitu koneksi dengan data spasial dan data non spasial. Setelah koneksi berjalan maka perangkat lunak akan melakukan preproses data untuk menggabungkan data spasial dengan data spasial agar bisa didapatkan suatu tabel preproses yang akan digunakan pada keseluruhan proses *rule mining*. Selanjutnya adalah proses *rule mining* yang terdiri dari dua bagian utama yaitu

pencarian *frequent itemset* dan *rule generation*, sehingga akan diperoleh *rule-rule* yang diinginkan. Interpretasi *rule* adalah proses terakhir yang akan dilakukan oleh perangkat lunak, yaitu seleksi *rule* yang telah terbentuk untuk mendapatkan pola asosiasi DBD.

Dibawah ini adalah gambar yang menjelaskan tentang arsitektur perangkat lunak. Untuk Basis data digunakan Microsoft SQL Server 2000 yang menyimpan tabel-tabel data spasial dan data non spasial, sedangkan untuk bahasa pemrograman dan antarmuka digunakan Microsoft Visual Basic 6.0.



Gambar 3.1. Arsitektur Perangkat Lunak.

3.3.2. Perancangan Data

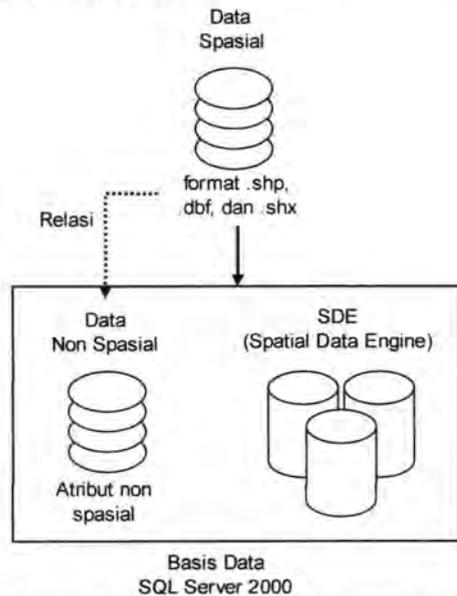
Perancangan data bertujuan untuk memilih dan mengatur kebutuhan data dari perangkat lunak yang akan dibangun. Pada bagian ini ditentukan data yang akan terlibat langsung dalam sistem. Perancangan data menjelaskan desain basis data global, desain basis data secara konsep, dan desain basis data secara fisik.

3.3.2.1. Desain Basis Data Global

Basis data (*database*) merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya dan digunakan perangkat lunak tertentu untuk memanipulasinya. Perancangan basis data adalah proses membangun basis data untuk *association rule mining*. Basis data yang digunakan untuk proses *association rule mining* ini terdiri dari dua macam, yaitu basis data yang berisi informasi spasial dalam bentuk *file* dengan format *.shp*, *.dbf*, dan *.shx* dan basis data yang memuat informasi atribut non-spasial yang disimpan di Microsoft SQL Server. Kedua basis data

tersebut diorganisasikan menjadi basis data terintegrasi yang disimpan di Microsoft SQL Server.

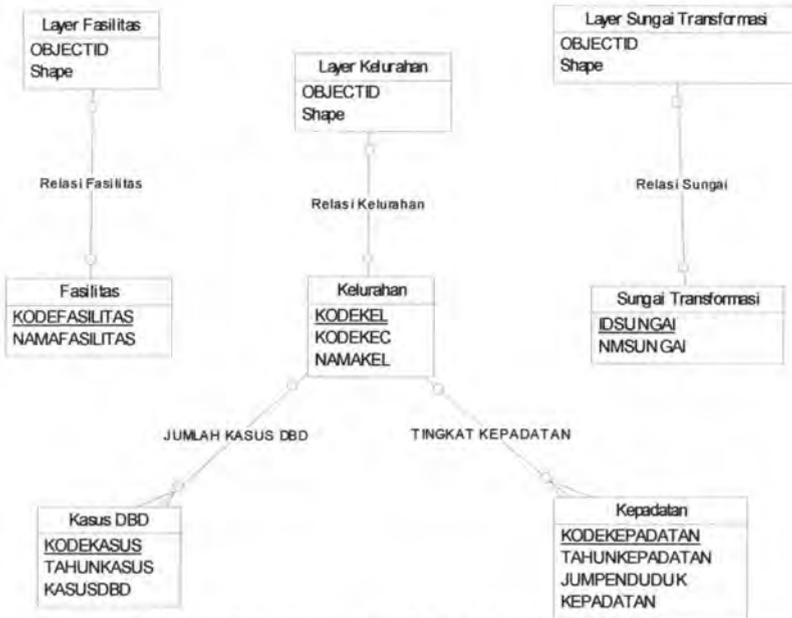
Untuk mengintegrasikan basis data spasial dan non-spasial dibangun suatu basis data yang disebut *Spatial Database Engine* (SDE). SDE dibuat di Microsoft SQL Server dan harus memenuhi dua kriteria. Pertama, basis data yang dibuat harus dengan nama *sde* dan ukuran minimal sebesar 140 Mb. Kedua, basis data yang telah dibuat harus memiliki hak untuk membuat tabel, membuat *view*, membuat *procedure*, dan membuat *function*. Setelah SDE berhasil dibuat, langkah selanjutnya adalah menempatkan basis data spasial yang berbentuk file dengan format *.shp*, *.shx*, dan *.dbf* ke dalam sistem basis data Microsoft SQL Server dan direlasikan dengan basis data non-spasial yang berisikan atribut dari data spasial.



Gambar 3.2. Desain Basis Data Global.

3.3.2.2. Desain Basis Data Konseptual

Perancangan basis data secara konseptual (CDM) merupakan upaya untuk membuat model yang masih bersifat konsep.



Gambar 3.3 Conceptual Data Model (CDM).

3.3.2.3. Desain Basis Data Fisik

Dari CDM di atas, konsep tersebut kemudian dituangkan ke dalam perancangan basis data fisik yang tersimpan pada media penyimpan eksternal (yang spesifik terhadap DBMS yang dipakai).

Rancangan basis data yang telah terintegrasi memuat atribut spasial dan non-spasial terdiri dari delapan buah tabel, antara lain:

1. Tabel Layer Fasilitas, memuat informasi spasial tentang fasilitas kesehatan.

2. Tabel Layer Kelurahan, memuat informasi spasial tentang kelurahan.
3. Tabel Layer Sungai Transformasi, memuat informasi spasial tentang Rawa.
4. Tabel Fasilitas, memuat informasi tentang fasilitas kesehatan.
5. Tabel Kelurahan, memuat informasi tentang kelurahan.
6. Tabel Sungai Transformasi, memuat informasi tentang rawa.
7. Tabel Kepadatan, memuat informasi tentang kepadatan penduduk.
8. Tabel Kasus DBD, memuat informasi tentang kasus DBD.

Masing-masing struktur data dari tabel diatas ditampilkan dalam tabel-tabel berikut:

Tabel 3.1 Struktur Data Layer Fasilitas

Nama Field	Tipe	Panjang	Keterangan
OBJECTID	Int	4	Kode shape fasilitas
Shape	Int	4	Shape fasilitas

Tabel 3.2 Struktur Data Layer Kelurahan

Nama Field	Tipe	Panjang	Keterangan
OBJECTID	Int	4	Kode shape kelurahan
Shape	Int	4	Shape kelurahan

Tabel 3.3 Struktur Data Layer Sungai Transformasi.

Nama Field	Tipe	Panjang	Keterangan
OBJECTID	Int	4	Kode shape sungai transformasi
Shape	Int	4	Shape sungai transformasi

Tabel 3.4 Struktur Data Fasilitas

Nama Field	Tipe	Panjang	Keterangan
KODEFASILITAS	Char	10	Kode Fasilitas
NAMAFASILITAS	Int	4	Nama Fasilitas

Tabel 3.5 Struktur Data Kelurahan

Nama Field	Tipe	Panjang	Keterangan
KODEKEC	Char	10	Kode kecamatan
KODEKEL	Char	10	Kode kelurahan
NAMAKEL	VarChar	30	Nama kelurahan

Tabel 3.6 Struktur Sungai Transformasi

Nama Field	Tipe	Panjang	Keterangan
IDSUNGAI	Char	10	Kode Sungai
NMSUNGAI	Char	10	Nama Sungai

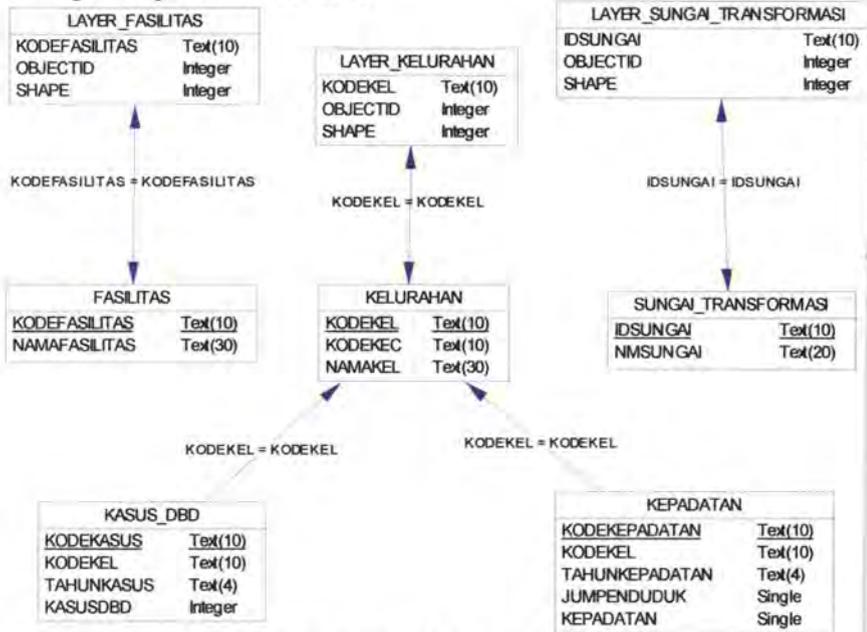
Tabel 3.7 Struktur Data Kepadatan

Nama Field	Tipe	Panjang	Keterangan
KODEKEPADATAN	Char	10	Kode Kepadatan
TAHUNKEPADATAN	Char	4	Tahun Kepadatan
JUMPENDUDUK	Float	8	Jumlah Penduduk
KEPADATAN	Float	8	Jumlah Kepadatan

Tabel 3.8 Struktur Kasus DBD

Nama Field	Tipe	Panjang	Keterangan
KODEKASUS	Char	10	Kode Kasus DBD
TAHUNKASUS	Char	4	Tahun Kasus DBD
KASUSDBD	Numeric	9	Jumlah Kasus DBD

Perancangan basis data secara fisik (PDM) akan ditampilkan pada Gambar 3.4.



Gambar 3.4 Physical Data Model (PDM)

Selain itu akan ada penambahan beberapa tabel seperti dibawah ini. Tabel tambahan berikut bersifat temporer, karena hanya digunakan selama proses *rule mining*, yang berfungsi untuk

menyimpan hasil preproses, *frequent itemset*, dan *rule* yang terbentuk.

Rancangan tabel tambahan yang nantinya akan dibutuhkan terdiri dari enam tabel, antara lain:

1. Tabel Preproses, penyimpanan hasil preproses.
2. Tabel ITEMSET1, merupakan tabel penyimpanan 1-*frequent itemset* sementara.
3. Tabel ITEMSET2, merupakan tabel penyimpanan 2-*frequent itemset* sementara.
4. Tabel ITEMSET3, merupakan tabel penyimpanan 3-*frequent itemset* sementara.
5. Tabel ITEMSET4, merupakan tabel penyimpanan 4-*frequent itemset* sementara.
6. Tabel KOMBINASI, merupakan tabel penyimpanan *rule-rule* yang terbentuk.

Masing-masing struktur data dari tabel diatas ditampilkan dalam tabel-tabel berikut:

Tabel 3.9 Struktur Data Preproses

Nama Field	Tipe	Panjang	Keterangan
TID	Char	10	Kode Transaksi
FASKES	VarChar	20	Status Fasilitas Kesehatan
DBD	VarChar	20	Jumlah kasus DBD
KEPADATAN	VarChar	20	Jumlah kepadatan penduduk
RAWA	VarChar	20	Status rawa

Tabel 3.10 Struktur Data ITEMSET1

Nama Field	Tipe	Panjang	Keterangan
-------------------	-------------	----------------	-------------------

TID	Char	5	Kode Transaksi
ITEM1	VarChar	50	Nama item
JUMLAH	Integer	4	Jumlah Item
SUPPORT	Numeric	8	Nilai support

Tabel 3.11 Struktur Data ITEMSET2

Nama Field	Tipe	Panjang	Keterangan
TID	Char	5	Kode Transaksi
ITEM1	VarChar	50	Nama item
ITEM2	VarChar	50	Nama item
JUMLAH	Integer	4	Jumlah Item
SUPPORT	Numeric	8	Nilai support

Tabel 3.12 Struktur Data ITEMSET3

Nama Field	Tipe	Panjang	Keterangan
TID	Char	5	Kode Transaksi
ITEM1	VarChar	50	Nama item
ITEM2	VarChar	50	Nama item
ITEM3	VarChar	50	Nama item
JUMLAH	Integer	4	Jumlah Item
SUPPORT	Numeric	8	Nilai support

Tabel 3.13 Struktur Data ITEMSET4

Nama Field	Tipe	Panjang	Keterangan
TID	Char	5	Kode Transaksi
ITEM1	VarChar	50	Nama item
ITEM2	VarChar	50	Nama item
ITEM3	VarChar	50	Nama item
ITEM4	VarChar	50	Nama item
JUMLAH	Integer	4	Jumlah Item
SUPPORT	Numeric	8	Nilai support

Tabel 3.14 Struktur Data KOMBINASI

Nama Field	Tipe	Panjang	Keterangan
Jika	VarChar	50	Nama itemset
Maka	VarChar	50	Nama itemset
STotal	Float	8	Nilai support total
SJika	Float	8	Nilai support jika
Conf	Float	8	Nilai confidence



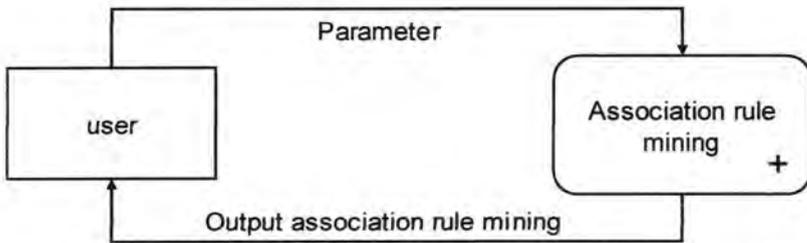
Dibawah ini merupakan desain fisik untuk tabel tambahan.



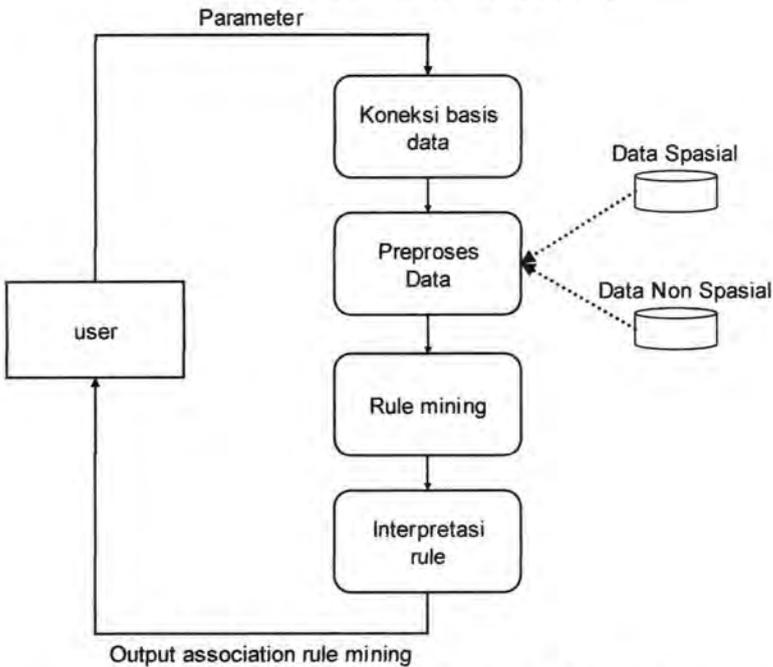
Gambar 3.5. Tabel tambahan.

3.3.3. Perancangan Proses

Perancangan proses menjelaskan alur proses yang dilakukan oleh sistem berdasarkan pada desain arsitektur perangkat lunak agar bisa mendapatkan keluaran sesuai dengan algoritma yang digunakan, yaitu algoritma apriori. Pada perancangan proses ini juga dibuat suatu alur pada sistem atau yang dikenal dengan diagram arus data (*data flow diagram* atau DFD). DFD merupakan alat yang biasa dipakai untuk mendokumentasikan proses dalam sistem. DFD konteks adalah DFD yang memperlihatkan sistem sebagai sebuah proses (DFD konteks biasa disebut *top level*), kemudian dari *top level* ini akan diperlihatkan bagian yang lebih rinci (*lower level*). *Top level* pada sistem *spasial data mining* ini akan diperlihatkan pada Gambar 3.6, kemudian penggambaran lebih rinci (*lower level*) akan ditunjukkan pada Gambar 3.7.



Gambar 3.6 DFD level 0 (*Top Level*)



Gambar 3.7 DFD level 1 (*Overview Diagram*)

Proses *association rule mining* terbagi menjadi 4 buah proses, yaitu koneksi basis data, preproses data, *rule mining*, dan menampilkan hasil *rule*.

3.3.3.1. Koneksi basis data

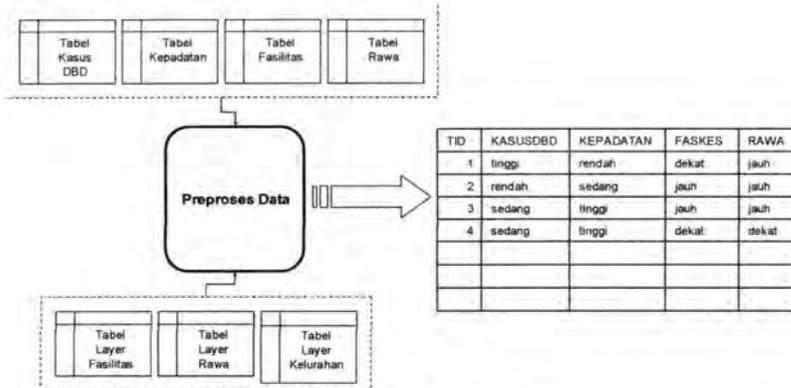
Proses koneksi basis data akan menjadi proses pertama di dalam association rule mining yang menghubungkan perangkat lunak dengan Microsoft SQL Server untuk akses basis data dan menampilkan objek peta. Proses ini akan membuat koneksi basis data dan melakukan pembentukan objek secara otomatis ke basis data berdasarkan input *User*, *Password*, dan *Server*.

3.3.3.2. Preproses Data

Preproses data bertujuan untuk menyesuaikan data agar siap untuk proses mining. Hasil dari preproses data ini adalah sebuah data transaksi atau biasanya juga disebut tabel relasi R yang menggabungkan data spasial dan non spasial. Berdasarkan perancangan data yang telah dibuat, data yang dipreproses adalah data dari tabel kelurahan, tabel kasus DBD, tabel kepadatan, tabel fasilitas kesehatan, dan tabel rawa.

Acuan dari preproses data ini adalah tabel kelurahan, sedangkan keempat tabel lainnya akan direlasikan sesuai dengan *key* yang telah ditentukan yaitu pada atribut kode kelurahan, selain itu juga dilakukan kategorisasi pada keempat tabel tersebut agar hasil *mining* nantinya lebih terperinci. Kategorisasi ini memiliki prosedur yang berbeda, untuk jumlah kasus DBD dan jumlah kepadatan yang nantinya mengisi *field* "jumlah kasus DBD" dan *field* "kepadatan" pada tabel preproses akan dikategorikan menurut jumlah, sedangkan fasilitas kesehatan yang nantinya mengisi *field* "fasilitas kesehatan" pada tabel preproses akan dikategorikan berdasarkan jauh tidaknya fasilitas kesehatan tersebut, begitu juga dengan rawa yang nantinya akan mengisi *field* "rawa" pada tabel preproses akan dikategorikan menurut jarak.

Gambar 3.8 dibawah ini menjelaskan tentang bagaimana preproses data dilakukan.



Gambar 3.8. Preproses data.

Prosedur untuk melakukan preproses data disajikan sebagai berikut

Input :

Tabel : KASUSDBD, KEPADATAN, FASKES, RAWA,
LayerKelurahan, LayerFasilitas,
LayerSungaitransformasi
Parameter : Nilai kategorisasi DBD, Kepadatan,
radius fasilitas kesehatan, dan radius rawa

Output :

Tabel Preproses R (KASUSDBD, KEPADATAN, FASKES,
RAWA)

Begin

Step 1 : Kategorisasi nilai jumlah kasus DBD dan jumlah kepadatan

- untuk setiap nilai jumlah kasus DBD dan jumlah kepadatan dikategorisasi sehingga menghasilkan kategori tinggi, sedang, atau rendah
- update nilai jumlah kasus DBD dan nilai kepadatan

Step 2 : Deteksi keberadaan fasilitas kesehatan terhadap kelurahan

- untuk setiap record LayerKelurahan cari kesamaan dengan setiap record pada LayerFasilitas berdasarkan field key shape kemudian dihitung jaraknya
- jika hasil hitung \leq parameter yang diberikan maka keberadaan fasilitas kesehatan bernilai 'dekat'
- jika hasil hitung $>$ parameter yang diberikan maka keberadaan fasilitas kesehatan bernilai

'jauh'

Step 3 : Deteksi kedekatan rawa dengan kelurahan
 - untuk setiap record LayerKelurahan cari kesamaan dengan setiap record pada LayerRawa berdasarkan field key shape kemudian dihitung jaraknya
 - jika hasil hitung \leq parameter yang diberikan maka status rawa bernilai 'dekat'
 - jika hasil hitung $>$ parameter yang diberikan maka status rawa bernilai 'jauh'

Step 4 : isi tabel preproses

End

Hasil dari preproses data adalah sebuah tabel transaksi R yang memiliki 6 buah *field* yaitu : TID, KELURAHAN, KASUSDBD, KEPADATAN, FASKES, RAWA. Tabel transaksi inilah yang nantinya akan menjadi sumber dari keseluruhan proses *rule mining*.

3.3.3.3. Rule Mining

Pencarian *frequent itemset* menggunakan algoritma apriori dilakukan pada sebuah tabel R yang berisi data hasil olahan preproses dengan parameter minimum *support* yang diinputkan oleh user. Hasil dari pencarian *frequent itemset* ini disimpan lagi kedalam tabel-tabel untuk memudahkan langkah selanjutnya yaitu membangun *rule*. Penjelasan lebih lengkap tentang algoritma akan dibahas pada sub bab algoritma apriori.

Proses membangun *rule* bisa dilakukan setelah didapatkan suatu *frequent itemset*. *Rule* yang terbentuk merupakan kombinasi diantara *item - item* pada *frequent itemset* dengan parameter *minimum confidence* yang diinputkan user. Jumlah *rule* yang terbentuk tergantung dari *frequent itemset* yang didapat, semakin banyak *item - item* yang menjadi anggota sebuah *frequent itemset* maka *rule* yang terbentuk juga akan semakin banyak.

Prosedur untuk membangun *rule* adalah seperti dibawah ini.

Input :

Frequent itemset $X = \{r\}$
 Minimum confidence c

Output :

Rule hasil kombinasi r

Begin

- Cari semua subset tak kosong dari X
- Pasangkan tiap tiap r subset tak kosong dari X
- r adalah rule jika $\geq c$

End

3.3.3.4. Algoritma Apriori

Algoritma apriori mengolah data preproses dengan parameter minimum *support* dan minimum *confidence* yang diinputkan oleh user. Langkah pertama adalah *scan* basis data untuk mendapatkan nilai *support* tiap-tiap *itemset*, kemudian inialisasi nilai k dengan 1 yang berarti *itemset* memuat 1 buah *item* saja, selanjutnya adalah membentuk C_k sehingga *frequent itemset* L_k bisa didapatkan melalui perbandingan nilai *support* tiap-tiap *itemset* anggota C_k terhadap nilai *minimum support* yang diberikan, jika *support* yang dimiliki oleh suatu *itemset* lebih besar atau sama dengan *minimum support* yang diberikan, maka tambahkan *itemset* tersebut ke dalam L_k . Selanjutnya adalah langkah *candidate generation* dimana dilakukan operasi *join* L_k dengan L_k itu sendiri untuk mendapatkan C_{k+1} , disini digunakan prinsip *prunning* dari apriori yaitu jika suatu *itemset* tidak termasuk sebagai *frequent itemset* maka supersetnya tidak akan ikut di dalam proses *candidate generation* selanjutnya.

Permasalahan kali ini yang melibatkan empat buah *field* utama sebagai acuan *rule* asosiasi yaitu kasus DBD, kepadatan per kelurahan, relasi spasial rawa, dan relasi spasial fasilitas kesehatan, menyebabkan jumlah *item* terbanyak dari sebuah *frequent itemset* adalah empat buah saja, sehingga dibutuhkan empat buah tabel untuk menyimpan masing-masing *frequent itemset* yang didapatkan, yaitu tabel 1-*frequent itemset*, 2-*frequent itemset*, 3-*frequent itemset*, dan 4-*frequent itemset*.

Prosedur apriori adalah sebagai berikut.

Input :

Tabel Preproses R, minimum support s , minimum confidence c

Output :

candidate itemset $C=\{k\}$, frequent itemset $X=\{r\}$

Begin

Step 1 : scan R

- untuk setiap t anggota R dihitung jumlah kemunculannya

Step 2 : Bentuk C_i dari setiap t

- inisialisasi $i = 1$

Step 3 : Temukan frequent itemset X_i dari $C(i)$ dengan parameter s

- untuk setiap k dengan support $\geq s$ maka menjadi anggota X

Step 4 : Bentuk $C(i+1)$ dari $X(i)$

- untuk setiap r anggota $X(i)$ dipasangkan, $r_1 \diamond r_2$

Step 5 : loop Step 3 dan Step 4 sampai $C(i+1)$ tidak dapat dibentuk lagi

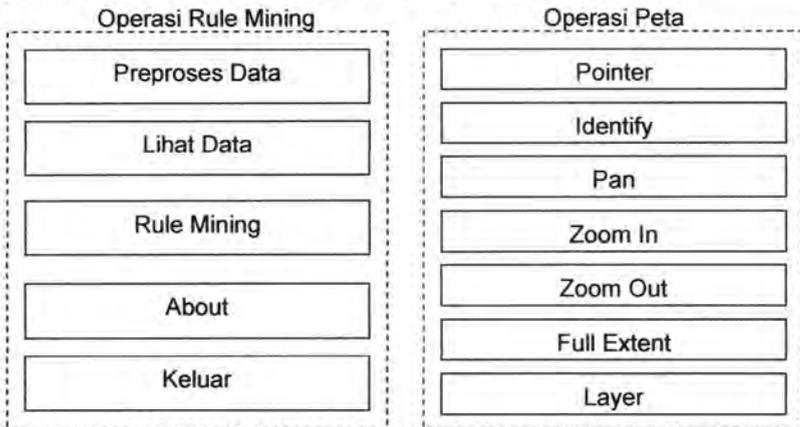
End

3.3.4. Perancangan Antarmuka

Pada sub bab ini dijelaskan struktur rancangan antarmuka dari sistem. Fasilitas-fasilitas minimal yang harus ada pada perangkat lunak ini antara lain, proses operasi dari tabel yang dipakai seperti view dan edit tabel, proses standart pada peta, seperti *Zooming*, *Pan* (menggeser peta), dan identifikasi pada peta, dan fasilitas yang terakhir adalah proses *rule mining*. Adapun rancangan hierarki menu pada perangkat lunak ditampilkan pada Gambar 3.9.

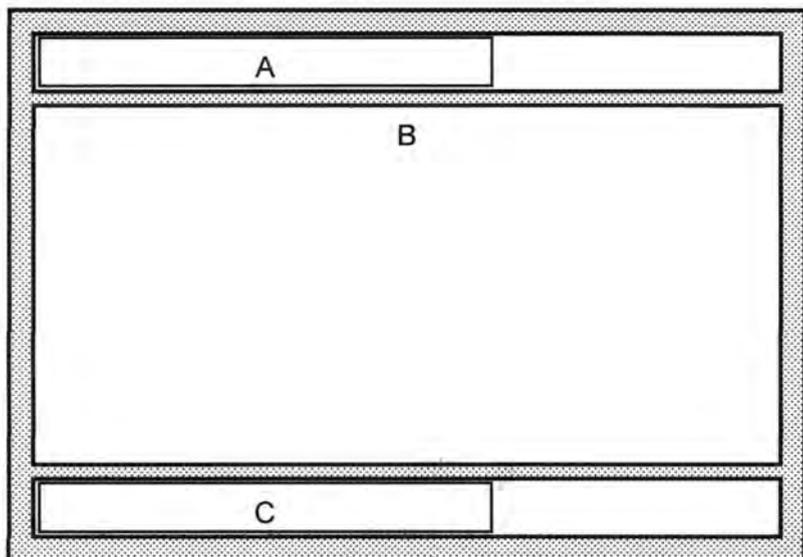
Perangkat lunak terdiri atas dua bagian utama, yaitu operasi *rule mining* dan operasi peta. Pada operasi *rule mining* menu yang disediakan adalah menu Preproses Data, menu lihat data, dan Proses *Rule Mining*. Sedangkan pada bagian operasi peta

disediakan menu untuk operasi peta seperti : *zoom out*, *zoom in*, *pan*, *full extent*, *identify*, *pointer*, dan pemilihan *layer*.



Gambar 3.9. Struktur Menu Perangkat Lunak

Rancangan antarmuka utama perangkat lunak ditampilkan seperti Gambar 3.10. Antarmuka utama terdiri dari tiga bagian. Pertama, bagian *Menu bar* yang terbagi atas *preproses data*, *lihat data*, *rule mining*, *About* dan menu *Keluar*. Kedua, bagian peta yang menampilkan layout kelurahan, fasilitas kesehatan, dan rawa di Surabaya. Ketiga adalah bagian operasi standar pada peta seperti *zooming*, *pan*, *full extent*, dan *identify*. Proses *zooming* dibagi menjadi dua, yaitu *zoom in* dan *zoom out* yang masing-masing digunakan untuk memperbesar dan memperkecil tampilan peta. *Pan* digunakan untuk menggeser posisi peta, sedangkan *full extent* digunakan untuk mengatur tampilan peta pada posisi awal secara otomatis. Proses *identify* adalah melakukan identifikasi salah satu objek dalam peta yang dipilih oleh *user* berdasarkan *layer* yang dipilih. Hasil dari proses *identify* berupa informasi tentang objek yang terpilih oleh *user* tersebut.



Gambar 3.10. Rancangan desain interface perangkat lunak.

Keterangan :

A. *Menu Bar*.

B. *Peta*.

C. *Tool* operasi pada peta.

BAB IV IMPLEMENTASI DAN UJI COBA

BAB IV

IMPLEMENTASI DAN UJI COBA

Dalam bab ini dijelaskan tentang pembuatan dan uji coba perangkat lunak. Dari uji coba yang dilakukan akan didapatkan hasil yang kemudian akan dilakukan analisis terhadap hasil uji coba tersebut.

4.1. Pembuatan Perangkat Lunak

Pada sub bab ini diuraikan tentang pembuatan perangkat lunak yang meliputi uraian tentang modul-modul utama perangkat lunak dan implementasi antarmuka. Dalam pembuatan perangkat lunak ini digunakan bahasa pemrograman visual berbasis Windows, yaitu Microsoft Visual Basic 6 dan objek pemetaan menggunakan MapObjects 2.2. Sedangkan sistem basis data yang digunakan untuk implementasi ini adalah Microsoft SQL Server.

Perangkat lunak diimplementasikan dengan menggunakan komputer dengan sistem operasi Microsoft Windows XP Professional SP 2 dengan spesifikasi perangkat keras prosesor Pentium 4 3.06 GHz dengan memori utama sebesar 512MB. Spesifikasi tersebut digunakan sebagai standar aplikasi dan pengujian perangkat lunak.

4.1.1. Modul-Modul Utama Perangkat Lunak

Berikut akan diuraikan modul-modul utama yang digunakan dalam implementasi perangkat lunak. Perangkat lunak yang dibangun akan mengintegrasikan komponen basis data yang telah disimpan dalam suatu sistem basis data, yaitu Microsoft SQL Server, yang kemudian ditampilkan ke *user* melalui Microsoft Visual Basic 6.0. Oleh karena itu, langkah pertama yang harus dilakukan adalah membangun hubungan antara sistem perangkat lunak dengan sistem basis data yang digunakan. Modul membangun koneksi terbagi menjadi dua, yaitu koneksi terhadap basis data sde yang merupakan perantara untuk mengakses data

spasial dan koneksi terhadap basis data non-spasial untuk mengakses data-data penunjang.

4.1.1.1. Modul koneksi basis data spasial

Langkah-langkah yang dilakukan oleh modul ini adalah membangun suatu variabel objek koneksi spasial yang bertipe `MapObjects2.DataConnection` sesuai dengan parameter yang telah diatur di dalam perangkat lunak, yaitu `NamaServer`, jenis basis data yang digunakan, dan nama komputer tempat basis data disimpan. Objek akan melakukan evaluasi terhadap parameter-parameter yang dimasukkan untuk membangun koneksi, jika evaluasi berhasil maka modul akan mengembalikan nilai 1 dan jika gagal akan mengembalikan nilai selain 1 sesuai dengan kesalahan yang terjadi. Kode untuk membangun koneksi dengan basis data spasial adalah sebagai berikut :

```
Public Function MulaiKoneksiSpasial(sServer As String,
sInstance As String, sDatabase As String, sUser As String,
sPassword As String) As Integer
    Set mKoneksiSpasial = New MapObjects2.DataConnection
    mKoneksiSpasial.Server = "SDEB1:" & sServer
    mKoneksiSpasial.Database = "instance=" & sInstance
    mKoneksiSpasial.User = sUser
    mKoneksiSpasial.Password = sPassword
    If sDatabase <> "" Then
        mKoneksiSpasial.Database = mKoneksiSpasial.Database
& _
        ";database=" &
sDatabase
    End If
    If mKoneksiSpasial.Connect Then
        MulaiKoneksiSpasial = 1
    Else
        MulaiKoneksiSpasial =
mKoneksiSpasial.ConnectError
    End If
End Function
```

4.1.1.2. Modul koneksi data non spasial

Langkah-langkah yang dilakukan oleh modul koneksi basis data non-spasial ini adalah membangun suatu variabel objek koneksi non-spasial yang bertipe `ADODB.Connection`. Dengan

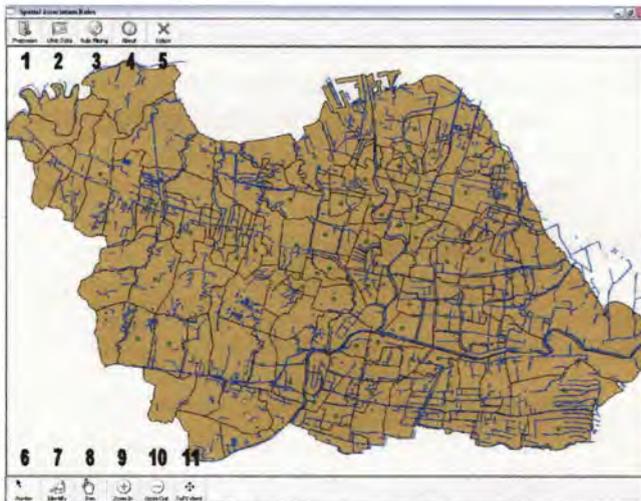
parameter yang diberikan objek tersebut membangun koneksi ke basis data, jika koneksi berhasil dibangun maka modul akan mengembalikan nilai True dan jika gagal akan mengembalikan nilai False. Berikut code untuk membangun koneksi untuk basis data non-spasial :

```
Public Function MulaiKoneksiNonSpasial(sServer As String,
sDatabase As String, _
sUser As String,
sPassword As String) As Boolean

    On Error GoTo Gagal
    Set mKoneksiNonSpasial = New ADODB.Connection
    mKoneksiNonSpasial.ConnectionString =
    "Provider=SQLOLEDB.1;Password=" & _
    sPassword & ";Persist Security Info=True;User
    ID=" & _
    sUser & ";Initial Catalog=" & sDatabase & ";Data
    Source=" & sServer
    mKoneksiNonSpasial.Open
    MulaiKoneksiNonSpasial = True
    Exit Function
Gagal:
    MulaiKoneksiNonSpasial = False
End Function
```

4.2. Implementasi Antarmuka

Implementasi antarmuka dibuat berdasarkan desain interface yang telah dibuat. Berikut tampilan *form* awal yang sudah membangun koneksi dengan basis data Microsoft SQL Server.



Gambar 4.1 Form tampilan awal

Keterangan :

1. Tombol Preproses Data
2. Tombol Lihat Data
3. Tombol Rule Mining
4. Tombol Deskripsi
5. Tombol Keluar
6. Tombol Pointer
7. Tombol Identify
8. Tombol Pan
9. Tombol Zoom In
10. Tombol Zoom Out
11. Tombol Full Extent

Setelah koneksi tersambung, *user* bisa memulai proses *association rule mining*. Sebelum melakukan proses, terlebih dahulu *user* harus melakukan preproses data untuk membuat tabel relasi *R* melalui menu 'preproses' yang telah disediakan pada perangkat lunak, pada form Preproses data ini *user* juga harus mengisi nilai untuk pengkategorian jumlah kasus DBD, pengkategorian kepadatan penduduk, dan pengkategorian untuk

pengkategorian kepadatan penduduk, dan pengkategorian untuk radius fasilitas kesehatan serta rawa. Berikut tampilan antarmuka untuk hasil preproses data (Gambar 4.2), setelah melakukan preproses data maka perangkat lunak siap untuk melakukan proses rule mining dengan menentukan parameter support dan confidence terlebih dahulu, tampilan antarmuka hasil dari pencarian frequent itemset ditampilkan pada Gambar 4.3 dan pada Gambar 4.4 adalah tampilan antarmuka rule-rule yang dibentuk berdasarkan frequent itemset yang didapat.

HASIL PREPROSES DATA								
KASUS ORG	ID	KELOMPOK	FASKES	ORG	KAPADATAN	RAWA		
Kasus Aler Misk Merasa ORG	2	Jawa	3579810001	Banyuwangi	FASKES Jauh	ORG Rendah	Kepadatan Rendah	Rawa_Jauh
			3579810002	Sumur Wilut	FASKES Jauh	ORG Sedang	Kepadatan Rendah	Rawa_Jauh
			3579810003	Lidah Wetan	FASKES Jauh	ORG Rendah	Kepadatan Rendah	Rawa_Dekat
Batas Rumah Misk Malasari ORG	5	Jawa	3579810005	Jerat	FASKES Dekat	ORG Rendah	Kepadatan Rendah	Rawa_Dekat
			3579810006	Lukarembi	FASKES Jauh	ORG Rendah	Kepadatan Rendah	Rawa_Dekat
			3579820002	Karang Wang	FASKES Jauh	ORG Sedang	Kepadatan Tinggi	Rawa_Dekat
			3579821003	Bales Kumpuk	FASKES Jauh	ORG Tinggi	Kepadatan Sedang	Rawa_Dekat
			3579822007	Wirang	FASKES Dekat	ORG Tinggi	Kepadatan Sedang	Rawa_Dekat
			3579822009	Pratakkakendal	FASKES Jauh	ORG Tinggi	Kepadatan Rendah	Rawa_Dekat
			3579822019	Dukuh Palis	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Dekat
			3579822022	Dukuh Kumpang	FASKES Dekat	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579820004	Simalan Kerto	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579830012	Bendulmaris	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579830013	Sudarmas	FASKES Dekat	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579831008	Kubung Ciri	FASKES Dekat	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579831007	Jambangan	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579831009	Karah	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Dekat
			3579832002	Dukuh Menanggal	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579832009	Kerintang	FASKES Jauh	ORG Tinggi	Kepadatan Sedang	Rawa_Dekat
			3579840013	Kedung Baruk	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579840014	Panunggan Sari	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579841010	Pragan	FASKES Jauh	ORG Sedang	Kepadatan Tinggi	Rawa_Jauh
			3579842002	Bungkul Menanggal	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Dekat
			3579850001	Kepunden Jangkungan	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579850008	Gubung Bathi	FASKES Jauh	ORG Sedang	Kepadatan Sedang	Rawa_Jauh
			3579851007	Masyur Sabrangan	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579851011	Mulyorejo	FASKES Dekat	ORG Tinggi	Kepadatan Sedang	Rawa_Jauh
			3579860002	Ponong Semo	FASKES Dekat	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579864004	Gubung	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579864005	Arlangga	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579870002	Widodarmas	FASKES Dekat	ORG Tinggi	Kepadatan Tinggi	Rawa_Dekat
			3579870005	Ngajati	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579890003	Tegalwan	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579890008	Banjarump	FASKES Dekat	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh
			3579890004	Kumpang Kragan	FASKES Jauh	ORG Tinggi	Kepadatan Tinggi	Rawa_Jauh

Gambar 4.2 Form Hasil Preproses



Hasil Itemset

Parameter
 Support (dalam persen) 20
 Confidence (dalam persen) 30

HASIL 1 ITEMSET

ID	ITEM	JUMLAH	SUPPORT
1	Kepadatan Tinggi	180	0.47
2	DBD Tinggi	117	0.49
3	FASKES Jauh	110	0.47
4	FASKES Dekat	93	0.33

HASIL 2 ITEMSET

ID	ITEM	ITEM2	JUMLAH	SUPPORT
1	Kepadatan Tinggi	DBD Tinggi	45	0.30
2	Kepadatan Tinggi	FASKES Jauh	68	0.42
3	Kepadatan Tinggi	FASKES Dekat	43	0.25
4	DBD Tinggi	FASKES Jauh	75	0.44
5	DBD Tinggi	FASKES Dekat	43	0.25

HASIL 3 ITEMSET

ID	ITEM	ITEM2	ITEM3	JUMLAH	SUPPORT
1	Kepadatan Tinggi	DBD Tinggi	FASKES Jauh	15	0.34
2	Kepadatan Tinggi	FASD Tinggi	FASKES Dekat	14	0.31

HASIL 4 ITEMSET

ID	ITEM	ITEM	ITEM	JUMLAH	SUPPORT
1					

Gambar 4.3 Tampilan Antarmuka Hasil Pencarian *frequent itemset*.

Hasil Rule yang terbentuk

RULE-RULE YANG TERBENTUK

IF	THEN	Support	Conf
DBD Tinggi	Kepadatan Tinggi, FASD 0.21	0.69	0.3
Kepadatan Tinggi	DBD Tinggi, FASKES D 0.21	0.67	0.31
DBD Tinggi	FASKES Dekat	0.69	0.30
Kepadatan Tinggi	FASKES Dekat	0.25	0.37
Kepadatan Tinggi, DBD/FASKES Dekat		0.21	0.55
DBD Tinggi	Kepadatan Tinggi, FASD 0.34	0.69	0.49
FASKES Jauh	Kepadatan Tinggi, DBD 0.38	0.67	0.51
Kepadatan Tinggi	DBD Tinggi, FASKES J 0.34	0.67	0.51
Kepadatan Tinggi, DBD/FASKES Jauh		0.34	0.55
FASKES Jauh	Kepadatan Tinggi	0.42	0.67
Kepadatan Tinggi	FASKES Jauh	0.42	0.63
DBD Tinggi	FASKES Jauh	0.44	0.64
FASKES Dekat	Kepadatan Tinggi, DBD 0.25	0.79	0.64
FASKES Jauh	DBD Tinggi	0.44	0.66
FASKES Dekat	DBD Tinggi	0.25	0.32
FASKES Dekat	Kepadatan Tinggi	0.25	0.32
DBD Tinggi, FASKES J	Kepadatan Tinggi	0.34	0.44
DBD Tinggi	Kepadatan Tinggi	0.35	0.69
Kepadatan Tinggi, FASD 0.34	DBD Tinggi	0.34	0.42
Kepadatan Tinggi	DBD Tinggi	0.35	0.67
DBD Tinggi, FASKES D	Kepadatan Tinggi	0.21	0.64
Kepadatan Tinggi, FASD 0.21	DBD Tinggi	0.21	0.25

Waktu Preproses 5.375 detik
 Waktu Pencarian frequent pattern 0.125 detik
 Waktu Pencarian Rule 0.0625 detik
 Waktu TOTAL 5.5 detik

Gambar 4.4 Tampilan Antarmuka Rule-Rule yang dibentuk. Hasil akhir dari perangkat lunak ini adalah asosiasi antara kasus DBD dengan faktor-faktor yang mempengaruhi, sehingga rule-rule yang telah didapat akan diseleksi untuk mendapatkan hasil

yang diinginkan, yaitu hanya rule-rule yang memuat item DBD saja yang akan ditampilkan, seleksi rule dilakukan dengan mengklik tombol seleksi rule pada form hasil rule, seperti yang bisa dilihat pada gambar 4.5 dibawah ini.

Hasil Rule yang terbentuk

RULE-RULE YANG TERBENTUK

Idra	Nama	StLog	Saka	Cor	
▶	Kepadatan Tinggi	DBD Tinggi , FASKES Dekat	0,23	0,82	0,28
	DBD Tinggi	kepadatan Tinggi , FASKES Dekat	0,23	0,69	0,33
	Kepadatan Tinggi	FASKES Dekat	0,29	0,82	0,36
	DBD Tinggi	FASKES Dekat	0,26	0,62	0,36
	Kepadatan Tinggi , DBD Tinggi	FASKES Dekat	0,22	0,64	0,28
	Kepadatan Tinggi	DBD Tinggi , FASKES Jauh	0,4	0,82	0,49
	DBD Tinggi	kepadatan Tinggi , FASKES Jauh	0,4	0,69	0,58
	FASKES Jauh	kepadatan Tinggi , DBD Tinggi	0,4	0,57	0,6
	Kepadatan Tinggi , DBD Tinggi	FASKES Jauh	0,4	0,64	0,62
	DBD Tinggi	FASKES Jauh	0,44	0,68	0,64
	Kepadatan Tinggi	FASKES Jauh	0,53	0,82	0,66
	FASKES Jauh	DBD Tinggi	0,44	0,67	0,66
	FASKES Dekat	kepadatan Tinggi , DBD Tinggi	0,23	0,93	0,7
	Kepadatan Tinggi , FASKES Jauh	DBD Tinggi	0,4	0,52	0,75
	FASKES Dekat	DBD Tinggi	0,25	0,53	0,76
	Kepadatan Tinggi	DBD Tinggi	0,44	0,82	0,78
	FASKES Jauh	kepadatan Tinggi	0,53	0,67	0,79
	Kepadatan Tinggi , FASKES Dekat	DBD Tinggi	0,23	0,29	0,79
	FASKES Dekat	kepadatan Tinggi	0,29	0,33	0,88
	DBD Tinggi , FASKES Jauh	kepadatan Tinggi	0,4	0,44	0,91
	DBD Tinggi , FASKES Dekat	kepadatan Tinggi	0,23	0,26	0,92
	DBD Tinggi	kepadatan Tinggi	0,44	0,45	0,93

Idra	Nama	StLog	Saka	Cor	
▶	Kepadatan Tinggi	DBD Tinggi , FASKES Dekat	0,23	0,82	0,28
	Kepadatan Tinggi	DBD Tinggi , FASKES Jauh	0,4	0,82	0,49
	FASKES Jauh	kepadatan Tinggi , DBD Tinggi	0,4	0,67	0,6
	FASKES Jauh	DBD Tinggi	0,44	0,67	0,66
	FASKES Dekat	kepadatan Tinggi , DBD Tinggi	0,23	0,33	0,7
	Kepadatan Tinggi , FASKES Jauh	DBD Tinggi	0,4	0,53	0,75
	FASKES Dekat	DBD Tinggi	0,25	0,53	0,76
	Kepadatan Tinggi , FASKES Dekat	DBD Tinggi	0,23	0,29	0,79

Waktu Preproses	5	detik.	SELEKSI RULE
Waktu Pemecahan frekuensi partisi	0,140625	detik.	
Waktu Pemecahan Nilai	0,09375	detik.	
Waktu TOTAL	5,140625	detik.	

Gambar 4.5 Tampilan Antarmuka Hasil Seleksi Rule.

4.3. Uji Coba Perangkat Lunak

Dalam bab ini dijelaskan hasil uji coba perangkat lunak dengan parameter berbeda.

4.3.1. Lingkungan Uji Coba

Perangkat keras yang dipakai untuk uji coba perangkat lunak ini adalah komputer dengan Processor Pentium 4 3.06GHz dengan memori utama 512 MB yang telah diinstall Microsoft SQL Server 2000 Enterprise Manager dengan Sistem Operasi Microsoft Windows XP Professional SP 2.

Spesifikasi Dataset

Untuk melakukan pengujian *Dataset* yang dipersiapkan antara lain :

- | | |
|-----------------------------|-------------------|
| a. Tabel Kecamatan | : 28 record data |
| b. Tabel Kelurahan | : 163 record data |
| c. Tabel LayKecamatan | : 28 record data |
| d. Tabel Laykelurahan | : 163 record data |
| e. Tabel Kepadatan penduduk | : 163 record data |
| f. Tabel Kasus DBD | : 163 record data |
| g. Tabel LayFasilitas | : 53 record data |
| h. Tabel Fasilitas | : 53 record data |

4.4. Hasil Uji Coba

Uji coba dilakukan dengan minimum support dan minimum confidence yang berbeda-beda. Berikut deskripsi hasil uji coba yang telah dilakukan:

4.4.1. Uji coba 1

- | | |
|---------------------------|--|
| <i>Minimum support</i> | : 10 |
| <i>Minimum confidence</i> | : 75 |
| <i>Frequent itemset</i> | : 4-frequent itemset (dua buah record) |
| Jumlah Rule Akhir | : 9 |
| Waktu Proses | : 7,640625 detik |

Hasil *frequent itemset* :

Tabel 4.1 Hasil semua *frequent itemset*

ITEMI	JUMLAH	SUPPORT			
Kepadatan Tinggi	109	0.67			
Kepadatan Sedang	25	0.15			
Kepadatan Rendah	29	0.18			
DBD Tinggi	119	0.73			

DBD Rendah	40	0.25			
FASKES Jauh	110	0.67			
FASKES Dekat	53	0.33			
Rawa Jauh	127	0.78			
Rawa Dekat	36	0.22			
ITEM1	ITEM2	JUMLAH	SUPPORT		
Kepadatan Tinggi	DBD Tinggi	92	0.56		
Kepadatan Tinggi	FASKES Jauh	68	0.42		
Kepadatan Tinggi	FASKES Dekat	41	0.25		
Kepadatan Tinggi	Rawa Jauh	93	0.57		
Kepadatan Tinggi	Rawa Dekat	16	0.1		
Kepadatan Sedang	DBD Tinggi	18	0.11		
Kepadatan Sedang	FASKES Jauh	19	0.12		
Kepadatan Rendah	DBD Rendah	20	0.12		
Kepadatan Rendah	FASKES Jauh	23	0.14		
Kepadatan Rendah	Rawa Jauh	19	0.12		
DBD Tinggi	FASKES Jauh	77	0.47		
DBD Tinggi	FASKES Dekat	42	0.26		
DBD Tinggi	Rawa Jauh	97	0.6		

DBD Tinggi	Rawa Dekat	22	0.13		
DBD Rendah	FASKES Jauh	32	0.2		
DBD Rendah	Rawa Jauh	26	0.16		
FASKES Jauh	Rawa Jauh	86	0.53		
FASKES Jauh	Rawa Dekat	24	0.15		
FASKES Dekat	Rawa Jauh	41	0.25		
ITEM1	ITEM2	ITEM3	JUMLAH	SUPPORT	
Kepadatan Tinggi	DBD Tinggi	FASKES Jauh	57	0.35	
Kepadatan Tinggi	DBD Tinggi	FASKES Dekat	35	0.21	
Kepadatan Tinggi	DBD Tinggi	Rawa Jauh	79	0.48	
Kepadatan Tinggi	FASKES Jauh	Rawa Jauh	59	0.36	
Kepadatan Tinggi	FASKES Dekat	Rawa Jauh	34	0.21	
Kepadatan Rendah	DBD Rendah	FASKES Jauh	17	0.1	
Kepadatan Rendah	FASKES Jauh	Rawa Jauh	16	0.1	
DBD Tinggi	FASKES Jauh	Rawa Jauh	63	0.39	
DBD Tinggi	FASKES Dekat	Rawa Jauh	34	0.21	
DBD Rendah	FASKES Jauh	Rawa Jauh	22	0.13	
ITEM1	ITEM2	ITEM3	ITEM4	JUMLAH	SUPPORT
Kepadatan Tinggi	DBD Tinggi	FASKES Jauh	Rawa Jauh	51	0.31

Kepadatan Tinggi	DBD Tinggi	FASKES Dekat	Rawa Jauh	28	0.17
------------------	------------	--------------	-----------	----	------

Tabel 4.2 Hasil rule

Jika	Maka	STotal	SJika	Conf
Rawa Jauh	DBD Tinggi	0.6	0.78	0.77
FASKES Dekat	DBD Tinggi	0.26	0.33	0.79
Kepadatan Tinggi , FASKES Dekat , Rawa Jauh	DBD Tinggi	0.17	0.21	0.81
Kepadatan Tinggi , FASKES Jauh	DBD Tinggi	0.35	0.42	0.83
FASKES Dekat , Rawa Jauh	DBD Tinggi	0.21	0.25	0.84
Kepadatan Tinggi	DBD Tinggi	0.56	0.67	0.84
Kepadatan Tinggi , FASKES Dekat	DBD Tinggi	0.21	0.25	0.84
Kepadatan Tinggi , Rawa Jauh	DBD Tinggi	0.48	0.57	0.84
Kepadatan Tinggi , FASKES Jauh , Rawa Jauh	DBD Tinggi	0.31	0.36	0.86

4.4.2. Uji coba 2

<i>Minimum support</i>	: 30
<i>Minimum confidence</i>	75
<i>Frequent itemset</i>	: 4-frequent itemset (1 buah record)
Jumlah Rule Akhir	: 5
Waktu Proses	: 6,515625 detik

Hasil *frequent itemset* :

Tabel 4.3 Hasil semua *frequent itemset*

ITEM1	JUMLAH	SUPPORT			
Kepadatan	109	0.67			

Tinggi					
DBD Tinggi	119	0.73			
FASKES Jauh	110	0.67			
FASKES Dekat	53	0.33			
Rawa Jauh	127	0.78			
ITEM1	ITEM2	JUMLAH	SUPPORT		
Kepadatan Tinggi	DBD Tinggi	92	0.56		
Kepadatan Tinggi	FASKES Jauh	68	0.42		
Kepadatan Tinggi	Rawa Jauh	93	0.57		
DBD Tinggi	FASKES Jauh	77	0.47		
DBD Tinggi	Rawa Jauh	97	0.6		
FASKES Jauh	Rawa Jauh	86	0.53		
ITEM1	ITEM2	ITEM3	JUMLAH	SUPPORT	
Kepadatan Tinggi	DBD Tinggi	FASKES Jauh	57	0.35	
Kepadatan Tinggi	DBD Tinggi	Rawa Jauh	79	0.48	
Kepadatan Tinggi	FASKES Jauh	Rawa Jauh	59	0.36	
DBD Tinggi	FASKES Jauh	Rawa Jauh	63	0.39	
ITEM1	ITEM2	ITEM3	ITEM4	JUMLAH	SUPPORT
Kepadatan Tinggi	DBD Tinggi	FASKES Jauh	Rawa Jauh	51	0.31

Tabel 4.4 Hasil rule

Jika	Maka	STotal	SJika	Conf
Rawa Jauh	DBD Tinggi	0.6	0.78	0.77
Kepadatan Tinggi , FASKES Jauh	DBD Tinggi	0.35	0.42	0.83
Kepadatan Tinggi	DBD Tinggi	0.56	0.67	0.84
Kepadatan Tinggi , Rawa Jauh	DBD Tinggi	0.48	0.57	0.84
Kepadatan Tinggi , FASKES Jauh , Rawa Jauh	DBD Tinggi	0.31	0.36	0.86

4.4.3. Uji coba 3

<i>Minimum support</i>	: 40
<i>Minimum confidence</i>	: 75
<i>Frequent itemset</i>	: 3-frequent itemset (1 buah record)
Jumlah Rule Akhir	: 3
Waktu Proses	: 6,375 detik

Hasil *frequent itemset* :

Tabel 4.5 Hasil semua *frequent itemset*

ITEM1	JUMLAH	SUPPORT		
Kepadatan Tinggi	109	0.67		
DBD Tinggi	119	0.73		
FASKES Jauh	110	0.67		
Rawa Jauh	127	0.78		
ITEM1	ITEM2	JUMLAH	SUPPORT	
Kepadatan Tinggi	DBD Tinggi	92	0.56	
Kepadatan Tinggi	FASKES Jauh	68	0.42	

Kepadatan Tinggi	Rawa Jauh	93	0.57	
DBD Tinggi	FASKES Jauh	77	0.47	
DBD Tinggi	Rawa Jauh	97	0.6	
FASKES Jauh	Rawa Jauh	86	0.53	
ITEM1	ITEM2	ITEM3	JUMLAH	SUPPORT
Kepadatan Tinggi	DBD Tinggi	Rawa Jauh	79	0.48

Tabel 4.6 Hasil rule

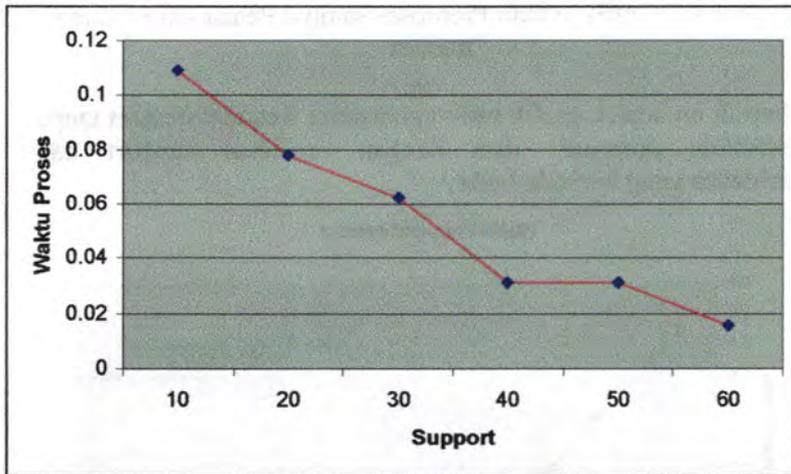
Jika	Maka	STotal	SJika	Conf
Rawa Jauh	DBD Tinggi	0.6	0.78	0.77
Kepadatan Tinggi	DBD Tinggi	0.56	0.67	0.84
Kepadatan Tinggi , Rawa Jauh	DBD Tinggi	0.48	0.57	0.84

Untuk melakukan analisa terhadap proses *rule mining*, perlu dilakukan uji coba dengan memasukkan berbagai macam variasi parameter support dan confidence yang dapat dilihat pada tabel berikut ini:

Tabel 4.7 Hasil uji

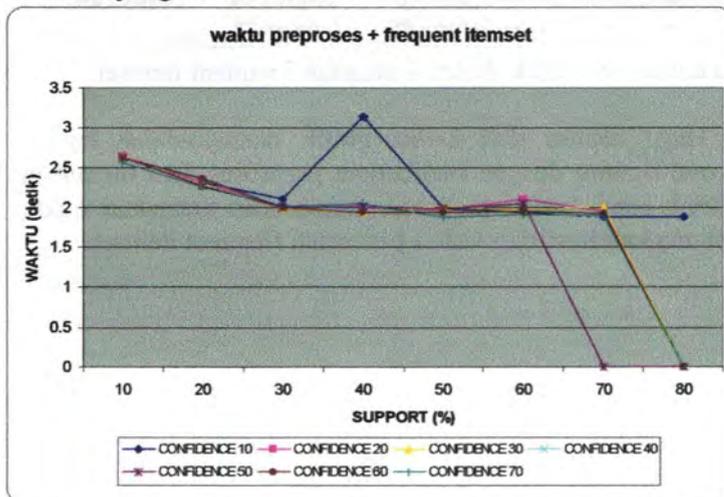
Support	Confidence	Waktu Pencarian frequent itemset	Waktu pencarian rule
10	80	1.65625	0.109375
20	80	0.640625	0.078125
30	80	0.21875	0.0625
40	80	0.140625	0.03125
50	80	0.140625	0.03125
60	80	0.140625	0.015625

Dibawah ini adalah grafik perbandingan waktu proses dengan support dan confidence yang diberikan.



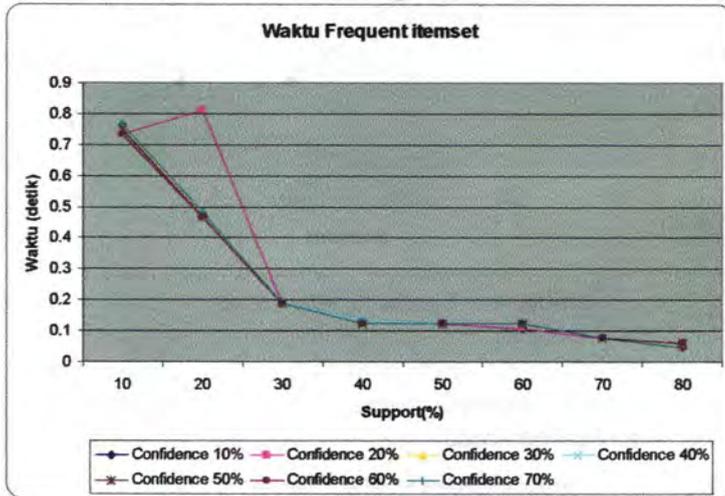
Gambar 4.6. Grafik Waktu Proses

Dibawah ini adalah grafik waktu preproses sampai dengan waktu pencarian frequent itemset dengan memakai support dan confidence yang berbeda-beda.



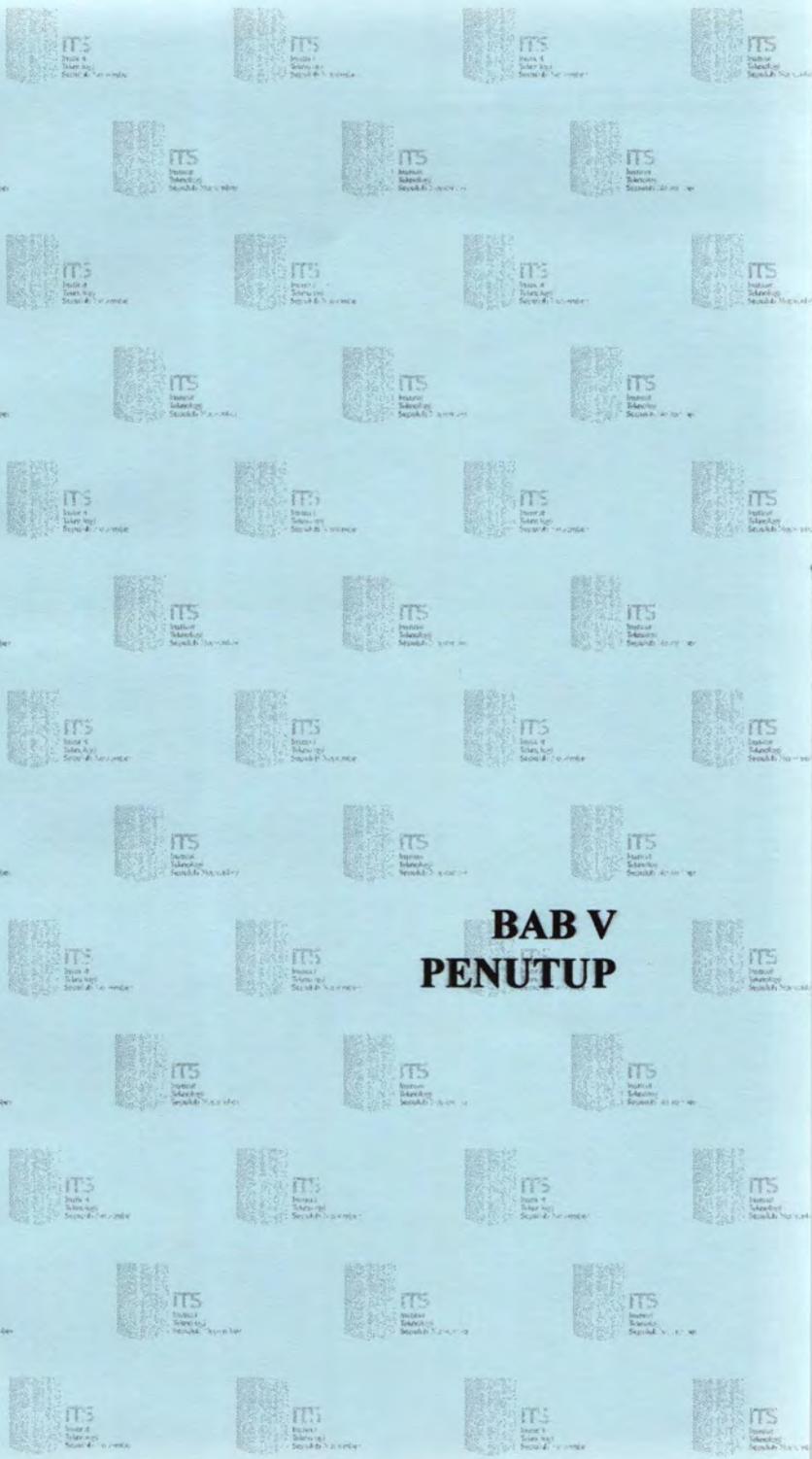
Gambar 4.7 Grafik Waktu Preproses sampai Pencarian Frequent Itemset.

Dibawah ini adalah grafik waktu pencarian frequent itemset tanpa melibatkan preproses data dengan memakai support dan confidence yang berbeda-beda



Gambar 4.8 Grafik Waktu Pencarian Frequent Itemset.

Hasil analisa dari kedua grafik diatas adalah bahwa proses rule mining dengan melibatkan preproses data memakan waktu lebih lama, karena preproses data sendiri memakan waktu minimal dua kali lipat dari waktu pencarian frequent itemset.



BAB V PENUTUP

BAB V PENUTUP

Pada bab ini akan dijelaskan tentang beberapa hal yang dapat disimpulkan dari penyelesaian Tugas Akhir ini beserta beberapa kemungkinan pengembangan yang bisa dilakukan.

5.1. Kesimpulan

Dari hasil uji coba perangkat lunak pada BAB IV dapat diambil beberapa kesimpulan sebagai berikut :

1. Pola asosiasi cenderung menunjukkan korelasi yang kuat antara “kepadatan Tinggi” dengan “DBD Tinggi”, hal ini menunjukkan bahwa jika suatu kelurahan dengan tingkat kepadatan yang tinggi akan memiliki jumlah kasus DBD yang tinggi juga.
2. Hasil preproses data dipengaruhi oleh inputan yang diberikan oleh *user* untuk kategorisasi jumlah kasus DBD, tingkat kepadatan, jarak kelurahan dengan rawa dan fasilitas kesehatan.
3. Proses pencarian *frequent itemset* sangat dipengaruhi oleh parameter *support* yang diberikan oleh *user*, semakin besar *support* yang diberikan maka semakin cepat waktu yang diperlukan.

5.2. Saran

Pengembangan yang dapat dilakukan pada perangkat lunak ini adalah memperbarui dan meningkatkan data yang dipakai, baik data spasial maupun data non spasial, selain itu ruang lingkup dari kasus yang bersangkutan juga bisa dikembangkan, apabila dalam tugas akhir kali ini yang menjadi obyek utama adalah kotamadya Surabaya maka untuk selanjutnya bisa diperluas sampai lingkup propinsi atau bahkan lingkup Indonesia.

DAFTAR PUSATAKA

DAFTAR PUSTAKA

- Agrawal, R., Imielinski, T. and Swami, A. 1993. "Mining Association Rules Between Sets of Items in Large Databases". *ACM SIGMOD International Conference on Management of Data*. Washington D. C., USA.
- Ester, M., Frommelt, A., Kriegel, H.-P. and Sander, J. 2000. "Spatial data mining: database primitives, algorithms and efficient DBMS support", *Data Mining and Knowledge Discovery*, 4, 193-216.
- Ester, M., Frommelt, A., Kriegel, H.-P. and Sander, J. 2001. "Algorithms and applications for spatial data mining". IN MILLER, H. J. & HAN, J. (Eds.) *Geographic Data Mining and Knowledge Discovery*. New York, USA, Taylor and Francis.
- Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996) "The KDD process for extracting useful knowledge from volumes of data", *Communications of the ACM* 39(11), 27-34.
- Jiawei Han, Micheline Kamber, 2001. *Data Mining: Concept and Techniques*, Morgan Kaufmann, Amerika Serikat.
- K.Koperski and J. Han. 1995. "Discovery of spatial association rules in geographic information databases". *Lecture Notes In Computer Science*, 951:47-66.
- Mennis, J. & Liu, J. W. 2005. "Mining Association Rules in Spatio-Temporal Data: An Analysis of Urban Socioeconomic and Land Cover Change". *Transactions in GIS*, 9, 5 - 17.
- Rakesh Agrawal, Ramakrishnan Srikan, 1994. "Fast Algorithms for Mining Association Rules", *Proc VLDB*.



BIOGRAFI



Penulis lahir di Mojokerto, Jawa Timur pada tanggal 22 Oktober 1981. Pendidikan formal yang telah ditempuh yaitu SDN Pongangan Indah Gresik, SLTPN 3 Gresik, dan SMUN 5 Surabaya.

Memulai jenjang pendidikan tinggi pada tahun 2000 di Institut Teknologi Sepuluh Nopember Surabaya sebagai mahasiswa jurusan Matematika dan memilih bidang studi Ilmu Komputer.

Aktif di Himpunan Mahasiswa Matematika, Karang Taruna, dan Ikatan Alumni SMUN 5 Surabaya.