

**ANALISIS KETERKAITAN DATA KEPENDUDUKAN UNTUK  
PERENCANAAN PEMBANGUNAN DAERAH DENGAN  
MENGUNAKAN ALGORITMA APRIORI**

*THE ANALYSIS OF DEMOGRAPHY DATA FOR PLANNING OF  
REGIONAL DEVELOPMENT WHICH IS USE APRIORI ALGORITHM*

**MUH. SAKIR  
P2700211403**



**PROGRAM PASCASARJANA  
UNIVERSITAS HASANUDDIN  
MAKASAR  
2013**

## DAFTAR ISI

Halaman

HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN.....	ii
DAFTAR ISI .....	iii
KATA PENGANTAR.....	iv
ABSTRAK .....	v
DAFTAR GAMBAR.....	vi
DAFTAR TABEL .....	vii
<b>BAB I PENDAHULUAN</b>	
A. Latar Belakang.....	1
B. Rumusan Masalah .....	3
C. Tujuan Penelitian .....	3
D. Manfaat Penelitian .....	4
E. Batasan Penelitian.....	4
<b>BAB II TINJAUAN PUSTAKA</b>	
A. Deskripsi Data Kependudukan.....	7
B. Perencanaan Pembangunan Daerah.....	8
C. <i>Association Rule</i> .....	10
D. Algoritma Apriori.....	11

E.	Penelitian Terkait.....	14
F.	Kerangka Konseptual .....	15

### **BAB III METODE PENELITIAN**

A.	Rancangan Penelitian .....	16
B.	Sumber Data.....	18
C.	Teknik Pengumpulan Data .....	19
D.	Metode Perancangan Sistem.....	25
E.	<i>Flow chart</i> sistem yang akan dibuat.....	29
F.	Alat dan Bahan Penelitian.....	33

### **BAB IV HASIL DAN PEMBAHASAN**

A.	Hasil Penelitian.....	34
B.	Pembahasan .....	45
C.	Pengujian Aplikasi .....	53
D.	Hasil Evaluasi Menggunakan Kuisisioner .....	54

### **BAB V KESIMPULAN DAN SARAN**

A.	Kesimpulan.....	58
B.	Saran .....	58

### **DAFTAR PUSTAKA**

### **LAMPIRAN**

## KATA PENGANTAR

Puji syukur Kehadirat Allah S.W.T yang telah memberikan kesehatan, dan kesempatan sehingga penulis dapat menyelesaikan laporan hasil penelitian dengan Judul **“Analisis Keterkaitan Data Kependudukan untuk Perencanaan Pembangunan Daerah dengan Menggunakan Algoritma Apriori”**

Dalam penyusunan dan penulisan hasil penelitian ini, Penulis menyadari bahwa tidak terlepas dari bantuan dan partisipasi dari berbagai pihak, oleh karena itu dalam kesempatan ini penulis menyampaikan rasa terimakasih kepada :

1. Bapak Dr. Ing. Faizal Arya Samman, ST., MT dan Bapak Dr. Eng. Armin Lawi, M.Eng selaku dosen pembimbing kami.
2. Bapak/ibu dosen pascasarjana UNHAS pada program studi Teknik Elektro.
3. Kepada kedua orang tua, saudara dan keluarga yang telah memberikan dukungan dalam proses menempuh pendidikan.
4. Kepada istri tercinta
5. Teman – teman mahasiswa program Pasca Sarjana Teknik Informatika Universitas Hasanuddin angkatan 2011.
6. Serta seluruh pihak yang telah memberi dukungan.

Dengan segala kerendahan hati kami senantiasa terbuka untuk menerima saran dan kritikan yang bersifat konstruktif demi perbaikan dalam penyusunan laporan hasil penelitian ini, semoga segala kebaikan bernilai ibadah disisi Allah, S.W.T, Aamiin.

Makasar, Juni 2013

Penulis

## Abstrak

**MUH. SAKIR** (Analisis Keterkaitan Data Kependudukan Untuk Perencanaan Pembangunan Daerah dengan Menggunakan Algoritma Apriori) Pembimbing : **Dr. Ing. Faizal Arya Samman, ST., MT,** dan **Dr. Eng. Armin Lawi, M.Eng**

Penelitian ini bertujuan untuk menganalisis keterkaitan data kependudukan untuk menemukan pola hubungan antara lapangan usaha dari pekerjaan utama dengan kondisi ekonomi, melalui proses *machine learning* menggunakan algoritma apriori agar dapat menjadi rekomendasi perencanaan pembangunan daerah. Metode yang digunakan dalam penelitian ini meliputi tiga tahapan yaitu input data kependudukan yang telah diseleksi, data-data tersebut akan diproses oleh *machine learning* berdasarkan *association rules* menggunakan algoritma apriori, data yang telah diproses akan menampilkan nilai *support* dan nilai *confidence* yang menjadi dasar pengambilan kesimpulan. Hasil dari penelitian ini menunjukkan bahwa 86,27 % penduduk di Desa/Kelurahan Malino yang mempunyai pekerjaan utama dalam bidang pertanian tanaman padi dan palawija masih dalam kondisi prasejahtera, 91,81 % penduduk di Kelurahan Limbung yang mempunyai pekerjaan utama dalam bidang industri pengolahan masih prasejahtera, 84,93 % penduduk di Kelurahan Sungguminasa yang mempunyai pekerjaan utama dalam bidang industri pengolahan masih prasejahtera. Kesimpulan dari penelitian ini menunjukkan bahwa program aplikasi komputer dengan algoritma apriori dapat menggali informasi keterkaitan data antara pekerjaan utama dan kondisi ekonomi suatu wilayah.

Kata kunci: *Data kependudukan, perencanaan pembangunan, association rule, algoritma apriori.*

### **Abstract**

**MUH. SAKIR** (*The Analysis of Demography Data for Planning of Regional Development Which is Use Apriori Algorithm*). Advisor: **Dr. Ing.FaizalAryaSamman, ST, MT** and **Dr. Eng. Armin Lawi, M.Eng.**

*This research aims to analyze demography data to find out the correlation between the major occupation and level of welfare through machine learning process. This process used apriorialgorithm for recommendation to regional development plan. The method consisted of three steps. The first step input the selected demography data. Second, process the data through machine learning based on association rules using apriori algorithm. Next, the data that have been processed will show the support and confidence value. The conclusion will make based on this value. The result shows that 86,27% of people who live in Malino, who have a major occupation as farmer still live under welfare line, 91,81 % of people who live in limbung who have occupation as processing industrystill live under welfare line. 84,93 % of people who live in Sungguminasa who have occupation as processing industry still live under welfare line. The conclusion of this research indicate that the computer application program with a priori algorithm can explore the data link information between main job and economic conditions of a region*

*Key words: Demography data, development plan, association rule, apriori algorithm,*

## DAFTAR GAMBAR

		Halaman
Gambar 1	Bagan kerangka konseptual .....	15
Gambar 2	Garis besar rancangan penelitian.....	16
Gambar 3	Aliran data dalam inferensi <i>machine learning</i> .....	18
Gambar 4	Diagram <i>use case</i> .....	26
Gambar 5	<i>Class diagram</i> .....	27
Gambar 6	<i>Activity diagram</i> .....	28
Gambar 7	Flow chart input data.....	29
Gambar 8	Flow chart pemilihan C1.....	30
Gambar 9	Flow chart proses <i>pruning</i> (pemangkasan) pada C1	30
Gambar 10	Flow chart pemilihan C2.....	31
Gambar 11	Flow chart proses <i>pruning</i> (pemangkasan) pada C2...	31
Gambar 12	Flow chart proses nilai <i>support, confidence, final rule</i> ...	32
Gambar 13	Form 1 pendataan penduduk oleh BPS.....	34
Gambar 14	Form 2 pendataan penduduk oleh BPS.....	35
Gambar 15	Hasil nilai <i>support</i> .....	40
Gambar 16	Hasil nilai <i>confidence</i> .....	41
Gambar 17	Prosentase pekerjaan utama penduduk.....	49

Gambar 18	<i>Pie chart</i> Kondisi ekonomi Kel. Malino.....	50
Gambar 19	Sebaran pekerjaan utama penduduk pada sampel 2....	51
Gambar 20	Kondisi ekonomi Desa/Kel. Limbung.....	51
Gambar 21	<i>Pie Chart</i> sebaran pekerjaan pada sampel 3.....	52
Gambar 22	Kondisi ekonomi Kel. Sungguminasa.....	52
Gambar 23	<i>Pie chart</i> prosentase kemudahan <i>user</i> dalam memahami tahapan tampilan pada program .....	55
Gambar 24	<i>Pie chart</i> prosentase kemudahan & kecepatan aplikasi ketika <i>user</i> menginput data .....	55
Gambar 25	<i>Pie chart</i> prosentase penilaian <i>user</i> terhadap tata letak tampilan.....	56
Gambar 26	<i>Pie chart</i> prosentase kejelasan dari setiap keterangan yang ditampilkan pada <i>interface</i> .....	56
Gambar 27	<i>Pie chart</i> prosentase penilaian <i>user</i> terhadap kegunaan program.....	57



## DAFTAR TABEL

		Halaman
Tabel 1	Data lapangan Usaha.....	19
Tabel 2	Data kondisi ekonomi.....	21
Tabel 3	Kategorisasi lapangan usaha dari pekerjaan utama.....	23
Tabel 4	Contoh tabel data awal.....	25
Tabel 5	Format data awal.....	36
Tabel 6	Pemilihan kandidat untuk satu data (C1).....	37
Tabel 7	Hasil <i>prune</i> (pemangkasan) setelah diberi <i>threshold</i> =1.....	38
Tabel 8	Pemilihan kandidat untuk <i>rule</i> dua data (C2).....	39
Tabel 9	Nilai <i>support</i> dan <i>confidence</i> kandidat dua data.....	42
Tabel 10	<i>Rule</i> pasangan data yang terseleksi.....	43
Tabel 11	Keterangan pada masing-masing <i>rule</i> .....	46
Tabel 12	Pengujian <i>black box</i> .....	53

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang Masalah**

Jumlah penduduk Indonesia yang banyak, membutuhkan tata kelola yang baik seperti pemetaan tingkat pendidikan, pekerjaan, kepadatan suatu wilayah dan data-data kependudukan lainnya sehingga pemerintah dalam mengambil kebijakan dapat lebih optimal dan tepat sasaran. Kebijakan yang tepat ditunjang oleh analisis permasalahan disertai data-data yang akurat sehingga perlu suatu metode untuk membantu pemerintah dalam optimalisasi kebijakan tersebut.

Keberadaan salah satu cabang pengetahuan yang disebut dengan data mining diharapkan dapat menjadi salah satu solusi untuk memanfaatkan data-data besar yang ada. Keberadaan data-data yang sangat besar dan akurat tentunya sangat potensial untuk dipergunakan dalam memberikan informasi-informasi penting bagi pemerintah dalam menentukan arah kebijakan daerah.

Pemerintah seharusnya mulai saat ini mengembangkan penggunaan teknologi basis data untuk melakukan pengelolaan data kependudukan. Jika pemerintah telah mempunyai data yang lengkap dan akurat maka data tersebut akan sangat bermanfaat untuk berbagai hal misalnya menganalisa pola hubungan antara lapangan usaha yang tersedia dengan kondisi ekonomi pada suatu wilayah, untuk mengetahui

apakah masyarakatnya sudah dominan sejahtera atau pra sejahtera, untuk mengetahui rata-rata bidang pekerjaan pada suatu wilayah, serta berbagai aspek lainnya yang dapat diteliti dari data-data kependudukan yang ada berdasarkan hasil sensus penduduk. Jika pemerintah telah mempunyai itu semua maka akan memudahkan pemerintah untuk melakukan perencanaan pembangunan agar pembangunan tepat sasaran.

Beberapa disiplin ilmu pengetahuan mengambil peranan untuk aktif dalam memberikan solusi untuk meminimalisir dampak negatif dari pertumbuhan penduduk yang sangat besar. Arah kebijakan pemerintah dapat lebih optimal jika didukung oleh sumber-sumber informasi yang tepat dan kebijakan yang optimal dapat memberikan solusi ditengah jumlah penduduk yang sangat besar.

Data yang lengkap akan menjadi dasar pemetaan, daerah mana dan butuh apa, dan daerah mana yang lebih prioritas, sehingga antara kebijakan dan kebutuhan setiap daerah sesuai. Namun permasalahan mendasarnya adalah terkadang data-data yang ada tidak dimanfaatkan sehingga kecenderungan pemetaan dan kebijakan belum optimal, sehingga berdampak pada tingkat kesejahteraan masyarakat yang berbanding terbalik dengan jumlah penduduk.

Teknik Informatika sebagai salah satu disiplin ilmu pengetahuan dapat berperan aktif dalam hal menganalisis data-data yang ada dalam jumlah besar untuk memberikan informasi penting dengan menggunakan

komputasi menggunakan perangkat lunak secara efisien dan optimal. Berdasarkan latar belakang tersebut maka penulis melakukan penelitian dengan judul “**Analisis Keterkaitan Data Kependudukan untuk Perencanaan Pembangunan Daerah dengan Menggunakan Algoritma Apriori**”

## **B. Rumusan Masalah**

Permasalahan yang diangkat dalam penelitian ini adalah:

1. Bagaimana menganalisis keterkaitan data kependudukan untuk menemukan pola hubungan antara lapangan usaha dari pekerjaan utama dengan kondisi ekonomi menggunakan algoritma apriori.
2. Bagaimana memberi kemudahan kepada pemerintah dan dinas-dinas di daerah agar dapat menggali informasi keterkaitan data secara *online* dari *output* sebuah aplikasi komputer menggunakan algoritma apriori.

## **C. Tujuan Penelitian**

Tujuan penelitian ini adalah :

1. Menganalisis keterkaitan data kependudukan untuk menemukan pola hubungan antara lapangan usaha dari pekerjaan utama dengan kondisi ekonomi menggunakan algoritma apriori.
2. Memberi kemudahan kepada pemerintah dan dinas-dinas di daerah agar dapat menggali informasi keterkaitan data secara *on line* dari *output* sebuah aplikasi komputer menggunakan algoritma apriori.

#### **D. Manfaat Penelitian**

1. Memudahkan dalam menganalisis keterkaitan antara pekerjaan utama penduduk dengan kondisi ekonomi suatu wilayah dengan menggunakan perangkat lunak pada penelitian ini.
2. Sebagai salah satu referensi pemerintah dalam pengambilan kebijakan yang didasarkan pada analisis keterkaitan data-data kependudukan dan kondisi ekonomi dengan bantuan *software* yang memanfaatkan algoritma apriori, khusus pada penelitian ini dibatasi pada keterkaitan antara pekerjaan utama penduduk dengan kondisi ekonomi
3. Memberikan gambaran tentang progresivitas kemajuan suatu wilayah
4. Membantu masyarakat dengan meningkatnya taraf kesejahteraan yang disebabkan oleh kebijakan yang tepat oleh pihak pemerintah.

#### **E. Batasan Penelitian**

Dari beberapa aspek yang tertulis pada form sensus penduduk, maka pada penelitian ini yang akan dikaji dibatasi pada data-data lapangan usaha dari pekerjaan utama dengan kondisi ekonomi masyarakat yang dikategorikan dalam masyarakat sejahtera dan pra sejahtera, dimana kondisi ekonomi didapatkan dari parameter keikutsertaan dalam program beras miskin (raskin).

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **A. Deskripsi Data Kependudukan**

Konstitusi mengamanatkan, bahwa negara berkewajiban memberikan pengakuan, jaminan perlindungan dan kepastian hukum yang adil serta perlakuan yang sama dihadapan hukum kepada penduduknya. Disamping itu, negara juga berkewajiban memberikan kesejahteraan kepada penduduk, serta memfasilitasi hak penduduk untuk berkomunikasi dan memperoleh informasi bagi pengembangan diri dan lingkungan sosialnya.

Tuntutan tersedianya data penduduk yang akurat, mutakhir dan lengkap telah diamanatkan dalam Undang-undang Nomor 25 Tahun 2000 tentang Program Pembangunan Nasional (Propenas) Tahun 2000 – 2004, pada Bidang Program Pengembangan dan Keserasian Kebijakan Kependudukan, dimana salah satu Kegiatan Pokoknya adalah melakukan pengkajian, pengembangan, dan penyediaan data dan informasi kependudukan yang akurat setiap saat dan lengkap serta menggambarkan karakteristik penduduk, baik pada tingkat makro maupun mikro.

Untuk melaksanakan amanat Propenas tersebut, pemerintah membentuk Direktorat Jenderal Administrasi Kependudukan Departemen Dalam Negeri, dengan tugas merumuskan dan melaksanakan kebijakan dan standarisasi teknis di bidang administrasi kependudukan, dan

fungsinya antara lain penyiapan perumusan kebijakan, pelaksana kebijakan, serta perumusan standar, norma, pedoman, kriteria dan prosedur di bidang informasi kependudukan.

Penyediaan data dan informasi kependudukan dimaksud, adalah data mulai dari tingkat kelurahan/desa, kecamatan, kabupaten/kota, propinsi sampai ke pusat, yang akurat, relevan, dan perekaman data langsung (real time) setelah dilakukan pembuktian secara hukum, sehingga data-informasi valid dan handal (reliable). Data dan informasi dimaksud mudah diakses oleh berbagai pihak berkepentingan (stakeholders) dengan tetap menerapkan sistem pengendalian dan jaminan keamanan data.

Dalam kaitan itu, Ditjen. Administrasi Kependudukan melalui Direktorat Informasi Kependudukan mengembangkan kebijakan Sistem Informasi Administrasi Kependudukan (SI AK) terpadu secara nasional dengan menerapkan Nomor Induk Kependudukan (NIK) dengan mendayagunakan teknologi informasi, yang diindikasikan terbangunannya data basis (bank data) penduduk nasional.

Data-informasi kependudukan yang terdapat dalam data basis penduduk, diharapkan didayagunakan untuk kepentingan perumusan kebijakan dan perencanaan pembangunan serta pelayanan publik. Khusus di lingkungan Ditjen. Administrasi Kependudukan, data-informasi kependudukan tersebut akan dikembangkan untuk mendukung perumusan kebijakan peningkatan kualitas penduduk, pengarahannya kuantitas penduduk, penataan persebaran penduduk, serta untuk

memenuhi kebutuhan proyeksi dan penyesuaian kebijakan kependudukan. (Anonim, 2002).

Data informasi kependudukan yang dipergunakan sebagai parameter dalam penelitian ini adalah: data keikutsertaan dalam program beras miskin, keikutsertaan dalam program beras miskin menjadi parameter untuk kondisi ekonomi suatu keluarga, kondisi ekonomi dalam penelitian ini terbagi menjadi dua yaitu sejahtera dan pra sejahtera.

Sesuai dengan pedoman umum penyaluran raskin oleh kementerian Koordinator bidang kesejahteraan rakyat republik indonesia tahun 2012 menyatakan bahwa tujuan program raskin adalah mengurangi beban pengeluaran Rumah Tangga Sasaran (RTS) melalui pemenuhan sebagian kebutuhan pangan pokok dalam bentuk beras, dimana sasaran penerima Raskin adalah berkurangnya beban pengeluaran RTS berdasarkan data PPLS-11 BPS dalam mencukupi kebutuhan pangan beras melalui pendistribusian beras bersubsidi sebanyak 180 Kg/RTS/tahun atau setara dengan 15 kg/RTS/bulan dengan harga tebus Rp1.600,00/kg netto di TD.

Rumah tangga sasaran yang dimaksudkan adalah rumah tangga miskin di Desa/Kelurahan yang berhak menerima Raskin dan terdaftar dalam Daftar Penerima Manfaat (DPM-1) yang ditetapkan oleh Kepala Desa/Lurah sebagai hasil Musyawarah Desa/Kelurahan dan disahkan oleh Camat sesuai hasil pendataan PPLS-11 BPS tahun 2011 (Kemenkokesra RI, 2012).



## **B. Perencanaan Pembangunan Daerah**

Perencanaan pembangunan daerah dimaksudkan untuk menghasilkan pembangunan yang dapat memnuhi kebutuhan masyarakat, salah satunya dengan melaksanakan Musrenbang (Musyawarah Perencanaan Pembangunan) mulai tingkat desa sampai provinsi. Melalui proses ini diharapkan program/kegiatan yang akan dilaksanakan oleh Pemda dapat memberikan manfaat kepada masyarakat secara optimal.

Beberapa permasalahan yang biasa dihadapi dalam perencanaan pembangunan daerah adalah:

- 1) Perhatian SKPD terhadap pentingnya dokumen perencanaan masih kurang.
- 2) Kemampuan SKPD dalam mengartikulasikan kebutuhan masyarakat masih kurang.
- 3) Kemampuan keuangan daerah dalam pembiayaan program pembangunan masih kurang.
- 4) Data pendukung perencanaan pembangunan kurang akurat.
- 5) Kemampuan masyarakat dalam mengidentifikasi kebutuhan pembangunan masih kurang.
- 6) Tingkat partisipasi masyarakat dalam musyawarah perencanaan pembangunan masih rendah.
- 7) Belum sinergisnya proses perencanaan pembangunan nasional dari pendekatan politik (proses politik) ke pendekatan teknokratik.

- 8) Dokumen perencanaan yang disusun belum menekankan pada perencanaan yang terfokus dan langsung dapat dilaksanakan, cenderung masih berupa *wishing list*, serta program dan kegiatan yang direncanakan masih belum disusun berdasarkan pada ketersediaan anggaran.
- 9) Adanya ego atau kepentingan antarsektor yang seringkali dinyatakan sebagai kesulitan untuk melakukan koordinasi, sehingga persoalan yang bersifat lintas sektor sering ditangani secara parsial dan terfragmentasi sehingga cenderung tidak menyentuh atau menyelesaikan persoalan yang sebenarnya.
- 10) Adanya anggapan atau asumsi bahwa dokumen perencanaan tersebut kurang mengakomodasi kebutuhan yang sebenarnya dari daerah dan antarwilayah.
- 11) Proses perencanaan teknokratik yang berbasis pada data sekunder dan primer, baik dari hasil monitoring dan evaluasi maupun hasil kajian/telaahan, dianggap masih belum memadai sehingga kekuatan data dan informasi dalam memproyeksikan arah pembangunan berikutnya masih lemah.
- 12) Masih terdapat kesulitan untuk memastikan adanya konsistensi antara perencanaan (program/kegiatan) pembangunan dan alokasi penganggarnya.
- 13) Belum optimalnya sistem pengendalian dan evaluasi pembangunan.

Isu strategis pada urusan perencanaan pembangunan adalah belum efektifnya perencanaan dari bawah (*bottom up planning*) yang disebabkan oleh kurang akuratnya data pendukung perencanaan pembangunan, kurangnya kemampuan masyarakat dalam mengidentifikasi kebutuhan pembangunan dan masih rendahnya tingkat partisipasi masyarakat dalam Musrenbang. (Abdullah, S, 2010).

### **C. Association Rules**

*Association rules* (aturan asosiasi) atau *affinity analysis* (analisis afinitas) berkenaan dengan studi tentang "apa bersama apa". Aturan asosiasi ingin memberikan informasi tersebut dalam bentuk hubungan "if-then" atau "jika-maka". Aturan ini dihitung dari data yang sifatnya probabilistik.

Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu prosentase kombinasi item tersebut. dalam database dan *confidence* (nilai kepastian) yaitu kuatnya hubungan antar item dalam aturan asosiatif. Analisis asosiasi didefinisikan suatu proses untuk menemukan semua aturan asosiatif yang memenuhi syarat minimum untuk support (minimum support) dan syarat minimum untuk confidence (minimum confidence). Jumlah besar aturan yang mungkin dikembangkan, perlu memiliki aturan-aturan yang cukup kuat tingkat ketergantungan antar item dalam antecedent dan consequent. Untuk mengukur kekuatan aturan asosiasi ini, digunakan

ukuran support dan confidence. Support adalah rasio antara jumlah transaksi yang memuat antecedent dan consequent dengan jumlah transaksi.

$$\text{Support} = \frac{\text{Jumlah transaksi yang mengandung antecedent dan consequent}}{\text{Jumlah transaksi}}$$

Jumlah transaksi yang meliputi semua item dalam *antecedent* dan *consequent* dengan jumlah transaksi yang meliputi semua item dalam *antecedent*

$$\text{Confidence} = \frac{\text{Jumlah transaksi dengan item dalam antecedent dan consequent}}{\text{Jumlah transaksi dengan item dalam antecedent}}$$

#### **D. Algoritma Apriori**

Algoritma apriori adalah algoritma untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola-pola item di dalam suatu database yang memiliki frekuensi atau support di atas ambang batas tertentu yang disebut dengan istilah minimum support (minsup) atau *threshold*. Pola frekuensi tinggi ini digunakan untuk menyusun aturan asosiatif dan juga beberapa teknik data mining lainnya.

Algoritma apriori dibagi menjadi beberapa tahap yang disebut iterasi. Tiap iterasi menghasilkan pola frekuensi tinggi dengan panjang yang sama dimulai dari pass pertama yang menghasilkan pola frekuensi tinggi dengan panjang satu. Pada iterasi pertama ini, support dari setiap item dihitung dengan men-scan database.

Setelah support dari setiap item didapat, item yang memiliki support diatas minimum support dipilih sebagai pola frekuensi tinggi dengan panjang 1 atau sering disingkat 1-itemset. Singkatan k-itemset berarti satu set yang terdiri dari k item.

Iterasi kedua menghasilkan 2-itemset yang tiap set-nya memiliki dua item. Pertama dibuat kandidat 2-itemset dari kombinasi semua 1-itemset. Lalu untuk tiap kandidat 2-itemset ini dihitung support-nya dengan men-scan database. Support disini artinya jumlah transaksi dalam database yang mengandung kedua item dalam kandidat 2-itemset. Setelah support dari semua kandidat 2-itemset didapatkan, kandidat 2-itemset yang memenuhi syarat minimum support dapat ditetapkan sebagai 2-itemset yang juga merupakan pola frekuensi tinggi dengan panjang 2.

Untuk selanjutnya pada iterasi ke-k dapat dibagi lagi menjadi beberapa bagian:

1. Pembentukan kandidat itemset

Kandidat k-itemset dibentuk dari kombinasi (k-1)-itemset yang didapat dari iterasi sebelumnya. Satu ciri dari algoritma Apriori adalah adanya pemangkasan kandidat k-itemset yang subset-nya yang berisi k-1 item tidak termasuk dalam pola frekuensi tinggi dengan panjang k-1.

2. Penghitungan support dari tiap kandidat k-itemset

Support dari tiap kandidat k-itemset didapat dengan men-scan database untuk menghitung jumlah transaksi yang memuat semua item di

dalam kandidat k-itemset tersebut. Ini adalah juga ciri dari algoritma Apriori dimana diperlukan penghitungan dengan scan seluruh database sebanyak k-itemset terpanjang.

### 3. Tetapkan pola frekuensi tinggi

Pola frekuensi tinggi yang memuat k item atau k-itemset ditetapkan dari kandidat k-itemset yang support-nya lebih besar dari minimum support. Bila tidak didapat pola frekuensi tinggi baru maka seluruh proses dihentikan. Bila tidak, maka k ditambah satu dan kembali ke bagian 1 (Noranita, B, dkk, 2010)

## **E. Penelitian Terkait**

**“Distribusi Basis Data Kependudukan Untuk Optimalisasi Akses Data: Suatu Kajian Pustaka”**, Edhy Sutanta, Ahmad Ashari, Mahasiswa Program S3 Ilmu Komputer, FMIPA, UGM. Penelitian ini mengungkap sebuah gagasan tentang skenario penerapan distribusi *database* kependudukan WNI yang meliputi skenario *database* master kependudukan nasional tunggal dan terpusat, desain logik pengembangan *database* kependudukan, desain fisik pengembangan *database* kependudukan, distribusi *database* kependudukan, replikasi dan fragmentasi *database*, serta aspek transparansi dalam DDBMS.

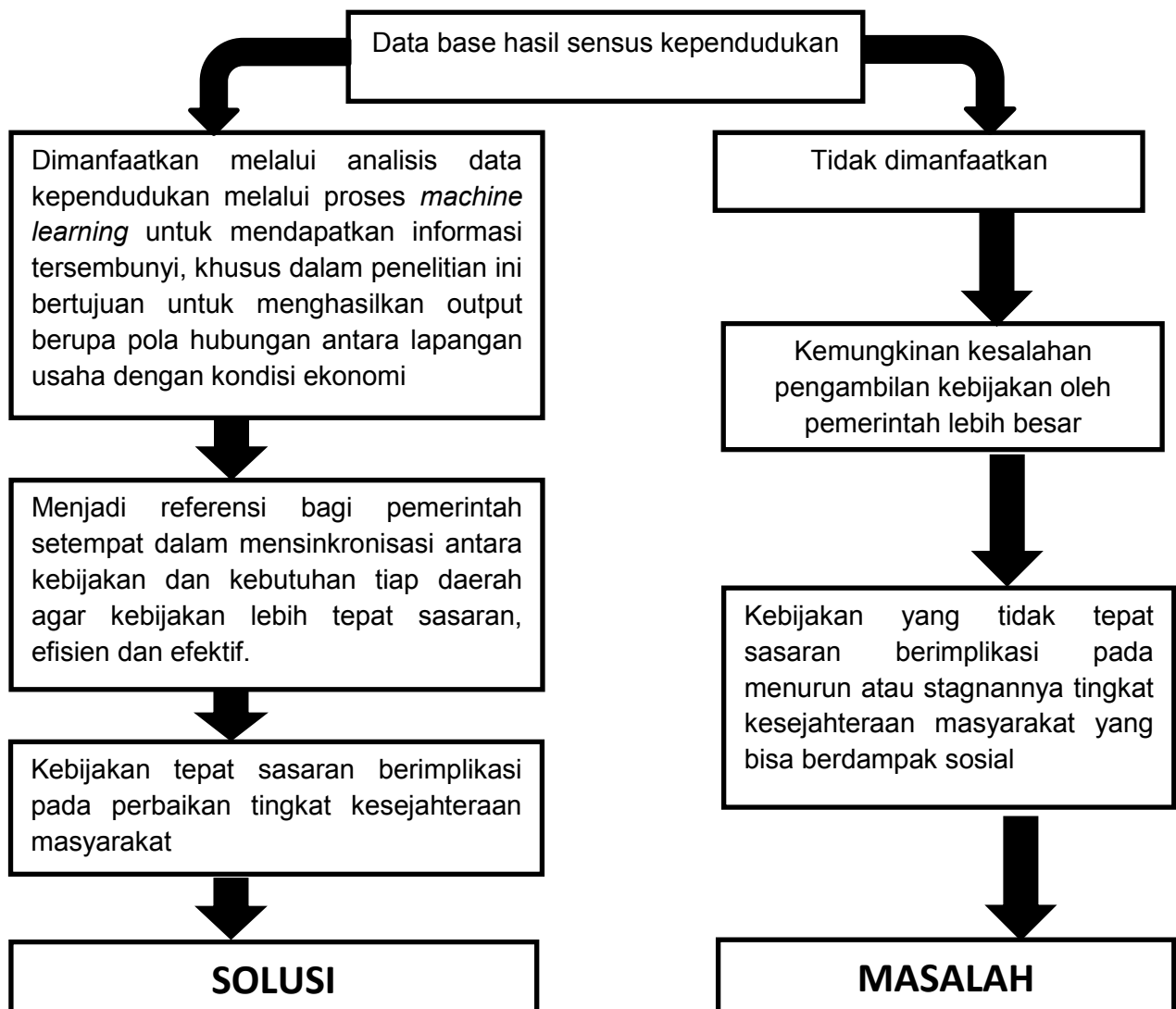
**“Rancang bangun *data warehouse* (studi kasus) di PDAM Kabupaten Cilacap”**, Isa Bahroni Dosen Tetap Program Studi Teknik Informatika, Politeknik Cilacap. Penelitian ini membuat sebuah rancang bangun *data warehouse*. sebagai upaya untuk membantu para pengambil kebijakan dalam SPK (Sistem Pengambilan Keputusan) di lingkup PDAM Kab. Cilacap agar kebijakan yang diambil tepat sasaran.

**“Model Integrasi Ragam Sistem Informasi Egov dengan Memanfaatkan Database Kependudukan Nasional”**, Retantyo Wardoyo, Khabib Mustofa, Mahasiswa Program S3 Ilmu Komputer, FMIPA, UGM. Makalah ini merupakan hasil kajian pustaka yang mengungkap problem dan solusi dalam integrasi database antar CBIS

(Computer Based Information Systems) dengan memanfaatkan data induk berupa data kependudukan nasional untuk beragam aplikasi dalam eGov dengan memanfaatkan nomor identitas tunggal (SIN).

## F. Kerangka Konseptual

Alur kerangka konseptual dalam usulan penelitian ini digambarkan sebagai berikut:



Gambar 1 . Bagan Kerangka Konseptual