

Pengembangan Sistem Informasi Data KB dan Analisis Pola Pemilihan Metode Kontrasepsi Menggunakan Algoritma *Sql-Based Fp-Growth*

Yumina Jumiati^{*1)} dan Nurdin Bahtiar²⁾

^{1,2}Jurusan Ilmu komputer/Informatika, Fakultas Sains dan Matematika, Universitas Diponegoro,
Jl. Prof. H. Soedharto, SH, Tembalang, Semarang, 50275, Indonesia

Abstract

One of Unit Pelaksana Teknis Pemberdayaan Masyarakat, Perempuan, dan Keluarga Berencana (UPT BAPERMASPER dan KB) Region XV District of Mijen's duty is to lead, plan, do, evaluate, and report the data management in service, management, and control of family planning programs and the empowerment of women in the district. The inadequate number offield officers (PLKB) often become spectacle in collecting and reporting the data KB. Family Planning Data Information System (SIDAK) is developed in order to support the mentioned task to make it easier to report, store, and manage the data. Additionally, SIDAK is equipped with a data analysis feature with association mining technique using SQL-Based FP-Growth algorithm. This algorithm analyze the data KB and create a relation (association rule) between the attributes used for analysis; the wife's age, the highest education achieved by the couple, the number of children, the welfare level, and the contraception method. The research analysis of 302 dataKBs resulted in maximum support value of 1.66%, maximum confidence value of 100%, and maximum lift ratio of 60.24.

Keywords: data mining, association rule, SQL-Based FP-Growth.

1. Pendahuluan

Unit Pelaksana Teknis Badan Pemberdayaan Masyarakat, Perempuan, dan Keluarga Berencana (UPT BAPERMASPER dan KB) Wilayah XV Kecamatan Mijen Semarang Jawa Tengah, bertugas memimpin, merencanakan, melaksanakan, mengevaluasi, dan melaporkan pengelolaan, sebagian fungsi Badan di bidang pelayanan, pengelolaan, dan pengendali program KB dan pemberdayaan perempuan di wilayah kecamatan tersebut. Dalam melaksanakan tugasnya, UPT tersebut dibantu oleh Petugas Penyuluh Keluarga Berencana Desa (PPKBD) pada setiap kelurahan yang melakukan pendataan keluarga dan melaporkan ke UPT secara manual. Data hasil pendataan kemudian dianalisis dan dievaluasi untuk menentukan kebijakan yang tepat pada pelaksanaan program KB yang akan dilaksanakan. Minimnya jumlah petugas lapangan sering menjadi kendala dalam hal pengumpulan dan pelaporan data.

Sistem pengelolaan berbasis komputer merupakan suatu hal yang penting dalam lembaga ini, kendala yang muncul merupakan tantangan tersendiri bagi pengelola agar menyediakan layanan sistem informasi yang akurat dan efisien. Suatu informasi dikatakan bernilai bila manfaatnya lebih efektif dibandingkan dengan biaya mendapatkannya (Yakub, 2012). Karenanya, perlu diupayakan pembangunan sistem informasi yang dapat memberikan manfaat lebih bagi penggunaannya, misalnyadengan menambahkan teknik data mining untuk membantu dalam pembuatan keputusan.

Berdasarkan analisis yang dilakukan oleh Radita Kusumaningrum, beberapa faktor yang mempengaruhi pemilihan metode kontrasepsi oleh peserta KB di antaranya adalah umur istri, jumlah anak, kesejahteraan, kepemilikan JAMKESMAS, pendidikan, pengetahuan, dukungan

*Correspondance : yuminajumiati@gmail.com

suami / istri, serta agama (Kusumaningrum, 2009). Tidak ada metode kontrasepsi yang aman dan efektif bagi semua klien karena masing-masing mempunyai kesesuaian dan kecocokan individual. Secara umum persyaratan metode kontrasepsi ideal adalah: aman, berdaya guna, dapat diterima, harga terjangkau, serta mengembalikan kesuburan bila tidak lagi digunakan. (Nasution, 2011).

Berdasarkan permasalahan di atas, makalah ini akan membahas tentang bagaimana mengembangkan teknik data mining asosiasi untuk membantu pengguna alat KB dalam membuat keputusan memilih metode kontrasepsi yang sesuai dengan karakteristik mereka. Dalam hal ini metode yang digunakan adalah algoritma SQL-Based FP-Growth.

2. Metode Penelitian

1. Data Mining

Menurut Jiawei Han dan Micheline Kamber (2006), *data mining* adalah kegiatan menemukan pola yang menarik dari data dalam jumlah yang sangat besar. Sumber data dapat berupa basis data, gudang data (*data warehouse*), atau penyimpanan informasi lainnya. *Data mining* berkaitan dengan bidang ilmu-ilmu lain, seperti *database system*, *data warehousing*, *statistics*, *machine learning*, *information retrieval*, dan komputasi tingkat tinggi (Han & Kamber, 2006).

Menurut Budi Santosa (2007), *data mining (Knowledge Discovery in Database /KDD)* adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Santosa B. , 2007)

Berdasarkan beberapa definisi di atas, maka dapat disimpulkan bahwa *data mining* adalah serangkaian proses penggalian sekumpulan besar data untuk menemukan aturan, pola atau hubungan menarik di dalamnya sebagai suatu pengetahuan lebih yang bermanfaat.

Tahapan *data mining* menurut Han dan Kamber (2006) adalah sebagai berikut:

- a. *Data Cleaning*
- b. *Data Integration*
- c. *Data Selection*
- d. *Data Transformation*
- e. *Data Mining*
- f. *Pattern Evaluation*
- g. *Knowledge Presentation*

2. Association Rule Mining (ARM)

Menurut Kusrini dan Emha Taufik Luthfi (2009), ARM adalah teknik *data mining* untuk menemukan aturan asosiasi antara suatu kombinasi *item*. *Interestingness measure* yang dapat digunakan dalam *data mining* adalah *support* dan *confidence*. *Support* adalah suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *itemset* dari keseluruhan transaksi. *Confidence* adalah suatu ukuran yang menunjukkan hubungan antar dua *item* secara *conditional*. Metodologi dasar analisis asosiasi terbagi menjadi dua tahap, yaitu tahap analisis pola frekuensi tinggi dan tahap pembentukan aturan asosiasi (Kusrini & Luthfi, 2009).

a. Analisis Pola Frekuensi Tinggi

Tahapan mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam basis data. Nilai *support* sebuah item A diperoleh dengan rumus (1).

$$Support(A) = \frac{\sum \text{transaksi mengandung A}}{\sum \text{transaksi dalam database}} \quad (1)$$

b. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, kemudian dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan $A \rightarrow B$ diperoleh dari rumus (2).

$$Confidence = \frac{\sum \text{transaksi mengandung A dan B}}{\sum \text{transaksi mengandung A}} \quad (2)$$

Selain menggunakan *confidence* untuk mengukur tingkat kebenaran, dikenal juga *lift ratio* untuk melihat kuat tidaknya aturan asosiasi dengan membandingkan dengan nilai *benchmark confidence*. Dimana diasumsikan kejadian item dari *consequent* dalam suatu transaksi adalah *independent* dengan kejadian dari *antecedent* dari suatu asosiasi (Santosa B., 2007)

$$Benchmark\ confidence = \frac{\sum \text{transaksi dengan item dalam consequent}}{\sum \text{transaksi dalam database}} \quad (3)$$

Lift ratio adalah perbandingan antara *confidence* untuk suatu aturan dibagi dengan *benchmark confidence*, dimana diasumsikan *consequent* dan *antecedent* saling *independent*. Nilai *lift ratio* lebih besar dari 1 menunjukkan adanya manfaat dari aturan tersebut. Lebih tinggi nilai *lift ratio*, lebih besar kekuatan asosiasi (Santosa B., 2007).

$$Lift\ Ratio = \frac{confidence}{benchmark\ confidence} \quad (4)$$

3. SQL-Based FP-Growth

Pembangunan *FP-Tree* pada algoritma *FP-Growth* pada suatu basis data yang besar memerlukan kapasitas memori yang sangat besar juga. Penelitian yang dilakukan oleh Xuequn Shang, Sattler Kai-Uwe, dan Geist Ingolf, menemukan suatu solusi permasalahan tersebut dengan memanfaatkan sistem manajemen basis data relasional (*Relational Database Management System/RDBMS*).

Salah satu alternatif adalah dengan menyimpan *FP-Tree* dalam struktur tabel (tabel FP). Terdapat dua pendekatan dalam pembangunan tabel FP, yaitu pendekatan pola frekuensi (*Frequent Pattern/FP*) dan pendekatan dengan perluasan pola frekuensi (*Expand Frequent Pattern/EFP*). Pendekatan FP memeriksa tiap item frekuensi satu per satu, apakah sesuai atau tidak untuk dimasukkan ke dalam tabel FP. Pendekatan EFP membangun sebuah tabel *FP-Tree* temporer (tabel EFP), kemudian membangun tabel FP dari tabel EFP (Shang, 2005).

Tahapan algoritma *SQL-Based FP-Growth* dijelaskan sebagai berikut (Schuiling, 2014):

a. Pembuatan Tabel FP

Tabel FP merupakan bentuk penerapan *FP-Tree* dalam bentuk tabel. Pembuatan tabel FP yang dijelaskan pada tahapan ini mengikuti pendekatan FP, dilakukan dalam tiga tahap sebagai berikut:

- (1.) *Item counting* (penghitungan item): memindai tabel transaksi satu kali dan hitung frekuensi kemunculan item (*support count*).

```
CREATE TEMPORARY TABLE TOTAL_COUNTS (Item str, Support int);
INSERT INTO TOTAL_COUNTS
SELECT Item, COUNT(*) AS Support FROM T
GROUP BY Item
HAVING COUNT(*) >= $minsupport
ORDER BY Item, COUNT(*) DESC;
```

- (2.) *Transaction ordering* (pengurutan transaksi): mengurutkan item-item frekuen dari tabel transaksi dalam urutan secara descending berdasarkan frekuensi kemunculan item tersebut.

```
CREATE TEMPORARY TABLE SORTED_T (TID int, Item str);
INSERT INTO SORTED_T
SELECT T1.TID, T1.Item FROM T AS T1, (SELECT Item, Support FROM
TOTAL_COUNTS) AS T2
WHERE T1.Item = T2.Item ORDER BY T1.TID, T2.Support DESC, T1.Item;
```

- (3.) *FP-Table construction* (pembuatan tabel FP): membuat tabel FP yang memuat semua transaksi dengan item-item yang memenuhi batasan frekuensi kemunculan item (*minimum support count*).

```
CREATE TABLE FP(PATH str, COUNT int);
INSERT INTO FP
SELECT PATH, COUNT(*)
FROM (SELECT ':'||CONCAT(Item), ':') AS PATH FROM SORTED_T
GROUP BY TID) AS GROUPED
GROUP BY PATH;
```

b. Menemukan Pola Frekuen dari Tabel FP

Setelah tabel FP selesai dibuat, langkah selanjutnya adalah membuat tabel *conditional FP* untuk setiap item sebelum menemukan semua pola frekuen dari tabel FP. Pembuatan tabel *conditional FP* dapat dikerjakan dengan mengeksekusi perintah SQL seperti berikut:

```
SELECT SUBSTRING(PATH, 0, POSITION(':'||Item||':' in PATH)) AS Pattern,
SUM(COUNT) FROM FP
WHERE PATH LIKE '%:'||Item||':%' AND PATH NOT LIKE ':'||Item||':%'
GROUP BY Pattern
```

3. Hasil dan Pembahasan

Pengujian fungsi *data mining* dengan masukan (*input*) tahun data 2013 (tabel datakb2013) dan variasi nilai *lift ratio*. Cuplikan data pada tabel awal (datakb2013) dan tabel transaksi hasil transformasi dan integrasi dari tabel datakb2013 ditunjukkan pada tabel 1.

Tabel 1. Tabel data KB 2013

id_kel	id	suami	umur_s	pend_s	pek_s	istri	umur_i	pend_i	pek_i	jml_anak	tk	mk
1001			40	B7	E41		43	B7	D87	3	D2	0
1001			33	B1	E19		27	B2	D01	2	D0	1
1001			25	B1	E14		23	B4	D14	1	D1	1
1001			35	B1	E14		38	B4	D14	2	D1	1
1001			28	B1	E14		27	B2	D01	2	D0	1
1001			35	B7	E14		34	B7	D14	1	D3	1
1001			32	B7	E14		31	B4	D14	2	D1	3
1001			32	B7	E14		36	B7	D14	2	D2	1
1001			57	B7	E14		51	B7	D04	4	D4	4
1001			31	B4	E14		36	B6	D14	2	D2	4

Data pada tabel datakb2013 diseleksi dan ditransformasikan ke dalam format untuk analisis, kemudian diintegrasikan untuk mendapatkan tabel transaksi. *Identifier* (ID) transaksi adalah nomor kartu keluarga Pasangan Usia Subur (PUS) peserta KB.

Tabel 2. Tabel transaksi (kanan) hasil integrasi tabel hasil transformasi dari tabel datakb2013 (kiri)

Id	umur_i	pend	jml_anak	tk	mk	id	antecedent	consequent	path
	A2	B4	C3	D2	G0		A2:B4:C3:D2	G0	A2:B4:C3:D2 => G0
	A2	B7	C0	D1	G0		A2:B7:C0:D1	G0	A2:B7:C0:D1 => G0
	A2	B4	C2	D1	G1		A2:B4:C2:D1	G1	A2:B4:C2:D1 => G1
	A2	B6	C2	D4	G4		A2:B6:C2:D4	G4	A2:B6:C2:D4 => G4
	A2	B7	C2	D2	G2		A2:B7:C2:D2	G2	A2:B7:C2:D2 => G2

Id	umur	i	pend	jml_anak	tk	mk	id	antecedent	consequent	path
A2	B6	C1	D2	G1			A2:B6:C1:D2	G1	A2:B6:C1:D2 => G1	
A2	B7	C1	D2	G0			A2:B7:C1:D2	G0	A2:B7:C1:D2 => G0	
A1	B3	C2	D0	G3			A1:B3:C2:D0	G3	A1:B3:C2:D0 => G3	
A2	B4	C1	D1	G3			A2:B4:C1:D1	G3	A2:B4:C1:D1 => G3	
A2	B7	C2	D2	G2			A2:B7:C2:D2	G2	A2:B7:C2:D2 => G2	

Nilai *minimum lift ratio* pada hasil pengujian 302 data PUS pada analisis *data mining* SIDA adalah 0.57 sedangkan nilai *maksimum lift ratio* adalah 60.24. Variasi nilai *lift ratio* dalam pengujian ini adalah 1, 10, dan 20. Masukan *minimum lift ratio* senilai 1, menghasilkan 199 aturan asosiasi. Masukan *minimum lift ratio* senilai 10, menghasilkan 14 aturan asosiasi. Sedangkan dengan masukan *minimum lift ratio* senilai 20, menghasilkan 3 aturan asosiasi. Berikut akan dijelaskan perhitungan nilai *support*, *confidence*, dan *lift ratio* untuk rule ke-6 :

Jumlah transaksi tabel transaksi = 302
Pola (*Path*) = A3:B1:C5:D2 => G7
Antecedent = A3:B1:C5:D2
Consequent = G7
Frekuensi kemunculan *path* = 4
Frekuensi kemunculan *antecedent* = 8
Frekuensi kemunculan *consequent* = 17

$$\text{Support (Path)} = \frac{\sum \text{transaksi mengandung Path}}{\sum \text{transaksi dalam tabel transaksi}} \times 100\% = \frac{4}{302} \times 100\% = 1.32\%$$

$$\begin{aligned} \text{Confidence(Path)} &= \frac{\sum \text{transaksi mengandung Path}}{\sum \text{transaksi mengandung Antecedent}} \times 100\% \\ &= \frac{8}{13} \times 100\% = 0.6667 \times 100\% = 66.67\% \end{aligned}$$

$$\text{Benchmark confidence} = \frac{\sum \text{transaksi dengan item dalam Consequent}}{\sum \text{transaksi dalam tabel transaksi}} = \frac{17}{302} = 0.056291$$

$$\text{Lift Ratio} = \frac{\text{confidence}}{\text{benchmark confidence}} = \frac{0.6667}{0.05629} = 11.84$$

Aturan asosiasi yang lain dihitung nilai *support*, *confidence*, dan *lift rationya* dengan cara yang sama menghasilkan nilai seperti ditunjukkan pada tabel 3.

Tabel 3. Hasil pengujian terhadap fungsi *data mining*

No.	Min. Lift Ratio	Rule ke-	Aturan Asosiasi	Support	Confidence	Lift Ratio
1.	10	1	A2:B1:C3:D3 => G7	0.33	100.00	17.76
		2	A2:B1:C4:D2 => G7	0.33	100.00	17.76
		3	A2:B4:C1:D3 => G2	0.33	100.00	15.90
		4	A2:B4:C3:D2 => G6	0.33	33.33	20.08
		5	A2:B7:C3:D1 => G7	0.33	100.00	17.76
		6	A3:B1:C5:D2 => G7	1.32	66.67	11.84
		7	A3:B3:C5:D3 => G7	0.33	100.00	17.76
		8	A3:B4:C3:D3 => G6	0.66	100.00	60.24
		9	A3:B6:C3:D4 => G7	0.33	100.00	17.76
		10	A3:B7:C3:D2 => G6	0.33	25.00	15.06
		11	A3:B7:C3:D4 => G7	0.33	100.00	17.76
		12	A3:B7:C4:D2 => G7	0.33	100.00	17.76
		13	A3:B7:C5:D4 => G7	0.33	100.00	17.76
		14	A3:B8:C2:D3 => G6	0.33	50.00	30.12
3.	20	1	A2:B4:C3:D2 => G6	0.33	33.33	20.08
		2	A3:B4:C3:D3 => G6	0.66	100.00	60.24
		3	A3:B8:C2:D3 => G6	0.33	50.00	30.12

Berdasarkan hasil pengujian fungsi *data mining* yang ditunjukkan pada tabel 1, didapatkan beberapa informasi dengan ukuran *interestingness* yang ditunjukkan dengan nilai *lift ratio*. Atribut yang digunakan adalah atribut umur istri (A), pendidikan tertinggi PUS (B), jumlah anak (C), tingkat kesejahteraan (D), dan metode kontrasepsi (G). Contoh analisis terhadap hasil pengujian adalah sebagai berikut:

a. *Minimum lift ratio 1*

Hubungan asosiasi yang terbentuk dijelaskan sebagai berikut:

- (1.) PUS dengan kriteria jumlah anak 0 (C0) cenderung untuk tidak menggunakan metode kontrasepsi (G0).
- (2.) PUS dengan kriteria umur istri kurang dari 20 tahun (A1) dan jumlah anak 1 cenderung untuk menggunakan metode kontrasepsi Suntik (G3).
- (3.) PUS dengan kriteria umur istri kurang dari 20 tahun (A1) dan jumlah anak 2 cenderung untuk menggunakan metode kontrasepsi Implant (G4).
- (4.) PUS dengan kriteria umur istri 20 – 39 tahun adalah kelompok umur istri dengan pilihan penggunaan metode kontrasepsi yang paling bervariasi.
- (5.) PUS dengan kriteria umur istri 20 – 39 tahun dengan jumlah anak 1 (C1) lebih cenderung menggunakan metode kontrasepsi jangka pendek (non-MKJP), yaitu Pill (G1), Kondom (G2), atau Suntik (G3).
- (6.) PUS dengan kriteria umur istri 20 – 39 tahun dengan jumlah anak 2 (C2) atau 3 (C3) lebih cenderung menggunakan non-MKJP Implant (G4) dan MKJP MOW (G5).
- (7.) PUS dengan kriteria umur istri 40 tahun atau lebih (A3) dan jumlah anak 1 (C1) cenderung menggunakan tidak menggunakan program KB (G0).
- (8.) PUS dengan kriteria umur istri 40 tahun atau lebih (A3), dan jumlah anak 2 (C2), cenderung menggunakan non-MKJP, yaitu Pill (G1), Suntik (G3), dan Implant (G4).
- (9.) PUS dengan kriteria umur istri 40 tahun atau lebih (A3) dengan jumlah anak 3 atau lebih, cenderung menggunakan MKJP, yaitu MOW (G5), MOP (G6), dan IUD (G7).

b. *Minimum lift ratio 10*

Hubungan asosiasi yang terbentuk dijelaskan sebagai berikut:

- (1.) Metode kontrasepsi MOP cenderung digunakan oleh PUS dengan kriteria umur istri 40 tahun atau lebih (A3) dengan jumlah anak 2 atau lebih, tingkat pendidikan minimum SLTA/Sederajat (B4), dan tingkat kesejahteraan keluarga minimum tingkat II (D2).
- (2.) PUS dengan kriteria umur istri 20 – 39 tahun (A2) atau lebih (A3) dan jumlah anak 3 (C3) atau lebih, cenderung menggunakan metode kontrasepsi MOW (G6) atau IUD (G7).

c. *Minimum lift ratio 20*

Hubungan asosiasi yang terbentuk menunjukkan bahwa metode kontrasepsi MOP cenderung digunakan oleh PUS dengan kriteria umur istri 40 tahun atau lebih (A3) dengan jumlah anak 2 atau lebih, tingkat pendidikan minimum SLTA/Sederajat (B4), dan tingkat kesejahteraan keluarga minimum tingkat II (D2).

Secara umum, jika mengacu pada perhitungan lift ratio tertinggi (yaitu 20), menggambarkan bahwa metode kontrasepsi yang populer adalah MOP dimana metode ini digunakan oleh pasangan usia subur dengan karakteristik sebagai berikut:

- a) umur istri 40 tahun atau lebih
- b) jumlah anak 2 atau lebih,
- c) tingkat pendidikan minimum SLTA/Sederajat,
- d) dan tingkat kesejahteraan keluarga minimum tingkat II.

4. Simpulan

SIDAK dapat digunakan untuk membantu mempermudah dalam pelaporan, penyimpanan, dan pengelolaan data KB serta mendukung proses analisis penggunaan metode kontrasepsi menggunakan teknik data mining asosiasi menggunakan algoritma *SQL-Based FP-Growth*.

Berdasarkan analisis hasil pengujian fungsi data mining dengan menggunakan algoritma *SQL-Based FP-Growth* terhadap sampel data KB tahun 2013, menghasilkan nilai *support* tertinggi 1.66%, nilai *confidence* tertinggi 100%, dan nilai *lift ratio* tertinggi 60.24. Semakin tinggi nilai parameter yang digunakan dalam pemrosesan dengan algoritma *SQL-Based FP-Growth*, semakin sedikit aturan yang dihasilkan, semakin kuat hubungan asosiasi antar atribut dalam aturan yang terbentuk.

Saran untuk pengembangan atau penelitian lebih lanjut yaitu dapat dilakukan pengujian terhadap semua atribut yang mempengaruhi pemilihan metode kontrasepsi dengan algoritma *SQL-Based FP-Growth* per atribut dan kombinasi gabungannya.

Daftar Pustaka

- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. San Fransisco: Morgan Kauffmann Publisher.
- Kadir, A. (2003). *Pengenalan Sistem Informasi*. Yogyakarta: Andi Offset.
- Kusrini, & Luthfi, E. T. (2009). *Algoritma Data Mining*. Yogyakarta: ANDI.
- Kusumaningrum, R. (-, -). *Faktor-faktor yang Mempengaruhi Pemilihan Jenis Kontrasepsi yang Digunakan pada Pasangan Usia Subur*. Dipetik May 13, 2014, dari [eprints.undip.ac.id: http://eprints.undip.ac.id/19194/1/Radita_Kusumaningrum.pdf](http://eprints.undip.ac.id/http://eprints.undip.ac.id/19194/1/Radita_Kusumaningrum.pdf)
- Nasution, S. L. (2011, - -). *faktor-faktor yang Mempengaruhi Penggunaan MKJP di Enam Wilayah Indonesia*. Dipetik May 12, 2014, dari [www.bkkbn.go.id: http://www.bkkbn.go.id/litbang/pusna/Hasil%20Penelitian/FAKTOR-FAKTOR%20YANG%20MEMPENGARUHI%20PENGUNAAN%20MKJP%20DI%20ENAM%20WILAYAH%20INDONESIA.pdf](http://www.bkkbn.go.id/http://www.bkkbn.go.id/litbang/pusna/Hasil%20Penelitian/FAKTOR-FAKTOR%20YANG%20MEMPENGARUHI%20PENGUNAAN%20MKJP%20DI%20ENAM%20WILAYAH%20INDONESIA.pdf)
- Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis* (1 ed.). Yogyakarta: Graha Ilmu.
- Schuilng, N. H. (2014, Agustus). *Scalable Frequent Pattern Mining using Relational Databases*. Dipetik September 20, 2015, dari dspace.library.uu.nl/bitstream/handle/1874/298560/Thesis.pdf?sequence=2
- Shang, X. (2005, February 14). *SQL Based Frequent Pattern Mining*. Dipetik Mei 15, 2015, dari diglib.uni-magdeburg.de/Dissertationen/2005/xueshang.pdf
- Yakub. (2012). *Pengantar Sistem Informasi*. Yogyakarta: Graha Ilmu.