IMPLEMENTASI ALGORITMA FP – GROWTH DAN ASSOCIATION RULE PADA PEMILIHAN BIDANG MINAT PROGRAM STUDI TEKNIK INFORMATIKA

Nama Mahasiswa	**:Tri Rahayu Wulandari**
NIM	**:1210651078**
Jurusan	**:Teknik Informatika**
Dosen Pembimbing 1	**:Lutfi Ali Muharrom, S. Si, M.Si**
Dosen Pembimbing 2	**: Daryanto S.Kom, M.Kom**

Abstrak

Pada saat ini, mahasiswa di Jurusan Teknik Informatika Universitas Muhammadiyah Jember belum memiliki sistem untuk membantu mahasiswa menentukan bidang minat. Pemilihan bidang minat ini sangat diperlukan karena sangat berkaitan dengan pengambilan Tugas Akhir dari mahasiswa itu sendiri. Sistem ini akan dibangun berdasarkan hasil analisa terhadap sejumlah data – data sebelumnya. Algoritma FP – Growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sekumpulan data. Algoritma *FP – Growth generate candidate* tidak dilakukan karena *FP - Growth* menggunakan konsep pembangunan *tree* dalam pencarian *frequent itemset.*

Dari paparan di atas diharapkan dengan adanya model sistem pendukung keputusan yang dimilki oleh Jurusan dapat membantu untuk menentukan pola nilai histori matakuliah dalam penentuan matakuliah pilihan yang akan ditawarkan dengan menerapkan teknik *data mining* dengan menggunakan *Association Rule* dan Algoritma *FP – Growth.*

Kata Kunci : Algoritma *Fp – growth, association rule mining, Fp – Tree*

Abstrak

At this time, a student in the Department of Information Engineering University of Muhammadiyah Jember not have a system to help students determine their areas of interest. Selection of areas of interest is very necessary because it is related to decision-Final of the students themselves. The system will be built based on the analysis of some data - previous data. Algorithm FP - Growth is one of the alternative algorithm that can be used to specify the data set that appears most frequently (frequent itemset) in a set of data. Algorithm FP - Growth generate candidate was not done because of FP - Growth uses the concept of development in the search for frequent itemset tree.

From the above explanation is expected by the model of decision support system that is owned by the Department can help to determine the pattern of historical value in determining the course that will be offered optional subjects by applying data mining techniques using Association Rule and algorithm FP - Growth.

Keywords: Algorithm *Fp - growth, association rule mining, fp - Tree,*

1.1 Latar Belakang

Algoritma FP – Growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sekumpulan data. Algoritma *FP – Growth generate candidate* tidak dilakukan karena *FP - Growth* menggunakan konsep pembangunan *tree* dalam pencarian *frequent itemset.*

Dalam model ini, hasil rekomendasi penetuan matakuliah pilihannya didapat dengan menggunakan metode *data mining* yaitu *Association Rule* dengan pendekatan algoritma *FP – Growth*. Untuk mendapatkan pola ini digunakan pendekatan *FP – Growth* untuk mempercepat mendapatkan *ekstrak frequent itemset*. Pada akhirnya model sistem pendukung keputusan ini dapat digunakan untuk mendapatkan rekomendasi penentuan bidang minat berdasarkan data acuan yang sesuai dengan kebijakan yang dikeluarkan pihak jurusan.

Pada saat ini, mahasiswa di Jurusan Teknik Informatika Universitas Muhammadiyah Jember belum memiliki sistem untuk membantu mahasiswa menentukan bidang minat. Pemilihan bidang minat ini sangat diperlukan karena sangat berkaitan dengan pengambilan Tugas Akhir dari mahasiswa itu sendiri. Sistem ini akan dibangun berdasarkan hasil analisa terhadap seumlah data – data sebelumnya

1.2 Tujuan

Menghasilkan sebuah model sistem pendukung keputusan dalam penentuan matakuliah pilihan di Jurusan Teknik Informatika dengan data acuan dari kebijakan yang dinamis, dan membuat sistem pencarian pola asosiasi konsentrasi dimana akan ditemukan aturan - aturan (*rule*).

2.1 Penggalian Data

Penggalian data (data mining) adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam basis data. Penggalian data adalah tentang memecahkan suatu masalah dengan menganalisis data yang telah ada, dan sudah di definiskan sebagai proses pola dalam data, dimana pola yang didapat harus memiliki beberapa keuntungang (Written dan Frank,2005)

2.2 Association Rule

Association rule merupakan suatu proses pada *data mining* untuk menentukan semua aturan *assosiatif* yang memenuhi syarat minimum untuk *support* (*minsup)* dan *confidence* (*mincof*) pada sebuah database. Menurut Agrawal dan Srikant dalam (Ridowati, 2003), masalah dalam pencarian seluruh *association rule* dapat dibagi menjadi dua, yaitu bagaimana menemukan seluruh barang – barang (*itemset*) yang memiliki nilai penunjang (*support)* di atas nilai penunjang terkecil yang telah ditentukan. Nilai penunjang untuk sebuah barang adalah jumalah transaksi yang terdapat dalam *itemset*, *itemset* dengan nilai terkecil disebut *large Itemset.* Masalah yang kedua adalah bagaimana menggunakan large itemsetbuntuk menghasilkan aturan – aturan yang diperlukan.

Aturan yang hendak dihasilkan adalah dalam bentuk a→ (I–a), yaitu perbandingan antara *support* (I) dengan *support* (a) haruslah lebih besar dari *mincof* (*minimum confidence*). Misakan [I] = $\{I_1, I_2, \ldots I_m\}$ adalah himpunan literal yang disebut barang. Himpunan *item – item* disebut dengan *Itemset.* D adalah himpunan transaksi, dimana setiap transaksi T itemset sehingga TE, I.

3.1 Algoritma FP - Growth

Algoritma Fp – Growth menggunakan konsep pembangunan tree yang biasa disebut FP – *Tree*, dalam pencarian *Frequent Itemset* bukan menggunakan generate candidate seperti yang dilakukan pada algoritma apriori. Dengan menggunakan konsep tersebut proses algoritma fp – growth menjadi lebih cepat dari pada algoritma apriori. Metode yang digunakan FP – growth ada tiga tahapan utama, yakni : 1 tahapan pembangkitan *Condational Pattern Base*, 2 tahapan pembangkitan *Condational FP – Tree* dan ke 3 Pencarian *Frequent Itemset.*

Analisis dari data keranjang dapat digunakan untuk mendapatkan aturan asosiasi dari kombinasi IPK mahasiswa per KBK dengan bidang minat yang diambil oleh mahasiswa seperti contoh tabel dibawah ini :

Tabel 1 Data mahasiswa

ID	KBJ	SBC	RPL	BM
0910651010	3,20	3,39	3,21	RPL
0910651011	2,79	3,05	3,00	RPL
0910651012	3,16	3,16	3,42	KBJ
0910651025	2,93	3,14	3,35	KBJ
0910651038	2,93	3,01	3,07	SBC
0910651040	3,00	3,13	3,28	SBC
0910651041	3,13	3,07	3,50	KBJ
0910651045	2,90	3,13	3,50	RPL
0910651050	2,93	3,16	3,42	SBC
0910651104	2,95	3,58	2,85	RPL

Untuk memudahkan analisis, maka IPK masing – masing KBK dibuat menjadi 3 interval :

IPK menjadi 3 interval :

IPK KBK KBJ
(minimal – 2,98) = KBJ 1
(3,00 – 3,19) = KBJ 2
(maximal – 3,20) = KBJ 3
IPK KBK SBC
(minimal - 2,98) = SBC 1
(3,00 – 3,38) = SBC 2
(maximal – 3,39) = SBC 3
IPK KBK RPL
(minimal – 3,00) = RPL 1
(3,01 – 3,41) = RPL 2
(maximal – 3,42) = RPL 3

Tabel 2 Data transaksi Mentah

ID	KBJ	SBC	RPL	BM
1010	KBJ-2	SBC-3	RPL-2	RPL
1011	KBJ-1	SBC-2	RPL-2	RPL
1012	KBJ-2	SBC-2	RPL-3	KBJ
1025	KBJ-1	SBC-2	RPL-2	KBJ
1038	KBJ-1	SBC-2	RPL-2	SBC
1040	KBJ-2	SBC-2	RPL-2	SBC
1041	KBJ-2	SBC-2	RPL-2	KBJ
1045	KBJ-1	SBC-2	RPL-2	RPL
1050	KBJ-1	SBC-2	RPL-3	SBC
1104	KBJ-1	SBC-2	RPL-1	RPL

Berikut adalah proses pencarian Frequent Itemset. Langkah pertama adalah menentukan *minimum support* dan *minimum confidence (Minsup)*, yaitu dengan *minsup* (0,3) dan *mincof* (0,9)

Tabel 3 Kemunculan Item

Itemset	*Support*
KBJ-2	4
SBC-3	1
RPL-2	7
KBJ	3
KBJ-1	6
SBC-2	9
RPL-1	1
RPL	4
KBJ-3	0

SBC-1	0
RPL-3	2
SBC	3

Setelah diperoleh *Frequent List*, kemudian ubah item sesuai frequent list dengan menghilangkan item SBC-3 , RPL-1, KBJ-3, SBC-1, RPL-3. Dan kemudian di urutkan dari yang terbesar pada setiap item transaksi.

Tabel 4 *Frequent List*

Itemset	*Support*
SBC-2	9
RPL-2	7
KBJ-1	6
KBJ-2	4
RPL	4
KBJ	3
SBC	3

Tabel 5 transaksi yang telah disesuaikan dengan *frequent list*

ID	ITEM
0910651010	RPL-2,KBJ-2,RPL
0910651011	SBC-2,RPL-2,KBJ-1,RPL
0910651012	SBC-2,KBJ-2,KBJ
0910651025	SBC-2,RPL-2,KBJ-1,KBJ
0910651038	SBC-2,RPL-2,KBJ-1,SBC
0910651040	SBC-2,RPL-2,KBJ-2,SBC
0910651041	SBC-2,RPL-2,KBJ-2,KBJ
0910651045	SBC-2,RPL-2,KBJ-1,RPL
0910651050	SBC-2,KBJ-1,SBC
0910651104	SBC-2,KBJ-1,RPL

Pembuatan FP – Tree diawali dari TID 1 sampai TID 10.

Sehingga kita menemukan daftar *Frequent Itemset* berdasarkan hubungan akhiran atau *Suffix*.

Tabel 6 *Suffix* atau Akhiran

Suffix	**Frequent Itemset**
{SBC}	{KBJ-1,SBC}, {SBC-2,SBC},{KBJ-2,SBC},{SBC-2,SBC}
{KBJ}	{KBJ-2,KBJ},{RPL-2,KBJ-2,KBJ},{SBC-2,KBJ-2,KBJ},{RPL-2,KBJ-1,KBJ},{SBC-2,KBJ-1,KBJ},{SBC-2,KBJ}
{RPL}	{KBJ-2,RPL},{RPL-2,KBJ-2,RPL},{KBJ-1,RPL},{RPL-2,KBJ-1,RPL},{SBC-2,KBJ-1,RPL},{SBC-2,RPL}
{KBJ-2}	{RPL-2,KBJ-2},{SBC-2,RPL-2,KBJ-2},{SBC-2,KBJ-2}
{KBJ-1}	{RPL-2,KBJ-1},{SBC-2,KBJ-1},{SBC-2,RPL-2,KBJ-1},{SBC-2,KBJ-1}
{RPL-2}	{SBC-2,RPL-2}
{SBC-2}	{SBC-2}

Proses Pencarian *Association Rule*

Proses pencarian association rule ini memanfaatkan minimum confidence yang telah ditentukan ,yaitu 0,9 atau 90%, dimana S→ $(I - s)$

R1 = IPK KBK RPL = (RPL-2) →IPK KBK KBJ = (KBJ-2)

Confidence = Support Count {IPK KBK RPL = (RPL – 2) ; IPK KBK KBJ (KBJ – 2) }/ *Support Count* {IPK KBK RPL = (RPL – 2)} = 4/7 = 0,6

R2 = IPK KBK RPL = (RPL – 2) → IPK KBK RPL

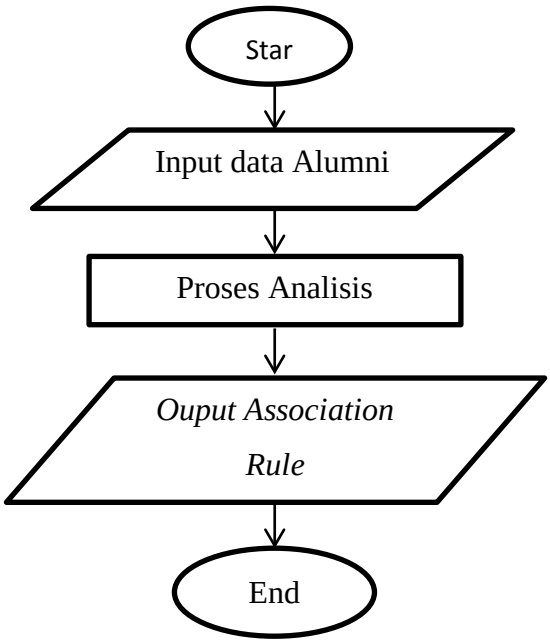
Confidence = Support Count {IPK KBK RPL = (RPL -2) ; RPL} / Support Count {IPK KBK RPL = (RPL -2) = 4/7 = 0,6

R3 = IPK KBK KBJ = (KBJ – 2) → RPL

Confidence = Support Count {IPK KBK KBJ = (KBJ – 2) ; RPL} / Support Count {IPK KBK KBJ = (KBJ -2) = 4/4 = 1

Jadi contoh rule yang terpilih adalah R3 yang ditandai dengan cetakan tebal, rule yang tidak terpilih atau ditolak dikarenakan tidak memenuhi syarat *minimum confidence.*

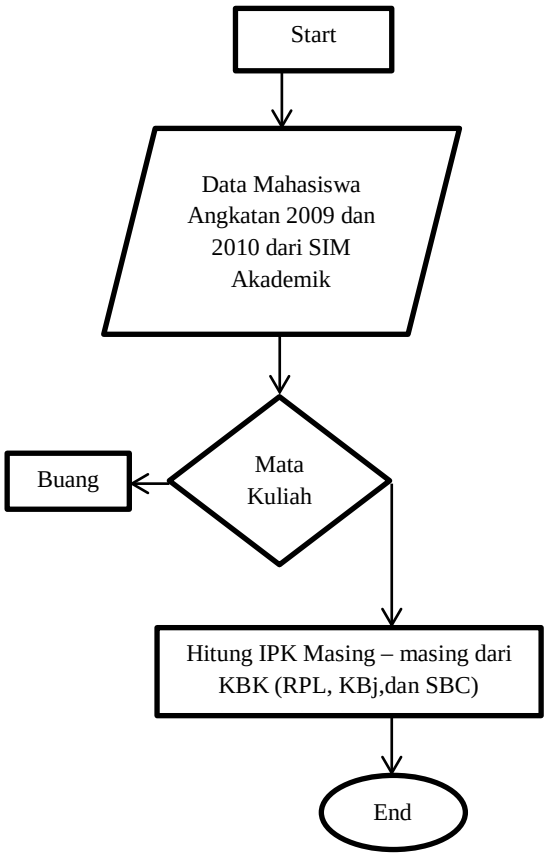
3.2 Flowchart Sistem



Sistem bidang minat ini menggunakn algoritma *FP – Growth* dalam menentukan bidang minat mahasiswa. Langkah – langkah dalam memproses dengan menggunkan algoritma *FP – Growth* ini adalah, yang pertama, mencari *rule* terlebih dahulu, yaitu dengan memsaukkan input data alumni yang telah terpilih, kemudian algoritma *FP – Growth* ini akan digunakn untuk menganalisis data. Saat memasukkan

data mahasiswa yang melalui tahap pro – processing atau dari data mentah sampai rule yang terpilih maka saat itu pula algoritma *FP – Growth* memulai melakukan pekerjaan dengan 2 tahapan, yaitu dengan melakukan pembentukan *frequent itemset* dan proses pembentukan *Association Rule*, sehingga keluarlah *Ouput Association Rule.*

3.3Perancangan Data

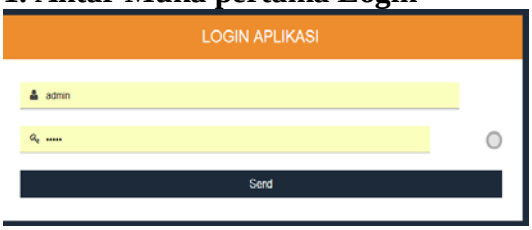


Berikut adalah penjelasan dari langkah – langkah perancangan data dia atas :

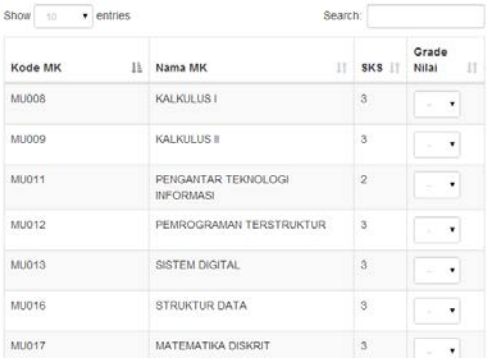
1. Pertama kita mengambil data mahasiswa angkatan tahun 2009 dan 2010 dari SIM akademik, kemudian data tersebut tidak dapat langsung digunakan karena kita harus terlebih dahulu memisah antara matakuliah umum dan matakuliah yang diwajib, karena kita hanya memerlukan matakuliah wajib saja,
2. Selanjutnya, setelah selesai memilih matakuliah yang diwajibkan kita akan memulai menghitung IPK dari KBK RPL, KBK KBJ, dan KBK dari SBC.
3. Kemudian ouput yang dihasilkan adalah ID dari mahasiswa dengan NIM 0910651010 mendapatkan IPK dari KBK RPL (3,21), IPK dari KBK KBJ (3,20) , dan IPK dari KBK SBC (3,39).

4.1 Pengujian

1. Antar Muka pertama Login



2.Antar muka Input data dengan seluruh Matakuliah yang telah ditempuh baik mata kuliah yang terdapat pada kelompok SBC,KBJ, maupun RPL yang menjadi acuan dalam penentuan penjurusan.



Show 10 entries Search:

Kode MK	Nama MK	SKS	Grade Nilai
MU008	KALKULUS I	3	-
MU009	KALKULUS II	3	-
MU011	PENGANTAR TEKNOLOGI INFORMASI	2	-
MU012	PEMROGRAMAN TERSTRUKTUR	3	-
MU013	SISTEM DIGITAL	3	-
MU016	STRUKTUR DATA	3	-
MU017	MATEMATIKA DISKRIT	3	-

3.**Proses Antarmuka Kombinasi IPK**

PROSES KOMBINASI

No	NIM	Nama Mahasiswa	KBJ	SBC	RPL	HASIL KOMBINASI IPK
1	910651001	FARIDA RAHAYU NINGTIAS	KBJ 3	SBC 3	RPL 3	SBC
2	910651002	Nur Arif Fitriadi	KBJ 2	SBC 2	RPL 3	KBJ
3	910651003	Zuhri Nurhuda	KBJ 3	SBC 2	RPL 2	KBJ
4	910651004	RESARDI ESA PRASETYO	KBJ 3	SBC 2	RPL 2	KBJ
5	910651009	Bertha Nindri Widianarta	KBJ 3	SBC 3	RPL 2	SBC

4.Proses antar muka Frekuensi itemset

Merupakan pencarian frekuensi berdasarkan minsuuport dan minconfidence yang telah diinputkan.

Frequesi

No	Itemset	Frequensi
1	KBJ 3	8
2	RPL 3	8
3	SBC 2	7
4	SBC 3	7
5	KBJ 2	6
6	SBC	5
7	RPL	4
8	RPL 1	3
9	RPL 2	3

Proses Eksekusi Itemset

Eksekusi

Nim	Nama Mahasiswa	KBJ	SBC	RPL
910651001	FARIDA RAHAYU NINGTIAS	KBJ 3	SBC 3	RPL 3
910651002	Nur Arif Fitriadi	KBJ 2	SBC 2	RPL 3
910651003	Zuhri Nurhuda	KBJ 3	SBC 2	RPL 2
910651004	RESARDI ESA PRASETYO	KBJ 3	SBC 2	RPL 2
910651009	Bertha Nindri Widianarta	KBJ 3	SBC 3	RPL 2
910651010	ENI TRISNAWATI	KBJ 3	SBC 2	RPL 3
910651011	Mohammad Hidayatullah	KBJ 3	SBC 2	RPL 3
910651012	Ihya Ulumudin	KBJ 2	SBC 2	RPL 3
910651015	ACHMAD AJI PRASETYO	KBJ 3	SBC 3	RPL 3
910651017	DENNY AGUS TRIANTO	KBJ 3	SBC 3	RPL 3

Hasil Suffix dari Frequent Itemset

Perhitungan Bidang Minat | INTERVAL | Frequent Eksekusi Itemset | Suffix Itemset

Frequents Itemset

No	alternatif	hasil kombinasi Itemset							
1	KBJ 3	KBJ 3,KBJ 3,KBJ 3	KBJ 3,KBJ 3,RPL 3	KBJ 3,KBJ 3,SBC 2	KBJ 3,KBJ 3,SBC 3	KBJ 3,KBJ 3,KBJ 2	KBJ 3,KBJ 3,SBC	KBJ 3,KBJ 3,RPL	K 3,K 3,RP
2	RPL 3	RPL 3,KBJ 3,KBJ 3	RPL 3,KBJ 3,RPL 3	RPL 3,KBJ 3,SBC 2	RPL 3,KBJ 3,SBC 3	RPL 3,KBJ 3,KBJ 2	RPL 3,KBJ 3,SBC	RPL 3,KBJ 3,RPL	R 3,K 3,RP
3	SBC 2	SBC 2,KBJ 3,KBJ 3	SBC 2,KBJ 3,RPL 3	SBC 2,KBJ 3,SBC 2	SBC 2,KBJ 3,SBC 3	SBC 2,KBJ 3,KBJ 2	SBC 2,KBJ 3,SBC	SBC 2,KBJ 3,RPL	S 2,K 3,RP
4	SBC 3	SBC 3,KBJ	SBC 3,KBJ	SBC 3,KBJ	SBC 3,KBJ	SBC 3,KBJ	SBC 3,KBJ	SBC 3,KBJ	S 3,K

Setelah melalui proses perhitungan dengan penerapan itemset (*association rule* dengan memanfaatkan *minimum confidence* yang telah di tentukan), kemudian didapat hasil akhir (yang terdapat pada menu proses eksekusi). Hasil yang dimaksud adalah sebagai berikut:

Hasil Rule yang terpilih

Hasil Eksekusi | Grafik Minat

Eksekusi

Copy | Excel | PDF | Column visibility | Search:

Nim	Nama Mahasiswa	Hasil
910651001	FARIDA RAHAYU NINGTIAS	KBJ 3
910651002	Nur Arif Fitriadi	RPL 3
910651003	Zuhri Nurhuda	KBJ 3
910651004	RESARDI ESA PRASETYO	KBJ 3
910651009	Bertha Nindri Widianarta	KBJ 3
910651010	ENI TRISNAWATI	KBJ 3
910651011	Mohammad Hidayatullah	KBJ 3

5.1 Kesimpulan

Dari analisis dan pengujian yang telah dilakukan, maka dapat disimpulkan bahwa metode algoritma FP – Growth dapat digunakan untuk membantu para mahasiswa dalam menentukan bidang minat mereka dengan melihat nilai IPK

masing- masing mahasiswa tersebut dengan menerapkan itemset *association rule* dengan memanfaatkan *minimum confidence* yang telah di tentukan, sehingga didapat hasil akhir yang terdapat pada menu proses eksekusi.

5.2 Saran

Saran – saran untuk pengembangan selanjutnya yaitu, peneliti selanjutnya dapat membandingkan metode algoritma FP – Growth dengan metode lain seperti Forward, Channing bayes, atau Fuzzy.

Daftar Pustaka

David Samuel, *Penerapan struktur FP – Tree dan Algoritma FP – Growth dalam Optimasi Penentuan Frequent Itemset*. Institut Teknologi Bandung, 2008.

Erwin, " Analisis Market Basket dengan Algoritma Apriori dan FP – Growth " , Juli, 2009

Han, J,. Pei,. J., & Yin, Yiwen. (2004) *Mining Frequent Patterns without Candidate Generation*. Simon Fraser Universit. Netherlands: Kluwer Academic.

Kusrini and Emha Taufiq Lutfhi, Algoritma Data Mining, 1st ed., Theresia Ari Prabawati, Ed. Yogyakarta: ANDI, 2009

Santoso, Budi. 2007. Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis. GrahaIlmu. Yogyakarta.