



**UNIVERSITAS INDONESIA**

**PENCARIAN POLA SEKUENSIAL  
MENGUNAKAN ALGORITMA *APRIORI ALL***



**SKRIPSI**

**FANDI  
1204000327**

**FAKULTAS ILMU KOMPUTER  
PROGRAM SARJANA ILMU KOMPUTER  
DEPOK  
JULI 2009**



**UNIVERSITAS INDONESIA**

**PENCARIAN POLA SEKUENSIAL  
MENGUNAKAN ALGORITMA *APRIORI ALL***

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Ilmu Komputer**

**FANDI**

**1204000327**

**FAKULTAS ILMU KOMPUTER  
PROGRAM SARJANA ILMU KOMPUTER  
DEPOK  
JULI 2009**

## HALAMAN PERNYATAAN ORISINALITAS

**Skripsi ini adalah hasil karya saya sendiri,  
dan semua sumber baik yang dikutip maupun dirujuk  
telah saya nyatakan dengan benar.**

Nama : Fandi

NPM : 1204000327

Tanda Tangan :

Tanggal : 21 Juli 2009



## HALAMAN PERSETUJUAN

Nama : Fandi  
NPM : 1204000327  
Judul Skripsi : Pencarian Pola Sekuensial Menggunakan Algoritma  
*AprioriAll*

Tugas Akhir ini telah diperiksa dan disetujui oleh:

Depok, 21 Juli 2009

**Yova Ruldeviyani, M.Kom**  
Pembimbing I

**Yudho Giri Sucahyo, Ph.D**  
Pembimbing II

## HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Fandi  
Program Studi : Ilmu Komputer  
Judul Skripsi : Pencarian Pola Sekuensial Menggunakan Algoritma  
*AprioriAll*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Ilmu Komputer pada Program Studi Sarjana Ilmu Komputer Fakultas Ilmu Komputer Universitas Indonesia.

### DEWAN PENGUJI

Pembimbing I : Yova Ruldeviyani, M.Kom (.....)

Pembimbing II : Yudho Giri Sucahyo, Ph.D (.....)

Penguji : Dr. Achmad Nizar Hidayanto (.....)

Penguji : Rizal Fathoni Aji, M.Kom (.....)

## KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa karena doa penulis telah dikabulkan oleh-Nya sehingga Tugas Akhir ini berhasil selesai tepat waktu sebelum tahun ajaran 2008/2009 berganti. Pada awal pengerjaan Tugas Akhir ini, penulis menemui banyak rintangan yang menyita waktu. Akibatnya penulis terpojok dalam situasi merasa kekurangan waktu untuk bisa menyelesaikan Tugas Akhir ini. Namun pada suatu malam, penulis mendapat kesempatan melihat bintang jatuh. Penulis menganggap hal tersebut sebagai pertanda dari-Nya bagi penulis untuk tidak menyerah dan tetap bersemangat melanjutkan pengerjaan Tugas Akhir ini. Berkat kehendak dan pertolongan-Nya, penulis berhasil menyelesaikan Tugas Akhir ini.

Terima kasih kepada semua pihak berikut ini yang telah memberikan bantuan yang sangat bermanfaat bagi penulis selama ini. Pihak-pihak tersebut yaitu:

- (1) Ibu Yova Ruldeviyani, M.Kom, sebagai dosen pembimbing I Tugas Akhir yang telah menyediakan waktu, tenaga dan pikiran untuk membantu penulis menyelesaikan Tugas Akhir ini.
- (2) Pak Yudho Giri Suchyo, Ph.D, sebagai dosen pembimbing II Tugas Akhir yang telah bersedia menyediakan waktu, tenaga dan pikiran untuk membantu penulis menyelesaikan Tugas Akhir ini.
- (3) Pak Dr. Achmad Nizar Hidayanto, sebagai dosen penguji Tugas Akhir yang telah bersedia menyediakan waktu, tenaga dan pikiran untuk menguji penulis menyelesaikan Tugas Akhir ini.
- (4) Pak Rizal Fahtoni Aji, M.Kom, sebagai dosen penguji Tugas Akhir yang telah bersedia menyediakan waktu, tenaga dan pikiran untuk menguji penulis menyelesaikan Tugas Akhir ini.

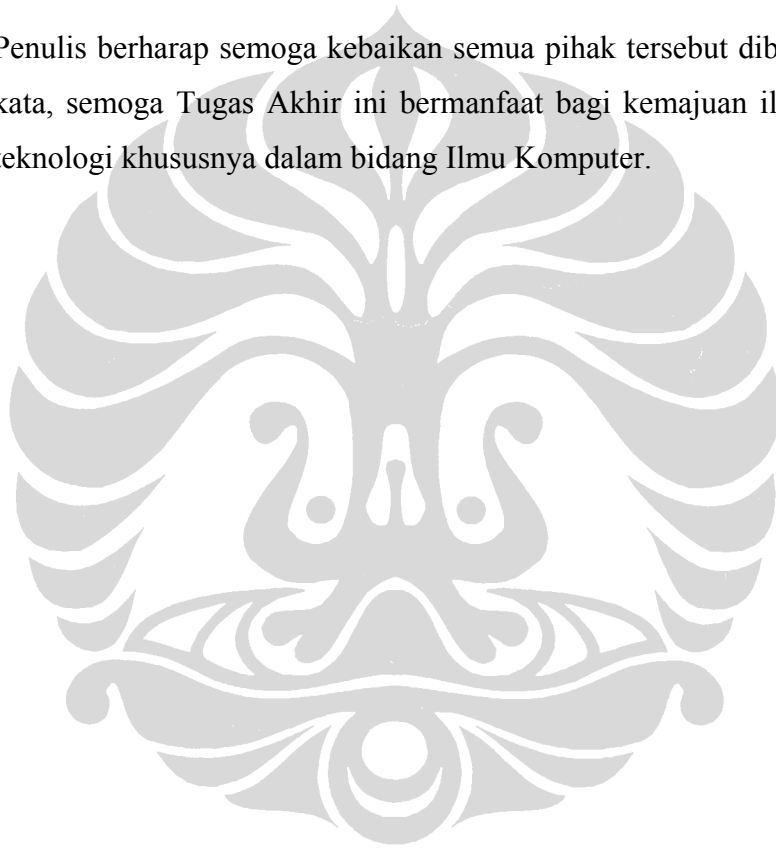
- (5) Ibu Dr. Aniati Murni, sebagai dosen pembimbing akademik yang telah menyediakan waktu, tenaga dan pikiran untuk membimbing penulis selama berkuliah di Fakultas Ilmu Komputer Universitas Indonesia.
- (6) Kedua orang tua dan seluruh anggota keluarga penulis yang telah memberikan kasih sayang dan dukungan kepada penulis selama ini.
- (7) Ibu Jenny Ong dan Ibu Siska Utoyo, sebagai orang tua asuh penulis yang telah memberikan banyak bantuan, nasehat dan dukungan kepada penulis sehingga penulis berhasil menyelesaikan kuliah di Fakultas Ilmu Komputer Universitas Indonesia.
- (8) Pawel Rendak dan Lukasz Barcikowski, sebagai pengarang salah satu referensi yang paling membantu penulis dalam memahami logika dan teknik implementasi pencarian pola sekuensial [REN09].
- (9) Martin Leonard Tangel, Iwan Prihatono, Maulana Imam, Reza Benaji, Andre Tampubolon, Candra Adhi dan Edwin Dwi sebagai teman angkatan 2004 yang bersama-sama saya melalui masa-masa terakhir berkuliah di Fakultas Ilmu Komputer Universitas Indonesia. Juga kepada seluruh teman-teman angkatan 2004 lainnya.
- (10) Anthony Steven (angkatan 2003) yang telah menjadi sumber inspirasi penulis untuk mengambil *Data Mining* sebagai topik Tugas Akhir. Juga kepada seluruh teman-teman angkatan 2003 lainnya.
- (11) Ryan Loanda (angkatan 2002), sebagai pemimpin kelompok kecil Persekutuan Oikumene (PO) yang telah memberikan bimbingan dan inspirasi bagi penulis. Juga kepada seluruh teman-teman angkatan 2002 lainnya.
- (12) Mika Permana (angkatan 2001), sebagai asisten dosen mata kuliah Dasar-Dasar Pemograman (DDP) yang telah menjadi sumber inspirasi penulis dalam mempelajari pemograman. Juga kepada seluruh teman-teman angkatan 2001 lainnya.

- (13) Semua teman-teman angkatan 2005, angkatan 2006, angkatan 2007 dan angkatan 2008 Fakultas Ilmu Komputer Universitas Indonesia. Penulis senang bisa kenal dengan teman-teman semua.
- (14) Semua dosen, staf, satpam, *janitor*, dan seluruh anggota keluarga besar Fakultas Ilmu Komputer Universitas Indonesia.
- (15) Semua pihak yang telah membantu penulis dalam pengerjaan Tugas Akhir ini namun belum disebutkan namanya.

Penulis berharap semoga kebaikan semua pihak tersebut dibalas oleh-Nya. Akhir kata, semoga Tugas Akhir ini bermanfaat bagi kemajuan ilmu pengetahuan dan teknologi khususnya dalam bidang Ilmu Komputer.

Depok, 21 Juli 2009

Penulis





**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI  
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS**

---

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama : Fandi  
NPM : 1204000327  
Program Studi : Ilmu Komputer  
Fakultas : Ilmu Komputer  
Jenis karya : Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul:

**Pencarian Pola Sekuensial Menggunakan Algoritma *AprioriAll***

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok  
Pada tanggal : 21 Juli 2009

Yang menyatakan

(.....)

# DAFTAR ISI

ABSTRAK .....	i
ABSTRACT .....	ii
DAFTAR ISI .....	iii
DAFTAR GAMBAR .....	v
DAFTAR TABEL .....	vi
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah .....	3
1.3 Tujuan .....	3
1.4 Ruang Lingkup .....	3
1.5 Metodologi .....	4
1.5.1 Studi Literatur .....	4
1.5.2 Implementasi .....	4
1.5.3 Pengujian .....	4
1.5.4 Analisis Hasil Pengujian .....	4
1.6 Sistematika Penulisan .....	5
BAB II LANDASAN TEORI .....	6
2.1 <i>Data Mining</i> .....	6
2.1.1 <i>Exploratory Data Analysis (EDA)</i> .....	6
2.1.2 <i>Descriptive Modeling</i> .....	7
2.1.3 <i>Predictive Modeling</i> .....	7
2.1.4 <i>Discovering Patterns and Rules</i> .....	7
2.1.5 <i>Retrieval by Content</i> .....	7
2.2 <i>Association rules</i> .....	8
2.3 <i>Algoritma Apriori</i> .....	11
2.4 <i>Algoritma AprioriAll</i> .....	15
2.5 <i>Mining Sequential Pattern (MSP)</i> .....	18
2.5.1 <i>Sort Phase</i> .....	20
2.5.2 <i>Litemset Phase</i> .....	21
2.5.3 <i>Transformation Phase</i> .....	22
2.5.4 <i>Sequence Phase</i> .....	23
2.5.5 <i>Maximal Phase</i> .....	23
BAB III ANALISIS DAN PERANCANGAN .....	26
3.1 Analisis Kebutuhan .....	26
3.2 Perancangan .....	29
3.2.1 Perancangan <i>Input</i> .....	29
3.2.2 Perancangan Struktur Data .....	30

3.2.3 Perancangan <i>MSP Class</i> .....	31
3.2.4 Perancangan <i>MSP_itemset Class</i> .....	33
3.2.5 Perancangan <i>Output</i> .....	35
BAB IV IMPLEMENTASI .....	39
4.1 Pembacaan <i>Input</i> .....	40
4.2 <i>Sort Phase</i> .....	42
4.3 <i>Litemset Phase</i> .....	44
4.4 <i>Transformation Phase</i> .....	46
4.5 <i>Sequence Phase</i> .....	47
4.6 <i>Maximal Phase</i> .....	50
4.7 Penampilan <i>Output</i> .....	51
4.8 Optimasi Hasil Implementasi .....	52
BAB V HASIL PENGUJIAN .....	53
5.1 Lingkungan Pengujian .....	53
5.2 Hasil Pengujian Untuk <i>Dataset Kecil</i> .....	54
5.3 Hasil Pengujian Untuk <i>Dataset Besar</i> .....	55
5.3.1 <i>Performance vs Support</i> .....	57
5.3.2 <i>Performance vs Transaction</i> .....	60
5.3.3 <i>Performance vs Itemset</i> .....	62
BAB VI PENUTUP .....	64
6.1 Kesimpulan .....	64
6.2 Saran .....	65
DAFTAR PUSTAKA .....	66

## DAFTAR GAMBAR

Gambar 2.1 – Algoritma <i>Apriori</i> [AGR94].....	11
Gambar 2.2 – <i>Function Apriori-Gen</i> [AGR94].....	11
Gambar 2.3 – Algoritma <i>AprioriAll</i> [AGR95].....	16
Gambar 2.4 – <i>Function AprioriAll-Gen</i> [AGR95].....	17
Gambar 2.5 – <i>Pseudocode Maximal Sequences</i> [AGR95].....	23
Gambar 2.6 – Diagram MSP .....	25
Gambar 3.1 – <i>FIKUI Mining</i> .....	26
Gambar 3.2 – Submenu <i>Sequential Patterns</i> .....	28
Gambar 3.3 – Tampilan <i>Input</i> .....	29
Gambar 3.4 – Struktur Data <i>Vector</i> .....	30
Gambar 3.5 – <i>MSP Class Diagram</i> .....	31
Gambar 3.6 – <i>MSP_itemset Class Diagram</i> .....	33
Gambar 3.7 – Tampilan <i>Output</i> .....	35
Gambar 3.8 – Jendela <i>Result</i> .....	37
Gambar 3.9 – <i>Output File</i> .....	38
Gambar 4.1 – Membuka <i>Input File</i> .....	40
Gambar 4.2 – <i>Pseudocode</i> Pembacaan <i>Input</i> .....	41
Gambar 4.3 – <i>Pseudocode Sort Phase</i> .....	42
Gambar 4.4 – <i>Pseudocode Litemset Phase</i> .....	44
Gambar 4.5 – Tahapan Algoritma <i>Apriori</i> .....	45
Gambar 4.6 – <i>Pseudocode Transformation Phase</i> .....	46
Gambar 4.7 – <i>Pseudocode Sequence Phase</i> .....	47
Gambar 4.8 – Tahapan Algoritma <i>AprioriAll</i> .....	48
Gambar 4.9 – <i>Candidate Generation Tree</i> .....	49
Gambar 4.10 – <i>Pseudocode Maximal Phase</i> .....	50
Gambar 4.11 – Penampilan <i>Output</i> .....	51
Gambar 5.1 – <i>Dataset Agrawal</i> .....	54
Gambar 5.2 – <i>Output Agrawal</i> .....	54
Gambar 5.3 – <i>Data Generator</i> .....	55
Gambar 5.4 – Hasil Pengujian <i>Performance vs Support</i> .....	59
Gambar 5.5 – Hasil Pengujian <i>Performance vs Transaction</i> .....	61
Gambar 5.6 – Hasil Pengujian <i>Performance vs Itemset</i> .....	63

## DAFTAR TABEL

Tabel 2.1 – <i>Transaction Database</i> .....	9
Tabel 2.2 – <i>1-itemset</i> .....	10
Tabel 2.3 – <i>large 1-itemset</i> .....	12
Tabel 2.4 – <i>2-itemset</i> .....	12
Tabel 2.5 – <i>large 2-itemset</i> .....	13
Tabel 2.6 – <i>3-itemset</i> .....	13
Tabel 2.7 – <i>large 3-itemset</i> .....	13
Tabel 2.8 – <i>All large itemsets</i> .....	14
Tabel 2.9 – <i>Customer Transaction Database</i> .....	20
Tabel 2.10 – <i>Customer Sequence</i> .....	20
Tabel 2.11 – <i>Perbedaan Support</i> .....	21
Tabel 2.12 – <i>Litemsets</i> .....	21
Tabel 2.13 – <i>Transformed Sequences</i> .....	22
Tabel 2.14 – <i>Large Sequences (mapping mode)</i> .....	23
Tabel 2.15 – <i>Maximal Sequences</i> .....	24
Tabel 2.16 – <i>Result</i> .....	24
Tabel 3.1 – <i>Deskripsi MSP Class</i> .....	32
Tabel 3.2 – <i>Deskripsi MSP_itemset</i> .....	34
Tabel 4.1 – <i>customer_database</i> .....	41
Tabel 4.2 – <i>customer_sequences</i> .....	42
Tabel 4.3 – <i>all_1_sequences</i> .....	43
Tabel 4.4 – <i>litemsets</i> .....	44
Tabel 4.5 – <i>transformed_sequences</i> .....	46
Tabel 4.6 – <i>large_Sequences</i> .....	47
Tabel 4.7 – <i>maximal_sequences</i> .....	50
Tabel 5.1 – <i>Karakteristik Data</i> .....	56
Tabel 5.2 – <i>Karakteristik Dataset untuk Support 90%</i> .....	57
Tabel 5.3 – <i>Karakteristik Dataset untuk Support 75%</i> .....	57
Tabel 5.4 – <i>Karakteristik Dataset untuk Support 50%</i> .....	57
Tabel 5.5 – <i>Karakteristik Dataset untuk Support 25%</i> .....	58
Tabel 5.6 – <i>Karakteristik Dataset untuk Support 10%</i> .....	58
Tabel 5.7 – <i>Hasil Pengujian Performance vs Support</i> .....	58
Tabel 5.8 – <i>Karakteristik Dataset untuk Performance vs Transaction</i> .....	60
Tabel 5.9 – <i>Hasil Pengujian Performance vs Transaction</i> .....	60
Tabel 5.10 – <i>Karakteristik Dataset untuk Performance vs Itemset</i> .....	62
Tabel 5.11 – <i>Hasil Pengujian Performance vs Itemset</i> .....	62