

**PROGRAM BANTU PREDIKSI PENJUALAN DENGAN
MENGUNAKAN ALGORITMA SEQUENTIAL PATTERN DISCOVERY
USING EQUIVALENT CLASS STUDI KASUS: RAHAYU MOTOR**

Skripsi



oleh

TONY SETIAWAN

22094649

**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI
INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA**

2014

**PROGRAM BANTU PREDIKSI PENJUALAN DENGAN
MENGUNAKAN ALGORITMA SEQUENTIAL PATTERN DISCOVERY
USING EQUIVALENT CLASS STUDI KASUS: RAHAYU MOTOR**

SKRIPSI



Diajukan kepada program Program Studi Teknik Informatika Fakultas Teknologi Informasi

Universitas Kristen Duta Wacana

sebagai Salah Satu Syarat dalam Memperoleh Gelar

Sarjana Komputer

Disusun oleh

TONY SETIAWAN

22094649

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS KRISTEN DUTA WACANA

2014

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

**PROGRAM BANTU PREDIKSI PENJUALAN DENGAN
MENGUNAKAN ALGORITMA SEQUENTIAL PATTERN DISCOVERY
USING EQUIVALENT CLASS
STUDI KASUS: RAHAYU MOTOR**

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni penaubutan gelar keserjanaan saya.

ogyakarta, 11 September 2014



TONY SETIAWAN
22094649

HALAMAN PERSETUJUAN

Judul Skripsi : PROGRAM BANTU PREDIKSI PENJUALAN
DENGAN MENGGUNAKAN ALGORITMA
SEQUENTIAL PATTERN DISCOVERY USING
EQUIVALENT CLASS
STUDI KASUS: RAHAYU MOTOR

Nama Mahasiswa : TONY SETIAWAN
NIM : 22094649
Matakuliah : Skripsi (Tugas Akhir)
Kode : TIW276
Semester : Genap
Tahun Akademik : 2013/2014

Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 11 September 2014

Dosen Pembimbing I



Budi Susanto, SKom., M.T.

Dosen Pembimbing II



Hendro Setiadi, M.Eng

HALAMAN PENGESAHAN

**PROGRAM BANTU PREDIKSI PENJUALAN DENGAN
MENGUNAKAN ALGORITMA SEQUENTIAL PATTERN
DISCOVERY USING EQUIVALENT CLASS
STUDI KASUS: RAHAYU MOTOR**

Oleh: TONY SETIAWAN / 22094649

Dipertahankan di depan Dewan Penguji Sarjana
Program Studi Teknik Informatika Fakultas Teknologi Informatika
Universitas Kristen Duta Wacana Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 12 Agustus 2014

Yogyakarta, 11 September 2014

Mengesahkan,

Dewan Penguji:

1. Budi Susanto, S.T., M.T.
2. Hendro Setiadi, M.T.
3. Ersa Dharma, S.Kom., M.Sc.
4. Widy Soeroto, Salarjo, S.Kom., M.Cs



Dekan

(Drs. Wimmie Handiwidjono, MIT.)

Ketua Program Studi

(Nugroho Agus Haryono, M.Si)

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa atas berkat, rahmat, dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “PROGRAM BANTU PREDIKSI PENJUALAN DENGAN MENGGUNAKAN ALGORITMA SEQUENTIAL PATTERN DISCOVERY USING EQUIVALENT CLASS STUDI KASUS: RAHAYU MOTOR” dengan baik dan tepat waktu.

Penulisan laporan ini merupakan kelengkapan dan pemenuhan dari salah satu syarat dalam memperoleh gelar Sarjana Komputer. Selain itu, penulisan laporan Tugas Akhir ini juga bertujuan untuk melatih mahasiswa agar dapat menghasilkan suatu karya yang dapat dipertanggungjawabkan secara ilmiah, sehingga dapat bermanfaat bagi penggunanya.

Dalam menyelesaikan penelitian dan laporan Tugas Akhir ini, penulis telah banyak menerima bimbingan, saran, dan masukan dari berbagai pihak, baik secara langsung maupun secara tidak langsung. Untuk itu dengan segala kerendahan hati, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada :

1. Bapak **Budi Santoro, SKom.,M.T.** selaku dosen pembimbing I yang selalu sabar dalam membimbing penulis dalam mengerjakan penelitian dan penyusunan laporan Tugas Akhir.
2. Bapak **Hendro Setiadi, M.Eng** selaku dosen pembimbing II yang selalu sabar dan baik membimbing penulis dalam mengerjakan penelitian dan penyusunan laporan Tugas Akhir.
3. Keluarga dan someone special yang selalu memberikan doa dan dukungannya kepada penulis dalam menyelesaikan Tugas Akhir.
4. Rekan-rekan penulis yang dengan senang hati memberikan arahan, saran, dan, berbagi dalam pengerjaan Tugas Akhir maupun penulisan laporan Tugas Akhir.

5. Pihak lain yang tidak dapat penulis sebutkan satu per satu, sehingga Tugas Akhir ini dapat terselesaikan dengan baik.

Penulis menyadari bahwa penelitian dan laporan Tugas Akhir ini masih jauh dari sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun dari pembaca sekalian, sehingga suatu saat nanti penulis dapat memberikan karya yang lebih baik lagi.

Akhir kata penulis meminta maaf bila ada kesalahan dalam penyusunan laporan maupun sewaktu penulis melakukan penelitian Tugas Akhir. Semoga penelitian dan laporan Tugas Akhir ini dapat berguna bagi kita semua.

Yogyakarta, 5 Agustus 2014

Penulis

INTISARI

PROGRAM BANTU PREDIKSI PENJUALAN DENGAN MENGGUNAKAN ALGORITMA SEQUENTIAL PATTERN DISCOVERY USING EQUIVALENT CLASS STUDI KASUS: RAHAYU MOTOR

Saat ini, persaingan dalam dunia bisnis semakin ketat. Banyaknya jumlah toko yang menjual jenis barang yang sama dengan harga yang tidak jauh berbeda menimbulkan persaingan yang tidak dapat dihindari. Supaya suatu toko yang menjual jenis barang yang sama dapat memiliki keunggulan yang lebih dengan toko yang lain, toko tersebut harus memiliki keunggulan dibanding toko yang lain. Salah satu cara untuk mendapatkan keunggulan tersebut adalah dengan mengetahui pola beli pelanggan. Dengan mengetahui pola beli pelanggan, suatu toko dapat mengetahui jenis barang yang dibeli secara berkala sehingga dapat memiliki keuntungan dari toko lain yang menjual jenis barang yang sama dan otomatis dapat meningkatkan keuntungan penjualan.

Sequence mining merupakan salah satu metode *data mining* yang sering digunakan pada saat ini. Banyak algoritma yang dapat digunakan dalam *sequence mining*, seperti algoritma apriori, GSP, dan SPADE (*Sequential Pattern Discovery using Equivalent Class*).

Hasil dari penelitian menggunakan algoritma SPADE yang diuji dengan data dari Rahayu motor menunjukkan bahwa dari tahun ke tahun, memang terjadi kesamaan jenis barang yang dibeli secara berkala oleh beberapa pelanggan yang berbeda. Hal ini menunjukkan bahwa memang barang tersebut merupakan jenis barang yang paling banyak laku dari tahun ke tahun dan merupakan jenis barang yang tidak boleh sampai habis stok nya.

Kata kunci : Sequence Mining, *Data Mining*, SPADE

DAFTAR ISI

HALAMAN JUDUL	
PERNYATAAN KEASLIAN TUGAS AKHIR.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
KATA PENGANTAR	vi
INTISARI	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
BAB 1	1
1.1 Latar Belakang Masalah.....	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Metode Penelitian	3
1.6 Sistem dan Keorisipan	4
BAB 2	5
2.1 Tinjauan Pustaka.....	5
2.2 Landasan Teori.....	6
2.1.1 Data Mining	6
2.1.2 Sequence Mining.....	6
2.1.3 Breadth First Search.....	7
2.1.4 Depth First Search.....	8
2.1.5 Algoritma SPADE.....	9
BAB 3	16
3.1 Kebutuhan Sistem	16
3.1.1 Use Case Diagram.....	16

3.1.2 Expanded Use Case.....	17
3.2. Diagram Blok Proses	21
3.3. Flowchart Sistem	22
3.3.1 Pencarian Frequent Sequence item 1 dan 2.....	23
3.3.2 Menghitung Frequent Sequence.....	25
3.3.3 Penggabungan id-list sementara.....	26
3.3.4 Analisis data penjualan berdasarkan algoritma SPADE	28
3.4. Perancangan Database.....	29
3.4.1 ER Diagram	29
3.4.2 Rancangan Tabel Database	30
3.5. Rancangan Antar Muka	32
3.6. Rancangan Pengujian Sistem.....	33
BAB 4	34
4.1. Implementasi Sistem	34
4.1.1 Implementasi Proses Extract, Transform, Load.....	34
4.1.2 Implementasi Proses Data Mining Dengan Algoritma SPADE.....	36
4.1.2.1 Implementasi Proses Menghitung Sekuens 1.....	36
4.1.2.2 Implementasi Proses Menghitung Sekuens 2.....	37
4.1.2.3 Implementasi Proses Enumerate Frequent Sekuens.....	37
4.1.2.4 Implementasi Proses Menghitung Confidence.....	39
4.1.2 Implementasi Algoritma Menggunakan Sampel Data	40
4.2. Hasil Analisis Menggunakan Data Sebenarnya	43
BAB 5	44
5.1. Kesimpulan	54
5.2. Saran	54
DAFTAR PUSTAKA	55
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2. 1 Contoh sequence mining pada www.amazon.com	7
Gambar 2. 2 Contoh BFS.....	8
Gambar 2. 3 Contoh DFS.....	9
Gambar 2. 4 Rumus algoritma SPADE	10
Gambar 2. 5 Contoh transformasi database vertikal	12
Gambar 2. 6 Pseudocode untuk BFS dan DFS	13
Gambar 2. 7 Penggabungan id-list sementara.....	14
Gambar 2. 8 Algoritma pemangkasan sekuens	15
Gambar 3. 1 Use case analisis sistem	16
Gambar 3. 2 Gambar diagram blok proses	21
Gambar 3. 3 Pencarian frequent sequence item 1 dan 2	23
Gambar 3. 4 Pencacahan frequent sequence	25
Gambar 3. 5 Penggabungan id-list.....	26
Gambar 3. 6 Flowchart analisis data penjualan berdasarkan algoritma SPADE	28
Gambar 3. 7 Rancangan entity relationship diagram	29
Gambar 3. 8 Rancangan antarmuka analisis algoritma SPADE	32
Gambar 3. 9 Rancangan antarmuka hasil analisis algoritma SPADE.....	33
Gambar 4. 1 Folder database awal.....	34
Gambar 4. 2 Proses penggantian ekstensi file JUALx.bin menjadi JUALx.dbf.....	35
Gambar 4. 3 File database JUALx.dbf yang dibuka menggunakan Microsoft Visual FoxPro.....	35
Gambar 4. 4 Contoh perhitungan algoritma SPADE.....	40
Gambar 4. 5 Hasil analisis sekuens 1 pada sistem dengan menggunakan contoh sampel data	41
Gambar 4. 6 Hasil analisis sekuens 2 pada sistem dengan menggunakan contoh sampel data	42
Gambar 4. 7 Hasil analisis sekuens 3 pada sistem dengan menggunakan contoh sampel data	42
Gambar 4. 8 Hasil analisis sekuens 4 pada sistem dengan menggunakan contoh sampel data	42
Gambar 4. 9 Hasil analisis sekuens 5 pada sistem dengan menggunakan contoh sampel data	43
Gambar 4. 10 Hasil analisis data untuk tahun 2009 dengan minimal support 4 dan confidence 70%	44
Gambar 4. 11 Hasil analisis data untuk tahun 2009 dengan minimal support 5 dan confidence 70%	45
Gambar 4. 12 Hasil analisis data untuk tahun 2009 dengan minimal support 4 dan confidence 20%	45
Gambar 4. 13 Hasil analisis data untuk tahun 2009 dengan minimal support 8 dan confidence 70%	47
Gambar 4. 14 Error creating window handle	48
Gambar 4. 15 Analisis data tahun 2010 bulan januari hingga maret dengan minimal support 4 dan confidence 70%	49

Gambar 4. 16 Analisis data tahun 2010 bulan april hingga juni dengan minimal support 4 dan confidence 70%	49
Gambar 4. 17 Analisis data tahun 2010 bulan juli hingga september dengan minimal support 4 dan confidence 70%	50
Gambar 4. 18 Analisis data tahun 2010 bulan oktober hingga desember dengan minimal support 4 dan confidence 70%	50

@UKDW

DAFTAR TABEL

Tabel 3. 1 Expanded use case untuk add user	17
Tabel 3. 2 Expanded use case untuk edit password	17
Tabel 3. 3 Expanded use case untuk delete user	18
Tabel 3. 4 Expanded use case untuk lihat histori transaksi	19
Tabel 3. 5 Expanded use case untuk analisis data mining dengan algoritma SPADE	19
Tabel 3. 6 Expanded use case untuk managemen user control	20
Tabel 3. 7 Expanded use case untuk login	21
Tabel 3. 8 Tabel pelanggan	30
Tabel 3. 9 Tabel penjualan	30
Tabel 3. 10 Tabel barang	31
Tabel 3. 11 Tabel jenis	31
Tabel 3. 12 Tabel user	31
Tabel 4. 1 Tabel perbandingan nilai minimal support dan confidence	46

@UKDWN

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Saat ini, persaingan dalam dunia bisnis semakin ketat. Banyaknya jumlah toko yang menjual jenis barang yang sama dengan harga yang tidak jauh berbeda menimbulkan persaingan yang tidak dapat dihindari. Supaya suatu toko yang menjual jenis barang yang sama dapat memiliki keunggulan yang lebih dengan toko yang lain, toko tersebut harus memiliki keunggulan dibanding toko yang lain. Salah satu cara untuk mendapatkan keunggulan tersebut adalah dengan mengetahui pola beli pelanggan. Dengan mengetahui pola beli pelanggan, suatu toko dapat mengetahui jenis barang yang dibeli secara bersamaan sehingga dapat memiliki keuntungan dari toko lain yang menjual jenis barang yang sama dan otomatis dapat meningkatkan keuntungan penjualan.

Sequence mining merupakan salah satu metode *data mining* yang sering digunakan pada saat ini. Pada mulanya, *sequence mining* digunakan untuk *market basket analysis* dimana relasi antar transaksi akan di analisis untuk memperoleh transaksi mana saja yang saling berkaitan satu dengan yang lain. Banyak algoritma yang dapat digunakan dalam *sequence mining*, seperti algoritma apriori, GSP, dan SPADE (*Sequential Pattern Discovery using Equivalent Class*). Kelebihan algoritma SPADE dibandingkan dengan algoritma apriori atau GSP adalah algoritma SPADE bekerja dengan menggunakan database vertikal (setiap barang di asosiasi kan dengan sebuah id pelanggan dan id transaksi) dimana algoritma apriori dan GSP menggunakan database horisontal (seorang pelanggan memiliki data transaksi yang telah dilakukannya). Kelebihan lain dari algoritma SPADE adalah sebuah masalah utama akan dipecah-pecah menjadi beberapa masalah kecil kemudian menyelesaikan masalah tersebut secara terpisah di dalam memori. Selain itu, algoritma SPADE juga dapat menemukan semua sequence yang ada hanya dalam tiga kali pencarian dalam database. Hal ini menjadikan algoritma SPADE menjadi lebih efisien daripada algoritma lain, terutama dalam ukuran database yang besar. Zaki (2001) juga mengatakan bahwa algoritma SPADE lebih efektif dan efisien jika dibandingkan dengan algoritma lain seperti GSP atau apriori.

Rencana yang akan dilakukan pada tugas akhir ini adalah dengan membuat sebuah sistem jual beli barang dan analisis data yang berdasarkan pada algoritma SPADE untuk menghasilkan informasi tentang seberapa sering orang membeli item bersama-sama dalam beberapa transaksi penjualan.

1.2. Perumusan Masalah

Permasalahan yang akan dibahas pada tugas akhir ini adalah bagaimana metode SPADE dapat diterapkan pada data penjualan barang dari tiap pelanggan dengan jenis barang tertentu. Dari hasil analisis sekuens penjualan menggunakan algoritma SPADE, dapat diketahui kombinasi jenis barang apa saja yang dibeli pelanggan dalam kurun waktu tertentu.

1.3. Batasan Masalah

Dalam penelitian ini, penulis memberikan batasan masalah untuk sistem yang akan dibuat:

1. Data yang akan diolah adalah data penjualan yang dimiliki oleh Rahayu motor periode tahun 2009-2010.
2. Minimum sequence yang dapat di proses adalah 4.
3. Maksimum jumlah data yang dapat di analisis tiap proses adalah data selama 6 bulan.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menghasilkan sebuah sistem yang dapat memberikan analisis tentang kombinasi jenis barang apa saja yang sering dibeli oleh pelanggan dalam rentang waktu tertentu dengan menggunakan algoritma SPADE.

Manfaat yang didapatkan dari penerapan sistem analisis data dengan menggunakan algoritma SPADE adalah mendapat "*competitive advantage*" dalam dunia bisnis sehingga akan memberikan keuntungan lebih banyak karena pengambilan keputusan dalam bisnis telah berdasarkan hasil dari analisa data.

1.5. Metode Penelitian

Dalam penulisan Skripsi ini, digunakan beberapa metode sebagai acuan dalam perancangan, implementasi dan penelitian terhadap sistem yang dibuat. Metode tersebut sebagai berikut:

1. Studi Pustaka

Studi Pustaka dilakukan dengan mempelajari teori-teori melalui buku, artikel, jurnal dan bahan lain yang mendukung yang berhubungan dengan algoritma *SPADE*, *data mining* dan metode-metode pendukung lainnya yang dibutuhkan.

2. Pengembangan system menggunakan metodologi

a. Pengumpulan data

Metode pengumpulan data dilakukan dengan cara melakukan kunjungan ke toko bayi Wijaya dengan tujuan untuk mendapatkan data penjualan tahun 2010-2011

b. Pembuatan sistem

Terdapat beberapa tahap dalam pembuatan sistem ini, pertama yaitu membuat sistem informasi toko yang meliputi *CRUD (Create Read Update Delete)*.

c. Analisis data

Analisis untuk mendapatkan pola penjualan barang akan menggunakan algoritma *SPADE*

d. Implementasi dan testing

Pada tahap ini akan dilakukan percobaan berupa memberikan *input* sesuai batasan masalah yang ada. Lalu akan dilihat *Output* yang diinginkan adalah berupa *data* barang yang muncul setelah dilakukan proses *data mining*.

1.6. Sistematika Penulisan

Skripsi ini disusun dalam sebuah laporan dengan sistematika atau spesifikasi terdiri dari 5 bab:

Bab 1 PENDAHULUAN yang berisi latar belakang masalah, perumusan masalah, batasan masalah, hipotesis, tujuan penelitian, metodologi, dan sistematika penulisan skripsi.

Bab 2 TINJAUAN PUSTAKA yang berisi gagasan-gagasan yang muncul dengan memberikan landasan teori yang akurat dari berbagai sumber dan konsep-konsep yang dibutuhkan dalam pembuatan sistem aplikasi *data mining* menggunakan algoritma SPADE.

Bab 3 ANALISIS DAN PERANCANGAN SISTEM yang berisi perancangan sistem yang akan memberikan gambaran sistem yang akan dibuat serta prosedur-prosedur yang digunakan dalam sistem.

Bab 4 IMPLEMENTASI DAN ANALISIS SISTEM yang berisi implementasi dari hasil perancangan sistem dan pengujian terhadap sistem yang telah dibuat.

Bab 5 KESIMPULAN DAN SARAN YANG berisi kesimpulan atas sistem yang telah dibuat serta saran-saran dalam pengembangan dari Skripsi ini agar dapat dikembangkan kembali.

BAB 5

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari penelitian yang telah dilakukan oleh penulis, terdapat beberapa kesimpulan yang diperoleh, yaitu sebagai berikut :

1. Algoritma SPADE memberikan gambaran tentang kombinasi jenis barang apa saja yang paling sering dibeli oleh pelanggan dengan batasan nilai *support* dan *confidence* yang dimasukkan oleh *user* dalam kurun waktu tertentu.
2. Dengan menggunakan algoritma SPADE, pengguna dapat mengetahui hasil analisis kombinasi barang apa saja yang dibeli oleh pelanggan tertentu dalam rentang waktu tertentu. Setelah hasil analisis didapat, dapat dibuat prediksi yang relatif akurat untuk penjualan pada tahun berikutnya.
3. Nilai minimal *support* dan *confidence* berpengaruh untuk menentukan hasil perhitungan dari algoritma SPADE. Semakin tinggi nilai minimal *support* dan *confidence*, semakin spesifik hasil analisis yang didapat.
4. Algoritma SPADE memiliki kelemahan yaitu proses data yang lama jika jenis data yang diuji sangat banyak. Hal ini dikarenakan algoritma SPADE akan mencari satu-persatu kombinasi yang mungkin dari tiap item di sid dan erit yang ada.

5.2. Saran

Untuk pengembangan sistem lebih lanjut, saran yang dapat diberikan oleh penulis adalah dengan lebih memelihara sistem dan selalu melakukan proses maintenance secara berkala untuk memastikan data yang masuk ke dalam sistem merupakan data yang sudah bersih sehingga proses perhitungan algoritma menjadi lebih optimal. Selain itu, jika menginginkan proses perhitungan yang lebih cepat dapat mengganti bahasa pemrograman menggunakan bahasa C++ karena C++ merupakan *native language* sehingga *memory* yang digunakan menjadi lebih ringan dibandingkan vb.net yang dibatasi oleh *framework* dari sistem.

DAFTAR PUSTAKA

- H. Cheng, X. Yan, and J. Han. IncSpan: Incremental mining of sequential patterns in large database. In Proc. 2004 ACM SIGKDD Int. Conf. Knowledge Discovery in Databases (KDD'04), pages 527–532, Seattle, WA, Aug. 2004.
- Jiawei, H and Kamber, M. Data Mining Concepts and Techniques Second Edition. San Francisco: Morgan Kaufmann Publishers, 2006
- Jiawei, H., Kamber, M., and Pei, J. Data Mining Concepts and Techniques Third Edition. USA: Morgan Kaufmann Publishers, 2012
- M. Garofalakis, R. Rastogi, and K. Shim. SPIRIT: Sequential pattern mining with regular expression constraints. In Proc. 1999 Int. Conf. Very Large Data Bases (VLDB'99), pages 223–234, Edinburgh, UK, Sept. 1999.
- NIZAR R. MABROUKEH and C. I. EZEIFE, A Taxonomy of Sequential Pattern Mining Algorithms, ACM Computing Surveys, Vol. 43, No. 1, Article 3, Publication date: November 2010
- R.Agrawal and R.Srikant. Mining sequential patterns. Proceedings of the Eleventh International Conference on Data Engineering, 1995.
- R. Srikant and R. Agrawal. Mining sequential patterns: Generalizations and performance improvements. In Proc. 5th Int. Conf. Extending Database Technology (EDBT'96), pages 3–17, Avignon, France, Mar. 1996.
- W.Wang and J. Wang. Mining Sequential Patterns from Large data Sets, Springer, 2005.
- Zaki, M. SPADE: An Efficient Algorithm for Mining Frequent Sequences. The Netherlands: Kluwer Academic Publishers, 2001