

***ASSOCIATION RULE MINING DATA PEMINJAMAN
PERPUSTAKAAN MENGGUNAKAN KOMBINASI
APRIORI DAN JACCARD SIMILARITY***

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Syarat Mencapai Gelar Strata Satu
Program Studi Informatika



Disusun oleh :

MUHAMMAD HEZBY AL HAQ

NIM. M0511038

**PROGRAM STUDI INFORMATIKA
FAKULTAS MATEMATIKA & ILMU PENGETAHUAN ALAM
UNIVERSITAS SEBELAS MARET**

2016

commit to user

***ASSOCIATION RULE MINING DATA PEMINJAMAN
PERPUSTAKAAN MENGGUNAKAN APRIORI DAN
JACCARD SIMILARITY***

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Syarat Mencapai Gelar Strata Satu
Program Studi Informatika



Disusun oleh :

MUHAMMAD HEZBY AL HAQ

NIM. M0511038

**PROGRAM STUDI INFORMATIKA
FAKULTAS MATEMATIKA & ILMU PENGETAHUAN ALAM
UNIVERSITAS SEBELAS MARET**

2016

commit to user

SKRIPSI

***ASSOCIATION RULE MINING DATA PEMINJAMAN
PERPUSTAKAAN MENGGUNAKAN APRIORI DAN JACCARD
SIMILARITY***

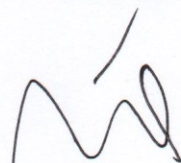
Disusun oleh:

MUHAMMAD HEZBY AL HAQ

M0511038

**Skripsi ini telah disetujui untuk dipertahankan di hadapan dewan penguji
pada tanggal: 25 Oktober 2016**

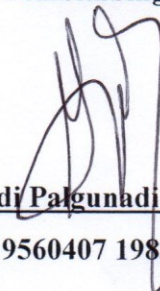
Pembimbing I



Ristu Saptono, S. Si., M. T.

NIP. 19790210 200211 2 001

Pembimbing II



Drs. Sarngadi Palgunadi Yohanes M.Sc.

NIP. 19560407 198303 1 004

SKRIPSI

**ASSOCIATION RULE MINING DATA PEMINJAMAN
PERPUSTAKAAN MENGGUNAKAN APRIORI DAN JACCARD
SIMILARITY**

Disusun oleh :

MUHAMMAD HEZBY AL HAQ

M0511038

Telah dipertahankan di hadapan Dewan Penguji
Pada tanggal 25 Oktober 2016

Susunan Dewan Penguji

1. Ristu Saptono, S. Si., M. T.
NIP. 19790210 200211 2 001

()

2. Drs. Sarngadi Palgunadi Yohanes M.Sc.
NIP. 19560407 198303 1 004

()

3. Dr. Wiranto, M.Kom., M.Cs.
NIP. 19661230 199302 1 001

()

4. Drs. Bambang Harjito M.App.Sc., Ph.D.
NIP. 19621130 199103 1 002

()

Disahkan oleh:

Kepala Program Studi Informatika



Drs. Bambang Harjito, M.App.Sc., Ph.D.
NIP. 19621130 199103 1 002

***ASSOCIATION RULE MINING DATA PEMINJAMAN PERPUSTAKAAN
MENGUNAKAN APRIORI DAN JACCARD SIMILARITY***

MUHAMMAD HEZBY AL HAQ

Program Studi Informatika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Sebelas Maret

ABSTRAK

UPT Perpustakaan UNS memiliki jumlah koleksi sebanyak 37.271 buah pustaka dan rata-rata 75.316 pertahun sirkulasi buku yang dikelola dalam sistem UNSLA (*UNS Library Automation*). Analisis diperlukan untuk menggali informasi berharga yang dapat digunakan untuk berbagai kepentingan. *Apriori* adalah algoritma yang paling sering digunakan dalam *association rule mining*. Penggunaan *Apriori* memiliki kelemahan pada data yang tersebar. *Jaccard Similarity* adalah algoritma yang digunakan untuk mencari kesamaan antar dua set. Penerapan *Jaccard Similarity* pada *association rule mining* diketahui dapat menemukan *association rule* pada data yang tersebar. Penelitian ini dilakukan untuk mencari tahu bagaimana konsistensi *association rule* yang dihasilkan kombinasi *Apriori* dan *Jaccard Similarity* dibandingkan dengan *Apriori* dan *Jaccard Similarity* pada data peminjaman buku UPT Perpustakaan UNS. Data buku yang digunakan dikelompokkan menjadi 10 kategori buku dan dipecah berdasarkan bulan dan tahun. *Association rule mining* dilakukan dengan menggunakan ketiga metode. *Association rule* yang dihasilkan dibandingkan konsistensinya pada bulan dan tahun yang diketahui. Sebagai hasil, diketahui bahwa *association rule mining* menggunakan kombinasi *Apriori* dan *Jaccard Similarity* lebih konsisten dibandingkan *Apriori* dan *Jaccard Similarity*. Walaupun begitu, *association rule mining* menggunakan *Jaccard Similarity* menghasilkan variasi lebih banyak dibanding *Apriori* dan kombinasinya.

Kata Kunci: Perpustakaan, Data Mining, Association Rule, Apriori, Jaccard Similarity

ASSOCIATION RULE MINING ON LIBRARY BOOKS LENDING DATA USING APRIORI AND JACCARD SIMILARITY

MUHAMMAD HEZBY AL HAQ

Department of Informatics, Mathematic and Natural Science Faculty,
Sebelas Maret University

ABSTRACT

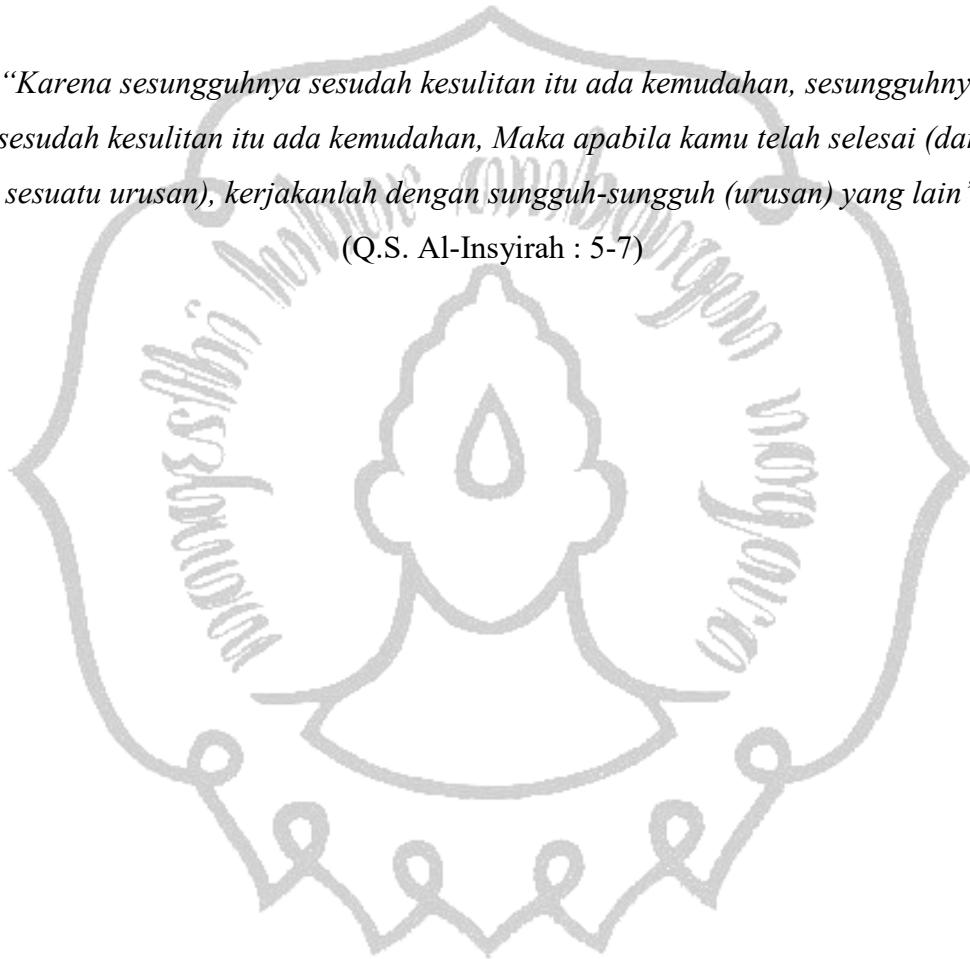
UPT Perpustakaan UNS has 37,271 collections and on average 75,316 annual circulations of the book that is managed by UNSLA (UNS Library Automation). An analysis is needed to discover valuable information that can be used for various purposes. Association rule mining is one of data mining techniques to look for relationship pattern in the market basket data. Apriori algorithm is commonly used in association rule mining. However, Apriori has limitations in conducting association rule mining on sparse data. Jaccard Similarity algorithm is used to find the similarities between the two sets. Application of Jaccard Similarity to the association rule mining can find association rule on sparse data. This research was conducted to determine the consistency of association rule generated by the combination of both Apriori and Jaccard Similarity compared to regular Apriori and Jaccard Similarity on the book lending data of UPT Library UNS. Data are grouped into ten different categories of books and split by month and year. Association rule mining is done by using all three methods. Association rules produced by each method compared for consistency in the known month and year. As a result, it is known that the association rule mining using a combination of Apriori and Jaccard Similarity is more consistent than the original Apriori and Jaccard Similarity. However, association rule mining using Jaccard Similarity generate more variation than Apriori and combination.

Keywords: Library, Data Mining, Association Rule, Apriori, Jaccard Similarity

MOTTO

“Sesungguhnya orang-orang yang mengatakan: “Tuhan kami ialah Allah” kemudian mereka meneguhkan pendirian mereka, maka malaikat akan turun kepada mereka dengan mengatakan: “Janganlah kamu takut dan janganlah merasa sedih dan bergembiralah kamu dengan jannah yang telah dijanjikan Allah kepadamu”” (Q.S. Al-Fushilat: 30)

“Karena sesungguhnya sesudah kesulitan itu ada kemudahan, sesungguhnya sesudah kesulitan itu ada kemudahan, Maka apabila kamu telah selesai (dari sesuatu urusan), kerjakanlah dengan sungguh-sungguh (urusan) yang lain” (Q.S. Al-Insyirah : 5-7)



PERSEMBAHAN

Karya ini penulis persembahkan kepada :

“Umi dan Abi serta keluarga. Semoga senantiasa dalam limpahan hidayah-Nya hingga di akhirat kelak kita dapat bertemu kembali sebagai hamba-Nya yang bersyukur.”

“Islam. Semoga diri yang hina ini senantiasa diberikan keistiqomahan dalam menyampaikan risalah Islam sebagai rahmatan lil alamin.”

“Tanah Merah, Indonesia tercinta. Semoga di Negeri ini makin baik dalam pengolahan data.”



KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah Subhanahu Wa Ta'ala yang senantiasa memberikan nikmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “*Association Rule Mining Data Peminjaman Perpustakaan Menggunakan Apriori dan Jaccard Similarity*” yang menjadi salah satu syarat wajib untuk memperoleh gelar Sarjana Informatika di Universitas Sebelas Maret (UNS) Surakarta.

Penulis menyadari akan keterbatasan yang dimiliki, begitu banyak bimbingan, bantuan, serta motivasi yang diberikan dalam proses penyusunan skripsi ini. Oleh karena itu, ucapan terima kasih penulis sampaikan kepada :

1. Bapak Drs. Bambang Harjito, M.App.Sc., Ph.D. selaku Ketua Program Studi S1 Informatika yang telah banyak memberi bimbingan dan pengarahan selama penulis menempuh studi di Program Studi Informatika FMIPA UNS,
2. Bapak Ristu Saptono, S. Si., M. T. selaku Dosen Pembimbing I yang penuh kesabaran membimbing, mengarahkan, dan memberi motivasi kepada penulis selama proses penyusunan skripsi ini,
3. Bapak Drs. Sarngadi Palgunadi Yohanes M.Sc. selaku Dosen Pembimbing II yang penuh kesabaran membimbing, mengarahkan, dan memberi motivasi kepada penulis selama proses penyusunan skripsi ini,
4. Bapak Abdul Aziz, S.Kom., M.Cs. selaku Dosen Pembimbing Akademik yang penuh kesabaran membimbing, mengarahkan, dan memberi motivasi kepada penulis selama penulis menempuh kuliah S1,
5. Bapak dan Ibu dosen serta karyawan di Program Studi Informatika FMIPA UNS yang telah mengajar penulis selama masa studi dan membantu dalam proses penyusunan skripsi ini,
6. Umi, Abi, dan keluarga, serta teman-teman yang telah memberikan bantuan dan dukungan sehingga penyusunan skripsi ini dapat terselesaikan.

Penulis berharap semoga skripsi ini dapat bermanfaat bagi semua pihak yang berkepentingan.

Surakarta, 25 Oktober 2016

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
ABSTRAK	iv
ABSTRACT.....	v
MOTTO	ivi
HALAMAN PERSEMBAHAN	ivii
KATA PENGANTAR	iviii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN.....	xviii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Landasan Teori.....	5

2.1.1	UNS Library Automation (UNSLA)	5
2.1.2	Database Management System (DBMS)	5
2.1.3	MySQL	5
2.1.4	Knowledge Discovery.....	5
2.1.5	Preprocessing	6
2.1.6	Data Mining	7
2.1.7	Association Analysis.....	7
2.1.8	Market Basket Data.....	8
2.1.9	Binary Representation.....	8
2.1.10	Itemset dan Support Count.....	9
2.1.11	Association Rule	9
2.1.12	Association Rule Discovery.....	9
2.1.13	Apriori.....	10
2.1.14	Jaccard Similarity.....	11
2.2	Penelitian Terkait	12
BAB III METODE PENELITIAN.....		16
3.1.	Pengumpulan Data	16
3.2.	Preprocessing	17
3.3.	Pengolahan Data	18
3.3.1.	Pengolahan Data Menggunakan Apriori.....	18
3.3.2.	Pengolahan Data Menggunakan Jaccard Similarity	19
3.3.3.	Kombinasi Apriori dan Jaccard Similarity	19
3.4.	Analisis Konsistensi.....	20
3.5.	Perbandingan.....	21

BAB IV PEMBAHASAN.....	22
4.1. Deskripsi Data.....	22
4.2. Preprocessing	23
4.3. Pengolahan Data	26
4.3.1. Pengolahan Data Menggunakan Apriori.....	26
4.3.1.1. Optimasi Frequent Itemset Generation	26
4.3.1.2. Optimasi Rule Generation.....	28
4.3.1.3. Hasil Pengolahan Data Menggunakan Apriori	29
4.3.2. Pengolahan Data Menggunakan Jaccard Similarity	32
4.3.2.1. Optimasi Prune Candidate	32
4.3.2.2. Hasil Jaccard Similarity	34
4.3.3. Pengolahan Data Menggunakan Kombinasi Apriori dan Jaccard Similarity.....	37
4.3.3.1. Optimasi Prune Candidate	37
4.3.3.2. Hasil Kombinasi.....	39
4.4. Analisis Konsistensi.....	41
4.4.1. Analisis Konsistensi Per Bulan pada Tahun Berbeda.....	41
4.4.1.1. Analisis Konsistensi Per Bulan pada Tahun Berbeda pada Apriori	41
4.4.1.2. Analisis Konsistensi Per Bulan pada Tahun Berbeda pada Jaccard Similarity	43
4.4.1.3. Analisis Konsistensi Per Bulan pada Tahun Berbeda pada Kombinasi Apriori dan Jaccard Similarity	46
4.4.2. Analisis Konsistensi Per Bulan dan Tahun.....	47

4.4.2.1. Analisis Konsistensi Per Bulan dan Tahun pada Apriori.....	48
4.4.2.2. Analisis Konsistensi Per Bulan dan Tahun pada Jaccard Similarity.....	50
4.4.2.3. Analisis Konsistensi Per Bulan dan Tahun pada Kombinasi Apriori dan Jaccard Similarity	52
4.5. Perbandingan.....	555
4.5.1. Perbandingan Konsistensi Association Rule Per Bulan Dalam Tahun Berbeda.....	555
4.5.2. Perbandingan Konsistensi Association Rule Per Bulan Dan Tahun	566
BAB V PENUTUP.....	599
5.1 Kesimpulan	599
5.2 Saran	599
DAFTAR PUSTAKA	60
LAMPIRAN 1	62

DAFTAR TABEL

Tabel 2.1 Contoh <i>Market Basket Data</i>	8
Tabel 2.2 <i>Binary Representation of Market Basket Data</i> (Tan, 2006)	9
Tabel 2.3 Perbandingan Penelitian.....	15
Tabel 4.1 Rincian Jumlah Data Transaksional.....	22
Tabel 4.2 Contoh Hasil <i>Data Selection</i>	23
Tabel 4.3 Contoh Hasil <i>Data Cleaning</i>	24
Tabel 4.4 Contoh <i>Data Transformation</i>	24
Tabel 4.5 Rincian Jumlah Data Hasil <i>Preprocessing</i>	25
Tabel 4.6 Rincian Kategori Buku Hasil <i>Preprocessing</i>	25
Tabel 4.7 Hasil Penghitungan <i>Similarity</i> Antar <i>Association Rule</i> pada Tahun yang Berbeda Menggunakan <i>Apriori</i>	41
Tabel 4.8 <i>Association Rule</i> yang Konsisten Ditemukan Setiap Tahun dalam Bulan yang Sama Menggunakan <i>Apriori</i>	43
Tabel 4.9 Hasil Penghitungan <i>Similarity</i> Antar <i>Association Rule</i> pada Tahun yang Berbeda Menggunakan <i>Jaccard Similarity</i>	44
Tabel 4.10 <i>Association Rule</i> yang Konsisten Ditemukan Setiap Tahun dalam Bulan yang Sama Menggunakan <i>Jaccard Similarity</i>	45
Tabel 4.11 Hasil Penghitungan <i>Similarity</i> Antar <i>Association Rule</i> pada Tahun yang Berbeda Menggunakan Kombinasi <i>Apriori</i> dan <i>Jaccard Similarity</i>	46
Tabel 4.12 <i>Association Rule</i> yang Konsisten Ditemukan Setiap Tahun dalam Bulan yang Sama Menggunakan Kombinasi <i>Apriori</i> dan <i>Jaccard Similarity</i>	47
Tabel 4.13 Hasil Penghitungan <i>Similarity</i> Antar <i>Association Rule</i> Menggunakan <i>Apriori</i>	48
Tabel 4.14 Hasil Penghitungan <i>Similarity</i> Antar <i>Association Rule</i> Menggunakan <i>Jaccard Similarity</i>	51
Tabel 4.15 Hasil Penghitungan <i>Similarity</i> Antar <i>Association Rule</i> Menggunakan Kombinasi <i>Apriori</i> dan <i>Jaccard</i>	53

Tabel 4.16 Perbandingan Konsistensi *Association Rule* Per Bulan Dalam Tahun Berbeda.....55

Tabel 4.17 Perbandingan Konsistensi *Association Rule* Per Bulan dan Tahun 57



DAFTAR GAMBAR

Gambar 2.1 <i>Knowledge Discovery</i> (Han & Kamber, 2006)	6
Gambar 2.2 Ilustrasi Prinsip Apriori (Tan, 2005).....	11
Gambar 3.1 Metodologi Penelitian.....	16
Gambar 3.2 <i>Preprocessing</i>	17
Gambar 3.3 Kombinasi <i>Apriori</i> dan <i>Jaccard Similarity</i>	19
Gambar 4.1 Desain Database.....	23
Gambar 4.2 Efek Perubahan Nilai <i>Minsup</i> Terhadap Nilai <i>Support</i>	26
Gambar 4.3 Efek Perubahan Nilai <i>Minsup</i> Terhadap Jumlah Itemset.....	27
Gambar 4.4 Efek Perubahan Nilai <i>Minsup</i> Terhadap Jumlah Itemset (Skala Besar)	27
Gambar 4.5 Efek Perubahan Nilai <i>Minconf</i> Terhadap Nilai <i>Confidence</i>	28
Gambar 4.6 Efek Perubahan Nilai <i>Minconf</i> Terhadap Jumlah <i>Association Rule</i>	28
Gambar 4.7 <i>Association Rule</i> Hasil Pengolahan Menggunakan <i>Apriori</i>	29
Gambar 4.8 Jumlah <i>Frequent Itemset</i> dan <i>Association Rule</i> Dihasilkan Setiap Bulan Menggunakan <i>Apriori</i>	30
Gambar 4.9 Efek Perubahan Nilai <i>Minsim</i> Terhadap Nilai <i>Jaccard Similarity Coefficient</i>	32
Gambar 4.10 Efek Perubahan Nilai <i>Minsim</i> Terhadap Jumlah <i>Association Rule</i> yang Dihasilkan	33
Gambar 4.11 Efek Perubahan Nilai <i>Minsim</i> Terhadap Jumlah <i>Association Rule</i> yang Dihasilkan (Skala Besar).....	33
Gambar 4.12 <i>Association Rule</i> Hasil Pengolahan Menggunakan <i>Jaccard Similarity</i>	34
Gambar 4.13 Jumlah <i>Association Rule</i> yang Dihasilkan Setiap Bulan Menggunakan <i>Jaccard Similarity</i>	36
Gambar 4.14 Efek Perubahan Nilai <i>Minsim</i> Terhadap Nilai <i>Jaccard Similarity Coefficient</i> Pada Metode Kombinasi	38
Gambar 4.15 Efek Perubahan Nilai <i>Minsim</i> Terhadap Jumlah <i>Association Rule</i> yang Dihasilkan Pada Metode Kombinasi.....	38

Gambar 4.16 *Association Rule* Hasil Pengolahan Menggunakan Kombinasi *Apriori* dan *Jaccard Similarity* 39

Gambar 4.17 Jumlah *Association Rule* yang Dihasilkan Setiap Bulan Menggunakan Kombinasi *Apriori* dan *Jaccard Similarit*..... 40



DAFTAR LAMPIRAN

LAMPIRAN 1 Contoh Data Transaksi.....62

