

**KAJIAN BELANJA KONSUMEN PADA SITUS
E-COMMERCE DENGAN METODE DATA MINING
BERBASIS ASOSIASI:
STUDI KASUS PESANDELIVERY.COM**

TESIS



Oleh:

**Ajang Sopandi
1111600241**

**PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

JAKARTA

2013

**KAJIAN BELANJA KONSUMEN PADA SITUS
E-COMMERCE DENGAN METODE DATA MINING
BERBASIS ASOSIASI:
STUDI KASUS PESANDELIVERY.COM**

TESIS

**Diajukan sebagai salah satu persyaratan
memperoleh gelar Magister Ilmu Komputer (M.Kom)**



**Oleh:
Ajang Sopandi
1111600241**

**PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

JAKARTA

2013



PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PERNYATAAN

Nama Mahasiswa : Ajang Sopandi
Nomor Induk Mahasiswa : 1111600241
Konsentrasi : Teknologi Sistem Informasi
Program : Pascasarjana

Menyatakan bahwa TESIS yang berjudul:

Kajian Belanja Konsumen pada Situs E-Commerce dengan Metode Data Mining Berbasis Asosiasi: Studi Kasus PesanDelivery.com

1. Merupakan hasil karya tulis ilmiah sendiri dan bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik oleh pihak lain.
2. Saya ijin untuk dikelola oleh Universitas Budi Luhur sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan penuh tanggung jawab dan saya siap menerima konsekuensi apapun sesuai aturan yang berlaku, apabila dikemudian hari pernyataan ini tidak benar.

Jakarta, 27 Agustus 2013





PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR

LEMBAR PENGESAHAN

Nama Mahasiswa : Ajang Sopandi
Nomor Induk Mahasiswa : 1111600241
Konsentrasi : Teknologi Sistem Informasi
Judul Tesis : Kajian Belanja Konsumen pada Situs E-Commerce
dengan Metode Data Mining Berbasis Asosiasi: Studi
Kasus PesanDelivery.com

Jakarta, 27 Agustus 2013

Tim Penguji:

Tanda Tangan:

Ketua,

Dr., Ir. Nazori A.Z., M.T.

Anggota,

Ir. Wendi Usino, M.Sc., M.M.

Pembimbing,

Dr. Ir. Prabowo Pudjo Widodo, MS.

Ketua Program Studi,

Dr. Ir. Nazori A.Z., M.T.

ABSTRAK

PesanDelivery adalah perusahaan kurir yang berbasis *e-commerce*, yang akan mendapatkan keuntungan dari setiap pengiriman barang yang dipesan oleh konsumen melalui situs PesanDelivery.com. Dalam upaya untuk meningkatkan order pesanan barang, penting bagi pihak manajemen mengetahui produk apa saja yang paling banyak dipesan oleh konsumen. Pada situs *e-commerce* yang setiap harinya terjadi transaksi penjualan, diperlukan sebuah aplikasi yang mampu menyaring data sehingga diperoleh informasi yang bermanfaat dengan mengimplementasikan data mining. Data mining dibutuhkan untuk memberikan solusi nyata bagi para pengambil keputusan dalam mengembangkan bisnis mereka. Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database atau sering disebut *Knowledge Discovery in Database (KDD)*. Teknik yang digunakan dalam aplikasi data mining ini adalah Aturan Asosiasi (*Association Rule*) dengan Algoritma *Apriori* dan *FP-Growth*. Algoritma *Apriori* adalah Algoritma Asosiasi yang paling mudah dipahami dan diimplementasikan namun memiliki kekurangan. Saat dibandingkan dengan algoritma lainnya dengan proses *association rule* untuk melakukan proses pencarian *frequent itemset*, Algoritma *Apriori* harus melakukan *scanning database* berulang kali untuk setiap kombinasi *item*, berakibat pada terlalu banyaknya waktu dan sumber daya sistem yang harus terbuang. Selain itu, dibutuhkan *generate candidate* yang besar untuk mendapatkan kombinasi *item* dari database. Untuk mengatasi hal tersebut maka digunakanlah Algoritma *FP-Growth*. Dalam penelitian ini dibahas penerapan Algoritma *Apriori* dan *FP-Growth* dalam proses pencarian *frequent itemsets*, sehingga dihasilkan *3-itemset* produk makanan dan minuman yang dibeli secara bersamaan dalam satu kali transaksi.

Kata Kunci: *E-Commerce, Data Mining, Aturan Asosiasi, Algoritma Apriori, Algoritma FP-Growth, Knowledge Discovery in Database (KDD).*

ABSTRACT

PesanDelivery is a e-commerce based courier company which benefited from order shipments by consumers through PesanDelivery.com website. In the effort to improving its orders, it is important for the management to know what products that is ordered most by the consumer. In most e-commerce websites that hold daily sale transactions, it is necessary to have an application that is able to filter the useful informations from sales and transactions data by data mining. Data mining is required to provide a real solution for decision makers in business development. Data mining is a term to describe the knowledge discovery in databases (KDD) process. Techniques used in this research is the Association Rules with Apriori and FP-Growth algorithm. Apriori is the most widely known and common algorithm, but has some shortcomings. When compared with other algorithms to perform frequent itemset search, Apriori must repeatedly scan the database for each item combination, resulting in too much time and system resources to be wasted in just one process. In addition, it needs a large generate candidate to get a item combinations from the database. To overcome this shortcomings of Apriori, FP-Growth is used. This research discussed the implementation of Apriori and FP-Growth algorithm in the process of finding frequent itemsets, resulting in 3-itemsets of food and beverages which simultaneously purchased in a single transactions.

Keywords: *E-Commerce, Data Mining, Association Rule, Apriori Algorithm, FP-Growth Algorithm, Knowledge Discovery In Database (KDD).*

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya penulis bisa menyelesaikan penelitian tesis yang berjudul *Kajian Belanja Konsumen pada Situs E-Commerce dengan Metode Data Mining Berbasis Asosiasi: Studi Kasus PesanDelivery.com*. Salawat serta salam penulis haturkan kepada Rasulullah Muhammad SAW. Tujuan dari penulisan penelitian tesis ini adalah sebagai salah satu syarat untuk menyusun tesis pada Program Studi Magister Ilmu Komputer, Universitas Budi Luhur Jakarta.

Rasa dan ucapan terima kasih penulis persembahkan kepada semua pihak yang telah membantu penulis dalam menyusun penelitian tesis ini:

1. Bapak Dr. Ir. Prabowo Pudjowidodo, MS., selaku dosen pembimbing tesis yang telah membimbing dan memotivasi penulis dalam mengerjakan penelitian tesis ini.
2. Bapak Dr. Moedjiono, M.Sc., selaku Dekan program studi Magister Ilmu KomputerProgram Pascasarjana di Universitas Budi Luhur yang telah memberikan kesempatan kepada penulis untuk belajar dan menyelesaikan Tesis.
3. Bapak Dr. Ir. Nazori Az., MT., selaku Ketua Jurusan Program Studi Magister Ilmu KomputerProgram Pascasarjana di Universitas Budi Luhur yang telah memberikan dukungan dalam penyelesaian Tesis.
4. Seluruh keluarga besar PesanDelivery.com, yang telah memberikan izin penelitian dan membantu penulis dalam mengumpulkan data dan informasi yang diperlukan terkait penelitian tesis ini.
5. Herlinah, SE., istri tercinta yang selalu memberikan semangat, dukungan, dan do'a bagi penulis.
6. Qaireen Sabrilla Azzahya, anak dan putri pertama penulis yang baru dilahirkan tepat pada hari kemerdekaan 17 Agustus 2013.
7. Kedua orangtua dan juga kedua mertua penulis yang telah memberikan dukungan dan do'anya.

8. Kakak, Dendi Supriadi dan Nana Suryana, M.Pd, yang telah memberikan inspirasi bagi penulis untuk melanjutkan kuliah di Strata Dua (S2).
9. Rekan-rekan mahasiswa MKOM Universitas Budi Luhur kelas XA Semester 1 dan 2, dan MKOM Semester 3 konsentrasi Teknologi Sistem Informasi, terima kasih atas kebersamaan, kerja keras, dan dukungan serta semangatnya.
10. Bapak dan Ibu Dosen pengajar mata kuliah di Program Studi Magister Ilmu Komputer Universitas Budi Luhur yang telah dengan sabar memberikan ilmu pengetahuan, pencerahan, dan bimbingan dalam belajar.

Penulis menyadari, sebagai makhluk Tuhan yang jauh dari kesempurnaan, masih banyak kekurangan dari penelitian tesis ini, oleh karena itu penulis mengharapkan kritik dan saran yang bersifat membangun demi kesempurnaan penelitian tesis nantinya. Semoga penelitian tesis ini masih dapat memberikan manfaat dari keterbatasannya. *Amiin ya Rabb.*

Jakarta, 27 Agustus 2013

Ajang Sopandi

HALAMAN PERSEMBAHAN

kususun jari jemari ku diatas *keyboard laptop* ku sebagai pembuka kalimat persembahan ku, diikuti dengan “Bismillahirrahmanirrahim” sebagai awal setiap memulai pekerjaanku.

sembah sujud serta syukur ku persembahkan kepada Allah SWT. taburan cinta dan kasih sayang-Mu telah memberikanku kekuatan, membekaliku dengan ilmu serta memperkenalkanku dengan cinta.

atas karunia serta kemudahan yang Engkau berikan, akhirnya TESIS ini dapat terselesaikan.

sholawat dan salam selalu terlimpahkan keharibaan Rasulullah Muhammad SAW, sebagaimana sabdamu ya Rasul:
*“Barang siapa yang menginginkan dunia, hendaklah ia berilmu,
Barang siapa yang menginginkan akhirat hendaklah ia berilmu,
Barang siapa yang menginginkan kedua sekaligus, ia pun harus berilmu.”*
Telah menginspirasi umatmu ini dalam menuntut ilmu.

~ Kupersembahkan untuk orang tua, saudara, keluarga besar, teman,
istri dan anakku tercinta. ~

DAFTAR ISI

	Halaman
ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
HALAMAN PERSEMBAHAN	v
DAFTAR ISI	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	3
1.2.1 Identifikasi Masalah	3
1.2.2 Pembatasan Masalah	3
1.2.3 Rumusan Masalah	4
1.3 Tujuan dan Manfaat Penelitian	4
1.4 Tata Urut Penulisan	5
1.5 Daftar Pengertian	5
BAB II LANDASAN TEORI DAN KERANGKA KONSEP	8
2.1 Tinjauan Pustaka	8
2.1.1 Data Mining	8
2.1.1.1 Tahapan Data Mining	10
2.1.1.2 Arsitektur Data Mining	11
2.1.1.3 KDD (<i>Knowledge Discovery in Databases</i>)	13
2.1.1.4 <i>Market Basket Analysis</i>	14
2.1.2 Aturan Asosiasi (<i>Association Rules</i>)	15
2.1.2.1 Bentuk Umum Aturan Asosiasi	15
2.1.2.2 Pembentukan Aturan Asosiasi	16
2.1.3 Algoritma <i>Apriori</i>	16
2.1.4 Algoritma <i>FP-Growth</i> (<i>Frequent Pattern-Growth</i>)	18
2.1.5 Software Quality Assurance (SQA)	24
2.2 Tinjauan Studi	27
2.2.1 Model Penelitian Cristidist	27
2.2.2 Model Penelitian Cavique	29
2.2.3 Model Penelitian Yu Li, Zuo Meiyun dan Bo Yang	30
2.2.4 Model Penelitian Goldie Gunadi, dan & Indra Sensuse	31
2.2.5 Rangkuman Penelitian Terkait	32
2.3 Tinjauan Objek Penelitian	33
2.3.1. Aspek Organisasi	33
2.3.1.1 Lokasi Penelitian	33
2.3.1.2 Gambaran Umum Perusahaan	33
2.3.1.3 Bisnis Proses Perusahaan	34
2.3.1.4 Visi dan Misi Perusahaan	35
2.3.2. Aspek Sistem	36
2.3.2.1. Hardware	36

	2.3.2.2. Software.....	37
	2.3.2.3. Data.....	37
	2.4 Kerangka Konsep	37
	2.5 Hipotesis	38
BAB III	METODOLOGI DAN RANCANGAN PENELITIAN	39
	3.1 Metode Penelitian.....	39
	3.2 Metode Pemilihan Sampel.....	39
	3.3 Metode Pengumpulan Data	40
	3.4 Instrumentasi	41
	3.5 Langkah-Langkah Penelitian.....	41
	3.5.1 <i>Domain Understanding and KDD Goals</i>	41
	3.5.2 <i>Selection and Addition</i>	42
	3.5.3 <i>Preprocessing and Data Cleaning</i>	44
	3.5.4 Transformation	45
	3.5.5 Pemilihan Algoritma Asosiasi	46
	3.5.6 Penambangan Data (Data Mining)	47
	3.5.7 Evaluasi dan Analisis.....	52
	3.5.8 Pemilihan Algoritma Terbaik	54
	3.5.9 <i>Graphical User Interface (GUI)</i>	54
	3.5.10 <i>Software Quality Assurance (SQA)</i>	54
	3.6 Jadwal Penelitian.....	56
BAB IV	PEMBAHASAN HASIL PENELITIAN	58
	4.1 Hasil Penelitian.....	58
	4.1.1 Perhitungan Algoritma <i>Apriori</i>	59
	4.1.2 Perhitungan Algoritma FP-Growth.....	66
	4.2 Evaluasi	81
	4.2.1 Evaluasi Aturan Asosiasi Algoritma Apriori.....	81
	4.2.2 Evaluasi Aturan Asosiasi Algoritma FP-Growth	82
	4.3 Analisis.....	82
	4.4 Penerapan Algoritma Terpilih.....	84
	4.4.1 Perancangan <i>Graphical User Interface (GUI)</i>	84
	4.4.1.1. Overview	84
	4.4.1.2. Preprocessing.....	86
	4.4.1.3. Flowchart.....	87
	4.4.1.4. Pembuatan <i>Graphical User Interface (GUI)</i> ...	87
	4.4.2 <i>Software Quality Assurance (SQA)</i>	88
	4.5 Implikasi Penelitian	89
	4.5.1 Aspek Sistem	90
	4.5.2 Aspek Manajerial.....	91
	4.5.3 Aspek Penelitian Lanjutan.....	91
BAB V	PENUTUP	92
	5.1 Kesimpulan.....	92
	5.2 Saran.....	92
	DAFTAR PUSTAKA	93
	LAMPIRAN-LAMPIRAN.....	96
	DAFTAR RIWAYAT HIDUP.....	119

DAFTAR GAMBAR

Gambar	Halaman
I-1 <i>Homescreen</i> website <i>e-commerce</i> PesanDelivery.com	2
II-1 Tahapan Data Mining [(Han, Jiawei, and Kamber, 2006), 9]	10
II-2 Arsitektur Data Mining [(Han, Jiawei, and Kamber 2006), 9]	13
II-3 Proses KDD (Knowledge Discovery in Databases) [(Maimon, 2010), 3]	13
II-4 Bentuk umum Aturan Asosiasi [Pathrecia, 2012), 15]	16
II-5 Transaksi 1: BEAD	20
II-6 Transaksi 2: BEC	21
II-7 Transaksi 3: BEAD	21
II-8 Transaksi 4: BEAC	22
II-9 Transaksi 5: BEACD	22
II-10 Transaksi 6: BCD	23
II-11 <i>frequent itemset</i> {DAE, DAEB, DAB, DEB, CB, CEB, CB, AE, AEB, AB, EB}	24
II-12 Bagian-bagian dari SQA [(Pressman, 2001), 1]	25
II-13 Model yang diusulkan Christidis	28
II-14 Model yang diusulkan Cavique	30
II-15 Model yang diusulkan Yu Li, Zuo mei dan Bo Yang	31
II-16 Proses order pesandelivery.com	35
II-17 Kerangka Konsep Pemecahan Masalah	38
III-1 Model yang diusulkan	46
III-2 Membuat proses baru pada RapidMiner	48
III-3 Menentukan tempat penyimpanan data	48
III-4 Menentukan dataset yang ingin digunakan	49
III-5 Pemberian tipe data atribut pada saat <i>import</i> data	49
III-6 Proses <i>pivot</i> data	50
III-7 Hasil <i>pivot</i> data	50
III-8 Proses <i>fp-growth</i>	51
III-9 Hasil <i>fp-growth</i>	51
III-10 Hasil <i>association rules</i> dengan algoritma <i>fp-growth</i>	52
IV-1. Pola-pola asosiasi yang terbentuk dengan menggunakan Algoritma <i>Apriori</i>	66
IV-2. Hasil pembentukan <i>FP-tree</i> setelah TID 1	71
IV-3. Hasil pembentukan <i>FP-tree</i> setelah TID 2	72
IV-4. Hasil pembentukan <i>FP-tree</i> setelah TID 3	72
IV-5. Hasil pembentukan <i>FP-tree</i> setelah TID 4	73
IV-6. Hasil pembentukan <i>FP-tree</i> setelah TID 5	73
IV-7. Hasil pembentukan <i>FP-tree</i> setelah TID 6	74
IV-8. Hasil pembentukan <i>FP-tree</i> setelah TID 7	75
IV-9. Hasil pembentukan <i>FP-tree</i> setelah TID 8	75
IV-10. Hasil pembentukan <i>FP-tree</i> setelah TID 9	76
IV-11. Hasil pembentukan <i>FP-tree</i> setelah TID 10	77
IV-12. <i>FP-Tree</i> untuk P8	78

IV-13. <i>FP-Tree</i> untuk P6	79
IV-14. Pola-pola asosiasi yang terbentuk dengan menggunakan Algoritma <i>FP-Growth</i>	81
IV-15. <i>Overview</i> Aplikasi <i>FP-Growth</i>	85
IV-16. Tahapan <i>Preprocessing</i> Aplikasi <i>FP-Growth</i>	86
IV-17. <i>Flowchart</i> Aplikasi <i>FP-Growth</i>	87
IV-18. GUI Aplikasi data mining berbasis Asosiasi dengan Algoritma <i>FP-Growth</i>	88



DAFTAR TABEL

Tabel	Halaman
II-1 Contoh data transaksi	19
II-2 Hasil dari semua metode	29
II-3 Perbandingan penelitian terkait	32
II-4 Perbandingan penelitian terkait dengan penelitian penulis	33
II-5 Spesifikasi komputer yang digunakan	36
III-1 Tabel Objek Penelitian	42
III-2 Data Awal	43
III-3 Data setelah diseleksi atribut	44
III-4 Spesifikasi komputer yang digunakan	47
III-5 <i>Metric of Software Quality Assurance (SQA)</i>	55
III-6 Hasil Evaluasi SQA	56
III-7 Jadwal Penelitian	57
IV-1 Data transaksi dengan persentase kemunculan paling tinggi pada algoritma <i>Apriori</i>	59
IV-2 Keterangan nama produk	60
IV-3 Keterangan kode produk untuk perhitungan <i>Apriori</i>	60
IV-4 10 dataset transaksi pada algoritma <i>Apriori</i>	61
IV-5 Binari tabel transaksi algoritma <i>Apriori</i>	62
IV-6 Data <i>frekuent 1-Itemset</i>	62
IV-7 Data <i>frekuent 2-Itemset</i>	63
IV-8 Data <i>frekuent 3-Itemset</i>	63
IV-9 Pendekatan Asosiasi yang terbentuk dari data <i>frekuent 3-Itemset</i>	64
IV-10 <i>Rule</i> dengan kandidat kemunculan paling sering	65
IV-11 10 data transaksi dengan persentase kemunculan paling tinggi pada algoritma <i>FP-Growth</i>	67
IV-12 Keterangan nama produk	67
IV-13 Keterangan kode produk untuk perhitungan <i>FP-Growth</i>	68
IV-14 10 dataset transaksi pada algoritma <i>FP-Growth</i>	68
IV-15 pengurutan data transaksi berdasarkan nilai <i>Support Count</i>	69
IV-16 Pengurutan data transaksi berdasarkan <i>minimum support count</i> ≤ 2	70
IV-17 Pengurutan <i>itemset</i> tiap transaksi berdasarkan frekuensi paling tinggi	70
IV-18 Evaluasi aturan <i>Apriori</i>	81
IV-19 Evaluasi aturan <i>FP-Growth</i>	82
IV-20 Analisa aturan <i>Apriori</i>	83
IV-21 Analisa aturan <i>FP-Growth</i>	83
IV-22 <i>Metric of Software Quality Assurance (SQA)</i>	88
IV-23 Hasil Evaluasi SQA	89
IV-24 Spesifikasi Hardware dan Software	90

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Surat Keterangan melakukan penelitian	96
2. Pengujian SQA (<i>Software Quality Assurance</i>) oleh Direktur PesanDelivery.com	97
3. Pengujian SQA (<i>Software Quality Assurance</i>) oleh Operator PesanDelivery.com	98
4. Pengujian SQA (<i>Software Quality Assurance</i>) oleh Administrator PesanDelivery.com	99
5. Pengujian SQA (<i>Software Quality Assurance</i>) oleh Customer PesanDelivery.com	100
6. Dataset Penelitian.....	101
7. Source Code	110



DAFTAR PUSTAKA

- [Berry 2004] Berry, Michael J.A, Linoff, Gordon S. *Data Mining Techniques For Marketing, Sales, Customer Relationship Management Second Editon*. United States of America: Wiley Publishing, Inc., 2004.
- [Berndtssom 2008] Berndtssom, M., Hansson, J., Olsson, B., Lundell, B., *Planning and Implementing your Final Year Project - with Success!: A Guide for Students in Computer Science and Information Systems*. London: Springer, 2008.
- [Bounsaythip2001] Bounsaythip, C., Rinta-Runsala, E. *Overview of Data Mining for Customer Behavior Modeling*. Research Report TTE1, 18, 2001.
- [Bramer 2007] Bramer, Max. *Principles of Data Mining*. London: Springer, 2007.
- [Cavique 2007] Cavique, L. *A Network Algorithm to Discover Sequential Patterns.ESCS-IPL, Portugal*, 2007.
- [Cho 2004] Cho, Y. H., Kim, J. K. *Application of Web usage mining and product taxonomy to collaborative recommendations in e-commerce*. Expert system with applications, 26, 2004.
- [Christidist2010] Christidist, K., Apostolou, D., and Mentzas, G. *Exploring Customer Preferences with Probabilistic Topics Models*. National Technical Univesity of Athens, 2010.
- [Dawson 2009] Dawson, Catherine. *Practical Research Method*. Oxford: HowtoBook, 2009.
- [Fenriana 2012] Fenriana, Indah. *Penerapan Algoritma K-Means dan Fp-Growth untuk Analisis Keranjang Pasar pada Penjualan Produk Alumunium*. Tesis Tidak Terpublikasi. Jakarta: MKOM STMIK Eresha, 2012.
- [Gunadi 2012] Gunadi, Goldie,dan Dana Indra Sensuse. *Penerapan Metode Data Mining Market Basket Analysis terhadap Data Penjualan Produk Bukudengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (fp-growth): Sudi Kasus Percetakan PT.Gramedia*. Jurnal Telematika MKom Vol.4 No.1, Maret 2012.
- [Han 2000] Han, Jiawei dan Micheline Kamber, *Data Mining: Concepts and Techniques : Chapter 6. Mining Association Rules in Large Databases*. Simon Fraser University, 2000.