

**IMPLEMENTASI ALGORITMA RULEGROWTH UNTUK
MEMPREDIKSI URUTAN PENJUALAN BARANG DARI
DATA PENJUALAN BARANG PUSAT PERKULAKAN**

Skripsi



oleh

**MICHAEL ROYALDO
71130086**

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

2017

**IMPLEMENTASI ALGORITMA RULEGROWTH UNTUK
MEMPREDIKSI URUTAN PENJUALAN BARANG DARI
DATA PENJUALAN BARANG PUSAT PERKULAKAN**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana
Sebagai Salah Satu Syarat dalam Memperoleh Gelar
Sarjana Komputer

Disusun oleh :

MICHAEL ROYNALDO
71130086

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS KRISTEN DUTA WACANA

2017

PERNYATAAN KEASLIAN SKRIPSI

PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

IMPLEMENTASI ALGORITMA RULEGROWTH UNTUK MEMPREDIKSI URUTAN PENJUALAN BARANG DARI DATA PENJUALAN BARANG PUSAT PERKULAKAN

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 pencabutan gelar keserjanaan saya.

Yogyakarta, 11 Juni 2017



MICHAEL ROYNALDO
71130086

HALAMAN PERSETUJUAN

HALAMAN PERSETUJUAN

Judul Skripsi : IMPLEMENTASI ALGORITMA RULEGROWTH
UNTUK MEMPREDIKSI URUTAN PENJUALAN
BARANG DARI DATA PENJUALAN BARANG
PUSAT PERKULAKAN

Nama Mahasiswa : MICHAEL ROYNALDO

N I M : 71130086

Matakuliah : Skripsi (Tugas Akhir)

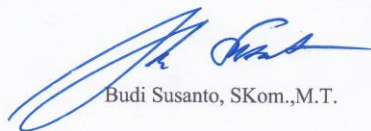
Kode : TIW276

Semester : Genap

Tahun Akademik : 2016/2017

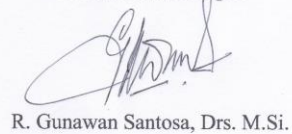
Telah diperiksa dan disetujui di
Yogyakarta,
Pada tanggal 9 Mei 2017

Dosen Pembimbing I



Budi Susanto, SKom.,M.T.

Dosen Pembimbing II



R. Gunawan Santosa, Drs. M.Si.

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

IMPLEMENTASI ALGORITMA RULEGROWTH UNTUK MEMPREDIKSI URUTAN PENJUALAN BARANG DARI DATA PENJUALAN BARANG PUSAT PERKULAKAN

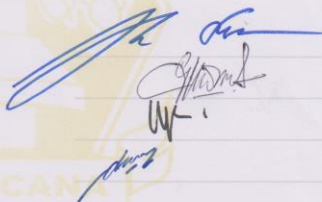
Oleh: MICHAEL ROYNALDO / 71130086

Dipertahankan di depan Dewan Penguji Skripsi
Program Studi Teknik Informatika Fakultas Teknologi Informasi
Universitas Kristen Duta Wacana - Yogyakarta
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar
Sarjana Komputer
pada tanggal 29 Mei 2017

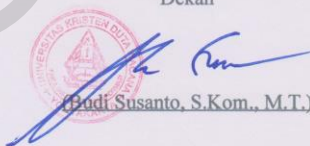
Yogyakarta, 11 Juni 2017
Mengesahkan,

Dewan Penguji:

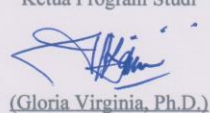
1. Budi Susanto, SKom., M.T.
2. R. Gunawan Santosa, Drs. M.Si.
3. Rosa Delima, S.Kom., M.Kom.
4. Danny Sebastian, S.Kom., M.M., M.T.



Dekan


(Budi Susanto, S.Kom., M.T.)

Ketua Program Studi


(Gloria Virginia, Ph.D.)

UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan kepada Tuhan yang Maha Esa atas berkat-Nya. Karena-Nya skripsi dengan judul “Implementasi Algoritma RuleGrowth Untuk Memprediksi Urutan Penjualan Barang Dari Data Penjualan Barang Pusat Perkulakan” dapat penulis selesaikan.

Penulis menyusun skripsi ini dalam rangka memenuhi salah satu persyaratan untuk mencapai gelar sarjana pada Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Duta Wacana, Yogyakarta.

Penulis menyadari bahwa terselesaikannya skripsi ini tak lepas dari campur tangan berbagai pihak. Oleh karena itu penulis ingin menyampaikan terima kasih kepada:

1. Bapak Budi Susanto, S.Kom., M.T., selaku dosen pembimbing I yang telah banyak memberikan bimbingan, saran dan kritik selama pembuatan sistem dan penulisan skripsi ini.
2. Bapak Drs. R. Gunawan, M.Si., selaku dosen pembimbing II yang juga telah memberikan masukan dan arahan yang membantu penulis dalam pembuatan skripsi.
3. Kepada keluarga terkasih, Bapak, Ibu, dan adik saya yang selalu memberikan dukungan, doa, nasehat dan motivasi kepada penulis dalam menyelesaikan studi.
4. Kepada Jeanette Aprillya yang selalu memberikan semangat dan cinta serta setia menemani penulis dalam menyelesaikan skripsi.
5. Kepada Bima Haryanto Putra, Ervan Kurnia Sanjaya dan Nico Raynaldo yang merupakan rekan satu tim dalam mengerjakan skripsi.

6. Kepada teman-teman PPLK yang selalu memberikan semangat dan hiburan selama penulis mengalami kesulitan dalam pembuatan skripsi.
7. Kepada teman-teman jurusan Teknik Informatika 2013, yang sama-sama berjuang dalam menyelesaikan tugas akhir.
8. Terakhir, penulis hendak berterimakasih pada setiap nama yang tidak dapat penulis catumkan satu per satu.

Tentunya penulis masih memiliki banyak kekurangan dalam melakukan penulisan skripsi ini. Oleh karena itu, penulis sangat menghargai jika ada berbagai masukan dari para pembaca baik berupa kritik maupun saran yang dapat membangun penulis dalam penyempurnaan penulisan-penulisan skripsi di masa yang akan datang. Penulis juga meminta maaf bila terdapat kesalahan dalam penulisan skripsi ini.

Terima Kasih.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan yang Maha Esa atas berkat dan rahmat-Nya, sehingga penulis dapat menyelesaikan tugas akhir ini.

Penulis tentu sangat berterima kasih kepada semua pihak atas semua dukungan yang telah diberikan kepada penulis dalam menyelesaikan tugas akhir ini. Penulis juga tak lupa ingin meminta maaf bila terdapat kesalahan dalam penulisan tugas akhir ini.

Penulis berharap, para pembaca dapat mendapatkan informasi-informasi yang dapat berguna bagi para pembaca sekalian.

Yogyakarta, 11 Mei 2017

Penulis

INTISARI

IMPLEMENTASI ALGORITMA RULEGROWTH UNTUK MEMPREDIKSI URUTAN PENJUALAN BARANG DARI DATA PENJUALAN BARANG PUSAT PERKULAKAN

RuleGrowth adalah salah satu algoritma yang dapat digunakan untuk mencari aturan-aturan dari data transaksi penjualan yang sudah dikelompokkan berdasarkan member. Dalam penelitian ini, penulis menggunakan algoritma *RuleGrowth* pada data transaksi Grosir X dari Juni 2014 - Agustus 2016. Pengujian pada penelitian ini menggunakan nilai *minimum support*, *minimum confidence* dan *lift ratio*. Data transaksi tahun 2016 akan digunakan sebagai data uji dan data latihnya adalah data transaksi tahun 2015. Aturan-aturan yang dihasilkan dari data latih dan data uji akan dibandingkan untuk mencari aturan yang berulang pada tahun 2015 - 2016, aturan yang hanya muncul pada tahun 2015 dan aturan yang baru muncul pada tahun 2016. Dan hasilnya, pada periode waktu Januari - April didapatkan jumlah aturan yang berulang pada tahun 2015 - 2016 sebanyak 136, jumlah aturan yang hanya muncul tahun 2015 sebanyak 61 dan jumlah yang baru muncul tahun 2016 sebanyak 669. Sedangkan untuk periode waktu Mei - Agustus didapatkan jumlah aturan yang berulang pada tahun 2015 - 2016 sebanyak 60, jumlah aturan yang hanya muncul tahun 2015 sebanyak 105 dan jumlah yang baru muncul tahun 2016 sebanyak 40. Hasil tersebut didapatkan dengan menggunakan nilai *minimum support* 26% dan *minimum confidence* 55%.

Kata kunci: *Sequential Rule Mining, RuleGrowth, Data Mining, Data Warehouse.*

DAFTAR ISI

HALAMAN JUDUL	ii
PERNYATAAN KEASLIAN TUGAS AKHIR	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PENGESAHAN	v
UCAPAN TERIMA KASIH	vi
KATA PENGANTAR	viii
INTISARI	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
BAB I. PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah	3
1.3. Batasan Sistem	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	4
1.6. Metode Penelitian.....	4
1.7. Sistematika Penulisan.....	5
BAB II. TINJAUAN PUSTAKA	7
2.1. Tinjauan Pustaka	7
2.2. Landasan Teori.....	8
2.2.1. Data Mining	8
2.2.2. Sequential Pattern Mining dan Sequential Rule Mining	11
2.2.3. Algoritma RuleGrowth	12
2.2.4. Studi Kasus RuleGrowth.....	17
2.2.5. <i>Data Warehouse</i>	20
2.2.6. Business Intelligence.....	21
2.2.7. SPMF	22
2.2.8. Java Server Pages (JSP)	22

BAB III. ANALISIS DAN PERANCANGAN SISTEM.....	24
3.1. Analisis Kebutuhan	24
3.1.1. Kebutuhan Fungsional	24
3.1.2. Kebutuhan Non Fungsional	27
3.2. Rancangan Environment	28
3.3. Arsitektur Sistem.....	28
3.4. Rancangan <i>Database</i>	29
3.5. Rancangan Proses.....	34
3.5.1. Flowchart Data Cleaning	35
3.5.2. Flowchart Data Filtering	36
3.5.3. Pseudocode RuleGrowth.....	37
3.6. Rancangan Antarmuka	40
3.7. Rancangan Pengujian.....	41
BAB IV. IMPLEMENTASI DAN ANALISIS	44
4.1. Implementasi Sistem	44
4.1.1. Konfigurasi Server	44
4.1.2. Konfigurasi <i>Hostname</i>	45
4.1.3. Konfigurasi <i>IP Address</i>	46
4.1.4. Konfigurasi <i>Hadoop</i>	47
4.1.5. Konfigurasi <i>Zookeeper</i>	52
4.1.6. Konfigurasi <i>HBase</i>	53
4.1.7. Implementasi Proses <i>Extract Transformation Load (ETL)</i>	55
4.1.8. Implementasi Proses <i>Data Cleaning</i>	58
4.1.9. Implementasi Proses <i>Data Filtering</i>	59
4.1.10. Implementasi Algoritma <i>RuleGrowth</i>	59
4.1.11. Implementasi Antarmuka.....	62
4.1.12. Halaman Login.....	62
4.1.13. Halaman Input RuleGrowth	62
4.1.14. Halaman Hasil dari RuleGrowth.....	63
4.1.15. Halaman Input Analisis RuleGrowth	64
4.1.16. Halaman Hasil Analisis RuleGrowth.....	65
4.2. Analisis Sistem.....	66

BAB V. KESIMPULAN DAN SARAN	71
5.1. Kesimpulan.....	71
5.2. Saran.....	72
DAFTAR PUSTAKA	73
LAMPIRAN	75

©UKDWN

DAFTAR GAMBAR

Gambar 2. 1. Tahapan Data Mining.	9
Gambar 2. 2. Contoh Penentuan Rule dari Sequence Database.	12
Gambar 2. 3. Rumus support dan confidence.	13
Gambar 2. 4 Algoritma RuleGrowth.	14
Gambar 2. 5. Fungsi ExpandLeft.	15
Gambar 2. 6. Fungsi ExpandRight.	16
Gambar 2. 7. Arsitektut Data Warehouse	21
Gambar 2. 8. Proses Kerja JSP.	23
Gambar 3. 1. Use Case Diagram.	25
Gambar 3. 2. Rancangan Arsitektur Sistem.	29
Gambar 3. 3. Rancangan skema pemodelan basis data.	34
Gambar 3. 4. Proses alur kerja sistem.	35
Gambar 3. 5. Proses Data Cleaning.	36
Gambar 3. 6. Proses Data Filtering.	37
Gambar 3. 7. Pseudocode RuleGrowth.	38
Gambar 3. 8. Pseudocode ExpandLeft.	39
Gambar 3. 9. Pseudocode ExpandRight.	39
Gambar 3. 10. Tampilan halaman login.	40
Gambar 3. 11. Tampilan halaman utama sistem.	41
Gambar 3. 12. Rumus Lift Ratio.	43
Gambar 4. 1. Topologi client-server.	44
Gambar 4. 2. Struktur Ekosistem Hadoop.	45
Gambar 4. 3. Contoh isi file /etc/hostname.	45
Gambar 4. 4. Pemberian IP static pada Server dan Client.	46
Gambar 4. 5. Contoh isi file /etc/hosts.	46
Gambar 4. 6. Hadoop environment variables.	48
Gambar 4. 7. Konfigurasi pada core-site.xml.	49
Gambar 4. 8. Konfigurasi pada hdfs-site.xml pada master.	49

Gambar 4. 9. Konfigurasi pada hdfs-site.xml pada slave1.	49
Gambar 4. 10. Konfigurasi pada yarn-site.xml (1).	50
Gambar 4. 11. Konfigurasi pada yarn-site.xml (2).	50
Gambar 4. 12. Konfigurasi pada mapred-site.xml.	50
Gambar 4. 13. Proses java yang berjalan pada master dan slave1.	51
Gambar 4. 14. Halaman hadoop namenode.	51
Gambar 4. 15. Halaman manajemen cluster.	52
Gambar 4. 16. HBase environment variables.	53
Gambar 4. 17. Konfigurasi pada hbase-env.sh.	53
Gambar 4. 18. Konfigurasi pada hbase-site.xml.	54
Gambar 4. 19. Proses java setelah HBase dijalankan.	54
Gambar 4. 20. Halaman HBase Master.	55
Gambar 4. 21. Halaman Hbase RegionServer.	55
Gambar 4. 22. Transformasi Tabel Dimensi dan Data Cleaning.	56
Gambar 4. 23. Transformasi Tabel Fakta.	56
Gambar 4. 24. Proses Load Tabel Dimensi dan Fakta.	58
Gambar 4. 25. Halaman Login.	62
Gambar 4. 26. Halaman Input Data untuk RuleGrowth.	62
Gambar 4. 27. Halaman Hasil dari RuleGrowth.	63
Gambar 4. 28. Tab Jenis Barang.	64
Gambar 4. 29. Halaman Input Analisis.	64
Gambar 4. 30. Halaman Hasil Analisis.	65

DAFTAR TABEL

Tabel 2. 1. Contoh <i>Sequence Database</i>	17
Tabel 2. 2. <i>Item</i> hasil langkah 1.	18
Tabel 2. 3. Semua <i>rule</i> yang terbentuk dari <i>item</i> pada tabel 2.2.....	18
Tabel 2. 4. Nilai <i>confidence</i> dari hasil tabel 2.3.	19
Tabel 2. 5. Tabel Rule setelah <i>expandLeft</i> dan <i>expandRight</i>	19
Tabel 2. 6. Nilai <i>confidence</i> dari hasil tabel 2.5.	20
Tabel 2. 7. Hasil akhir <i>Sequential Rule</i>	20
Tabel 3. 1. Tabel <i>tbtr_jualdetail</i>	30
Tabel 3. 2. Tabel <i>tbmaster_prodmast</i>	30
Tabel 3. 3. Tabel <i>tbmaster_customer</i>	30
Tabel 3. 4. Tabel <i>tbmaster_divisi</i>	31
Tabel 3. 5. Tabel <i>tbmaster_departement</i>	31
Tabel 3. 6. Tabel <i>tbmaster_kategori</i>	31
Tabel 3. 7. Dimensi barang.	33
Tabel 3. 8. Dimensi divdepkat.	33
Tabel 3. 9. Dimensi member.	33
Tabel 3. 10. Dimensi waktu.	33
Tabel 3. 11. <i>Sequential Rule</i> dari data latih dan data uji.	42
Tabel 3. 12. <i>Sequential Rule</i> yang sama dari data latih dan data uji.	42
Tabel 4. 1. Pengujian <i>MinSupport</i> 1 Januari - 30 April.....	66
Tabel 4. 2. Pengujian <i>MinSupport</i> 1 Mei - 31 Agustus.	67
Tabel 4. 3. Pengujian <i>MinConfidence</i> 1 Januari - 30 April.	68
Tabel 4. 4. Pengujian <i>MinConfidence</i> 1 Mei - 31 Agustus.....	68
Tabel 4. 5. <i>Sequential Rule</i> 1 Januari - 30 April dan 1 Mei - 31 Agustus... ..	69

INTISARI

IMPLEMENTASI ALGORITMA RULEGROWTH UNTUK MEMPREDIKSI URUTAN PENJUALAN BARANG DARI DATA PENJUALAN BARANG PUSAT PERKULAKAN

RuleGrowth adalah salah satu algoritma yang dapat digunakan untuk mencari aturan-aturan dari data transaksi penjualan yang sudah dikelompokkan berdasarkan member. Dalam penelitian ini, penulis menggunakan algoritma *RuleGrowth* pada data transaksi Grosir X dari Juni 2014 - Agustus 2016. Pengujian pada penelitian ini menggunakan nilai *minimum support*, *minimum confidence* dan *lift ratio*. Data transaksi tahun 2016 akan digunakan sebagai data uji dan data latihnya adalah data transaksi tahun 2015. Aturan-aturan yang dihasilkan dari data latih dan data uji akan dibandingkan untuk mencari aturan yang berulang pada tahun 2015 - 2016, aturan yang hanya muncul pada tahun 2015 dan aturan yang baru muncul pada tahun 2016. Dan hasilnya, pada periode waktu Januari - April didapatkan jumlah aturan yang berulang pada tahun 2015 - 2016 sebanyak 136, jumlah aturan yang hanya muncul tahun 2015 sebanyak 61 dan jumlah yang baru muncul tahun 2016 sebanyak 669. Sedangkan untuk periode waktu Mei - Agustus didapatkan jumlah aturan yang berulang pada tahun 2015 - 2016 sebanyak 60, jumlah aturan yang hanya muncul tahun 2015 sebanyak 105 dan jumlah yang baru muncul tahun 2016 sebanyak 40. Hasil tersebut didapatkan dengan menggunakan nilai *minimum support* 26% dan *minimum confidence* 55%.

Kata kunci: *Sequential Rule Mining, RuleGrowth, Data Mining, Data Warehouse.*

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Dalam kehidupan sehari-hari, kegiatan transaksi jual beli sudah menjadi hal yang wajar untuk dilakukan. Pada umumnya, setiap *data* transaksi jual beli yang terjadi akan disimpan oleh perusahaan untuk menentukan keuntungan atau kerugian perusahaan tersebut. *Data* transaksi juga biasanya digunakan untuk melihat perkembangan perusahaan dari waktu ke waktu. Intinya, *data-data* tersebut akan diolah agar dapat digunakan perusahaan untuk dasar pengambilan keputusan. Misalnya hasil dari *data* tersebut menunjukkan kerugian (margin negatif, uang modal tidak kembali), maka perusahaan pun harus segera mengambil keputusan untuk meningkatkan pemasaran, dan sebagainya. Hal ini menunjukkan bahwa dengan dilakukannya pemanfaatan *data-data* transaksi dengan baik dapat meningkatkan perusahaan tersebut. Dalam penelitian ini, peneliti akan menggunakan *data* transaksi penjualan barang kebutuhan pokok seperti pangan pada sebuah perusahaan perkulakan (grosir), Grosir X.

Grosir X ini adalah salah satu perusahaan pusat perkulakan (grosir) yang cukup banyak di Indonesia. *Data* yang akan digunakan adalah data transaksi penjualan barang di Grosir X Semarang dari tahun Juni 2014 hingga Agustus 2016. *Data-data* tersebut kemudian akan diolah untuk menentukan sebuah aturan pembelian barang oleh member (*sequential rule*) di Grosir X. Tentunya pengolahan *data-data* tersebut tidak dilakukan secara manual. Bila dikerjakan secara manual maka akan dibutuhkan waktu yang sangat lama dan hasilnya pun menjadi tidak relevan karena proses transaksi penjualan barang yang selalu bertambah dari waktu ke waktu. Oleh karena itu, pada penelitian ini akan dibuat sistem yang dapat membantu perusahaan dalam menentukan aturan pembelian member di Grosir X.

Pengolahan *data* transaksi penjualan barang pada Grosir X akan diolah menggunakan proses *data mining*. Proses *data mining* adalah analisis dari *data* yang besar untuk menemukan hubungan yang tidak terduga dan merangkumnya sehingga mudah dimengerti serta berguna untuk pemilik data tersebut (Hand et al., 2001). Dengan dilakukannya proses *data mining*, perusahaan dapat mengolah *data* mentah menjadi sebuah informasi (*knowledge*) yang akan membantu perusahaan dalam mengambil keputusan.

Untuk melakukan proses *data mining* tersebut, diperlukan *Data Warehouse* agar dapat mengolah *data* transaksi penjualan barang yang jumlahnya sangatlah banyak (*Big Data*) dengan efektif dan efisien. Dalam penelitian ini digunakan ekosistem *Hadoop* untuk menampung *Big Data* tersebut. Sedangkan pengambilan *data* (*query*) digunakan bantuan aplikasi *Apache HBase*.

Data-data tersebut akan diproses menggunakan salah satu proses *data mining* yaitu algoritma *sequential rule mining*. Algoritma *sequential rule mining* ini bermacam-macam, namun pada penelitian ini akan digunakan algoritma *RuleGrowth*. Algoritma *RuleGrowth* ini digunakan untuk mencari aturan antar barang (*rule*) dari sebuah *sequence database*. Dalam penelitian ini *sequence database* akan dibedakan berdasarkan member konsumen yang ada pada *data* transaksi penjualan pada Grosir X. Algoritma *RuleGrowth* akan diimplementasikan menggunakan bahasa pemrograman *Java*. Hasil yang didapatkan dari algoritma *RuleGrowth* tersebut kemudian akan dianalisis dan dapat digunakan sebagai bahan pertimbangan perusahaan X dalam mengambil keputusan. Misalnya, diketahui *rule data* penjualan yang menyatakan kebanyakan pembeli akan membeli susu bila sebelumnya membeli keju. Dari hal tersebut perusahaan dapat melihat perilaku konsumen, dan dapat digunakan sebagai dasar mengajukan promosi untuk meningkatkan penjualan, misal membuat promosi pembelian 2 (dua) keju akan mendapatkan 1 (satu) susu gratis dan sebagainya.

1.2. Rumusan Masalah

Rumusan masalah dalam penelitian yang akan dilakukan penulis antara lain:

1. Bagaimana mengimplementasikan algoritma *RuleGrowth* pada data transaksi penjualan Grosir X?
2. Berapakah nilai *minimum support* dan *minimum confidence* yang baik untuk digunakan pada data transaksi penjualan Grosir X tahun 2015 dan 2016?
3. Bagaimanakah perbandingan hasil *sequential rule* yang dihasilkan dari data transaksi pada tahun 2015 - 2016?

1.3. Batasan Sistem

Batasan masalah yang digunakan oleh penulis dalam pembuatan sistem adalah sebagai berikut.

1. *Data* yang digunakan adalah *data* transaksi penjualan barang Grosir X cabang Semarang dalam jangka waktu kurang lebih 2 (dua) tahun terhitung dari Juni 2014 hingga Agustus 2016.
2. *Sequence database* didapatkan berdasarkan member konsumen yang terdapat pada *data* transaksi penjualan barang Grosir X.
3. *Rule* yang dihasilkan bukan dari nama tiap barang, namun dari kategori barang tersebut.
4. Proses *Data Cleaning* dalam menentukan data penjualan yang *valid* ditentukan oleh pihak Grosir X.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah membuat suatu aplikasi yang dapat menghasilkan *sequential rule* dari *data* transaksi penjualan menggunakan algoritma *RuleGrowth*. Kemudian dari *sequential rule* yang ditemukan tersebut dapat digunakan untuk membantu Grosir X dalam pengambilan keputusan, contohnya dalam menentukan penempatan barang yang dijual.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini bagi peneliti adalah untuk menerapkan dan memperluas wawasan dalam menerapkan algoritma *data mining*. Selain itu, peneliti dapat belajar mengelola *data warehouse* pada kasus yang nyata. Peneliti juga dapat menyelesaikan tugas akhir perkuliahan.

Manfaat bagi Grosir X adalah untuk mempermudah Grosir X dalam mengolah data transaksi penjualan. Grosir X juga dapat melakukan promosi untuk penjualan jenis barang tertentu sesuai dengan aturan yang didapatkan.

Manfaat bagi penelitian selanjutnya, penelitian ini dapat menjadi bahan atau dasar bagi penelitian-penelitian selanjutnya yang masih berkaitan dengan *data mining*. Selain itu, penelitian ini juga dapat dikembangkan kembali menggunakan algoritma yang berbeda, atau pada kasus yang berbeda.

1.6. Metode Penelitian

Metode-metode yang digunakan penulis dalam menyelesaikan penelitian ini antara lain:

1. Metode Pengumpulan Data
 - a. Studi Pustaka

Penulis melakukan studi pustaka dengan cara mengumpulkan dan mempelajari teori-teori dan literatur yang dapat membantu penyelesaian penelitian. Khususnya teori-teori yang berkaitan dengan *Data Mining*, pembangunan dan pengelolaan *Data Warehouse*, *Sequential Rule Mining*, algoritma *RuleGrowth* serta *software-software* yang akan digunakan untuk membangun sistem.

- b. Pengambilan *Data*

Penulis akan mengumpulkan *data* transaksi penjualan barang dari Grosir X Semarang dari Juni 2014 hingga Agustus 2016.

2. Metode Perancangan Sistem

Tahap ini adalah tahap dimana penulis melakukan perancangan sistem yang akan dibuat, mulai dari *Data Warehouse* yang digunakan sebagai tempat penyimpanan *Big Data*, pembersihan *data* transaksi penjualan barang yang tidak digunakan, serta perancangan antarmuka pada sistem yang akan dibuat.

3. Metode Pengembangan Sistem

Tahap ini adalah tahap dilakukannya implementasi algoritma *RuleGrowth* terhadap *data* transaksi penjualan barang kebutuhan pokok Grosir X. Pengembangan sistem ini dilakukan hingga sistem dapat menghasilkan *sequential rule* dari *data* transaksi Grosir X.

4. Metode Evaluasi

Tahap evaluasi akan dilakukan dengan membuat *data* transaksi tahun 2015 sebagai *data* latih, kemudian *data* transaksi tahun 2016 akan digunakan untuk *data* uji. Kemudian dilakukan pengujian apakah hasil *rule* penerapan algoritma pada *data* latih terjadi perulangan dengan *rule data* uji. Hasil yang didapatkan tersebut kemudian dianalisis dengan berbagai *minimum support* dan *confidence* untuk mencari *sequential rule* yang paling banyak ditemukan kesamaan *rule* dari tahun 2015 dan 2016. Selain itu, *sequential rule* yang dihasilkan bukan hanya ditentukan berdasarkan *support* dan *confidence*, namun juga akan dihitung menggunakan *lift ratio*.

1.7. Sistematika Penulisan

Sistematika penulisan tugas akhir ini akan terbagi dalam 5 (lima) bab dengan urutan penulisan sebagai berikut.

Bab I. Pendahuluan

Bab ini berisi mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, metodologi penelitian serta sistematika penulisan laporan skripsi.

Bab II. Tinjauan Pustaka

Bab ini berisi mengenai tinjauan pustaka dan landasan teori. Tinjauan pustaka menguraikan beberapa teori yang menjelaskan mengenai algoritma *Sequential Rule Mining*, serta penelitian yang serupa. Landasan teori berisi mengenai penjelasan konsep-konsep yang diperlukan dalam memecahkan masalah pada penelitian ini.

Bab III. Perancangan Penelitian

Bab ini berisi mengenai perancangan sistem. Perancangan sistem meliputi pengolahan data ke *database*, instalasi *server*, perancangan implementasi algoritma *RuleGrowth* dan juga perancangan antarmuka sistem yang akan digunakan.

Bab IV. Implementasi dan Analisis Sistem

Bab ini berisi mengenai implementasi algoritma yang digunakan dalam pembuatan sistem, serta menganalisis hasil dari sistem yang dibuat.

Bab V. Kesimpulan dan Saran

Bab ini berisi mengenai kesimpulan yang didapatkan setelah menganalisis hasil sistem yang telah dibuat. Dan juga berisi saran yang dapat digunakan untuk mengembangkan sistem ini lebih lanjut.

Selain berisi bab-bab tersebut, tugas akhir ini juga dilengkapi dengan adanya daftar isi, daftar gambar, daftar tabel, daftar pustaka dan lampiran.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dalam penelitian ini, algoritma *RuleGrowth* sudah dapat diimplementasikan untuk menghasilkan *sequential rule* yang berasal dari data transaksi penjualan barang Grosir X. Sebelum data transaksi tersebut dapat digunakan, data-data tersebut harus melalui beberapa proses antara lain proses ETL, *data cleaning* dan *data filtering*. Data yang sudah dipersiapkan akan dapat menghasilkan *sequential rule* menggunakan algoritma *RuleGrowth* yang digunakan. Sistem yang dibuat juga sudah mengimplementasikan perhitungan *lift ratio* pada *sequential rule* yang akan disimpan. Jadi *sequential rule* yang dihasilkan lebih kuat dibandingkan *sequential rule* yang hanya didapatkan dari *minimum support* dan *minimum confidence*.

Dari proses pengujian yang sudah dilakukan, didapatkan nilai *minimum support* 26% dan *minimum confidence* 55% untuk presentase nilai terbesar dari perbandingan antara *rule* yang berulang tahun 2015-2016 dan *rule* dari data latih. Dengan menggunakan kedua nilai tersebut, pada periode waktu 1 Januari - 30 April didapatkan presentase sebesar 69,04% dan pada periode waktu 1 Mei - 31 Agustus didapatkan presentase sebesar 36,36%. Dengan jumlah *sequential rule* yang berulang pada tahun 2015 - 2016 sebanyak 136 *rule* untuk periode waktu 1 Januari - 30 April dan 60 *rule* untuk periode waktu 1 Mei - 31 Agustus.

Dari *sequential rule* yang terjadi pada tahun 2015 hingga 2016 tersebut, Grosir X dapat mengetahui aturan-aturan apa saja dilakukan oleh memernya selama tahun 2015 dan 2016. Dengan adanya *sequential rule* tersebut, Grosir X juga yang menggunakan informasi tersebut untuk proses

pengambilan keputusan, seperti mempersiapkan stok barang, atau bisa juga sebagai pertimbangan dalam melakukan promosi penjualan.

Dalam membandingkan *sequential rule* yang muncul pada tahun 2015 - 2016, Selain didapatkannya *sequential rule* yang berulang, Grosir X juga dapat melihat jumlah *sequential rule* yang hanya berlaku pada tahun 2015 saja (*rule lama*) dan jumlah *sequential rule* yang baru muncul pada tahun 2016 (*rule baru*). Untuk periode waktu 1 Januari - 30 April didapatkan jumlah *rule baru* sebanyak 669 dan *rule lama* sebanyak 61. Sedangkan untuk periode waktu 1 Mei - 31 Agustus didapatkan *rule baru* sebanyak 105 dan *rule lama* sebanyak 40. Hal tersebut menunjukkan bahwa pada periode waktu 1 Januari - 30 April, banyak terjadi perubahan perilaku sebagian besar member dalam melakukan transaksi dari pada periode waktu 1 Mei - 31 Agustus.

5.2. Saran

Saran-saran yang dapat digunakan dalam pengembangan aplikasi selanjutnya adalah:

1. Data transaksi penjualan yang digunakan memiliki jangka waktu yang lebih lama, contoh 5 tahun. Dengan begitu proses pengujian *rule* akan menjadi lebih detil karena data latihnya akan semakin banyak.
2. Algoritma selain *RuleGrowth* seperti algoritma *ERMiner* juga dapat dicoba dalam memproses data transaksi dari penjualan Grosir X.

DAFTAR PUSTAKA

- Fournier-Viger, P., Wu, C., Tseng, V.S., Cao, L., & Nikambou, R. (2015). Mining Partially-Ordered Sequential Rules Common to Multiple Sequences. *IEEE Transactions on Knowledge and Data Engineering*. Retrieved from <http://dx.doi.org/10.1109/TKDE.2015.2405509>
- Fournier-Viger, P., Gomariz, A., Gueniche, T., Soltani, A., Wu, C., & Tseng, V.S. (2014). SPMF: A Java Open-Source Pattern Mining Library. *Journal Machine Learning Research*, 15, 3569-3573.
- Fournier-Viger, P., Faghihi, U., Nkambou, R., & Nguifo, E.M. (2012). CMRULES: An Efficient Algorithm for Mining Sequential Rules Common to Several Sequences. *Knowledge-Based Systems*, 25(1), 63-76.
- Fournie-Viger, P., Nikambou, R., & Tseng, V.S. (2011). RuleGrowth: Mining Sequential Rules Common to Several Sequences by Pattern-Growth. *Proceedings of the 26th Symposium on Applied Computing(ACM SAC 2011)*. ACM Press, pp. 956-961.
- Hand, D. J., Mannila, H., & Smyth, P. (2001). *Principles of Data Mining*. Cambridge, MA: MIT Press. Retrieved from ftp://gamma.sbin.org/pub/doc/books/Principles_of_Data_Mining.pdf.
- Husni, Z.N., & Mukhlash, I. (2014). Implementasi Business Intelligence Pada Manajemen Report Bank XYZ. *Journal Sains dan Senni Pomits*, 3(2), A16-A21.
- Inmon, W. (2002). *Building the Data Warehouse*. Canada: Wiley & Sons.
- Juliastio, R., & Gunawan. (2015). *Sequential Pattern Mining Dengan Spade Untuk Prediksi Pembelian Spare Part Dan Aksesoris Komputer Pada Kedatangan Kembali Konsumen*. Seminar Nasional “Inovasi dalam Desain dan Teknologi”. IDEaTech.
- Sunento. (n.d.). *Module 1: Pengenalan Dasar Tentang JSP*. Retrieved from <http://kambing.ui.ac.id/onnopurbo/library/library-ref-ind/ref-ind-2/application/programming/Module%201%20-%20Pengenalan%20Dasar%20JSP.pdf>

Yulianton, H. (Januari 2008). *Data Mining untuk Dunia Bisnis. Jurnal Teknologi Informasi DINAMIK*, 13(1), 9-15.

©UKDWN