

Optimasi Seleksi Aturan Untuk Rekomendasi *Bundling* Produk Melalui Kombinasi Algoritma Apriori dan *Utility Weighted Score* (UWS)

Ichsan Fahmi, Hadi Suyono, Moehammad Sarosa

Abstract— This research not only aims to improve Apriori Algorithm (AA) in the method of Association Rules Mining (ARM) but also proposes an alternative solution for bundling recommendation models for retail products. Binary mapping problem with pruning rules that often eliminates the potential product try to overcome with additional parameters UWS. These improvements at the same impact on the quality of product bundling recommended. Optimization is done on the rules generated by the AA with parameters *U_gain* and *W_gain* which is a representation of various product attributes. Quantity and profit is part of the attributes of the product items to be used in this study. The test results using data Point of Sales (POS) showed an improvement of rules. If using conventional AA, the resulting rules often do not contain potential items. This is because the selection process repeated with a more critical level of confidence that often disappear rules with potential items. On the other hand, a number of rules produced often do not contain the items quality product. Optimization is done by changing the parameters of the transaction minimum, minimum support and minimum confidence combined with calculation *U_gain* and *W_gain* to get the rules that generate the maximum value.

Index Terms: Apriori Algorithm, *Bundling* Produk, *Confidence*, *Utility Weighted Score*

Abstrak— Penelitian ini tidak hanya bertujuan untuk memperbaiki Algoritma Apriori (AA) dalam metode Association Rules Mining (ARM) tetapi juga mengusulkan suatu solusi alternatif untuk model rekomendasi *bundling* produk bagi ritel. Problem pemetaan biner dengan pemangkasan aturan yang sering menghilangkan produk yang potensial coba diatasi dengan tambahan parameter UWS. Perbaikan ini sekaligus berdampak pada kualitas *bundling* produk yang direkomendasikan. Optimasi dilakukan pada aturan yang dihasilkan oleh AA dengan parameter *U_gain* dan *W_gain* yang merupakan representasi dari berbagai atribut produk[1]. Quantity dan profit adalah bagian dari atribut item produk yang akan digunakan dalam penelitian ini. Hasil pengujian dengan menggunakan data *Point of Sales* (POS) menunjukkan adanya perbaikan aturan. Jika menggunakan AA konvensional, aturan yang dihasilkan sering tidak mengandung item yang potensial. Hal ini karena proses seleksi ulang dengan level *confidence* yang lebih kritis sering menghilang aturan dengan item yang potensial. Pada sisi yang lain, sejumlah aturan yang dihasilkan sering tidak mengandung item produk yang berkualitas. Optimasi dilakukan dengan mengubah

parameter minimum transaksi, minimum *support* dan minimum *confidence* yang dikombinasikan dengan perhitungan *U_gain* dan *W_gain* untuk mendapatkan aturan yang menghasilkan nilai maksimal.

Kata Kunci—Algoritma Apriori, *Bundling* Produk, *Confidence*, *Utility Weighted Score*

I. PENDAHULUAN

DATA transaksi ritel merupakan representasi perilaku pelanggan yang tersimpan dalam semua *Point of Sales* (POS). Banyak keputusan bisnis dan aktifitas ritel sangat dipengaruhi oleh data POS. Sebagian ritel menggunakan data POS sebagai strategi pemasaran dalam menghadapi berbagai persaingan bisnis. *Bundling* produk adalah salah satu strategi penjualan yang telah banyak digunakan dalam berbagai bidang usaha dan jasa termasuk juga oleh ritel. Stremersch dan Tellis [1] mendefinisikan *bundling* sebagai sebuah strategi pemasaran yang melibatkan penawaran dua produk atau lebih yang terpisah untuk dijual sebagai satu kesatuan/paket. Pada prinsipnya ritel melakukan *bundling* produk lebih berorientasi untuk membantu meningkatkan penjualan meski secara tidak langsung ada beberapa pertimbangan lain yang menjadi motif dilakukannya *bundling* produk. Ritel yang memiliki ratusan bahkan ribuan item produk seharusnya telah memiliki sistem yang baik untuk mendesain *bundling* produk secara otomatis. Belum adanya sistem yang mendukung untuk pemilihan produk sebagai *bundling* produk menjadi penyebab digunakannya metode konvensional yang lebih bersifat subyektif meski tetap mengacu pada data transaksi POS. Sebagian ritel mencoba alternatif lain dengan berdasarkan pada quisioner yang disebar secara acak pada suatu wilayah target. Pada sisi yang lain, kehadiran Data Mining (DM) dengan berbagai metodenya menjadi solusi alternatif dalam pengambilan keputusan. *Metode Association Rules Mining* (ARM) dengan Algoritma Apriori (AA) merupakan sebuah konsep alternatif untuk membantu proses *bundling* produk. AA identik dengan *bundling* produk sebab AA mencoba menemukan semua kombinasi produk yang tersimpan dalam data POS. Kombinasi produk ini menjadi solusi alternatif untuk rekomendasi *bundling* produk.

AA termasuk salah satu diantara 10 algoritma DM yang paling populer[2]. Meskipun demikian, AA juga menyisakan beberapa problem baik dalam tahap *pre*

Ichsan Fahmi, Hadi Suyono, Moehammad Sarosa are with the Electrical Engineering Department of Brawijaya University, Malang, Indonesia (corresponding author provide phone 0341-554166; email fahmiski2011@gmail.com).

processing maupun tahap *post processing*. Aturan yang dihasilkan oleh AA dalam tahap *post processing* pada hakekatnya bersumber dari pemetaan biner yang meniadakan semua atribut dari produk misalnya quantity, profit dan harga. Sementara bagi ritel, proses yang berkaitan dengan *bundling* produk tidak terlepas dari unsur atribut tersebut yang menjadi sumber referensi pertimbangan utama. Problem berikutnya adalah seleksi ulang dengan menggunakan level *confidence* yang lebih kritis sering menghilangkan produk yang potensial dari data POS. Adanya kontradiksi ini menjadi salahsatu pertimbangan diajukannya konsep tambahan parameter *Utility Weighted Score* (UWS yang akan mengoptimasi aturan yang dihasilkan oleh AA sehingga *bundling* produk yang direkomendasikan akan lebih berkualitas.

II. LANDASAN TEORI

Pada bab ini berisi dasar teori yang relevan yang mendukung penelitian ini.

A. Algoritma Apriori (AA)

Algoritma Apriori (AA) adalah salah satu algoritma dari metode ARM yang telah banyak diterapkan dalam berbagai bidang. AA juga terus mengalami perkembangan dengan berbagai modifikasi dan perbaikan yang disesuaikan dengan lingkungan implementasinya.

Konsep dasar AA adalah proses iterasi yang berulang dengan dibatasi oleh suatu nilai minimum untuk seleksi ke tahap iterasi berikutnya. Target AA adalah mencari relasi diantara berbagai atribut yang terjadi dalam suatu kejadian waktu dengan pendekatan pemetaan biner. Hasil dari setiap iterasi akan dijadikan acuan untuk iterasi berikutnya. Pada tahap akhir iterasi dihasilkan sejumlah aturan yang terdiri dari kombinasi berbagai item dalam relasi *antecedent* dan *consequent*. Di dalam proses seleksi aturan in, AA bersandar pada dua parameter utama yakni minimum *support* dan minimum *confidence*. Sedangkan untuk proses evaluasi terhadap aturan tersebut sering digunakan parameter *Interest Measurement* (IM). Dalam berbagai literature, IM dibedakan menjadi *Objective IM* dan *Subjective IM*. *Objective IM* lebih berkaitan dengan unsur statistik sedangkan *Subjective IM* lebih berkaitan dengan persepsi user yang menggunakan aturan tersebut. Adapun rumus perhitungan minimum *support* dan minimum *confidence*[3], [4] adalah sebagai berikut:

$$Support A = \frac{\text{Jumlah transaksi mengandung A}}{\text{Total transaksi}} \quad (1)$$

Sedangkan untuk dua item maka rumus *support*nya adalah:

$$Supp A \cap B = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Total transaksi}} \quad (2)$$

untuk *confidencenya* adalah:

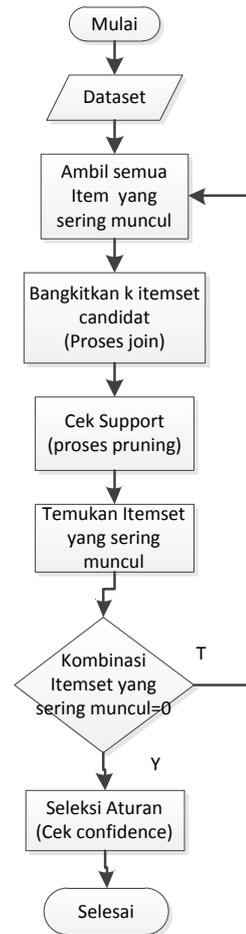
$$Conf P(B|A) = \frac{\text{Jumlah transaksi mengandung A dan B}}{\text{Jumlah transaksi mengandung A}} \quad (3)$$

Parameter *Support* dan *confidence* adalah bagian dari *Objective IM* juga[5]. Parameter lain yang juga populer digunakan untuk evaluasi aturan adalah *lift*[6]. Adapun rumus untuk *lift* adalah :

$$lift(A \Rightarrow B) = \frac{supp(A \cup B)}{supp(B) \times supp(A)} \quad (4)$$

Konsep dasar AA terlihat pada gambar 1

Gambar 1. Flow chart Algoritma Apriori



Gambar 1. Flow chart Algoritma Apriori

B. Utility Weighted Score (UWS)

UWS merupakan kombinasi dari parameter *weighted* dan *utility*[7]. Sebagian referensi memandangnya sebagai *utility internal* dan *utility external*[8]. Data Mining (DM) dengan metode Association Rule Mining (ARM) yang berdasarkan *utility* adalah pintu masuk area baru bagi penelitian. *Utility* suatu item adalah istilah subyektif yang tergantung pada user dan aplikasi. Khusus untuk data POS, hal ini berkaitan dengan sejumlah atribut yang melekat pada setiap item produk. Beberapa atribut tersebut adalah quantity, profit, harga, discount dan sebagainya. Gambar 2. adalah flow chart sederhana untuk proses UWS. Atribut item produk yang digunakan sebagai *utility* untuk menghitung nilai *U_gain* adalah profit. Sedangkan untuk menghitung *W_gain*, atribut yang dipilih adalah quantity.

III. METODE PENELITIAN

Secara umum metode penelitian yang dilakukan berkaitan dengan beberapa hal berikut ini:

1. Studi literatur dan survei
2. Pemodelan software
3. Analisis kebutuhan software

4. Desain software
5. Implementasi
6. Evaluasi

Studi literature dan survey dilakukan untuk mengumpulkan semua informasi yang berkaitan dengan penelitian yang dilakukan. Survei dilakukan untuk mendapatkan informasi tambahan menyangkut aktifitas bisnis, standar kebutuhan user, *stakeholder* yang terlibat, prosedur dan mekanisme sistem yang sedang berjalan, spesifikasi software dan hardware yang ada dan sebagainya. Penelitian ini menggunakan dua jenis data. Data sintesis dan data real POS yang diambil dari salah satu ritel.

Dari hasil survey lapangan diketahui bahwa *bundling* produk sering digunakan oleh ritel untuk membantu meningkatkan penjualan item produk. Selain itu *bundling* produk juga menjadi solusi untuk mengatasi item produk yang bermasalah. *Bundling* produk juga berfungsi sebagai media promosi yang efektif. Pemilihan item produk untuk *bundling* produk secara prosedur akan mempertimbangkan quantity dan profit sebagai salah satu referensinya.

IV. HASIL DAN PEMBAHASAN

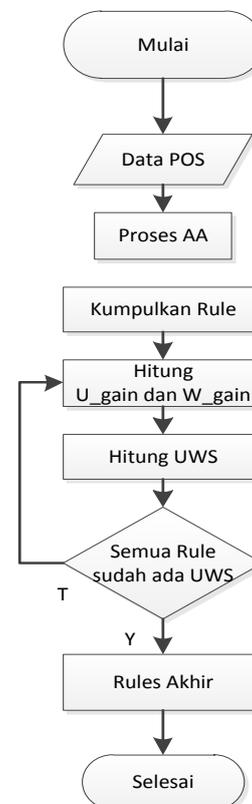
TABEL I
ATURAN YANG DIPROSES DENGAN AA

Item1	Item2	Item3	Confidence
233455	276937	308907	100%
233455	276937	320522	100%
233455	276937	322315	100%
233455	276937	323996	100%
233455	308907	276937	100%
233455	308907	320522	100%
233455	308907	322315	100%
233455	308907	323996	100%
233455	320522	276937	100%
233455	320522	308907	100%

Aturan yang merupakan output dari sistem, baik yang menggunakan AA atau UWS akan dievaluasi menggunakan beberapa model sesuai yang tampak pada gambar 3. Dengan menggunakan data sintesis transaksi POS yang telah diketahui karakteristiknya yakni yang mengandung aturan dengan level *confidence* 100%, terdapat 7 item produk yang ditransaksikan berulang sebanyak 10 kali maka aturan yang dihasilkan oleh Algoritma Apriori (AA) adalah tampak pada tabel 1. Pada percobaan berikutnya digunakan data POS yang tidak memiliki item produk yang ditransaksikan berulang dan sistem meresponnya dengan tidak menampilkan satupun aturan. Hal ini karena tidak ada item produk yang memenuhi nilai minimum *frequent*. Penggunaan data sintesis ini hanya untuk memudahkan proses analisa sistem. Dari percobaan awal ini diketahui bahwa sistem yang dibuat dapat membedakan data transaksi yang mengandung aturan dan yang tidak mengandung aturan.

Point berikutnya adalah membandingkan aturan yang dihasilkan sistem dengan perhitungan manual. Adapun hasilnya menunjukkan bahwa keduanya tidak memiliki

perbedaan aturan satupun. Tahap berikutnya adalah membandingkan aturan yang dihasilkan oleh sistem yang menggunakan AA dengan suatu tools DM. Dalam percobaan ini tools DM yang digunakan adalah Weka versi 3.6. Hasilnya menunjukkan bahwa keduanya baik sistem maupun Weka tidak memiliki perbedaan yang signifikan untuk aturan yang mengandung item produk



Gambar 2. Flow chart UWS



Gambar 3. Model evaluasi penelitian

Percobaan selanjutnya adalah menggunakan data real POS. Sistem yang didesain ini juga menampilkan beberapa karakteristik data POS yang tersedia, yakni jumlah quantity total setiap item, total profit setiap item dan total harga yang ditransaksikan. Karakteristik ini akan menjadi pelengkap dari evaluasi yang dilakukan. Salah satu contoh percobaan POS dengan AA dan kombinasi AA dengan UWS beserta hasil aturannya dapat dilihat pada tabel 2.

Jika kedua aturan ini dibandingkan maka tampak terlihat ada perbedaan item produk yang digunakan, termasuk juga jumlahnya. Misalkan hanya diambil 3 aturan saja, terlihat bahwa untuk AA terdapat item 308907 yang jika dikembalikan pada data transaksinya secara quantity hanya berjumlah 104 quantity transaksi. Nilai quantity ini lebih kecil dibandingkan item yang digunakan pada aturan yang dihasilkan oleh UWS, misalkan item 323996 yang memiliki 124 quantity transaksi penjualan.

TABEL II
ATURAN YANG DIPROSES DENGAN UWS

Item1	Item2	Item3	UWS
233455	276937	320522	20
233455	320522	276937	20
233455	276937	323996	19
233455	323996	276937	19
276937	320522	323996	19
276937	323996	320522	19
233455	276937	322315	18
233455	320522	323996	18
233455	322315	276937	18
233455	323996	320522	18

Hal yang sama tampak juga pada kontribusi profit yang diberikan oleh masing masing item, UWS lebih memberikan kontribusi profit dibandingkan dengan AA. Misalkan nilai *confidence* dalam tabel.1 tidak 100% tetapi bervariasi dengan urutan ranking yang demikian, maka ada kemungkinan item produk 323996 akan terpangkas oleh seleksi ulang dengan level minimum *confidence* yang lebih kritis sedangkan item produk 308907 yang menempati urutan teratas akan tetap terpilih meski kontribusinya kecil dibandingkan item produk 323996. Kombinasi AA dengan parameter UWS yang mempertimbangkan unsur quantity dan *profit* dalam menentukan aturan mempunyai kemiripan dengan prosedur *bundling produk*. Untuk mendapatkan nilai maksimal dari kombinasi item produk pada suatu aturan maka optimasi perlu dilakukan dengan mengganti nilai minimum transaksi dari kandidat item, minimum *support* dan minimum *confidence*. Seleksi terakhir dari aturan yang dihasilkan dengan UWS adalah menggunakan model peringkat. Aturan yang berada pada peringkat pertama mempunyai nilai UWS yang terbesar. Nilai ini sesungguhnya adalah representasi dari unsur quantity dan *profit*. Jika terjadi pemangkasan aturan pun item yang hilang mempunyai nilai quantity dan *profit* yang lebih kecil dibandingkan yang terpilih.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian dan analisis yang menggunakan beberapa data transaksi POS yang berbeda karakteristiknya maka dapat disimpulkan bahwa:

1. Parameter UWS dapat memperbaiki kualitas aturan yang dihasilkan oleh metode ARM yang hanya menggunakan AA.
2. Hasil analisis menunjukkan bahwa untuk semua data transaksi yang digunakan, jumlah item yang terkandung dalam aturan oleh AA *dapat* berbeda dengan UWS
3. Optimasi aturan dari sistem ini dapat digunakan sebagai alternatif untuk rekomendasi *bundling* produk

DAFTAR PUSTAKA

- [1] Stremersch dan Tellis ,2002. Strategic bundling of products and Prices: A new synthesis for marketing. Journal of Marketing; ProQuest pg. 55
- [2] Wu, X, etc*. Top 10 algorithms in data mining, 2008. J.Knowledge Inf Syst 14:1– 37. Springer-Verlag
- [3] Kusriani dan Luthfi,E. 2009 Algoritma Data Mining, Penerbit Andi.Yogyakarta
- [4] Nengsih, Warnia,2015, A Comparative Study on Market Basket Analysis and Apriori Association, International Conference on Information and Communication Technology (ICoICT), IEEE 978-1-4799-7752-9 Page(s):461 - 464
- [5] Sahoo, Kumar and Goswami ,2015,An efficient approach for mining association rules from high utility itemsets, Expert Systems with Applications 42 (2015) 5754–5778
- [6] Nong Ye,2003,THE HANDBOOK OF DATA MINING,Arizona State University,Lawrence Erlbaum Associates, Inc., Publishers
- [7] Sandhu, Dhaliwal, Panda dan Bisht,2010,An Improvement in Apriori algorithm Using Profit And Quantity, An Second International Conference on Computer and Network Technology. IEEE,978-0-7695-4042-9, page 3-7
- [8] Shankar, Purusothaman and Jayanthi,2008,Novel Algorithm for Mining High Utility Itemsets, Proceedings of the International Conference on Computing, Communication and Networking (ICCCN 2008), IEEE 978-1-4244-3595-1 Page(s):1 - 6