# Cross-Selling's Product Determination in the Context of Analytical CRM Based on Association Rules 

Eko K. Budiardjo ${ }^{1}$, Bayu Adhi Tama ${ }^{2}$<br>${ }^{1}$ Fakultas Ilmu Komputer - Universitas Indonesia<br>${ }^{2}$ Fakultas Ilmu Komputer - Universitas Sriwijaya<br>eko@cs.ui.ac.id,bayu@ilkom.unsri.ac.id


#### Abstract

Cross-selling's product determination achieves through out an analysis of purchasing data by using business intelligence tools such as data mining. Data mining is part of Analytical CRM used to discover data pattern, while market basket analysis method with association rules technique are a data mining task used to discover the candidate of crossselling's product combinations, based on frequent products in purchasing data.


Index Terms-Cross-Selling, Customer Relationship Management (CRM), Association Rules.

## I. Introduction

Business competition has encouraged a trading company to be more selective on implementing their marketing strategy. Cross-selling's marketing strategy, as part of Customer Relationship Management (CRM) could be used as an alternative way. Sales transaction is possible to be increased by selling additional products from the main product that customer has purchased.

The determination of cross-selling product from sales transaction data as well as the techniques that will be used for Analytical CRM. Steps include source data identification, cross-selling modeling, choosing algorithm, correlation analysis, and result presentation through internal website.

## II. TRANSACTION DATA

Identification is the first step in cross-selling product determination is to know business process that currently run in the company. As mention in previous chapter, PT XYZ is a trading company which has widely range products. Sales and distribution has supported by ERP (enterprise resource planning) which integrate all of company activities.

Implementation of cross-selling product determination will increase revenue in a row with selling of additional product that is derived from its base product. The final result of this research is cross-selling model through sales transaction data analysis by using association rule in order to support decision making process as well as determine best marketing strategy effectively and also enhance customer loyalty.

Data source identification must be performed in next step. Database is relatively uniform since has extracted from ERP's database server. In ERP system, there are many databases which also contain thousand tables and attributes. Each database handles divisions or business process in a organization, such as customer database, inventory, finance, sales and marketing, human resource, etc. Database that will be used is extracted from subject area sales.

Sales transaction data in 2007 is used to determine cross-selling strategy. One branch of PT XYZ is choosen. However, this branch have not implemented business intelligence tools such as data warehouse. Therefore, data mining task will directly mining database that came from OLTP (online transaction processing) (Tang and MacLennan, 2005).

The company's product offering are very diverse (selling more than 1000 different items) which is divided into five division. They are Industrial/Graphic Art, Telecomunication, Photography and Photofinishing, Ricoh, and Medical. Sales area covers South Sumatera, Bengkulu, Jambi, and Bangka Belitung. Their major customer include agent, sub-agent, and other business-tobusiness customer such as printing company, hospital, publishing company, and so on that in round numbers 400 customers. Some products are listed in the table below that are categorized by division.

For this study, not all sales transaction data will be involved in determining cross-selling product, however, only products that contained one of the top 50 products (how many times purchased, not how many quantity that have been sold). In 2007, the company had sold approximately 4674 different products to 400 customers. Sales transaction will be divided into two periods of time. Data in first semester (January-June) will be used as experiment and data in second semester will be used as evaluation. Top 50 product in first semester

## III. ANALYsis of Cross-SELLING's Product DETERMINATION

After we choose the products that will be used, we try to determine possible combination of all products. The method of Market basket analysis and product affinity
analysis will be performed. Actually, association rule technique will be used in order to find product association or combination through transaction data (Bugher, 2000).

In association technique, there are three different indicators such as support score, confidence score, and improvement score. Support score specify frequent of two product appears in a transaction, confidence score indicate one product is dependent to other product, and improvement score specify customer to be more likely buy cross-selling's product from base product.

It is likely to produce the best cross-selling product from the total of transaction each customer in certain period is the item with high confidence score. As an example the first product is the combination of A and B with confidence score of $x \%$ mean that when product $A$ is bought, possibilty product B is also bought equal to $\mathrm{x} \%$ from the entire period of transaction. The second indicator that should be decided is support score which decides how many products of A and B appear in the entire of transactions simultaneously. As an example, the combination of Product C and D have support score y\% indicate that product C and D appear simultaneously egual to $\mathrm{y} \%$ from the entire of the total of transaction, whereas to combinate product E and F . It has an improvement score $\geq 1$. indicate that product E and F is a positive correlated, means that if the customer buys product E , the customer also agrees to buy product F , whereas if improvement score $<1$, product E and F is a negative correlated. Product with negative correlated can not be concluded as a suggestion in determining crossselling product although having support score and high confidence score.

## IV. SQL SERVER 2005 ANALYSIS SERVICE

Tools which are used to analyze is SQL Server 2005. SQL Server 2005 Analysis Service has an ability in making data mining with knowledge discovery task such association rule and other tasks which have predictive character. First step in processing data mining in SQL Server 2005. It is used to compare tables data transaction result is copied into table format so easy to be known by SQL Server 2005. It only recognizes the concept of nested table to present a variable length where a group of attribute related to the same case. Kind of representation table concept is nested from SQL Server 2005 as follows:


Figure 1: Example of nested table in SQL Server 2005 (Iyer and Jesper, 2004)

In this PT XYZ study case, selling data transaction on each customer is grouped into tables of owning three columns that is transaction column id, type of goods and type of good's name. this data is grouped on each customer transaction in making purchase at PT XYZ during range of time. Here is a represtation nested from purchasing transaction on each customer which is used in this study case

| TRA ID - | TYPE BARANG | NAMA TYPE BARANG |
| :---: | :---: | :---: |
|  | AA13161 | CRYSTAL $805 \times 610 \mathrm{G}$ |
|  | AA13163 | CRYSTAL $805 \times 610$ L |
|  | AA13165 | CRYSTAL $805 \times 610 \mathrm{M}$ |
| 1 | AA13169 | CRYSTAL $806 \times 610$ L |
|  | AA13173 | CRYSTAL $808 \times 295 \mathrm{~L}$ |
| 1 | AA3141 | 14834332 SUPER CONDITIONER 100 TABLET |
|  | AA3278 | 850086 CN16L/Q N4R 4X4 L |
|  | AA3284 | CP49E REPL |
|  | AA70075 | KAMERA AVEC |
|  | AA70116 | KAMERA BONA ZOOM +2 NS 400/36+BATT |
|  | AA70122 | KAMERA BENEV SILVER +2 NS 400/36+BATT+FB |
|  | AA7120 | FILM WARNA NEW SUPERIA 200/36 |
|  | AA7195 | FILM WARNA NEW SUPERIA X-TRA 400/36 |
|  | AA8024 | SLING BAG |
|  | AA9171 | AMPLOP FIS |
|  | AK4997 | PAKET VATICA 3R (12 BUAH) |
|  | AA13161 | CRYSTAL $805 \times 610 \mathrm{G}$ |
|  | AA13173 | CRYSTAL $808 \times 295 \mathrm{~L}$ |
|  | AA3015 | 850123 CN-16Q NQ1R 10LTX2 |
|  | AA3016 | 850045 CN-16Q NQ2R 8LTX2 |
|  | AA3284 | CP49E REPL |

Figure 2: Nested Table of Transaction Data of PT XYZ

## V. CONCLUSION

Investigating in order to determine the cross-selling's product, in which explained in this paper, is valid for a trading company in the fields of medical, printing, telecommunication, photography, and photocopy. The Transaction data used is sales transaction data for 2007, in which extracted from their ERP system. Combinations of best cross-selling's products were chosen by order ranking of all possible products in descending order by their confidence score. Evaluation is discussed by comparing the cross-selling's product determination using January-June transaction data and validate each product combinations by comparing July-December data.

## REFERENCES

[1] Bugher, G. "Market Basket Analysis - Knowledge Discovery in Transactional Data (white paper),"Megaputer Intelligence Inc., 1999. (tersedia online di www.megaputer.com diakses pada tanggal 10 Juni 2008).
[1] Bugher, G. "Market Basket Analysis of Sales Data for a client of Cambridge Technology Partner,"Megaputer Intelligence Inc., 2000. (tersedia online di www.megaputer.com/mba_ctp.pdf diakses pada tanggal 10 Juni 2008).
[2] Berry, M. J. A. and Linoff G. S. Mastering Data Mining - The Art and Science of Customer Relationship Management, New York: Jhon Wiley and Sons, 2000.
[3] Brown, Stanley A. Customer Relationship Management : A Strategic Imperative in the World of e-Business, Kanada : Jhon Wiley and Sons Ltd, 2000.
[4] Cashin, Jefrey R. Implementation of A Cross-Selling Strategy for A Large Midwestern Healthcare Equipment Company, Ph.D Dissertation, Department od Psychology, Southern Illinois University at Carbandole, February 2003.
[5] Greenberg, Paul. CRM At the Speed of Time - Capturing and Keeping Customer in Internet Real Time, McGraw-Hill, 2000.
[6] IBM. "Retail Selling Optimization - Strategic cross-selling and upselling for survival in the volatile retail market (white paper)," IBM Retail Industry Solution, 2004. (tersedia online di www-03.ibm.com/industries/retail/doc/content/bin/retail-sellingoptimization.pdf diakses pada tanggal 10 Juni 2008).
[7] Iyer, R and Jesper Lind. "SQL Server 2005 Data Mining - Create a Web Cross-sell Application" (tersedia online di http://www. aspnetpro.com/newsletter-article/2004/10/asp200410ri_1/as200410 ri_1.asp) [Diakses pada tanggal 30 Mei 2008].
[8] Kälakota, Ravi et al. e-Business 2.0: Roadmap for Success, Addison-Wesley Publishing Company, 2000.
[9] Microsoft. "Microsoft Business Solutions - Axapta 3.0 Reference Guide,"Denmark: Microsoft Business Solutions, 2004.
[10] Neuman, W. L. Social Research Methods: Qualitative and Quantitative Approaches $5^{\text {th }}$ Ed., Allyn and Bacon, 2003
[11] Rao, Srinivasa and Saurabh, S. "Business Intelligence and Retailing (white paper),"Wipro Technology, 2001 (tersedia online di http://www.wipro.com/webpages/insights/businessintelligence retail.htm diakses pada tanggal 10 Juni 2008)
[12] SageCRM. "Turning Customer Loyalty into Profit (white paper)," Irlandia: Sage, 2006 (tersedia online di www.caplus.com/brochures /Turning_Customer_Loyalty_into_Profit.pdf diakses pada tanggal 10 Juni 2008)
[13] Swift, Ronald. S. Accelerating Customer Relationships - Using CRM and Relationship Technologies, Prentice Hall Inc., 2000.
[14] Tang, ZhaoHui and Jamie McLennan. Data Mining with SQL Server 2005, Indianapolis: Wiley Publishing Inc, 2005.
[15] Todman, Chris. Designing a Data Warehouse - Supporting Customer Relationship Management, Hewlett Packard, 2001
[16] Witten, Ian H. and Eibe Frank. Data Mining - Practical Machine Learning Tools and Techniques $2^{\text {nd }}$ Edition, Morgan Kaufmann Publisher, 2005.
[17] Xu, M. and Walton J. "Gaining Customer Knowledge Through Analytical CRM," Industrial Management \& Data System, Vol. 105 No. 7 (2005), pp 955-977.

