MARKET BASKET ANALYSIS FOR ANALYZING JAGO MINI-MARKET SALES

TUGAS AKHIR

Disusun Oleh :

Veronica Christiani

Nim : 22064042

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS KRISTEN DUTA WACANA
YOGYAKARTA
2011

MARKET BASKET ANALYSIS FOR ANALYZING JAGO MINI-MARKET SALES

TUGAS AKHIR

Diajukan Kepada Fakultas Teknik Program Studi Teknik Informatika
Universitas Kristen Duta Wacana Yogyakarta
Untuk Memenuhi Sebagian Syarat-Syarat
Guna Memperoleh Gelar
Sarjana Komputer

Disusun Oleh:

<u>Veronica Christiani</u>
22064042

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS KRISTEN DUTA WACANA
YOGYAKARTA
2011

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa tugas akhir dengan judul :

"Market Basket Analysis for Analyzing Jago Mini-Market Sales"

Yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan sarjana Program Studi Teknik Informatika, Fakultas Teknik Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi kesarjanaan 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 skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia menerima sanksi berupa pencabutan gelar kesarjanaan saya.

Yogyakarta, 18 Juli 2010

(Veronica Christiani)

22064042

i

HALAMAN PERSETUJUAN

Judul	:	Market Basket Analysis for Analyzing Jago Mini-Market Sales
Nama	:	Veronica Christiani
NIM	:	22064042
Mata Kuliah	:	Tugas Akhir
Kode	:	TI2126
Semester	:	Genap
Tahun Akademik	:	2011 / 2012

Telah diperiksa dan disetujui
Di Yogyakarta,
Pada tanggal ...19 Juli 2011....

Dosen Pembimbing I   Dosen Pembimbing II

(Lucia Dwi Krisnawati, S.S., M.A)   (Ir. Harianto Kristanto, M.T., M.M.)

HALAMAN PENGESAHAN

SKRIPSI

Market Basket Analysis for Analyzing Jago Mini-Market Sales
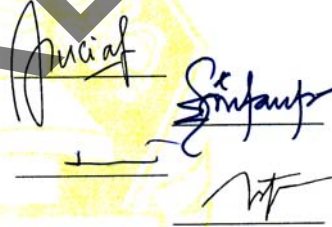
Oleh : Veronica Christiani / 22064042

Dipertahankan di depan Dewan Penguji Tugas Akhir/Skripsi

Fakultas Teknologi Informasi Program Studi Teknik Informatika

Universitas Kristen Duta Wacana Yogyakarta

dan diterima untuk Memenuhi Sebagian

Syarat-syarat Guna Memperoleh Gelar

Sarjana Komputer

pada tanggal

27 Juli 2011

Yogyakarta, 8 Agustus 2011

Mengesahkan,

Dewan Penguji :

1. Lucia Dwi Krisnawati, S.S., M.A

2. Ir. Harianto Kristanto, M.T., M.M

3. Ir. Sri Suwarno, M.Eng

4. Restyandito, S.Kom., M.Sis

Dekan,

Drs. Wimmie Handiwidjojo, MIT

Ketua Program Studi,

Nugroho AgusH,S.Si,M.Si.

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa yang telah membimbing selama masa pengerjaan tugas akhir dengan judul Market Basket Analysis for Analyzing Jago Mini-Market Sales

Penulisan laporan Tugas Akhir ini merupakan kelengkapan dan pemenuhan dari salah satu syarat dalam meyelesaikan studi Teknik Informatika dan memperoleh gelar Sarjana Komputer. Selain itu, tulisan ini untuk melatih mahasiswa untuk dapat menghasilkan suatu karya yang dapat dipertanggungjawabkan secara ilmiah, sehingga dapat bermanfaat bagi penggunanya.

Dalam menyelesaikan pembuatan program dan laporan ini, penulis telah banyak menerima bimbingan, saran, dan bantuan dari berbagai pihak. Untuk itu, dengan penuh syukur dan kerendahan hati, perkenankan penulis untuk menyampaikan ucapan terima kasih kepada :

1. Lucia Dwi Krisnawati, S.S., M.A selaku dosen pembimbing I yang telah membimbing dalam mengerjakan laporan ini.

2. Ir. Harianto Kristanto, M.T., M.M. selaku dosen pembimbing II yang telah membimbing dan memberikan saran selama pengerjaan Tugas Akhir.

3. Keluarga tercinta yang telah mendukungku untuk boleh menyelesaikan skripsi ini.

4. Ibu Uli dan Ibu Maria selaku pemilik Jago swalayan yang telah rela memberikan ijin untuk melakukan analisis pada data penjualan.

5. Heri Setiawan yang telah banyak memberi support.

6. Deddy Wijaya Sulianto yang banyak memberi masukan tentang system design diagram.

7. Anak-anak kos Ecco yang telah menghiburku ketika aku mengalami kejenuhan dalam mengerjakan tugas akhir ini.

8. Teman – teman lain dan pihak – pihak lain yang telah membantu dalam proses pengerjaan tugas akhir ini.

Penulis menyadari bahwa laporan dan program yang dibuat jauh dari sempurna. Dengan kerendahan hati, penulis mengharapkan kritik dan saran pengembangan untuk tugas akhir ini. Diharapkan dengan adanya kritik dan saran dari teman – teman dapat membantu dalam penyempurnaan tugas akhir ini.

Akhir kata, penulis meminta maaf apabila terdapat kesalahan selama proses pengerjaan tugas akhir ini. Semoga tugas akhir ini juga dapat berguna bagi teman-teman dalam memperoleh pengetahuan lain.

Yogyakarta, 18 Juli 2011

Penulis

Abstract

The presence of the Information Technology, especially database inside a company has been very common, it probably becomes the main need of a corporation. Nowadays, corporation does not only need database but also the knowledge of data. In Final Year proejct 2, the student is expected to work a project which is related with data mining (mining the knowledges) and local business, then the author worked in a project of Market Basket Analysis for Jago Mini-Market. Jago Mini-Market is a mini-market which sells food, beverages, clothes, and sewing kits material. There are so many transactions occuring in a day with various combinations. Even if there are many transactions in a day, but some items available on the shelves are expired too. Then the author created a system to analyze Jago Mini-Market's sewing kits sales, to find out which items correlated each others.

Market Basket Analysis is a modeling technique which is based on a theory that there is a correlation between items bought by customer. MBA is the common rule method which can be applied in the field of retail such as, mini–market, or as well as restaurant. MBA refers to methodology that studies the composition of a shopping basket within one shopping period.

The MBA result can be useds as a reference for manager to decide marketing strategy, such as discount items, items package, and pricing strategy. From the analysis, the support and confidence value of items pairs are so low because there are too many item variations in limited transaction records. This is caused by data itself which is a raw data with unstructured id, and unstructured item categorized.

Table of Contents

List of Table and Figures

List of Symbols and Abbreviations

CRM – Customer Relationship Management

DFD – Data Flow Diagram

IT – Information Technology

MBA – Market Basket Analysis

MEC – Mountain Equipment Co-op

RAD – Rapid Application Development

SQL – Structured Query Language

UKDW – Universitas Kristen Duta Wacana

Chapter 1 : Introduction

1.1. Project Summary

The project is related to the application that can support final-year project for IT undergraduate students along the process of developing the final documentation for their projects. Along this project, the application is also developed with lecturers of final-year project and project supervisor in mind. This project is intended to analyse the store's sales and also give suggestions to manager so she can make marketing strategies to increase store's sales. This project is implemented in dekstop application which is very user friendly. In begining, the author made observations and discussions with the owner and manager, so the requirements needed and model diagrams can be created. The administrator will do user management task, then the manager will analyse and create marketing strategy based on receipt data by market basket analysis results.

1.2. Company Background

Jago Mini-market was established in 1965. It is located in Urip Sumoharjo Street no.24 A. In Jogjakarta, Urip Sumoharjo Street is one of Jogjakarta big streets where many retail shops, super–market and mall are located in these areas. Its location is closed to Duta Wacana Christian University (UKDW), Budya Wacana Elementary School and Bethesda hospital.

Jago Mini-market is a family business, thus the crucial decisions relating to operation of Jago mini – market will be taken during the family discussions. Jago mini-market is owned by Ms. Maria and Ms. Uli. Ms. Maria is responsible for taking decisions related to finance while Ms. Uli acts as its manager. Ms Uli is responsible for the daily transaction at Jago Mini-market and also helps her elder-sister make decisions.

Jago Mini-market provides daily needs such as soaps, shampoos, detergents, liquid-cleansers, and etc. It also provides sewing kits such as needles, threads, ribbonss, and etc, household utensil, food and beverages too. Because it is located on strategic area, Jago Mini-market's customers are focused mainly on students, university students, and local customers.

In the future, Jago Mini-market expects to expand its business by opening branches or outlets in different parts of Jogjakarta. In spite of its tradional mini-market system, Jago Mini-market is able to compete with multinational supermarket because it has all products categories with low prices.

1.3. Issues with Current Data

Since it is run traditionally, Jago Mini-market does not have an Information System to support the transaction process. They still use hand-writing manual receipts and cash register in each division. Using hand-writing manual receipts may be more authentic than computerize print-out, but using hand-writing manual receipts also have disadvantages too.

With hand-writing manual receipts, the staffs work twice, they need to make copies of the data in the receipts after they finished their work time. They will summarize all of items sold in the wrap-up sheet. Each employee will hand-write the wrap-up sheet. Both receipts and wrap-up sheet then will be submitted to the manager. The author also scan some wrap-up sheet and hand-writing manual receipts in the Appendix A

The disadvantages of manual data keeping are the process is very time-consuming, and also problem in the validity of the data. The possibility of mistaken data input, unstructured data, and also redudant data.

If we use manual data keeping, we need to provide several times/hours to do the copies of the receipts. If we do it in rush, it might cause mistakes in inputing data so it is just wasting time. The next problem which always occur in manual data keeping is validity of the data that is less reliable. Some mistakes usually happen, such as writing inconsistently on item name, item price, how

many items which have been purchased and sometimes there happens wrong calculation too.

Another issue is unstructured data which occurred in Jago Mini-market. They do not have specific standardization in item name for example ribbon item has five colors which each color consist of five meters long. In Jago Mini-market, for item which has several colors they not use unique code to define the item in each color, such as rbn-p as pink ribbon, rbn-r as red ribbon and etc. They only give general name for all items, such as ribbon. This issue causes data redundancy. This mistake will make it difficult for Jago Mini-market to view the trends and predict the stock availability.

Beside the uncoded data, Jago Mini-market items placement was unstructured, for example, the placement of liquid-cleansers which is placed closed to placement of snacks, the placement of ribbons which is placed closed to baby stuffs. This situation can annoy the customers who come to the mini-market itself. On observation phase, the author found some items which are already dirty and obsolete. This situation is affected by the stock management and item placement. Sometimes we need to move the placement of item because item's sales depend also on item placement too.

The issues should be solved as soon as possible because Jago Mini-market's owners would expand their store to another place. We feel anxious if they would get trouble when the issues are not solved soon.

The latest issue is Jago Mini-market management only share sewing kits materials receipts and all receipts still in raw data. These problems might prevent the author to finish the project on schedule, because working with raw data was so challenging. Start from input, filter, clean, convert and process the data.

1.4. Formulation of Issues

The author also discuss several things in this project:

- How to migrate the raw data into Microsoft Access 2007?
- How to prepare the data before doing MBA (cleaning the data)?

- How to analyze the data using MBA concept to get the correlation between certain items?
- How the analysis result help in marketing plan?

1.5. Project Scope

The agreeable scope of the project involves developing a desktop system which covers a total of six use cases. These six functionalities serve the manager and admin. The system implementation will develop six use cases. For each actor who uses the system must login first that login use case is included in the system as "include" part (Fowler, 2004). In this system, the author will develop account access only for manager and admin. According to data warehouse characteristic, the author restrict the scope to focus on cleaning, filtering and analysis the data only. Because of limited data which only a month data or around 9,000 records, the author didn't use time period to filter the data. The author analysis the data which given from the Jago Mini-Market only. The MBA only analyze correlation between two categories and in each category will contains an item or more (maximum four items for each categories).

1.6. Benefits and Constraints of Proposed System

The proposed system (Market Based Analysis System) gives benefits for the owner (also as administrator) and the manager. For the owner, the system helps the owner to decide crucial decisions to Jago Mini-market. The owner capable to viewing the transaction by period, the items pair and etc. The owner also can view easily the purchasing patterns from certain period transactions. The owner also cooperates with manager to make marketing strategies for Jago Mini-market.

The manager is capable to view the items which have highest sales frequency, items which sales together or items pair (have correlation), and also help the owner to decide the marketing strategies manually by MBA results, such

as the promotion (coupons, cutting price, etc), special packages (Valentine packages, Holiday packages, and etc), discount items, and the placement of the items such as jam location near with bread, eggs near powder and etc. The system also helps owner and manager to print out the the item data with certain condition and also receipts based on certain conditions, such as period, and item name.

1.7. Project Description

The end-users of the Market Basket Analysis system consist of administrator and manager. The system is to be accessed via desktop. All users must register an account prior to accessing the system. A total of seven functionalities are included in the system.

The manager has five functionalities on the system. These functionalities are as follows:

- ✓ Log into the system (include)
- ✓ Manage information of the account
- ✓ Search items or receipts depend on certain criterias such as category of the item, period, and item itself based on item's sales or item support.
- ✓ Analyze the desired items from certain category.
- ✓ Reports based on items pair, by category, by period and etc.

The administrator has three functionalities on the system. These functionalities are as follows:

- ✓ Log into the system (include)
- ✓ Manage the users
- ✓ Search items or receipts depend on certain criterias such as category of the item, period, and item itself based on item's sales or item support.

The system will provide Graphical User Interface which contains the combo boxes, buttons and radio buttons to custom the certain condition, for

example limit support or confidence based on user input, minimum count of item's appearance, category of items and also the items itselves.

The report will be saved as *.rpt (Crystal Report extension). The report will help the owner and manager to decide several marketing strategies.

1.8. Project Aims

The aim of this project is to analyze the sales of sewing kit materials in Jago Mini-Market. This project is designed to find out items with the highest sales, the lowest sales and also correlation between the items. From analysis results, we can get the unique patterns to decide marketing strategy for that place.It also help the manager to improve the promotion, and pricing strategies to maximize the profit.

1.9. Project Objectives

The objectives for this project are as follows:

- To show the manager the items pair (for example: 'Renda Air' and 'Benang Extra', 'Jarum Jahit' and 'Benang Astra', and many more)
- To show the manager the item's support, confidence and lift based on certain transaction data
- To avoid the big number of unsold items that are available in the warehouse or stores
- To help the owner and the manager in deciding the appropriate pricing strategy

1.10.   Software and Hardware Requirement

Hardware requirements: laptop, USB Flash drive to back up the data

Software requirements: Microsoft Excel 2007, Microsoft Access 2007, Microsoft Visual Studio 2010, Crystal Report Viewer 2008, Star UML 5.02, Microsoft Project 2007, Microsoft Visio 2007, Microsoft Word 2007, Microsoft Power Point, Notepad, FastStone Capture, Balsamiq Mockups and Paint.

1.11. Research Methodology

The used methodoly to build the system is Market Basket Analysis. The author also studied for the methodology by :

- Book, On-line Serial, and E-Book

  The author use several books to gain the information about data mining, association rules and market basket analysis. The most of books discuss about theory of MBA and association rules so it really help the author to outline the literature review. The author also used on-line serial source to help her in coding concecpt and how to migrate the data into appropriate data structure.

- Consultation with The Supervisors

  The consultation activity is useful activity for the author to solve the problem. The supervisors always give suggestions for system development and also in report writing style.

- Observation and Testing

  This method is used when the author is finishing the system. The author did some tests to find out weaknesses from the system during system development.

1.12. Vision

As mentioned before, the latest issue is that Jago Mini-market management only gave the author sewing kits receipts and all receipts are still in raw data. These problems might prevent the author to finish the project on schedule, because working with raw data was so hard. The author need to input,

filter, clean, convert and process the data by herself. The author is working hard to input several data around 9000 rows to complete Market Basket Analysis standard. The author change the software for database from SQL Server 2008 to Microsoft Access 2007 because the author found out it work easier than SQL Server 2008. The author also use Crystal Report Viewer to generate the report.

Because of limited time to do this project, the second supervisor, Mr. Harianto Kristanto advised the author to restrict the requirement that we focus on data cleaning and specific analysis about correlation item between two categories only. He also suggest to build Market Basket Analysis system only. In this project, data cleaning phase is focused on manual cleaning because the author work with raw data which very unstructure and hard to clean.

From Final Year 1, the author already updated some components such as, as a result the title changes into "Analyzing" not for increasing, the DFD diagram, Use cases, and also revise the database design. The revisions are included in monthly reports and interim presentations.

Chapter 5 :   Conclusion

5.1. Summary

From this implementation, the author conclude several things from Market Basket Analysis system for analyzing Jago Mini-Market, there are :

1) Many variation of items in a receipt will be affected the value of support and confidence combinations. More items means less value will created

2) Lift value also help to prove the rules which created (based on combination of item, support and confidence value). If the lift value is less than 1, then X and Y is negatively correlated. If the lift value is greater than 1, then X and Y is positively correlated. If the lift value is equal to 0, then X and Y are independent and there is no correlation between them.

3) The system only analyze items pair because so many item variation but less speard of data.

4) The result of MBA showed minimum value of support, and confidence. This causes of unstructured item categorized, the speard of data is not balance, and receipt data is too small compare with variation of items.

5) From business view, this MBA system can not describe all knowledges inside data because the data structure not good enough to mining eventhough the author also cleaning and prepare the data manually. The data type is sewing kits which not too suit with market basket analysis. Sewing kits's price are so cheap, so the probabilities of combination are very high.

6) At least, we can view the items pair, eventhough with minimum value. From items pair, the manager can decided the placement of items, and also pricing strategy, for example : if we buy one piece of "Benang Jahit", then it will cost IDR 1000. If we buy 4 pieces of it, then it will cost IDR 3500. It is IDR 500 cheaper than we buy a piece.

Based on technical feasibility, there are some issues related with implementation the system at Jago Mini-Market,

- There is no technology which needed such as computer, operation system and also system information.
- There is no IT resource beside the author

Based on economic feasibility, there are some issues related with system implementation at Jago Mini-Market,

- The budget to develop the system itself is very limited
- Tax consideration, Jago Mini-Market's assumption that using information system will ruin the income tax calculation and also they are affraid of the tax increase because of the automatic report results.

5.2. Future Work

Based on the issues related with technical feasibility, the author suggest several things,

- Before Jago Mini-Market implemented information system, they should make the clear items list and their spesific characters or categories, so it can help in building effective information system.
- Jago Mini-Market must provide the technology, such as computer, OS and the information system which the information system can be implemented in daily needs mini-market and also the information system can be provided by IT or IS student who take 'KP' project to minimize the cost.
- Jago Mini-Market need some staffs to inputed the raw data into information system databases.
- Jago Mini-Market should train an employee or a manager to operate information system and also the current system.

Based on the issues related with economical feasibility, the author suggest several things,

- The budget for information system can be reduced if the system is build by IT students but it also needed human resource to input and maintain the system. The human resource can be the manager or staffs.

- For tax problem, the information system can be splitted into two authorization. First for general view which only show general transaction and it should show less receipts than the second ones which contains the real detail.

- After implemented the information system, MBA system will be more useful to analyze the correlation item in daily needs section such as food and beverages, and household utensil. From MBA, the manager can get references which items can be paired in packages and which item's price will be reduced and many more.

References

Ambler, S. W. (2010). *UML 2 Sequence Diagram*. [On-line serial]. Retrieved at 14 July 2011 from http://www.agilemodeling.com/ artifacts/sequenceDiagram.htm

Anonymous. (2010). *Real – time*. [On-line serial]. Retrieved at 27 August 2010 from http://www.audioenglish.net/dictionary/real_time.htm.

Berry, Michael J. A. (2000). *Mastering Data Mining*. United States : Wiley.

Berry, Michael J. A., & Linoff, G. (1997). *Data Mining Techniques : For Marketing, Sales & Customer Support*. United States : Wiley.

Blue Martini Software. (2002). *"Blue Martini Business Intelligence at Work : Charting the Terrains of MEC Website Data"*. [pdf]. California, CA : Author. Retrieved at 10 October 2010 from website http://www.bluemartini.com/bi.

Case Maker. (2000). *"Rapid Application Development"*. [pdf]. Retrieved at 15 September 2010 from website http://www.casemaker.com.

CMS. (2008, March 27). *"Selecting The Development Approach"*. [pdf]. Retrieved at 10 September 2010 from http://www.cms.gov/ systemlifecycleframework/downloads/selectingdevelopmentapproach.p df.

Han, J. & Kamber, M. (2006). *Data Mining : Concepts and Techniques.* 2nd Edition. San Fransisco,CA : Morgan Kaufman Publisher.

Inmon, Bill. (1995). "What is a Data Warehouse?". Prism Tech Topic, Vol. 1, No. 1. Retrieved from http://www.intranetjournal.com/ features/datawarehousing.html.

Ittelkom. (2010). *"Data Warehouse"*. Retrieved from http://www.ittelkom.ac.id/ myhut.org/public/datawarehouse.doc.

Microsoft. (2010). *Microsoft Visual Studio 2010.* Retrieved at 25 November 2010 from http://www.microsoft.com/visualstudio/en-us/products/2010-editions.

Mountain Equipment Co-Op. (2010). *"About MEC"*. [On-line serial]. Retrieved at 14 October 2010 from http://www.mec.ca/Main/home.jsp.

Noulo. (2010) *RAD (Rapid Application Development)*. [On-line serial]. Retrieved at 10 September 2010 from http://www.novulo.com/ AboutNovulo.aspx.

Olson, DL, & Delen, D. (2008). *Advanced Data Mining Techniques*. Germany : Springer – Verlag.

Sin, Thant. (2006). *"Usability of Analysis Sequence Diagram"*. [pdf] Florida, FL : Florida International University. Retrieved at 22 October 2010 from http://www.fia.edu.

Trust System & Software. (2010). *RAD Diagram.* [online]. Retrieved at 10 September 2010 from http://softtrust.com.

Yudianton. (2004). *Implementasi Market Basket Analysis Untuk Mendapatkan Pengetahuan Baru Bagi Pembeli Buku : Studi Kasus Penjualan Buku di Toga Mas Yogyakarta.* [On-line serial]. Retrieved at 15 February 2011 from SinTA UKDW website : http://sinta.ukdw.ac.id/sinta/search.jsp?query=market+basket+analysis&btnsearch=Cari.

Blibiography

Clipart Guide. (2010). Retrieved at 12 November 2010 from http://www.clipartguide.com/_pages/0511-1009-2316-4637.html.

Cylogy. (2010). "Perzonalization Overview". [pdf] Retrieved at 20 November 2010 from http://www.cylogy.com.

Effexis. (2009). UML Sequence Diagram Tutorial. [On-line serial]. Retrieved at 14 July 2011 from http://www.sequencediagrameditor.com/ uml/sequence-diagram.htm

The Fact Point Group. (2008). "Leading Practice in Market Basket Analysis". [pdf]. Los Altos, CA. Retrieved at 19 October 2010 from http://www.factpoint.com.

Fowler, Martin. (2004). UML Distilled. 3rd Edition. San Fransisco, CA : Pearson.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). "From Data Mining to Knowledge Discovery in Databases". [On-line serial]. Rhode Island, RI : AAAI. Retrieved at 14 October 2010 from http://www.aaai.org/ conferences/national/1997/aaai97.html.

Giudici, P., & Figini, S. (2009). Applied Data Mining for Business and Industry. 2nd Edition. United Kingdom : Wiley.

SAS Institute, Research and Development, 2004. "On Analytical Tools for Market (Associations) Analysis". [Power-Point Slides]. New York, NY. Retrieved at 17 October 2010.

Schmuller, Joseph. (2006). "Sams Teach Yourself UML in 24 Hours". 3rd Edition. United States : Sams.

Thearling. (2010). "An Introduction to Data Mining". [WWW page]. Retrieved at 24 November 2010 from http://www.thearling.com/ text/dmwhite/dmwhite.htm.

Vercellis, Carlo. (2009). Business Intelligence : Data Mining and Optimization for Decision Making. United Kingdom : Wiley.

Witten, I.H., & Frank, E. (2005). Data Mining : Practical Machine Learning Tools and Techniques. 2nd Edition. San Francisco, CA : Morgan Kaufmann.