SHIV SHAKTI

International Journal of in Multidisciplinary and Academic Research (SSIJMAR)

Vol. 4, No. 3, June 2015 (ISSN 2278 – 5973)

CONCEPTUALIZATION OF DATA MINING TECHNIQUE USING APRIORI ALGORITHM+ IN THE ANALYSIS OF SALES AND CONSUMER’S PURCHASING

Ms. Nitika Siger
M. Tech Student
Department of Computer Science & Engineering
MDU University Harayana, India

Dr. Pushpender Sarao
Dean, S.I.T.M., Rewari

Impact Factor = 3.133 (Scientific Journal Impact Factor Value for 2012 by Inno Space Scientific Journal Impact Factor)

Global Impact Factor (2013)= 0.326 (By GIF)

Indexing:

CONCEPTUALIZATION OF DATA MINING TECHNIQUE USING APRIORI ALGORITHM+ IN THE ANALYSIS OF SALES AND CONSUMER’S PURCHASING

Ms. Nitika Siger
M. Tech Student
Department of Computer Science & Engineering
MDU University Harayana, India

Abstract : The analysis of sales and consumer’s purchasing is the hottest topic in the current research scenario. Before developing this system, we assume that there is a supermarket that wishes to identify its frequently purchased item-sets i.e. they want to know that if customers purchase a product X, what is their tendency or probability to buy another product Y or product Z along with it. In this way, the supermarket can keep all those products along side in order to increase their sales. In other words, let’s assume that there is a super mart with 1,000 products. A lot of customers purchase a lot of products every day. There could be some products which are at very far-off places for which the customers have to walk around this whole super mart in order to get them and there could be some products which are kept together but they are not actually required by the customers at the same time or in other words, they don’t make for pair-purchasing. Hence, the supermarket authorities want to find & analyze what similar products should be kept in one single shelf that match with the common purchasing-tendency or mentality of the customers so that the customers do not waste their invaluable time in walking around a large super mart and can purchase all the required items in a minimum possible timeframe. This will also eventually increase the sales & goodwill of the super mart for bringing in ease & efficiency in product purchasing.

Keywords : Java GUI ,JFC, AWT,Microsoft Office Access, Visual Basic, DAO, VBA, Microsoft SQL Server , Microsoft Access Jet Database (accdb and mdb formats).

1. Introduction :

Apriori is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database: this has applications in domains such as market basket analysis. Apriori employs an iterative approach known as level-wise search, where **k-itemsets** are used to explore **(k+1)-itemsets**. First, the set of frequent **1-itemsets** is found by scanning the database to accumulate the count for each item, and collecting those items that satisfy minimum support. The resulting set is denoted as **L1**. Next, L1 is used to find **L2**, the set of frequent **2-itemsets**, which is used to find L3, and so on, until no more frequent **k-itemsets** can be found. The finding of each Lk requires one full scan of the database.

At final iteration you will end up with many **k-itemsets** which is basically called **association rules**. To select interesting rules from the set of all possible rules various constraint **measures** such as **support** and **confidence** is applied.

Let I = {i1, i2, ..., in} be a set of n binary attributes called items. Let D = {t1, t2, ..., tn} be a set of transactions called the database. Each transaction in D has a unique transaction ID and contains a subset of the items in I. A rule is defined as an implication of the form $X \rightarrow Y$ where $X, Y \subseteq I$ and $\cap = \emptyset$. The sets of items (for short itemsets) X and Y are called antecedent (left-hand-side or LHS) and consequent (right-hand-side or RHS) of the rule respectively.

1.1 The Foundations of Data Mining

Data mining techniques are the result of a long process of research and product development. This evolution began when business data was first stored on computers, continued with improvements in data access, and more recently, generated technologies that allow users to navigate through their data in real time. Data mining takes this evolutionary process beyond retrospective data access and navigation to prospective and proactive information delivery. Data mining is ready for application in the business community because it is supported by three technologies that are now sufficiently mature:

- Massive data collection
- Powerful multiprocessor computers
- Data mining algorithms

Commercial databases are growing at unprecedented rates. A recent META Group survey of data warehouse projects found that 19% of respondents are beyond the 50 gigabyte level, while 59% expect to be there by second quarter of 1996.1 In some industries, such as retail, these numbers can be much larger. The accompanying need for improved computational engines can now be met in a cost-effective manner with parallel multiprocessor computer technology. Data mining algorithms embody techniques that have existed for at least 10 years, but have only recently been implemented as mature, reliable, understandable tools that consistently outperform older statistical methods.

In the evolution from business data to business information, each new step has built upon the previous one. For example, dynamic data access is critical for drill-through in data navigation applications, and the ability to store large databases is critical to data mining. From the user's point of view, the four steps listed in Table 1 were revolutionary because they allowed new business questions to be answered accurately and quickly.

The core components of data mining technology have been under development for decades, in research areas such as statistics, artificial intelligence, and machine learning. Today, the maturity of these techniques, coupled with high-performance relational database engines and broad data integration efforts, make these technologies practical for current data warehouse environments.

1.2 The Scope of Data Mining

Data mining derives its name from the similarities between searching for valuable business information in a large database — for example, finding linked products in gigabytes of store scanner data — and mining a mountain for a vein of valuable ore. Both processes require either sifting through an immense amount of material, or intelligently probing it to find exactly where the value resides. Given databases of sufficient size and quality, data mining technology can generate new business opportunities by providing these capabilities:

Evolutionary Step	**Business Question**	**Enabling Technologies**	**Product Providers**	**Characteristics**
Data Collection (1960s)	"What was my total revenue in the last five years?"	Computers, tapes, disks	IBM, CDC	Retrospective, static data delivery
Data Access (1980s)	"What were unit sales in New England last March?"	Relational databases (RDBMS), Structured Query Language (SQL), ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Retrospective, dynamic data delivery at record level
Data Warehousing & Decision Support (1990s)	"What were unit sales in New England last March? Drill down to Boston."	On-line analytic processing (OLAP), multidimensional databases, data warehouses	Pilot, Comshare, Arbor, Cognos, Micro strategy	Retrospective, dynamic data delivery at multiple levels
Data Mining (Emerging Today)	"What's likely to happen to Boston unit sales next month? Why?"	Advanced algorithms, multiprocessor computers, massive databases	Pilot, Lockheed, IBM, SGI, numerous startups (nascent industry)	Prospective, proactive information delivery

Table 1. Steps in the Evolution of Data Mining.

1.3 How Data Mining Works

How exactly is data mining able to tell you important things that you didn't know or what is going to happen next? The technique that is used to perform these feats in data mining is called modeling. Modeling is simply the act of building a model in one situation where you know the answer and then applying it to another situation that you don't. For instance, if you were looking for a sunken Spanish galleon on the high seas the first thing you might do is to research the times when Spanish treasure had been found by others in the past. You might note that these ships often tend to be found off the coast of Bermuda and that there are certain characteristics to the ocean currents, and certain routes that have likely been taken by the ship's captains in that era. You note these similarities and build a model that includes the characteristics that are common to the locations of these sunken treasures. With these models in hand you sail off looking for treasure where your model indicates it most likely might be given a similar situation in the past. Hopefully, if you've got a good model, you find your treasure.

1.4 An Architecture for Data Mining

To best apply these advanced techniques, they must be fully integrated with a data warehouse as well as flexible interactive business analysis tools. Many data mining

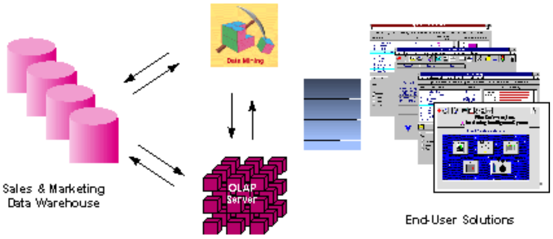


Figure 1 - Integrated Data Mining Architecture

Conclusion

Comprehensive data warehouses that integrate operational data with customer, supplier, and market information have resulted in an explosion of information. Competition requires timely and sophisticated analysis on an integrated view of the data. However, there is a growing gap between more powerful storage and retrieval systems and the users' ability to effectively analyze and act on the information they contain. Both relational and OLAP technologies have tremendous capabilities for navigating massive data warehouses, but brute force navigation of data is not enough. A new technological leap is needed to structure and prioritize information for specific end-user problems. The data mining tools can make this leap.

1.5 Introduction to apriori algorithm in data mining

In computer science and data mining, Apriori is a classic algorithm for learning association rules. Apriori is designed to operate on databases containing transactions (for example, collections of items bought by customers, or details of a website frequentation). Other algorithms are designed for finding association rules in data having no transactions (Winepi and Minepi), or having no timestamps (DNA sequencing).

1.5.1 ALGORITHM

Association rule mining is to find out association rules that satisfy the predefined minimum support and confidence from a given database. The problem is usually decomposed into two sub problems.

- Find those itemsets whose occurrences exceed a predefined threshold in the database; those itemsets are called frequent or large itemsets.
- Generate association rules from those large itemsets with the constraints of minimal confidence.

Suppose one of the large itemsets is $L_k = \{I_1, I_2, ..., I_k\}$; association rules with this itemsets are generated in the following way: the first rule is $\{I_1, I_2, ..., I_{k-1}\} = > \{I_k\}$. By checking the confidence this rule can be determined as interesting or not. Then, other rules are generated by deleting the last items in the antecedent and inserting it to the consequent, further the confidences of the new rules are checked to determine the interestingness of them. This process iterates until the antecedent becomes empty.

Since the second subproblem is quite straight forward, most of the research focuses on the first subproblem. The Apriori algorithm finds the frequent sets L in Database D.

- Find frequent set L_{k-1}.
- Join Step.
  - ○ C_k is generated by joining $L_{k-1}$with itself
- Prune Step.
  - ○ Any $(k-1)$ -itemset that is not frequent cannot be a subset of a frequent k -itemset, hence should be removed.

where

- (C_k: Candidate itemset of size k)
- (L_k: frequent itemset of size k)

Apriori Pseudo code

```
Apriori $(T, \varepsilon)$
$L_1 \leftarrow \{$ large 1-itemsets that appear in more than $\varepsilon$ transactions }
```

```
$k \leftarrow 2$
*while* $L_{k-1} \neq \varnothing$
    $C_k \leftarrow$ *Generate*($L_{k-1}$)
    *for transactions* $t \in T$
        $C_t \leftarrow$ Subset($C_k$,$t$)
        *for candidates* $c \in C_t$
            $\mathrm{count}[c] \leftarrow \mathrm{count}[c] + 1$
    $L_k \leftarrow \{c \in C_k | \ \mathrm{count}[c] \geq \varepsilon\}$
    $k \leftarrow k + 1$
        $\bigcup L_k$
*return* $k$
```

LITERATURE SURVEY

2.1 Problem Description

A lot of customers purchase a lot of products every day. There could be some products which are at very far-off places for which the customers have to walk around the whole super mart in order to get them and there could be some products which are kept together but they are not actually required by the customers at the same time or in other words, they don't make for pair-purchasing. Suppose, if customer purchases bread, then there is a lot of tendency or probability that he/she can purchase Butter, Eggs or Milk or all three along with it. But there is less probability that he/she will purchase a toothbrush along with it because that does not match with the common purchasing tendency or mentality of any customer. Even worse, if the customers find that Bread & Butter are kept at two extreme ends of the supermarket, they may purchase either of them or may not purchase any of them due to reluctance which is a common human behaviour and this will surely affect the sales of the supermarket.

2.2 LITERATURE SURVEY ON INTEGRATION OF INDUCTION AND DEDUCTION FOR MINING SUPERMARKET SALES DATA

The objective of this paper is precisely to demonstrate how a suitable integration of deductive reasoning, such as that supported by logic database languages, and inductive reasoning, such as that supported by association rules, provides a viable solution to many high-level problems in market basket analysis. We briefly present Datasift, a prototype system for the analysis of supermarket sales data, based on an architecture that integrates the deductive capabilities of a logic-based database language, LDL++ [1], with the inductive capabilities of diverse data mining algorithms and tools, notably association rules.

2.3 LITERATURE SURVEY ON IMPLEMENTATION OF WEB USAGE MINING USING APRIORI AND FP GROWTH ALGORITHM

Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from Web data, in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site. Web usage mining itself can be classified further depending on the kind of usage data considered. They are web server data, application server data and application level data. Web server data correspond to the user logs that are collected at Web server.

2.4 LITRATURE SURVEY ON FREQUENT DATA ITEMSET MINING USING VS APRIORI ALGORITHM

The organization, management and accessing of information in better manner in various data warehouse applications have been active areas of research for many researchers for more than last two decades. The work presented in this paper is motivated from their work and inspired to reduce complexity involved in data mining from data warehouse. A new algorithm named VS_Apriori is introduced as the extension of existing Apriori Algorithm that intelligently mines the frequent data itemset in large scale database. Experimental results are presented to illustrate the role of Apriori Algorithm, to demonstrate efficient way and to implement the Algorithm for generating frequent data itemset. Experiments are also performed to show high speedups.

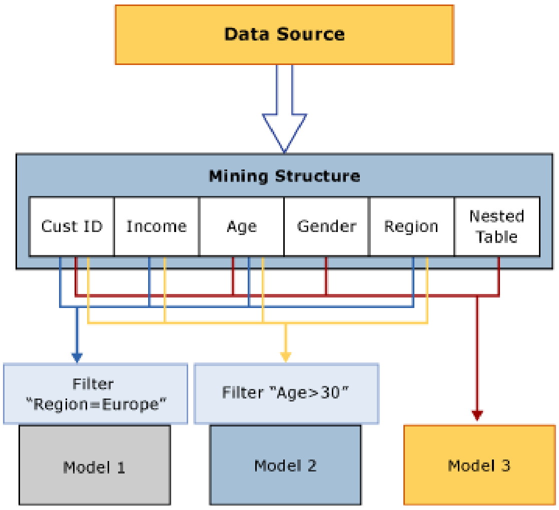
2.5 LITERATURE SURVEY ON DATA MINING: A COMPETITIVE TOOL IN THE BANKING AND RETAIL INDUSTRIES

Data might be one of the most valuable assets of any corporation—but only if it knows how to reveal valuable knowledge hidden in raw data. Data mining allows to extract diamonds of knowledge from the historical data, and predict outcomes of future situations. It helps optimise business decisions, increase the value of each customer and communication, and improve customer satisfaction

3. DEVELOPMENT OF AN DATA MINING BY USING APRIORI ALGORITHM

3.1 Mining Structures (Analysis Services - Data Mining)

The mining structure defines the data from which mining models are built: it specifies the source data view, the number and type of columns, and an optional partition into training and testing sets. A single mining structure can support multiple mining models that share the same domain. The following diagram illustrates the relationship of the data mining structure to the data source, and to its constituent data mining models.



The mining structure in the diagram is based on a data source that contains multiple tables or views, joined on the CustomerID field. One table contains information about customers, such as the geographical region, age, income and gender, while the related nested table contains multiple rows of additional information about each customer, such as products the customer has purchased. The diagram shows that multiple models can be built on one mining structure, and that the models can use different columns from the structure.

Model 1 Uses CustomerID, Income, Age, Region, and filters the data on Region.

Model 2 Uses CustomerID, Income, Age, Region and filters the data on Age.

Model 3 Uses CustomerID, Age, Gender, and the nested table, with no filter.

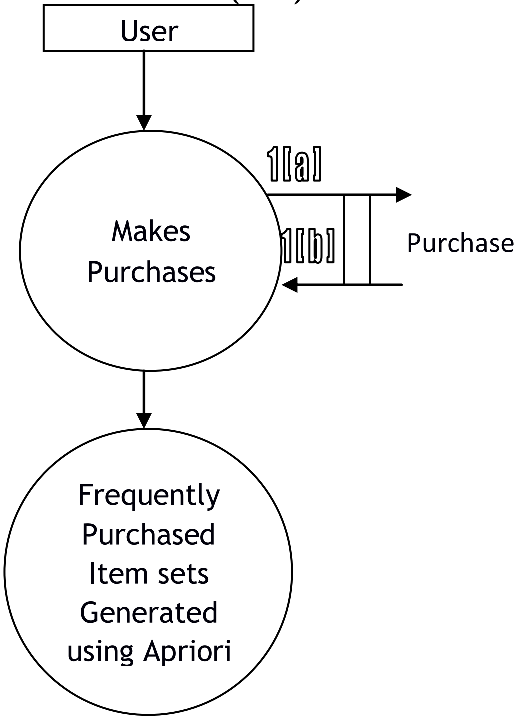
3.2 STEPS USED FOR RESULT

Under the implementation of this application, we have followed the below points:

- First we enabled the user to select the appropriate industrial sector from which he/she wishes to analyze the sales of the products purchased. This was achieved through the concept implementation of Java Swings in which the interface provided a drop-down list that empowered the user to select the industrial sector.
- Thereafter, depending upon the selection preference, we presented 10 products along with the check boxes to make the user select from those in order to purchase them and make a single unique transaction. We achieved this through the use of Delegation Event Model employed GUI programming within Java using Swings.
- Then, all the purchases that the user made were committed to the database using JDBC through which we ensured that every single transaction goes into the database with a unique transaction ID and with the products being purchased by the user.
- Once the user planned to exit from the super market and has finished his purchasing, we triggered our algorithm which fetched all the records from the database one by one.
- Then, according to the algorithm, we made the use of Java collection framework, which is essentially the dynamic data structures in Java in order to create dynamic tokens and candidate elements to recursively create the sets of frequently purchased item-sets in order to make the user, which is the super market head in our case, to determine the products that were frequently purchased in pairs or sets of 2, 3, 4, etc.

Here also, we employed the use of Java Swings along with Collections framework

3.3 DATA FLOW DIAGRAM (DFD)

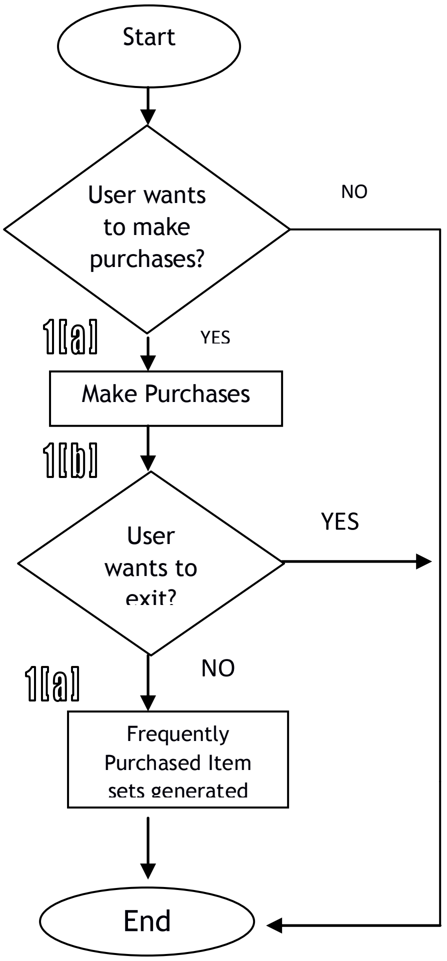


Description of inputs & outputs above:

1[a]: Items Purchased

1[b]: Transaction id of the purchase made

3.4 FLOWCHART


Start
User wants to make purchases?
NO
1[a]
YES
Make Purchases
1[b]
User wants to exit?
YES
1[a]
NO
Frequently Purchased Item sets generated
End

1[a]: Items Purchased

1[b]: Transaction id of the purchase made

4 EXPERIMENTAL RESULTS

In our project, we have planned to implement the above idea by taking the example of 10 computer hardware items like Monitor, CPU, Printer etc. from which the user can make a purchase.
After the user selects any item(s) for purchasing, it is recorded in the database. The user is allowed to make any no. of purchases that he/she wishes. Once he/she has finished making the purchases, the Apriori algorithm will mine the purchase-data stored in the database and will generate the n frequently-purchased item-sets for us i.e. in this case, it will generate 1 frequently-purchased item-set, 2 frequently-purchased item-sets.... up to 10 frequently-purchased item-sets since there are 10 items that we assume

4.1 GUI TOOL USED

Swing is the primary Java GUI widget toolkit. It is part of Oracle's Java Foundation Classes (JFC) — an API for providing a graphical user interface(GUI) for Java programs.
Swing was developed to provide a more sophisticated set of GUI components than the earlier Abstract Window Toolkit (AWT). Swing provides a native look and feel that emulates the look and feel of several platforms, and also supports a pluggable look and feel that allows applications to have a look and feel unrelated to the underlying platform. It has more powerful and flexible components than AWT. In addition to familiar components such as buttons, check box and labels, Swing provides several advanced components such as tabbed panel, scroll panes, trees, tables and lists.
Unlike AWT components, Swing components are not implemented by platform-specific code. Instead they are written entirely in Java and therefore are platform-independent. The term "lightweight" is used to describe such an element

4.2 DATABASE USED

Microsoft Office Access, previously known as Microsoft Access, is a database management system from Microsoft that combines the relational Microsoft Jet Database Engine with a graphical user interface and software-development tools. It is a member of the Microsoft Office suite of applications, included in the Professional and higher editions or sold separately. Software developers and data architects can use Microsoft Access to develop application software, and "power users" can use it to build software applications. Like other Office applications, Access is supported by Visual Basic for Applications, an object-oriented programming language that can reference a variety of objects including DAO (Data Access Objects), ActiveX Data Objects, and many other ActiveX components. Visual objects used in forms and reports expose their methods and properties in the VBA programming environment, and VBA code modules may declare and call Windows operating-system functions.

Uses

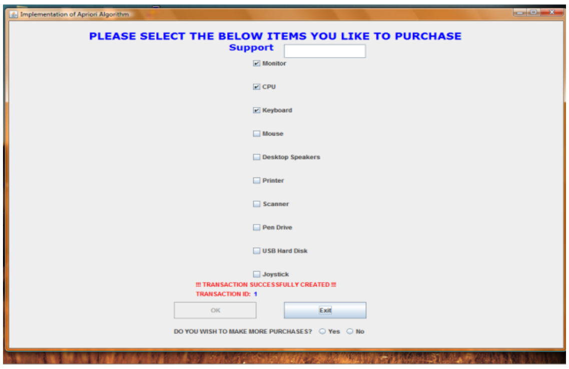
In addition to using its own database storage file, Microsoft Access may also be used as the 'front-end' with other products as the 'back-end' tables, such as Microsoft SQL Server and non-Microsoft products such as Oracle and Sybase. Multiple backend sources can be used by a Microsoft Access Jet Database (accdb and mdb formats).

Features

Users can create tables, queries, forms and reports, and connect them together with macros. Advanced users can use VBA to write rich solutions with advanced data manipulation and user control. Access also has report creation features that can work with any data source that Access can "access".

4.3 SCREENSHOTS

Making 7 Purchase Transactions



PLEASE SELECT THE BELOW ITEMS YOU LIKE TO PURCHASE

Support

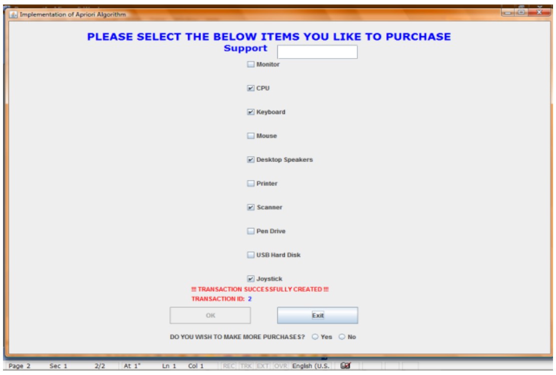
☑ Monitor
☑ CPU
☑ Keyboard
□ Mouse
□ Desktop Speakers
□ Printer
□ Scanner
□ Pen Drive
□ USB Hard Disk
□ Joystick

!!! TRANSACTION SUCCESSFULLY CREATED !!!

TRANSACTION ID: 1

DO YOU WISH TO MAKE MORE PURCHASES? ○ Yes ○ No

Generating Frequently Purchased Item-Sets with Support = 8


Implementation of Apriori Algorithm

PLEASE SELECT THE BELOW ITEMS YOU LIKE TO PURCHASE

Support

□ Monitor
☑ CPU
☑ Keyboard
□ Mouse
☑ Desktop Speakers
□ Printer
☑ Scanner
□ Pen Drive
□ USB Hard Disk
☑ Joystick

!!! TRANSACTION SUCCESSFULLY CREATED !!!

TRANSACTION ID: 2

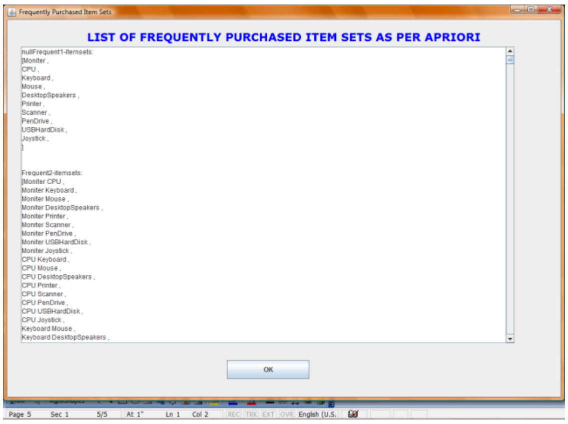
DO YOU WISH TO MAKE MORE PURCHASES? ○ Yes ○ No

Page 2 Sec 1 2/2 At 1" Ln 1 Col 1 REC TRK EXT OVR English (U.S.

List of frequently purchased item-sets in the GUI

Note: The below screenshots have been taken through scrolling the window for viewing Frequent**n-**itemsets (for the above 7 transactions) where n = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

Frequent 1 & 2 item-sets



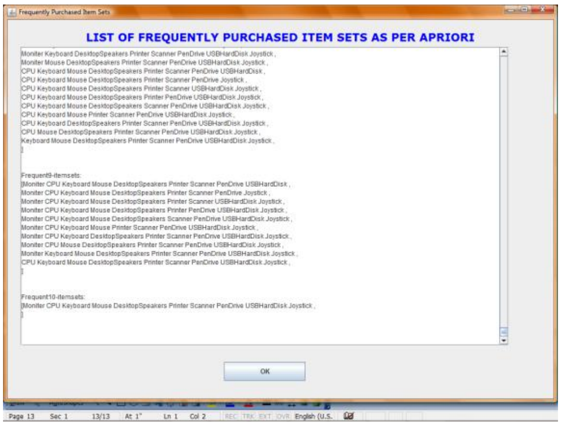
LIST OF FREQUENTLY PURCHASED ITEM SETS AS PER APRIORI

```
nullFrequent1-itemsets:
[Moniter ,
CPU ,
Keyboard ,
Mouse ,
DesktopSpeakers ,
Printer ,
Scanner ,
PenDrive ,
USBHardDisk ,
Joystick ,
]


Frequent2-itemsets:
[Moniter CPU ,
Moniter Keyboard ,
Moniter Mouse ,
Moniter DesktopSpeakers ,
Moniter Printer ,
Moniter Scanner ,
Moniter PenDrive ,
Moniter USBHardDisk ,
Moniter Joystick ,
CPU Keyboard ,
CPU Mouse ,
CPU DesktopSpeakers ,
CPU Printer ,
CPU Scanner ,
CPU PenDrive ,
CPU USBHardDisk ,
CPU Joystick ,
Keyboard Mouse ,
Keyboard DesktopSpeakers ,
```

OK
Page 5 Sec 1 5/5 At 1" Ln 1 Col 2 REC TRK EXT OVR English (U.S.

Frequent 9 & 10 item-sets



LIST OF FREQUENTLY PURCHASED ITEM SETS AS PER APRIORI

```
Moniter Keyboard DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
Moniter Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
CPU Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk ,
CPU Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive Joystick ,
CPU Keyboard Mouse DesktopSpeakers Printer Scanner USBHardDisk Joystick ,
CPU Keyboard Mouse DesktopSpeakers Printer PenDrive USBHardDisk Joystick ,
CPU Keyboard Mouse DesktopSpeakers Scanner PenDrive USBHardDisk Joystick ,
CPU Keyboard Mouse Printer Scanner PenDrive USBHardDisk Joystick ,
CPU Keyboard DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
CPU Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
]


Frequent9-itemsets:
[Moniter CPU Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk ,
Moniter CPU Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive Joystick ,
Moniter CPU Keyboard Mouse DesktopSpeakers Printer Scanner USBHardDisk Joystick ,
Moniter CPU Keyboard Mouse DesktopSpeakers Printer PenDrive USBHardDisk Joystick ,
Moniter CPU Keyboard Mouse DesktopSpeakers Scanner PenDrive USBHardDisk Joystick ,
Moniter CPU Keyboard Mouse Printer Scanner PenDrive USBHardDisk Joystick ,
Moniter CPU Keyboard DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
Moniter CPU Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
Moniter Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
CPU Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
]


Frequent10-itemsets:
[Moniter CPU Keyboard Mouse DesktopSpeakers Printer Scanner PenDrive USBHardDisk Joystick ,
]
```

OK
Page 13 Sec 1 13/13 At 1" Ln 1 Col 2 REC TRK EXT OVR English (U.S.

4.4 METHODOLOGY

The basic objective of software engineering is to develop methods and procedures for software development that can scale up for large systems and that can be used to consistently produce high quality software at low cost and with a small cycle time. That is, the key objectives are consistency, low cost, high quality, small cycle time, and scalability.

We have developed this system after duly spending time on each software development phase individually and freezing the status before we move on to the next phase i.e. we have used **Linear Sequential** model in this application

CONCLUSION AND FUTURE WORK

We can easily correlate the above prototype with a large super mart with thousands of products in which our algorithm would generate the item-sets that are frequently-purchased by the customers in the pairs of 1, 2, 3 or more. And using that data, those items can be kept along side in order to increase the sales & goodwill of the super mart through saving the customer's time & effort during purchasing. Apriori algorithm offers a reliable technique for accessing frequent data itemset. It helps in managing transaction in controlled manner. It also helps to manage various services, like monitoring, planning and execution of transaction for frequent data itemset mining in intelligent manner.

Future Work

- We can utilize the algorithm for the benefit and sales analysis of different industrial sectors.

- In addition to sales, we can create an inbuilt intelligence that not only tells us about most-purchased products but also predict its sales for future considering the present statistics.

- We can further enhance this application so that it can be deployed over web and could be executed globally. This way the users shall become aware of the most prominent products sold globally thereby enhancing the brand value.

- This application could also employ the hybrid use of multiple algorithms in which apart from sales analysis, various other metrics like productivity, storage, processing, etc. could also be analyzed.

- This application holds a considerable level of enhancement in terms of integration with hardware, biometrics and other automated systems wherein once the product is product and sales is finalized, the application could gather that data and use it for fruitful business intelligence.

- The application can further be enhanced by inclusion of cloud mechanisms in same wherein the data from different locations could be fetched from the remote systems located across the world and could be analyzed from different geographical locations for an understanding about the people's purchasing preference and interests.

References

[1] Gosling, James, A brief history of the Green project. Java.net, no date ca. Q1/1998]. Retrieved April 29, 2007.

[2] Gosling, James; Joy, Bill; Steele, Guy L., Jr.; Bracha, Gilad (2005). The Java Language Specification (3rd ed.). Addison-Wesley. ISBN 0-321-24678-0.

[3] Lindholm, Tim; Yellin, Frank (1999). The Java Virtual Machine Specification (2nd ed.). Addison-Wesley. ISBN 0-201-43294-3.

[4] java.com - Java for end-users

[5] Oracle's Developer Resources for Java Technology.

[6] Java SE 7 API Javadocs

[7] Oracle's Beginner's tutorial for Java SE Programming

[8] A Brief History of the Green Project

[9] Michael O'Connell: Java: The Inside Story, SunWord, July 1995.

[10] Patrick Naughton: Java Was Strongly Influenced by Objective-C (no date).

[11]David Bank: The Java Saga, Wired Issue 3.12 (December 1995).

[12] Shahrooz Feizabadi: A history of Java in: Marc Abrams, ed., World Wide Web – Beyond the Basics, Prentice Hall, 1998.

[13] Patrick Naughton: The Long Strange Trip to Java, March 18, 1996

[14] Open University (UK): M254 Java Everywhere (free open content documents).

[15]is-research GmbH: List of programming languages for a Java Virtual Machine.

[16] How Java's Floating-Point Hurts Everyone Everywhere, by W. Kahan and Joseph D. Darcy, University of California, Berkeley.

[17] http://www.coderanch.com/forums - A good community for discussing java concerns.

[18]R.Agrawal, S.Sarawagi, S.Thomas. Integrating Association Rule Mining with Relational Database Systems: Alternatives and Implications. In Procs. of ACM-SIGMOD'98, 1998.

[19]R.Agrawal, R.Srikant. Fast Algorithms for Mining Association Rules. In Procs. of 20th Int'l Conference on Very Large Databases, 1994.

[20]E.Baralis, G.Psaila. Incremental Refinement of Association Rule Mining. In Procs. Of SEBD'98, 1998.

[21]M. J. A. Berry, G.Linoff. Data Mining Techniques for Marketing Sales, and Customer Support. Wiley, 1997.