 ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Analysis and Prediction of Future Manufacturing Requirements of Pharmaceuticals Industry

[1]Nishtha Kesswani, [2]Sanjay Kr. Garg, [3]Mohit Jamer
[1,2,3]Central University of Rajasthan, BandarSindri, Ajmer, Rajasthan, India

Abstract
This paper presents the use of data mining technique to do the prediction for pharmaceutical Industry. It is very important to gather data from different data sources store and maintain the data, generate information and knowledge so that prediction can be done for future growth. To examine this vast amount of data and drawing fruitful conclusion and inferences it needs the special tools called data mining tools. This paper reviews the prediction on some real data, analyze and conclude that manufacturing of medicines done in pharmaceutical industry should according to constraint in market.

Keywords
Prediction, Pharmaceutical

I. Introduction
This paper presents the use of data mining technique to do the prediction for pharmaceutical Industry. It is very important to gather data from different data sources store and maintain the data, generate information and knowledge so that prediction can be done for future growth. To examine this vast amount of data and drawing fruitful conclusion and inferences it needs the special tools called data mining tools. This paper reviews the prediction on some real data, analyze and conclude that manufacturing of medicines done in pharmaceutical industry should according to constraint in market.

II. Data Mining in Pharmaceutical Industry
The pharmaceutical industry uses data mining techniques in a variety of ways, including profiling, classification, clustering, prediction and data analysis. However, due to the limited scope of this comment, only prediction and data analysis will be discussed. Prediction in the pharmaceutical context partly involves using past prescribing history of doctors to better predict future behavior. Data analysis in the pharmaceutical context uses the inferences drawn from the data mining research to determine the optimal course of action regarding future business decisions. At the outset from the data warehouse some previous years of data will be taken. Now a prediction can be done for determining the sales of product with respect to their particular quarter (period). With this a prediction will be achieved that a particular product sold maximum in that quarter.

III. Manufacturing Trends in Pharmaceutical Industry
Generally manufacturing unit is the one which work for all the time and engage in manufacturing of all types of products of an Industry, whether the product is in requirement or not for next many months the manufacturing unit keeps processing and the medicine is prepared and packed. There are many diseases which are seasonal and generally effect for some period of time only. So medicines required for these types of diseases are also for some short period of time not for whole year. If these types of medicines will be manufacture for a particular time of time there will be many advantages.

Using Data mining technique constraints can reach efficient manufacturing goals:-

- To maximize capacity utilization
- To minimize Expiry Time problem
- To minimize flow time of manufacturing
- To minimize of production costs
- To minimize unusual Labor cost

IV. What is Data Mining?
Data mining is a new term for what many analysts working in the field of sales and marketing analytics have been doing for years. It covers a broad range of techniques and methodologies and incorporates many tools to discover patterns and relationships in data. Data mining is usually associated with large amounts of data that is often housed in a data warehouse. There are five basic data mining tasks: profiling, clustering, classification, prediction, decision analysis.

V. What Data Mining Can Do?
Data mining can allow you to know your data. It can help you make better decisions and design more effective programs with improved targeting. And finally, it can uncover problems in your data.

VI. What Data Mining Can't Do?
Data mining cannot define your marketing objective or explain why an outcome occurs. It cannot correct problems in your data. And it cannot promise perfect results.

A sample data is collected for some seasonal diseases and respective medicines. These products will be differentiated between High Requirement Months and Low Requirement Months on the basis of requirement of product.

Graph of High Requirement Months

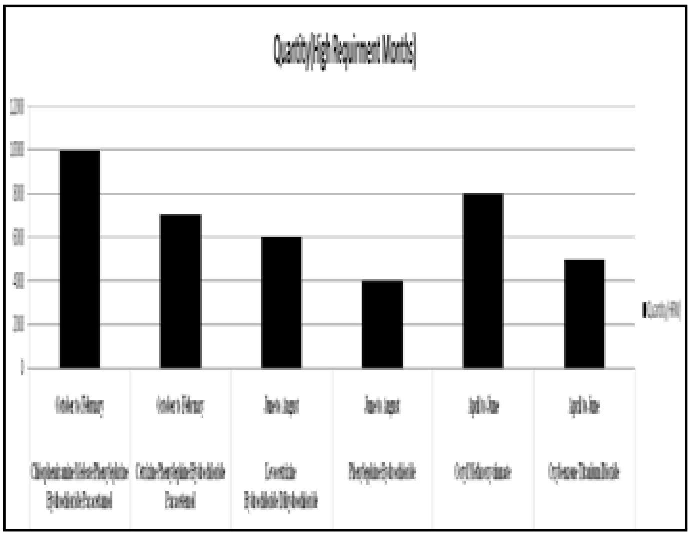

Fig. 1: High Requirement Analysis

Graph of Low Requirement Months

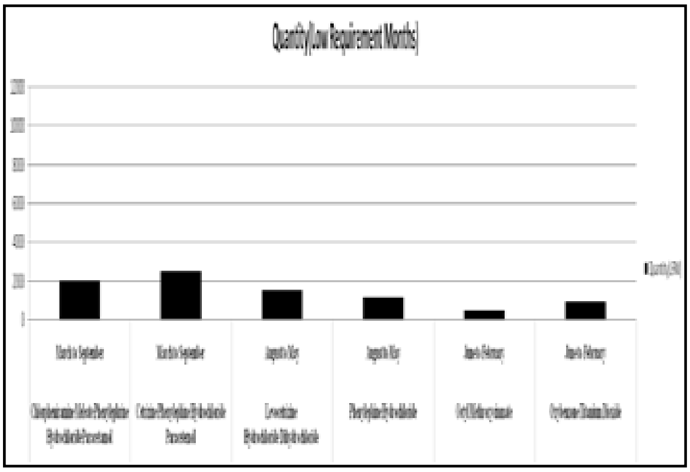

Fig. 2: Low Requirement Month Analysis

This graph differentiates the Sale of product in a year on the basis of High Requirement Month and Low Requirement. So for prediction large amount of data i.e. Data Warehouse will be taken and find out the High requirement Months and Low Requirement Months using Data Mining tool. If this analysis will be done that which product will sell more in particular months than the manufacturing can be done accordingly leads to major benefits. Now some real data is taken of salts and

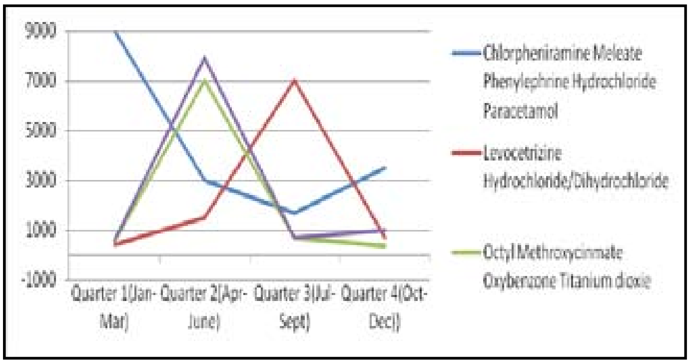

Fig. 3: Analysis Demand in an Year

VII. Applications of Data Mining in the Pharmaceutical Industry

Most healthcare institutions lack the appropriate information systems to produce reliable reports with respect to other information than purely financial and volume related statements (Prins & Stegwee, 2000). The management of pharmaceutical industry starts to recognize the relevance of the definition of drugs and products in relation to management information. In the uncertainty between costs, care-results and patient satisfaction the right balance is needed and can be found in upcoming information and Communication technology.

The delivery of healthcare has always been information intensive, and there are signs that the industry is recognizing the increasing importance of information processing in the new managed care environment (Morrisey, 1995). Most automated systems are used as a tool for daily work: they are focused on ‘production’ (daily registration). All the data, which are used to keep the organization running, operational data, are in these automated systems. These systems are also called legacy systems. There is a growing need to do more with the data of an organization than to use them for administration only.

Given the size of the databases being queried, there is likely to be a trade-off in accuracy of information and processing time. Sampling techniques and tests of significance may be satisfactory to identify some of the more common relationships; however, uncommon relationships may require substantial search time. Of course, the real data mining challenge comes when the user supplies only a minimal amount of information. Then, based upon the information in the databases and the relevant data entered by the user, a list of warnings or known reactions (accompanied by probabilities) should be reported.

Note that user profiles can contain large amounts of information, and efficient and effective data mining tools need to be developed to probe the databases for relevant information. Secondly, the patient’s (anonymous) profile should be recorded along with any adverse reactions reported by the patient, so that future correlations can be reported. Over time, the databases will become much larger, and interaction data for existing medicines will become more complete. The amount of existing pharmaceutical information (pharmacological properties, dosages, contraindications, warnings, etc.) is enormous; however, this fact reflects the number of medicines on the market, rather than an abundance of detailed information about each product.

One of the major problems with pharmaceutical data is actually a lack of information. For example, a food and drug administration department estimated that only about 1% of serious events are reported to the food and drug administration department. Fear of legal action may be a contributing factor; however, most health care providers simply don’t have the time to fill out reports of possible adverse drug reactions.

Furthermore, it is expensive and time consuming for pharmaceutical companies to perform a thorough job of data collection, especially when most of the information is not required by law. Finally, one should note that the food and drug administration department does not require manufacturers to test new medicines for potential interactions. There are in general three stages of drug development namely in adding of new drugs, development tests and predicts drug behavior, clinical trials test the drug in humans and commercialization takes drug and sells it to likely consumers (doctors and patients).

VIII. Benefits of Data Mining at Various Stages

Research Stage – instead of trial and error, data mining can help find drugs that have desirable activity

Development Stage – data mining can help predict who will benefit from drug

Clinical Trials Stage – data mining protects patients and helps regulate drug testing

Commercialization Stage – data mining can optimize use of sales resources like manpower, advertising

Drug-safety programs in both large and small pharmaceutical companies are now incorporating data mining processes for spontaneous reports on safety issues, as a supplement to traditional case-by-case medical review. While data mining is not a regulatory requirement, it can contribute valuable information. It is viewed favorably by regulators as a component of a proactive approach to safety.

IX. Approach - Maximal Frequent Sets

Since the collection of all frequent sets is downward closed, it can be rep-resented by its maximal elements, the so called maximal frequent sets. Most algorithms that have been proposed to find the maximal frequent sets rely on the same general structure as

ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

the Apriori and Eclat algorithm. The main additions' are the use of several look ahead techniques and efficient subset checking. The Max-Miner algorithm, proposed by Bayardo (1998), is an adapted version of the Apriori algorithm to which two look ahead techniques are added. Initially, all candidate k + 1-sets are partitioned such that all sets sharing the same k-prefix are in a single part. Hence, in one such part, corresponding to a prefix set X, each candidate set adds exactly one item to X. Denote this set of 'added' items by I. When a superset of X [I is already known to be frequent, this part of candidate sets can already be removed, since they can never belong to the maximal frequent sets anymore, and hence, also their supports don't need to be counted anymore. This subset checking procedure is done using a similar hash-tree as is used to store all frequent and candidate sets in Apriori.

First, during the support counting procedure, for each part, not only the support of all candidate sets is counted, but also the support of X [I. If it turns out that this set it frequent, again none of its subsets need to be generated anymore, since they can never belong to the maximal frequent sets. All other

k+1-sets that turn out to be frequent are added to the collection of maximal sets unless a superset is already known to be frequent, and all subsets are removed from the collection, since, obviously, they are not maximal. A second technique is the so called support lower bounding technique. That is, after counting the support of every candidate set X [f if , it is possible to compute a lower bound on the support its supersets using the following inequality:

Support(X [J) ¸ support(X) ¡ Xi2 J support(X) ¡ support(X):

For every part with prefix set X, this bound is computed starting with J containing the most frequent item, after which items are added in frequency de-creasing order as long as the total sum remains above the minimum support threshold. Finally, X [J is added to the maximal frequent sets and all its subsets are removed. Obviously, these techniques result in additional pruning power on top of the Apriori algorithm, when only maximal frequent sets are needed. Later, several other algorithms used similar look ahead techniques on top of depth-first algorithms such as Eclat. Among them, the most popular are GenMax (Gouda and Zaki, 2001) and MAFIA (Burdick et al., 2001), which also use more advanced techniques to check whether a superset of a candidate set was already found to be frequent. Also the FP-tree approach has shown to be effective for maximal frequent set mining (G. Grahne, 2003; Liu et al., 2003).

A completely different approach, called Dualize and Advance, was proposed by Gunopulos et al. (2003). Here, a randomized algorithm finds a few maximal frequent sets by simply adding items to a frequent set until no extension is possible anymore. Then, all other maximal frequent sets can be found similarly by adding items to sets which are so called minimal hyper graph transversals of the complements of all already found maximal frequent sets. Although the algorithm has been theoretically shown to be better than all other proposed algorithms, until now, extensive experiments have only shown otherwise

X. Conclusion

In the era of Information Technology, using technology we can achieve great success in every domain. In this paper I had tried to predict the consumption of goods on the basis of Months/Quarters. Using this prediction a decision can be taken that says what product should manufacture in any particular month/quarter. This leads to bring profits in company and reduce the problem of expiry in medicine types of product.

References

[1] Jayanthi Ranjan,"Applications of Data Mining Techniques", In 2005 - 2007 JATIT Pharmaceutical Industry

[2] Mr. S. P. Deshpande, Dr. V. M. Thakare,"Data Mining System and Applications: A Review", International Journal of Distributed and Parallel systems (IJDPS) Vol. 1, No. 1, September 2010.

[3] J.scott Armstrong, Kestem C.Green,"Demand Forecasting: evidence-Based Methods".

[4] Dr Robert D. Small, Herbert A.Edelstein,"Data Mining In The Pharmaceutical Industry", Informatics.

[5] J.A.Harding, M.Shahbaz, Shrinivas, A.Kusiak,"DATA Mining in Manufacturing: A Review", In 2006 Journal. of Manufacturing Science and Engineering.

[6] A.Kusiak,"Data Mining:manufacturing and Service Application", International Journal of Production Research Vol. 44, No's 18-19, 2006

[7] John J.Cohen. Zeneca Pharmaceuticals, C. Olivia Parr Rud, Data Square,"Data Mining of Market Knowledge in the Pharmaceutical Industry".

[8] Pavel Vazan, Pavol Tanuska, Michal Kebisek,"The Data Mining Usage in Production System Management", World Academy of Science, Engineering and Technology 2011.