

Data mining in medicines production system – tablet as a product

Miral Kothari ¹, Priti Sadaria ²

¹ Department of MCA, AITS, Yogidham Gurukul, Rajkot, Gujarat, India

² Department of Computer Science, Virani Science College, Yogidham Gurukul, Rajkot, Gujarat, India

¹ Gujarat Technological University

² Saurashtra University

¹ miralkothari@gmail.com

Abstract- As the information technology and internet grows, industries are able to store more and more amount of data. With the use of data mining, it is easy to analyze or process the data or to generate knowledge based on the data. Pharmaceutical industry is a huge industry which manufactures drugs in various formats such as tablets, capsules, and ointment or injection materials. Production of drugs or medicines includes a huge deal with various kinds of data. This paper outlines the application of data mining in production of medicines.

Keywords – Data mining, Tablet – medicine manufacturing process, knowledge discovery

I. INTRODUCTION

Industries always seek for the innovation in their business. Out of all the innovative concepts that industry has thought about, some are useful in product design and manufacturing of the product. Talking about the designing and manufacturing phase, the data conventionally flows as illustrated in Figure 1.

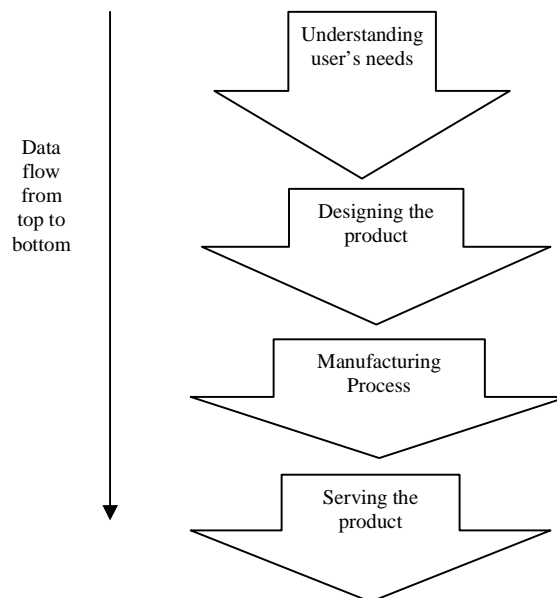


Figure 1. Flow of the data

Data mining in medicines production system – tablet as a product

The process of production may become imperfect by use of any narrowed flow of information which is not as per the main flow of information. The local redirection such as changes in pre-decided design or errors generated while manufacturing process is being carried out. Data mining techniques contributes to generate knowledge from the data and knowledge is applied again on the first phase i.e. understanding user’s need and trends (Kusiak & Smith, 2007)[1]. This cycle is illustrated in Figure 2.

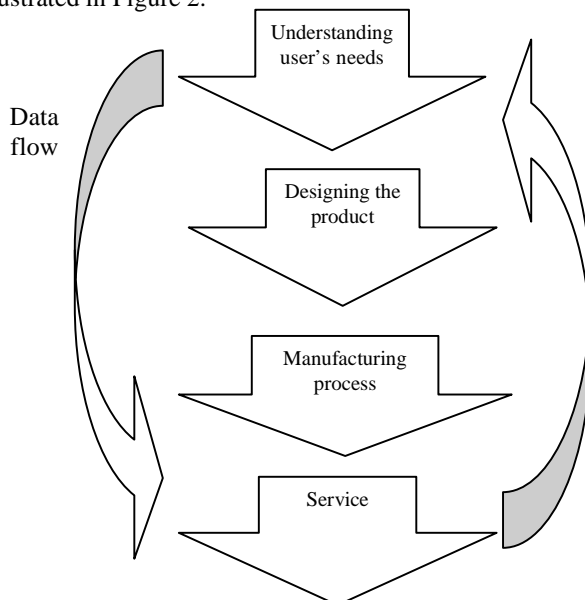


Figure 2. Cycle of data flow and knowledge

A large volume of data is generated during the lifecycle of the product. The data comes from every stage of the product life cycle such as market analysis (data collected from doctor, patients consuming the medicines) , making blueprint phase, manufacturing process and providing product in the hand of consumers. After the generation of data, knowledge is extracted from the filtered data called information. The key factor is knowledge produced from the data with the help of data mining. The Figure 3 illustrates how the data grows to knowledge.

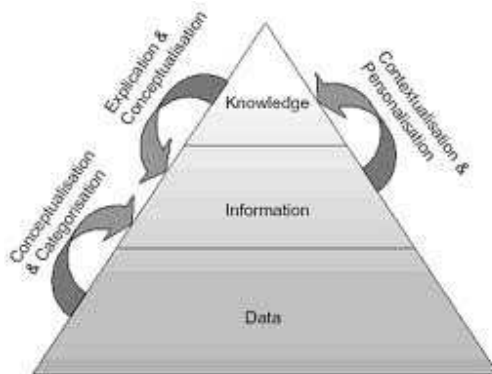


Figure 3. The way data becomes knowledge

The improvement in the design of the next generation product or in the manufacturing process is lead by or is based on the produced knowledge. Definitely the achievement of business lies in discovering the correct and relevant knowledge. Information from which the knowledge is to be produced must be timely and up to date.

A. Understanding user’s demand-

If the collected data remains only data and unprocessed, many facts are hidden behind it. But with the application of data mining, it is possible to discover some facts related to customer's approach towards the product which is not directly represented by the data.

For e.g. Data mining applied to doctor and patient domain may provide the answers to few questions such as

- Which qualities of our product emphasize customers to purchase our product?
- How the new customers turn to our product?
- Why some customers never turn to our product?

Most of the database management systems, such as MS-Access and Oracle, come with query facilities providing answers to some of these higher-level questions. However, for in-depth analysis, data-mining algorithms are needed (Witten & Frank, 2005) [2].

B. Sensing Market Trends -

With the development of larger and larger industries, concept of data ware house also grows. All industries want to keep their data for future reference and study through storing the data in data warehouse. By mining the data from the data warehouse, companies can dig out trends and changing tastes of customers. Mining time series is an active research area (Kusiak & Song, 2006)[3]. Trends are important to be understood because it reflects changes in tastes of customers, technology and results of strategies set by the company.

C. Diminish data complexity -

Database stores data in form of rows and columns, where rows represent number of records which represents the data and columns represent categories under which data is stored. Size of data is correlated with size of the company. The larger company may store huge amount of data. To avoid ambiguity and mess irrelevant data can be removed from the database. This ultimately results in easy data – mining.

D. Increase understandability -

Human being can understand the data easily if it is presented in the form of charts, graphs or arranged table. Moreover, relationship can be also easily understood if more than one kind of data is presented in a single pane. Knowledge about the data is needed to understand the data fully.

II. KNOWLEDGE DISCOVERY

Knowledge Discovery is the process of digging out the knowledge from the available large amount of data. Data collected by the industry can be useful but knowledge produced based on that data is not only useful but important as well as relevant and filtered. Data mining is used to mine the data and put it to the level which can provide some knowledge. There are two main classification of data mining. i. Descriptive and ii. Predictive. Descriptive data mining helps in describing the matter what data represents while the predictive data mining helps in predicting (expecting results) results. Along with the depth and richness of the data processing, data mining also supports the needs of an individual object of the database. In all the areas like engineering, medical, pharmaceutical and business applications, data mining is proved successful technique for knowledge discovery. (Da Cunha, Agard, & Kusiak, 2006 [4]; Harding, Shahbaz, Srinivas, & Kusiak, 2006[5]; Kusiak, 2006[6]).

The area of designing the product and manufacturing process are fully appropriate for applying the data mining as it contains varieties of data in very large amount. As data mining helps in generating knowledge from the data, it not only provides conclusive data but also helps in decision making by providing light on the business risks. Following are the few data-mining algorithms that are commonly used (Witten & Frank, 2005 [2]):

- Decision-tree algorithms.
- Decision-rule algorithms.
- Clustering.
- Regression.

Data mining in medicines production system – tablet as a product

Data mining can be useful not only in providing overview of the matter involved in the data but also for prediction based on the data or create any precise model from the data. For e.g. Data mining can be used to: (Kusiak & Smith, 2007)[1]

- Describe the data
- Categorize the data
- Conceptualize the data
- Relationships between the data

Data mining only cannot be used for all kinds of deals with data. It needs some supportive other tools and technologies to work together and provide exact results. Technologies or techniques such as data warehousing useful for providing data to data mining, visualization techniques helpful for presenting the data in understandable format, computer network for distributing the data. Such techniques are to be integrated along with the data mining to have the efficient effect of the data mining.

Next Section of this paper discusses application of data mining in manufacturing process of the product.

III. PRODUCTION PROCESS OF MEDICINES - TABLET

Medicines of various kinds, before coming into patients or doctors hand, passes through many stages. Figure 4 illustrates the process of how a material becomes edible tablet or medicine.[7]

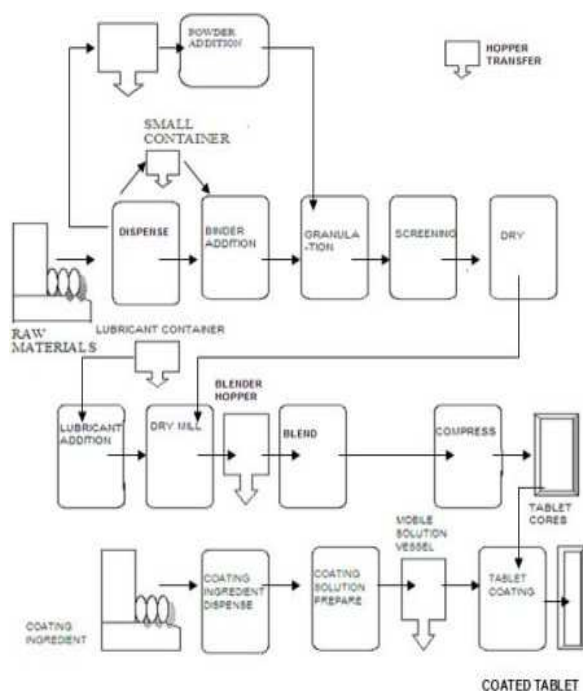


Figure 4. Manufacturing Process of tablet - medicine

During the process, the physical characteristics of starting initial material is changed number of times before being formed as a product. The active ingredients and excipients are mixed to form the granules. Moreover, many other

unit processes are involved in making tablets, including particle size reduction, sizing of the tablets, blending, drying the material up to required temperature, compaction and coating.

The sequence of unit operations [7] is illustrated in Figure 5.

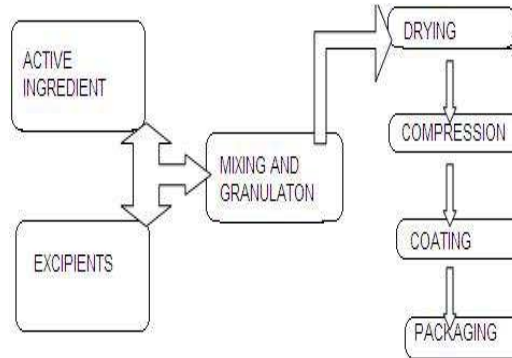


Figure 5. Sequence of operations in tablet mfg.

Various factors related with tablet can affect content uniformity, bioavailability or stability.

A. Dispensing -

The first step of the tablet manufacturing process is dispensing. Dispensing is the stage where the weight of RM (raw material) and other ingredients is determined according to the dosage and batch required to be prepared. Dispensing can be done manually or with the help of weigh scale tool. Some possible issues like weighing accuracy, dust control during manual handling and quantity control of each ingredient may arise. Referring to the earlier data stored in the company, the above possible issues may be solved or controlled.

B. Sizing -

During manufacturing of tablets, it is easier to blend the material if the all the ingredients are of the same size. It improves the uniformity to the tablet. The sizing is the stage which does size reduction, milling, crushing or grinding to the mixture and makes almost fine particle size. Sizing promotes controlled particle size distribution of dry granulation which is useful to promote better flow of mixture in tablet machine.

C. Power blending -

Power blending is the most difficult unit operation because it cent per cent accuracy is practically unachievable. The blending is the stage of pre granulation of tablet manufacturing. The mixing process for every product has optimum mixing time and prolonged mixing may result in a spoiled or unformed product. Along with mixing time, mixing speed is most affecting factor to the product. The “V” blender, Tumbling blender, Agitated powder blender, Container blender etc. are the widely used types of blender.

D. Granulation -

The next stage after blending is the granulation. The prepared mixture of material is granulated to provide uniformity (homogeneity).

E. Drying -

Drying is the most important phase in the tablet manufacturing. In this phase, dryer is used to dry the material and to remove unwanted moisture from the product. The care has to be taken in keeping the low enough moisture to ensure the free flowing. The widely used dryers include Vacuum tray dryer, Microwave dryer, Spray dryer.

F. Tablet Compression -

Data mining in medicines production system – tablet as a product

As the granules are prepared, they are compressed to be converted into final product. A high speed mechanical device is used to compress the tablet which squeezes the granules with extreme precision in particular shape. The punches and dies are fixed for the desired shape. Figure 6 illustrates the common stages of compression. The ingredients are pressed forcefully and they turn into tablet.

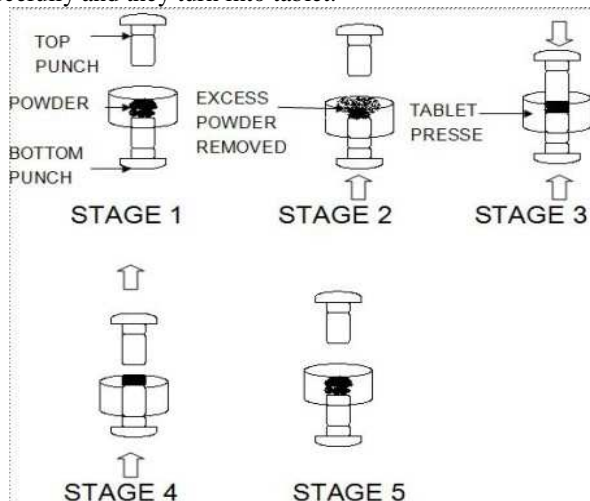


Figure 6. Common stages of compression

G. Supplementary process -

Along with the above mentioned stages, tablet manufacturing process includes some supplementary processes which turn it to the fine product such as tablet weight monitoring device (monitors force at each compression), tablet deduster (removes the extra powder which comes with every tablet coming out of compression), Fette machine (chills the compression device).

H. Coating and packing -

As soon as the tablet is ready, coating is done to give the finishing touch to the tablet or to cover the bitter taste of the ingredients. It is not necessary that all kind of tablets are being coated. When coating is done, product is ready for the packing. The packing can be in aluminum foil or blister strips. Blister packs are common and are being used for wide variety of product.

These above mentioned six stages convert a mixture of raw material and ingredients in to a medicine. Data mining is useful as and when the reference of existing data is needed for any decision making.

IV. DATA INVOLVED IN PRODUCTION

The data mining techniques can be applied when some data is already in hand. Data mining methods refer to the tools provided by the data mining. Figure 7 illustrates how the figures of AM (active material) and Excipients fall or increase as the medical weight of the tablet to be produced is to be changed.

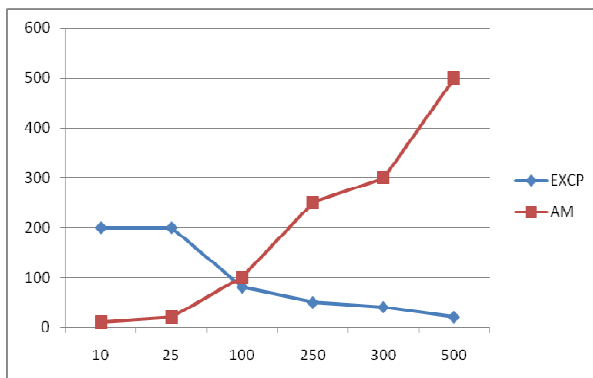


Figure 7. Increase and fall of AM and Excipients in context to mg.

Data mining categorization according to its function can be classification, regression, clustering, summarization, link analysis (Fayyad, Piatetsky – Shapiro and Smyth, 1996[8], Fayyad, Piatetsky – Shapiro, Smyth and Uthurusamy, 1996[9], Berry and Linoff, 1997[10]). The data involved in production process of tablet is the medical weight of the tablet, quantity of active material, quantity of excipients (relevant with active material) and physical size of the tablet. Each of the data may change in context to batch to be manufactured or medicine to be produced. The two different tablets of same medical weight may not contain same quantity of active material and excipients. The Table 1 illustrates the comparative study of increase or fall of ingredients and size as the milligram (medical weight of the tablet) varies. The data mining techniques can be applied at the stage of sizing particularly. If any problem occurs during sizing or blending, the data is to be reviewed again using data mining. The same loop can be applied in the stage of tablet compression. If any problem arises during compression of the tablet, the data is to be mined again from the previous stages. The previous data reflects the low enough moisture to be kept in the material or size of the granules which directly affects the quality of the tablets.

Tablet weight (in mgs.)	Active material (app. In mgs.)	Excipients (app. In mgs.)	Size (in mm)
10	10	200%	7.14
25-50	25-50	200%	7.94
100	100	80%	8.73
250	250	50%	9.52
300	300	40%	11.11
500	500	20%	12.7

Table 1. Showing comparative study between AM, Excipients, Size in context to Tablet medical weight

V. CONCLUSION

The data mining procedure emphasized the set of data by repeating interaction activities. The data set to be changed continuously as the changes occurs in market and generations of the medicines. (Chih-Hung Hsu, 2009)[11]Up to date industrial standards are to be maintained to remain in the market and to reach up to the satisfaction mark of doctors and customers. But such business decision can be made only by data mining. By applying proper data mining techniques, the possible problems during manufacturing process can be solved and ultimately it results into accurate and consistent production of medicines.

VI. REFERENCE

- [1]Kusiak, A., & Smith, M. (2007) Data mining in design of products and production systems. Elsevier
- [2]Witten, I. H., & Frank, E. (2005). Data mining: Practical machine learning tools and techniques. New York: Elsevier.
- [3]Kusiak, A., & Song, Z. (2006). Combustion efficiency optimization and virtual testing: A data-mining approach. IEEE Transactions on Industrial Informatics,2(3), 176–184.

Data mining in medicines production system – tablet as a product

- [4] Da Cunha, C., Agard, B., & Kusiak, A. (2006). Data mining for improvement of product quality. *International Journal of Production Research*, 44(18/19), 4027–4041.
- [5] Harding, J. A., Shahbaz, M., Srinivas, S., & Kusiak, A. (2006). Data mining in manufacturing: A review. *ASME Transactions: Journal of Manufacturing Science and Engineering*, 128(4), 969–976.
- [6] Kusiak, A. (2006). Data Mining: Manufacturing and service applications. *International Journal of Production Research*, 44(18/19), 4175–4191.
- [7] www.pharmainfo.net
- [8] Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34
- [9] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*. Cambridge AAI/MIT Press.
- [10] Berry, M. J. A., & Linoff, G. (1997). *Data mining techniques for marketing, sales, and customer support*. John Wiley & Sons Press.
- [11] Chih-Hung Hsu (2009), *Data mining to improve industrial standards and enhance production and marketing: An empirical study in apparel industry*, Elsevier.