PRAGMATIC APPLICATION OF DATA MINING IN HEALTHCARE

[1]Saravanan N, [2]Malini C.S
[1]Assistant Professor, [2]II MCA Student
[1 2]Department of Computer Applications,
[1 2]Priyadarshini Engineering College, Vaniyambadi, Vellore, Tamilnadu, India

Abstract: Data mining holds a lot of potential in the healthcare industries. It can enable the health systems in using the data and the analytics for identification of the best possible practices along with inefficiencies. It reduces the cost and improves the overall care structure. Data mining can help in predicting the timeline of future events. It can easily help in predicting the future diseases or deficiencies that can occur in the patient's lifetime. Data mining provides the methodology and technology to transform these mounds of data into useful information for decision making. This article explores data mining applications in healthcare. It also gives an illustrative example of a healthcare data mining application involving the identification of risk factors associated with the onset of diabetes. Finally, the article highlights the limitations of data mining and discusses some future directions.

IndexTerms **- Data mining, Healthcare, data analytics, life science**

I. INTRODUCTION

Data mining holds great potential for the healthcare industry to enable health systems to systematically use data and analytics to identify inefficiencies and best practices that improve care and reduce costs. Some experts believe the opportunities to improve care and reduce costs concurrently could apply to as much as 30% of overall healthcare spending. This could be a win/win overall [1]. But due to the complexity of healthcare and a slower rate of technology adoption, our industry lags behind these others in implementing effective data mining and analytic strategies.

Like analytics and business intelligence, the term data mining can mean different things to different people. The most basic definition of data mining is the analysis of large data sets to discover patterns and use those patterns to forecast or predict the likelihood of future events [2]. All analyses is not of large quantities of data constitute data mining. We generally categorize analytics as follows:

- Descriptive analytics—Describing what has happened
- Predictive analytics—Predicting what will happen
- Prescriptive analytics—Determining what to do about it

It is to the middle category, predictive analytics that data mining applies. Data mining involves uncovering patterns from vast data stores and using that information to build predictive models. Many industries successfully use data mining. It helps the retail industry model customer response [3]. It helps banks predict customer profitability. It serves similar use cases in telecom, manufacturing, the automotive industry, higher education, life sciences, and more. However, data mining in healthcare today remains, for the most part, an academic exercise with only a few pragmatic success stories. Academicians are using data-mining approaches like decision trees, clusters, neural networks, and time series to publish research. Healthcare, however, has always been slow to incorporate the latest research into everyday practice [4]. The leading warehouse practitioners are asking themselves to narrow the adoption time from the bench (research) to the bedside (pragmatic quality improvement) and affect outcomes.

The implement the analytics foundation to mine the data and they have the best practices and organizational systems in place to make data mining insights actionable, they are now ready to use predictive analytics in new and innovative ways. a health system trying to succeed in risk-based contracts while still performing well under the fee-for-service reimbursement model. The transition to value-based purchasing is a slow one [5]. Until the flip is switched all the way, health systems have to design processes that enable them to straddle both models. This client is using data mining to lower its census for patients under risk contracts, while at the same time keeping its patient volume steady for patients not included in these contracts. We are mining the data to predict what the volumes will be for each category of patient. Then, the health system develops processes to make sure these patients receive the appropriate care at the right place and at the right time. This would include care management outreach for high-risk patients [6,7].

There is an unstoppable growth in the amount of electronic health records or EHRs being collected by healthcare facilities. It has been the norm for nurses to take responsibility in handling patient data input that was traditionally recorded in paper-based forms. Accuracy is extremely important when it comes to patient care and computerizing this massive amount of data enhances the quality of the whole system [8]. Data mining has proven to be extremely effective and had been used to uncover patterns from the large amount of stored information and then used to build predictive models. Since the early 90s, this practice has been used to help with fraud detection, credit scoring and maintenance scheduling but it's finally being utilized in healthcare programs around the country. Improving the quality of patient care and reducing healthcare costs are the ideal goals of many programs. Data mining has helped these programs succeed.

II. THREE SYSTEMS APPROACH

The most effective strategy for taking data mining beyond the realm of academic research is the three systems approach. Implementing all three systems is the key to driving real-world improvement with any analytics initiative in healthcare. Unfortunately, very few healthcare organizations implement all three of these systems [9]. The three systems are:

1. **The analytics system.** This system includes the technology and the expertise to gather data, make sense of it and standardize measurements. Aggregating clinical, financial, patient satisfaction, and other data into an enterprise data warehouse (EDW) is the foundational piece of this system.
2. **The best practice system.** The best practice system involves standardizing knowledge work—systematically applying evidence-based best practices to care delivery. Researchers make significant findings each year about clinical best practices, but, as I mentioned previously, it takes years for these findings to be incorporated into clinical practice. A strong best practice system enables organizations to put the latest medical evidence into practice quickly.
3. **The adoption system.** This system involves driving change management through new organizational structures. In particular, it involves implementing team structures that will enable consistent, enterprise-wide adoption of best practices. This system is by no means easy to implement. It requires real organizational change to drive adoption of best practices throughout an organization.

Implementing all three enables a healthcare organization to pragmatically apply data mining to everyday clinical practice.

III. DATA MINING AND ITS EFFECT ON PRIVACY

Data mining possesses great potential for the healthcare industry, but it also comes with a few privacy concerns. Massive amounts of patient data being shared during the data mining process may leave some patients worried that their personal information could fall into the wrong hands. However, experts argue that this is a risk worth taking. "There will be criminals. There will be people who are bad actors. At some point, something is going to get out," Thomas Graf, chief medical officer at Geisinger Health System told The Washington Post in a 2014 article. "It's not an irrational fear. At the same time, people die driving every year and we still choose to drive cars, or most of us do. It's a risk every person has to decide where they fall on the line."

3.1 Future of Data Mining

The shift from written to electronic health records has played a huge part in the push to use patient data to improve areas of the healthcare industry. The adoption of electronic health records have allowed healthcare professionals to distribute the knowledge across all sectors of healthcare, which in turn, helps reduce medical errors, provide thorough documentation and improve patient care and satisfaction. Data mining is also projected to help cut costs [10]. The future of healthcare may well depend on using data mining to decrease healthcare costs, identify treatment plans and best practices, measure effectiveness, detect fraudulent insurance and medical claims, and ultimately, improve the standard of patient care. Data mining is used successfully and extensively in healthcare today. For example, I was part of a project that mined healthcare claims to determine best providers and procedures for conditions, diagnostic aids for certain procedures and protein analysis for drug development.

3.2 Advantages of Data Mining Tools

Data mining tools that are interactive, visual, understandable, well-performing and work directly on the data warehouse/mart of the organization could be used by front line workers for immediate and lasting business benefit. There are numerous, accessible data mining techniques that are more effective than most simply because they will be used by so many within an organization. With little investment, they can draw attention to significant anomalies that deserve further investigation. Data mining tools help customers analyze data by executing a series of actions and returning results that provide visibility into behaviors surrounding the dimensions of the company's business. SQL Server 2005, for example, provides seven "out of the box" algorithms that can assist a company in obtaining insight into their business [11]. Each algorithm works differently to produce an output of results. In all cases, the algorithms are "trained" by exposing them to the customers' existing data sets. The training set might include sets such as order history, payable/receivables, web navigation logs, or customer demographic information.

3.3 Disadvantages of Data Mining Tools

The techniques deployed by some tools are generally well beyond the understanding of the average business analyst or knowledge worker. This is because the tool was generally designed for expert statisticians involved in the detailed science of predictive modeling. This would be the disadvantage of data mining today [12]. If this advanced level of analysis is reserved for the few, instead of for the masses, the full value of data mining in the organization cannot be realized.

IV. PROVEN APPLICATIONS OF DATA MINIING

Data mining has been used in many industries to improve customer experience and satisfaction, and increase product safety and usability [13]. In healthcare, data mining has proven effective in areas such as predictive medicine, customer relationship management, detection of fraud and abuse, management of healthcare and measuring the effectiveness of certain treatments.

Here is a short breakdown of two of these applications:

- **Measuring Treatment Effectiveness** – This application of data mining involves comparing and contrasting symptoms, causes and courses of treatment to find the most effective course of action for a certain illness or condition. For example, patient groups who are treated with different drug regimens can be compared to determine which treatment plans work best and save the most money. Furthermore, the continued use of this application could help standardize a method of treatment for specific diseases, thus making the diagnosis and treatment process quicker and easier.

- **Detecting Fraud and Abuse** – This involves establishing normal patterns, then identifying unusual patterns of medical claims by clinics, physicians, labs, or others. This application can also be used to identify inappropriate referrals or prescriptions and insurance fraud and fraudulent medical claims.

V. BENEFITS OF DATA MINING IN HEALTHCARE INDUSTRY

The solutions might favor healthcare providers or insurance companies, data mining benefits everyone concerned, from healthcare organizations to insurers to patients [14, 15].

- Patients receive more affordable and better healthcare services. This happens when healthcare officials use data mining programs to identify and observe high-risk patients and chronic diseases and design the right interventions needed. These programs also reduce the number of claims and hospital admissions, further streamlining the process.
- Healthcare providers use data mining and data analysis to find best practices and the most effective treatments. These tools compare symptoms, causes, treatments and negative effects and then proceed to analyze which action will prove most effective for a group of patients. This is also a way for providers to develop the best standards of care and clinical best practices.
- Insurers are now able to better detect medical insurance abuse and fraud because of data mining. Unusual claims patterns are easier to spot with this tool and it can identify inappropriate referrals and fraudulent medical and insurance claims. When insurers reduce their losses due to fraud, the cost of health care also decreases.
- Healthcare facilities and groups use data mining tools to reach better patient-related decisions. Patient satisfaction is improved because data mining provides information that will help staff with patient interactions by recognizing usage patterns, current and future needs, and patient preferences.

Electronic health records (EHR) are quickly becoming more common among healthcare facilities. With increased access to a large amount of patient data, healthcare providers can now optimize the efficiency and quality of their organizations using data mining.

VI. CONCLUSION

In healthcare, data mining has proven effective in areas such as predictive medicine, customer relationship management, detection of fraud and abuse, management of healthcare and measuring the effectiveness of certain treatments. The purpose of data mining, whether it's being used in healthcare or business, is to identify useful and understandable patterns by analyzing large sets of data. These data patterns help predict industry or information trends, and then determine what to do about them. In the healthcare industry specifically, data mining can be used to decrease costs by increasing efficiencies, improve patient quality of life, and perhaps most importantly, save the lives of more patients.

REFERENCES

[1] Bellazzi, R., Zupan, B.: Predictive data mining in clinical medicine: Current issues and guidelines. International Journal of Medical Informatics 77, 81–97 (2008).

[2] Canlas Jr., R.D.: Data Mining in Healthcare: Current Applications and Issues (2009).

[3] Chen, H., Fuller, S., Friedman, C., Hersh, W.: Medical Informatics. Knowledge Management and Data Mining in Biomedicine. Springer Science (2005).

[4] Cios, K.J., Moore, G.W., Uniqueness of medical data mining. Artificial Intelligence in Medicine 26, 1–24 (2002).

[5] Kolce, E., Frasheri, N., A Literature Review of Data Mining Techniques Used in Healthcare Databases (2012).

[6] Healthcare Information and Management Systems Society. Electronic Health Records. A Global Perspective. White paper. HIMSS Enterprise Systems Steering Committee and the Global Enterprise Task Force (2010).

[7] Janarthanan Y, Balajee J.M, and Srinivasa Raghava S. "Content based video retrieval and analysis using image processing: A review." International Journal of Pharmacy and Technology 8, no.4 (2016): 5042-5048.

[8] Jeyakumar, Balajee, MA Saleem Durai, and Daphne Lopez. "Case Studies in Amalgamation of Deep Learning and Big Data." In HCI Challenges and Privacy Preservation in Big Data Security, pp. 159-174. IGI Global, 2018.

[9] Kamalakannan, S. "G., Balajee, J., Srinivasa Raghavan.,"Superior content-based video retrieval system according to query image"." International Journal of Applied Engineering Research 10, no. 3 (2015): 7951-7957.

[10] Kambatla, K., Kollias, G., Kumar, V. and Grama, A., Trends in big data analytics. *Journal of Parallel and Distributed Computing*, *74*(7), pp.2561-2573, 2014.

[11] Ranjith, D., J. Balajee, and C. Kumar. "In premises of cloud computing and models." International Journal of Pharmacy and Technology 8, no. 3 (2016): 4685-4695.

[12] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." *Neural networks* 61 (2015): 85-117.

[13] Ushapreethi P, Balajee Jeyakumar and BalaKrishnan P, Action Recongnition in Video Survillance Using Hipi and Map Reducing Model,International Journal of Mechanical Engineering and Technology 8(11), 2017,pp. 368–375.

[14] Ushapreethi, P. and Lakshmipriya, G.G."Survey on Video Big Data: Analysis Methods and Applications." International Journal of Applied Engineering Research,12(10), pp.2221-2231.2017.

[15] Sethumadhavi R, Balajee J "Big Data Deep Learning in Healthcare for Electronic Health Records," International Scientific Research Organization Journal, vol. 2, Issue 2, pp. 31–35, Jul. 2017.