International Journal of
Advances in Scientific Research and Engineering (ijasre)

E-ISSN : 2454-8006

DOI: 10.31695/IJASRE.2018.32881

Volume 4, Issue 9
September - 2018

Data Mining in Healthcare

Matthew N. O. Sadiku, Kelechi G. Eze, and Sarhan M. Musa

Roy G. Perry College of Engineering

Prairie View A&M University

Prairie View, TX 77446

U.S.A

ABSTRACT

The data mining has played a significant role in healthcare industry because of its descriptive and predictive power. It has been used in predicting various types of diseases.
The data mining helps in planning healthcare activities and reducing the number of inpatients in hospital. It can also be used for decision-making at different levels of the healthcare sector. This paper provides a brief introduction to data mining and its applications in healthcare industry.

Key words: Data mining, Healthcare.

1. INTRODUCTION

The need to extract useful knowledge hidden in large amount of data and to act on the knowledge is becoming increasingly important in all the various fields of technology, business, marketing, science, and healthcare. Healthcare organizations store large volumes of sensitive personal data, such as patient names, dates of birth, and personal medical records. They lack effective analysis tools to discover hidden relationships and trends. Their value is greatly undermined by our inability to translate them into useful knowledge and actions. Data mining is the process of taking the raw material from the data warehouse and converting it to information that can be used in decision making. It involves digging of useful information or knowledge from previous data [1].

The need to understand large data sets has increased in all the fields of technology, business, and science. There is a wealth of massive data available within the healthcare industry because healthcare involves collecting data at every encounter of each individual with the care provider, pharmacy, payer, or government agency. However, there is a lack of effective analysis tools to discover hidden relationships and trends in data.

Data mining (or knowledge discovery in databases) has been used in many industries to improve customer experience and satisfaction. It has been used to uncover patterns from the large amount of data and then used to build predictive models. Data mining technology can help explore, detect, and discover new knowledge from big data. In healthcare, data mining has been effective in areas such as effective treatment, healthcare management, customer relation management, predictive medicine, detection of fraud and abuse, decision making, improving patient quality of life, and saving the lives of more patients. Data mining can help researchers gain insights and can uncover new healthcare knowledge for clinical and administrative decision making [2].

2. DATA MINING BASICS

Data mining, as a relatively new concept, emerged in the middle of 1990s. It is the practice of extracting useful information from big data. It is the process of sorting through large data sets to identify patterns and establish relationships. Data mining tools allow one to discover patterns and to use those patterns to predict future trends. The more organized or structured the data is, the easier it is to analyze.

Data mining should be regarded as a process which involves the following steps [3]:
(1) Efficient data storage and data processing.

(2) Decide on the number of variables to be investigated.

(3) Data needs to be visualized and summarized.

(4) Apply statistics such as mean, percentiles, standard deviation, and correlation.

(5) Apply analysis methods such as regression, nearest neighbor methods, k-mean clustering, etc.

(6) Implement insights gained from the analysis.

Data mining can also be regarded as an assortment of algorithmic techniques for generating previously unknown from the raw data. Common data mining techniques used in healthcare include Artificial Neural Network, Decision trees, Genetic Algorithms, Fuzzy logic, Fuzzy based Neural Networks, Bayesian Networks, and Support Vector Machines. The data mining techniques are effective for the following reasons [4].

- It works by learning from data that is past history.
- It is easy to use and test for results as prediction is based on the past history.
- Data from various resources is managed and the required data alone is extracted for the technique being used.
- Models are easily updated by re-learning, past information, and change in trends.

Today, data mining has grown so vast that they can be used in several applications such as risk management, financial analysis, insurance, and healthcare.

3. HEALTHCARE DATA MINING

The healthcare industry generates massive amounts of data including patient Electronic Health Record (EHR) or Electronic Medical Records (EMR), drug development data, and patient survey data. Data mining in healthcare is used mainly for predicting various diseases and assisting advising doctors in making decisions. Using data mining, the healthcare industry can benefit in many areas such as medical research, pharmaceuticals, medical devices, hospital management, health care insurance, detection and prevention of fraud.

The healthcare field faces some pressure to reduce costs while increasing the quality of services. Data-mining techniques are becoming more widely used in the fields of diabetes, heart diseases, and heart attacks. Some applications of data mining in healthcare include the following [5]:

- Effective management of hospital resource
- Hospital infection control
- Hospital ranking
- Better customer relation
- Smarter treatment techniques
- Improved patient care
- Recognize high-risk patient
- Decrease insurance fraud

3.1 Benefits and Challenges

The benefits of data mining in the healthcare industry include [6]:

- Patients receive more affordable and better healthcare services.
- Healthcare providers use data mining and data analysis to find best practices.
- Insurance organization can now better detect medical insurance abuse and fraud.
- Healthcare provider can reach better patient-related decisions.

Data mining is getting popular in the healthcare industry because it offers benefits to all stakeholders: patients, care providers, healthcare organizations, researchers, and insurers. Data mining predictive models can be used as a second opinion for the doctor's decision on a treatment. It may help in grouping the patients having similar type of diseases or health issues.

Using data mining techniques in healthcare is very complex and is faced with some challenges. This is partly due to the idiosyncrasies of the medical profession. Privacy and security of patient data is a big challenge because of the sensitivity of healthcare data. Since the health data contains personal, sensitive information, there is the risk of invading privacy [7]. Other challenges include noise, high dimensionality, sparseness, fragmentation, and nonlinear relationships among data elements.

DOI: 10.31695/IJASRE.2018.32881

4. CONCLUSION

Data mining is a newly developed technology, which has its own methods, procedures and techniques. It is the task of discovering useful patterns from a large amount of data. In healthcare, data mining is becoming increasingly popular and a necessity. Its applications in healthcare can have tremendous potential and usefulness. It plays an important role for uncovering new trends in healthcare industry.

Although healthcare data mining is still in its infancy, the healthcare data mining literature is very rich. The future of healthcare will depend on using data mining to decrease healthcare costs and improve the standard of patient care.

REFERENCES

[1] U. Shafique et al., "Data mining in healthcare for heart diseases," *International Journal of Innovation and Applied Studies*, vol. 10, no. 4, March. 2015, pp. 1312-1322.

[2] I. Yoo et al., "Data mining in healthcare and biomedicine: A survey of the literature," *Journal of Medical System*, vol. 36, 2012, pp. 2431–244.

[3]. M. N. O. Sadiku, A. E. Shadare, and S. M. Musa, "Data mining: A brief introduction," *European Scientific Journal*, July 2015, vol. 11, no. 21, pp. 509-513.

[4] S. Rahaman and S. Bij, "Data mining facilitates e-patients through e-healthcare: An empirical study," *Proceedings of International Conference on New Trends in Information and Service Science,* 2009, pp. 1158-1165.

[5] D. Tomar and S. Agarwal, "A survey on data mining approaches for healthcare," *International Journal of Bio-Science and Bio-Technology*, vol.5, no.5, 2013, pp. 241-266.

[6]"How data mining is helping healthcare," https://the-modeling-agency.com/how-data-mining-is-helping-healthcare/

[7] M. H. Tekieh and B. Raahemi, "Importance of data mining in healthcare: A survey," *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2015, pp. 1057-1062.

AUTHORS

Matthew N.O. Sadiku, is a professor at Prairie View A&M University, Texas. He is the author of several books and papers. He is an IEEE fellow. His research interests include computational electromagnetics and computer networks.

Kelechi G. Eze, is a doctoral student at Prairie View A&M University, Texas. He is a student member of IEEE. His research interests include Internet of things security, data security and privacy, blockchain technology, wireless sensor networks, and machine learning.

Sarhan M. Musa, is a professor in the Department of Engineering Technology at Prairie View A&M University, Texas. He has been the director of Prairie View Networking Academy, Texas, since 2004. He is an LTD Sprint and Boeing Welliver Fellow.