IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

Data Analytics in Healthcare Industry

Dr. K. Sharmila[1] and Mrs. N. Shyamala Devi[2]
Associate Professor, Department of Computer Science[1]
Assistant Professor, Department of Computer Science[2]
Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, India
sharmila.scs@velsuniv.ac.in[1] and nshyamaladevi.scs@velsuniv.ac.in[2]

Abstract: *The Healthcare industry is one of the most challenging industries when it comes to data analytics and management. The massive influx of complex data in the healthcare industry makes it almost intimidating for organizations that don't have the right data analytics and management tools to handle it. This also results in the failure of such organizations to squeeze out meaningful insights from the data, which otherwise can be used to improve the quality of healthcare offered to patients, as well as, to increase the efficiency of the organization. There are many powerful analytical tools that are available in the market that can help us leverage the volumes of data available. The goal of this paper is to provide a comprehensive use of health data analytics which allows the modern world for the improvements of patient care, faster and more accurate diagnoses, preventive measures, more personalized treatment and more informed decision-making. At the business level, it can lower costs, simplify internal operations and more.*

Keywords: Healthcare, data analytics, decision making

I. INTRODUCTION

Data Science has been established as an important emergent scientific field and paradigm driving research evolution in such disciplines as statistics, computing science and intelligence science, and practical transformation in such domains as science, engineering, the public sector, business, social sci-ence, and lifestyle. The field encompasses the larger ar-eas of artificial intelligence, data analytics, machine learning, pattern recognition, natural language understanding, and big data manipulation. Data Science deals with the slicing and dicing of the big chunks of data, as well as finding insightful patterns and trends from them using technology, mathematics, and statistical techniques[5].

Data Analyticsis the science of examining raw data with the aim of finding hidden patterns and eliciting useful information. Some of the key areas were Business analytics can be used to benefit an organization are determining root cause for failures, finding buying patterns of customers, identifying fraudulent transactions, recalculations risk portfolios etc. Data Analytics can be classified it to

- Descriptive statistics – quantitatively describe features of the collected information
- EDA (Exploratory data analysis) – Find out new features through the data
- CDA (confirmatory data analysis) – verify a hypothesis[2].

The Healthcare industry is one of the most challenging industries when it comes to data analytics and management. The massive influx of complex data in the healthcare industry makes it almost intimidating for organizations that don't have the right data analytics and management tools to handle it[1].

he healthcare data may just seem like a jumble of numbers and words that is until this data is converted into meaningful information by using advanced data analytics tools. According to one survey, over 90% of healthcare organizations strongly agreed to the growing influence of data analytics on their organizations in the next few years. This shows that just like every other industry, data analytics is slowly but surely starting to influence the healthcare industry for the good, transforming not just the clinical and operational side of the practice but also the administrative and financial aspects of the healthcare industry.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

II. HEALTH CARE DATA

Health data is any data relating to the health of an individual patient or collective population. This information is gathered from a series of health information systems (HIS) and other technological tools utilized by health care professionals, insurance companies and government organizations. The health care industry is awash with valuable data in the form of detailed records. Industry regulations stipulate that health care providers must retain many of these records for a set period of time[6].

There are a variety of tools and systems used to collect, store, share and analyze health data gathered through various means. These tools include:

- Electronic Health Records (EHRs)
- Personal Health Records (PHRs)
- Electronic Prescription Services (E-prescribing)
- Patient Portals
- Master Patient Indexes (MPI)
- Health-Related Smart Phone Apps and more[3].

With digital data collection, there is more and more health care data to be analyzed every second. With the increase of electronic record keeping, applications and other electronic means of data collection and storage, there is a significant amount of data being collected in real time.

These data sets are so complex that traditional processing software and storage options cannot be used. Cloud storage is a necessity when dealing with "Big Data." Cloud storage is built to be secure, an absolute must when dealing with sensitive patient information. It is also very cost-efficient and has been helpful in lowering the increasing cost of health care.

III. DATA ANALYTICS TOOLS

Data Analysts are making use of this data on a regular basis and thus the advent of Data Analytics. To solve the purpose of converting or transforming raw data into valuable information and to work on data analysis projects, there exist multiple Data Analytics tools in today's world. The information could be vital in making data-driven business decisions. A data analyst's key responsibility is to gather insights from data and to do so; they must make use of the various data analytics tools available to them.

Data Analytics Tools

- R Programming
- Python
- Excel
- Tableau
- QlikView
- Power BI
- SAS
- KNIME
- Apache Spark[4]

3.1 R Programming Language

R is one of the most popular languages for statistical modelling, visualization, and data analysis. It is an open-source programming language. With the help of R, it is easy to perform data manipulation with packages such as plyr, dplyr, and tidy. It is excellent when it comes to data visualization and analysis with packages such as ggplot2, lattice, plotly etc.

3.2 Python

Python is one of the most powerful Data Analytics tools that is available to the user. It comes with a wide set of package/libraries. Python is a free, open-source software that can be used for a high level of visualisation and comes with packages such as Matplotlib, Seaborn. Pandas is one of the widely used data analytics library that

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

comes with Python.

3.3 Excel

Microsoft Excel is a simple yet powerful tool for data collection and analysis. It is a part of the Microsoft Office tool suite, and is readily available, widely used and easy to learn. Microsoft Excel can be considered as a great starting point for data analysis. The Data Analytics Tool pack in Excel offers a variety of options to perform statistical analysis of your data.

3.4 Tableau

Tableau is a Business Intelligence Tool developed for data analysts where one can visualize, analyze, and understand their data. Tableau provides fast analytics, can explore wide types of data like spreadsheets, databases, data on Hadoop and cloud services. It is easy to use, as it has a powerful GUI.

3.5 QlikView

QlikView is a Self-Service Business Intelligence, Data Visualization, and Data Analytics tool. It aims to accelerate business value through data by providing features such as Data Integration, Data Literacy, and Data Analytics.

3.6 Power BI

Power BI is yet another powerful business analytics solution by Microsoft. You can visualize your data, connect to many data sources and share the outcomes across your organization. Power BI can be integrated with other Data Analytics Tools, including Microsoft Excel.

3.7 SAS

SAS is a statistical software suite widely used for data management and predictive analysis.

3.8 KNIME

KNIME or Konstanz Information Miner is an open-source, free data analytics, reporting, and integration platform. It is built for analytics on a GUI workflow and helps in gathering data as well as creating models used for deployment and production.

3.9 Apache Spark

Apache Spark is one of the most active apache projects. Spark is a framework that supports applications while maintaining MapReduce's scalability and fault tolerance. Resilient distributed datasets (RDDs), a read-only set of items partitioned over a set of devices to suit user needs, are provided by Spark technologies.

3.10 Importance of Health Care Data Analytics

The collection of data in health care settings has become more streamlined in recent years. Both historical information and current information can be used for datasets to track trends and make predictions.

The fee-for-service style of health care is becoming a thing of the past. There is a growing demand for patient-centric, or value-based, medical care which has led to a considerable shift towards predictive and preventive measures in regards to public health in recent years. Data makes this possible. Instead of simply treating the symptoms as they present, practitioners are able to identify patients at high risk of developing chronic illnesses and help to treat an issue before it surfaces. This helps to lower costs for the practitioner, insurance company and patient as the preventive treatment may help to stave off long-term issues and expensive hospitalizations[4].

3.11 Potential Benefits of using Data Analytics in Healthcare

Modern analytics gives possibilities not only to have insight in historical data, but also to have information necessary to generate insight into what may happen in the future. The emphasis on reform has prompted payers and suppliers to pursue data analysis to reduce risk, detect fraud, improve efficiency and save lives. Everyone—payers,

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

providers, even patients—are focusing on doing more with fewer resources. Thus, some areas in which enhanced data and analytics can yield the greatest results include various healthcare stakeholders[7]. The benefits and effects from Data in healthcare can be divided into four groups [8]:

1. Improving the quality of healthcare services.
2. Supporting the work of medical personnel
3. Supporting scientific and research activity
4. Business and management

IV. CONCLUSION

Data Analytics has applications in all the major fields like Business, Stock Market, Education, Health care etc. The health care industry is one of the leading sectors where huge revenue will be generated across globe as the numbers of patients are increasing drastically with the population.Streaming data acquisition is demanding efficient real time analytic frameworks for the future. Data Analytics could be used along with machine learning techniques for studies related to the spread of pandemics, the efficacy for the treatment of acute diseases to identify earlier[9,11], or emotion recognition[10].

REFERENCES

[1]. https://digitalhealth.folio3.com/blog/what-is-the-role-of-data-analytics-in-healthcare/.
[2]. Sweety Bakyarani.E and Dr. Srimathi.H, "A Survey of Tools for Data Analytics", International Journal of Advance Engineering and Research Development Volume 5, Issue 02, February -2018 ISSN (O): 2348-4470.
[3]. https://online.shrs.pitt.edu/blog/data-analytics-in-health-care/
[4]. https://www.mygreatlearning.com/blog/top-data-analytics-tools/
[5]. Rahul Reddy Nadikattu Research on Data Science, Data Analytics and Big Data", International Journal Of Engineering, Science AND - Volume 9, Issue 5, May 2020 Pages: 99-105.
[6]. https://www.coursera.org/articles/healthcare-analytics.
[7]. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00553-4.
[8]. Batko K. Możliwości wykorzystania Big Data w ochronie zdrowia. Roczniki Kolegium Analiz Ekonomicznych. 2016;42:267–82.
[9]. Corsi A, de Souza FF, Pagani RN, et al. Big data analytics as a tool for fighting pandemics: a systematic review of literature. J Ambient Intell Hum Comput. 2021;12:9163–80.
[10]. Jain N, Gupta V, Shubham S, et al. Understanding cartoon emotion using integrated deep neural network on large dataset. Neural Comput Appl. 2021.
[11]. Wu J, et al. Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations. J Med Internet Res. 2020;22(10): e21980.