# Health Care Systems using Machine Learning

**Asha Abraham, Sumithra M**

***Abstract*****:** ***The study of different age people from different lifestyle, what all symptoms came for the disease, at what stage, what measures taken to get rid of, later changes, what were the side effects, how it decreased or increased will help in future prediction of possible chance for illness in others. A System which having the above details, suggestions from good and expert practitioners, can be used to give warnings to people about the possibility to get affected after 5 or 10 years and to take pre-cautions . Combinational outcome of Descriptive, Predictive and Prescriptive Data Analytics methods on past and present structured and unstructured Big Data can be used to predict future after effects which should be prioritized and treated. Suggestions can be given to people to take appointment with dieticians, change food habits, perform exercises, practice remedial measures etc. Right step at correct time save lives, gives happiness to families, reduces medical expenditures. The relevant information hidden in massive amount of data are made available by the AI assistants to make better clinical decisions in the functional areas of healthcare. To make such a system, a detailed study on big data, analytics methods, health care, practicing methods, electronic health records etc is required. A preventive guidance and less cost expert system which is helpful for common man and experts, for immediate solution and care can be developed in future. Deep machine learning algorithms to detect later possibility of occurrence should be developed. For this a study on Big Data and Health Care Analytics is done.***

***Index Terms*****:** ***Big Data, Data Analytics, Descriptive Analytics, Predictive Analytics, Prescriptive Analytics, Volume, Velocity, Veracity.***

## I. INTRODUCTION

Nowadays, a huge no: of researches focus on data analysis or data mining for healthcare data on technical details, but one of the greatest challenges is how to develop a comprehensive healthcare system for effectively manage multisource heterogeneous healthcare data with particular technical features. Thus, a detailed design of intelligent healthcare systems assisted by data analytics make the following contributions: It proposes a unified data collection layer for integrating the healthcare data from public sources and personal devices. It establishes a cloud-enabled and data-driven platform for multisource heterogeneous healthcare data storage and analysis for users[1]. In recent years, the healthcare system in the United States has been rapidly adopting electronic health records, which will greatly increase the amount of clinical data available electronically. Simultaneously, rapid progress has also been made in clinical analysis techniques to analyze large amounts of data and to gather new insights from the analyses. As a result, we have unprecedented opportunities to reduce the cost of healthcare by using mass data[1].

Today's healthcare industries are moving from volume based business into value based business, which requires an overwork from doctors and nurses to be more productive and efficient. This will improve healthcare practice, changing individual life style and driving them into longer life, prevent diseases, illnesses and infections.[13].

Take advantage of the massive amounts of data and provide right intervention to the right patient at the right time[12].

Improving the health of patients while decreasing the costs of care. [12]

Develop algorithms to predict the number of days a patient will spend in a hospital in the next year by making use of combination of several predictive models.[12]

With the adoption of EHRs and other digital tools, much more structured and unstructured data is now available to be processed and analyzed.[11]

Emergence of Machine Learning and Analytics tools that can make useful inferences and predictions based on available data.[11]

The dramatically positive impact Analytics can have on the pressures health systems face to be more efficient and improve clinical outcomes[11].

Computing and Cloud infrastructure that allows for large quantities of data processing in real time, pushed to the right operational and clinical decision makers at the right time. A decade ago, ideas such as video calls with doctors, doctor-on-demand and Wi-Fi-enabled blood pressure monitors were a fantasy. Today, they are real — and mainstream. Both caregivers and consumers are adopting these tools at a rapid pace to increase efficiency, lower costs and improve patient satisfaction. And they are producing vast amounts of data — data with which traditional warehouses simply cannot keep up.[11]

Leadership of health systems that is more

**Revised Manuscript Received on December 22, 2018.**

**Asha Abraham** Computer Science Department, Veltech Rangarajan Dr, SagunthalaR&D Institute of Science and Technology, Avadi, Chennai, Tamil-Nadu, India

**Sumithra M** Computer Science Department, Veltech Rangarajan Dr, SagunthalaR&D Institute of Science and Technology, Avadi, Chennai, Tamil-Nadu, India

technology-friendly, willing to adopt new tools and open to adopting new technologies. In fact, a lot of them recognize that new-world Data Analytics is no longer optional; it's a must-have for healthcare institutions to survive.[11]

Connectivity between Healthcare Analytics System and Individuals (Medical Staff and Patients)

Connectivity approaches generate thoughts and ideas from connected networks of minds and leverages prior experiences with the utilization of technologies in our everyday life. (Siemens 2004) Moreover, Mc Horney (2009) has added that healthcare analytics is not solely regarding technology and the knowledge however; it is also regards how much individuals are attached to and familiar with medical care systems and their personal skills such as ability to learn and adopt such systems in their life, as different people have different attitudes and reasons for not accepting such technologies, especially older people[13].

Understanding and predicting diseases require an aggregated approach where structured and unstructured data stemming from a myriad of clinical and nonclinical modalities are utilized for a more comprehensive perspective of the disease states. An aspect of healthcare research that has recently gained traction is in addressing some of the growing pains in introducing concepts of big data analytics to medicine. Researchers are studying the complex nature of healthcare data in terms of both characteristics of the data itself and the taxonomy of analytics that can be meaningfully performed on them[10].

Principally, to perform good data analytics, first of all we should teach individuals how to understand and realize the importance of dealing with such data, for instance how to deal with breast cancer (Hanoch 2012). However, Miron et al (2011) believe d that whatever and how much our patients are educated and skilled to provide us with the data we expect, medical professionals still highly need to test and clarify this data to consider it and keep it on record. Also, he added that once when the data has been tested and Australasian Conference on Information Systems Al Khatib et al 2015, Sydney Healthcare Data Analytics

are clarified, we then need to find out how to change an individual's behavior, starting with parents and guardians who are responsible for raising their children[13].

**Operational Efficiency Through Predictive Analytics**[3]

Let's focus on a little bit deeper: Enabling healthcare providers to do more with less time and effort. This is a critical need for hospitals to remain viable because of a number of forces:

An aging population is increasing the demand for expensive, critical care.

There is an ever-increasing shortage of qualified physicians and nurses and caregivers.

The move toward value-based care has increased the need for accountability and transparency.

These forces are turning out to be so strong and powerful that many healthcare organizations are being forced to consolidate, and this trend will continue. Efficiency is therefore at the top of agenda, and Data Analytics has an important seat at the table. In fact, hospitals today face the same cost and revenue pressure that retail, transportation and airlines have faced for years. Healthcare providers can't keep spending their way out of trouble by investing in more and more infrastructure; instead, they must optimize their use of the assets currently in place. To do this, providers need to consistently make excellent operational decisions, as these other industries have, through Predictive Analytics that learn continually and use optimization algorithms.

## II. BACKGROUND

Big Data

Big Data means all kind of data ie structured, semi-structure and unstructured( text form), consists of 3 Vs- Volume, Velocity and Variety, Veracity and Value.

3-Vs[4]: Volume: The quantity of data generated every second is too large which could only be handled only by distributed systems connected by networks brought together by softwares

Velocity: The speed at which vast amount of data being generated, collected and analyzed.

Veracity: Veracity is the quality or trustworthiness of data, the reliability and accuracy of the content.

Variety: Variety is the different types of data like text, video, image etc.

Value: Value refers the worth of data being extracted. In health care field based on the value of data, government can predict the infectious diseases to give alertness among people.

Variability[2]:this refers to the variations happened when data is taken from life cycle. To provide valuable information hidden and unforeseen variables should be find out

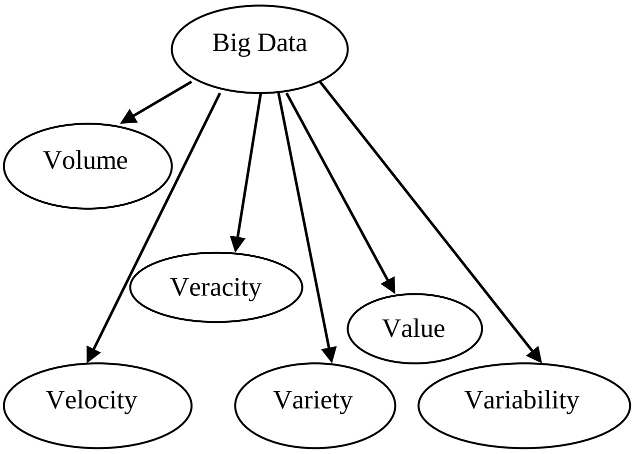


Figure 1.6-Vs of Big Data

Data Science is a discipline that utilizes a combination of computational, mathematical and statistical tools to acquire, analyse and process Big Data.

Machine Learning is a discipline under Data Science that imparts and empower machines that makes them act for themselves.

## III. DATA ANALYTICS

Analytics is the analysis of data, typically large sets of data, by the use of mathematics, statistics and computer software. It is the science of using data to build models that leads to better decisions that in turn add values to individuals, companies and institutions.

Statistics is a branch of mathematics that deals with the collection, organization, analysis, and interpretation of numerical data. It is especially useful in drawing general conclusions about a set of data from a sample of the data[16]. This summarize a process that an analyst uses to characterize a data set. If the data set depends on a sample of a larger population, then the analyst can develop interpretations about the population primarily based on the statistical outcomes from the sample. A form of mathematical analysis that uses quantified models, representations and synopses for a given set of experimental data or real-life studies . Study the methodologies to gather, review, analyze and draw conclusions from data. Some statistical measures include mean, regression analysis, skewness, kurtosis, variance and analysis of variance.

Statistical analysis involves the process of gathering and evaluating data and then summarizing the data into a mathematical form. Statistical methods analyze large volumes of data and their properties. Statistical data is gathered using a sample procedure or other method. Two types of statistical methods are used in analyzing data: One is descriptive statistics which is used to synopsize data from a sample exercising the mean or standard deviation. Second is the inferential statistics in which data is viewed as a subclass of a specific population.

A mean is the mathematical average of a group of two or more numerals. The mean for a specified set of numbers can be computed in multiple ways, including the arithmetic mean, which shows how well a specific commodity performs over time, and the geometric mean which shows the performance results of an investor's portfolio invested in that same commodity over the same period.

Regression analysis determines the extent to which specific factors such as interest rates, the price of a product or service, or particular industries or sectors influence the price fluctuations of an asset. This is depicted in the form of a straight line called linear regression.

Skewness describes the degree a set of data varies from the standard distribution in a set of statistical data. Most data sets, including commodity returns and stock prices, have either positive skew, a curve skewed toward the left of the data average, or negative skew, a curve skewed toward the right of the data average.

## IV. CONCLUSION

Kurtosis measures whether the data are light-tailed (less outlier-prone) or heavy-tailed (more outlier-prone) than the normal distribution. Data sets with high kurtosis have heavy tails, or outliers, which implies greater investment risk in the form of occasional wild returns. Data sets with low kurtosis have light tails, or lack of outliers, which implies lesser investment risk.

Variance is a measurement of the span of numbers in a data set. The variance measures the distance each number in the set is from the mean. Variance can help determine the risk an investor might accept when buying an investment.

## REFERENCES

1. Intelligent Healthcare Systems Assisted by Data Analytics and Mobile Computing Xiao Ma,1 Zie Wang,1 Sheng Zhou,1 Haoyu Wen,1 and Yin Zhang 1,2 Received 9 January 2018; Accepted 20 May 2018; Published 3 July 2018 Academic Editor: Javier Prieto Wireless Communications and Mobile Computing Volume 2018, Article ID 3928080, 16 pages https://doi.org/10.1155/2018/3928080
2. Big Data in Healthcare Management: A Review of Literature
3. Senthilkumar SA1, Bharatendara K Rai2, Amruta A Meshram2, Angappa Gunasekaran3, Chandrakumarmangalam S4American Journal of Theoretical and Applied Business 2018; 4(2): 57-69 http://www.sciencepublishinggroup.com/j/ajtab doi: 0.11648/j.ajtab.20180402.14 ISSN: 2469-7834 (Print); ISSN: 2469-7842
4. A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining, Md Saiful Islam, Md Mahmudul

Hasan, Xiaoyi Wang, Hayley D. Germack, and Md Noor-E-Alam,Available:https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6023432/#B8-healthcare-06-00054

5. Big Data and Its Challenges V. Maria Antoniate Martin*1, Dr. K. David2, A.Vignesh3 . Volume 3, Issue 3 | March-April-2018 | ISSN : 2456-3307. http:// ijsrcseit.com
6. A Case Study based OpenEHR .
7. Available:https://www.sciencedirect.com
8. Health Analytics Types, Functions and Levels: A Review of Literature Mohamed KHALIFA1. Data, Informatics and Technology: An Inspiration for Improved Healthcare A. Hasman
9. Artificial intelligence in healthcare: past, present and future. Fei Jiang1, Yong Jiang2, Hui Zhi3, Yi Dong4, Hao Li5, Sufeng Ma6, Yilong Wang7, Qiang Dong4, Haipeng Shen8, Yongjun Wang9 https://svn.bmj.com/content/2/4/230
10. Rajesh, M., and J. M. Gnanasekar. "Path Observation Based Physical Routing Protocol for Wireless Ad Hoc Networks." Wireless Personal Communications 97.1 (2017): 1267-1289.
11. Rajesh, M., and J. M. Gnanasekar. "Sector Routing Protocol (SRP) in Ad-hoc Networks." Control Network and Complex Systems 5.7 (2015): 1-4.
12. Rajesh, M. "A Review on Excellence Analysis of Relationship Spur Advance in Wireless Ad Hoc Networks." International Journal of Pure and Applied Mathematics 118.9 (2018): 407-412.
13. Rajesh, M., et al. "SENSITIVE DATA SECURITY IN CLOUD COMPUTING AID OF DIFFERENT ENCRYPTION TECHNIQUES." Journal of Advanced Research in Dynamical and Control Systems 18.
14. Rajesh, M. "A signature based information security system for vitality proficient information accumulation in wireless sensor systems." International Journal of Pure and Applied Mathematics 118.9 (2018): 367-387.
15. Rajesh, M., K. Balasubramaniaswamy, and S. Aravindh. "MEBCK from Web using NLP Techniques." Computer Engineering and Intelligent Systems 6.8: 24-26.