# ANALYTICS IN HOSPITAL HEALTH-CARE DATA

# LITERETURE SURVEY

TEAM ID                    :   PNT2022TMID26228

TEAM MEMBERS               :   V.NANDITHA

                               M.VISHALI

                               R.DIVYA DARSHANI

                               B.VAIDEGI

 DEPARTMENT                :   IT

COLLEGE NAME               :   PANIMALAR INSTITUTE OF

                               TECHNOLOGY,

                               CHENNAI-123

**M.S.SIMI AND K.S.NAYAKI GAVE A REVIEW ON "DATA ANALYTICS IN HEALTH CARE" AT INTERNATIONAL CONFERENCE ON CIRCUIT, POWER AND COMPUTING TECHNOLOGIES(2019).** The knowledge discovery in database (KDD) is alarmed by the advancement of strategies and systems for making the utilization of data. A standout amongst the most vital stride of the KDD is the data analytics. Data analytics is one of the extensively researched areas inferable from the wide impact showed by this computational strategy on differing fields, such as, Artificial Intelligence (AI), databases, statistics, and visualization. It has unlimited applications and ways to deal with analytics the data in suitable ways. Both the data analytics and medicine have raised some of dependable early discovery frameworks and different medical services related frameworks from the medical data. Medical data analytics is a dynamic interdisciplinary area of research that is viewed as the result of applying artificial intelligence and data analytics concepts to the field of clinical and medical services. We have reviewed the different papers intricate in this field in terms of technique, algorithms and results. The aim of this research work is to give a review on the foundation benchmarks in analytics of infertility, and present the findings and results of past researches on utilizing data analytics procedures to analyze electronic health records.

**M.AMBIGAVATHI AND D.SRIDHARAN,"BIG DATA ANALYTICS IN HEALTH CARE,"AT 2018 TENTH INTERNATIONAL CONFERENCE ON (ICOAC).** The pace of

both digital innovation and technology disruption is refining the healthcare industry at an exponential rate. The large volume of healthcare data continues to mount every second, making it harder and very difficult to find any form of useful information. Recently, big data is shifting the traditional way of data delivery into valuable insights using big data analytics method. Big data analytics provides a lot of benefits in the healthcare sector to detect critical diseases at the initial stage and deliver better healthcare services to the right patient at the right time so that it improves the quality of life care. Big data analytics tools play an essential role to analyze and integrate large volumes of structured, semi-structured and unstructured vital data rapidly produced by the various clinical, hospitals, other social web sources and medical data lakes. However, there are several issues to be addressed in the current health data analytics platforms that offer technical mechanisms for data collection, aggregation, process, analysis, visualization, and interpretation. Due to lack of detailed study in the previous literature, this article inspects the promising field of big data analytics in healthcare. This article examines the unique characteristics of big data, big data analytical tools, different phases followed by the healthcare economy from data collection to the data delivery stage. Further, this article briefly summarizes the open research challenges with feasible findings, and then finally offers the conclusion.

**Z.FEI, Y.RYEZNIK, A.SVERDLOW, C.W.TAN AND W.K.WONG GAVE "AN OVERVIEW OF HEALTH CARE DATA ANALYTICS WITH APPLICATION TO THE COVID-19 PANDEMIC",2020.** In the era of big data, standard analysis tools may be inadequate for making inference and there is a growing need for more efficient and innovative ways to collect, process, analyze and interpret the massive and complex data. We provide an overview of challenges in big data problems and describe how innovative analytical methods, machine learning tools and metaheuristics can tackle general healthcare problems with a focus on the current pandemic. In particular, we give applications of modern digital technology, statistical methods,data platforms and data integration systems to improve diagnosis and treatment of diseases in clinical research and novel epidemiologic tools to tackle infection source problems, such as finding Patient Zero in the spread of epidemics. We make the case that analyzing and interpreting big data is a very challenging task that requires a multi-disciplinary effort to continuously create more effective methodologies and powerful tools to transfer data information into knowledge that enables informed decision making.

**S.KAPOOR, L.KASAR, A.MANDOLE AND J.MAHAJAN GAVE AN APPROACH IN "HEART DISEASE PREDICTION USING MACHINE LEARING AND DATA ANALYTICS APPROACH"(2019).** In recent times, Machine Learning has played a significant role in the healthcare industry and amongst all of the major diseases, heart disease is one of the significant and most critical diseases to predict. There is a rapid increase in the number of cases each day. It has been observed that in every minute, 4 people between the age group of 30-50 get a stroke, so we are using machine learning algorithms to mitigate this problem. Kaggle used the heart disease dataset used for this project. This paper demonstrates the prediction of heart disease using multiple machine learning classification algorithms such as Naive Bayes, Random Forest, SVM etc., and compares their accuracy scores. Later on, Stacking Ensemble Learning Technique is used to boost our classification models' performance.

**S.NOUREDDINE,Z.BAARIR,A.TOUMI,A.BETKA,N.KAZAR AND A.N.BEHARKAT. "SMART BREAST CANCER PREDICTION USING DATA MINING PROCESS".(2017)** In terms of mortality, breast cancer occupies the 5th place in the world. According to the World Health organization (WHO), breast cancer is the principal cause of cancer death in the last years, especially among women. Following this very alarming finding, research was developed using predictive

approaches based on the analysis of a large volume of data. In the literature, data analysis was carried out through data mining using several techniques, especially metaheuristics, to aid in the preventive or predictive decision of breast cancer. In this paper, a recent metaheuristic called Symbiotic Organisms Search (SOS) based on a bio-inspired phenomenon is combined with an Artificial Neural Network (ANN), will be utilized in datamining for a predictive study of breast cancer disease. This paper contains also an intelligent application using the SOS algorithm. The experiments generated very confident results in terms of predictive decision following the classification of data.

## A.T.NAGI, M.JAVED AWAN, R.JAVED AND N.AYESHA. "A COMPARISION OF TWO-STAGE CLASSIFIER ALGORITHM WITH ENSEMBLE TECHNIQUES ON DETECTION OF DIABETIC RETINOPATHY"(2018).

The Diabetic retinopathy is disease of the human eye that causes retinal damage in diabetic patients. It further leads to the blindness. The machine learning techniques plays an important rule to predict the early diabetic retinopathy which avoided from the intensive labor. In this paper we used the novel technique, the Two Stage Classifier, an ensemble technique which combines various machine learning algorithms for classification. In the subject paper, the classifier is applied to predict Diabetic retinopathy (DR), a disease of the human eye that causes retinal damage in diabetic patients and ultimately lead to complete blindness. The problem lies in the fact that it is time consuming to detect this disease but an early detection of the disease is essential to avoid complete blindness. We apply machine learning algorithms to determine the existence of DR and compare the accuracies of the applied techniques. The Two Stage Classifier, turns out to be better not only in terms of parallelism but also in terms of accuracy.

## A.ARJUN, A.SRINATH AND B.R.CHANDAVARKAR. "PREDICTIVE ANALYTICS AND DATA MINING IN HEALTHCARE".AT 2021 12TH INTERNATIONAL TECHNOLOGIES.

Machine Learning and Data Mining for healthcare. There has been an enormous growth in the field of HIT (health information technology) in the recent years. Be it detection of certain diseases, scanning of organs, finding tumors, these machine oriented operations without human intervention, have certainly increased the quality of medical attention one can get, and the technology required has come a long way. Health data tends to be inherently complex with exceptions in almost all cases. Data mining is the technique of converting raw data into a meaningful format. Analysis and prediction on such data, although computationally and algorithmically complex, is an emerging technology that is a small step to more proactive and preventive automated treatment options. There are various data mining techniques such as classification, clustering, association, regression, prediction, pattern recognition . Even the efficiency of certain medicines can be found using machine learning techniques, which is a life saving and cost effective method. In this paper, we are going to use machine learning as a tool for predictive analysis to predict chronic kidney diseases based on the Chronic disease dataset taken from UCI ML repository. We will be applying machine learning algorithms, specifically decision trees, to build a classifier to predict if a person has the disease or not. This paper shows the issue that specific machine learning algorithms need to be tailor-made to specific nature of medical data.

## "ANALYSIS OF RESEARCH IN HEALTHCARE DATA ANALYTICS"(2017):

The main aim of this paper is to provide a deep analysis on the research field of healthcare data analytics., as well as highlighting some of guidelines and gaps in previous studies. This study has focused on searching relevant papers about healthcare analytics by searching in seven popular databases such as google scholar and springer using specific keywords, in order to understand the healthcare topic and conduct our literature review. The paper has listed some data analytics tools and techniques that have been used to improve healthcare performance in many areas such as: medical

operations, reports, decision making, and prediction and prevention system. Moreover, the systematic review has showed an interesting demographic of fields of publication, research approaches, as well as outlined some of the possible reasons and issues associated with healthcare data analytics, based on geographical distribution theme.