

# ANALYTICS FOR HOSPITAL'S HEALTH-CARE DATA LITERATURE SURVEY

**DOMAIN NAME** : DATA ANALYTICS  
**TEAM ID** : PNT2022TMID29546  
**BATCH** : B7-1A3E  
**TEAM LEADER** : THIRUVIKRAMAN G K  
**TEAM MEMBER 1** : VIGNESH S  
**TEAM MEMBER 2** : VINOTH S  
**TEAM MEMBER 3** : HEMANTH S  
**TEAM MEMBER 4** : TAMIL SELVAN G

## **PAPER 1:** Big Data Analytics in Healthcare Systems

**Published Year** : 2019  
**Author** : Lidong Wang  
**Journal Name** : International Journal of Mathematical, Engineering and Management Sciences

**Summary** : Big Data analytics can improve patient outcomes, advance and personalize care, improve provider relationships with patients, and reduce medical spending. This paper introduces healthcare data, big data in healthcare systems, and applications and advantages of Big Data analytics in healthcare. We also present the technological progress of big data in healthcare, such as cloud computing and stream processing. Challenges of Big Data analytics in healthcare systems are also discussed.

## **PAPER 2** : Analysis of Research in Healthcare Data Analytics

**Published Year** : 2015  
**Author** : Mohammad Ahmad Alkhatib  
**Journal Name** : Australasian Conference on Information Systems

**Summary** : 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.

## **PAPER 3** : Big data in healthcare: management, analysis and future prospects

**Published Year** : 2019  
**Author** : Sabyasachi Dash  
**Jounal Name** : 3B's Research Group, Headquarters of the European Institute of Excellence on Tissue Engineering and Regenerative Medicine.

**Summary :** ‘Big data’ is massive amounts of information that can work wonders. It has become a topic of special interest for the past two decades because of a great potential that is hidden in it. Various public and private sector industries generate, store, and analyze big data with an aim to improve the services they provide. In the healthcare industry, various sources for big data include hospital records, medical records of patients, results of medical examinations, and devices that are a part of internet of things. Biomedical research also generates a significant portion of big data relevant to public healthcare. This data requires proper management and analysis in order to derive meaningful information. Otherwise, seeking solution by analyzing big data quickly becomes comparable to finding a needle in the haystack. There are various challenges associated with each step of handling big data which can only be surpassed by using high-end computing solutions for big data analysis. That is why, to provide relevant solutions for improving public health, healthcare providers are required to be fully equipped with appropriate infrastructure to systematically generate and analyze big data. An efficient management, analysis, and interpretation of big data can change the game by opening new avenues for modern healthcare. That is exactly why various industries, including the healthcare industry, are taking vigorous steps to convert this potential into better services and financial advantages. With a strong integration of biomedical and healthcare data, modern healthcare organizations can possibly revolutionize the medical therapies and personalized medicine.

#### **PAPER 4 : Optimizing the Electronic Health Records Through Big Data Analytics: A Knowledge-Based View**

**Published Year :** 2019

**Author :** Caifeng Zhang

**Journal Name :** Special section on Advanced Optical Imaging for Extreme Environments-IEEE XPLORE.

**Summary :** Many hospitals are suffering from ineffective use of big data analytics with electronic health records (EHRs) to generate high quality insights for their clinical practices. Organizational learning has been a key role in improving the use of big data analytics with EHRs. Drawing on the knowledge-based view and big data lifecycle, we investigate how the three modes of knowledge can achieve meaningful use of big data analytics with EHRs. To test the associations in the proposed research model, we surveyed 580 nurses of a large hospital in China in 2019. Structural equation modelling was used to examine relationships between knowledge mode of EHRs and meaningful use of EHRs. The results reveal that know-what about EHRs utilization, know-how EHRs storage and utilization, and know-why storage and utilization can improve nurses’ meaningful use of big data analytics with EHRs. This study contributes to the existing digital health and big data literature by exploring the proper adaptation of analytical tools to EHRs from the different knowledge mode in order to shape meaningful use of big data analytics with EHRs.

#### **PAPER 5 : BIG DATA for Healthcare: A Survey**

**Published Year :** 2019

**Author :** Safa Bahri

**Journal Name :** Special section on Advanced Optical Imaging for Extreme Environments-IEEE XPLORE.

**Summary :** Recently, the massification of new technologies, which has been adopted by a large majority of the world population, has accumulated a tremendous amount of data, including clinical data. This clinical data have been gathered up and interpreted by medical organizations in order to gain insights and knowledge useful for clinical decisions, drug recommendations, and better diagnoses, among many other uses. This paper highlights the enormous impacts of big data on medical stakeholders, patients, physicians, pharmaceutical and medical operators, and healthcare insurers, and also reviews the different challenges that must be taken into account to get the best benefits from all this big data and the available applications.

## **PAPER 6 : Data Analysis and Forecasting of Tuberculosis Prevalence Rates for Smart Healthcare Based on a Novel Combination Model**

**Published Year :** 2018

**Author :** Jiyang Wang

**Journal Name :** Applied Science-Multidisciplinary Digital Publishing Institute

**Summary :** In recent years, healthcare has attracted much attention, which is looking for more and more data analytics in healthcare to relieve medical problems in medical staff shortage, ageing population, people living alone, and quality of life. Data mining, analysis, and forecasting play a vital role in modern social and medical fields. However, how to select a proper model to mine and analyze the relevant medical information in the data is not only an extremely challenging problem, but also a concerning problem. Tuberculosis remains a major global health problem despite recent and continued progress in prevention and treatment. There is no doubt that the effective analysis and accurate forecasting of global tuberculosis prevalence rates lay a solid foundation for the construction of an epidemic disease warning and monitoring system from a global perspective. In this paper, the tuberculosis prevalence rate time series for four World Bank income groups are targeted. Kruskal–Wallis analysis of variance and multiple comparison tests are conducted to determine whether the differences of tuberculosis prevalence rates for different income groups are statistically significant or not, and a novel combined forecasting model with its weights optimized by a recently developed artificial intelligence algorithm—cuckoo search—is proposed to forecast the hierarchical tuberculosis prevalence rates from 2013 to 2016. Numerical results show that the developed combination model is not only simple, but is also able to satisfactorily approximate the actual tuberculosis prevalence rate, and can be an effective tool in mining and analyzing big data in the medical field.

## **PAPER 7 : Big data analytics in healthcare: promise and potential**

**Published Year :** 2014

**Author :** Wullianallur Raghupathi

**Journal Name :** Health Information Science and Systems

**Summary :** The paper describes the nascent field of big data analytics in healthcare, discusses the benefits, outlines an architectural framework and methodology, describes examples reported in the literature, briefly discusses the challenges, and offers conclusions. The healthcare industry historically has generated large amounts of data, driven by record keeping, compliance & regulatory requirements, and patient care. While most data is stored in hard copy form, the current trend is toward rapid digitization of these large amounts of data. The paper provides a broad overview of big data analytics for healthcare researchers and practitioners. In the future we'll see the rapid, widespread implementation and use of big data analytics across the healthcare organization and the healthcare industry. To that end, the several challenges highlighted above, must be addressed. Big data analytics in healthcare is evolving into a promising field for providing insight from very large data sets and improving outcomes while reducing costs. Its potential is great; however there remain challenges to overcome.

## **PAPER 8 : Exploring big data analytics in health care**

**Published Year :** 2020

**Author :** V. Santhi

**Journal Name :** International Journal of Intelligent Networks

**Summary :** Health care Industries are facing lot of challenges in maintaining patient information across various databases due to storage issues. In order to extract patient information, preprocessing techniques can be applied in the process of data mining across databases. But as the data is growing enormously with rapid speed, data mining techniques are becoming obsolete due to issues such as Storage, Speed. So, cost optimization has become one of the major requirements in health industry as there is huge burden in maintaining large volumes of patient's information using traditional databases. Here Big Data plays a vital role in storing huge volumes of patient information using storage mechanisms such as HDFS, HBase. Many issues in health care are discussed in this paper such as prediction of diseases, getting patients information across databases as a single view.

**PAPER 9 : Disease Prediction by Machine Learning Over Big Data From Healthcare Communities**

**Published Year :** 2017

**Author :** Min Chen

**Journal Name :** Special Section on Health Care Big Data – IEEE XPLORE

**Summary :** With big data growth in biomedical and healthcare communities, accurate analysis of medical data benefits early disease detection, patient care, and community services. However, the analysis accuracy is reduced when the quality of medical data is incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of disease outbreaks. In this paper, we streamline machine learning algorithms for effective prediction of chronic disease outbreak in disease-frequent communities. We experiment the modified prediction models over real-life hospital data collected from central China in 2013–2015. To overcome the difficulty of incomplete data, we use a latent factor model to reconstruct the missing data. We experiment on a regional chronic disease of cerebral infarction. We propose a new convolutional neural network (CNN)-based multimodal disease risk prediction algorithm using structured and unstructured data from hospital. To the best of our knowledge, none of the existing work focused on both data types in the area of medical big data analytics. Compared with several typical prediction algorithms, the prediction accuracy of our proposed algorithm reaches 94.8% with a convergence speed, which is faster than that of the CNN-based unimodal disease risk prediction algorithm.

**PAPER 10 : Data Analytics and Its Advantages for Addressing the Complexity of Healthcare: A Simulated Zika Case Study Example**

**Published Year :** 2019

**Author :** Lily Popova Zhuhadar

**Journal Name :** Applied Science-Multidisciplinary Digital Publishing Institute

**Summary :** The need to control rising costs in healthcare has led to an increase in the use of data analytics to develop more efficient healthcare business models. This article discusses a simulation that uses data analytics to minimize the number of physicians and nurses needed in healthcare facilities during a crisis situation. Using a hypothetical emergency scenario, the hospital uses a healthcare analytical system to predict the necessary resources to govern the situation. Based on historical data regarding the flow of patients through the facility, a discrete-event simulation estimates resource scheduling and the resulting impact on both wait times and personnel demand. Furthermore, the value of multiple replications for discrete-event simulation models is discussed and defined, along with factors that enable greater control of multiple design points with this simulated experiment. The results of this study demonstrate the value of simulation modeling in effective resource planning. The addition of only a single doctor significantly reduced predicted wait times for patients during the crisis. Further, the findings support the use of data analytics and predictive modeling to mitigate rising healthcare costs in the United States through efficient planning and resource allocation.