



## BIG DATA ANALYTICS IN HEALTHCARE SERVICES

**Gunasekar Thangarasu\***

Faculty of Engineering and Technology  
Linton University College  
Malaysia

**Kayalvizhi Subramanian**

Faculty of Fundamental and Applied Sciences  
Universiti Teknologi PETRONAS  
Malaysia

**S. Kayalvizhi**

Faculty of Built Environment  
Linton University College  
Malaysia

**\*Corrospoding author's Email:** [gunasekar97@gmail.com](mailto:gunasekar97@gmail.com), [kayalvizhi\\_16005735@utp.edu.my](mailto:kayalvizhi_16005735@utp.edu.my),  
[skayalvizhi2012@gmail.com](mailto:skayalvizhi2012@gmail.com)

*Peer-review under responsibility of 3<sup>rd</sup> Asia International Multidisciplanry Conference 2019 editorial board  
(<http://www.utm.my/asia/our-team/>)*

*© 2019 Published by Readers Insight Publisher,  
lat 306 Savoy Residencia, Block 3 F11/1,44000 Islamabad. Pakistan,  
[info@readersinsight.net](mailto:info@readersinsight.net)*

*This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).*





## RESEARCH HIGHLIGHTS

This study addresses the healthcare services problems which focus on the upcoming and promising areas of medical research and proposed a novel approach integrating in big data analytics and Apache. The proposed approach will improve the healthcare services fastly and efficiently. The big data analytics can continually evaluate clinical data in order to improve the effective practices of physicians and improved patient care.

## RESEARCH OBJECTIVES

The patient community always expects to avail a broad range of healthcare services at affordable cost with personalized recommendation. In addition to physician's clinical diagnosis, they have an opportunity to gain more medical knowledge through digital platform. While a variety of research exists to support healthcare services in the healthcare organizations, still the services are delayed or not provides on right time especially for developing countries. This study intends to investigate the problems and proposed a novel approach integrated with emerging technologies in big data analytics and Apache Spark to improve the healthcare services.

## MATERIALS AND METHODS

The core value of big data has been effectively utilized for the identification of behavioural patterns of the consumers to develop innovative business services and solution. In the healthcare sector, the implication of big data serves predictive analytical techniques and machine learning platforms for the provision of sustainable solutions such as the implementation of treatment plan and personalized medical care. (1-3) The research design which includes, clinical data acquisition, data normalization, Feature extraction, Applied into proposed data analytic method then, perform the analysis and data visualization. The data samples have been used in this study is from a secondary data source which is acquired from California Behavioral Risk Factor Surveillance Survey in data. Gov website. The total acquired sample data sets are 550. After the data normalization received 505 complete data sets, it can be used for big data analysis. Apache Spark is utilized for batch and streaming process of big data analytics and machine learning to predict different human diseases. The collected data is stored in Map R-Database. This R-Database provides easy, measurable, fast read and writing of data. Apache drill is used for data exploration and preprocessing of the data in a schema-free Structured Query Language (SQL) query engine. Open Database Connectivity (ODBC) with apache drill provides facilitating tools for handling existing big data analysis. Napier Technologies enterprise capabilities provide for global data centre replication.

## RESULTS

Many big data analytics tools focus on either analytics or data processing. Some frameworks, like Apache Spark, support both. These enable developers and data scientists to use the same tools for real-time processing; complex extract, transform and load tasks; machine learning; reporting; and SQL. (4, 5) This is important because data science is a highly iterative process. A data scientist might create 100 models before arriving at one that is put into production. This iterative process often involves enriching the data to improve the results of the models.





## FINDINGS

Big data analytics involve a complex process that can span healthcare and business management, data scientists, developers and production teams. The data visualization report is the best solutions always cater to the comprehensive requirement of various stakeholders involved in the healthcare domain. With impact of big data, healthcare domain was revamped and offered intensive solutions for handling diversified big data source that range from patient health record to medical images. This study reviews various research attempts in establishing healthcare services and technologies. Big data analytics has the potential to transform the way healthcare providers use sophisticated technologies to gain insight from their clinical and other data repositories and make informed decisions. The proposed methodology will ensure a successful big data analytics implementation program, organizations should cultivate analytics architects from their existing talent pools of solution architects with analytics skills or data scientists with engineering and technical skills. Conversely, practitioners in this profession should develop analytic skills to fill the demand for this vital role.

## REFERENCES

1. Yichuan Wang, LeeAnn Kung and Terry Anthony Byrd. Big data analytics: Understanding its capabilities and potential benefits for healthcare organization. *Technological Forecasting and Social Change*. 2018; 126(1): 3-13.
2. Wil van der Aalst. *Data Science in Action*. In *Process Mining*, Springer: Berlin Heidelberg; 2016.
3. Ashwin Belle, Raghuram Thiagarajan, Reza Soroushmehr, Fatemeh Navidi, Daniel Beard and Kayvan najarian. Big Data Analytics in Healthcare. *BioMed Research International*. 2015; 3(1): 1-16.
4. Archenna. J and Mary Anita E.A. A survey of Big Data Analytics in Healthcare and Government. *Journal of Procedia Computer Science*, 2015; 50(1): 408-413.
5. Kirk Borne. Collaborative Annotation for Scientific Data Discovery and Reuse. *Bulletin of the Association for Information Science and Technology*. 2013; 39(4): 44-45.

