

Big Data Analytics in Healthcare Using Machine Learning Algorithms: A Comparative Study

<https://doi.org/10.3991/ijoe.v16i13.18609>

Sai Hanuman Akundi (✉), Soujanya R, Madhuri PM
Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India
a_saihanuman@hotmail.com

Abstract—In recent years vast quantities of data have been managed in various ways of medical applications and multiple organizations worldwide have developed this type of data and, together, these heterogeneous data are called big data. Data with other characteristics, quantity, speed and variety are the word big data. The healthcare sector has faced the need to handle the large data from different sources, renowned for generating large amounts of heterogeneous data. We can use the Big Data analysis to make proper decision in the health system by tweaking some of the current machine learning algorithms. If we have a large amount of knowledge that we want to predict or identify patterns, master learning would be the way forward. In this article, a brief overview of the Big Data, functionality and ways of Big data analytics are presented, which play an important role and affect healthcare information technology significantly. Within this paper we have presented a comparative study of algorithms for machine learning. We need to make effective use of all the current machine learning algorithms to anticipate accurate outcomes in the world of nursing.

Keywords—Big data, Big data Analytics, Predictive Analytics, Machine Learning, Apache Spark

1 Introduction

Big data on patient healthcare, compliance and numerous regulatory demands are created rapidly in all fields, including safety. As the population of the world continues to increase with human lifespan, models for treatment are evolving rapidly and data are required to support those decisions underlying such rapid change [4]. Recent frameworks for comprehensive data analysis of health information have been developed across a broad variety of contexts, such as the examination of patient preferences and the assessment of care costs and outcomes to determine the safest and most cost-effective therapies. In the study of health information, health-care information informatics is described as the absorption of medical sciences. Health computing includes the acquisition, storage and compilation of information to enhance healthcare providers' performance.

In today's technical environment, access, research, the most often spoken words are protecting and processing big data. Big data analysis is a tool for collecting data from

various sources and then extracting data and then analyzing it in order to find useful facts and statistics within this data recovery. Such data analysis not only helps to locate the secret facts and statistics of most big data, but it also categorizes the data or ranks the data in respect to the relevant data it contains. The method of extracting information from a wide variety of data in the brief big data analysis [14].

Predictive analytics in this industry will deliver outstanding results by improving service quality. There is a need for quantitative work in the healthcare industry [5]. However, the most important definitions include predictive analytics, including mathematical approaches like data mining and machine learning to forecast the present and the past. Predictive approaches used to determine the risk of re-admission to the hospital population of patients today. Such details allow physicians to make better choices on patient treatment. Predictive work calls for broad computer education awareness and use [4].

In the remainder we discussed firstly the common meanings and features, their use cases, and the description and types of analytics of big data. Section 3 we have addressed how big data analytics are useful for the maintenance of health data and how to predict them. We have discussed in the fourth section the different methods of Big Data and Algorithms for machine learning. Finally, there will be challenges, guidelines for the future and conclusions. This paper also addresses the open-source platform Apache Spark. Spark is an incubator-status Apache cluster computing framework which is designed to accelerate data analysis, run fast programs and write data. Spark supports the in-memory processing mechanism that allows data to be queried much faster than disk-based drives like Hadoop, as well as general model execution that optimizes arbitrary operator graphs. [14].

2 Big Data

The word big data refers to technical advancement and implementation to give the right person the right information from a great number of data that has grown gradually over a long period of time in our society.

Nevertheless, there are also several other apps. Big data with a 3V platform was described by Doug Laney (Gartner). This addressed length, pace and variety increase. In Apache Hadoop (2010), "Big data" were described as "datasets that cannot be collected, managed and processed within a reasonable range by general computers." "Big data is high volume, high speed and high information assets that require a new form of processing in order to improve decision taking, knowledge discovery and process optimisation" was redefined in 2012 by Gartner [15].

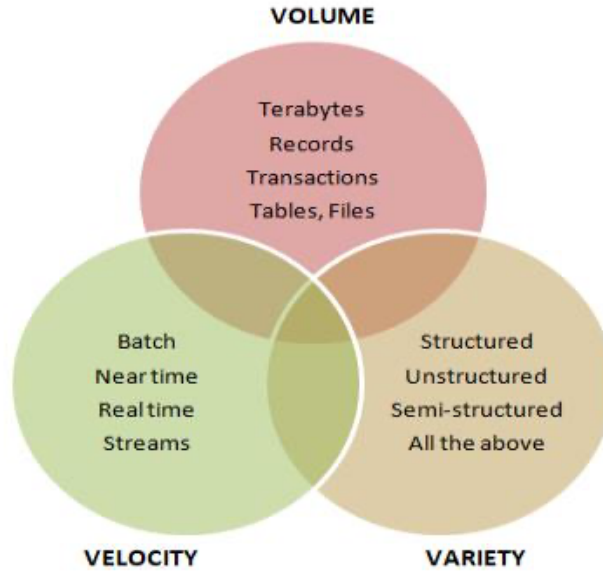


Fig. 1. 3V's in Big Data

Amir Gandomi et al. [3] specified the following for the 3V's:

- **Volume** is indicated by density of the data. Big sizes in terabyte or in petabyte or exabyte are usually huge. Beaver Doug et al. [5] Facebook actually holds some 2.8 trillion petabytes of data, storing more than one million images a second.
- **Variety** in a dataset this applies to the diversity of structures. During technological development, we may use different data types which have different formats. These types of data are voice, sound, text, images, log files etc., big data is found by three parties. The data are organized, unstructured and semicircular. The figure below shows it [15].
- **Velocity** is the speed of generation and processing of data is the time. As previously stated, since the introduction of digital devices, such as smartphones and sensors, we have developed these data formats in an unparalleled way.

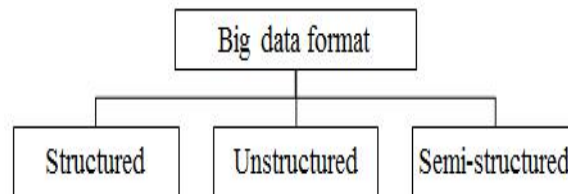


Fig. 2. Types of Big Data

Time refers to the pace of the creation and analysis of data. As described, digital devices like smartphones and sensors are coming up, we have produced these data formats in an unprecedented way.

2.1 Use cases of big data

Large hospital data refers to patient data in the individual health care facilities, such as medical records, Lab results, X-ray reports, case histories, diets, the physicians' and nurses' lists in a specific hospital. Health systems rely on big emerging technologies to collect all this patient details in order to have a more detailed view of care delivery and result-based payment models, management of health care and patient engagement [17].

2.2 Big data analytics

The Big Data Analytics is used to capture, organize, analyze and evaluate massive data sets to identify various patterns and other useful information. Big data analysis is a range of technology and techniques that involve new ways of integration to expose massive, bigger, more complicated and huge-scale, secret figures from large data sets. Relative to the complexity and speed of the data being processed, the analysis of relational database is changing.

Big data analysis is characterized as the extraction by large quantities of stored data of relevant information and useful insights. The main purpose of these analyzes is to promote researchers' decision taking, such as presenting dashboards, graphs or operational reporting for threshold and IPC monitoring. This includes the use of data interpretation and scenarios using analytical and statistical approaches, validating theories and making accurate estimates of possible events. Data mining is perhaps a fundamental concept of Big Data Analytics to analyse and explore large data systems of order to identify valid and useful data patterns [9].

2.3 Types of big data analytics

Big Data Analytics is primarily used in 4 research categories: descriptive analysis, predictive analysis, prescriptive analytics and diagnostic analysis. Descriptive assessments make the data collected meaningful information, through statistic graphic resources such as maps, diagrams, bar diagrams and dashboards, for analyzing, reporting, tracking and visualization purposes. Predictive analytics are typically characterized as extrapolation of data for better decision making based on available information purposes. Descriptive and predictive analytics are related to prescriptive analytics [17] as shown in Table 1.

Table 1. Types of Big Data Analytics

S.No	Type of Analytics	It consists of asking the question like:	Functionality
1	Descriptive Analytics	What is happening?	For collection of historical data and give you a sense of what is going to take place in the future and the opportunities
2	Diagnostic Analytics	Why did it happen?	The root cause of a problem is examined in diagnostic analysis. It's used to figure out why something happened.
3	Predictive Analytics	What is likely to happen?	It uses data from the past to predict the future. The entire thing is prediction. Predictive analytics use many techniques such as artificial intelligence and data mining to analyze current knowledge and construct scenarios.
4	Prescriptive Analytics	What should be done?	It is committed to finding the right steps Descriptive analytical systems include historical data, which can be predicted by predictive analytical data. These parameters are used to find the best solution in prescriptive analysis.

We need Big Data Analytics to boost healthcare productivity by delivering patient-centric resources, predicting disease spread sooner, tracking hospital quality and enhancing treatment methods [17].

3 Big Data Analytics in Healthcare

3.1 Big data analytics in maintaining healthcare data

Electronic health record [4] (EHR) holds the data on individual patients in electronic health. The EHR documents a range of data on health informatics, along with race, health status, allergies as well as drugs, vaccinations, clinical test results, and private data such as fitness level. Centered on IOT, Sanjeevani EHR has been used to hold clinical data from patients [2].

3.2 Predictive analytics in healthcare

Predictive analysis has been recognized for the past two years as an important approach to business intelligence but its applications in the real world far outweigh the market sense. Different approaches, including text and multimedia analytics, form part of big data analysis. Predictive analytics, including statistical methods like data extraction and machine learning, which look at present and past to predict the future, however, is among the most important categories. Predictive approaches used to assess whether patients are at risk of re-admission today in the hospital setting. Such data will help physicians make informed decisions on patient care. Predictive Analysis includes a widely applicable comprehension and use of machine learning [4].

The prediction of future probabilities with very accurate results is based on these predictors. Machine learning and regression approaches can be separated into research methods. Predictive analysis has become popular with machine learning techniques because of their excellent success in managing large-scale data sets with consistent features and noisy results. Clinical studies demonstrated that machine education is ideal for predictive models by deleting large patterns. [5].

Predictive Analytics serves numerous life sciences and healthcare providers divisions. It seeks to reliably identify diseases, enhance patient care, maximize resources and increase clinical outcomes. Predictive Analytics help companies, by cost optimization, plan themselves for healthcare [5].

4 Machine Learning Algorithms in Healthcare

There are a lot of machine learning methods used in different fields to predict large-scale analytics. Health data analysis focuses on the systematic use of this health data base by various mathematical, Predictional and statistical models and approaches to support market analysis, decision making, preparation, learning, early disease diagnosis and disease monitoring [1].

Machine learning has recently dramatically increased computer capacity for images identification and marking, speech recognition and translation, skillful playgrounds, higher IQ, disease prediction and better decision making over data. For such machine learning applications, the goal is generally to train a computer as people or better than people. The model with marked data is used for training typically supervised learning algorithms, and then test results are used for evaluation using test data [1].

4.1 Machine learning

Machine Learning: the iconic concept is-a software program can learn from what happened of E with some task class T and output measure P even if progress in T tasks, as measured by P, increases with E. Machine learning is an artificial intelligence division that uses a range of mathematical, predictive and performance analysis that processors can benefit again from previous example and identify patterns from enormous, noisy or complex information sets that are difficult to discern. Machine Learning is a method of data analysis that integrates the development of a proposed method. Machine learning enables computers for discovering secret insights without being explicitly configured through procedures that learn from data [5].

Machine learning is an artificial intelligence branch in computer science that often uses analytical tools to allow computers to "read" data without being specifically programmed (i.e. gradually enhance performance for the specific task). The term machine learning invented by Arthur Samuel in 1959. Mechanism research examines the analysis and creation of algorithms, which you can learn and forecast data – these algorithms can be mastered by strictly static system instructions, by making information-led predictions or decisions and by creating a sample-input model. Machine learning is used for a variety of computational activities, in which it is difficult or impossible to design

and program specific algorithms with good efficiency; E-mail servers, intruders, or malicious insiders working against data infringement provide examples of applications, OCR, learning level, and computer vision are among examples for applications.

The prediction models allow clinicians, data scientists, engineers and experts "to produce accurate, repeatable decision and outcome" and to reveal "concealed insight" by observing from historical connections and data patterns [4].

4.2 Steps for applying data to machine learning

Machine learning mission can be broken down to the steps below:

1. Collection of data
2. Exploration and preparation of data
3. Training of a data model
4. Evaluating performance of the model
5. Improving performance of the model

The data should be collected in a digital form appropriate for research purposes, second stage in the cycle in machine learning requires a lot of human interaction. The third phase is critical to determine how much the proposed method has heard of similar knowledge. The fourth stage will be used to determine the performance of the model which can be tested with the validation data set. The last stage will be used if the efficiency of the model is to be enhanced, the advanced techniques must be used [5].

Once these measures are completed, the model can be used for its intended purpose if it seems to be working acceptably. The model should be used to provide forecast score data, financial data estimates, to produce correct marketing or analysis knowledge, or to perform tasks automatically. The successes and shortcomings of that same method used could also provide useful details for the preparation of the new model.

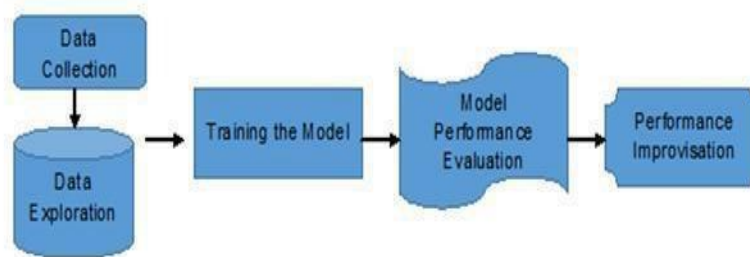


Fig. 3. Process of Machine Learning

4.3 Types of algorithms

Machine learning includes various algorithms and is categorized into three broad groups based on the nature of the learning process.

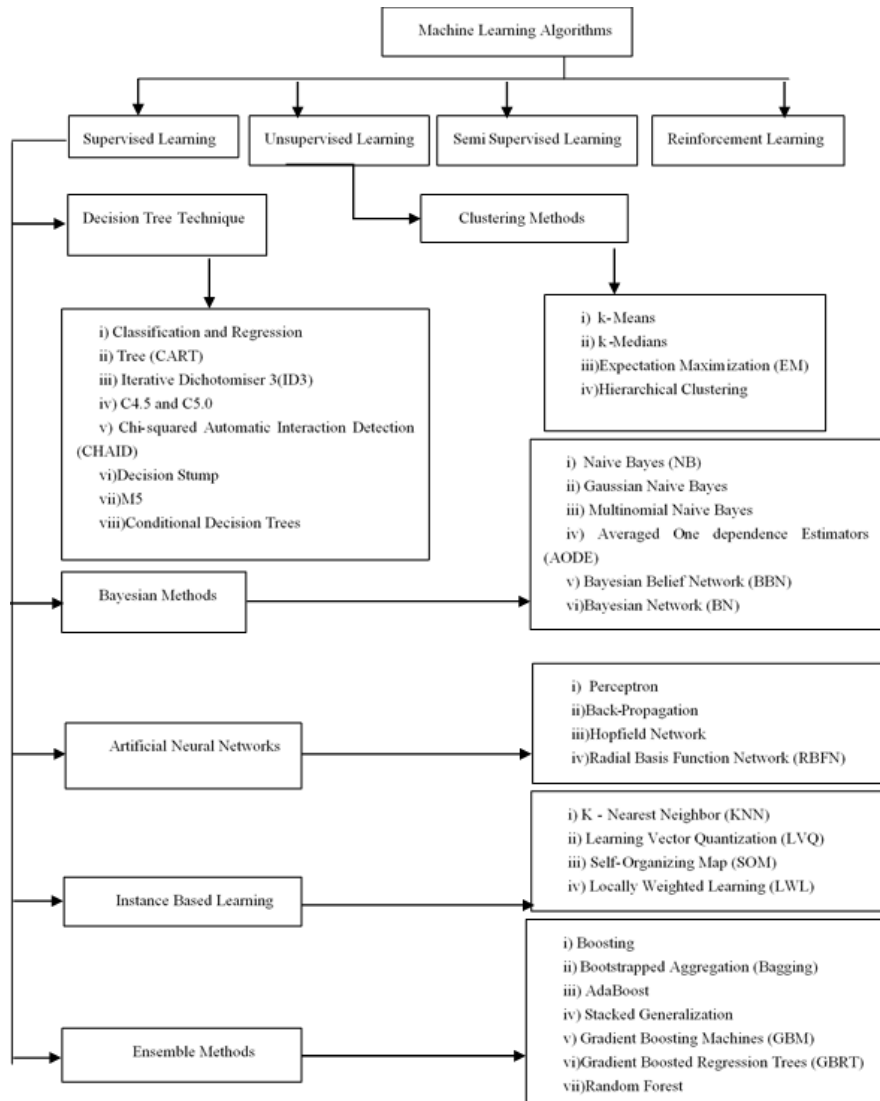


Fig. 4. Types of Machine Learning Algorithms

4.4 Comparative study of machine learning algorithms:

Comparative study of some machine learning algorithms is tabled in the Table 2.

Table 2. Comparative Study of Machine Learning Algorithms

S.No	Type of Machine Learning Algorithm	Algorithm	Functionality	Advantage	Disadvantage	Ref/Author/Year
1	Supervised Learning	Decision Tree (DT)	Decision Trees are a supervised, Univariate learning approach being used classification and regression work.	Simple, fast and gives accurate result	Traditional algorithm for the decision-making tree with a lot could be consuming.	[3] 2016 Dharavath Ramesh et al.
		Naive Bayes (NB)	Naive Bayes employs a common approach to estimate various class probabilities based on different attributes.	Naive Bayesian Classification has very significantly high reliability for large data classification.	If the attribute is independent, then only the Naive Bayesian classification gives it total precision adaptation is the entire attribute.	[3] 2016 Dharavath Ramesh et al.
		K Nearest Neighbours (KNN)	It is a classification algorithm that operates on the classification of the collected data based on the closest neighbor to the already trained model class name.	Highest accuracy for predicting data set. In the case of text mining, the result is better than the Bayesian approach.	The KNN algorithm is that is a lazy learner, i.e. it learned nothing from training samples and just uses specification of the training samples itself.	[1] 2018 Azeem Sarwar et al.
		Support Vector Machine (SVM)	In an N-dimensional area (N — the number of features) this algorithm will find a hyperplane that defines the data points	Highest accuracy for predicting data set. Seamlessly controls a broad data set and sort it into defined groups. It is the latest and most popular algorithm prediction very much faster in algorithm.	SVM performance will weakens when a large number of noisy data are available in the data set.	[3] 2016 Dharavath Ramesh et al. [1] 2018 Azeem Sarwar et al.
		Neural Network (NN)	The Neural Network is indeed an adaptive learning model that changes the weight of the neural connections.	Neural Network model gives a accurate results.	It takes a great deal of time for computing massive data sets to deal with big data. And with the rise in data size, the memory demand for the model increases.	[3] 2016 Dharavath Ramesh et al.
2	Un-Supervised Learning	Partitioned Clustering	The entire set of data is divided into a predetermined number of bits, ensuring that	In K-Means The distance or similitude measure	Clusters of different size and different density (in the source data)	[3] 2016 Dharavath Ramesh et al.

			no one else is null and each point is part of only one cluster.	between the cluster center and the data point is determined. K-Me-doids the me-doids of the clusters are used in place of centroid.	may not perform effectively.	
		Hierarchic al Clustering		This method is beneficial with regard to big data, since it can com-pactly com-prise the rele-vant data by establishing appropriate groups with nearly no data loss.	Agglomerative: This cannot be re-versed until a combination of two clusters is completed. Split-ting: It is also the downside of not allowing one clus-ter to be further merged until split into two.	[3] 2016 Dharavath Ramesh et al

4.5 Analysis of different machine algorithms in healthcare sector

Various algorithms for machine learning have been used to analyze health data in this empirical analysis. Table 1 shows the comparative study of various algorithms in machine learning.

Azeem Sarwar and Nasir Kamal in [1] presented a Prediction of Diabetes Using Machine Learning Algorithms in Healthcare. Six machine learning algorithms were used by the authors in the national institutes of diabetes, digestive and renal diseases. SVM, KNN, LR, DT, RF and NB are the following algorithms. Diabetes predictions on the Indian PIMA dataset have been made. The SVM and KNN have shown to be highly accurate for diabetes prediction. These two algorithms give an accuracy of 77 percent, which is higher than the other four. SVM and KNN can be inferred to be ideal for predicting the disease.

P. Saranya and Dr. P. Asha in [2] presented a Lung cancer, which It is a lethal disease which has an elevated mortality rate, and it raises the need for disease prediction. Hybrid vector support system K means, profound preparation, supervised learning models and fusion are other algorithms that are commonly used to improve accuracy and sensitivity.

In [3], presented Some research in the field of healthcare related to machine learning such as the Artificial Network was used for classification of chest conditions. Diagnosis was used for chronic illness diagnosis. The conventional approach to filtering was being modified to quickly pursue data high-dimensional.

In [5], discussed throughout the study of medical data sets for cardiovascular prediction, techniques Naive Bayes, Neural network and Decision Tree algorithms are used. A theoretical study has introduced machine learning techniques such as the decision tree (C4.5), the SVM support system (Support Vector Machine) and the Artificial

Neural Network (ANN). Only three algorithms were employed but the author concluded that the SVM classification model predicted recurrence with lower error and higher accuracy.

Auto Regressive (AR) Modeling, a series of extensive feature sets has been generated recently. The Hybrid Firefly and Particle Swarm Optimization (FFPSO) is used explicitly to modify the raw ECG signal, not to extract characteristics via the extraction procedures [17].

Support Vector Machine is a classification supervised machine learning process. This is an example for a binary linear classifier based on Statistical learning theory. This has high accuracy and is able to handle sequences of small dimensions [18].

The mining algorithm association rule often leads to useless policies. The values determined by the assistance as well as the trust value must be optimized before classification to prevent this form of concern. And the real defects can be identified by using the Artificial Neural Network (ANN) [19].

Software based on components can be helpful to improve software quality and to analyze efficiency. Neural networks were an active research area relevant for various fields and a small number attempted to find suitable reusable components by integrating neural networks with software engineering [20].

The confidentiality and protection of user data while creating such communications. Many inspectors have developed several sophisticated algorithms for user protection [21].

Machine learning is something like data mining [4], which both seek pattern data. This data is used in machine learning to better understand the program than to collect data based on human understanding, as in data mining. Computer analysis detects data patterns and accordingly modifies the function of the system [16].

Healthcare still is in the early stages of taking advantage of the new possibilities presented by Big Data and making successful decisions. In order to predict the accurate results in health care domain we have to make good use of all the above traditional machine learning algorithms.

5 Tools Used to Analyze Healthcare Data-Apache Spark

Apache Spark is a substitute to open source Hadoop. It is a holistic data processing tool that includes libraries with a higher level of support for SQL query (Spark SQL), Streaming Data, Machine Learning (MLlib) and Graph Processing (GraphX) – distributed data processing. These libraries help to increase the efficiency of developers, as the application server involves fewer scripting and that can be easily incorporated into more advanced forms of computing. Storage data management which can make Spark quicker and easier than Hadoop in cross-pass analysis (on small datasets) is facilitated either by implementation of Resilient distributed data sets (RDDs). This applies more if the data is smaller than the memory available. This shows that it would take a huge amount of memory to process real big data with Apache Spark. Since the cost of memory is greater than the value of the hard drive, MapReduce is considered to be

somewhat economical than Apache Spark for larger data sets. This also offers decent consistency and integrated-in versatility for large-scale data processing [10].

1. **Spark**: for an in-memory computing group, built through its sophisticated libraries for rapid data computing
2. **Spark streaming**: enables in real time data collection and analysis enables data to be obtained and analysed in real time
3. **Spark SQL**: will perform queries for SQL
4. **Spark MLlib**: is a useful and efficient library to solve problems related to machine learning. It includes various classification, regression, clustering and optimisation algorithms for machine learning
5. **Spark R**: It does provide the same computer language capabilities for numerical computation as well as 'R' visuals. However, its distributed architecture reduces time consumption
6. **Spark GraphX**: is devoted for graphical analysis.

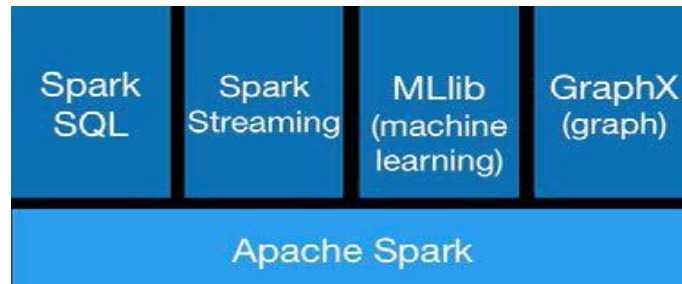


Fig.5. Apache Spark

5.1 Analyzing healthcare data using apache spark

For healthcare applications Apache Spark may be used for tracking remote health, diagnostic aids, medication comparisons, retrieval of information, clinical decision assistance systems, for searching patients for similarities, for predicting and visualizing personalizing public health surveillance outcomes for personalized medicine. Furthermore, the Spark Machine Learning Library can also perform large-scale classification, clustering, predictive modeling and associative rule mining. Effectively managed and analyzed with Spark's streaming library, real-time streaming data from wearable sensors and IoT-enabled computers. Spark's computer technology would be most critically suitable for centralized health and real-time monitoring to reduce massive health-care costs [11].

SparkML is used in structured and unstructured information retrieval for the recommendation and diagnosis. Spark's Vector Support Module, Random Forests and K-Means cluster libraries are used in large-scale classifications. However, SparkML is used to measure, warn, inform and monitor the health. Information is obtained via Spark SQL and SPARQL, where Spark GraphX functions as a visual analysis of the queried data.

6 Conclusion

In this paper we presented a brief overview of Big Data and the features and forms of Big Data Analytics that play an important role and affect the healthcare system. Within this paper we have also suggested comparative analysis of algorithms for machine learning. In order to predict the accurate results in health care domain we have to make good use of all the traditional machine learning algorithms. The difference between the traditional machine learning models was that it's not versatile, but that too long for massive data sets in certain cases. We must therefore change the algorithm or adjust it to make it appropriate for a data management crisis. To overcome this problem Apache Spark will be used which is an alternate to Hadoop open source. It's a unified data processing engine which includes libraries with a higher standard to support SQL query (Spark SQL), Streaming Data, Machine Learning (MLlib) and Graph Processing (GraphX) – distributed data processing the Apache Spark Machine Learning Library can also perform large-scale classification, clustering, predictive modeling and associative rule mining.

7 References

- [1] Muhammad Azeem Sarwar, Nasir Kamal, Wajeeda Hamid and Munam Ali Shah, "Prediction of Diabetes Using Machine Learning Algorithms in Healthcare" in proceedings of the 24th International Conference on Automation & Computing, Newcastle University, Newcastle upon Tyne, UK, 6-7 September 2018. <https://doi.org/10.23919/iconac.2018.8748992>
- [2] P.Saranya, Dr.P.Asha, "Survey on Big Data Analytics in Health Care" in Second International Conference on Smart Systems and Inventive Technology (ICSSIT 2019) IEEE Xplore Part Number: CFP19P17-ART; ISBN:978-1-7281-2119-2. <https://doi.org/10.1109/ics-sit46314.2019.8987882>
- [3] Dharavath Ramesh, Member, IEEE, Pranshu Suraj, and Lokendra Saini, "Big data Analytics in Healthcare: A Survey Approach" in IEEE transactions 978-1-4673-6621-2/16/\$31.00 © 2016 IEEE.
- [4] Sunil Kumar and Maninder Singh, "Big Data Analytics for Health care Industry: Impact, Applications, and Tools" in Big Data Mining and Analytics ISSN222096-0654 1105/0611pp48–57 Volume 2, Number 1, March 2019 DOI: 10.26599/BDMA.2018.9020031.
- [5] B. Nithya and Dr. V. Ilango, "Predictive Analytics in Health Care Using Machine Learning Tools and Techniques " in International Conference on Intelligent Computing and Control Systems ICICCS 2017 ;978-1-5386-2745-7/17/\$31.00 ©2017 IEEE. <https://doi.org/10.1109/iccons.2017.8250771>
- [6] Ying Yu, Min Li, Liangliang Liu, Yaohang Li, Jianxin Wang, " Clinical big data and deep learning: Applications, challenges, and future outlooks" in. <https://doi.org/10.26599/bdma.2019.9020007>
- [7] Pariwat Ongsulee , Veena Chotchaung, Eak Bamrunsi and Thanaporn Rodcheewit, " Big Data, Predictive Analytics and Machine Learning" in Sixteenth International Conference on ICT and Knowledge Engineering 2018. <https://doi.org/10.1109/ictke.2018.8612393>
- [8] Nambiar Jyothi Ravindran and Dr. Prakash Gopalakrishnan, " Predictive Analysis for Healthcare Sector Using Big Data Technology" in IEEE transactions 978-1-5386-5657-0/18/\$31.00 c 2018IEEE.

- [9] Safa bahri, Nesrine Zoghliami, Mourad Abed, and João Manuel R.S. Tavares, "Big Data for Healthcare: A Survey" in IEEE Access. January 2019 DOI: 10.1109/ACCESS.2018.2889180. <https://doi.org/10.1109/access.2018.2889180>
- [10] Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma and Sandeep Kaushik, "Big data in healthcare: management, analysis and future prospects" in the Journal of Big Data available at. <https://doi.org/10.1186/s40537-019-0217-0>
- [11] Md. Rezaul Karim, Ratnesh Sahay and Dietrich Rebholz-Schuhmann, "A Scalable, Secure and Realtime Healthcare Analytics Framework with Apache Spark" in Conference Paper October 2015 DOI: 10.13140/RG.2.1.3331.7208.
- [12] Venkatesh Palaniswamy and Ramkumar Tirunakavukarasu, "Implications of big data analytics in developing healthcare frameworks- A Review" in Journal of King Saud University - Computer and Information Sciences Volume 31, Issue 4, October 2019, Pages 415-425 available at. <https://doi.org/10.1016/j.jksuci.2017.12.007>
- [13] Abdul Ghaffar Shoro and Tariq Rahim Soomro, "Big Data Analysis: An Spark Perspective" in Global Journal of Computer Science and Technology: C Software & Data Engineering Volume 15 Issue 1 Version 1.0 Year 2015 ISSN: 0975-4172 & Print ISSN: 0975-4350.
- [14] Gemson Andrew Ebenezer J. and Durga S, "Big Data Analytics in Healthcare: A Survey" in ARPN Journal of Engineering and Applied Sciences VOL. 10, NO. 8, MAY 2015 ISSN 1819-6608.
- [15] Youssra Riahi, Sara Riahi, "Big Data and Big Data Analytics: Concepts, Types and Technologies" in International Journal of Research and Engineering ISSN: 2348-7860 (O) | 2348-7852 (P) | Vol. 5 No. 9 | September-October 2018 | PP. 524-528 <https://doi.org/10.21276/ijre.2018.5.9.5>
- [16] J.Archenaa and E.A.Mary Anita, "A Survey Of Big Data Analytics in Healthcare and Government "in 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15) <https://doi.org/10.1016/j.procs.2015.04.021>
- [17] Kora, P, "ECG based Myocardial Infarction detection using Hybrid Firefly Algorithm " in Computer Methods and Programs in Biomedicine, 2017, 152, pp. 141-148. 2-s2.0-85030103123. <https://doi.org/10.1016/j.cmpb.2017.09.015>
- [18] Padmavathi, K., Krishna, K.S.R., "Myocardial infarction detection using magnitude squared coherence and Support Vector Machine" in International Conference on Medical Imaging, m-Health and Emerging Communication Systems, Med Com 2014, art. no. 7006037, pp. 382-385. <https://doi.org/10.1109/medcom.2014.7006037>
- [19] Dhanalaxmi, B., Apparao Naidu, G., Anuradha, K., "Adaptive PSO based association rule mining technique for software defect classification using ANN " in Procedia Computer Science, 2015, 46, pp. 432-442. <https://doi.org/10.1016/j.procs.2015.02.041>
- [20] Krishna Chythanya, N., Rajamani, L., "Neural Network Approach for Reusable Component Handling " Proceedings - 7th IEEE International Advanced Computing Conference, IACC 2017, art. no. 7976764, pp. 75- 79. <https://doi.org/10.1109/iacc.2017.0030>
- [21] Gayathri, P., Umar, S., Sridevi, G., Bashwanth, N., Srikanth, R. Hybrid cryptography for random-key generation based on ECC algorithm (2017) International Journal of Electrical and Computer Engineering, 7 (3), pp. 1293-1298. <https://doi.org/10.11591/ijece.v7i3.pp1293-1298>

8 Authors

Mr. Sai Hanuman Akundi, Ms. Soujanya R, Ms. Madhuri PM all authors are from Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India.

Article submitted 2020-09-16. Resubmitted 2020-10-10. Final acceptance 2020-10-11. Final version published as submitted by the authors.