Journal of Harbin Engineering University
ISSN: 1006-7043
Vol 44 No. 7
July 2023

A Comprehensive survey of big data Analytics and State-of-the-art Chronic Kidney Disease Prediction Systems in the health care

P. Ravikumaran[1] ,

Assistant Professor, Dept. of Computer Science & Engineering., Fatima Michael College of Engg & Tech, Madurai- 625020, Tamil Nadu, India

K. Vimala Devi[2],

AssociateProfessor, School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Vellore Campus, Vellore- 632014, India

K. Valarmathi[3],

Professor, Dept of Electronics & Communication Engineering, P.S.R Engineering College, Sivakasi- 626140, Tamil Nadu, India

Abstract- The quantum of data generated in the health care sector has raised a lot over the past few years. That's when Big Data Analytics oriented methodologies have come in handy. Furthermore, the application of deep leaning concepts like Deep Belief Networks (DBN) is found to be much beneficial, especially in predicting the risks associated with chronic condition diseases. So, we make a comprehensive survey of the state-of-the-art Chronic Kidney Disease Prediction Systems in the health care and deduce our outcomes,

Keywords: Big Data Analytics, DBN, Chronic Kidney Disease, and chronic condition.

I. INTRODUCTION

Smarter big data handling is much needed in this developed age that we live today. As the measure of data increases, the architectures and technologies needed for handling those vast data also need to get improvised day by day [1]. As these architectures and technologies evolve, the data extraction process becomes smoother and efficient. According to [2], the following are the properties associated with big data:

- Volume
- Complexity
- Variety
- Variability
- Value
- Velocity

Owing to the advantages that the big data handling gained in various sectors, its deployment has increased considerably.

This section is then sub-divided into two as follows:

A. Main Contributions of survey

With our survey made in this paper, the following are the objectives:

- To realize the importance of handling enormous amount of data by using big data analytics.
- To pinpoint and investigate the role and growth of big data analytics in the arena of healthcare.
- To compare and contrast the predicting systems pertaining to the chronic kidney disease conditions.
- To finally deduce the need for effective chronic kidney disease prediction systems.

B. Pictorial representation of survey

The sections in our comprehensively made survey have been clearly indicated in the following fig. 1:

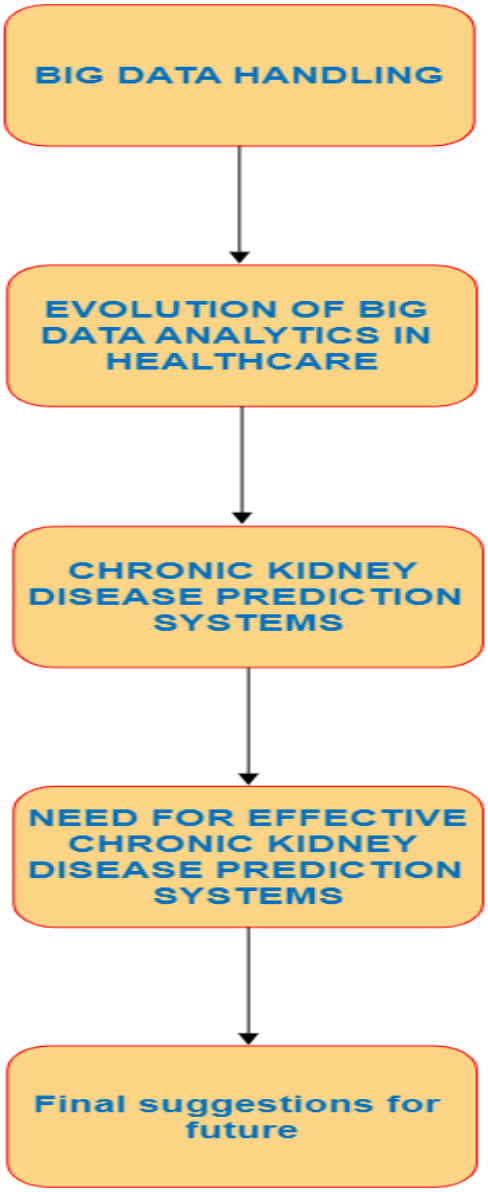

Fig. 1 Pictorial representation of sections inclucated in our survey

II. BIG DATA HANDLING

This section will be sub-divided into three as follows:

A. Evolution of Cloud Methodologies

As the digital transformations are getting increased day by day, various data might necessarily have to be handled in cloud methodologies, which is solely the reason for efficient data handling. But these cloud methodologies could never exist without any technical as well as non-technical constraints. As a result, the evolution of these cloud methodologies becomes essential as the scope for developmental rises up. Some of those evolving cloud methodologies are discussed below.

An education sector-aware cloud methodologies were investigated by [3] to spot improvisations in the cloud tech utilizations, interaction intermediary to the Human and computer, and big data in various engineering-based universities and colleges. This study was not only a theoretical study, but also a chronological study investigating about the necessity of deploying diversified cloud methodologies depending upon the varying needs.

[4] studied about the various factors contributing to the efficient data handling pertaining to the placement methods and data storage. They deduced the areas of improvement or evolution that the cloud technologies should make to contribute towards the efficient data handling.

Likewise, many such works have considered the importance of cloud technological evolution. Those works are briefed below.

- Cloud tech as well as big data specific challenges along with improvisation areas were discussed [5].
- The positive role of Cloud tech as well as big data leading to towards the development of Organization were studied [6].
- The evolvement of cloud tech with respect to its Data integrity verification was comprehensively studied with the help of outsourced vast quantity of data [7].
- Evolved version of cloud technology was proposed and deployed for the sake of data storage and forwarding [8].
- Remote sensing specific Google Earth-based cloud tech towards the investigation of vast quantity of data was studied [9].
- Another evolved version of cloud tech was proposed to improvise the data quality [10].
- Modified cloudlets-oriented cloud tech was proposed and experimented with Health care sector [11].
- Scheduling schema was devised related to the task in the cloud based on the dynamic batch mode cost saving [12].

B. Secured Handling

As the quantity of data gets bigger, the security of data being handled also becomes vulnerable and thus the quality of the extracted information or features also gets reduced. Thus, the secured handling of data becomes indispensable. In particular, data processed by the cloud technology should also be secured as there will be a greater number of devices that are connected with each other. Various instances of works concentrating on the security aspect of the cloud handled methodologies are being discussed in the following tabulation in table 1:

Table 1 Summary of secured cloud-oriented methodologies

Works	**Methodology/ Consideration**	**Inference/ Contribution**
[13]	RBSEE- Request-based, secured and energy efficient method for big handling under IoT environment.	Security of data getting handled between the source and destination was ensured by the utilization of "Twofish" cryptographic method.
[14]	DDDG- Duplicate Divergence-Different characteristics powered dragon Genetic method.	The secureness of the data getting handled was achieved with the deployment of dragon operator-oriented genetic method and k-DDD anonymization to ensure higher accuracy and lower data loss.
[15]	Watermarking and Encryption powered security aware system was proposed for the sake of cloud computation of health care data.	Raspberry Pi platform and hard disc was deployed towards the effective hardware management to achieve the data secureness by yielding favourable robustness against various attacks.
[16]	Associated big data classification with secureness towards a better mobility cloud data.	Associated classification cum secureness of mobility big data would be done even before the execution of data investigation,

		mobility, and duplication. The data gets categorized into two risks criteria namely, public and secret data.
[17]	Risk minimizing protected data sharing rules set were defined both in big data and cloud computing.	The importance of cloud computing in mobile means was deduced with reference to the limitations like Availability, Changes, Compliance, and Privacy.
[18]	Intelligent and secured Medi care alerting as well as monitoring structure for big data and IoT system.	Intelligent and secured Medi care system was formulated with the deployment of GC- Grouping cum choosing method as well as MF-R-Meta Fog-Redirection along with MapReduce for forecasting the heart-related ailments. Better desired outcomes in terms of sensitivity, throughput, f-measure, and accuracy were achieved.
[19]	Multi cloud-oriented hybridized encryption methodology towards the secured storage for big data.	The data would get secured even before it gets stored in the multi-cloud and prevents the insider vulnerabilities by executing operations like data uploading, indexing, slicing, distribution, encryption, decryption, merging, and receiving.

Fig. 2 Risks tackled when securing the cloud data [20]

The above fig. 2 shows the risks that get tackled by securing the cloud-based data.

C. Big Data in Diseases Prediction

As the diseases prediction involves the vast quantity of data in it, the early or accurate prediction of any diseases is only possible with the efficient handling of those vast data. This disease prediction is not only prevalent for non-chronic conditions, but also valid for chronic disease like heart attack, kidney failure, diabetes, etc.

Determining the pathological characteristics that contribute to the chronic diseases should be done in order to avoid ending up in the chronic or more severe health conditions, just like identifying

the interventions of drug to be saver for sudden heart failure [21].

Even the neonatal diseases could be forecasted and prevented if the vast quantity of data taken in a NICU- Neonatal Intensive Care Unit could be handled efficiently [22]. Likewise, many similar chronic/ severe medical conditions could be treated or prevented by proper handling of the vast quantity of patient's data.

III. Evolution of Big Data Analytics In Healthcare

As the application of big data analytics in health care has become more prevalent because of increasing diseases and improper lifestyle. As a matter of fact, even the lifestyle of sportsmen is being monitored so as to maintain the health in good shape [23]. Thus, we give a comprehensive review of the big data analytics being deployed in the healthcare as follows:

A deep transfer learning-oriented myocardial infarction forecasting methodology was proposed by [24] with special consideration given to the non-rural/ city heath care systems. As the most efficient tool for diagnosing myocardial infarction was ECG, they deduced the computer dependent diagnosing system to depict the risks of a myocardial infarction by analyzing the images of ECG by deploying VGG-MI1 and VGG-MI2. They were able to achieve better outcomes in terms of 99.22 % accuracy, 99.49 % specificity, and 99.15% sensitivity for VGG-MI2. However, this model was only void for myocardial infarction and doesn't hold good for other heart ailments.

An improved and intelligent system of wearable health monitoring system was proposed by [25] by deploying ontologies and data mining techniques along with Bi-LSTM- Bidirectional Long Short-Term Memory. This monitoring system was able to forecast any after side effects of drugs taken related to the ailments like mental state of mind, Blood Pressure, and Diabetes. The accuracy yielded by this monitoring system was claimed to be favorable. But this method had limitations like deploying SentiWordNet and non-multi-modal data.

A big data focused comprehensive investigation was done by [26] by keeping in mind the computational intelligence methods from the perspective of interconnected machines and customized devices. They made use of HSTSM- Hierarchical Spatial-Temporal State Machine in their comprehensive investigation so as to sort out the several limitations taking place in a computational operation.

An Innovative three staged IoT-based methodology for early diagnosis of heart related ailments was devised by [27] with the deployment of Apache HBase and Apache Mahout in stages 2 and 3 respectively. The Apache HBase in stage 2 was utilized for accumulating the vast data collected wearables and Apache Mahout in stage 3 used for logistic regression-oriented forecasting of heart related ailments. Finally, ROC was plotted to identify the best clinical characteristics to receive the heart ailment. However, the energy conservative node choosing could have been done.

A SMSI- Social Media Search Indexes of certain pre-defined factors was deployed by [28] to forecast the count of COVID-19 cases reported from December 2019 to February 2020. Out of the six methods deployed for sorting out the overfitting issue, subset selection was the most efficient one. This SMSI-based outcomes were found to provide an early caution to all pertaining to the future COVID-19 cases.

Another instant of COVID-19 work was proposed by [29] in order to pose a fight against this deadly disease by using the concept of IoT. They concluded that the intelligent health care monitoring methods would be effective as the digital transition of health records were high.

A method of CloudDTH- Digital Twin Health maintaining was proposed by [30] towards the personalized health maintenance of elderly human beings. Their methodology was able to execute operations like diagnosing, monitoring, and forecast the health well-being of the elderly people. Finally, the application perspective cases studies were made. But the accuracy of the model was not favorable enough.

Another instant of work considering the personalized health care system was proposed by [31] by prioritizing the data rather than the method itself. They believed that only the un-biased data could possibly be aiding the personalized health care system.

Various other works dealing with big data analytics in health care for the health well-being of human beings are further discussed below.

- The existence of multiple disease-aware system in health care was comprehensively reviewed by [32] correlating between the big data and AI-Artificial Intelligence.
- An improvised and modified large scale classification in healthcare was achieved by [33] by deploying Optimized Gabor Feature Extraction process.
- The typical Medi data was characterized and reviewed by [34] to deduce the benefits, limitations, and further course of action whenever dealing with vast quantum of data.
- Another instant of big data-oriented management and investigation along with futuristic benefits were reviewed comprehensively by [35].
- A survey perspective investigation considering 224 answers was made by [36] to know about the suitability of big data in healthcare.
- The relationship of automated system enabled decision making and customized services in health care was reviewed and suggested by [37].
- The inter-relationship intermediary to both the big data analytics and health care sector was reviewed by [38] with reference to its limitations and diversified applications.
- A standard operative review methodology was followed by [39] with the consideration of 41 earlier findings to deduce their perspectives.
- The role of big data analytics in the drug developmental purpose was comprehensively reviewed by [40] to contribute towards the precision medicine.
- The inter-relationships between IoT, cloud computing, and big data was reviewed by [41] with reference to the health care and Industry 4.0 (4^{th} industrial development).
- Another instant of standard operative review methodology was adhered by [42] with the consideration of 29 most crucial earlier findings to deduce their perspectives.

Now, let's see the big data analytics in the arena of healthcare from the below table 2

Table 2 Summary of big data analytics in healthcare

Works	**Methodology/ Consideration**	**Inference/ Contribution**
[23]	Lifestyle of sportsmen.	Maintenance of good health.
[24]	Deep transfer learning-oriented myocardial infarction forecasting methodology.	Deduced the computer dependent diagnosing system to depict the risks of a myocardial infarction by best model VGG-MI2.

[25]	Ontologies and data mining techniques along with Bi-LSTM- Bidirectional Long Short-Term Memory.	Forecast any after side effects of drugs taken related to the ailments like mental state of mind, Blood Pressure, and Diabetes.
[26]	HSTSM.	Sort out the several limitations taking place in a computational operation.
[27]	Innovative three staged IoT-based methodology.	Logistic regression-oriented forecasting of heart related ailments with best clinical characteristics.
[28]	SMSI.	Forecasting the count of COVID-19 cases reported from December 2019 to February 2020 with efficient subset selection.
[29]	Fight against COVID-19 by using the concept of IoT.	Intelligent health care monitoring methods would be effective because of higher transition rates.
[30]	CloudDTH.	The application perspective cases studies were made with less accuracy.
[31]	Prioritization of data rather than the method.	Un-biased data could possibly be aiding the personalized health care system
[32]	Existence of multiple disease-aware system.	Correlated between the big data and AI-Artificial Intelligence.
[33]	Optimized Gabor Feature Extraction process.	Improvised large scale classification in healthcare.
[34]	Typical Medi data.	Deduced the benefits, limitations, and further course of action whenever dealing with vast quantum of data.

[35]	Big data-oriented management and investigation.	Futuristic benefits were reviewed comprehensively.
[36]	Survey perspective investigation considering 224 answers	The suitability of big data in healthcare is known.
[37]	Relationship of decision making and customized services.	Relationships suggested successfully.
[38]	Limitations and diversified applications of big data analytics and health.	Inter-relationship intermediary to both the big data analytics and health care sector.
[39]	A standard operative investigation considering 41 earlier findings.	Deduced various perspectives.
[40]	Precision Medicine.	The role of big data analytics in the drug developmental purpose was comprehensively reviewed.
[41]	Comprehensive review with reference to the health care and Industry 4.0.	Inter-relationships between IoT, cloud computing, and big data was reviewed.
[42]	A standard operative investigation considering 29 most crucial earlier findings.	Deduced various perspectives.

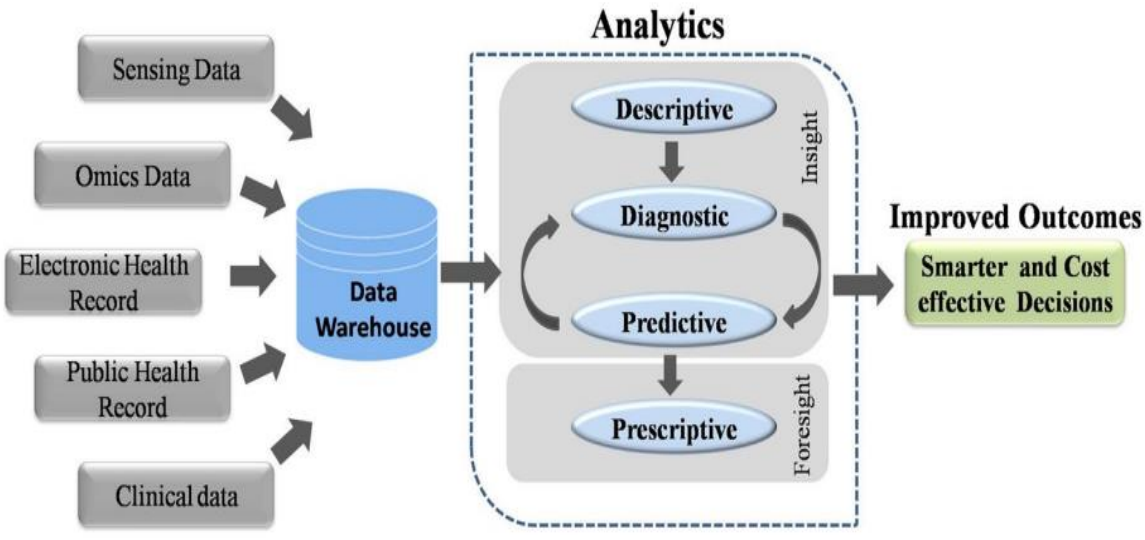

Fig. 3 Working architecture of big data in cost effective healthcare [43]

The above fig. 3 shows the typical working architecture of big data in cost effective sense in the healthcare.

IV. CHRONIC KIDNEY DISEASE PREDICTION SYSTEMS

Now a days, even the kidney problems in the elderly have started increasing. The danger of older people getting chronic/ severe kidney diseases have also increased. So, there is an indispensable need for any medical professionals to diagnose the kidney problems well in advance in order to avoid the chronic level kidney ailment. Now, we are going to review the related works carried towards the CKD- Chronic Kidney Disease in the below.

The relationship between the total volume of the kidney and Kidney developmental rate causing the Chronic Kidney Disease was identified and studied by [44] on the basis of an earlier radiological study consisting of 241 patients. According to them, if both of these parameters could be monitored, it would serve as a bio-marker for chronic kidney disease.

[45] studied the clinically important forecasting of the rate of mortality with both the Cardio-vascular and chronic kidney diseases. GFR- Glomerular Filtration Rate was determined with the help of 4,40,526 patients in order to aid the forecasting of mortality rate because of Cardio-vascular disease.

Both the acute level and chronic level kidney injuries/ diseases were studied by [46] to identify the bio-markers to prevent and treat these kidney diseases for avoiding the kidney disease in future. The pathophysiological relationship between the healthier tubular of kidney along with its various bio-markers were identified through this comprehensive review investigation.

A Temporal modified gradient boost machine was devised by [47] towards the Longitudinal risk forecasting of the chronic level condition in kidney in patients, who were suffering from diabetes diseases. They utilized data of 14,039 adults, who are suffering type 2 diabetic condition. Higher outcome yielded was ROC of 0.83. However, the pre-defined exclusions in the work could result in wrong false positive values being estimated for diabetic patients with kidney problems, who haven't screened for diabetes readings. Some of the other works which have considered the chronic kidney diseases will be delt below.

- Chronic level of kidney damage was studied by [48] with reference to the Pathological scoring of Fibrosis with the help of six deep learning processes.
- Kidney functionality forecasting model was devised by [49] with the help of inputs of ultrasound imaging by using the transfer learning-based ResNet method.
- Forecasting of Metabolic ailment was done by [50] with the help of VAI- Visceral Adiposity Index and LAP- Lipid Accumulation Product in patients who have chronic level of kidney disease.
- [51] identified the association between the chronic level of kidney disease and COVID-19 by their brief review.
- A cardiac dependent prediction methodology was devised by [52] to find the association existing between each of the following: coronary flow reserve; mortality in case of cardiovascular risks, and chronic kidney disease.
- [53] utilized the SVM- Support Vector Machine categorical neural networks to classify between the chronic and non-chronic conditions in the kidney diseases.
- Cases-wise reasoning and neural networks were deployed by [54] to provide the considerable forecasting of the chronic level of kidney disease with the consideration given to the population of the country Columbia during the period of 2018.
- Practical data-aware prediction model was developed by [55] to forecast the chronic level of kidney diseases in the patients, who were suffering from the diabetes diseases.
- A comprehensive comparative study considering the SVM categorical neural networks toward the forecasting of the chronic kidney disease [56].
- Diversified machine learning and statistical processes were deployed by [57] to forecast the rate of severeness towards the chronic kidney disease.
- DFS- Density-aware Feature Selection method along with ACO- Ant Colony

Optimization was deployed by the [58] in three stages for the sake of smartly predicting the chronic kidney diseases.

- Kidney specific inflammations-aware forecasting methodologies were reviewed comprehensively by [59] in order to predict the chronic level of kidney disease.

Now, let's see the summary of chronic kidney disease prediction systems in the following table 3:

Table 3 Summary of Chronic Kidney Disease Prediction Systems

Works	**Methodology/ Consideration**	**Inference/ Contribution**
[44]	Identifying the relationship between the total volume of the kidney and Kidney developmental rate causing the Chronic Kidney Disease.	If both of the parameters could be monitored, it would serve as a bio-marker for chronic kidney disease.
[45]	Clinically important forecasting of the rate of mortality with both the Cardio-vascular and chronic kidney diseases.	Aided the forecasting of mortality rate because of Cardio-vascular disease.
[46]	Identifying the bio-markers to prevent and treat these kidney diseases for avoiding the kidney disease in future.	The pathophysiological relationships were identified.
[47]	Longitudinal risk forecasting of the chronic level condition in kidney in patients.	Higher outcome yielded was ROC of 0.83.
[48]	Pathological scoring of Fibrosis with the help of six deep learning processes.	Successful kidney disease forecasting model was developed.
[49]	Transfer learning-based ResNet method.	Successful kidney functionality forecasting model was developed by using the inputs of ultrasound imaging.
[50]	VAI- Visceral Adiposity Index and LAP-Lipid Accumulation Product in chronic kidney disease patients.	Forecasting of Metabolic ailment was done.
[51]	Brief review with reference to COVID-19.	Association between the chronic level of kidney disease and COVID-19 was made.

[52]	Cardiac dependent prediction methodology.	Found the association of all involved parameters.
[53]	SVM- Support Vector Machine categorical neural networks.	Successful classification made between the chronic and non-chronic conditions of kidney diseases.
[54]	Cases-wise reasoning and neural networks.	Forecasting of the chronic level of kidney disease in line with the Columbian population was made.
[55]	Practical data-aware prediction model.	Forecasted the chronic level of kidney diseases in the diabetic patients.
[56]	Comparative study with SVM categorical neural networks.	Forecasted the chronic level of kidney diseases successfully.
[57]	Diversified machine learning and statistical processes.	Successfully forecasted the rate of severeness in the chronic level of kidney diseases.
[58]	Three layered DFS method along with ACO.	Smartly prediction of the chronic kidney diseases was made.
[59]	A comprehensive review on Kidney specific inflammations-aware forecasting methodologies.	Forecasted the chronic level of kidney diseases successfully.

V. NEED FOR EFFECTIVE CHRONIC KIDNEY DISEASE PREDICTION SYSTEMS

They were many factors contributing towards the chronic level of kidney disease in human beings, however, the untreated and uncared minor issues left in the kidney could cause a situation where the patient could find themselves in a chronic condition or even more severe situation of ending up in sudden death. Thus, diagnosing and treating the kidney ailments become indispensable.

These precautionary diagnosing methodologies adopted to predict the chronic as well as non-chronic kidney conditions have evolved from time to time ranging from the adoption of machine learning methodologies and

cloud technologies to the adoption of deep learning methodologies.

Some of the areas of improvement that are needed to be tackled by these effective chronic kidney disease prediction systems including but not limited to:

- Global as well as natural trend of predicting the chronic kidney diseases are grabbing the attention owing to the higher mortality/ morbidity, by also including but limited to diabetes, retinal, and cardiac risks, etc.
- Tacking the reduced GFR rates
- Classifying methodologies should be able account for Dimensionality reduction.
- The progression of chronic level of kidney disease needs to be controlled so as to prevent parenchymal cell loss, fibrosis, and chronic inflammation, etc. to prevent the end stage kidney disease or sudden death.

VI. COMPARATIVE INVESTIGATION OF THE COMMON PARAMETERS USED BY THE EXISTING WORKS

In this section, we will pin point out a few existing works realized in the literature related to chronic disease investigation and take and compare the common parameters considered by those works in order to analyze their performance in the below table 4.

Table 4 Comparative Investigation of the Parameters used by the existing works

Works	**Common Parameters involved for chronic investigation**	**Performance discussion**
[45, 52]	Associated Mortality rates	In most scenarios, the mortality rates were realized to be more whenever the chronic conditions were left carelessly.
[25, 47, 55]	Conditions of Diabetes	Much real-time prediction of chronic condition was possibly by considering the readings of diabetes.
[44, 46]	Role of knowing the bio- markers	For earlier diagnosis of any chronic conditions in human beings, bio- markers serves a significant signing tool that alerts about any abnormalities.
[21, 46, 48]	Role of knowing Pathology consequence	Pathological consequence could often result in chronic condition, especially, causes CKD.

To know how much the above seen 4 common parameters influence/ affect the chronic conditions in human beings, we are plotting a chart showing the significance of those parameters by rating from 1 to 4 (with 1 being the highest significance and 4 being the lowest significance) depending on our observation in the below fig. 4.

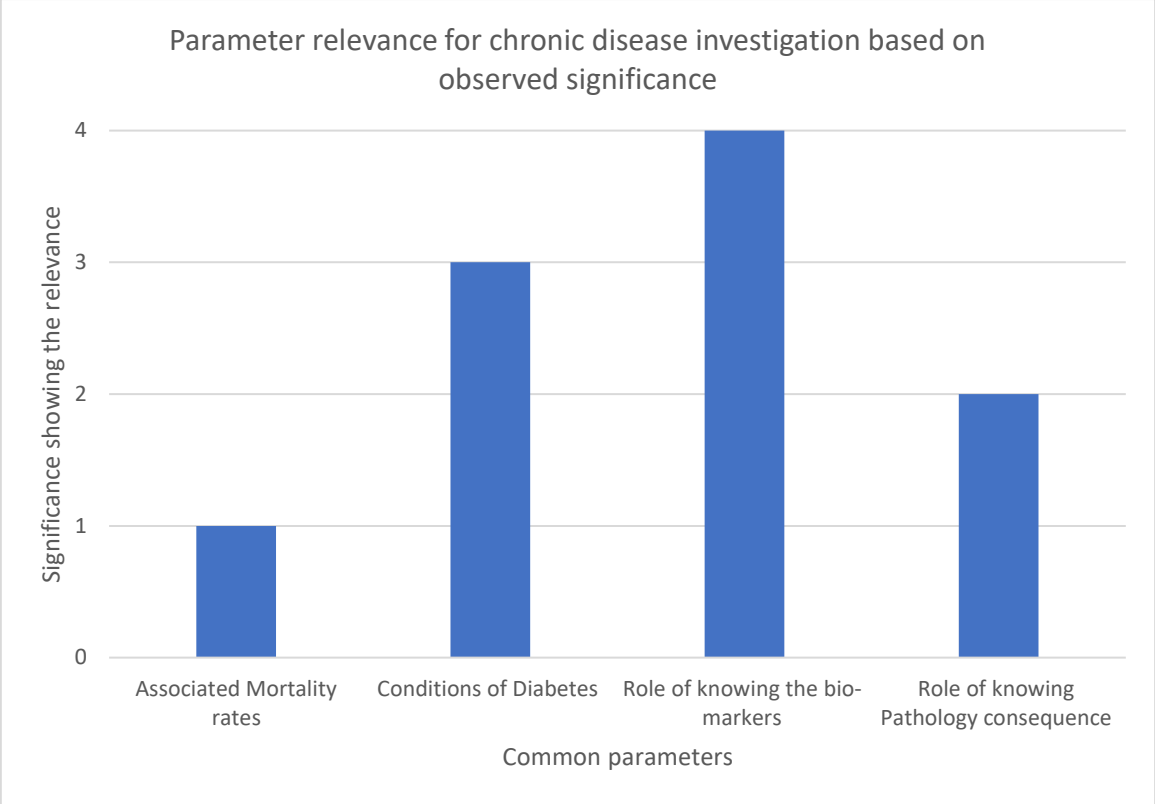

Fig. 4 Chart showing the parameter relevance for chronic disease investigation based on observed significance

VII. Conclusion and Future Work

With the comprehensive survey being made in this work, we have arrived at the following conclusions which could serve as the good source for the futuristic works:

- With the help of big data analytics, the way in which the large quantum of data gets processed have eased a lot in the recent years.
- As the role of big data analytics in healthcare was found to be immense, the possibilities for preserving the health of the patients could be considerably improvised.
- With the advent of chronic kidney disease prediction methodologies, we have been given a chance to treat it and get ourselves saved from the sudden deaths/ adverse health effects caused because of that chronic condition.
- The existence of those chronic kidney disease prediction is not alone enough, but the effective ones can actually reduce the unforeseen deaths of unattended or overlooked patients, thereby, the mortality rate could be lowered down.

If the chronic prediction methodologies are effectively proposed with deep learning concepts, it will be of great importance as they are way better than the conventional machine learning prediction systems. In particular, the Deep Belief Networks (DBNs) could be deployed by the future researchers for predicting any chronic level kidney cases with considerable accuracy to reduce any adverse health effects provided that it adopts to prevailing clinical conditions.

References

[1] A. Katal, M. Wazid, and R. H. Goudar, "Big data: issues, challenges, tools and good practices," in *2013 Sixth international conference on contemporary computing (IC3)*, 2013, pp. 404-409: IEEE.

[2] Katal, Wazid, and R. Goudar, H., "Big data: issues, challenges, tools and good practices," in *2013 Sixth international conference on contemporary computing*

(IC3), 2013, pp. 404-409: 2013 Sixth international conference on contemporary computing (IC3).

[3] Paul and Ghose, "A Novel Educational Proposal and Strategies Toward Promoting Cloud Computing, Big Data, and Human–Computer Interaction in Engineering Colleges and Universities," in *Advances in Smart Grid and Renewable Energy*: Springer, 2018, pp. 93-102.

[4] Mazumdar, Seybold, Kritikos, and Verginadis, "A survey on data storage and placement methodologies for cloud-big data ecosystem," *Journal of Big Data,* vol. 6, no. 1, pp. 1-37, 2019.

[5] Yang, Huang, Li, Liu, and Hu, "Big Data and cloud computing: innovation opportunities and challenges," *International Journal of Digital Earth,* vol. 10, no. 1, pp. 13-53, 2017.

[6] M. Abiodun, K. *et al.*, "Cloud and big data: a mutual benefit for organization development," *Journal of physics: conference series,* vol. 1767, no. 1, p. 012020, 2021.

[7] Zhou, Fu, Yu, Su, and Kuang, "Data integrity verification of the outsourced big data in the cloud environment: A survey," *Journal of Network and Computer Applications,* vol. 122, pp. 1-15, 2018.

[8] Mubarakali, A. Durai, D., Alshehri, AlFarraj, Ramakrishnan, and Mavaluru, "Fog-based delay-sensitive data transmission algorithm for data forwarding and storage in cloud environment for multimedia applications," *Big Data,* 2020.

[9] Amani *et al.*, "Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 13, pp. 5326-5350, 2020.

[10] Gianniti, Ciavotta, and Ardagna, "Optimizing quality-aware big data applications in the cloud," *IEEE Transactions on Cloud Computing,* vol. 9, no. 2, pp. 737-752, 2018.

[11] Lo'ai and Saldamli, "Reconsidering big data security and privacy in cloud and mobile cloud systems," *Journal of King Saud University-Computer and Information Sciences,* vol. 33, no. 7, pp. 810-819, 2021.

[12] R. K. Devi, K. V. Devi, and S. J. I. J. A. S. C. A. Arumugam, "Dynamic batch mode cost-efficient independent task scheduling scheme in cloud computing," vol. 8, no. 2, 2016.

[13] Ahad and Biswas, "Request-based, secured and energy-efficient (RBSEE) architecture for handling IoT big data," *Journal of Information Science,* vol. 45, no. 2, pp. 227-238, 2019.

[14] Madan and Goswami, "k-DDD measure and mapreduce based anonymity model for secured privacy-preserving big data publishing," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* vol. 27, no. 02, pp. 177-199, 2019.

[15] Boussif, Aloui, and Cherif, "Secured cloud computing for medical data based on watermarking and encryption," *IET Networks,* vol. 7, no. 5, pp. 294-298, 2018.

[16] Hababeh, Gharaibeh, Nofal, and Khalil, "An integrated methodology for big data classification and security for improving cloud systems data mobility," *IEEE Access,* vol. 7, pp. 9153-9163, 2018.

[17] Rao, Raju, Kumar, and G. Babu, C., "A Secured Data Sharing Protocol For Minimization Of Risk In Cloud Computing And Big Data," *NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal,* pp. 5574-5583, 2021.

[18] Manogaran, Varatharajan, Lopez, P. Kumar, M., Sundarasekar, and Thota, "A new architecture of Internet of Things and big data ecosystem for secured smart healthcare monitoring and alerting system," *Future Generation Computer Systems,* vol. 82, pp. 375-387, 2018.

[19] Viswanath and Krishna, "Hybrid encryption framework for securing big data storage in multi-cloud environment," *Evolutionary Intelligence,* vol. 14, no. 2, pp. 691-698, 2021.

[20] D. D. Rao, S. S. Raju, B. S. Kumar, G. C. J. N.-N. V. Babu, E. O. Journal, and NVEO, "A Secured Data Sharing Protocol For Minimization Of Risk In Cloud Computing And Big Data," pp. 5574-5583, 2021.

[21] Zhou *et al.*, "Modeling methodology for early warning of chronic heart failure based on real medical big data," *Expert Systems with Applications,* vol. 151, p. 113361, 2020.

[22] Kumar, Akangire, Sullivan, Fairchild, and Sampath, "Continuous vital sign analysis for predicting and preventing neonatal diseases in the twenty-first century: big data to the forefront," *Pediatric research,* vol. 87, no. 2, pp. 210-220, 2020.

[23] Zhan, "Sports and health big data system based on 5G network and Internet of Things system," *Microprocessors and Microsystems,* vol. 80, p. 103363, 2021.

[24] Alghamdi *et al.*, "Detection of myocardial infarction based on novel deep transfer learning methods for urban healthcare in smart cities," *Multimedia tools and applications,* pp. 1-22, 2020.

[25] Ali *et al.*, "An intelligent healthcare monitoring framework using wearable sensors and social networking data," *Future Generation Computer Systems,* vol. 114, pp. 23-43, 2021.

[26] Iqbal, Doctor, More, Mahmud, and Yousuf, "Big Data analytics and Computational Intelligence for Cyber–Physical Systems: Recent trends and state of the art applications," *Future Generation Computer Systems,* vol. 105, pp. 766-778, 2020.

[27] Kumar and Gandhi, "A novel three-tier Internet of Things architecture with machine learning algorithm for early detection of heart diseases," *Computers & Electrical Engineering,* vol. 65, pp. 222-235, 2018.

[28] Qin *et al.*, "Prediction of number of cases of 2019 novel coronavirus (COVID-19) using social media search index," *International journal of environmental research and public health,* vol. 17, no. 7, p. 2365, 2020.

[29] Rahman, N. Peeri, C., Shrestha, Zaki, Haque, and H. Ab, S. H., "Defending against the Novel Coronavirus (COVID-19) outbreak: How can the Internet of Things (IoT) help to save the world?," *Health policy and technology,* vol. 9, no. 2, p. 136, 2020.

[30] Liu *et al.*, "A novel cloud-based framework for the elderly healthcare services using digital twin," *IEEE access,* vol. 7, pp. 49088-49101, 2019.

[31] E. Cahan, M., Hernandez-Boussard, Thadaney-Israni, and D. Rubin, L., "Putting the data before the algorithm in big data addressing personalized healthcare," *NPJ Digital Medicine,* vol. 2, no. 1, pp. 1-6, 2019.

[32] L. Majnarić, T., Babič, O'Sullivan, and Holzinger, "AI and big data in healthcare: towards a more comprehensive research framework for multimorbidity," *Journal of Clinical Medicine,* vol. 10, no. 4, p. 766, 2021.

[33] Khan, Khan, Maqsood, Aadil, and M. Ghazanfar, A, "Optimized gabor feature extraction for mass classification using cuckoo search for big data e-healthcare," *Journal of Grid Computing,* vol. 17, no. 2, pp. 239-254, 2019.

[34] Shilo, Rossman, and Segal, "Axes of a revolution: challenges and promises of big data in healthcare," *Nature medicine,* vol. 26, no. 1, pp. 29-38, 2020.

[35] Dash, S. Shakyawar, K., Sharma, and Kaushik, "Big data in healthcare: management, analysis and future prospects," *Journal of Big Data,* vol. 6, no. 1, pp. 1-25, 2019.

[36] Shahbaz, Gao, Zhai, Shahzad, and Hu, "Investigating the adoption of big data analytics in healthcare: the moderating role of resistance to change," *Journal of Big Data,* vol. 6, no. 1, pp. 1-20, 2019.

[37] Galetsi, Katsaliaki, and Kumar, "Values, challenges and future directions of big data analytics in healthcare: A systematic review," *Social science & medicine,* vol. 241, p. 112533, 2019.

[38] Mehta and Pandit, "Concurrence of big data analytics and healthcare: A systematic review," *International journal of medical informatics,* vol. 114, pp. 57-65, 2018.

[39] Khanra, Dhir, A. Islam, N., and Mäntymäki, "Big data analytics in healthcare: a systematic literature review," *Enterprise Information Systems,* vol. 14, no. 7, pp. 878-912, 2020.

[40] Qian, Zhu, and Hoshida, "Use of big data in drug development for precision medicine: an update," *Expert review of precision medicine and drug development,* vol. 4, no. 3, pp. 189-200, 2019.

[41] Aceto, Persico, and Pescapé, "Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0," *Journal of Industrial Information Integration,* vol. 18, p. 100129, 2020.

[42] Pashazadeh and Navimipour, "Big data handling mechanisms in the healthcare applications: a comprehensive and systematic literature review," *Journal of biomedical informatics,* vol. 82, pp. 47-62, 2018.

[43] S. Dash, S. K. Shakyawar, M. Sharma, and S. J. J. o. B. D. Kaushik, "Big data in healthcare: management, analysis and future prospects," vol. 6, no. 1, pp. 1-25, 2019.

[44] Alan *et al.*, "Baseline total kidney volume and the rate of kidney growth are associated with chronic kidney disease progression in Autosomal Dominant Polycystic Kidney Disease," *Kidney international,* vol. 93, no. 3, pp. 691-699, 2018.

[45] Lees, C. Welsh, E., C. Celis-Morales, A., Mackay, Lewsey, and S. Gray, R. ... Mark P. B., "Glomerular filtration rate by differing measures, albuminuria and prediction of cardiovascular disease, mortality and end-stage kidney disease," *Nature medicine,* vol. 25, no. 11, pp. 1753-1760, 2019.

[46] Zhang and Parikh, "Biomarkers of acute and chronic kidney disease," *Annual review of physiology,* vol. 81, pp. 309-333, 2019.

[47] Song, L. Waitman, R., S. Alan, L., D. Robbins, C., Hu, and Liu, "Longitudinal risk prediction of chronic kidney disease in diabetic patients using a temporal-enhanced gradient boosting machine: retrospective cohort study," *JMIR medical informatics,* vol. 8, no. 1, p. e15510, 2020.

[48] V. Kolachalama, B. *et al.*, "Association of pathological fibrosis with renal survival using deep neural networks," *Kidney international reports,* vol. 3, no. 2, pp. 464-475, 2018.

[49] Kuo *et al.*, "Automation of the kidney function prediction and classification through ultrasound-based kidney imaging using deep learning," *NPJ digital medicine,* vol. 2, no. 1, pp. 1-9, 2019.

[50] Biyik and Guney, "Lipid accumulation product and visceral adiposity index: two new indices to predict metabolic syndrome in chronic kidney disease," *Eur Rev Med Pharmacol Sci,* vol. 23, no. 5, pp. 2167-2173, 2019.

[51] Henry and Lippi, "Chronic kidney disease is associated with severe coronavirus disease 2019 (COVID-19) infection," *International urology and nephrology,* vol. 52, no. 6, pp. 1193-1194, 2020.

[52] D. Charytan, M. *et al.*, "Coronary flow reserve is predictive of the risk of cardiovascular death regardless of chronic kidney disease stage," *Kidney international,* vol. 93, no. 2, pp. 501-509, 2018.

[53] Ravindra, Sriraam, and Geetha, "Classification of non-chronic and chronic kidney disease using SVM neural networks," *Int. J. Eng. Technol,* vol. 7, no. 1, pp. 191-194, 2018.

[54] Vasquez-Morales, M.-M. , S. M., Moreno-Ger, and J. Recio-Garcia, A., "xplainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning," *Ieee Access,* vol. 7, pp. 152900-152910, 2019.

[55] Ravizza *et al.*, "Predicting the early risk of chronic kidney disease in patients with diabetes using real-world data," *Nature medicine,* vol. 25, no. 1, pp. 57-59, 2019.

[56] N. Almansour, A. *et al.*, "Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study," *Computers in biology and medicine,* vol. 109, pp. 101-111, 2019.

[57] Xiao *et al.*, "Comparison and development of machine learning tools in the prediction of chronic kidney disease progression," *Journal of translational medicine,* vol. 17, no. 1, pp. 1-13, 2019.

[58] Elhoseny, Shankar, and Uthayakumar, "Intelligent diagnostic prediction and classification system for chronic kidney

disease," *Scientific reports,* vol. 9, no. 1, pp. 1-14, 2019.

[59] Mihai *et al.*, "Inflammation-related mechanisms in chronic kidney disease prediction, progression, and outcome," *Journal of immunology research,* vol. 2018, 2018.