**Health Monitoring System: A Review on Diabetes Prediction and Providing Medical Assistance to the Patients Using Big Data and Cloud IoT.**

**Sakshi Tyagi**

Student

Department of Computer Applications

KIET Group of Institutions, Delhi-NCR, Ghaziabad

[sakshi.2023mca1049@kiet.edu](mailto:sakshi.2023mca1049@kiet.edu)

**Shivani Chauhan**

Student

Department of Computer Applications

KIET Group of Institutions, Delhi-NCR, Ghaziabad

[shivani.2023mca1008@kiet.edu](mailto:shivani.2023mca1008@kiet.edu)

**Mukul Dhama**

Student

Department of Computer Applications

KIET Group of Institutions, Delhi-NCR, Ghaziabad

[mukul.2023mca1019@kiet.edu](mailto:mukul.2023mca1019@kiet.edu)

***Abstract***

We are conducting a scoping review of the academic literature on Big Data and diabetes care in order to review current Big Data applications in diabetes care and consider future potential. Data is being generated at an increasing rate in the healthcare industry, and this data has the potential to revolutionize the way diabetic care is delivered. Through data study, Diabetes care is beginning to be impacted by big data. Big Data's applicability in normal health treatment is yet in the future. Massive volumes of healthcare data are already being generated; the trick is to figure out how to use it to generate actionable insights. These objectives will necessitate a significant amount of development work.

Predictive analytics for diabetics are very important as it helps diabetics, their families, doctors, and medical researchers make decisions about diabetics based on a high amount of data. Diabetes is a condition that occurs when the body cannot use glucose normally. Glucose is the usual source of Body Cell Essence. Blood levels of glucose are caused by a hormone called insulin produced by the pancreas. There are basically two types of diabetes. Patients with type 1 diabetes and patients with type 2 diabetes. In patients with type 1 diabetes, the pancreas is persistently unable to produce insulin, and in patients with type 2 diabetes, the pancreas produces insulin and does not carry sufficient insulin or does not function. Today, type 1 diabetes is a widespread and prominent clinical problem. Current methods use methods other than snooping, and patient-tolerant data is typically sent to specialists using the IoT. Therefore, the proposed approach captures the blood glucose levels of diabetics, especially for robotized diabetics who check the structure in the ICU.

**1. Introduction**

Available clinical records show that type 1 diabetes is a notable clinical problem worldwide. There are about 2.6 million adults over the age of 18 suffering from diabetes, and the severity of diabetes should increase in Malaysia. Ketones are artificial substances that appear in the body when the muscle-to-fat ratio is used, as opposed to glucose due to urgency. This indicates that the body's cells cannot absorb enough sugar (glucose) from the blood, especially if the body's insulin is too low. Insulin is used by the body to use glucose, which focuses on what is important. In this sense, which screens ketone zookeepers one by one, helps control and screen the condition of diabetics with a vast number of ketones that remain wild in diabetics. Anyway, the two steps were considered intrusive, horrifying, and anomalous. CH3) 2CO is abstractly recognized by the method of diabetes biomarkers. CH3) 2CO stands for traditional stuff and knows the technology for Ketobetix and the fragrant scent of breathing in diabetics. The combination of breath CH3) 2CO is associated with glucose absorption and lipolysis. This is the method by which attachment to CH3) 2CO respiration is presented in the elevated modern form of diabetes and can be used to study the development of diabetes is called the easiest way and diabetes. Prevents rapid detection of sex ketoacidosis Type 1 diabetes. A method of checking ketone levels is performed using breath estimation. Atem presents a simple portable. Today, there are several chronic illnesses such as heart disease, stroke, cancer, chronic respiratory illness, and diabetes. It is a dangerous disease and has recently become one of the leading causes of death worldwide and requires careful monitoring to maintain patient health. The biggest challenge for diabetics is to raise or lower blood glucose within a specific interval, as diabetes is caused by insulin resistance and inadequate insulin production can lead to level up or level down of blood glucose. If these conditions are not met, some patients may need urgent medical care to avoid exacerbations [1] human management for checking diabetes levels. The strategy presented demonstrated advances in the hardware relationship with the Internet of Things (IoT) system to enhance patient evaluation and individual observation methods. The Arduino board will be used to study the sensor with breath detection capability. Breath observation levels are recorded on the system using distant correspondence. Data collection is linked to the site page. Ketone levels are assessed as the percentage of CH3) 2CO in exhaled breath accumulated when the patient breathes into a mouthpiece containing a gas sensor. This assessment is based on detecting the patient's blood glucose level by separating CH3) 2CO levels from respiration at and sending data to clinical authorities by the WIFI method via a message. These devices are a new way of continuous monitoring. They provide real-time information about blood glucose levels. This article introduces an intelligent diabetes monitoring system using the node MCU and machine learning algorithms. The MCU node is connected to a glucose meter and periodically records the blood glucose levels of diabetics. This collected data can be used by caregivers (patients, researchers, and doctors) to remotely monitor patients. As a result, patients and physicians alike need to process multiple records, interpret vast amounts of data, adjust insulin doses, and bring blood glucose levels as close to normal as possible

Diabetes is a life-threatening disease that has no cure. If you get this sickness once, it will be with you for the rest of your life. At the same time, having too much glucose in your blood might cause health problems. Kidney illness, heart disease, stroke, vision problems, dental problems, foot problems, and nerve damage are just a few examples. so that you can keep track of your diabetes and avoid complications

Kind 1 diabetes is the most well-known type of diabetes.

Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type 2 Diabetes Type

Type 1 diabetes occurs when the body is unable to manufacture insulin. It has a negative impact on children and young adults. It can also affect people of any age. People with this kind of diabetes must take insulin on a daily basis.

Type 2 diabetes occurs when the body is unable to manufacture or utilise insulin.

This kind of diabetes primarily affects people in their forties and fifties.

Diabetes During Pregnancy

This kind of diabetes primarily affects women. During pregnancy, this kind of diabetes develops. High blood sugar levels caused by gestational diabetes can harm your pregnancy and your baby's health.

**A CHALLENGES**

Diabetes can also lead to eyesight issues. It lowers blood glucose levels in the retina of diabetics who are older. It causes cataracts in diabetics in the future, and it causes bad vision quite readily. Patients' vision problems cause them a lot of trouble and interfere with their regular activities.

Hearing loss was induced by diabetes. Long-term high blood glucose levels can disrupt the delivery of blood and oxygen to the inner ear's tiny nerves and blood vessels, resulting in hearing loss. The nerves and blood arteries in the ear become damaged with time, limiting the person's ability to hear. It causes misunderstandings between people. As a result, every diabetic patient should assess their hearing ability.

Diabetic patients are physically weak and have a poor quality of life. As a result, diabetics must raise and maintain their daily activity levels. Controlling one's eating habits is crucial, as is physical activity. Diabetic patients must be encouraged to exercise on a daily basis. Brisk walking, bicycling, swimming, housework, and gardening are just a few examples.

**2. Big Data and the Future of Diabetes**

Diabetes affects 9% of the population in the United States, according to the Centres for Disease Control and Prevention. The demand for useful information on how to treat and manage the disease is greater than ever. A wide approach to data analysis can aid healthcare providers in gaining a better understanding of the condition, its prognoses, and its complications.

Great amounts of data are required for big data analysis, and persons with diabetes generate a large quantity of data just by going about their regular lives. Wearable exercise monitors, smart blood pressure cuffs, Bluetooth-enabled bathroom scales, and smart insulin pens all generate data.

For example, the FreeStyle Libre flash glucose monitoring system makes it simple for persons with diabetes to check their glucose levels in real-time.

Three basic types of big-data oriented solutions are presently accessible and progressively accepted from an IT standpoint. To begin with, cloud computing offers cost-effective ways to achieve great computational performance. **[2]** Second, parallel programming is becoming increasingly simple and efficient: "MapReduce," a programming model that allows algorithms to be implemented in distributed contexts, is becoming a very popular and frequently used paradigm. **[3]** Finally, new database technologies, such as No-SQL databases, are now available to address both the scalability and variety issues. **[4**] Better healthcare is explained by various big data technology stacks and research into health care mixed with efficiency, cost savings, and other factors. The use of Hadoop in health care has become more essential as a means of processing data and doing large-scale data management tasks. Hadoop's cost effectiveness can be improved by employing analytics on coupled compute and storage **[5].**

**3. Cloud computing with IoT architecture in prediction of diabetes**

The Internet of Things (IoT) and mobile health care (m-healthcare) applications provide numerous dimensionalities and online services. These applications have offered a new platform for millions of individuals to benefit from health suggestions on a regular basis in order to live a healthy life. The numerous aspects of these healthcare online applications were reinforced after the introduction of IoT technology and related devices that are employed in the medical industry. IoT devices in the healthcare context generate a massive amount of big data. Cloud computing technology is used to manage massive amounts of data while also being simple to use. Cloud-based applications are playing an increasingly important role in today's fast-paced environment.

For safe storage and access, these medical apps make advantage of Cloud Computing technologies. We propose a novel Cloud and IoT-based Mobile Health Care application for monitoring and detecting critical diseases in order to provide better services to people through online healthcare applications. A new framework for the general public is being established here. In this study, a new systematic strategy for diabetes disorders and related medical data is developed using the UCI Repository dataset and medical sensors in order to forecast persons who are seriously impacted by diabetes. In addition, for detecting the disease and its severity, we present a new classification technique called Fuzzy Rule-based Neural Classifier. The studies were carried out using the standard UCI Repository dataset as well as real health records obtained from various hospitals. The experimental results suggest that the proposed approach beats existing disease prediction systems in terms of performance.

**4. Predictive analytics system for Diabetes data**

Diabetes mellitus is one of the most common noncommunicable diseases today, with a significant impact on human life. It is currently regarded as one of the worst diseases on the planet. If diabetes is not addressed, it can lead to a variety of health problems.The healthcare business collects a massive amount of data, much of which is unstructured. The information must be organised into basic values. Medical intelligence systems will be created by applying computer analytics to the huge amounts of data collected in the healthcare system, which will aid medical prediction. This will result in a patient-centered healthcare system that lowers medical costs. In health care, predictive analytics is largely used to identify individuals who are in the early stages of diabetes, asthma, heart disease, and other serious life-threatening diseases. To predict type 2 diabetes, the suggested technique PDD employs data mining algorithms.

**5. Data collection**

The system receives the raw diabetes large data or data set as input. The unstructured voluminous input data can be obtained from a variety of Electronic Health Record (EHR) / Patient Health Record (PHR), Clinical systems, and external sources (government sources, laboratories, pharmacies, insurance companies, and so on), in a variety of formats (flat files,.csv, tables, ASCII/text, and so on), and from a variety of locations **[6].**

**6. Cloud-based patient profile analytics system**

The computer or system is educated with clinical data obtained from healthcare organisations in order to forecast and analyse patient profiles. Sensors are attached to persons in rural regions to monitor their health on a monthly or weekly basis, and these sensors use IoT devices to detect their health **[7].** The IoT gadget, on the other hand, is linked to the service provider (hospitals, doctors, clinics, etc.). Basically, the cloud-based data structure has many modelling systems, thus the cloud structure components have saved clinical or medical profile data for each and every patient, which are then trained to the machine throughout the machine learning process **[8].** Patient profiles, which include patient health graphs, bio data, and other information.2

The research of a TB patient's body condition is more crucial in this criterion in order to examine severity measurements and supply medical instructions **[9].** Several strategies, such as feature selection and various segmentation processes, should be followed during the data analytic process. In addition, the analytic approach could comprise classification, regression, feature learning, and other prediction methods.

Cloud computing has become one of the most often used expressions to describe a type of on-demand computer service provided by companies such as Amazon, Google, and Microsoft. It is a virtual model on which computations are done and services are provided to consumers, hence the name "cloud." Infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service are three main types of cloud services offered by cloud providers (SaaS). The notion of virtualization, which allows virtual machines (VM) to run on top of existing hardware to meet the needs and demands of users, lies at the heart of cloud computing **[10**]. In hospitals, cloud computing opens up new opportunities for exchanging information with patients about their health, treatment alternatives, and the treatment process. The usage of IT is influenced by continuous change in both supply and demand in the healthcare sector, and this is the primary driver for cloud computing adoption.Diabetes is a disease that affects a large number of individuals nowadays. Diabetes is caused by a shortage of insulin in the blood. Diabetes Mellitus **[11]** is the type of diabetes that is the most common.

**7.Limitation and Strength**

There is no doubt that big data has had a positive impact on the healthcare system. The main problem with unstructured and structured data in healthcare is a lack of data, which makes disease prediction and symptom analysis difficult. However, the data contained in the healthcare system is both uneven and balanced. Furthermore, healthcare patient profile data is maintained in legacy systems (often electronic medical record systems) that have limited interoperability. However, incorporating health information into big data is difficult, as there are numerous formats, metadata, schema, and standards that affect the data. Moral issues such as data privacy, discretion, and control of access to patients, as well as the commercialization of deidentified patients' data; and the ownership and control of patients' data are all factors that stymie the effective exchange of profile details between patients and healthcare providers. As a result, combining healthcare data from several sources becomes difficult. As a result, getting a detailed and complete image of a patient in a timely manner throughout care becomes an issue. Standards are agreed-upon specifications that allow disparate systems, tools, equipment, and platforms to communicate with each other. Healthcare organisations, on the other hand, do not all follow the same pattern.Case reports, medications, disorders, and examinations, for example, have distinct titles and codes in different hospitals. Healthcare data should be sufficiently protected and secured, according to healthcare providers and big data analytics developers. When studying big data in healthcare, tools that assure the confidentiality, integrity, and availability of protected health information should be employed. Physical security, data encryption, user authentication, and application security should all be used to protect healthcare data. It's also a good idea to encourage the adoption of audit trail systems.

**8. Diabetes data analytics tools and techniques**

A variety of techniques and approaches are utilised to assess patient data and explore the symptoms that accompany it. IoT with diabetetic systems is a hot topic in big data analytics right now, and AI is also being used in prediction of diabetes for machine learning purposes. . A description of the proposed structures and the used algorithms on these paintings are given. In their paintings, the authors supplied a smart structure for the surveillance of diabetic ailment that displays the fitness of diabetic sufferers through sensors incorporated into smartphones [12]. In some other paintings, the prediction of diabetes sorts the use of evaluation algorithms and Hadoop map-lessen, prediction of complications, and the prediction of the kind of remedy had been investigated [13], at the same time as in [14], the authors proposed a machine which could carry out predictions for Mellitus that's a form of diabetes the use of Hadoop/MapReduce. A new machine for the prediction of glucose awareness has been proposed in. Where the facts generated via way of means of the Continuous Glucose Monitoring may be analysed via way of means of the glucoSim software program the use of the Kalman Filter (KF) to reducing noise. Many researchers have used different data mining techniques to develop and implement different analytical and predictive models. In [15], the author found a pattern in the Diabetes dataset using the classification method by Naive Bayes and the decision tree algorithm using the Weka tool. In [16], the author uses naive Bayes and decision trees in a model classification technique to examine the hidden patterns in the diabetes dataset. The authors of [17] predicted patient diabetes using the C4.5 decision tree algorithm, Neural network algorithm, Kmeans clustering algorithm, and visualizations.

**8.1. Support Vector Machine**

SVM stands for supervised knowledge representation model. Furthermore, it is the associated algorithm model that is used to study patient data for regression and classifications.

A Support Vector Machine creates a hyperplane or group of hyperplanes in high layered space, maps all of the models in a guide, and divides the examples by an identifiable hole about as broad as could be expected, with each side presenting one class. We should adjust the regularisation boundary C, which determines the model's complexity, in this strategy.

Greater C indicates that the misclassification will be punished more severely, implying that the model will be more likely to overfit. The SVM classifier is used to divide the data into a particular number of categories. It is quite tough to use AI and data mining to each and every exploration investigation in order to deconstruct diabetes. We'll deconstruct multiple approaches and apply them to the dataset. We will make every effort to achieve the best possible result. The present enhancement focuses on increasing characterisation precision while reducing execution time.

**8.2. Decision Tree**

The supervised learning approach, which is used to solve problems with arrangements. A decision tree is a mechanism that iteratively divides a given dataset into at least two examples. The technique's goal is to predict the class worth of the objective variable. The decision tree will assist in separating the informative collection and constructing the choice model in order to predict the obscure class marks. A decision tree can be built for both parallel and continuous factors. The root hub is ideally observed by a decision tree based on the most significant entropy value. As a result, choice tree has the advantage of selecting the most stable theory from the preparation dataset. A dataset including a few credits and occasions esteems is a contribution to the Decision tree, and the result is the choice model. Choosing the parting characteristic, parts, halting rules, and pruning, as well as preparing tests, quality and quantity, and the request for parts, are all issues that must be addressed while developing a choice model.

**8.3. Naive Bayes (NB)**

Prior research has shown that Induction algorithms based on Naive-Bayes have shown surprising accuracy in many classification tasks even when the assumption of conditional independence on which they are founded is broken. The majority of studies, however, were conducted on small databases. We demonstrate this in certain larger databases, Naive-Bayes accuracy does not scale up as well as decision trees. Following that, we suggest a new algorithm, NBTree, which creates a hybrid of decision-tree and Naive Bayes classifiers: As with ordinary decision trees, the nodes of the decision tree include univariate splits, whereas leaves contain Naive-Bayesian classifiers. The approach combines the interpretability of Naive-Bayes and decision trees with the performance of classifiers that outperform both components, especially when applied to large databases.

**9.Solution foreseen**

Prior to implementing algorithms in patient risk prediction, it is necessary to examine previously existing algorithms, approaches, and events. The use of Cloud IoT in the health-care system is an effective technique of predicting and assessing high risk. The suggested health-care disease prediction system combines traditional health informatics with cloud IoT platform-based big data analytics. The disease in patients can be sensed by the patients themselves, and the information about the diseases is sent to the hospital's doctors, who study the data and then assess the patients based on their illnesses.

**10. Structures includes in Diabetes analytics**

Several models in the domains of healthcare informatics and big data analytics are described in this study. Table3.1 summarises the classification of these platforms, including all tools and approaches utilised in the comparison.

|  |  |  |  |
| --- | --- | --- | --- |
| **Author** | **Source and Year** | **Disease Diagnoses method** | **Outcome** |
| Z. Mian *et al* [1] | **Source:** Continuous Glucose Monitoring: Review of an Innovation in Diabetes Management.  **Year:** 2019 | Continuous glucose monitoring and sensor-enabled pump technology are used. | This technology eliminates the need for frequent blood glucose monitoring, which is often inconvenient for patients, and instead offers them a more convenient option. |
| Lin *et al* [2] | **Source:** Enabling large-scale biomedical analysis in the cloud.  **Year:** 2013 | Explains the data-intensive computing system and lists available cloud-based bioinformatics resources. | To make a large amount of variety of data understandable and usable for biomedical research, we need to make it easier. |
| Mohammed *et al* [3] | **Source:** Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends.  **Year:** 2014 | The Hadoop platform and the MapReduce programming framework's potential applications. | To process large amounts of clinical data in sectors connected to medical health informatics |
| Lee *et al* [4] | **Source:** Alternatives to relational database: comparison of NoSQL and XML approaches for clinical data storage  **Year:** 2013 | The feasibility of three database technologies - NoSQL, XML-enabled, and native XML - for structured clinical data is evaluated. | The greatest choice for query performance is a NoSQL database, although XML databases are better in terms of scalability, flexibility, and extensibility, all of which are necessary to deal with the features of clinical data. |
| D. Peter Augustine [5] | **Source:** Leveraging Big Data analytics and Hadoop in Developing India’s Health Care Services.  **Year:**2014 | Analyze and demonstrate the benefits of Big Data Analytics and Hadoop in healthcare applications. | The use of Big Data Analytics and Hadoop illustrates the significance of these technologies in providing healthcare services to everyone at the lowest possible cost. |
| Wullianallur Raghupathi *et al* [6] | **Source:** Big data analytics in healthcare: promise and potential.  **Year:**2014 | To describe big data analytics' promise and potential in healthcare. | For healthcare academics and practitioners, it provides a wide understanding of big data analytics. |
| Kalet *et al* [8] | **Source:** Quality assurance tasks and tools: The many roles of machine learning.  **Year:**2019 | Describe some study topics and some of the particular obstacles each one faces. | Improvements in planning time, plan quality, advanced dosimetric QA, predictive machine maintenance, higher safety checks, and advancements are all important for new adaptive planning-driven QA paradigms. |
| Pande, Tripti, *et al* [9] | **Source:** Prevalence of diabetes mellitus amongst hospitalized tuberculosis patients at an Indian tertiary care center: A descriptive analysis.  **Year:**2018 | Calculate the prevalence of Diabetes Mellitus in adulthood. | Age, type of TB, and undernutrition were all found to be significant predictors of TB-DM co-prevalence. |
| Ahmad M. Manasrah [10] | **Source:** A Variable Service Broker Routing Policy for data center selection in cloud analyst.  **Year:**2017 | Variable Service Broker Routing Policy is used to reduce the processing and response time of customer requests while staying within a reasonable cost range. | The proposed policy alters the old policy's sorting and selection equations. |
| A. Rghioui *et al* [11] | **Source:** A Smart Glucose Monitoring System for Diabetic Patients.  **Year:** 2020 | A compact portable device capable of detecting blood glucose levels and body temperature in diabetics. | Diabetes disease surveillance would allow doctors to remotely monitor their patients' health using sensors included in smartphones and smart portable devices. |
| N. M. S. Kumar *et al* [12] | **Source:** Predictive Methodology for Diabetic Data Analysis in Big Data.  **Year:** 2015 | Hadoop/predictive MapReduce's analytical method Reduce the environment to forecast the types of diabetes that are common and the complications that come with it. | This approach enables patients to be cured and cared for in a more efficient manner, with improved outcomes like as affordability and accessibility. |
| Ahmed, H.B *et al* [13] | **Source:** Effects of External Factors in CGM Sensor Glucose Concentration Prediction.  **Year:**2016 | To develop a blood glucose prediction system for usage in conjunction with a continuous glucose monitoring (CGM) device. | This technology eliminates the need for frequent blood glucose monitoring, which is often inconvenient for patients, and instead offers them a more convenient option. |
| Dr Saravana Kumar N M *et al* [14] | **Source:** Predictive Methodology for Diabetic Data Analysis in Big Data.  **Year:** 2015 | Hadoop/predictive MapReduce's analytical method Reduce the environment to forecast the types of diabetes that are common and the complications that come with it. | This approach enables patients to be cured and cared for in a more efficient manner, with improved outcomes like as affordability and accessibility. |
| Dost Muhammad Khan1 *et al* [15] | **Source:** An Integration of K-means and Decision Tree (ID3) towards a more Efficient Data Mining Algorithm.  **Year:** 2011 | To create a more efficient data mining approach employing intelligent agents, we combine the K-means clustering algorithm with the Decision tree (ID3) algorithm. | Data mining algorithms are used to uncover hidden patterns and connections across variables in large datasets. |
| Dost Muhammad Khan1 *et al* [16] | **Source:** An Integration of K-means and Decision Tree (ID3) towards a more Efficient Data Mining Algorithm.  **Year:**2011 | Integrating K-means clustering algorithm  with Decision tree (ID3) algorithm. | To develop a more efficient data mining technique that makes use of an intelligent agent called Learning Intelligent Agent (LIAgent), which can conduct classification, grouping, and interpretation tasks on datasets. |
| Deepti Sisodia *et al* [17] | **Source:** Prediction of Diabetes Using Classification Algorithm.  **Year:**2018 | The goal is to create a model that can accurately predict the likelihood of diabetes in people. | The designed system can be used to predict or diagnose diabetes using machine learning classification methods. |

**11.Conclusion**

This research looked at machine learning classification techniques for better diabetes illness prediction. The accuracy of this literature analysis in the SVM classification technique was the greatest. Different measures are used to compute the various performance values of categorization algorithms. Pima Indians Diabetes Dataset can be used to train and test data. The categorization algorithm reached the highest level of testing precision. This research gathered a variety of categorization techniques and combined them to improve accuracy, specificity, and sensitivity.

Using elastic net regression, the challenges that are studied in the previous works for enhancing accuracy for diabetes prediction and diagnosis will be worked out further. Elastic net regression is a hybrid of LASSO and Ridged Regression approaches that allows data in category, numerical, and picture form to be fed into the regression. The AdaBoost classifier was found to be the best model, with an accuracy of 98.8%. With two separate datasets, we compared the accuracies of machine learning algorithms. When compared to the existing dataset, it is obvious that the model enhances the accuracy and precision of diabetes prediction. This research could be expanded to see how likely non-diabetic persons are to develop diabetes in the coming years.

**References:**

[1] Z. Mian, K. L.Hermayer, A. Jenkins, "Continuous Glucose Monitoring: Review of an Innovation in Diabetes Management", The American Journal of the Medical Sciences Vol. 358, pp: 332-339, Issue 5, November 2019

[2] Lin, YC, Yu, CS, Lin, YJ. Enabling large-scale biomedical analysis in the cloud. Biomed Res Int. 2013;2013:185679.

[3]   
Mohammed, EA, Far, BH, Naugler, C. Applications of the MapReduce programming framework to clinical big data analysis: current landscape and future trends. BioData Min. 2014;7:22.

[4]   
Lee, KK, Tang, WC, Choi, KS. Alternatives to relational database: comparison of NoSQL and XML approaches for clinical data storage. Comput Methods Programs Biomed. 2013;110(1):99-109.

[5]. D. Peter Augustine, “Leveraging Big Data analytics and Hadoop in Developing India’s Health Care Services”, International Journal of Computer Applications, vol 89(16), pp 44-50, 2014.

[6]. Wullianallur Raghupathi, and Viju Raghupathi, “Big data analytics in healthcare: promise and potential”, Health Information Science and Systems, vol. 2(3) pp. 2-10, 2014.

[7] Liu, Ruiqing, et al. "Profile of Consecutive Fecal Calprotectin Levels in the Perioperative Period and Its Predictive Capacity for Early Endoscopic Recurrence in Crohn’s Disease." Diseases of the Colon & Rectum 62.3 (2019): 318-326.

[8] Kalet, Alan M., Samuel MH Luk, and Mark H. Phillips. "Quality assurance tasks and tools: The many roles of machine learning." Medical physics (2019).

[9] Pande, Tripti, et al. "Prevalence of diabetes mellitus amongst hospitalized tuberculosis patients at an Indian tertiary care center: A descriptive analysis." PloS one 13.7 (2018): e0200838.

[10] Ahmad M. Manasrah, “A Variable Service Broker Routing Policy for data center selection in cloud analyst”, Journal of King Saud University – Computer and Information Sciences 29, 365–377, 2017.

[11] P. Suresh Kumar, “Diagnosing Diabetes using Data Mining Techniques”, International Journal of Scientific and Research Publications, Vol 7, June 2017.

[12] A. Rghioui, J. Lloret, M. Harrane, A. Oumnad. “A Smart Glucose Monitoring System For Diabetic Patient”. Electronics.9 (4), 678, 2020.

[13] N. M. S. Kumar, T. Eswari, P. Sampath, and S. Lavanya, “Predictive Methodology for Diabetic Data Analysis in Big Data,” ProcediaComput. Sci., vol. 50, pp. 203–208, Jan. 2015

[14] Ahmed, H.B.; Serener, A. Effects of External Factors in CGM Sensor Glucose Concentration Prediction. Procedia Compute. Sci. 2016, 102, 623–629

[15] Dr Saravana Kumar N M, Eswari T, Sampath P and Lavanya S,” Predictive Methodology for Diabetic Data Analysis in Big Data”, 2nd International Symposium on Big Data and Cloud Computing,2015

[16] Dost Muhammad Khan1, Nawaz Mohamudally2, “An Integration of K-means and Decision Tree (ID3) towards a more Efficient Data Mining Algorithm”, Journal Of Computing, Volume 3, Issue 12, December 2011

[17] Deepti Sisodia, Dilip Singh Sisodia,” Prediction of Diabetes Using Classification Algorithm”, www.elsevier.com/locate/procedia, Procedia computer science 132(2018) 1578-1585.