

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN DIABETES CARE: A SYSTEMATIC REVIEW

Rupesh Jain

Research Scholar, University of Technology, Jaipur

ABSTRACT

Artificial intelligence (AI) is a rapidly developing field, and its applications to diabetes, a global pandemic, have the potential to radically alter how this chronic disease is diagnosed and treated. Diabetes management with the help of digital therapeutics has become a well-established intervention for lifestyle therapy. Diabetes self-management is becoming more common as a result of patients and doctors benefiting from clinical decision support. Remote monitoring of a patient's symptoms and biomarkers can now be performed indefinitely and without the patient's participation thanks to AI. Patients are more likely to be actively involved in their diabetes care when they use social media and online communities. Improvements in technology have allowed for more efficient management of diabetes's resources. Smart technical changes have improved glycemic control overall, resulting in lower fasting and postprandial glucose levels, fewer glucose excursions, and less glycosylated hemoglobin. Diabetes management will undergo a dramatic change as AI ushers in the era of data-driven precision care rather than the more traditional approaches to disease management.

Keywords: Artificial intelligence, Machine Learning, Diabetes, Clinical.

1. INTRODUCTION

Type 2 diabetes is very common among the older population in the United States. According to the International Diabetes Federation, there were 451 million cases of diabetes globally in 2017. In just 26 years, the population of Earth is expected to grow by more than 693 million people. However, both hereditary and environmental factors have been hypothesised to have a role in the development of diabetes, the exact nature of which remains unknown. Although there is currently no cure for it, medicines and therapy may help alleviate the symptoms. Cardiovascular collapse and organ failure are only two of the complications that diabetics face. Diagnosis and treatment of diabetes mellitus (DM) may also aid in the prevention of complications and the mitigation of serious health problems. [1]

Diabetic retinopathy can be diagnosed either manually by a clinician or automatically by an automated diagnostic system. Indicators of DM can be measured in a number of ways. Manual diagnosis has various benefits, one of which is that it does not require the use of machinery, freeing clinicians to concentrate on their area of expertise. It is possible that not even a highly experienced physician will be able to recognise all of the early signs of DM. A more accurate and timely illness diagnosis is possible now than ever before because to advancements in machine learning and artificial

intelligence (AI). With less work on their plates, doctors and nurses will make fewer mistakes. Use of computer-based decision support systems may aid in accurate diagnosis and cost-effective management. Test findings, patient reports, therapy and follow-ups, and medication are only a few of the many types of data available in the DM field. Data compilation is tedious and prone to mistakes when done manually. Poor data management has led to a breakdown in data organisation. The increasing volume of data necessitates a more refined approach to data extraction and analysis. Devices like this are utilised in cutting-edge medical facilities to streamline the collection and dissemination of huge volumes of data. Compared to human detection and diagnosis, automated identification and diagnosis of DM and abnormalities is far simpler and more reliable. Consequently, automated methods are required for identifying the presence of diabetes mellitus. Using machine learning and AI, we can build fully automated DM systems.

2. ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is defined as “a discipline of computer science that tries to construct systems or methods that evaluate information and allow the management of complexity in a wide variety of applications.” For effective data management and the creation of tools and technologies for managing diabetes, the application

of AI is both possible and desired. Having safe designs, safety reserves, and procedural safeguards in place, with all uncertainties identified for all prospective technological systems, is advised for providing safer technology through AI. The introduction of wearables, cellphones, and other devices has allowed for the constant surveillance of patients' symptoms and illness state due to technological advancements. Experts in the medical field should provide diabetic patients the option of using AI-assisted therapy. [2]

Artificial intelligence (AI) has the potential to enhance diabetes treatment in three primary areas: for people with diabetes, for healthcare providers, and for healthcare systems. Artificial intelligence (AI) has improved the self-care of diabetic patients in new ways, improved the decision-making speed and accuracy of doctors, made it easier for them to

keep in touch with their patients, and made better use of healthcare resources.

The IDx-DR gadget analyses digital retinal pictures with an AI algorithm to help diagnose retinopathy at an early stage. The device has received FDA approval. AI is being encouraged for application in diabetic treatment by the American Diabetes Association. The American Diabetes Association (ADA) has approved the use of self-driving AI to identify diabetic retinopathy and macular edema. Patient education and agency are enhanced by AI's availability. The prevalence and severity of co-morbidities, patient behaviours, length of stay in healthcare facilities, and frequency of visits to and interactions with healthcare practitioners are all influenced by digital solutions' impact on health care systems. Patient intake and internal transfers have also benefited from AI.[3]

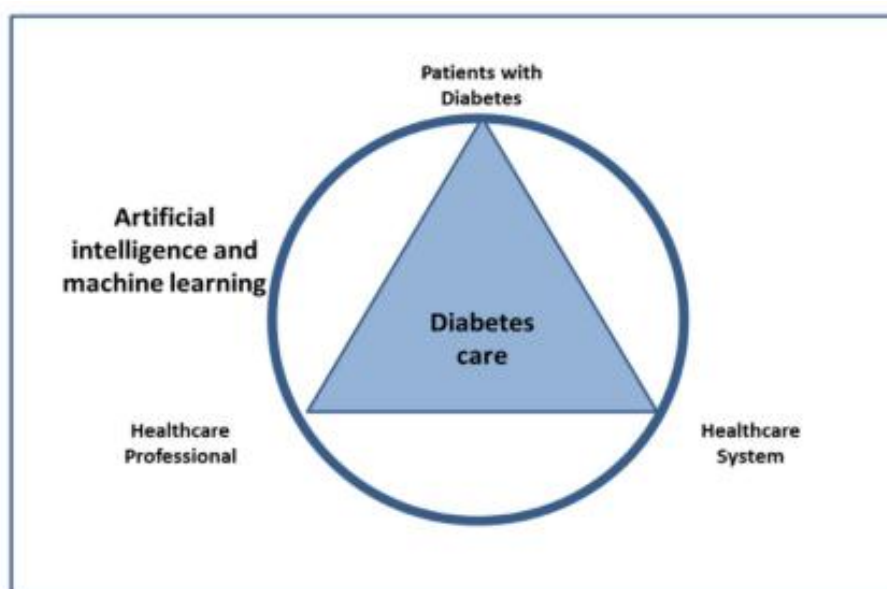


Figure 1: Scope of artificial intelligence and machine learning in diabetes care.

2.1 Techniques

Several artificial intelligence (AI) methods have been used to improve diabetes treatment. Diagnosing diabetes now involves a battery of tests in addition to the traditional ones, such as blood glucose and glycosylated haemoglobin, thanks to the introduction of AI. [4]

i. Case-based reasoning

CBR, an AI method for tackling novel problems by drawing on experience with similar cases, is seeing widespread application in the field of diabetes management. CBR has been implemented in diabetes care through tools like the Diabetes Support System. The system's intended function is to automatically recognise when a patient's blood

glucose levels are out of whack, make suggestions for how to fix the problem, and keep track of what works and what doesn't for that particular In diabetes, CBR has been used to tailor insulin treatment to each individual's metabolic profile and optimally manage insulin levels during mealtimes.

ii. Machine learning and deep learning

Diabetes digital support has been developed using a number of different machine learning processes. Among these are the k-nearest neighbour algorithm, the decision tree, the random forest, the classification and regression trees, the artificial neural network, and the naive Bayes classifier. Automated screening for blood glucose variability has been developed using machine learning. Feature selection methods (such as random forests,

logistic regressions, mutual information, principal component analyses, analysis of variance, and Fisher's discriminant ratio), outlier removal methods, cross-validation protocols, and classifiers (such as linear and quadratic discriminant analyses, naive Bayes, Gaussian process classifiers, support vector machines, artificial neural networks, Adaboost, and logistic regressions) are all part of the foundation of machine learning.

iii. Artificial neural networks

Connecting and analysing data from multiple sources to develop individualised answers is the goal of neural networks. Among the many areas where neural network methodology has been put to use, diabetes diagnosis stands out as particularly relevant and important. A number of factors can affect glycemic indices, and this effect has been studied using intelligent algorithms.[5]

2.2 Applications

i. Automated retinal screening

In order to speed up and improve accuracy of diabetic retinopathy diagnosis, deep learning algorithms have been created. Retinal screening using AI is a practical, reliable, and well acknowledged approach for detecting and monitoring diabetic retinopathy. Automated retinal screening has been shown to have a high sensitivity (92.3%) and specificity (93.7%). 96% of patients surveyed reported being happy or extremely satisfied with automated screening. Retinal haemorrhages, microaneurysms, exudates, neovascularization, and normal appearance have all been the subject of CNN training on small datasets.

ii. Clinical decision support

Clinical decision support tools using supervised machine learning for predicting short- and long-term HbA1c response following insulin initiation in patients with type 2 diabetes mellitus have been developed. Instruments like these also aid in the identification of clinical factors that may affect an individual patient's HbA1c response. It has been reported that the HbA1c response to insulin initiation can be reliably predicted using a generalised linear model based on baseline HbA1c and estimated glomerular filtration rate using elastic net regularization. HbA1c response over the short and long term were reported to have areas under the curve (AUC) of 0.80 (95% CI 0.780.83) and 0.81 (95% CI 0.790.84), respectively. An intuitive

method for tailoring interventions in medication adherence and predicting the risk of hospitalisation due to diabetes has been developed with the help of machine learning. Machine learning produced adherence thresholds of 46% to 94% as most discriminatory in a retrospective cohort study (n = 33,130) for risk of all-cause hospitalization. This research lends credence to the notion that predictive adherence thresholds vary by patient characteristics and medication complexity. [6]

iii. Predictive population risk stratification

By analysing a patient's way of life, physical health factors, mental health factors, and social network activities, a machine learning-based healthcare recommendation system (HRS) helped predict the risk for a disease like diabetes. With the use of big data analytics, forecasting models have been developed to help physicians and their patients with diabetes prepare for the unforeseen. Mobile apps that have been trained to interpret images of feet are now being used to monitor the progress of diabetic foot ulcers in patients who have diabetes.

Decision tree models have been developed using machine learning to predict whether or not a pregnant woman with gestational diabetes would go on to develop type 2 diabetes mellitus. This approach of prediction outperformed the standard practise of continuously monitoring fasting glucose levels, with a discriminative power of 83.0% in the training set and 76.9% in an independent testing set.

iv. Genomics

The creation of digital biomarkers, together with the use of advanced molecular phenotyping, genomics, epigenetic changes, and other techniques, represents a significant step forward in the detection and treatment of disease. This is useful for studying diabetes, a disease characterised by its heterogeneity and its chronic, protracted course, and for which enormous data sets are created. Information gleaned from the microbiome has been utilised to compile a bank of microbial marker genes that may be used to foretell diabetes risk and direct therapy in people with the disease. More than four hundred signals have been found by genome-wide association studies to possibly establish the genetic predisposition to diabetes. Pancreatic islet regulatory variations for refining diabetes-related signals have been predicted using convolutional neural network models trained on several genomewide mapping and regulatory epigenomic annotations.

v. Patient self-management

One of the most important parts of treating diabetes is teaching patients how to control their condition on their own. Patients may now monitor their own glucose levels, track their own statistics, and act as their own health experts thanks to advances in artificial intelligence.

- **Increased awareness:** People with diabetes can receive customised instruction through digital mediums. Web-based programmes and mobile and smartphone apps have increased public awareness and provided information about healthy eating and exercise routines. That's been especially helpful for keeping pregnant women's diabetes under control. It has been reported that 21 pregnant women who participated in a web-based intervention for the education and management of their diabetes benefited greatly.
- **Self-treatment:** Artificial intelligence empowers people with diabetes to make their own lifestyle choices in regards to diet and exercise. Patients have been able to monitor their diet quality and calorie intake with the help of apps. When patients with diabetes take pictures of their meals and evaluate their eating habits, they take greater responsibility for their diabetes care. [7]

3. BIG DATA

Big Data is an integration of social media, mobile apps, data analytics, and cloud storage, hence the acronym SMAC. Massive amounts of information are produced in many fields of inquiry, including scientific and industrial investigation. It's challenging to store and evaluate such rapidly growing data using conventional technology within the allotted time limit. Systems for managing and analysing data that can keep up with data deluge from various scientific and commercial applications are in great demand. Using the techniques, processes, and resources connected with big data, a corporation may create, change, and manage enormous amounts of data. Big Data is difficult for humans to process in bulk, hence a fast system is required to handle it. Big data technologies are being used to speed up data analysis and get users the insights they want for making decisions and generating projections.

Data that has been sitting in a warehouse for a while is updated and cleaned up. Raw data is compiled from a wide range of sources for the Big

Data system, such as reports, website monitoring, and real-time information. Analytics utilises Big Data to get insights from information stored in databases and other digital mediums. Conventionally derived data is only collected through other database systems like billing, point-of-sale, consumer marketing, and the company's finances. Data from sensors, e-mail, mobile devices, and other similar sources are used by Big Data systems. Analytics on large amounts of data can help solve problems that are becoming more complex.[8]

3.1 Big Data in Healthcare

Data volumes have expanded considerably in recent years as a result of the growing use of digital technology. Multiple factors, like the accessibility of high-speed mobile networks and the substantial digitalization of patient information, add up to an exponential increase in health care data. Variety of data types is a crucial part of healthcare Big Data. Medical prescriptions and hospital photographs are two examples of the kinds of data that might be found in these files. It is considered the Big Data problem when researchers try to manually examine and understand massive volumes of complicated data. Data management is crucial in the healthcare sector. Big data analytics in healthcare is particularly valuable to both patients and clinicians as it supports in the right treatment of patients. Big data presents difficulties due to the four V's: volume, velocity, variety, and veracity. Particularly helpful sources of data will be people with diabetes.

Due to the sheer number of available healthcare data, you'll require effective and dependable methods for handling this information. Researchers have shown that Big Data in healthcare may be used to foresee epidemics. Corporations are analysing the challenges of handling massive amounts of data. Gaining a deeper insight into physiological processes through data collection may pave the way for significant medical progress.[9]

4 DIABETES AND ITS COMPLICATIONS

Diabetes mellitus is a long-term condition brought on by either insufficient insulin production by the pancreas or resistance to the body's own insulin. The hormone insulin controls glucose levels. Increased glucose levels, or hyperglycemia, is another name for this condition. It's a common side effect of untreated diabetes that causes long-term

damage to many bodily systems, including the veins and nerves.

There are four distinct kinds of diabetes: Type 1, Type 2, prediabetes, and gestational diabetes. The inability of the body to produce insulin is what causes type 1 diabetes. In order for the body to use glucose from meals as an energy source, this hormone is required. Most people of this age range and younger are susceptible to this kind.

Type 2 diabetes is the most common form of the disease, accounting for 90–95 percent of all cases. Pre diabetes, also known as impaired glucose tolerance, is a less severe form of diabetes. When a

woman develops diabetes during pregnancy, it is called gestational diabetes. A mother's risk of developing diabetes after gestational diabetes is extremely elevated. The risk of obesity and diabetes in the newborn is also increased.[10]

In India, where over 62 million individuals have been diagnosed with diabetes, the disease has reached epidemic proportions. With 31.7 million cases of diabetes mellitus in 2000, India was far and away the leader, followed by China (20.8) and the United States (17.7). India's uncertain future is akin to the burden diabetes might one day place on the country.

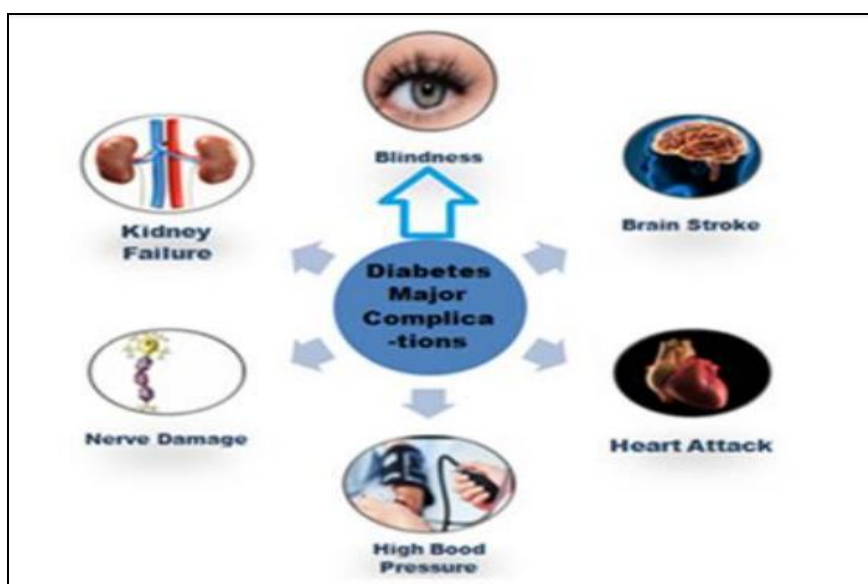


Figure 2: Diabetes major Complications:

5. BIG DATA ANALYTICS

Big Data is the massive amounts of information created daily by medical sources like doctors' notes, prescriptions, clinical reports, machine reading, and body sensors like IoT devices. In reality, healthcare workers face little resistance when it comes to Big Data. Extracting useful insights from massive data sets is difficult. The healthcare sector will be profoundly affected by big data analytics. The use of Big Data analytics in the medical field has the potential to help clinicians make more informed decisions, such as determining whether or not a patient will benefit from a given course of treatment. Healthcare sector analysis now necessitates Big Data computing environments hosted in the cloud. Semi-structured, structured, and unstructured Big Data.[11]

i. Cloud Computing

The cloud is a precious asset since it may be used by both current and future generations. Those working in IT and the sciences rely heavily on various computer infrastructure platforms. The internet has led to a Big Data explosion in many areas of research and development. Information collected from sources like the internet and scientific studies is typically stored as unstructured data.

Hardware and software are provided by service providers in data centres. Cloud computing services are made available through the Internet and are ideal for doing parallel data processing in a distributed setting in a short amount of time and at a cheap cost. Through the use of Cloud computing, large data sets may be seamlessly stored, gathered, and transported.

Amazon Web Services provides a remote service known as EMR, Elastic Map Reduce, to store and process massive volumes of data using the MapReduce system on the Amazon Cloud.

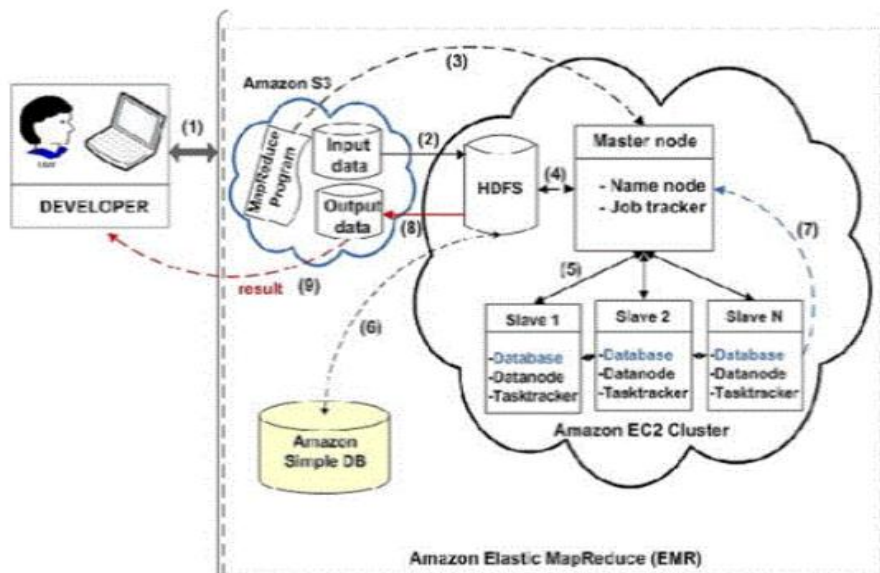


Figure 3: Amazon Elastic MapReduce (EMR)

ii. Hadoop Map Reduce

It takes a lot of computing power and storage space to extract useful information from massive amounts of data. Parallelism, also known as distributed computing, is the practise of utilising several central processing units (CPUs) or CPU cores to do many operations simultaneously. Essentially, the point of both parallel and distributed computing is to break down large problems into more manageable chunks that can be worked on in parallel, regardless of whether or not they are identical. Parallelization is affected by execution time, memory use, and concurrency. Pipelining, multiple arithmetic logic units, and several processors, each of which works on a distinct part

of the problem in the background, are all ways to speed up execution. To avoid bottlenecks while processing massive data sets, organisations are looking to parallel or distributed processing to utilise the combined memory of several computers. [12]

Having quick execution and ample storage space is essential for making sense of large amounts of data. It takes the massive amount of data and splits it up into thousands of pieces to be processed simultaneously by a large number of computers. Hadoop's HDFS and MapReduce are used to swiftly process and analyse massive volumes of data, whereas MapReduce is used to store just a small fraction of that data.

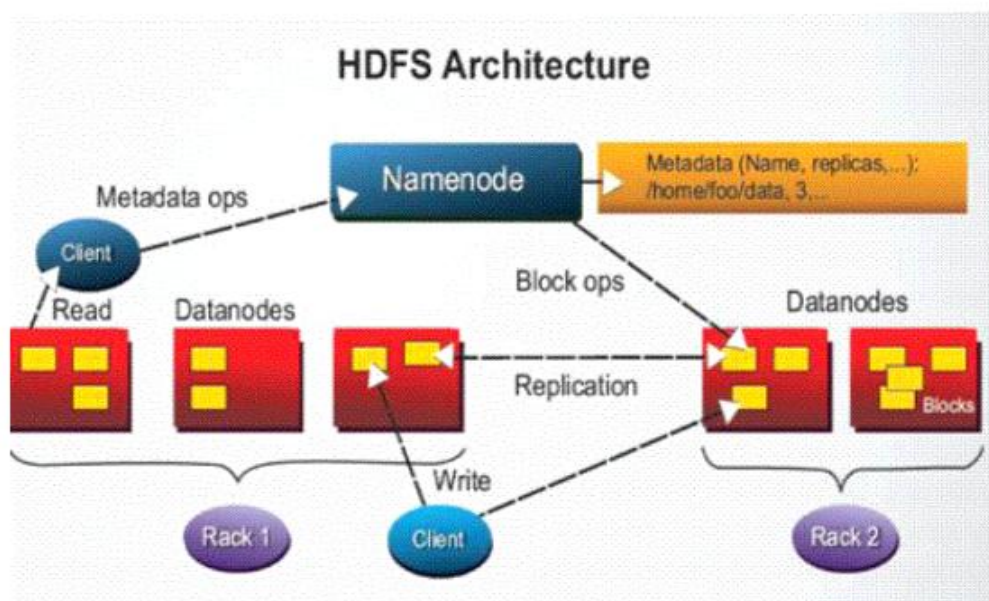


Figure 4: HDFS Architecture

In the Hadoop Distributed File System, as depicted in the preceding Figure 4, there are blocks of fixed size for each file. These blocks are kept in a computer or computers that make up a cluster. Using a Master/Slave setup, the system has one

NameNode (Masternode) and treats all other nodes as Data Nodes. Given these realities, the current investigation has utilised the Hadoop platform to arrange Big Data in an effort to solve the problem.[13]

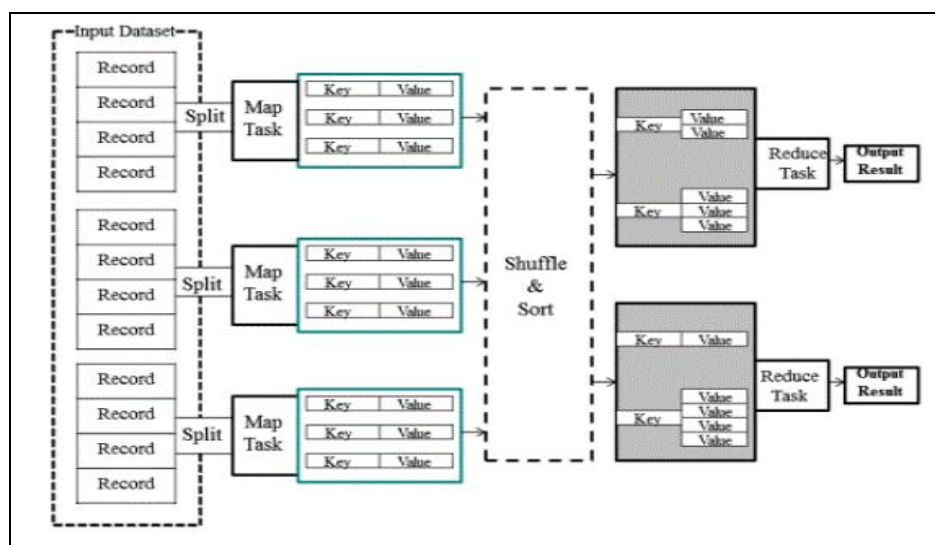


Figure 5: The MapReduce programming model

6. CONCLUSION

The healthcare industry may make extensive use of artificial intelligence and machine learning. Helping hospitals save time and money by automating administrative tasks, personalising patient care, and combating infectious diseases is a new and exciting use of AI and machine learning. Medical professionals and patients alike stand to benefit from data science studies that aim to improve diabetes mellitus diagnosis and diabetes type

prediction. Time is conserved throughout the process of creating a machine learning model for diabetes diagnosis. Patients are gaining more control over their health care, and doctors have better tools at their disposal than ever before to intervene quickly and precisely. These developments are efficient because they allow data to be gathered remotely, and virtual management is replacing frequent in-person clinic visits.

REFERENCES

1. Vikram Phaneendra, S. & E. Madhusudhan Reddy (2015). Big Data- solutions for RDBMS problems-A survey. In 12th IEEE/IFIP Network Operations & Management Symposium.
2. Ramirez, C., Nagappan, M. & Mirakhorli, M. (2015). Studying the impact of evolution in R libraries on software engineering research. Software Analytics (SWAN), IEEE 1st International Workshop on, pp. 29-30.
3. Kitcharoen, N., Kamolsantisuk, S., Angsomboon, R. & Achalakul, T. (2016). Rapid Miner framework for manufacturing data analysis on the Cloud. Computer Science and Software Engineering (JCSSE), 2013 10th International Joint Conference on, pp. 149-154.
4. Vale ZA, Ramos C, Ramos S, Pinto T. (2019). Data mining applications in power systems: Case-studies and future trends. 2009 Transm Distrib Conf Expo Asia Pacific: pp. 1-4. doi:10.1109/TD-ASIA.2009.5356830.
5. Johann, M. Kraus and Hans A Kestler (2017). "A Highly Efficient Multi-core Algorithm for Clustering Extremely Large Datasets." BMC Bioinformatics, 11: (169).
6. Kanchan M. Tarwani, Saleha S. Saudagar, Harshal D. Misalkar (2015). "Machine Learning in Big Data Analytics: An Overview", International Journal of Advanced Research in Computer Science and Software Engineering. 5 (4): pp. 270-274.
7. Sadhana and Savitha Shetty (2018). "Analysis of Diabetic Data Set Using Hive and R", International Journal of Emerging Technology and Advanced Engineering, 4(7).
8. Kiran kumara Reddi & DnvsI Indira (2017). "Different Technique to Transfer Big Data: survey" IEEE Transactions. 52(8).

9. Sadhana and Savitha Shetty (2017). Analysis of Diabetic Data Set Using Hive and R”, International Journal of Emerging Technology and Advanced Engineering, 4(7).
10. Guariguata, Leonor & Weil, Clara & Shaw, Jonathan. (2018). IDF Diabetes Atlas: Global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes research and clinical practice. 94. 311-21. 10.1016/j.diabres.2011.10.029.
11. Seuring T, Archangelidi O, Suhrcke M. (2015) The economic costs of type 2 diabetes: A global systematic review. Pharmacoeconomics, 33(8): 811–31.
12. Milley, A. (2018). Healthcare and data mining. Health Management Technology, 21(8), 44-47.
13. Christy, T. (2018). Analytical tools help health firms fight fraud. Insurance & Technology, 22(3), 22-26.