History of Medicine, 2022, 8(2): 628-637
DOI:10.48047.v8.2.2024.61

Transforming Diabetes Care: Leveraging Ensemble Machine Learning and Cloud Analytics for Personalized Diet Recommendations

Dr. Arun Elias[1], Chakka Balasruthi[1]
[1]Department of Computer Science Engineering,

[1]Sree Dattha Group of Institutions, Sheriguda, Hyderabad, Telangana

ABSTRACT

Millions of individuals across the world suffer with diabetes, a chronic medical condition. To avoid difficulties, it has to be managed and monitored continuously. Healthcare is experiencing a paradigm change as a result of technological advancements, utilizing data-driven techniques to improve illness identification, management, and customized treatment options. Big Data analytics offers the chance to examine enormous volumes of patient data and produce insightful findings for improved diabetic care, especially when applied to healthcare clouds. Traditionally, rule-based algorithms or basic machine learning models have been used in diabetes diagnosis and diet planning. These systems might not be able to appropriately represent the complex linkages found in the data or change to reflect the dynamic nature of patient care. Furthermore, they could not make the most of the potential of the vast amounts of healthcare data that are stored in cloud systems. However, the current approaches to diabetes detection and diet planning are frequently insufficiently sophisticated to manage the heterogeneity and complexity of patient data. Furthermore, analyzing and extracting useful information from cloud systems is challenging due to the massive amount of healthcare data accessible. A more comprehensive and complex system is required to improve diabetes management's accuracy, efficiency, and customisation. As a result, there is a rising demand for sophisticated analytics and machine learning methods to enhance diabetes detection accuracy and offer individualized meal plans catered to the requirements of each patient. Therefore, the goal of this study is to develop a cloud model and user interface for diabetes diagnosis and diet planning by utilizing an ensemble architecture, which has the potential to significantly transform healthcare analytics. Ensemble frameworks are superior to individual models in collecting complex patterns in data, which leads to better diabetes prediction accuracy and more customized diet programs. Ensemble models are especially well-suited for the dynamic and varied nature of healthcare big data clouds due to their greater durability and flexibility. Ensemble frameworks' scalability makes it possible to analyze massive amounts of healthcare data effectively, enabling real-time analytics and decision-making. Fundamentally, the importance resides in the revolutionary capacity to improve the accuracy, flexibility, and effectiveness of diabetes treatment, which will eventually result in improved patient outcomes and care.
Keywords: Patient data, healthcare data, machine learning, ensemble framework, diabetes management, personalized diet plans, and healthcare clouds.

1. INTRODUCTION

Diabetes has been a persistent health concern throughout human history, with records dating back to ancient civilizations. Early medical texts, such as those from ancient Egypt and Greece, describe symptoms resembling diabetes and propose treatments involving dietary modifications and herbal remedies. However, it was not until the 20th century that significant advancements were made in understanding the physiology and treatment of diabetes. The discovery of insulin in the 1920s marked a turning point in diabetes management, allowing individuals with diabetes to regulate their blood sugar levels effectively. Over the decades, medical research and technological innovation have continued to shape the landscape of diabetes care. The advent of glucometers in the mid-20th century revolutionized blood glucose monitoring, providing patients with a means to track their blood sugar levels at home. Subsequent developments in pharmaceuticals, such as oral hypoglycemic agents and insulin analogs, have further improved diabetes management and control.

In recent years, the proliferation of digital health technologies and the rise of big data analytics have ushered in a new era of diabetes care. Electronic health records (EHRs), wearable devices, and mobile health applications have enabled the collection of vast amounts of patient data, offering unprecedented insights into disease patterns and treatment outcomes. Additionally, advances in machine learning and artificial intelligence have opened up new possibilities for leveraging healthcare data to optimize disease detection, management, and personalized treatment strategies. The motivation behind the development of an ensemble framework-based diabetes detection and diet plan suggestion system stems from several key factors. Firstly, the rising prevalence of diabetes globally underscores the need for innovative approaches to disease detection and management. With millions of individuals affected by diabetes, there is a growing imperative to harness the power of technology and data analytics to improve patient outcomes and quality of life. Furthermore, the limitations of existing diabetes detection and diet planning methods highlight the need for more sophisticated and personalized approaches. Rule-based systems and simple machine learning models may not adequately capture the complexity of diabetes data or adapt to individual patient needs. As such, there is a compelling motivation to explore advanced analytics techniques, such as ensemble frameworks, to enhance the accuracy and efficacy of diabetes management strategies.

2. LITERATURE SURVEY

Kaur et al. have introduced a cloud IoT-based framework named CI-PDF for diabetes prediction considering accuracy, sensitivity, and specificity as evaluative parameters on the PIDD dataset and claimed to have achieved 94.5% of prediction accuracy by combining neural network (NN) and DT approaches [1]. Priyadarshini et al. have presented DeepFog, a fog computing-based deep neural architecture for predicting stress type, diabetes, and hypertension attacks using standard datasets and open-source software tools, and claimed to have achieved a superior and competitive method in comparison to others . Fernández-Caramés and Fraga-Lamas have introduced an IoT continuous glucose monitor- (CGM-) based system that claims to offer a translucent and truthful blood sugar data source from a population in a quick, flexible, scalable, and low-cost manner by accessing the collected blood sugar samples and warning them in the case of a dangerous situation being detected [2]. Barik et al. have introduced FogLearn, a fog computing-based framework for the application of -means clustering in Ganga River Basin Management and real-world feature data for detecting diabetes patients suffering from diabetes mellitus and found that fog computing holds a lot of promise for medical and geospatial big data analysis [3]. Fernández-Caramés et al. have created and implemented a system that improves commercial CGMs in terms of IoT capabilities, allowing them to monitor patients remotely and alert them about the severity of their conditions. And they claimed to have developed a better technique for diagnosing patients' illnesses remotely in real time [4]. Gia et al. have developed a fog-based structure for remote health monitoring and fall detection. The system provides numerous progressive amenities such as ECG feature extraction, security, and locally distributed storage. In addition, the system operates accurately, and the wearable sensor node is energy efficient [5]. Devarajan et al. proposed an energy-efficient fog-assisted healthcare system that manages glucose levels based on evaluative measures such as energy efficiecy, prediction accuracy, computational complexity, and latency on two datasets from the UCI repository diabetes dataset and the Physical Activity Monitoring Dataset (PAMAP2). The experimental results show that fog over cloud computing has increased bandwidth efficiency, reduced latency, and enhanced accuracy [6]. Abdel-Basset et al. have suggested a novel framework based on computer-propped diagnosis and IoT to detect and observe type 2 diabetes patients and indicated the validity and robustness of the proposed algorithms considering accuracy and execution time as the performance evaluators [7]. Haq et al. have developed a filter method based on the DT-ID3 (Iterative Dichotomiser 3) model for essential feature selection in comparison to two ensemble learning algorithms, Ada Boost and RF, using prediction accuracy and computation time as evaluative measures, and found that the DT algorithm based on selected features improves the classifier's performance [8]. Kumari et al. have proposed an ensemble voting classifier that uses the ensemble of three ML algorithms, viz., LR, NB, and RF for the classification considering the evaluative measures like accuracy, precision, recall, and -score on PIDD and claimed to have achieved comparatively enhanced results on binary classifications [9]. Geetha and Prasad have built a hybrid model named T2DDP that doctors can effectively use to treat diabetic patients by employing supervised classification algorithms such as NB and ensemble algorithms like bagging with RF and AdaBoost for DT and found that the forecast will be submitted to the patient's cell phone at an early stage to make the immediate decisions about the health risk [10].

3. PROPOSED SYSTEM

The Ensemble Framework-based Diabetes Detection and Diet Plan Suggestion system aims to provide a user-friendly interface for individuals to monitor their health status and receive personalized diet recommendations. By leveraging machine learning algorithms and cloud computing, the system offers accurate predictions regarding the likelihood of type 2 diabetes based on user data.

User Side:

The user side interface enables users to upload their health data files and receive real-time predictions from the cloud server. Upon selecting a file, the system reads the data and sends it to the server via a socket connection. It also retrieves a diet plan from a local file for users with abnormal health indicators.

The interface displays the sent data and predictions received from the server, allowing users to track their health status and receive dietary recommendations if necessary. This component provides a seamless experience for users to engage with the system and take proactive measures towards improving their health.

Cloud Side:

The cloud side application serves as the backend of the system, handling data processing, model training, prediction, and communication with the user side. It offers functionalities for preprocessing datasets, training machine learning models, and evaluating their performance.

The system supports multiple machine learning algorithms, including Decision Tree, SVM, and ANN, to analyze user data and make predictions regarding type 2 diabetes. An Ensemble Classifier model is also implemented, which combines the predictions of individual models to improve accuracy.

The cloud side application acts as a server, listening for incoming connections from the user side. Upon receiving data, it performs predictions using the trained models and sends back the results. It

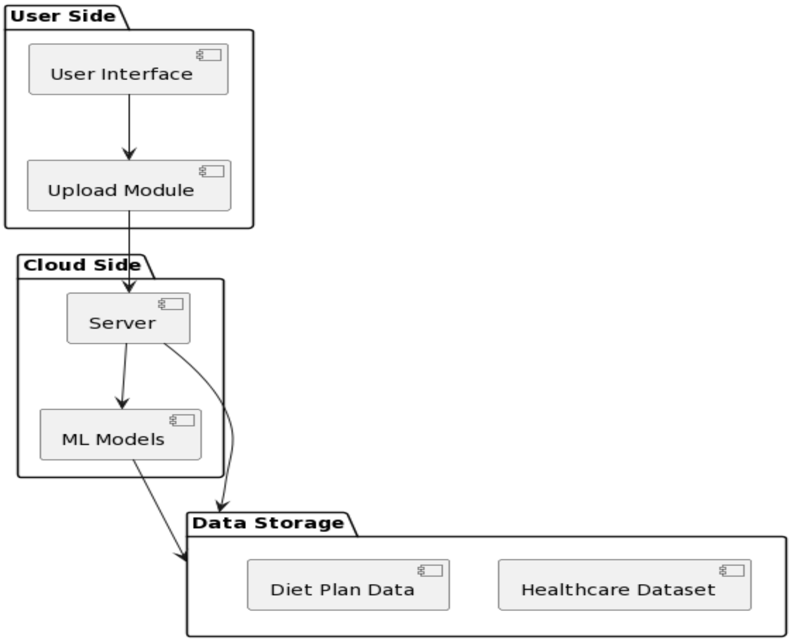

Figure 1 Presents the Block Diagram of Proposed System.

Support Vector Machine Algorithm

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

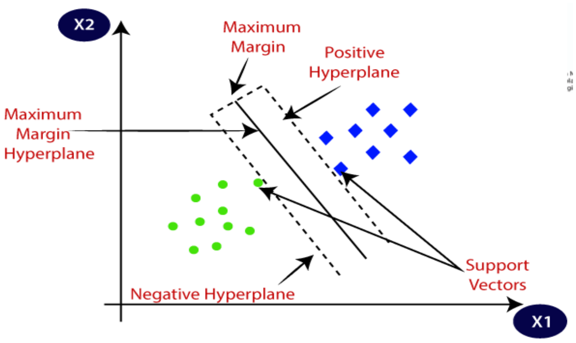

Figure 2 Analysis of SVM

SVM working

Linear SVM: The working of the SVM algorithm can be understood by using an example. Suppose we have a dataset that has two tags (green and blue), and the dataset has two features x1 and x2. We want a classifier that can classify the pair (x1, x2) of coordinates in either green or blue. Consider the below image:

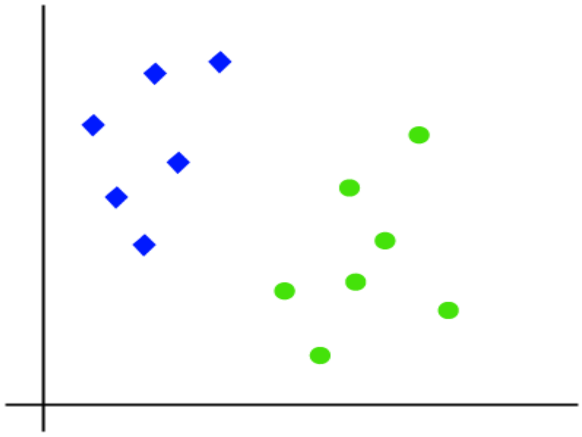

Figure 3 Linear SVM

So as it is 2-d space so by just using a straight line, we can easily separate these two classes. But there can be multiple lines that can separate these classes. Consider the below image:

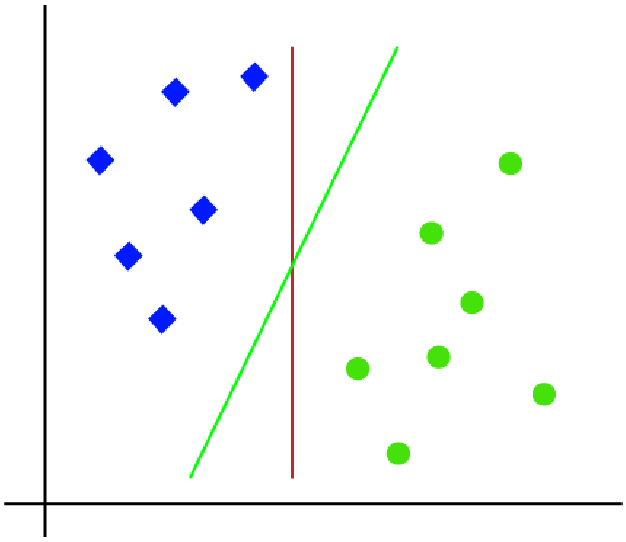

Figure 4. Test-Vector in SVM

Hence, the SVM algorithm helps to find the best line or decision boundary; this best boundary or region is called as a hyperplane. SVM algorithm finds the closest point of the lines from both the classes. These points are called support vectors. The distance between the vectors and the hyperplane is called as margin. And the goal of SVM is to maximize this margin. The hyperplane with maximum margin is called the optimal hyperplane.

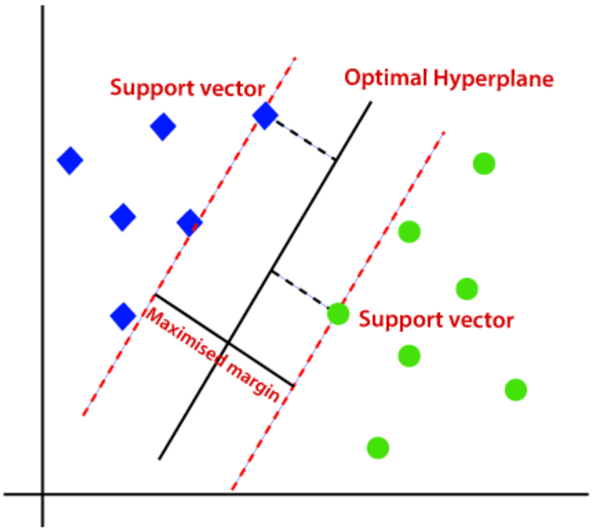

Figure 5. Classification in SVM

Non-Linear SVM: If data is linearly arranged, then we can separate it by using a straight line, but for non-linear data, we cannot draw a single straight line. Consider the below image:

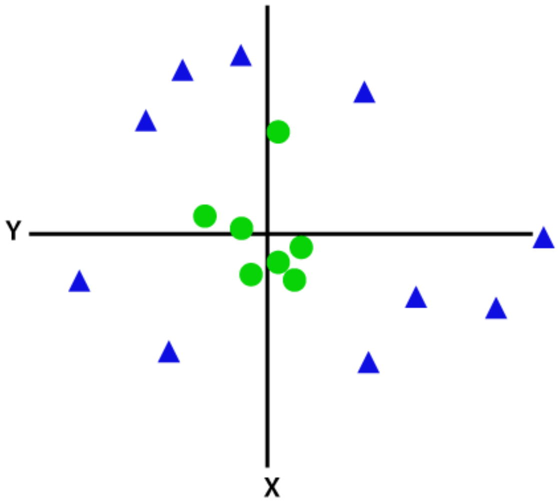

Figure 6 Non-Linear SVM

Hence, we get a circumference of radius 1 in case of non-linear data.

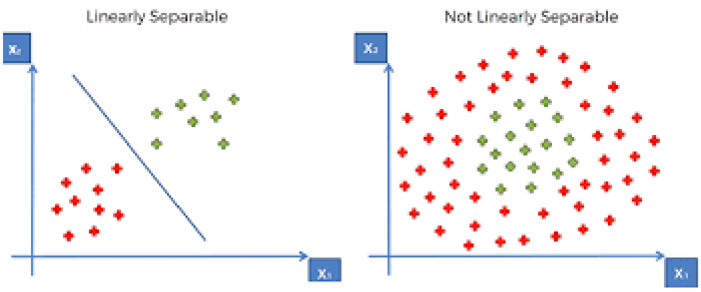

Figure 7: Difference between linear & Non-linear

4. RESULTS

Figure 8 displays the accuracy of various machine learning models utilized in the Diabetes Detection system: Decision Tree Classifier (DTC), Support Vector Machine (SVM), Artificial Neural Network

(ANN), and an Ensemble model. Accuracy is a crucial performance metric that measures the proportion of correct predictions made by each model. In this figure, accuracy values for each model are presented, allowing users to assess and compare their effectiveness in predicting diabetes. A higher accuracy value indicates better performance, highlighting the reliability and predictive power of the model.

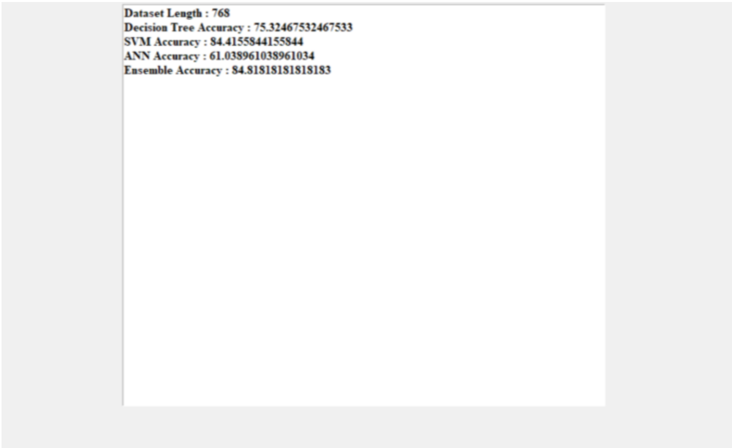

Figure 8: Presents accuracy of DTC, SVM, ANN, Ensemble models.

Figure 9 presents a performance comparison graph that visually depicts the performance metrics of all machine learning models used in the Diabetes Detection system. This graph provides a comprehensive overview of each model's performance, enabling users to identify trends and disparities among them. Performance metrics such as accuracy, precision, recall, and F1-score are typically evaluated and plotted on the graph. By presenting these metrics in a graphical format, users can easily interpret and compare the effectiveness of different models in predicting diabetes.

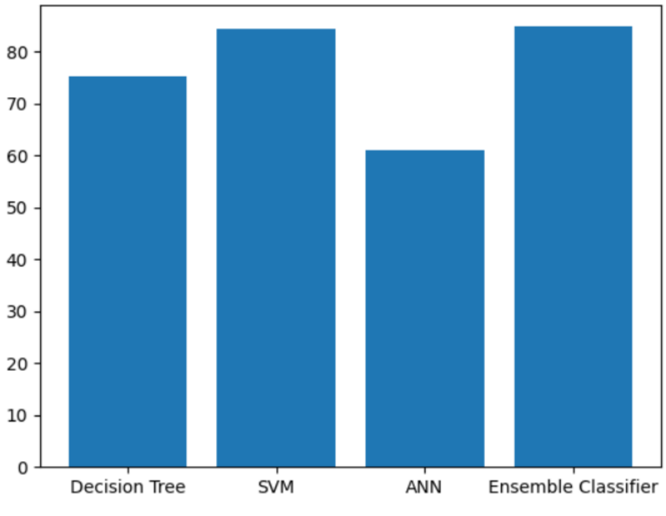

Figure 9: Presents the Performance Comparison graph of all models.

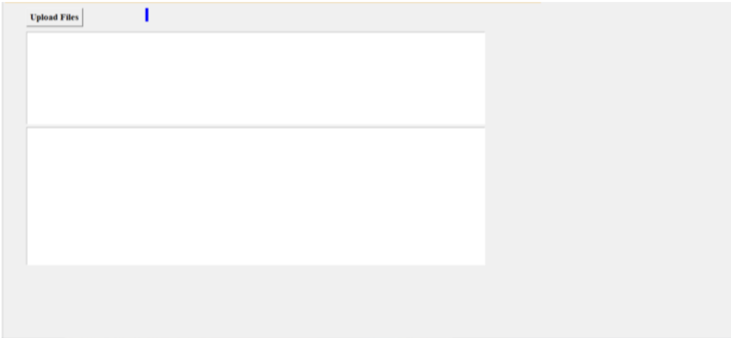

Figure 10: Shows the User side GUI.

Figure 10 showcases the graphical user interface (GUI) of the user side component in the Diabetes Detection system. The user side GUI serves as the primary interaction platform for users to upload their healthcare data, initiate predictive analysis, and view results. It is designed with user-friendly elements such as buttons, text fields, and labels arranged in an intuitive layout. This GUI ensures a seamless and accessible user experience, guiding users through the process of data upload and analysis effectively.

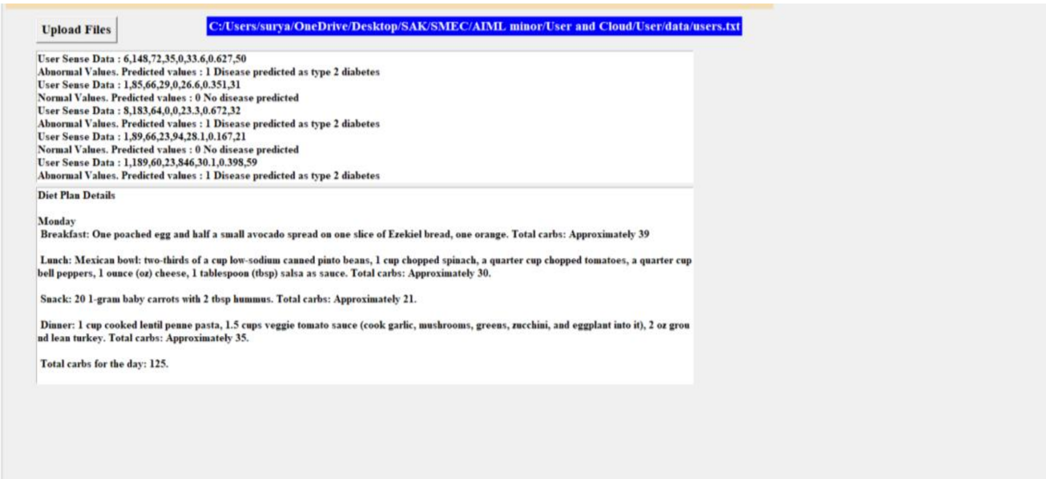

Figure11: Model prediction on uploaded test data in user side GUI.

Figure 11 demonstrates the model prediction process on the uploaded test data within the user side GUI. Once the user uploads their healthcare dataset, the system utilizes machine learning models to predict the likelihood of diabetes based on the provided data. Prediction results are displayed in real-time on the GUI, allowing users to instantly access and interpret the outcomes. This figure provides users with actionable insights into their health status, empowering them to make informed decisions about their well-being.

5. CONCLUSION

In conclusion, the development of an ensemble framework-based system for diabetes detection and diet plan suggestion represents a significant advancement in healthcare analytics. Diabetes, being a chronic health condition affecting millions worldwide, requires continuous monitoring and personalized management to prevent complications. With the evolution of technology, particularly the advent of Big Data analytics and cloud computing in healthcare, there emerges an unprecedented opportunity to leverage vast amounts of patient data for better disease management.

Traditionally, diabetes detection and diet planning have relied on rule-based systems or simplistic machine learning models, which not adequately capture the complexity of patient data or adapt to dynamic health conditions. However, the adoption of ensemble frameworks in this context offers a promising solution. By integrating multiple models and algorithms, ensemble frameworks excel in capturing nuanced data patterns, leading to improved accuracy in diabetes prediction and the creation of personalized diet plans tailored to individual patient needs.

Furthermore, the scalability of ensemble frameworks enables the efficient processing of large volumes of healthcare data, facilitating real-time analytics and decision-making. This transformative potential holds profound significance in revolutionizing healthcare analytics, ultimately leading to enhanced patient care and outcomes in diabetes management.

REFERENCES

[1] T. M. Fernández-Caramés and P. Fraga-Lamas, “Design of a fog computing, blockchain and IoT-based continuous glucose monitoring system for crowdsourcing mHealth,” Multidisciplinary Digital Publishing Institute Proceedings, vol. 4, no. 1, p. 37, 2018.

[2] P. Kaur, N. Sharma, A. Singh, and B. Gill, “CI-DPF: a cloud IoT based framework for diabetes prediction,” in 2018 IEEE 9th annual information technology, Electronics and Mobile Communication Conference (IEMCON), pp. 654–660, Vancouver, BC, Canada, 2018.

[3] R. K. Barik, R. Priyadarshini, H. Dubey, V. Kumar, and K. Mankodiya, “FogLearn,” International Journal of Fog Computing (IJFC), vol. 1, no. 1, pp. 15–34, 2018.

[4] T. N. Gia, I. B. Dhaou, M. Ali et al., “Energy efficient fog-assisted IoT system for monitoring diabetic patients with cardiovascular disease,” Future Generation Computer Systems, vol. 93, pp. 198–211, 2019.

[5] M. Devarajan, V. Subramaniyaswamy, V. Vijayakumar, and L. Ravi, “Fog-assisted personalized healthcare-support system for remote patients with diabetes,” Journal of Ambient Intelligence and Humanized Computing, vol. 10, no. 10, pp. 3747–3760, 2019.

[6] M. Abdel-Basset, G. Manogaran, A. Gamal, and V. Chang, “A novel intelligent medical decision support model based on soft computing and IoT,” IEEE Internet of Things Journal, vol. 7, no. 5, pp. 4160–4170, 2019.

[7] A. U. Haq, J. P. Li, J. Khan et al., “Intelligent machine learning approach for effective recognition of diabetes in E-healthcare using clinical data,” Sensors, vol. 20, no. 9, p. 2649, 2020.

[8] S. Kumari, D. Kumar, and M. Mittal, “An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier,” International Journal of Cognitive Computing in Engineering, vol. 2, pp. 40–46, 2021.

[9] G. Geetha and K. M. Prasad, “An hybrid ensemble machine learning approach to predict type 2 diabetes mellitus,” Webology, vol. 18, no. Special Issue 02, pp. 311–331, 2021.

[10] P. G. Shynu, V. G. Menon, R. L. Kumar, S. Kadry, and Y. Nam, “Blockchain-based secure healthcare application for diabetic-cardio disease prediction in fog computing,” IEEE Access, vol. 9, pp. 45706–45720, 2021.