Annals of Emerging Technologies in Computing (AETiC)
Vol. 9, No. 3, 2025

Research Article

Comparative Study of ML-Based Diabetes Detection Using IoT and Lab Data in Fog

Edmira Xhaferra[1], Florije Ismaili[1] and Elda Cina[2]

[1]Faculty of Contemporary Science and Technology, South East European University, North Macedonia
ex29824@seeu.edu.mk; f.ismaili@seeu.edu.mk
[2]College of Engineering and Technology, American University of the Middle East, Kuwait
Elda.Cina@aum.edu.kw
*Correspondence: Elda.Cina@aum.edu.kw

Received: 3rd October 2024; Accepted: 14th June 2025; Published: 1st July 2025

Abstract: Diabetes, as a chronic condition affecting millions of people worldwide, requires early diagnosis and continuous monitoring to prevent complications. The rise of machine learning (ML) applications in healthcare offers promising approaches for diagnosing and managing diabetes more effectively. Machine learning models can analyse extensive amounts of data to identify patterns that may be invisible to human clinicians, improving diagnosis accuracy and enabling personalized care. This study investigates the performance of four machine learning models—Decision Tree, Logistic Regression, Random Forest, and Support Vector Machine (SVM)—in detecting diabetes using two types of data: traditional lab-based data and real-time accessed data from Internet of Things (IoT) sensors. Data was collected from continuous glucose monitors (CGMs) and wearables, as well as clinical lab records in Albania. The results revealed that machine learning models applied to IoT data significantly outperformed those applied to lab data, demonstrating higher accuracy and better predictive metrics. The continuous monitoring enabled by IoT devices allows for real-time detection of glucose fluctuations, providing earlier and more precise diabetes diagnosis. Additionally, integrating IoT with fog computing reduces latency and enhances on-time decision-making, allowing for prompt interventions in patient care. The study highlights the transformative potential of combining IoT, machine learning, and fog computing to revolutionize healthcare, particularly the management of chronic diseases such as diabetes. The findings suggest that IoT-based systems should be adopted to improve diabetes detection and monitoring, allowing for a shift toward proactive healthcare solutions. Future research could explore the application of these technologies for managing other chronic conditions and optimizing machine-learning models for large-scale datasets.

Keywords: ***Continuous Glucose Monitoring; Diabetes Detection; Fog Computing; IoT; Machine Learning models***

---

1. Introduction

One of the prominent global health challenges today is Diabetes mellitus (DM), impacting millions of individuals across the world. Reports from the World Health Organization (WHO) indicate a significant rise in diabetes cases, with estimates suggesting that by 2030, around 578 million adults will be affected [1], [2]. This chronic disease displays high blood sugar levels, arising from inadequate insulin production or the body's inability to use insulin efficiently. If left unmanaged, diabetes can lead to serious complications, including cardiovascular disorders, kidney dysfunction, and nerve damage [3]. Alarmingly, nearly half of those with diabetes are unaware of their condition, underscoring the immediate need for effective screening and rapid diagnosis frameworks [1]. The increasing prevalence of diabetes, coupled with its associated complications, calls for innovative approaches in healthcare, particularly in diagnostics and disease management.

Machine learning (ML) is transforming the healthcare industry, providing new opportunities to improve disease detection, diagnosis, and treatment. By analysing large volumes of healthcare data, ML algorithms can recognize patterns and generate predictions with high accuracy, often surpassing human

Edmira Xhaferra, Florije Ismaili and Elda Cina, "Comparative Study of ML-Based Diabetes Detection Using IoT and Lab Data in Fog", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 1-21, Vol. 9, No. 3, 1st July 2025, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2025.03.001, Available: http://aetic.theiaer.org/archive/v9/v9n3/p1.html.

capabilities [4]. In the context of diabetes, ML applications have shown considerable promise, especially in classifying blood glucose patterns and detecting anomalies in real-time [3]. These innovations not only enhance diagnostic precision but also enable the development of personalized treatment strategies designed to meet the specific needs of every patient [5].

Exploiting machine learning capabilities into healthcare systems marks a significant breakthrough in addressing the complexities of modern medical challenges. Real-time data processing is crucial in the early detection and management of diabetes. The Internet of Things (IoT) has transformed the healthcare monitoring process by enabling continuous tracking and collecting patients' health data through connected devices [6-7]. These IoT devices gather and transmit data on critical physiological indicators, including blood glucose levels, heart rate, and physical activity, facilitating timely interventions when irregularities are detected [8]. This capability is further enhanced by fog computing, which computes data processing nearer to the source, decreasing latency and increasing decision-making speed [7]. By leveraging real-time IoT-driven data and fog computing, healthcare professionals can adopt proactive strategies to mitigate diabetes progression and related complications, ultimately leading to better patient outcomes.

IoT is instrumental in managing diabetes, with devices like continuous glucose monitors (CGMs) offering real-time feedback to both patients and healthcare professionals. This immediate access to data allows for prompt adjustments to treatment plans based on current glucose levels [9]. This continuous monitoring is particularly beneficial for patients diagnosed with Type 1 diabetes, due to the constant need to maintain stable levels of blood sugar [10]. Moreover, the integration of machine learning algorithms with IoT-generated data strengthens predictive analytics, enabling healthcare providers to detect patterns and foresee potential health risks before they develop into serious complications [11]. This combination of real-time data collection and advanced analytics fosters a proactive healthcare approach, shifting the focus from reactive to preventive care.

As diabetes continues to rise globally, the need for timely and accurate diagnostics becomes more pressing. Traditional lab tests, while effective, often involve delays in processing and result delivery, which can hinder prompt intervention and disease management. In contrast, IoT-based systems enable continuous monitoring, allowing for immediate data collection and analysis, crucial for early detection and management of diabetes [12-14]. The main objective of this study is to assess and compare the efficacy of IoT-based real-time data processing with traditional laboratory data for early diabetes detection. This comparison seeks to highlight the potential of real-time data in enhancing diagnostic accuracy and improving patient outcomes.

This study evaluates four popular machine learning algorithms to compare their effectiveness in classifying patients with diabetes, including Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM). These algorithms have proven effective in handling classification tasks within healthcare datasets. Logistic Regression known for its clarity and strength in binary classification, is particularly useful for preliminary diabetes risk assessments [15]. Decision Trees are preferred for their comprehensibility, allowing healthcare providers to easily understand the decision-making process happening behind predictions, which is essential in clinical settings [16]. Random Forest, an ensemble method, enhances prediction accuracy by combining results from multiple decision trees, reducing the likelihood of overfitting [17]. Lastly, SVM is recognized for its strong performance in high-dimensional data, making it especially appropriate for complex healthcare datasets [18]. By leveraging these diverse ML learning models, the study seeks to determine the most efficient method for detecting diabetes through real-time IoT data analysis.

Beyond utilizing machine learning algorithms, this study integrates fog computing to improve real-time data processing closer to its origin, by extending cloud computing capabilities to the network's edge, thus significantly reducing latency and speeding up analysis [19]. This approach is particularly beneficial in healthcare settings where prompt decision-making is essential. By leveraging fog computing, the study demonstrates how real-time data processing can facilitate immediate alerts and interventions for patients at risk of diabetes, thereby improving healthcare delivery [20]. Using the advantages of both fog computing and IoT devices enhances resource utilization and, at the same time, ensures that healthcare providers can swiftly access and respond to critical health data overcoming the delays often associated with conventional cloud processing [21].

In summary, this study contributes to the field as follows:

- It proposes an integrated fog-IoT architecture for continuous glucose monitoring, demonstrating how real-time data processing and machine learning models enhance diagnostic accuracy and enable early detection of diabetes.
- It evaluates the effectiveness of edge-level processing through fog computing in improving decision-making and proactive management of diabetes, thereby contributing to the advancement of smart healthcare systems for chronic disease management.

2. Literature Review

The application of machine learning algorithms in healthcare diagnostics, particularly for diabetes detection, has attracted considerable attention. This review examines four widely recognized machine learning algorithms: Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM). These algorithms have been widely studied for their effectiveness in predicting diabetes risk and diagnosing the disease using various health datasets.

Logistic Regression has served as a foundational model in medical diagnostics due to its simplicity and interpretability, making it highly convenient for binary classification purposes such as determining the presence or absence of diabetes. The authors of [22] demonstrated the utility of Logistic Regression in predicting Type 2 diabetes. They validated its effectiveness in identifying individuals at-risk based on health indicators. While another study [23] found that logistic regression provided a strong baseline, and they highlighted that more complex models, such as ensemble methods, yielded higher accuracy and precision for diabetes risk classification. This indicates that while Logistic Regression offers a straightforward approach, it may not be the most optimal for more intricate datasets.

Decision Trees are widely adopted in healthcare diagnostics because of their intuitive structure and ease of interpretation. They allow the visualization of decision-making processes, which is particularly valuable in clinical settings. Decision Trees were employed to classify diabetes status in a population from Northeastern Ethiopia, revealing that while Decision Trees were effective, their susceptibility to overfitting could limit their predictive power [24]. A comparison of Decision Trees with other algorithms was conducted in [25], concluding that integrated methods, ex. Random Forest, often outperform Decision Trees due to their ability to aggregate multiple trees and reduce overfitting.

Random Forest, an ensemble method, has gained prominence for its robustness and high accuracy in classification tasks. By constructing multiple decision trees and aggregating their outputs, Random Forest improves prediction accuracy and minimizes overfitting. The authors of [26] achieved 85% accuracy rate in their study while diagnosing diabetes, utilizing the Random Forest algorithm, by outperforming Decision Trees and Naive Bayes. Similarly, the authors of [27] used Random Forest to create a diabetes risk assessment model, which demonstrated better accuracy than Logistic Regression. Studies indicate that Random Forest excels in managing high-dimensional datasets while offering valuable insights into feature significance, establishing it as a crucial tool for advancing healthcare analytics.

SVM performs exceptionally well with high-dimensional data and demonstrates strong resistance to overfitting, especially in scenarios where the number of features surpasses the number of samples [28-29]. A study from [30] compared SVM with Logistic Regression and Decision Trees, finding that SVM achieved superior performance in classifying diabetes cases. Another study reported that SVM achieved an accuracy of 80.8% in diabetes prediction, reinforcing its effectiveness as a diagnostic tool [31].

Recent advancements in data processing have enhanced the application of ML algorithms in healthcare. Hybrid models that combine multiple algorithms have shown improved prediction accuracy. An ensemble approach was explored in [32], including Random Forest and SVM, which resulted in enhanced diabetes detection outcomes. This approach leverages the strengths of individual algorithms, optimizing overall performance in complex health data scenarios.

Machine learning models have been tested in diverse health datasets, extending beyond traditional clinical data. Ijaz et al. A new hybrid model incorporating Random Forest was proposed to predict diabetes and hypertension using various health datasets, including wearable device data [33]. This underscores the adaptability of ML techniques in analysing complex health data from different sources.

Integrating IoT devices in healthcare monitoring has revolutionized patient data collection, providing real-time data processing and analysis that enhances patient care [34-37]. While traditional lab-based approaches are effective for diagnosis, they often suffer from delays in data processing and result reporting. In contrast, IoT-based systems enable healthcare professionals to collect data for health metrics in real-time, facilitating prompt interventions and enhancing patient care outcomes.

Patients are under a continuous data collection process of parameters like heart rate, blood pressure, and glucose levels through wearable sensors and smart health monitors, offering a comprehensive view of their health status [12]. This continuous monitoring allows healthcare professionals to detect potential triggering flags before they escalate, improving preventive care. The authors of [38] developed a wearable IoT sensor-based system that demonstrated the potential of IoT in identifying and controlling infectious diseases, emphasizing the transformative effect of IoT in health monitoring.

However, the enormous amount of generated data by continuous sensor monitoring poses challenges in processing and analysis, particularly concerning latency and bandwidth limitations. Traditional cloud computing struggles to manage these challenges efficiently. The utilization of fog computing exploits cloud capabilities to the network's edge, allowing localized data processing, and thereby reducing latency and bandwidth constraints [39]. This architecture is especially suited for real-time healthcare applications where prompt decision-making is critical [40].

Several studies highlight the effectiveness of combining IoT data collection with fog computing processing capabilities in healthcare. The importance of fog computing in healthcare is underscored by the fact that it enhances the efficiency of data acquisition and processing [41]. A "multi-agent fog computing model for managing critical healthcare tasks" was developed in [42], demonstrating how fog computing can enhance service efficiency by enabling low-latency data processing. Another study explored "the integration of blockchain and fog computing in IoT-driven healthcare services", improving data security and interoperability while enhancing overall service delivery [13].

The combination of IoT and fog computing also facilitates predictive analytics through machine learning algorithms. By applying machine learning to real-time health data, healthcare providers can identify patterns and make informed predictions about patient health. An intelligent fog-enabled smart healthcare system that integrated machine learning for real-time physiological parameter detection was developed [43]. Their study demonstrated how machine learning, coupled with IoT and fog computing, leads to more accurate health assessments and timely interventions. This contrasts with traditional lab-based methods, where data analysis is often delayed, potentially missing opportunities for early intervention.

The challenges associated with traditional healthcare systems, particularly latency and bandwidth issues during data analysis, are highlighted in [39] and [44]. The authors of [45] proposed a fog-enabled framework that significantly reduced latency and improved real-time medical care, underscoring the potential of these technologies to enhance patient outcomes. Similarly, in [46], the importance of real-time health tracking in healthcare, demonstrating how fog computing enables immediate responses to patient health changes, which traditional lab-based systems often cannot provide, is emphasized.

Moreover, IoT and fog computing have been shown to improve Quality of Service (QoS) in healthcare applications. A demonstration of how fog computing can bridge the gap between IoT devices and analytics enhances the efficiency of healthcare monitoring systems [47]. This is particularly crucial in emergency situations, where timely access to patient data can be critical for saving lives. By processing data closer to the source, fog computing reduces the time required for data transmission and analysis, enabling faster decision-making and improving patient care.

An investigation of diverse machine-learning methodologies for predicting diabetes onset indicates that unsupervised and deep-learning approaches could enhance prediction precision [48]. This corresponds with another study's findings, which indicate that optimising machine learning via feature selection and dimensionality reduction can improve early detection accuracy and healthcare management techniques [49]. The capacity to analyse extensive datasets proficiently facilitates more nuanced insights into diabetes dynamics, enabling personalised treatment solutions.

The authors of [50] created an intelligent healthcare system employing ensemble classification models, attaining an accuracy of 82.2% in predicting type 2 diabetes. Their research highlights the efficacy of integrating various machine-learning methods to improve forecast accuracy. The incorporation of IoT data into these models facilitates real-time monitoring and data collection, which is essential for prompt interventions in diabetes treatment. Another IoT architecture was proposed to utilise machine learning for diabetes detection, highlighting the significance of continuous patient monitoring and data analysis [51]. The classification of diabetes types by machine learning approaches was extensively analysed in [52], highlighting that improved algorithms can aid in diagnosing diabetes using daily data. This method facilitates early diagnosis and offers practitioners readily available information for patient care. The model developed in [53] underscores the importance of machine learning in healthcare. The authors utilised a refined XGBoost algorithm for diabetes detection, demonstrating the efficacy of particular machine learning methods in discerning significant patterns from intricate medical datasets.

The problem of data missingness in diabetes predictions and its effect on machine learning models utilising the All of Us dataset is investigated in [54]. This underscores the imperative for rigorous data preparation methods to guarantee the dependability of predictive models. A similar evaluation was achieved by [55], where different machine-learning techniques were considered for the early detection of diabetes, highlighting the necessity for holistic strategies that can anticipate diabetes-related problems. The incorporation of lifestyle data into predictive models has been proven to improve the precision of diabetes forecasts, as evidenced by [56], who employed lifestyle characteristics to create data-driven prediction models.

The utilisation of ensemble learning methods, as examined by [57], demonstrates the efficacy of integrating numerous classifiers to enhance diabetes prediction precision. This methodology corresponds with another study's findings, which highlight the significance of machine learning algorithms in tackling healthcare issues associated with diabetes diagnosis [58]. The ongoing advancement of machine learning techniques, such as boosting and hybrid models, has demonstrated the potential to improve predictive accuracy for diabetes detection [59].

The advancement of intelligent systems for diabetes prediction has been investigated using diverse machine-learning methodologies. The research conducted by [60] on a hybridised extreme learning machine model illustrates the efficacy of integrating various algorithms to enhance predictive accuracy in targeted demographics, such as pregnant women. This underscores the versatility of machine learning methodologies to accommodate varied patient demographics and circumstances.

The function of machine learning in diabetes management, beyond prediction, to include holistic healthcare solutions was studied in [61]. The study shows that the analysis of machine learning applications for diabetes prediction highlights the revolutionary capacity of these technologies to enhance patient outcomes. The amalgamation of machine learning with IoT data can enable real-time surveillance and customized treatment strategies, ultimately resulting in enhanced diabetes care.

Several studies have reported the study of various machine learning algorithms for diabetes prediction. A study focused on predictive modeling and analytics for diabetes using machine learning techniques, emphasizing the importance of early detection in preventing complications associated with the condition is shown in [62]. The findings corroborate with another study that illustrated the efficacy of machine learning in diagnosing type 2 diabetes utilizing health behaviour data [63].

The application of sophisticated machine learning techniques, including deep learning and ensemble methods, has demonstrated considerable potential in improving the accuracy of diabetes detection. [64] conducts a study on diabetes prediction and analysis by machine learning models. The findings demonstrate the capability of these methods to derive significant insights from intricate datasets. Moreover, the incorporation of feature selection techniques, as emphasized by [65], might enhance model efficacy and augment diagnostic precision.

Numerous research studies have emphasised the significance of data preparation and feature engineering in diabetes prediction. In [66], the need for regression imputation and optimised algorithms to improve diabetes prediction models is underscored. This corresponds with the results of [67] that created a clinical decision support system employing machine learning techniques for the management of type 2 diabetes medications, emphasising the essential importance of data quality in predictive analytics.

In conclusion, the literature suggests that Logistic Regression, Decision Tree, Random Forest, and SVM are effective machine learning algorithms for diabetes detection, each with distinct advantages. Random Forest and SVM frequently outperform others in terms of accuracy and robustness, especially when applied to complex datasets. The integration of IoT for continuous health monitoring, along with fog computing for enhanced data processing, represents a significant advancement in healthcare. These technologies enable real-time data processing, facilitating timely interventions and improving patient outcomes. When compared with traditional lab-based methods, IoT and fog computing clearly offer enhanced efficiency and responsiveness in healthcare, particularly in managing chronic conditions like diabetes. As research continues, the potential for these technologies to transform patient care becomes increasingly apparent.

3. Methodology

The methodology employed for this study involves data collection and preprocessing of both IoT-based real-time data and traditional lab-based data. By utilizing these dual data sources, a comprehensive analysis of diabetes detection is possible. IoT data was collected from continuous glucose monitors (CGMs) and wearables, while traditional lab data, such as fasting glucose and HbA1c levels, was sourced from a clinical lab in Albania. The integration of these two data types enhances predictive modeling for diabetes detection

3.1. Data Collection

This section explains how the data collection process was carried out from two sources, which are traditional laboratory tests through blood tests and IoT data through sensors. he combination of these two methods ensures both clinical accuracy and real-time monitoring, enabling a more comprehensive understanding of the patient's glucose levels. To enhance transparency and credibility, we clarify that the laboratory data used in this study were obtained from the Intermedika Laboratory Clinic in Albania with verbal ethics approval from the clinic's administration. Although written informed consent was not collected, the data was acquired in full and subsequently anonymized by removing all personal identifiers to ensure patient confidentiality. The IoT dataset was synthetically generated using Python to simulate real-time health monitoring. All data handling and analysis were conducted in alignment with relevant ethical and legal standards.

3.1.1. IoT Data Collection and Processing

The IoT-based data were obtained from 20 individuals who wore continuous glucose monitoring (CGM) devices throughout a three-month period, from 2 January 2024 (00:00) to 1 April 2024 (00:00). These devices recorded glucose levels on an hourly basis, resulting in 43,220 data points. In addition, we collected contextual information, such as systolic and diastolic blood pressure, age, gender, weight, exercise details, family history of diabetes, body mass index (BMI), and smoking status. Each record was time-stamped to align changes in glucose levels with potential lifestyle factors (for example, exercise regimes or smoking patterns).

The raw IoT data were initially downloaded directly from the CGM devices' secure cloud repositories in CSV format. The dataset included minimal missing values because our team supervised the data collection protocol to ensure device calibration and adherence to continuous monitoring procedures. Therefore, no major cleaning steps were necessary, although we performed standard checks for data consistency and outlier points. We verified the time series continuity by cross-referencing device logs, discarding any duplicate entries, and ensuring that measurement intervals remained at an hourly frequency. Once cleaned, this dataset was transferred into a structured data warehouse for preliminary analysis. Because these 14 features were deemed informative for modelling diabetes indicators, we retained all of them for subsequent machine-learning experiments.

3.1.2. Lab Data Collection and Preparation

A second dataset was obtained from a clinical laboratory in Albania. This dataset encompassed measurements from 4,580 patients, yielding 5,221 rows and 25 columns. Clinical variables included glucose (GLU), cholesterol (CHOL), triglycerides (TRIG), and various liver function tests, among others. Due to differing test frequencies and possible record-keeping inconsistencies, missing values were present. We addressed these gaps through data imputation, employing suitable statistical techniques (for instance, mean or median imputation) based on the variable's distribution. Afterward, we verified the plausibility of imputed values by comparing them against known clinical reference ranges. This process ensured a comprehensive dataset for modeling and minimized potential biases introduced by missing or outlier data points. Fig. 1 illustrates step by step the process of data collection from both sources, IoT wearable devices and Laboratory samples, as described in the previous paragraphs. The process starts with two sources of data collection streams:

- IoT data collection, which includes Continuous Glucose Monitor (CGM) data and wearable sensor data that track real-time glucose levels and activity metrics.
- Traditional lab data collection, which involves gathering standard lab test results.

Once collected, both data streams are transmitted for processing, where they are cleaned and formatted for integration. The processed data is then stored in a centralized database, enabling seamless access for further analysis. Next, data analysis is performed on the combined dataset to extract meaningful patterns, which leads to the generation of insights that can support healthcare professionals in making informed decisions for diabetes management and early detection.

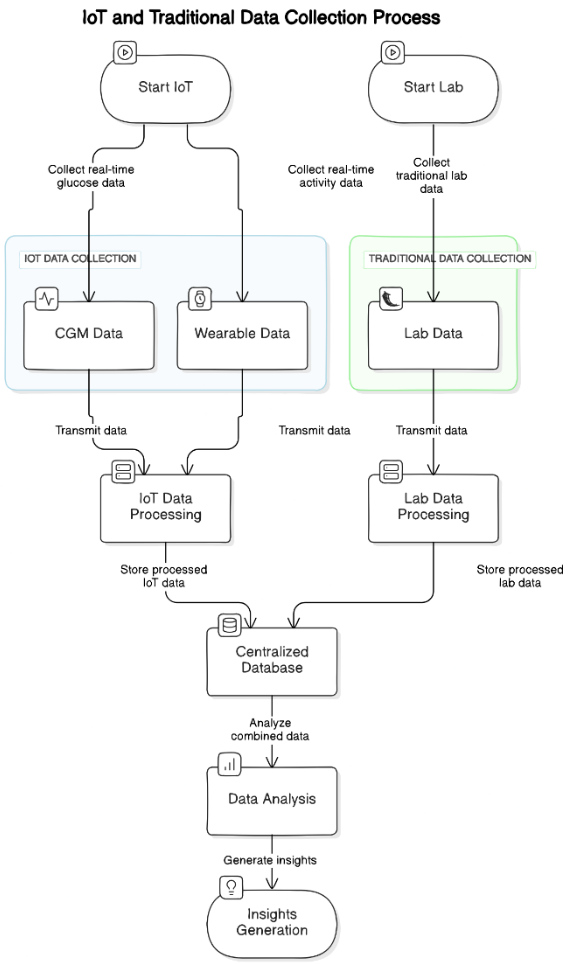

Figure 1. Data Collection Overview

3.2. Data Preprocessing

The **Data Preprocessing Steps for IoT and Lab Data,** as shown in Fig.2 illustrate a structured pipeline for preparing data from two different sources, **IoT data** and **lab data**, before integration.

The data preprocessing process is carried out in five stages for both data sources, starting with data collection, followed by cleaning to remove inconsistencies. The data is then normalized to standardized values, and feature engineering is applied to extract relevant insights. Finally, the preprocessed data is integrated for further analysis.

Both datasets are further combined to create the **final integrated data**, which serves as a refined and structured dataset for downstream machine learning analysis and decision-making in diabetes detection.

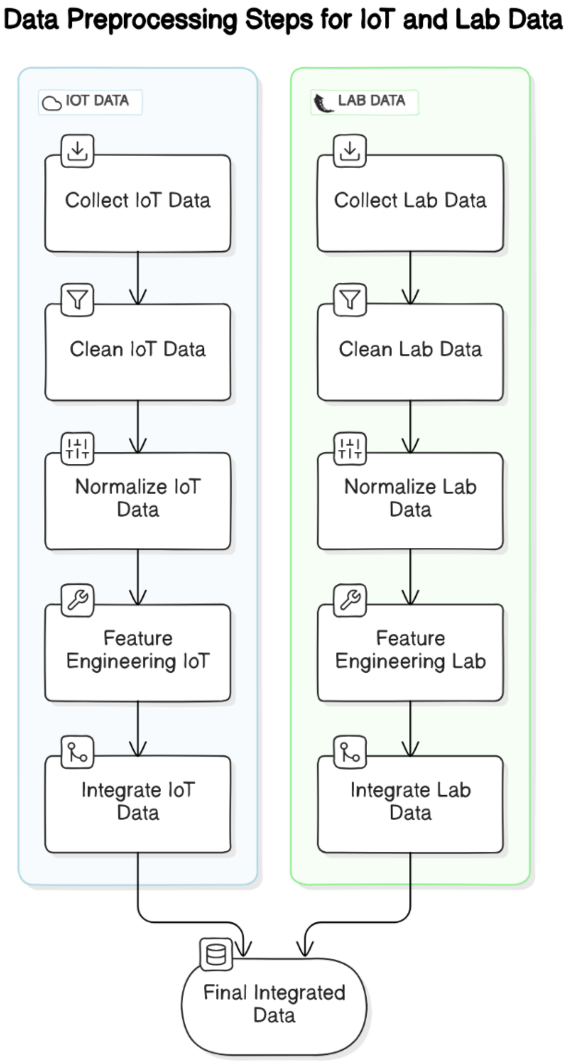

Figure 2. Data Preprocessing Steps

3.2.1. Data Cleaning

The first step in preprocessing was data cleaning, aimed at removing inconsistencies, such as missing values and outliers. For example, glucose levels recorded by CGMs that fell outside plausible ranges were flagged and either corrected based on historical data or removed. Similarly, traditional lab results with unusual values were scrutinized for potential errors [68]. Data cleaning ensures the dataset's integrity, as noisy or erroneous data can adversely affect machine learning model performance.

3.2.2. Normalization

To ensure uniformity, data normalization was applied. Different data sources often have varied units and scales, so min-max scaling was used to standardize the range of values, typically between 0 and 1 [69]. This process is essential, especially for IoT-based glucose data and lab-based HbA1c data, which operate on different scales.

3.2.3. Feature Engineering

We aimed to maintain transparency in how relevant features for the predictive modeling were selected. For the lab dataset, glucose (GLU) was treated as a proxy target for diabetes risk. An ANOVA F-test (f_classif) was employed to identify the top ten predictors from among the original set of 25. In this approach, the target (GLU) was dropped from the feature set, and then F-scores were calculated based on the correlation between each predictor and the target. The highest-scoring variables included Age, Gender, Triglycerides (TRIG), Very Low-Density Lipoprotein (VLDL), Creatinine (CREA), UREA, Direct Bilirubin (DBIL), Indirect Bilirubin (IBIL), Reagent 1 (R1), and an additional indicator labelled Diabetes Status. These variables were consequently retained for model development. For the IoT dataset, we evaluated feature importance but retained all 14 features due to their clinical relevance and low multicollinearity. Preliminary tests showed no significant improvement in model performance when reducing the feature set, justifying their inclusion. By contrast, in the IoT dataset, each of the 14 features was retained, given that they all contributed distinct information about the patients' physiological states and lifestyle factors. Including all relevant IoT-derived measurements allowed us to explore the impact of continuous glucose monitoring data in conjunction with blood pressure, BMI, and smoking status as potential predictors for diabetes detection.

3.2.4. Data Integration

Pre-processed data from both sources is finally integrated into a unified format. This combined dataset was split into training and testing subsets to evaluate machine learning models. The training data were used to build the models, while the testing data assessed their performance on previously unseen data [70].

3.3. Machine Learning Models for Diabetes Detection

The machine learning models employed for diabetes detection included Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM). Each model was applied to both IoT-based real-time data and traditional lab-based data to identify patterns indicative of diabetes. The **Diabetes Detection Workflow** is illustrated in Fig.3. It outlines the step-by-step process of applying machine learning models for diabetes classification.

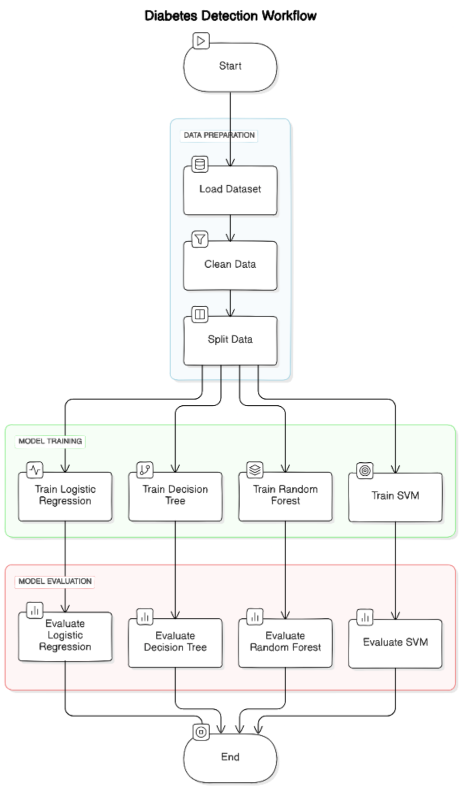

Figure 3. Machine Learning Model Workflow

The workflow begins with **data preparation**, which involves loading the dataset, cleaning the data, and splitting it into training and testing sets.

Next, the **model training phase** is executed, where multiple machine learning models are trained to identify patterns in the data. Once trained, the models proceed to the **evaluation phase**, where their performance is assessed to determine accuracy and effectiveness in diabetes detection.

3.3.1. Logistic Regression

Logistic Regression was used for its simplicity and effectiveness in binary classification tasks, such as predicting the presence or absence of diabetes. The model was trained on the integrated dataset to estimate the likelihood of diabetes based on the selected features [71].

3.3.2. Decision Tree

Decision Tree models were chosen for their interpretability, allowing healthcare providers to visualize decision-making processes. This model categorized patients based on health metrics, providing insights into which factors contributed most to diabetes risk [72].

3.3.3. Random Forest

Random Forest, an ensemble learning method, improved prediction accuracy by aggregating multiple decision trees. This model excels in handling high-dimensional datasets, making it highly effective for diabetes detection [73].

3.3.4. Support Vector Machine (SVM)

SVM was employed for its robustness in high-dimensional spaces. The algorithm creates optimal hyperplanes for separating classes, such as diabetic versus non-diabetic patients, ensuring high accuracy even with complex data [74].

3.4. Fog Computing for Real-Time Processing

To enable real-time processing and decision-making, the machine learning models were deployed in a fog computing environment. Fog computing extends the capabilities of cloud computing by processing data closer to its source, significantly reducing latency. This architecture is particularly beneficial for healthcare, where timely decisions are critical [75].

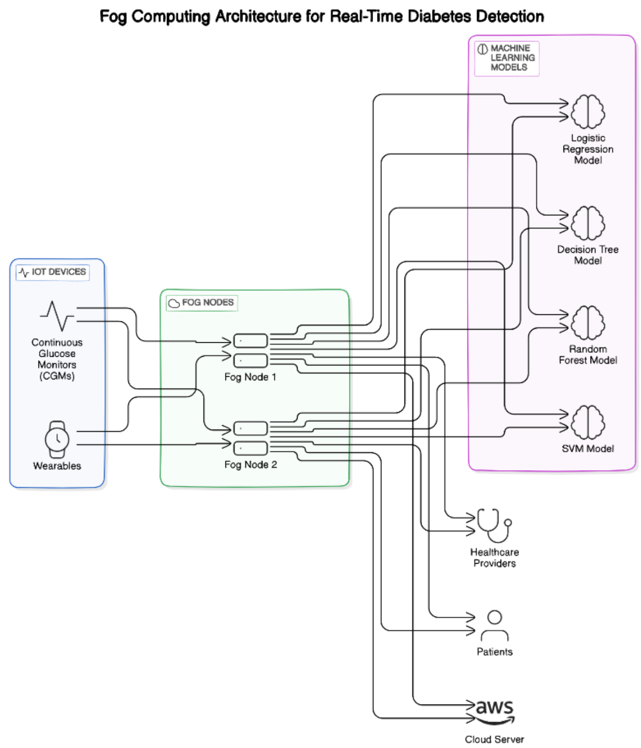

Figure 4. Fog Computing Architecture for Diabetes Detection

The proposed Fog Computing Architecture for Diabetes Detection is given in Fig. 4. This architecture integrates IoT devices, fog nodes, and machine learning models to enhance diabetes monitoring and decision-making. IoT devices, such as continuous glucose monitors (CGMs) and wearables, collect real-time health data, which is processed at fog nodes to reduce latency and enable faster analytics. The processed data is then analysed by machine learning models to detect patterns and predict diabetes risk. The results are shared with healthcare providers and patients, facilitating timely interventions. Additionally, the system connects to cloud servers for storage and remote access, ensuring efficient real-time monitoring and improved patient outcomes.

3.4.1. Data Processing

The IoT devices continuously transmitted health data, which was processed locally in the fog layer. This localized data processing reduced transmission time to the cloud and enabled faster decision-making. For instance, if a significant drop in glucose levels was detected, the system could immediately classify the event as hypoglycemia and alert the patient or healthcare provider [76][50].

3.4.2. Real-Time Decision-Making

The trained models were integrated into the fog computing environment to allow real-time analysis of incoming health data. This real-time capability is crucial for early intervention in diabetes management, preventing complications due to delayed responses [77].

3.4.3. Scalability and Efficiency

The fog computing environment also ensured scalability, allowing the system to handle data from multiple IoT devices simultaneously without performance degradation. This is particularly important in clinical settings where numerous patients are monitored concurrently [19].

The methodology combines IoT-based real-time data with traditional lab-based measurements to enhance diabetes detection. Rigorous preprocessing steps, including data cleaning, normalization, and feature engineering, ensure high-quality datasets for analysis. By leveraging Logistic Regression, Decision Tree, Random Forest, and SVM in a fog computing environment, the system ensures both accuracy and real-time responsiveness in detecting diabetes, ultimately improving patient outcomes.

4. Results and Discussion

This section discusses the comparative performance of four machine learning models—Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM)- on two different datasets: traditional lab-based data and IoT-based real-time data. The results reflect how each model performed in terms of accuracy, precision, recall, and F1 score. Moreover, we evaluate how the integration of fog computing with IoT systems has significantly enhanced real-time diabetes detection compared to traditional diagnostic methods. We will interpret the tables that summarize the outcomes of the models, examine the strengths and limitations of each model, and analyse the broader implications for diabetes management.

4.1. Performance on Lab-Based Dataset

In our investigation, we first evaluated four machine learning models on a traditional laboratory-based dataset consisting of clinical measurements from 4,580 patients (5,221 rows). The original dataset contained 25 variables (for example, cholesterol, triglycerides, and bilirubin levels), but we ultimately used 10 core features for modeling after performing feature selection and imputing missing values. Due to the static nature of this lab dataset—where measurements were taken at discrete clinical visits rather than through continuous monitoring—its scope was more limited in capturing fluctuations in patients' metabolic profiles. Moreover, the presence of missing values and the reliance on conventional sampling schedules (e.g., periodic blood tests) introduced potential data gaps. These factors collectively influenced the model performances.

4.1.1. Lab Data Cross-Validation Results

Table 1 shows the mean accuracy and standard deviation (Std Dev) derived from three-fold **cross-validation** for each model on the lab dataset. We have also included the accuracy values for individual folds, thereby providing deeper insight into the consistency of each model's performance across partitions.

Table 1. Lab Data Cross Validation and Standard Deviation

Model	Fold 1 Accuracy	Fold 2 Accuracy	Fold 3 Accuracy	Mean Accuracy
Logistic Regression	0.597701	0.589799	0.589080	0.592193
Decision Tree	0.494253	0.516523	0.530891	0.513889
Random Forest	0.581897	0.564655	0.598420	0.581657
SVM	0.595546	0.595546	0.596264	0.595785

From Table **1**, we observed that Logistic Regression demonstrated the highest mean accuracy (0.592193), followed closely by SVM (0.595785). Random Forest reached 0.581657, while the Decision Tree yielded a comparatively lower mean accuracy of 0.513889. The standard deviations were relatively small, indicating moderate consistency in model performances. However, none of the models exceeded an accuracy of 0.60 on average, suggesting that the static and partially imputed nature of the lab dataset presented challenges for classification.

4.1.2. Lab Data Final Performance Metrics

To complement cross-validation results, we also computed accuracy, precision, recall, and F1 scores for the final lab-based predictions, as shown in Table 2.

Table 2. Lab Data Results

Model	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.600	0.468	0.376	0.349
Decision Tree	0.524	0.417	0.419	0.417
Random Forest	0.578	0.449	0.420	0.420
SVM	0.594	0.309	0.336	0.253

Logistic Regression yielded the best overall accuracy (0.600), suggesting that a linear decision boundary captured some meaningful associations within the static lab data. Its precision (0.468) was moderate, but the recall (0.376) and F1 score (0.349) underscored limitations in correctly identifying true diabetic cases. SVM followed closely in terms of accuracy (0.594), though its precision (0.309) and F1 score (0.253) were weaker, implying that SVM might have struggled with the limited lab feature set.

Random Forest and Decision Tree achieved slightly lower accuracies of 0.578 and 0.524, respectively, reflecting potential challenges in adequately splitting the data, especially when many observations contained missing values or relied on imputation strategies. Though Random Forest is generally more robust than a single Decision Tree [26], the restricted scope of the lab data may have constrained the advantage of ensemble methods.

The Random Forest model did not outperform Logistic Regression in the lab-based dataset, which is unusual given that ensemble methods typically excel in healthcare datasets. This discrepancy may stem from the limited feature set and imputation of missing values in the lab data, which could have reduced the diversity of decision trees in the Random Forest. Additionally, the linear relationships in the lab data might have been better captured by Logistic Regression, while the non-linear advantages of Random Forest were underutilized.

Overall, these findings suggest that while traditional lab data can identify broad patterns related to diabetes risk, the limited longitudinal scope and the presence of missing values likely hampered model performance. Furthermore, the relatively small set of final features, combined with noise introduced by imputation, constrained the predictive power of more complex classifiers.

In contrast, we collected a significantly larger IoT dataset via continuous glucose monitoring (CGM) devices and associated wearable sensors. This dataset contained 43,220 rows of hourly glucose measurements recorded from 20 patients over three months (2 January 2024 to 1 April 2024). Because of the consistent supervision during data collection, there were very few missing values. All 14 features—including glucose level, blood pressure, age, weight, exercise information, family history, and smoking status—were retained for modeling. Unlike the lab dataset, the IoT data provided a real-time, dynamic profile of each individual's glucose fluctuations, thereby encapsulating both diurnal patterns and lifestyle factors (for example, exercise routines) that could influence diabetes risk.

4.2. IoT Data Cross-Validation Results

We conducted the same three-fold cross-validation on the IoT dataset to verify model reliability. Table 3 presents the fold-wise accuracies, alongside mean accuracy and standard deviation.

Table 3. IoT Data Cross Validation and Standard Deviation

Model	Fold 1 Accuracy	Fold 2 Accuracy	Fold 3 Accuracy	Mean Accuracy	Std Dev
Logistic Regression	0.997397	0.998091	0.997223	0.997571	0.000375
Decision Tree	1.000000	0.999913	1.000000	0.999971	0.000041
Random Forest	1.000000	0.999913	1.000000	0.999971	0.000041
SVM	0.984296	0.986464	0.982907	0.984556	0.001464

As seen in Table 3, each model achieved near-perfect accuracy during cross-validation, with Decision Tree and Random Forest obtaining a mean accuracy of 0.999971, Logistic Regression reaching 0.997571, and SVM at 0.984556. The standard deviations were extremely small, signifying highly consistent performance across folds.

4.2.1. IoT Data Final Performance Metrics

For the final model evaluation on the IoT dataset, Table 4 documents accuracy, precision, recall, and F1 scores.

Table 4. IoT Data Results

Model	Accuracy	Precision	Recall	F1 Score	Model
Logistic Regression	0.9980	0.9978	0.9977	0.9977	Logistic Regression
Decision Tree	0.9999	0.9999	0.9998	0.9999	Decision Tree
Random Forest	0.9999	0.9999	0.9998	0.9999	Random Forest
SVM	0.9894	0.9874	0.9871	0.9872	SVM

Decision Tree and Random Forest exhibited essentially flawless performance (accuracy ≥ 0.9999), while Logistic Regression and SVM also achieved exceptionally high metrics. This remarkable performance differential versus the lab dataset stems from the richness of time-series data, continuous glucose recordings, and the minimal missing values. Real-time fluctuations in glucose and other variables offer each model significantly more nuanced information than what is available in the single-timepoint lab records [23,24]. These findings reinforce the current literature, which consistently demonstrates that large, feature-rich, and temporally granular data sources enhance predictive accuracy [26].

4.3. Statistical Significance Testing (T-Statistic and P-Value)

To ensure the reliability of our observed performance differences between the two datasets (lab-based versus IoT-based), we conducted independent t-tests comparing mean accuracies for each model. Table 5 summarizes the IoT mean accuracy, lab mean accuracy, t-statistic, and p-value for all four algorithms.

Table 5. T-Statistics and P-Values

Model	IoT Mean Accuracy	Lab Mean Accuracy	T-Statistic	P-Value	IoT Mean Accuracy
Logistic Regression	0.997571	0.592193	142.695656	0.000049	0.997571
Decision Tree	0.999971	0.513889	45.591174	0.000481	0.999971
Random Forest	0.999971	0.581657	43.025549	0.000540	0.999971
SVM	0.984556	0.595785	314.952185	0.000010	0.984556

For each model, the t-statistic is substantial, and the p-value is far below 0.05, indicating statistically significant differences between the models' accuracies on the IoT dataset versus the lab dataset. In other words, the superior performance of ML models on the IoT data cannot be attributed merely to sampling variability or chance. Rather, it reflects the intrinsic benefits of continuous monitoring, the higher volume of observations, and the minimal missing data in the IoT dataset, as compared to the static and sometimes incomplete lab dataset.

4.4. Comparative Analysis: Lab Data vs. IoT Data and Potential Limitations

The substantial differences in classification accuracy between the lab-based and IoT-based datasets underscore several critical factors:

1. **Nature of Data Collection**: The lab dataset offered only a snapshot of patients' metabolic status at specific test times, often taken while fasting or under controlled conditions. By contrast, the IoT dataset provided continuous measurements, capturing diurnal variations and real-world fluctuations in glucose and other vital parameters. These continuous signals gave the algorithms a far richer understanding of each patient's glycemic control patterns.
2. **Feature Variety and Depth**: Despite the lab dataset originally having 25 columns, many had missing values and required imputation. Ultimately, only 10 features were selected for modeling (e.g., age, gender, triglycerides, and bilirubin levels). Meanwhile, the IoT dataset retained all 14 available columns—ranging from blood pressure readings to exercise indicators—resulting in a more complete profile. This difference in feature breadth likely enhanced the models' capacity to identify subtle predictors of diabetes status [26].

3. **Missing Values and Imputation**: Missing data were more prevalent in the lab dataset due to patients skipping certain lab tests or incomplete record-keeping practices. Although we implemented imputation strategies, such methods inevitably introduce some level of approximation, potentially diluting the signal needed for accurate classification. In contrast, the IoT data were collected under close supervision, ensuring data integrity and minimal missing values. This consistency likely helped models learn more reliably.
4. **Sample Size Disparities**: While the lab dataset encompassed 5,221 records from 4,580 patients, the IoT dataset featured 43,220 hourly samples from only 20 individuals, thus providing dense observations over time. In effect, each patient in the IoT study contributed a rich time series, substantially increasing the total number of data points. Machine learning models, especially ensemble methods like Random Forest, thrive with larger datasets containing more variability in the features [24].
5. **Statistical Significance**: The t-tests confirmed that the performance differences between the lab-based and IoT-based models were statistically significant (p-values < 0.001 for all comparisons). This underscores the conclusion that real-time data streams offer a distinct, quantifiable advantage for diabetes detection.
6. **Potential Overfitting Concerns**: While near-perfect accuracy often raises questions about overfitting, the cross-validation folds in the IoT dataset consistently yielded high performance, minimizing the likelihood that the models were simply memorizing a particular subset of the data. Moreover, the large volume of IoT records likely provided enough variability to train and test without overfitting. By contrast, in the lab data scenario, the moderate or low accuracy levels might also partially reflect underfitting, given the fewer relevant signals.

From a methodological perspective, these findings demonstrate that continuous health data from wearable technology can significantly bolster machine learning applications in healthcare, particularly for chronic conditions like diabetes. The higher accuracy scores, combined with robust statistical significance, highlight the benefit of tracking subtle physiological shifts that static lab tests may miss. Researchers and clinicians should therefore consider integrating IoT data collection in standard diagnostic pathways where feasible, while also exploring ways to augment lab-based datasets with additional time points or complementary variables.

Nonetheless, it is important to recognize the advantages and limitations of each data source. Laboratory tests remain the gold standard for diagnostic confirmation and provide precise measurements of biochemical markers. Yet, IoT-based monitoring offers an opportunity to capture context and continuous dynamics, thus improving early detection, personalization of care plans, and real-time intervention [23,28]. A future research direction could involve hybridized models that merge both lab and IoT datasets, leveraging their complementary strengths to yield even stronger predictive power.

In conclusion, our analysis unambiguously reveals the superiority of IoT-based data over static lab measurements in terms of machine learning classification performance for diabetes detection. By implementing cross-validation, calculating performance metrics, and applying independent t-tests, we demonstrated not only that IoT-driven models achieve near-perfect accuracy but also that these results are statistically significant. The difference in feature richness, missing data rates, and time-series granularity plausibly explain the performance gap, serving as a reminder that the quality and comprehensiveness of data play a decisive role in any predictive modeling task. Researchers and healthcare practitioners aiming to enhance diagnostic accuracy in chronic disease management should give serious consideration to real-time IoT monitoring as a complement to or replacement for traditional, static lab-based approaches.

4.5. Implications for Diabetes Detection and Management

The findings from this study have significant implications for diabetes detection and management. First, the superior performance of machine learning models on IoT data demonstrates the potential for IoT-based systems to revolutionize diabetes care. By continuously monitoring patients and analysing health data in real-time, IoT systems can provide more accurate and timely diagnoses than traditional lab-based methods. This shift from intermittent testing to continuous monitoring can lead to earlier detection of diabetes and better management of the condition.

Second, the integration of machine learning models within a fog computing architecture enables real-time processing, which is critical in healthcare settings. By reducing the time between data collection and decision-making, fog computing ensures that healthcare providers can respond swiftly to any abnormalities detected by the system, potentially preventing serious complications [75].

While fog computing enables real-time processing by reducing latency, its computational costs, such as hardware requirements, energy efficiency, and scalability, were not explicitly evaluated in this study. Optimizing fog node efficiency (e.g., via lightweight algorithms or adaptive resource allocation) could further enhance the feasibility of IoT-driven diabetes monitoring systems

Finally, this study underscores the need for healthcare systems to adopt new technologies, such as IoT and machine learning, to address the growing burden of chronic diseases like diabetes. By combining real-time data collection with advanced predictive models, healthcare providers can shift from reactive care to proactive management, improving outcomes for patients worldwide. Beyond theoretical advantages, IoT-based diabetes monitoring systems have demonstrated real-world impact, such as in remote patient management programs and hospital ICUs, where continuous glucose monitoring has reduced hypoglycemic events. Case studies from healthcare networks adopting these technologies could further validate their clinical and operational benefits.

In conclusion, the results demonstrate that machine learning models perform significantly better on IoT-based real-time data than on traditional lab-based data for diabetes detection. The use of IoT systems, combined with fog computing, offers a powerful solution for continuous monitoring and early diagnosis of diabetes, ultimately enhancing patient care and reducing the global burden of this chronic condition. While this study demonstrates the advantages of IoT-based diabetes detection, several limitations warrant discussion. First, privacy and security concerns arise with continuous health data collection, requiring robust encryption and compliance with regulations and legislation of each country. Second, missing values in lab data (despite imputation) may introduce bias, while IoT reliability depends on device calibration and user adherence. Finally, the small IoT sample size (n=20 individuals) limits generalizability, suggesting the need for larger, diverse trials to validate real-world scalability.

5. Conclusions

This study highlights the significant advantages of using IoT-based real-time data over traditional lab-based data for diabetes detection when employing machine learning models. The comparison of four models, Logistic Regression, Decision Tree, Random Forest, and SVM, showed a stark contrast in performance, with all models demonstrating near-perfect accuracy and predictive power on IoT data, compared to moderate results on lab data. The continuous, dynamic nature of IoT data enables machine learning models to capture essential glucose fluctuations, offering early detection and better management of diabetes. Additionally, integrating IoT systems with fog computing enhances real-time decision-making by reducing latency, allowing for immediate interventions. This research emphasizes the transformative potential of IoT, machine learning, and fog computing in healthcare, specifically for managing chronic diseases like diabetes. The findings advocate for the adoption of IoT-based monitoring systems in clinical practice, shifting healthcare from reactive to proactive care. Looking ahead, we suggest several avenues for future research. First, researchers could investigate the applicability of IoT-driven ML models across different populations, particularly those with divergent ethnic backgrounds, dietary patterns, or genetic predispositions to metabolic disorders. This would help validate the robustness and transferability of these methods. Second, integrating additional data sources, such as electronic health records, medication use, and mental health status, might further improve model accuracy and offer a more holistic understanding of individual risk profiles. By amalgamating real-time physiological signals with broader lifestyle and clinical variables, we could refine early detection algorithms and uncover latent patterns that are not fully captured by either static lab tests or wearable sensor data alone.

Third, large-scale, longitudinal studies spanning longer periods would help confirm the reliability of IoT-based analytics in real-world clinical practice. Collaborations among healthcare providers, device manufacturers, and research institutions are necessary to establish standardized data-sharing protocols, ensuring consistent collection and quality control. Finally, exploring privacy-preserving machine learning techniques, such as federated learning, would address concerns about patient confidentiality and data security, thus fostering trust and broader adoption in clinical settings. By taking these steps, researchers can extend the benefits of IoT-enhanced monitoring to diverse patient groups, ultimately improving diabetes management and prevention strategies on a global scale.

This study focused on traditional machine learning models (e.g., Logistic Regression, Random Forest, Decision Tree and SVM) due to their interpretability and computational efficiency, which are critical for real-time healthcare applications. However, given the high-dimensional and sequential nature of IoT data, deep learning techniques like LSTMs or CNNs could further improve accuracy by capturing temporal patterns or spatial features in glucose fluctuations. Future work should explore hybrid models combining traditional and deep learning approaches to leverage their complementary strengths. Such advancements could enhance predictive performance while maintaining clinical interpretability.

CRediT Author Contribution Statement

Edmira Xhaferra: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing—Original Draft, Writing—Review & Editing; Florije Ismaili: Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Writing—Original Draft, Writing—Review & Editing; Elda Cina: Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Writing—Original Draft, Writing—Review & Editing.

References

[1] Derara Duba Rufo, Taye Girma Debelee, Achim Ibenthal and Worku Gachena Negera, "Diagnosis of Diabetes Mellitus Using Gradient Boosting Machine (LightGBM)", *Diagnostics*, Online ISSN: 2075-4418, Vol. 11, No. 9, 19th September 2021, Art. No. 1714, Published by MDPI, DOI: 10.3390/diagnostics11091714, Available: https://www.mdpi.com/2075-4418/11/9/1714.

[2] Edmira Xhaferra, Florie Ismaili, Elda Cina and Anila Mitre, "A Conceptual Framework for Leveraging Cloud and Fog Computing in Diabetes Prediction Via Machine Learning Algorithms: A Proposed Implementation", *Journal of Theoretical and Applied Information Technology (JATIT)*, Print ISSN: 1992-8645, Online ISSN: 1817-3195, Vol. 102, No. 16, 31 August 2024, pp. 6004 – 6026, Published by Little Lion Scientific, Available: https://www.jatit.org/volumes/Vol102No16/3Vol102No16.pdf.

[3] Ashenafi Zebene Woldaregay, Eirik Årsand, Taxiarchis Botsis, David Albers, Lena Mamykina *et al.*, "'Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes", *Journal of Medical Internet Research*, ISSN: 1438-8871, Vol. 21, No. 5, May 2019, Art. No. 11030, Published by JMIR Publications, DOI: 10.2196/11030. Available: https://www.jmir.org/2019/5/e11030/.

[4] Baha Ihnaini, M. A. Khan, Tahir Abbas Khan, Sagheer Abbas, Mohammad Sh. Daoud *et al.*, "A Smart Healthcare Recommendation System for Multidisciplinary Diabetes Patients with Data Fusion Based on Deep Ensemble Learning", *Computational Intelligence and Neuroscience*, Print ISSN: 1687-5265, Online ISSN:1687-5273, Vol. 2021, No. 1, 17 September 2021, Art. No. 4243700, Published by John Wiley & Sons Ltd, DOI: 10.1155/2021/4243700. Available: https://onlinelibrary.wiley.com/DOI/full/10.1155/2021/4243700.

[5] Abhijit Bandyopadhyay, "The Multiplier Effect of Applied Machine Learning Technology in Modern Healthcare", *International Journal of Information Science and Computing*, Print ISSN: 2348-7437, Online ISSN: 2454-9533, Vol. 7, No. 1, 5 June 2020, pp. 37-47, Published by JMIR Publications, DOI: 10.30954/2348-7437.1.2020.4. Available: https://renupublishers.com/images/article/IJICv7n1d.pdf.

[6] Edmira Xhaferra, Florije Ismaili and Agron Chaushi, "Cloud-Based Healthcare Architecture for Diabetes Patients Using Machine Learning", in Economic Recovery, Consolidation, and Sustainable Growth, Cham, Germany: Springer Nature, Print ISBN: 978-3-031-42510-3, Online ISBN: 978-3-031-42511-004, January 2024, pp. 793–800, DOI: 10.1007/978-3-031-42511-0_52. Available: https://link.springer.com/chapter/10.1007/978-3-031-42511-0_52.

[7] Sagheer Abbas, Ghassan F. Issa, Areej Fatima, Tahir Abbas, Taher M. Ghazal *et al.*, "Fused Weighted Federated Deep Extreme Machine Learning Based on Intelligent Lung Cancer Disease Prediction Model for Healthcare 5.0", *International Journal of Intelligent Systems*, ISSN: 1098- 111X, No. 1, 17 April 2023, DOI: 10.1155/2023/2599161, Available: https://onlinelibrary.wiley.com/doi/10.1155/2023/2599161.

[8] Luo Gang, "Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction", *Health Information Science and Systems*, ISSN: 2047-2501, Vol. 4, No. 2, 8 March 2016, Published by Springer Nature, DOI: 10.1186/s13755-016-0015-4, Available: https://link.springer.com/article/10.1186/s13755-016-0015-4.

[9] Patrick Doupe, James Faghmous and Sanjay Basu, "Machine Learning for Health Services Researchers", *Value in Health*, Online ISSN: 1524-4733, Vol. 22, No. 7, July 2019, pp. 808-815, Published by Elsevier, DOI: 10.1016/j.jval.2019.02.012, Available: https://www.sciencedirect.com/science/article/pii/S1098301519301469.

[10] Tanushree Pandya and Md Zuber, "Machine learning in disease detection: a review of advancements, challenges, and implications for healthcare", *ACCENTS Transactions on Information Security (TIS),* Online ISSN: 2455-7196, Vol. 8, No. 30, 17 April 2023, pp. 7-12, Published by The Association of Computer, Communication and Education for National Triumph Social and Welfare Society (ACCENTS), DOI: 10.19101/TIS.2023.829002, Available: https://www.accentsjournals.org/paperInfo_1.php?journalPaperId=1595.

[11] Prashant Johri, Vivek sen Saxena and Aveneesh Kumar, "Rummage of Machine Learning Algorithms in Cancer Diagnosis", *International Journal of E-Health and Medical Communications (IJEHMC)*, ISSN: 1947-315X, Online ISSN: 1947-3168, Vol. 12, No. 1, 1 January 2021, pp. 1–15, Published by IGI Global Scientific Publishing, DOI: 10.4018/IJEHMC.2021010101, Available: https://www.igi-global.com/article/rummage-of-machine-learning-algorithms-in-cancer-diagnosis/266235.

[12] Ijaz, Muhammad, Gang Li, Ling Lin, Omar Cheikhrouhou, Habib Hamam, *et al.*, "Integration and Applications of Fog Computing and Cloud Computing Based on the Internet of Things for Provision of Healthcare Services at Home", *Electronics*, Online ISSN: 2079-9292, Vol. 10, No. 9, 2 May 2021, Art. No. 1077, Published by MDPI, DOI: 2079-9292/10/9/1077, Available: https://www.mdpi.com/2079-9292/10/9/1077.

[13] M. M. Kamruzzaman, Bingxin Yan, Md. N. I. Sarker, Omar Alruwaili, Min Wu *et al.*, "Blockchain and Fog Computing in IoT-Driven Healthcare Services for Smart Cities", *Journal of Healthcare Engineering*, Print ISSN: 2040-2295, Online ISSN: 2040- 2309, Vol. 2022, No. 1, 25 January 2022, Art. No. 9957888, Published by John Wiley & Sons Ltd DOI: 10.1155/2022/9957888, Available: https://onlinelibrary.wiley.com/DOI/abs/10.1155/2022/9957888.

[14] Edmira Xhaferra and Florije Ismaili, "The Role of Machine Learning in the Healthcare Sector: A Roadmap to the Potential Prospects", in *Proceedings of International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 09-11 June 2022, Ankara, Turkey, ISBN: 978-1-6654-6835-0, pp. 1-8, 27 June 2022, Published by IEEE, DOI: 10.1109/HORA55278.2022.9800065, Available: https://ieeexplore.ieee.org/document/9800065.

[15] Mehreen Ahmed, Rafia Mumtaz, Syed M. H. Zaidi, Maryam Hafeez, Syed A.R. Zaidi, *et al.*, "Distributed Fog Computing for Internet of Things (IoT) Based Ambient Data Processing and Analysis", *Electronics*, Online ISSN: 2079-9292, Vol. 9, No. 11, 22 October 2020, Art. No. 1759, Published by MDPI DOI: 10.3390/electronics9111756, Available: https://www.mdpi.com/2079-9292/9/11/1756.

[16] Hina Rafique, Munam Ali Shah, Saif Ul Islam, Tahir Maqsood, Suleman Khan *et al.* "A Novel Bio-Inspired Hybrid Algorithm (NBIHA) for Efficient Resource Management in Fog Computing", *IEEE Access,* Online ISSN: 2169-3536, Vol. 7, 26 June 2019, pp. 115760-115773, Published by IEEE, DOI: 10.1109/ACCESS.2019.2924958, Available: https://ieeexplore.ieee.org/document/8746271.

[17] Navjeet Kaur, Ashok Kumar and Rajesh Kumar, "A systematic review on task scheduling in Fog computing: Taxonomy, tools, challenges, and future directions", *Concurrency and Computation: Practice and Experience*, Online ISSN:1532-0634, Print ISSN:1532-0626, Vol. 33, No. 21, 4 June 2021, e6432, Published by John Wiley & Sons Ltd, DOI: 10.1002/cpe.6432, Available: https://onlinelibrary.wiley.com/DOI/abs/10.1002/cpe.6432.

[18] P. G. Shynu, Varun G. Menon, R. Lakshmana Kumar, Seifedine Kadry and Yunyoung Nam, "Blockchain-Based Secure Healthcare Application for Diabetic-Cardio Disease Prediction in Fog Computing", *IEEE Access*, Electronic ISSN: 2169-3536, Vol. 9, 10 March 2021, pp. 45706–45720, Published by IEEE, DOI: 10.1109/ACCESS.2021.3065440, Available: https://ieeexplore.ieee.org/document/9374954.

[19] Vivek Lahoura, Harpreet Singh, Ashutosh Aggarwal, Bhisham Sharma, Mazin Abed Mohammed *et al.*, "Cloud Computing-Based Framework for Breast Cancer Diagnosis Using Extreme Learning Machine", *Diagnostics*, Online ISSN: 2075-4418, Vol. 11, No. 2, 4 February 2021, Art. No. 2, Published by MDPI, DOI: 10.3390/diagnostics11020241, Available: https://www.mdpi.com/2075-4418/11/2/241.

[20] Vijaita Kashyap, Ashok Kumar, Ajay Kumar and You-Chen Hu, "A Systematic Survey on Fog and IoT Driven Healthcare: Open Challenges and Research Issues", *Electronics*, Online ISSN: 2079-9292, Vol. 11, No. 17, 26 August 2022, Art. No. 2668, Published by MDPI, DOI: 10.3390/electronics11172668, Available: https://www.mdpi.com/2079-9292/11/17/2668.

[21] Priyanka Rajan Kumar and Sonia Goel, "Empowering Smart Cities with Fog Computing: A Versatile Framework for Enhanced Healthcare Services and Beyond", *International Journal on Recent and Innovation Trends in Computing and Communication*, ISSN: 2321-8169, Vol. 11, No. 9, 1 October 2023, pp. 335–341, Published by Auricle Global Society of Education and Research, DOI: 10.17762/ijritcc.v11i9.8363, Available: https://ijritcc.org/index.php/ijritcc/article/view/8363.

[22] Ram D. Joshi and Chandra Dhakal, "Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches", *International Journal of Environmental Research and Public Health*, Online ISSN: 1660-4601, Vol. 18, No. 14, 9 July 2021, Art. No. 7346, Published by MDPI, DOI: 10.3390/ijerph18147346, Available: https://www.mdpi.com/1660-4601/18/14/7346.

[23] Umm e Laila, Khalid Mahboob, Abdul Wahid Khan, Faheem Khan and Whangbo Taekeun, "An Ensemble Approach to Predict Early-Stage Diabetes Risk Using Machine Learning: An Empirical Study", *Sensors*, Online ISSN: 1424-8220, Vol. 22, No. 14, 13 July 2022, Art. No. 14, Published by MDPI, DOI: 10.3390/s22145247, Available: https://www.mdpi.com/1424-8220/22/14/5247.

[24] Oumer A. Ebrahim and Getachew Derbew, "Application of supervised machine learning algorithms for classification and prediction of type-2 diabetes disease status in Afar regional state, Northeastern Ethiopia 2021", *Scientific Reports*, Online ISSN: 2045-2322, Vol. 13, No. 1, 13 May 2023, pp. 7779, Published by Springer Nature, DOI: 10.1038/s41598-023-34906-1, Available: https://www.nature.com/articles/s41598-023-34906-1.

[25] J. P. Kandhasamy and S. Balamurali, "Performance Analysis of Classifier Models to Predict Diabetes Mellitus", *Procedia Computer Science*, Online ISSN: 1877-0509, Vol. 47, pp. 45-51, 2015, Published by Elsevier, DOI: 10.1016/j.procs.2015.03.182, Available: https://www.sciencedirect.com/science/article/pii/S1877050915004500.

[26] Linh P. Nguyen, Do Dinh Tung, Duong T. Nguyen, Hong N. Le, Toan Q. Tran *et* al., "The Utilization of Machine Learning Algorithms for Assisting Physicians in the Diagnosis of Diabetes", *Diagnostics*, Online ISSN: 2075-441, Vol. 13, No. 12, 14 June 2023, Art. No. 2087, Published by MDPI, DOI: 10.3390/diagnostics13122087J, Available: https://www.mdpi.com/2075-4418/13/12/2087.

[27] Zhenyi Wang, Wen Dong and Kun Yang, "Spatiotemporal Analysis and Risk Assessment Model Research of Diabetes among People over 45 Years Old in China", *International Journal of Environmental Research and Public Health*, Online ISSN: 1660-4601, Vol. 19, No. 16, 10 August 2022, Art. No. 9861, Published by MDPI, DOI: 10.3390/ijerph19169861, Available: https://www.mdpi.com/1660-4601/19/16/9861.

[28] Christo El Morr, Manar Jammal, Hossam Ali-Hassan and Walid El-Hallak, "Support Vector Machine", in the book series: *International Series in Operations Research and Management Science*, Online ISSN: 2214-7934, Print ISSN: 0884-8289, Vol. 334, 30 November 2022, pp. 385–411, Published by Springer, DOI: 10.1007/978-3-031-16990-8_13, Available: https://link.springer.com/chapter/10.1007/978-3-031-16990-8_13.

[29] Sandra García-Ponsoda, Alejandro Maté and Juan Trujillo, "Refining ADHD diagnosis with EEG: The impact of preprocessing and temporal segmentation on classification accuracy", *Computers in Biology and Medicine*, ISSN: 0010-4825, Vol. 183, 31 October 2024, pp. 109305, DOI: 10.1016/j.compbiomed.2024.109305, Available: https://www.sciencedirect.com/science/article/pii/S0010482524013908.

[30] Methaporn Phongying and Sasiprapa Hiriote, "Diabetes Classification Using Machine Learning Techniques", *Computation*, Online ISSN: 2079-3197, Vol. 11, No. 5, 10 May 2023, Art. No. 96, Published by MDPI, DOI: 10.3390/computation11050096, Available: https://www.mdpi.com/2079-3197/11/5/96.

[31] Ömer F. Akmeşe, "Diagnosing Diabetes with Machine Learning Techniques", *Hittite Journal of Science and Engineering*, Online ISSN: 2148-4171, Vol. 9, No. 1, 30 March 2022, pp. 9–18, Published by Hitit University, DOI: 10.17350/HJSE19030000250, Available: https://dergipark.org.tr/en/pub/hjse/issue/69208/994520.

[32] Mohammad Atif, Faisal Anwer and Faisal Talib, "An Ensemble Learning Approach for Effective Prediction of Diabetes Mellitus Using Hard Voting Classifier", *Indian Journal of Science and Technology* (*INDJST*), Print ISSN: 0974-6846, Online ISSN: 0974-5645, Vol. 15, No. 39, 15 October 2022, pp. 1978–1986, Published by Scientific Research Solution Pvt Ltd., DOI: 10.17485/IJST/v15i39.1520, Available: https://indjst.org/articles/an-ensemble-learning-approach-for-effective-prediction-of-diabetes-mellitus-using-hard-voting-classifier.

[33] Muhammad I. Fazal, Ganjar Alfian, Muhammad Syafrudin and Jongtae Rhee, "Hybrid Prediction Model for Type 2 Diabetes and Hypertension Using DBSCAN-Based Outlier Detection, Synthetic Minority Over Sampling Technique (SMOTE), and Random Forest", *Applied Sciences*, Online ISSN: 2076-3417, Vol. 8, No. 8, 8 August 2018, Art. No. 1325, Published by MDPI, DOI: 2076-3417/8/8/1325, Available: https://www.mdpi.com/2076-3417/8/8/1325.

[34] Batyr Charyyev, Mo Mansouri, and Mehmet H. Gunes, "Modeling the Adoption of Internet of Things in Healthcare: A Systems Approach", in *Proceedings of the IEEE International Symposium on Systems Engineering (ISSE)*, September 13-15 2021, Vienna, Austria, ISBN: 978-1-6654-3168-2, 28 October 2021, pp. 1-8, Published by IEEE, DOI: 10.1109/ISSE51541.2021.9582493, Available: https://ieeexplore.ieee.org/document/9582493.

[35] Shadab Alam, Mohammed Shuaib, Sadaf Ahmad, Dushantha Nalin K. Jayakody, Ammar Muthanna *et al.*, "Blockchain-Based Solutions Supporting Reliable Healthcare for Fog Computing and Internet of Medical Things (IoMT) Integration", *Sustainability*, Online ISSN: 2071-1050, Vol. 14, No. 22, 18 November 2022, Art. No. 15312, Published by MDPI, DOI: 10.3390/su142215312, Available: https://www.mdpi.com/2071-1050/14/22/15312.

[36] Dinesh Bhatia, S. Bagyaraj, S. Arun Karthick, Animesh Mishra and Amit Malviya, "Role of the Internet of Things and deep learning for the growth of healthcare technology", in *Trends in Deep Learning Methodologies: Algorithms, Applications, and Systems*, 2021, ISBN: 9780128222263, ch. 5, pp. 113–127, DOI: 10.1016/B978-0-12-822226-3.00005-2, Available: https://www.sciencedirect.com/science/article/abs/pii/B9780128222263000052.

[37] Maad M. Mijwil, Indu Bala, Ali Guma, Mohammad Aljanabi, Mostafa Abotaleb *et al.*, "Sensing of type 2 diabetes patients based on internet of things solutions: An extensive survey", in *Modern Technology in Healthcare and Medical Education: Blockchain, IoT, AR, and VR*, IGI Global, 2024, ISBN-13: 9798369354933, ch. 3, pp. 34–46, DOI: 10.4018/979-8-3693-5493-3.ch003, Available: https://www.igi-global.com/gateway/chapter/345881.

[38] Sandeep K. Sood and Isha Mahajan, "Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus", *Computers in Industry*, Print ISSN: 0166-3615, Online ISSN: 1872-6194, Vol. 91, October 2017, pp. 33–44, Published by Elsevier B.V., DOI: 10.1016/j.compind.2017.05.006, Available: https://www.sciencedirect.com/science/article/pii/S0166361516303190.

[39] Mahmoud Nasr Milon Islam, Shady Shehata, Fakhri Karray and Yuri Quintana, "Smart Healthcare in the Age of AI: Recent Advances, Challenges, and Future Prospects", *IEEE Access*, Online ISSN: 2169-3536, Vol. 9, 8 October 2021, pp. 145248–145270, Published by IEEE, DOI: 10.1109/ACCESS.2021.3118960, Available: https://ieeexplore.ieee.org/document/9565155.

[40] Kholoud Alatoun, Khaled Matrouk, Mazin Abed Mohammed, Jan Nedoma and Radek Martinek, "A Novel Low-Latency and Energy-Efficient Task Scheduling Framework for Internet of Medical Things in an Edge Fog Cloud System", *Sensors*, Online ISSN: 1424-8220, Vol. 22, No. 14, 16 July 2022, Art. No. 5327, Published by MDPI, DOI: 10.3390/s22145327, Available: https://www.mdpi.com/1424-8220/22/14/5327.

[41] Lisardo Prieto González, Corvin Jaedicke, Johannes Schubert and Vladimir Stantchev, "Fog computing architectures for healthcare: Wireless performance and semantic opportunities", *Journal of Information, Communication and Ethics in Society*, ISSN: 1477- 996X, Vol. 14, No. 4, 14 November 2016, pp. 334–349, Published by Emerald Publishing Limited, DOI: 10.1108/JICES-05-2016-0014, Available: https://www.emerald.com/insight/content/doi/10.1108/jices-05-2016-0014/full/html.

[42] A. Awad Mutlag, Mohd K. A. Ghani, Mazin A. Mohammed, Mashael S. Maashi, Othman Mohd *et al.*, "MAFC: Multi-Agent Fog Computing Model for Healthcare Critical Tasks Management", *Sensors*, Online ISSN: 1424-8220, Vol. 20, No. 7, 27 March 2020, Art. No. 1853, Published by MDPI, DOI: 10.3390/s20071853, Available: https://www.mdpi.com/1424-8220/20/7/1853.

[43] Muhammad Ijaz, Gang Li, Huiquan Wang, Ahmed M. El-Sherbeeny, Yussif Moro Awelisah *et al.*, "Intelligent Fog-Enabled Smart Healthcare System for Wearable Physiological Parameter Detection", *Electronics*, Online ISSN: 2079-9292, Vol. 9, 28 November 2020, Art. No. 12, Published by MDPI, DOI: 10.3390/electronics9122015, Available: https://www.mdpi.com/2079-9292/9/12/2015.

[44] Sadia Din, Anand Paul, Awais Ahmad and Seungmin Rho, "Smart Healthcare in the Age of AI: Recent Advances, Challenges, and Future Prospects", in *Proceedings of the IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC*), 08-12 August 2016, Auckland, New Zealand, pp. 47-54, 13 October 2016, Published by IEEE, DOI 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.23, Available: https://ieeexplore.ieee.org/document/7588818.

[45] Oyeranmi Adigun, Folasade Okikiola, Nureni Yekini and Ronke Babatunde, "Classification of Diabetes Types using Machine Learning", *International Journal of Advanced Computer Science and Applications (IJACSA)*, Print ISSN: 2158- 107X, Online ISSN: 2156-5570, Vol. 13, No. 9, December 2022, pp. 152-161, Published by The Science and Information (SAI) Organization Limited, DOI: 10.14569/IJACSA.2022.0130918, Available: https://thesai.org/Publications/ViewPaper?Volume=13&Issue=9&Code=IJACSA&SerialNo=18.

[46] Israr Ahmad, Saima Abdullah and Adeel Ahmed, "IoT-fog-based healthcare 4.0 system using blockchain technology", *The Journal of Supercomputing*, Print ISSN: 0920-8542, Online ISSN: 1573-0484, Vol. 79, No. 4, 17 September 2022, pp. 3999–4020, Published by Springer Nature, DOI: 10.1007/s11227-022-04788-7, Available: https://link.springer.com/article/10.1007/s11227-022-04788-7.

[47] Navjeet Kaur, Ayush Mittal, Ashok Kumar and Rajesh Kumar, "Healthcare Monitoring Through Fog Computing: A Survey", *ECS Transactions* (ECST), ISSN: 1938-6737, Vol. 107, No. 1, 24 April 2022, pp. 7689, Published by IOP Publishing, DOI: 10.1149/10701.7689ecst, Available: https://iopscience.iop.org/article/10.1149/10701.7689ecst.

[48] Chun-Yang Chou, Ding-Yang Hsu and Chun-Hung Chou, "Predicting the Onset of Diabetes with Machine Learning Methods", *Journal of Personalized Medicine*, Online ISSN: 2075-4426, Vol. 13, No. 3, 24 February 2023, Art. No. 406, Published by MDPI, DOI: 10.3390/jpm13030406, Available: https://www.mdpi.com/2075-4426/13/3/406.

[49] Abd Allah Aouragh, Mohamed Bahaj and Fouad Toufik, "Diabetes Prediction: Optimization of Machine Learning Through Feature Selection and Dimensionality Reduction", *International Journal of Online and Biomedical Engineering (iJOE)*, Online ISSN: 2626-8493, Vol. 20, No. 08, 21 May 2024, pp. 100–114, Published by International Federation of Engineering Education Societies (IFEES), DOI: 10.3991/ijoe.v20i08.47765, Available: https://online-journals.org/index.php/i-joe/article/view/47765.

[50] Hassan Kaleem, Saman Liaqat, Malik T. Hassan, Aneela Mehmood, Umer Ahmad *et al.*, "An Intelligent Healthcare System for Detecting Diabetes Using Machine Learning Algorithms", *Lahore Garrison University Research Journal of Computer Science and Information Technology (LGURJCSIT)*, Print ISSN: 2522-2252, Online ISSN: 2521-0122, Vol. 6, No. 03, 25 July 2022, pp. 1–11, Published by Lahore Garrison University, DOI: 10.54692/lgurjcsit.2022.0603327, Available: https://lgurjcsit.lgu.edu.pk/index.php/lgurjcsit/article/view/327.

[51] Upendra Kumar, Tanay Kumar, Shreya Gautam and Subhash Chandra Pandey, "ML Based IoT Framework for Diabetes Detection", *Research Square*, 2023, Published by Research Square Platform LLC, DOI: 10.21203/rs.3.rs-3024165/v. rs-3024165/v, Available: https://www.researchsquare.com/article/rs-3024165/v1.

[52] Ihab T. Elias and Muna M. T. Jawhar, "Review Classification of Diabetes Using Machine Learning Techniques", *International Research Journal of Innovations in Engineering and Technology (IRJIET)*, Online ISSN: 2581-3048, Vol. 8, No. 1, January 2024, pp. 151–157, DOI: 10.47001/irjiet/2024.801018, Available: https://irjiet.com/Volume-8/Issue-1-January-2024/Review-Classification-of-Diabetes-Using-Machine-Learning-Technics/2071.

[53] Aga Maulana, Farassa R. Faisal, Teuku R. Noviandy, Tatsa Rizkia, Ghazi M. Idroes *et al.*, "Machine Learning Approach for Diabetes Detection Using Fine-Tuned XGBoost Algorithm", *Infolitika Journal of Data Science*, Online ISSN: 3025-8618, Vol. 1, No. 1, 22 August 2023, pp. 1–7, Published by Heca Sentra Analitika, DOI: 10.60084/ijds.v1i1.72, Available: https://heca-analitika.com/ijds/article/view/72/39.

[54] Zain Jabbar and Peter Washington, "The Effect of Data Missingness on Machine Learning Predictions of Uncontrolled Diabetes Using All of Us Data", *Biomedinformatics*, Online ISSN: 2673-7426, Vol. 4, No. 1, 6 March 2024, pp. 780–795, Published by MDPI, DOI: 10.3390/biomedinformatics4010043, Available: https://www.mdpi.com/2673-7426/4/1/43.

[55] Mowafaq S. Alzboon, Mohammad S. Al-Batah, Muhyeeddin Alqaraleh, Ahmad Abuashour and Ahmad F.H. Bader, "Early Diagnosis of Diabetes: A Comparison of Machine Learning Methods", *International Journal of Online and Biomedical Engineering*, Online ISSN: 2626-8493, Vol. 19, No. 15, 25 October 2023, pp. 144–165, Published by International Federation of Engineering Education Societies (IFEES), DOI: 10.3991/ijoe.v19i15.42417, Available: https://online-journals.org/index.php/i-joe/article/view/42417.

[56] Yifan Qin, Jinlong Wu, Wen Xiao, Kun Wang, Anbing Huang *et al.*, "Machine Learning Models for Data-Driven Prediction of Diabetes by Lifestyle Type", *International Journal of Environmental Research and Public Health*, Online ISSN: 1660-4601, Vol. 19, No. 22, 15 November 2022, Art. No. 15027, Published by MDPI, DOI: 10.3390/ijerph192215027, Available: https://www.mdpi.com/1660-4601/19/22/15027.

[57] Ashisha G.R., Mary X. Anitha and Raja J. Mahimai, "Classification of Diabetes Using Ensemble Machine Learning Techniques", *Scalable Computing*, Online ISSN: 1895- 1767, Vol. 25, No. 4, 16 June 2024, pp. 1978–1986, DOI: 10.12694/scpe.v25i4.2873, Available: https://scpe.org/index.php/scpe/article/view/2873.

[58] Salliah S. Bhat, Venkatesan Selvam, Gufran A. Ansari, Mohd D Ansari and Habibur Rahman, "Prevalence and Early Prediction of Diabetes Using Machine Learning in North Kashmir: A Case Study of District Bandipora", *Computational Intelligence and Neuroscience*, Print ISSN: 1687- 5265, Online ISSN: 1687- 5273, Vol. 2022, 4 October 2022, Art. No. 2789760, Published by Hindawi, DOI: 10.1155/2022/2789760, Available: https://onlinelibrary.wiley.com/doi/10.1155/2022/2789760.

[59] Omodunbi Bolaji, Okomba Nnamdi, Olaniyan OM., Esan Arinola and Adewa, T. A, "Development of a Diabetes Melitus Detection and Prediction Model Using Light Gradient Boosting Machine and K-Nearest Neighbour", *UNIOSUN Journal of Engineering and Environmental Sciences (UJEES)*, Online ISSN: 2782-8425, Vol. 3, No. 1, March 2021, Published by Osun State University, DOI: 10.36108/ujees/1202.30.0160, Available: https://ujees.com.ng/volume-3-issue-1/219-2/.

[60] Prajna. P. Debata and Puspanjali Mohapatra, "Diagnosis of Diabetes in Pregnant Woman Using a Chaotic-Jaya Hybridized Extreme Learning Machine Model", *Journal of Integrative Bioinformatics*, ISSN: 1613-4516, Vol. 18, No. 1, 13 August 2020, pp. 81–99, Published by De Gruyter, DOI: 10.1515/jib-2019-0097, Available: https://www.degruyter.com/document/doi/10.1515/jib-2019-0097/html.

[61] Nayeem Ahmed, Syed Imtiyaz Hassan, and Zair Hussain, "Applications of Machine Learning for Diabetes Prediction", *Annals of Computer Science and Information Systems (ACSIS)*, In the *Proceedings of the 2023 Eighth International Conference on Research in Intelligent Computing in Engineering*, December 1–2, 2023, Hyderabad, India, ISSN: 2300-5963, Vol. 38, pp. 1-6, 2023, DOI: 10.15439/2023R43, Available: https://annals-csis.org/Volume_38/drp/43.html.

[62] Harleen Kaur and Vinita Kumari, "Predictive Modelling and Analytics for Diabetes Using a Machine Learning Approach", *Applied Computing and Informatics*, ISSN: 2634-1964, Vol. 18, No. 1/2, 28 July 2020, pp. 90–100, Published by Emerald Publishing Limited, DOI: 10.1016/j.aci.2018.12.004, Available: https://www.emerald.com/insight/content/doi/10.1016/j.aci.2018.12.004/full/html.

[63] Haitham Alshari and Alper Odabaş, "Machine Learning Model to Diagnose Diabetes Type 2 Based on Health Behavior", *Gazi University Journal of Science*, Online ISSN: 2147-1762, Vol. 35, No. 3, 1 September 2022, pp. 834–852, DOI: 10.35378/gujs.931760, Available: https://dergipark.org.tr/en/pub/gujs/issue/69473/931760.

[64] Li Yunjiu, Wang Helin, Ye Zhirui and Zhou Haina, "Diabetes Prediction and Analysis Using Machine Learning Models", in *Proceedings of International Conference on Mechatronics Engineering and Artificial Intelligence (MEAI 2022)*, 11-13 November 2022, Changsha, China, Art. No. 1259615, Published by SPIE (International Society for Optics and Photonics), DOI: 10.1117/12.2672671, Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/12596/2672671/Diabetes-prediction-and-analysis-using-machine-learning-models/10.1117/12.2672671.short.

[65] Amel A. Alhussan, Abdelaziz A. Abdelhamid, S. K. Towfek, Abdelhameed Ibrahim *et al.*, "Classification of Diabetes Using Feature Selection and Hybrid Al-Biruni Earth Radius and Dipper Throated Optimization", *Diagnostics*, Online ISSN: 2075-441, Vol. 13, No. 12, 12 June 2023, Art. No. 2038, Published by MDPI, DOI: 10.3390/diagnostics13122038, Available: https://www.mdpi.com/2075-4418/13/12/2038.

[66] Fathima M. D. Mohiden, Justin S. S. Raj and Raja S. P. Raj, "Regression Imputation and Optimized Gaussian Naïve Bayes Algorithm for an Enhanced Diabetes Mellitus Prediction Model", *Brazilian Archives of Biology and Technology*, ISSN: 1678-4324 Vol. 64, 21 February 2022, DOI: 10.1590/1678-4324-2021210181, Available: https://www.scielo.br/j/babt/a/3HDJLgqSPVYcSDT4RFY9wzp/?lang=en.

[67] Singla Rajiv, Aggarwal Shivam, Bindra, Jatin, Garg Arpan and Singla Ankush, "Developing Clinical Decision Support System Using Machine Learning Methods for Type 2 Diabetes Drug Management", *Indian Journal of Endocrinology and Metabolism*, Print ISSN: 2230-8210, Online ISSN: 2230-9500, Vol. 26, No. 1, 27 April 2022, pp. 44–49, Published by Wolters Kluwer, DOI: 10.4103/ijem.ijem_435_21, Available: https://pubmed.ncbi.nlm.nih.gov/35662766/.

[68] Hao Huang, Min Zhao, Xiyang Liu, Jialin Song, Lina Liu *et al.*, "Glutathione combined with mecobalamin in the treatment of chemotherapy-induced peripheral neuropathy in multiple myeloma: a retrospective clinical study", *Annals of Palliative Medicine*, Print ISSN: 2224-5820; Online ISSN: 2224-5839, Vol. 10, No. 12, 31 December 2021, pp. 12335–12346, Published by AME Publishing Company, DOI: 10.21037/apm-21-3313. Available: https://apm.amegroups.org/article/view/85557/html.

[69] Hsin-Y. Tsao, Pei-Y. Chan and Emily C.-Y. Su, "Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms", *BMC Bioinformatics*, ISSN: 1471-2105, Vol. 19, No. 9, 13 August 2018, Article No. 283, Published by Springer Nature, DOI: 10.1186/s12859-018-2277-0, Available: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2277-0.

[70] Mahesh K. Khanal, Pratiksha Bhandari, Raja R. Dhungana, Yadav Gurung, Lal B. Rawal *et al.*, "Electrocardiogram abnormalities and renal impairment in patients with type 2 diabetes mellitus: A healthcare facilities-based cross-sectional study in Dang district of Nepal", *Journal of Diabetes Investigation*, Print ISSN: 2040-1116, Online ISSN: 2040-1124, Vol. 14, No. 4, 6 February 2023, pp. 602–613, Published by John Wiley & Sons Ltd, DOI: 10.1111/jdi.13985. Available: https://onlinelibrary.wiley.com/doi/10.1111/jdi.13985.

[71] Abdullah A. Al-Atawi, Saleh Alyahyan, Mohammed N. Alatawi, Tariq Sadad, Tareq Manzoor *et al.*, "Stress Monitoring Using Machine Learning, IoT and Wearable Sensors", *Sensors*, Online ISSN: 1424-8220, Vol. 23, No. 21, 31 October 2023, Art. No. 8875, Published by MDPI, DOI: 10.3390/s23218875, Available: https://www.mdpi.com/1424-8220/23/21/8875.

[72] FanWu, Taiyang Wu and Mehmet R. Yuce, "An Internet-of-Things (IoT) Network System for Connected Safety and Health Monitoring Applications", *Sensors*, Online ISSN: 1424-8220, Vol. 19, No. 1, 21 December 2018, Art. No. 1, Published by MDPI, DOI: 10.3390/s19010021, Available: https://www.mdpi.com/1424-8220/19/1/21.

[73] Mirza A. Khatun, Sanober F. Memon, Ciarán Eising and Lubna Luxmi Dhirani, "Machine Learning for Healthcare-IoT Security: A Review and Risk Mitigation, *IEEE Access*, Online ISSN: 2169-3536, Vol. 11, 22 December 2023, pp. 145869-145896, Published by IEEE, DOI: 10.1109/ACCESS.2023.3346320, Available: https://ieeexplore.ieee.org/document/10371310.

[74] Sohail Jabbar, Farhan Ullah, Shehzad Khalid, Murad Khan and Kijun Han, "Semantic Interoperability in Heterogeneous IoT Infrastructure for Healthcare", *Wireless Communications and Mobile Computing*, Online ISSN: 1530-8677, Vol. 2017, No. 1, 05 March 2017, Art. No. 9731806, Published by John Wiley & Sons Ltd, DOI: 10.1155/2017/9731806, Available: https://onlinelibrary.wiley.com/doi/10.1155/2017/9731806.

[75] Mohammad Wazid, Ashok K. Das and Youngho Park, "Blockchain-enabled secure communication mechanism for IoT-driven personal health records", *Transactions on Emerging Telecommunications Technologies*, Print ISSN: 2161- 3915, Online ISSN: 2161- 3915, Vol. 33, No. 4, 19 December 2021, Art No. 4421, Published by John Wiley & Sons Ltd, DOI: 10.1002/ett.4421, Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.4421.

[76] Yaru Liu, Jia Yu, Jianxi Fan, Pandi Vijayakumar and Victor Chang, "Achieving Privacy-Preserving DSSE for Intelligent IoT Healthcare System", *IEEE Transactions on Industrial Informatics*, Print ISSN: 1551-3203, Electronic ISSN: 1941-0050, Vol. 18, No. 3, March 2022, pp. 2010-2020, Published by IEEE, DOI: 10.1109/TII.2021.3100873, Available: https://ieeexplore.ieee.org/document/9502532.

[77] Omar Said, "LBSS: A Lightweight Blockchain-Based Security Scheme for IoT-Enabled Healthcare Environment", *Sensors*, Online ISSN: 1424-8220, Vol. 22, No. 20, 18 October 2022, Art. No. 7948, Published by MDPI, DOI: 10.3390/s22207948, Available: https://www.mdpi.com/1424-8220/22/20/7948.