Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Hybrid Weighed Quantum Lion Optimization With Enhanced Fast Track Gram Matrix-Independent Component Analysis To Improve The Performance Analysis Of Predicting Diabetics At Early Stage

Arun. D[1], Dr. R. Annamalai Saravanan[2]

[1] *Research Scholar, Department of Computer Science, Sankara College of Science and Commerce, Bharathiar University, Saravanampatti, Coimbatore.*
[2] *Associate Professor, Department of Information Technology, Hindusthan College of Arts & Science, Coimbatore.*
[1] *Corresponding Author : arun.divakaran@hotmail.com*

One of the most widely prevalent ailments across the globe is DM. Diabetic sufferers in India have increased drastically in the last few decades. Recent data indicates that 72.96 million young men and women have Diabetics Mellitus (DM). As a result, early detection of DM is crucial. Sometimes referred to as prolonged ailment, is a collection of digestive disorders brought on by persistently elevated blood glucose levels. If accurate early prediction is feasible, the probability parameter and intensity of the disease could be greatly decreased. There are few annotated information points and aberrations (or missing values) in the illness databases is very difficult to make a valid and precise Diabetics prognosis. A novel approach, the hybrid Weighed Quantum Lion Optimization with Enhanced Fast Track Gram Matrix-Independent Component Analysis (WQLO-EF-GM-ICA), aimed at enhancing the performance analysis of predicting DM at an early stage. DM, a chronic metabolic disorder, poses significant health risks if not detected and managed promptly. Traditional methods of prediction often rely on ML techniques, but challenges remain in terms of accuracy and efficiency. In response, this study proposes a hybrid methodology that combines WQLO, a bio-inspired optimization algorithm, with EF-GM-ICA, an enhanced version of the Gram Matrix-Independent Component Analysis approach. WQLO offers robust optimization capabilities, while EF-GM-ICA streamlines feature extraction from high-dimensional data. By integrating these techniques, the hybrid approach aims to improve prediction accuracy and efficiency in identifying individuals at risk of DM at an early stage. Through rigorous experimentation and performance analysis, the effectiveness of the proposed methodology is evaluated, offering insights into its potential contributions to DM prediction and early intervention strategies.

Keywords: Hybrid optimization; Weighed Quantum Lion Optimization; Enhanced Fast Track Gram Matrix-Independent Component Analysis; Diabetics prediction; Early stage detection; Performance analysis; Bio-inspired algorithms; Feature extraction; Chronic metabolic disorder.

1. Introduction

Since several centuries ago, researchers have been investigating a chronic disease namely DM. The term "Diabetics" was originally used by the physician Aretaeus in the fifth century, who described the illness as a "melting down of limbs and flesh into urine" [1]. The sweet, honey-like flavor of polyuric patients' urine which attracted insects like ants, etc. was identified by Indian doctors in the fifth century BC; nevertheless, the term "Mellitus," which is Latin for "honey," was added in the seventeenth century. DM is a chronic, non-contagious circumstance characterized by partial or total shortages of insulin. Although the causes of this illness might vary greatly, they are always related to either the body's cells' poor response to insulin or the pancreas' inability to secrete enough insulin, or both at some stage throughout the illness [2].

The development of DM is strongly linked to the main problems in the metabolism of fat, protein, carbohydrate, and glucose. Diabetics is a major public health and societal problem today due to its prevalence, distinctive features, and the frequency of other illnesses that commonly coexist with it [3]. Suitable treatments for DM have grown into a top priority in the field of medicine. It is critical to take action to diagnose diseases early on and to prevent them in addition to treating them. Prognostic modeling in DM detection is achieved using the methods used in data mining, ML, or any other AI sector. This involves using information and expertise to derive potential results based on historical data [4]. To reduce the risk of sickness, these strategies are used for the earlier prognosis and identification of DM. The most common symptoms of DM include hyperglycemia, abnormal metabolic processes, and an increased risk of specific issues about the eyes, kidneys, and nervous system—all major organs in human anatomy shown in Figure 1. These indicators are used to gather medical information, and indicators are then used to classify the information obtained [5].

Numerous medical professionals and academics have put out various DDS that use cutting-edge technology to identify DM. These systems are growing increasingly effective in design with technology improvements [6]. Forecasting the onset of DM usually proves a challenging endeavor. These methods take advantage of the improvements in technology that enable robots to learn. The difficulty in building these kinds of systems is that, whether or not human interaction is involved, one must concentrate on a distinct set of data patterns (i.e., the data that already exists) and inevitably draw new conclusions from the preliminary statistics [7].

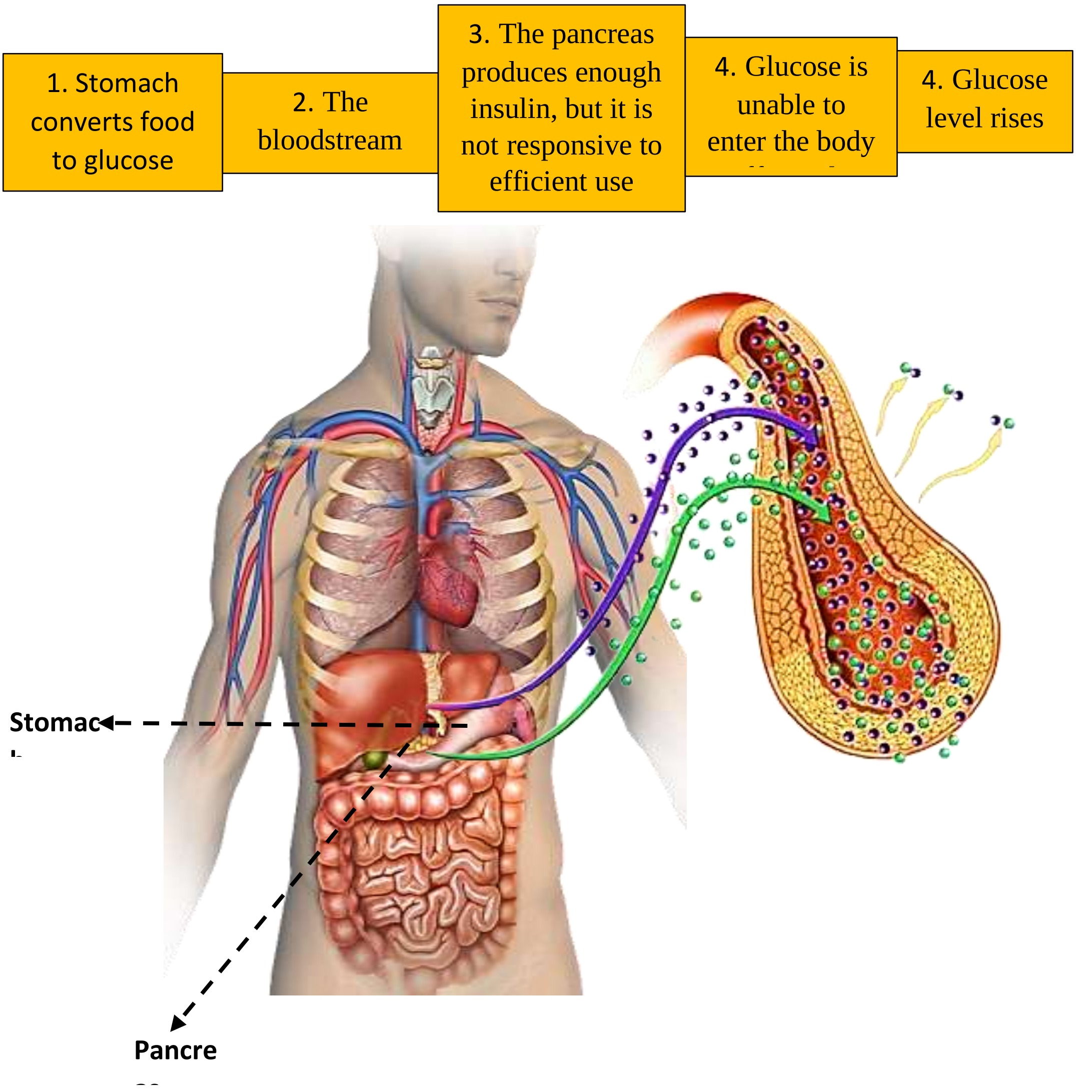

Figure 1: Diabetic mellitus

The Internet of Medical Things, Machine Learning (ML), data mining, artificial intelligence, and other innovative technologies have had an enormous effect on the medical field. The aforementioned approaches have proven successful in predicting diseases from vast amounts of medical data [8]. Preventative simulation, or the use of statistics and data to predict future outcomes based on historical experiences is shown in Figure 2. DM care can be made more successful by extending its reach through the application of ML techniques. In four key areas of DM care—medical assistance with decisions, prognostic population risk stratification, automatic vision examination, and self-management for patient tools—ML techniques are widely applied [9]. It resulted in a change in the way data-driven healthcare was provided to control DM. It has changed how DM is detected, treated, and averted, which may contribute to a decrease in the 8.8% global distribution of the disease.

Effective ML techniques are being utilized to create programs that assist frameworks for the early diagnosis of DM or its associated issues. A continuous, error-free method for tracking biomarkers and symptoms is made possible by ML. Many ML techniques, such as supervised, reinforcement learning, and unsupervised techniques, have been put forth. Given that ML approaches rely on data, this is viable [10]. ML may spare a significant amount of human labor because of the massive number of health-related information that is given to the diagnostic engine. Using such information as training, algorithms have been instructed in the ML approach to provide output that is more precise than the input information. Certain people may make use of the individual's blood sample info, while others might examine characteristics of clinical scans [11]. The criteria vary according to the multiple indications of the disease. Using several well-established methods, researchers have experimented with multiple algorithms and adjusted a great deal of hyperparameters to produce results that appear most suitable for real-time applications [12].

There are two stages to the SVM-based classification process: learning and testing. Using training data samples, this sorting method finds the predicted characteristics during the learning stage. In the stage of testing, the system assigns tags to the corresponding group and uses those attributes to classify the just received data set. An effective non-parametric classifier, SVM can handle both linear and non-linear functions to solve regression and classification problems [13]. Because of its great ability to deal with non-linear, massive databases in the field of medicine, this algorithm is a discriminative classification tool that can be used to identify biological sign anomalies. The grouping method's fundamental idea is to classify newly discovered testing datasets into appropriate groups based on identified learning datasets [14].

With today's technologies, making a diagnosis is a major issue. Diabetics is the colloquial term for DM. Diabetes is a tough and burdensome disorder that impacts individuals worldwide. A person's typical metabolic reactions before and after meals are measured to determine the DM's degree of severity [15]. Food usually passes along through the gut after it's taken up into the circulatory system and used as a source of energy. A component called blood sugar is utilized to indicate the presence of an endocrine pancreas, which secretes insulin, a hormone. Almost every kind of cell in the human anatomy requires insulin for the absorption of glucose, which is then stored in the liver, muscles, and fat tissues. Diabetics cannot be cured, but their complications can be reduced with adequate nutrition, physical activity, and blood glucose management medicine [16].

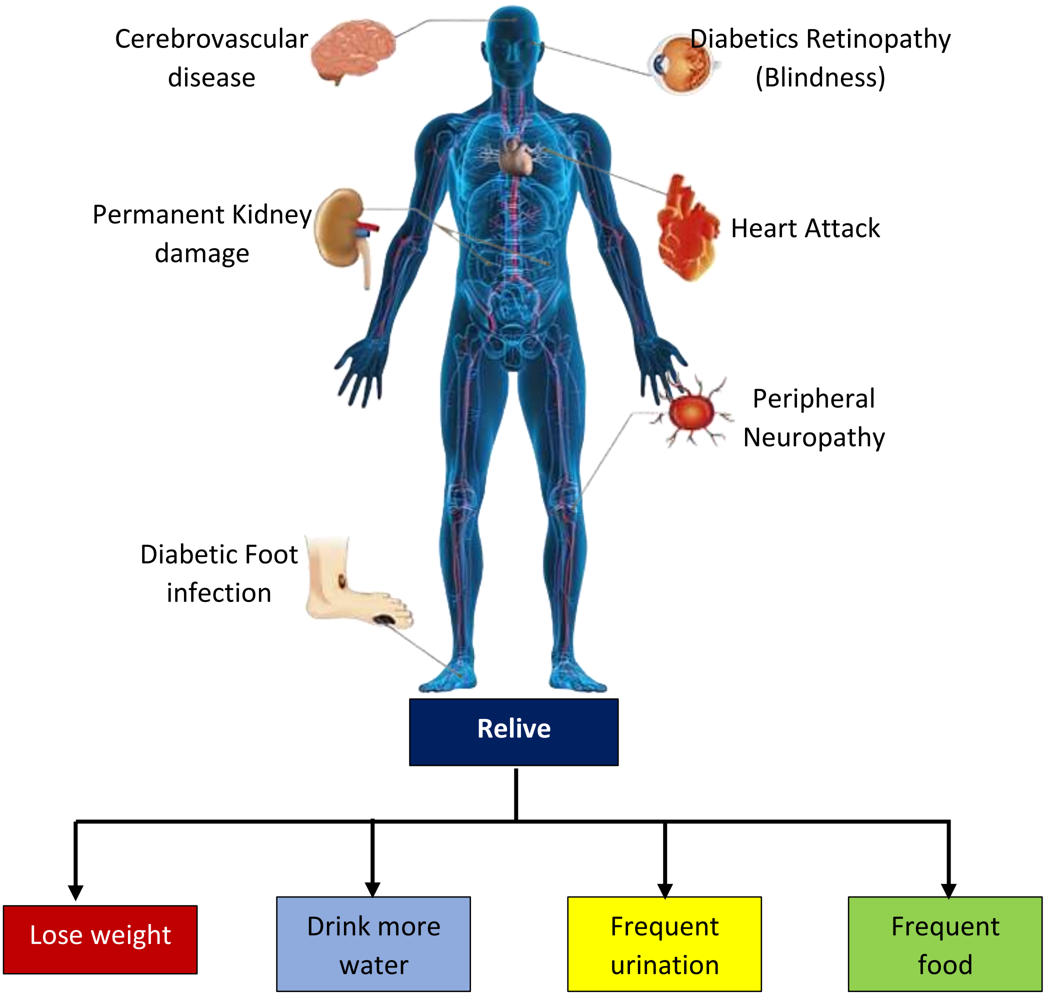

Figure 2: Symptoms and causes

The majority of what we eat is transformed into sugar, commonly referred to as glucose, and then absorbed into the bloodstream of the body. An increase in glucose levels causes the pancreas to secrete insulin. Insulin is either not used efficiently by the human body or is not produced in sufficient amounts. Because of this, glucose does not enter cells and instead stays in our circulation. Over some time, issues related to high blood glucose levels can arise [17].

ML algorithms used in medicine are primarily concerned with diagnosing and detecting diseases. ML and CI techniques can enhance healthcare by providing decision-making instruments to assist physicians in diagnosing and predicting medical issues. In recent years, ML has been used more often in the recognition of patterns and healthcare imaging for computerized detection. Reported that state-of-the-art ML approaches are extensively being employed to address issues employing deep neural networks and variants of monitored, semi-supervised, and unsupervised learning techniques [18].

1.1 Problem Statement

The problem statement revolves around the challenge of accurately predicting diabetes at an early stage, a critical aspect of managing the condition and reducing associated health risks. DM characterized by abnormal blood sugar levels, poses significant health complications if left undiagnosed or untreated. Early detection allows for timely intervention and lifestyle modifications, which can mitigate the progression of the disease and its complications. However, existing prediction methods often lack the accuracy and efficiency required for early-stage detection. ML techniques have shown promise in this domain, but challenges remain in terms of achieving high prediction accuracy while minimizing false positives and false negatives. Therefore, the problem statement involves developing advanced predictive models or algorithms capable of accurately identifying individuals at risk of diabetes at an early stage, thereby facilitating timely intervention and improving health outcomes.

1.2 Motivation

Improved Health Outcomes: Early detection of diabetes can lead to better health outcomes by enabling timely intervention and management strategies, which can help prevent or delay the onset of diabetes-related complications.

Preventive Healthcare: Early prediction allows for the implementation of preventive measures, lifestyle modifications, and medical interventions tailored to individual risk profiles, promoting healthier behaviors and reducing the burden of disease.

Optimized Patient Care: Healthcare providers can tailor interventions and treatment plans based on individual risk profiles, optimizing patient care and resource allocation to effectively manage and mitigate the progression of diabetes.

Public Health Impact: Early detection of diabetes contributes to disease surveillance, epidemiological studies, and healthcare planning, enabling policymakers and healthcare authorities to implement targeted prevention and intervention programs to address the growing burden of diabetes on healthcare systems.

Technological Advancements: Advancements in technology and data analytics present opportunities to develop more accurate and efficient predictive models for diabetes, leveraging machine learning algorithms, bioinformatics, and big data analytics to explore novel biomarkers and genetic factors contributing to early detection.

Interdisciplinary Collaboration: Collaboration between healthcare professionals, researchers, and technology experts facilitates interdisciplinary approaches to diabetes prediction, fostering innovation and the development of novel strategies to improve early detection methods.

Personalized Medicine: Early prediction of diabetes allows for personalized medicine approaches, where interventions and treatments can be tailored to individual risk factors and genetic predispositions, optimizing patient outcomes and reducing healthcare costs.

Quality of Life: Early detection and management of diabetes improve the quality of life for individuals at risk by empowering them with knowledge, resources, and support to make informed decisions about their health and well-being.

Cost-Effectiveness: Early prediction and intervention strategies are cost-effective in the long term, as they reduce the need for expensive medical treatments and hospitalizations associated with diabetes-related complications.

Public Awareness and Education: Promoting awareness and education about the importance of early detection of diabetes encourages individuals to seek preventive healthcare services, leading to early diagnosis and improved management of the condition.

2 Related Works

Medical experts are shocked by the huge and unexpected rise of patients during this COVID-19 outbreak, especially doctors and nurses. AI techniques could be used in these situations to determine whether a person has a dangerous condition. Particularly, conditions like DM, elevated BP, and cardiovascular disease that increase the risk of cardiac arrest and death in COVID-infected patients should be detected early. Diabetic diagnosis and treatment involve a lot of statistics, thus applying ML techniques to enhance outcomes and choose cutting-edge approaches makes sense [19]. Multiple methods of ML were employed by clinicians and researchers in the chosen primary DDS research to provide increased efficiency and precise outcomes. Nonetheless, a poll needs to be conducted to determine the best approach. This study looks for areas where ML methods might be applied to classify DM and identify possibilities. Additionally, it examines the most effective assessment metrics, the information systems used in creating models, and the optimization techniques used to boost categorization process efficacy [20].

DM has become one of the most significant societal and health hazards of our day. A significant part of the therapeutic endowment is played by the instruments and methods utilized to obtain an accurate and timely diagnosis of the diabetic condition. Additionally, this might have a big impact on the welfare of society and life pleasure. Various DDS have been recommended by several academics and physicians to diagnose diabetic mellitus utilizing cutting-edge technologies [21].

A poll is still necessary to determine the best Diabetic diagnosis strategy, as several have been developed and applied employing AI and ML techniques contrary to Diabetic information available in the literature. It is often a difficult undertaking to predict the onset of Diabetics. The most recent revolutionary innovations that enable machines to learn are exploited by modern DDSs. The difficulty in creating such systems is that, whether or not human interaction is involved, one must concentrate on a distinct set of data patterns and inevitably derive new conclusions from the preliminary statistics. The best ML algorithm for creating DDS is usually chosen based on how well it performs and how accurately it classifies data [22].

AI techniques have been used in several research endeavors to develop computer-assisted DDS. Certain illnesses have similar symptoms, the process of diagnosis is frequently troubling, and optimization techniques are necessary to deal with these problems successfully. The method of diagnosing DM is an issue with the categorization of info. Categorization, as seen via the perspective of ML, is the task of assigning an array of data to suitable groups according to the learning outcome of a subset of the dataset whose proper class is identified [23]. For better accuracy of classification, metaheuristic optimization techniques are often used to produce effective DDS. Most metaheuristic techniques draw inspiration from physical, biological, or natural principles and try to emulate the same, using various limitations at an elementary level. Any metaheuristic planner starts with a random candidate solution and iteratively works its way up to the best outcome [24].

Scientists and doctors are currently very interested in ML methods because of the availability of large resources of info that have been merged with improved algorithms and machines with higher computational capacities. ML is the technical field that deals with how computers learn from previous performance. Although they aren't synonymous, many academics have employed the phrases "ML" and "artificial intelligence" interchangeably because intelligence is primarily defined by an object's capacity for acquiring knowledge. The goal of artificial intelligence is to develop a smart helper or agent that uses different ML approaches [25]

At the moment, ML techniques are successfully used for large, especially high-dimensional information regression, clustering, classification, or dimensionality reduction procedures. ML has proven to be exceptionally skilled in several fields, such as image analysis, driverless vehicles, and online gaming. Therefore, effective ML approaches power many aspects of our daily lives, including identifying fraudulent transactions, credit history reporting, photo and audio recognition, searches on the internet, email/spam sorting, etc. [26]. According to [27], there are three main categories into which ML approaches fall: (i) predictive or supervised learning, in which a given function is inferred by the machine using a categorized training database; (ii) descriptive or unsupervised learning, the tutoring stage aims to infer the structure of an unidentified sample of information; and (iii) reinforcement learning, in which data is gathered through system communication with an evolving setting.

For the diagnosis of DM, SVM-based classifiers are used in multiple studies. Additionally, these scientists reported that SVM-based models performed better in categorization. Presented a painless and bloodless SVM-based classifier to identify ocular abnormalities caused by diabetic macular degeneration. The generated model's rate of classification is 91.67% for blood vessels and discharge, and 83.33% for optic discs and micro aneurysms. Using the periodically validated longitudinal database, proposed an organized result Support Vector Machine (SVM) method for determining the early stage of T1DM [28].

Of the eight initial characteristics provided in the collection of data, they have achieved 97.5% precision with six beneficial characteristics. Additionally, they contrasted their system against other similar techniques [29]. Using increased accuracy, produced a forecasting device that facilitates early Diabetics identification and prediction. Numerous in-depth investigations of data mining methods such as DT, NB, ANN, and DL were conducted on the PIMA datasets. They originally acquired the varied categorization model, and then they evaluated the model. By employing DL, they were able to achieve a precision rate of 98.07% [30].

2.1 Research Gap

Limited Integration of Optimization Techniques: While there has been significant research on optimization algorithms and techniques for predictive modeling, there is a gap in integrating multiple optimization approaches into a cohesive framework for diabetes prediction. Existing studies often focus on individual optimization methods without exploring the potential synergies of combining them to enhance prediction accuracy.

Sparse Utilization of Advanced Feature Extraction Techniques: Although feature extraction techniques such as Independent Component Analysis (ICA) have shown promise in extracting relevant features from high-dimensional data, their integration into predictive modeling for diabetes prediction is relatively limited. There is a research gap in exploring and optimizing the use of advanced feature extraction techniques within the context of early-stage diabetes prediction.

Lack of Focus on Early-Stage Diabetes Prediction: Many existing studies on diabetes prediction primarily focus on diagnosing diabetes at later stages or managing the condition rather than predicting it at an early stage. There is a gap in research specifically targeting early-stage prediction, which is crucial for timely intervention and prevention of diabetes-related complications.

Insufficient Validation Studies: While some studies propose novel predictive models or algorithms for diabetes prediction, there is often a lack of comprehensive validation studies using real-world datasets and robust evaluation metrics. There is a research gap in conducting rigorous validation and evaluation of proposed methodologies to assess their effectiveness and generalizability in real-world clinical settings.

Limited Application of Hybrid Optimization Techniques in Healthcare: Although hybrid optimization techniques have been explored in various domains, their application in healthcare, particularly for predictive modeling in chronic diseases like diabetes, is relatively sparse. There is a research gap in investigating the feasibility and efficacy of hybrid optimization approaches specifically tailored to healthcare applications, including early-stage diabetes prediction.

Addressing Computational Complexity: Predictive modeling for early-stage diabetes prediction often involves analyzing large and complex datasets, which can pose computational challenges. There is a research gap in developing efficient algorithms and optimization techniques to address the computational complexity associated with analyzing such datasets while maintaining prediction accuracy and scalability.

Addressing these research gaps can lead to the development of more accurate, efficient, and scalable predictive models for early-stage diabetes prediction, ultimately contributing to improved healthcare outcomes and reduced burden of diabetes-related complications.

3 Proposed System

The proposed pipeline's crucial step is the preliminary processing of the data in its raw form since this allows the classifiers to learn directly from the data quality. This approach is demonstrated in Figure 3 of the literature.

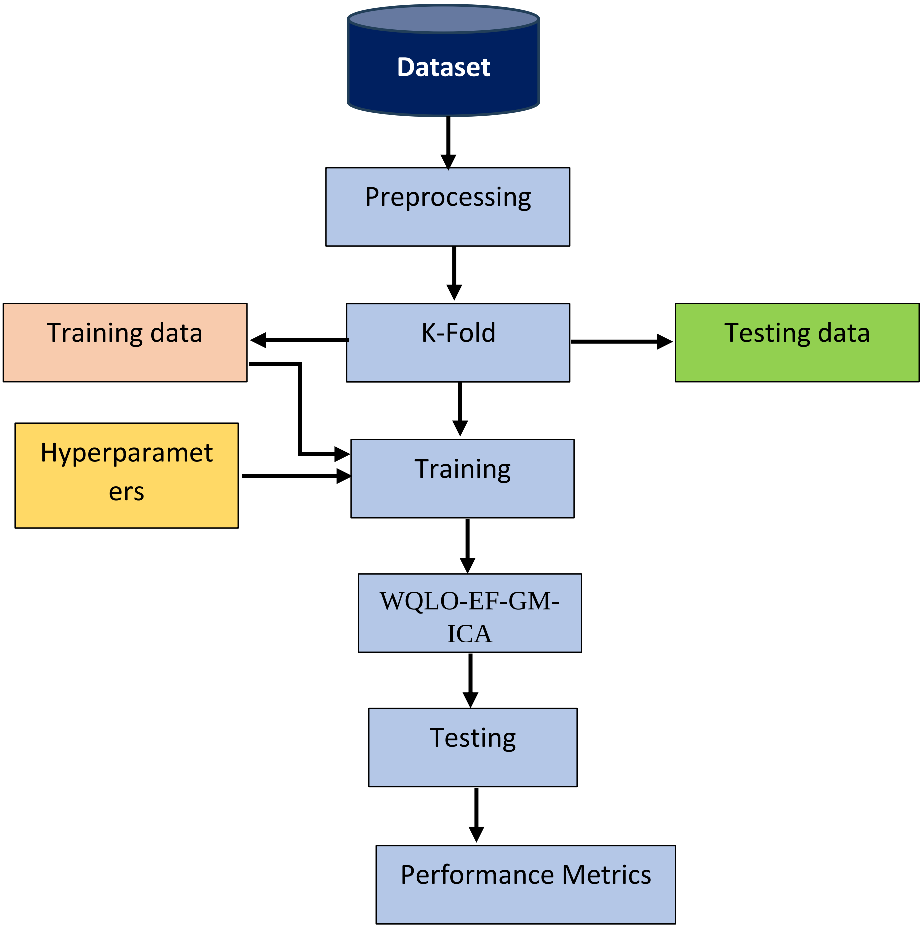

Figure 3: The proposed architecture to detect and predict DM

3.1 Dataset

The first and foremost prerequisite for carrying out the ML method is the collection of data. Scholars have recently used a range of information for their experiments and studies, including pre-existing ones and data that they gathered from different medical agencies and healthcare service providers. Numerous Diabetic databases are accessible for various Diabetic prediction purposes. The most extensively utilized dataset currently in existence is the Pima Indians Diabetics Dataset (PIDD). In this study, an experimental assessment is conducted using a pair of databases. The information set's purpose was to forecast a patient's likelihood of having DM using diagnostic information gathered from the database. This dataset includes 8000 patients without Diabetics (negative) and 589 patients with Diabetics (positive), each with eight distinct features. Table 1 displays the dataset attributes. DM Pedigree Function (DPF) computed shown in Equation (1).

$$\text{Pedigree} = \frac{\sum_x k_x(88-\text{ADM}_x)+20}{\sum_y k_y(14-\text{ALC}_y)+50} \quad (1)$$

Where the relatives who had and had not got Diabetes are indicated by x and y respectively.

Table 1: PIDD dataset features

S.No	Attribute name	Description	Value
1	Class	Label used in class for prediction	0- DM not present 1- DM present
2	np	Pregnant count	0-18
3	Age	Age	21-80 years
4	DPF	Diabetics pedigree function	0.0079-1.721
5	PGC	Plasma glucose concentration (after 2 hrs)	45-198 mg/dl
6	BMI	Body mass index (kg/m^2)	18.5-59 kg/m^2
7	dbp	Diastolic blood pressure (mmHg)	25-123 mmHg
8	si	2-h serum insulin (mU/ml)	15-512 mU/M1
9	test	Triceps skinfold thickness (mm)	

3.2 Pre-processing

The information needs to be processed in advance to remove any missing values or zeros once it is successfully collected. The preliminary processing is seen to be the very crucial stage since it establishes the viability or failure of the proposed model. The following is a list of crucial elements to take into account before beginning the pre-processing stage.

i) Variable and comparable outcomes

ii) Missing or null values

iii) Several examples of the same type

iv) An unparalleled variation

It has been noted that only a small number of entries were absent from the databases, and in certain instances, multiple values for attributes remained absent. In this instance, the attribute's average value is used to fill in the missing value. Individuals without a diagnosis of Diabetics and those with a diagnosis were assessed individually for mean values. The correctness of the dataset's predictions is impacted when the average value is substituted, as every parameter has a distinct gap.

$$\text{Replaced value} = \frac{\text{Old}-\text{value(feature)}-\text{Min(feature)}}{\text{Max(feature)}-\text{Min(feature)}} \times (\text{new\_Max(feature)} - \text{new\_Min(feature)}) + \text{new\_Max(feature)} \quad (2)$$

With every increment in the overall size of the feature, the classifiers' accuracy rises. As the dimension of the characteristic rises with no increase in sample sizes, the detectors' efficiency will nonetheless begin to decline. Owing to the constraint of dimensionality, the characteristic area gets increasingly less dense and drives the overfitting of the detectors by reducing their capacity for generalization.

3.3 Weighed Quantum Lion Optimization

Weighed Quantum Lion Optimization (WQLO) is the main tool used in this study to improve Diabetic prognosis estimation. DM is an ongoing metabolic condition that carries serious issues with health if it fails to be diagnosed and treated early. To avoid or postpone the emergence of Diabetics and its related challenges, preliminary identification of the disease is essential for putting prompt therapies and lifestyle changes into place.

The stalking and capturing techniques of lions in their habitat served as the model for the bio-inspired optimization algorithm known as WQLO. It quickly explores complex domains of search and finds the best possible outcomes by utilizing quantum computing concepts and quantum-inspired evolutionary algorithms. The study attempts to increase the precision, effectiveness, and resilience of predictive models by using WQLO to the challenge of early Diabetic prediction.

Improving the accuracy of predictions with WQLO, lowering computing difficulty, facilitating swift action, and optimizing the choice of features are the goals of the project. The algorithm will be used to examine sizable, diverse databases that include demographic, way of life, hereditary, and diagnostic information about Diabetics. The goal of the investigation is to find pertinent characteristics, sequences, and indicators connected to the initial stages condition by utilizing WQLO's optimization powers.

Several steps will be involved in the study, such as an evaluation of the literature, the construction of a procedure, the setup of datasets, experimenting, approval, and results analysis. The forecasting algorithms will be trained and assessed using actual data sets, and the results will be compared with current methods to determine whether WQLO is effective in enhancing performance analysis for Diabetic prognosis.

The results of the study will strengthen advanced forecasting in the medical field, especially as it relates to preliminary Diabetics prognosis. The goal of the project is to increase the standard of life for those who are in danger of acquiring DM, lower the cost of consequences connected to the disease, and boost healthcare results by creating a useful and efficient tool for preliminary Diabetic prognosis.

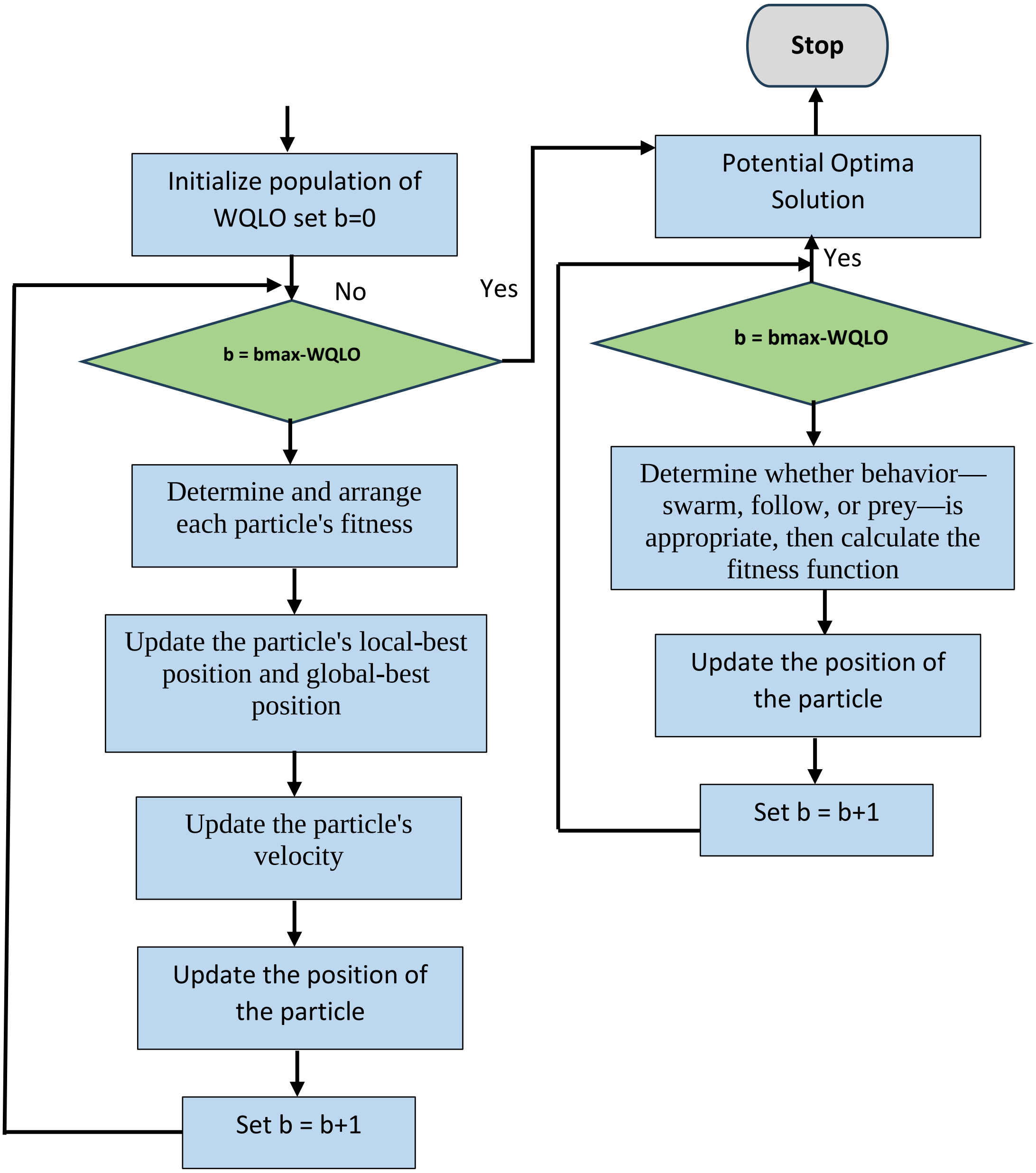

Figure 4: Flowchart of WQLO

The WQLO method is employed to start the worldwide hunt for Type-II Diabetics. The WQLO method starts the worldwide hunt with the ultimate optimal number of people from WQLO and ends it with a local search. Figure 4 illustrates the proposed WQLO method.

Devised a bio-inspired metaheuristic technique that is the standard Lion optimization process. WQLO imitates the commanding structure and hunting strategies of lions. They belong to the Felines family and have a demanding social hierarchy. Their method of bullying, which centers on lion intelligence, is widely recognized. In pride of five to twelve lions, they prowl in pride and are social predators that use cooperation to find their prey or enemies. The control is divided into four hierarchies inside the cluster: omega (ω), beta (β), delta (δ), and alpha (α) Lions. The top animal in the pride, the alpha (dominant) lion leads the pride in performance. It is responsible for all the decisions regarding punishment, waking time, napping location, and other aspects of the stalking operation. The entire group is informed about its choices and obliged to abide by them.

These stages are illustrated in Figure 5. Equations (3) and (4) are utilized to depict the lions' neighborhood conduct.

$$\vec{N} = \left|\vec{K}.\vec{P}_{prey}(x) - \vec{P}_{SA}(x)\right| \quad (3)$$

Where x denotes the present iteration. $\vec{P}_{prey}$ and $\vec{P}_{SA}$ are the position vectors of the enemy and searching agent, respectively. The terms $\vec{M}$ and $\vec{K}$ are constant vectors computed using Equations (5) and (6)

$$\vec{P}_{Lion}(x+1) = \vec{P}_{prey}(x) - \vec{M}.\vec{N} \quad (4)$$

$$\vec{M} = 2.\vec{m}.\vec{a_1} - \vec{m} \quad (5)$$

$$\vec{K} = 2.\vec{a_2} \quad (6)$$

Where $\vec{a_1}$ and $\vec{a_2}$ have random values in [0, 1]. The term acts as a regulating component and linearly decreases from 2 to 0 all along with the iterations.

Given that the exploratory agents can locate the opponent, all they need to do is surround it. Alpha oversees the entire procedure. All the lions in the group pride each other and evaluate their places based on where α, β, and δ are most optimally positioned. The following Equations (7) – (14) describe the surrounding activity of exploratory agents:

$$\vec{N_\alpha} = \left|\vec{K}_1.\vec{P}_\alpha - \vec{P}\right| \quad (7)$$

$$\vec{N_\alpha} = \left|\vec{K}_1.\vec{P}_\alpha - \vec{P}\right| \quad (8)$$

$$\vec{N_\beta} = \left|\vec{K}_2.\vec{P}_\beta - \vec{P}\right| \quad (9)$$

$$\vec{N_\delta} = \left|\vec{K}_3.\vec{P}_\delta - \vec{P}\right| \quad (10)$$

$$\vec{P_1} = \left|\vec{P}_\alpha . M_1 N_\alpha\right| \quad (11)$$

$$\vec{P_2} = \left|\vec{P}_\beta . M_2 N_\beta\right| \quad (12)$$

$$\vec{P_3} = \left|\vec{P}_\delta . M_3 N_\delta\right| \quad (13)$$

$$\vec{P}_{SA}(x+1) = \frac{\vec{P_1}+\vec{P_2}+\vec{P_3}}{3} \quad (14)$$

The lions encounter the enemy only when it becomes stationary. This task is modelled based on the value of $\vec{m}$ and it is estimated as given in Equation (15)

$$\vec{m} = 2 - \frac{2x}{X_{max}} \quad (15)$$

As a result, $|\vec{m}| < 1$ indicates that the lions must approach the victim to come into contact with it; if $|\vec{m}| > 1$, the lions will leave the victim and seek out a different target. Based on the ideal positions of α, β, and δ, searching agents search.

Figure 5: Lions (pride) searching and hunting

3.4 Hybrid Weighed Quantum Lion Optimization with Enhanced Fast Track Gram Matrix-Independent Component Analysis (WQLO-EF-GM-ICA)

The goal of the study is to improve the efficiency assessment of preliminary Diabetics forecasting via the integration of two cutting-edge optimization methods: Enhanced Fast Track Gram Matrix-Independent Component Analysis (EF-GM-ICA) and Weighed Quantum Lion

Optimization (WQLO). DM is a long-term metabolic condition that can be prevented or delayed in its onset, and its associated health problems can be lessened with early detection and care. The goal of the project is to increase the scalability, accuracy, and efficiency of predictive models for early-stage Diabetics prediction by utilizing the synergies between WQLO and EF-GM-ICA.

Robust optimization skills and quick exploration of extensive search spaces are provided by WQLO, a bio-inspired optimization process that draws inspiration from the scavenging habit of lions in the wild. The upgraded Gram Matrix-Independent Component Analysis method, or EF-GM-ICA, on the other hand, is designed to extract features from massive datasets. The investigation attempts to get around the drawbacks of conventional modeling and forecasting methodologies and produce greater accuracy for preliminary diabetic forecasting by combining these two strategies into a hybrid framework.

The combination of the WQLO-EF-GM-ICA methodology is being developed and validated to improve the accuracy of predictions, lowering computing difficulty, facilitate prompt intervention, and optimize the choice of characteristics. With a focus on the initial stages of diabetic forecasting, the project intends to develop predictive modeling in healthcare by utilizing sophisticated optimization approaches and extraction of feature procedures.

High-dimensional database extraction of features and reduction of dimensionality is accomplished with the sophisticated method known as Enhanced Fast Track Gram Matrix-Independent Component Analysis (EF-GM-ICA). It improves upon the conventional Gram Matrix-Independent Component Analysis (GM-ICA) technique, increasing both its efficacy and computing efficiency in identifying significant frameworks and trends in the information. In the traditional GM-ICA method, statistically autonomous elements are extracted using Independent Component Analysis (ICA) after calculating the Gram matrix, which is the inner product of the data matrix itself. Despite its efficacy, this method can be computationally demanding, particularly for extensive databases because of the high dimensionality of the Gram matrix.

To overcome this difficulty, EF-GM-ICA applies optimizations and improvements that expedite the computing procedure while maintaining the exactness of the output. These improvements could take the form of parallel computing plans, algorithmic tweaks, and optimization methods designed to specifically target the properties of the data.

A crucial feature of EF-GM-ICA is its capacity to accelerate the Gram matrix computation, which is frequently regarded as the main obstacle in conventional GM-ICA. The computational challenge related to the Gram matrix calculation is decreased by EF-GM-ICA by utilizing fast-track methods, parallel computing, or matrix factorization methodologies. This results in quicker analysis times and enhanced scalability for large databases.

Furthermore, to obtain distinct elements with greater efficacy, EF-GM-ICA might involve sophisticated methods for Independent Component Analysis (ICA). This minimizes processing costs and guarantees that the derived characteristics appropriately reflect the underlying framework of the data.

All things considered, upgraded rapid track gram matrix-independent component analysis provides a strong device for high-dimensional database extraction of features and dimensionality reduction. Because of its high computational effectiveness and efficiency, it is a good fit for several applications, such as forecasting, computation of signals, and recognition of patterns, where it is necessary to extract useful features from complicated data to perform correct analysis and interpretation.

Algorithm WQLO-EF-GM-ICA

Step 1. Initialization

Step 1.1: Initialize the population of lions with random positions and weights in the search space.

Step 1.2: Randomly initialize the unmixing matrix W for GM-ICA.

Step 2: Compute the Gram Matrix

Step 2.1: Compute the Gram matrix G from the data matrix X using the equation: $G = XX^T$.

Step 3: Enhanced Computation of Gram Matrix

Step 3.1: Utilize optimized algorithms or techniques to expedite the computation of the Gram matrix.

Step 4: Independent Component Analysis (ICA)

Step 4.1: Perform Independent Component Analysis using the EF-GM-ICA algorithm to extract independent components from the Gram matrix G.

Step 4.2: Decompose G into the unmixing matrix W and the matrix of independent source signals S.

Step 5: Weighed Quantum Lion Optimization

Step 5.1: Apply quantum gate operations to the positions and weights of the lions to simulate quantum behavior.

Step 5.2: Update the positions and weights of the lions based on their fitness values and quantum behavior.

Step 6: Fitness Evaluation:

Step 6.1: Evaluate the fitness of each lion based on the extracted features and their effectiveness in predicting diabetics at an early stage.

Step 6.2: The fitness function may include metrics such as accuracy, sensitivity, specificity, and AUC- ROC.

Step 7: Weighted Update

Step 7.1: Update the weights of the lions based on their fitness values and the outcome of the optimization process.

Step 7.2: Adjust the weights to reflect the importance of each lion in the optimization process.

Step 8: Sorting

Step 8.1: Sort the lions based on their fitness values and weights in descending order to prioritize lions with higher fitness and weights.

Step 9: Reproduction

Step 9.1: Reproduce new lions through crossover and mutation operations performed on the fittest lions.

Step 9.2: Crossover combines genetic information from two parent lions to generate offspring, while mutation introduces random variations.

Step 10: Survival Selection

Step 10.1: Select lions for the next generation based on a survival selection strategy, such as tournament selection or elitism.

Step 10.2: Elitism preserves a certain percentage of the fittest lions from the current generation for the next generation.

Step 11: Convergence Check

Step 11.1: Check for convergence criteria, such as reaching a maximum number of iterations or achieving a target fitness threshold.

Step 11.2: If the convergence criteria are met, terminate the algorithm.

Step 12: Iteration

Step 12.1 Repeat steps 5 to 11 for a predefined number of iterations or until convergence is achieved. Each iteration represents a cycle of exploration, evaluation, and adaptation.

Step 13: Output

Step 13.1: Return the best solution found by the algorithm, representing optimized parameters or features for predicting diabetics at an early stage.

4. Results and Discussions

Earlier forecasting of DM would assist in the use of ML along with additional analytics to diagnose Diabetics whether or not it exists in the patient. Numerous studies in various fields have been conducted on this topic. The primary concept underlying this proposed paradigm is the use of MATLAB software to identify Diabetics in its prestige. There are nine distinct characteristics and 768 rows of samples in this database. It is separated more clearly into simulations for testing and training. The number of tuples taken is 768. Table 2 describes the PIDD configuration. Where "Glucose" indicates the blood sugar level (mmol/L) and "Pregnancies" indicates the number of times a specific person becomes pregnant. Nine attributes serve as the training model for calculating the feature importance score. When classifying, the feature importance scores with the lowest values are disregarded.

Precision is a crucial parameter that might be used to determine the model's accuracy and efficacy. The cost curve known as loss is used to analyse the specifics of the contributions made by the various characteristics in the PIMA dataset shown in Tables 3-5. There are no characteristics omitted in #iteration 40.

Table 2: Statistical analysis of PIMA (UCI repository 2023)

Index	Attribute	Mean	Min/Max
F1	Plasma sugar level	121.02	0/199
F2	Two-hour serum insulin (mu U/ml)	78.9	0/846
F3	Diastolic blood pressure (mmHg)	69.5	0/120
F4	Triceps skinfold thickness (mm)	21.4	0/100
F5	No. of times pregnant	3.9	0/18
F6	Diabetics nutrition function	0.6	0.079/2.24
F7	Age	34.1	22/82

F8	Body mass index (kg/m2)	33	0/67.2

Table 3: Analysis of WQLO-EF-GM-ICA for #Iteration 20

Population	F1	F2	F3	F4	F5	F6	F7	F8	COST
10	0	1	0.432	0	0	1	1	1	0.227
20	0.816	0.168	0.177	0.982	0.086	0.636	0.066	0.367	0.236
30	1	0.979	0.270	0.054	0	0.735	0.309	0.619	0.220
40	0.740	0.487	0.510	0.096	0.065	0.928	0.938	0.464	0.224
50	0.952	0.448	0.134	0.027	0.004	0	0.394	0.45	0.224
60	0.921	0.519	0.270	0	0	0.835	0	0.610	0.220
70	0.795	1	0.598	0.168	0	1	0	0.631	0.223
80	0.125	0.351	0.313	0.382	0.227	0.871	0.803	0.977	0.218
90	0.876	0.663	0.216	0	0	0.615	0.139	0.316	0.216
100	0	1	0.412	0.005	0	0.842	0.499	1	0.222

Table 4: Analysis of WQLO-EF-GM-ICA for #Iteration 80

Population	F1	F2	F3	F4	F5	F6	F7	F8	COST
10	0.920	0.498	0.831	0.117	0.015	0.876	0.934	0.354	0.227
20	0.565	0.337	0.330	0.016	0	0.852	0.012	0.234	0.220
30	0.76	0.88	0.196	0	0	0.690	0.399	0.564	0.216
40	0.542	0.220	0.144	0	0	0.269	0.880	0.140	0.214
50	0.941	0.338	0.145	0.053	0.042	0.598	0.315	0.356	0.214
60	0.686	0.477	0.350	0.022	0.055	0.602	0.717	0.422	0.214
70	0.504	0.741	0.885	0.315	1	1	1	1	0.224
80	0.543	0.159	0.096	0	0	0.304	0.325	0.134	0.210
90	0.842	0.187	0	0	0	0.756	0.168	0.733	0.216
100	0.868	0.433	0.069	0	0	0.640	0.290	0.343	0.212

Table 5: Analysis of WQLO-EF-GM-ICA for #Iteration 100

Population	F1	F2	F3	F4	F5	F6	F7	F8	COST
10	0.188	0.373	0.413	0.328	0.60	1	0.670	0.837	0.218
20	1	1	0	0	0.608	0.766	0.764	0.812	0.222
30	1	0.476	0.377	0.081	0.053	1	0.949	0.946	0.218
40	0.439	0.788	0.160	0.244	0	1	0	0.872	0.219
50	0.814	0.335	0.144	0	0.029	0.553	0.256	0.440	0.215
60	0.656	0.231	0.089	0.010	0.040	0.706	0.318	0.498	0.211
70	0.950	0.757	1	0.004	0	1	1	0.389	0.215
80	0.997	0.621	0.343	0.065	0.072	0.991	0.747	0.592	0.214
90	0.352	0.314	0.274	0.168	0.305	0.99	0.134	0.996	0.218

100	0.887	0.249	0.279	0	0	0.503	0.957	0.247	0.2098

To assess the developed approach's performance more intuitively, the proposed algorithm's convergence curve is contrasted with those of other similar techniques. Figures 6 (a), (b), and (c) illustrate it. Compared to other similar algorithms from the previous stage, WQLO-EF-GM-ICA convergence time is significantly shorter, indicating that the developed algorithm is more capable of global exploration than other algorithms from the beginning stage. However, the algorithm has a drawback in that it can become trapped in the local optimal solution. Furthermore, there is still a lot of search potential left in the WQLO-EF-GM-ICA algorithm for the future.

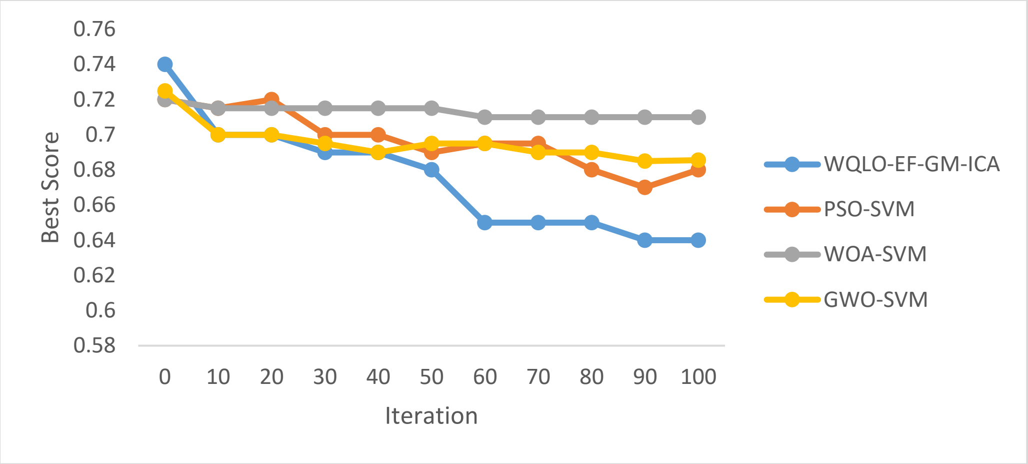

Figure 6: (a) Convergence curves of WQLO-EF-GM-ICA model vs other methods

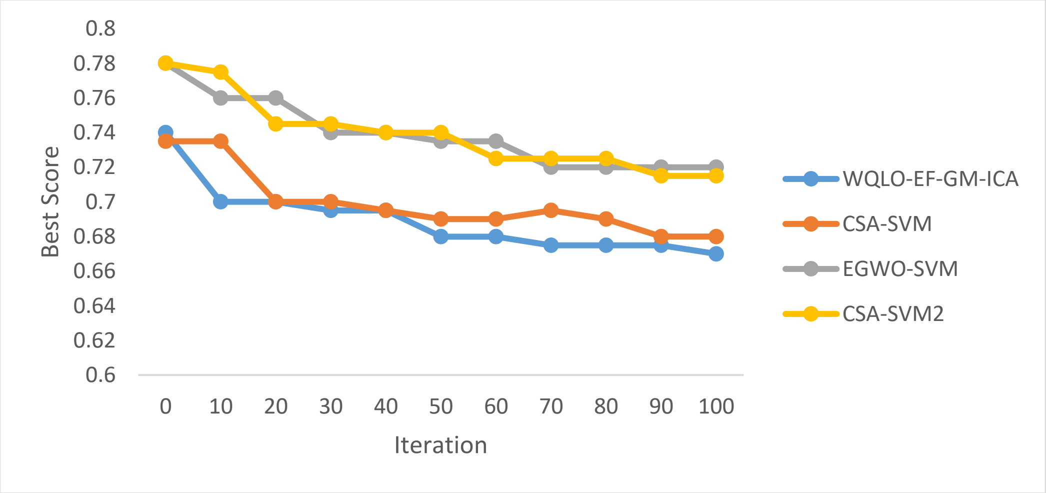

Figure 6 (b) Convergence curves of WQLO-EF-GM-ICA model vs other methods

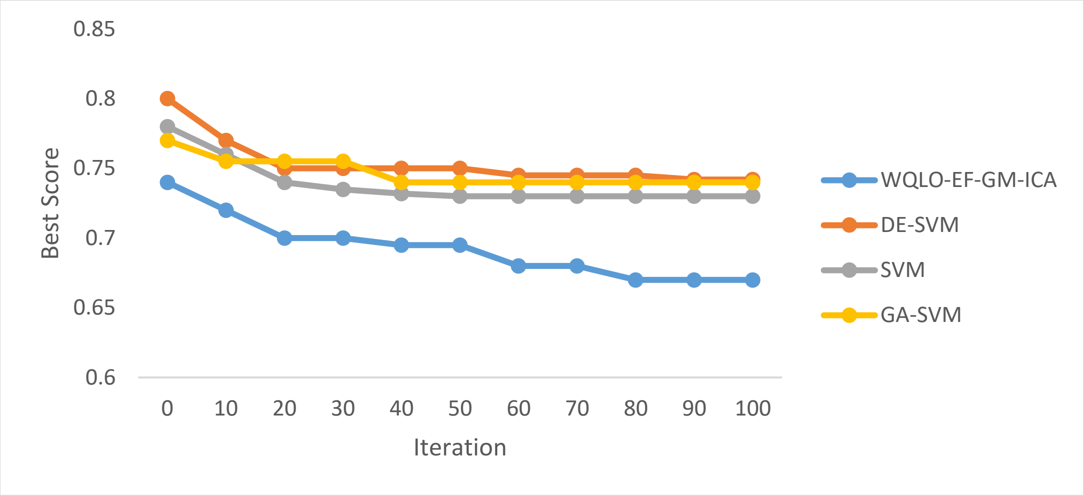

Figure 6 (c): Convergence curves of WQLO-EF-GM-ICA model vs other methods

As illustrated in Figure 7, the proposed WQLO-EF-GM-ICA approach has attained an excellent specificity rate in comparison to the existing methods.

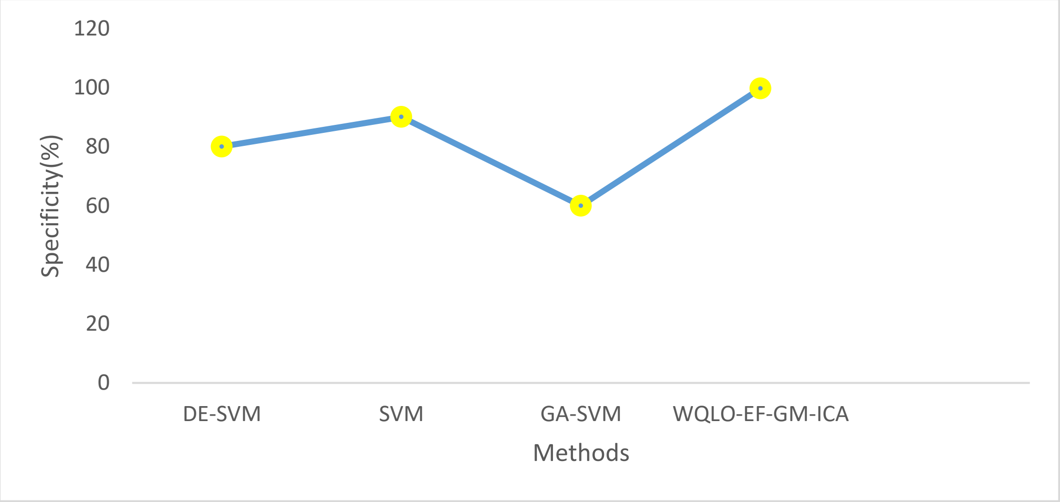

Figure 7: Comparison of proposed and existing methods (specificity (%))

The proposed approach of WQLO-EF-GM-ICA acquired a high sensitivity rate; the sensitivity and the specificity measure are represented graphically in Figure 8.

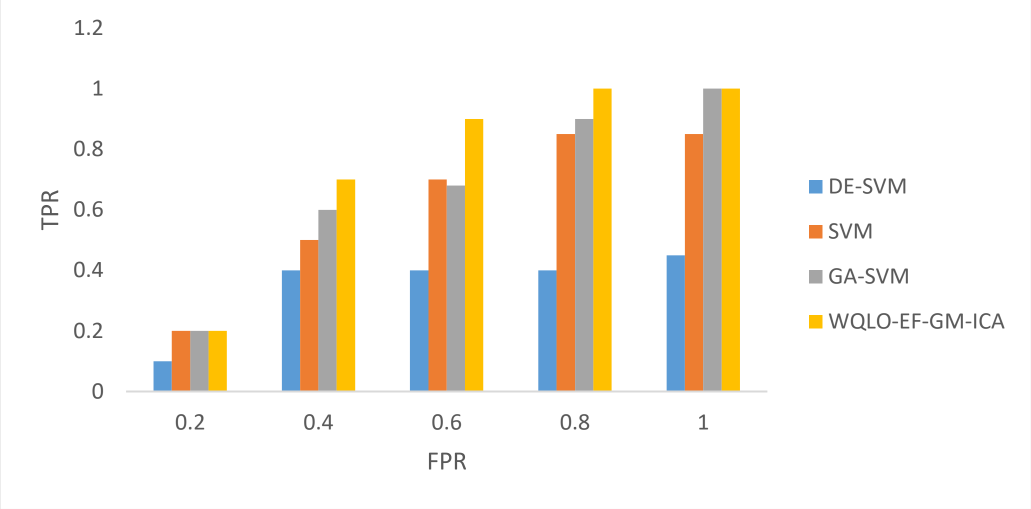

Figure 8. Comparison of proposed and existing methods sensitivity (FPR) and specificity (TPR)

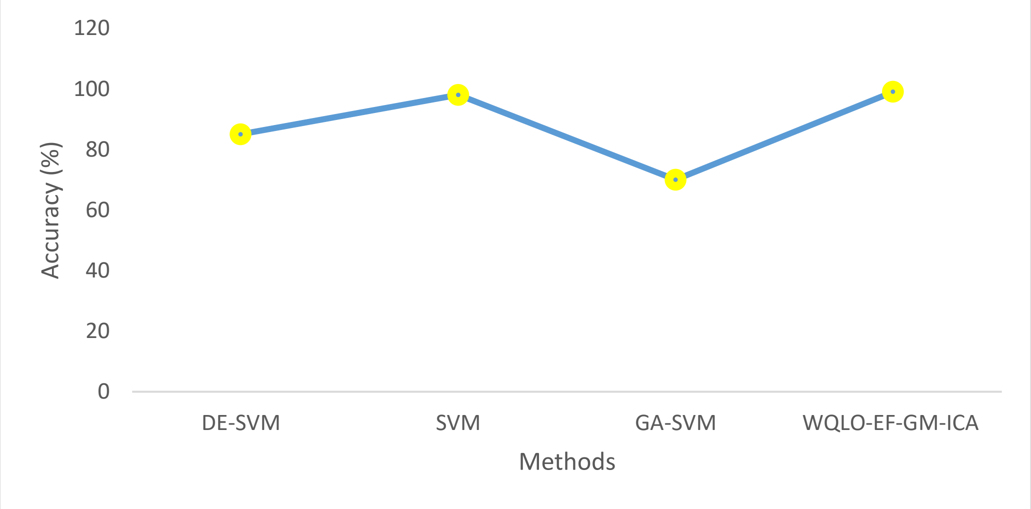

Figure 9: Comparison of proposed and existing methods (accuracy)

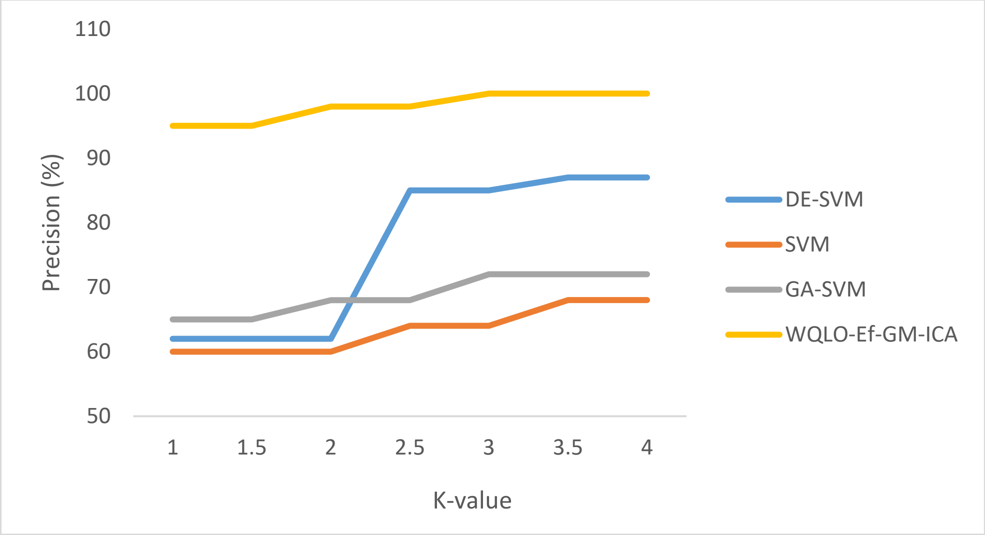

Figure 10: Comparison of proposed and existing methods (precision)

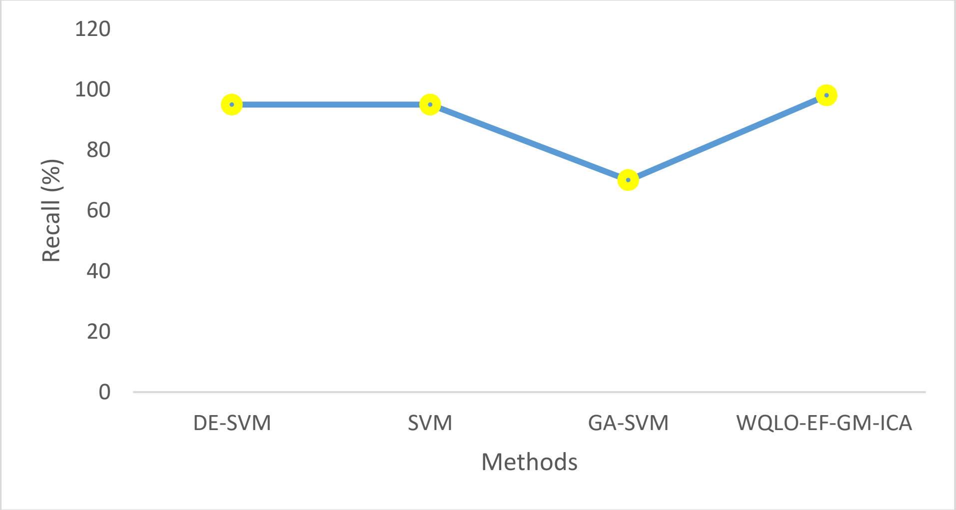

Figure 11: Comparison of proposed and existing methods (recall)

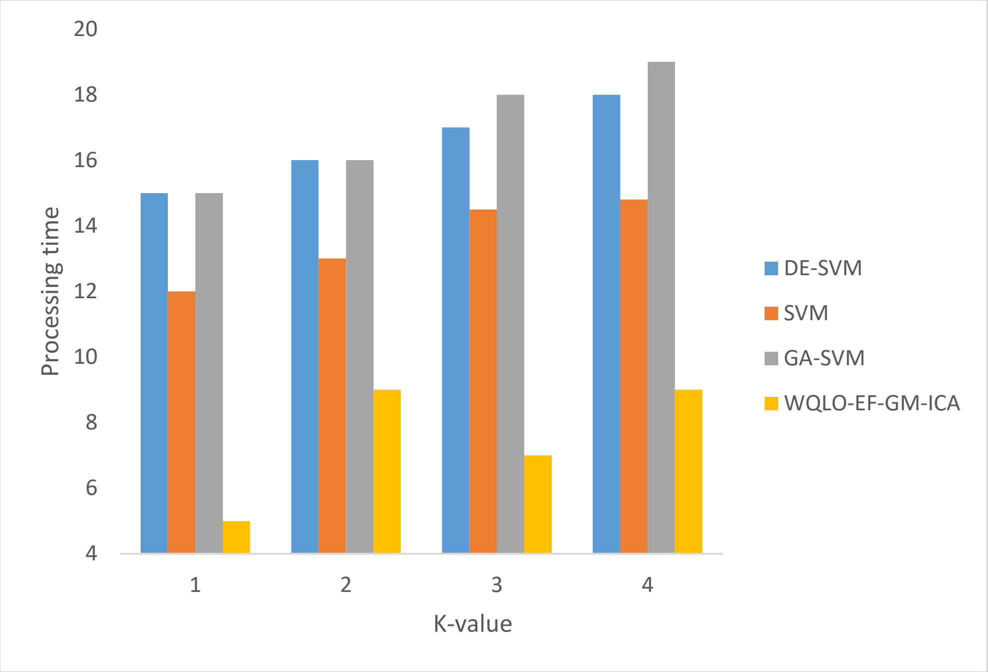

Figure 12. Processing time (s) with various classifiers

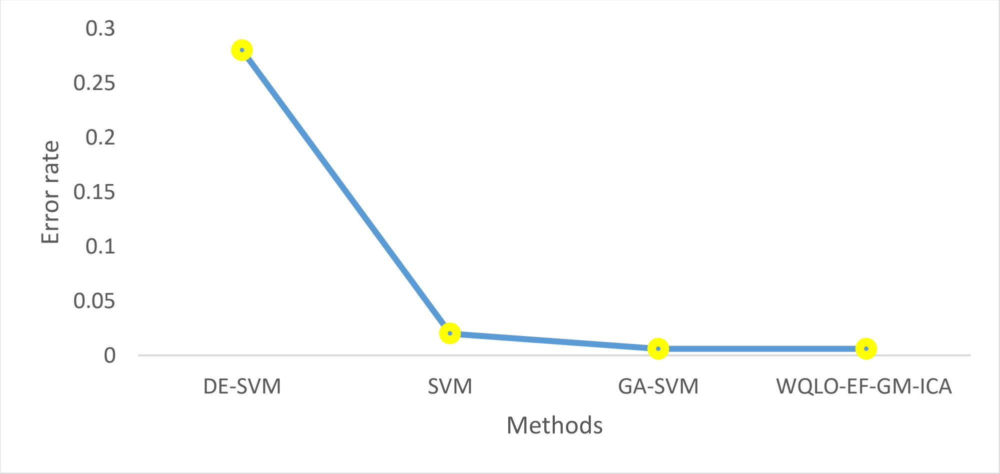

Figure 13: Comparison of proposed and existing methods (Error percentage)

One of the key measures used to assess the system's efficiency is precision computation. The greatest precision for predicting diseases achieved by the proposed WQLO-EF-GM-ICA technique is 98.98%, as demonstrated in Figure 9 when compared to other existing methods.

As a result, the accuracy of the proposed WQLO-EF-GM-ICA (100%) is contrasted with that of the current techniques, which are represented in Figure 10. Since sensitivity and recall are comparable, it also be used to estimate the recall value. The recall of the proposed WQLO-EF-GM-ICA (98) is contrasted with the current techniques, which are represented in Figure 11.

Compared to the existing methods, the proposed WQLO-EF-GM-ICA method efficiently distinguishes the illness characteristics in a 10 s shorter time, as demonstrated in Figure 12. As a result, the proposed approach to illness prediction effectively examines a vast quantity of data while minimizing logical complexity. Based on the k value, the execution duration is calculated (0 to 5). When contrasted to traditional methods, depicted in Figure 13, the failure rate of the proposed forecasting of diseases in the medical field is 0.0098 %. Additionally, a comparison is made between the general interpretation of the WQLO-EF-GM-ICA given and current techniques.

5 Conclusions and Future Enhancement

Finally, an intriguing technique for enhancing the accuracy of the assessment of premature Diabetic prediction is provided by the combination of WQLO-EF-GM-ICA. This combination of methods seeks to improve the scaling, precision, and effectiveness of diabetic prediction algorithms by fusing cutting-edge methods for extracting features with optimization strategies driven by the theory of quantum mechanics. Using an iterative improvement procedure, the algorithm efficiently detects pertinent patterns and characteristics in highly dimensional records, facilitating the discovery of significant information and enabling the forecasting of the initial stages of Diabetes. WQLO is used to provide strong and effective research and EF-GM-ICA is used to optimize computations and improve self-sufficient element removal in the information being analyzed. Furthermore, because the approach is a hybrid model, WQLO and EF-GM-ICA can work in concert, utilizing their unique advantages to get past the drawbacks of conventional optimizing and feature extraction methods. Continuously improving the algorithm's accuracy and flexibility to different information and forecasting assignments, the process refines the models for forecasting according to input from health assessments. Ultimately, a breakthrough in modeling predictions for the initial stages of Diabetic detection has been made with the development of a combination of WQLO with EF-GM-ICA. This strategy may help to enhance medical conclusions, prevent complications associated with Diabetics, and eventually enhance the standards of existence for those who are at a higher risk of acquiring Diabetics by streamlining feature selection, cutting down on computational difficulty, and facilitating immediate intervention. To fully capitalize on this new strategy's significance in medical practice and health-related initiatives, more investigation and validating efforts are necessary.

Future directions

Integration of Multi-Omics Data: Multi-omics knowledge, such as the field of transcriptomics, the field of proteomics, the field of metabolomics, and genomes, can be incorporated to gain a thorough knowledge of the molecular processes driving Diabetics. To be able to obtain a more comprehensive picture of this illness and increase the accuracy of predictions, future research may emphasize including such different sets of information within a predictive model.

Explainable AI Techniques: By using explicable artificial intelligence (AI), the model used for forecasting will become more interpretable, making it easier for medical practitioners to comprehend the variables that affect diabetic diagnosis. To clarify how the combined algorithm makes decisions, subsequent stages can include implementing strategies like significance evaluation as well as interpreting the model algorithms.

Real-Time Monitoring and Intervention: Patients who are in danger of acquiring Diabetes might benefit from proactive healthcare administration if the algorithm is extended to continuous surveillance and intervention systems. Further investigation may focus on developing portable gadgets and smartphone apps that use a combination of algorithms to deliver real-time, customized guidance and support.

Clinical Validation and Deployment: This is necessary to carry out extensive medical evaluations to assess the effectiveness as well as the efficacy of the combination of algorithms in practical medical facilities. Further approaches ought to center on working together with providers and healthcare organizations to use the algorithm in hospitals and evaluate its effects on outcomes for patients and medical expenses.

Adaptation to Population Diversity: Given the wide range of biological and environmental variables that affect risks for Diabetics in various communities, subsequent investigations ought to concentrate on modifying the combination of algorithms to account for features unique to certain groups. The process might entail gathering and examining information relating to various demographics to guarantee the algorithm's efficacy and applicability to various categories of people.

Continuous Algorithm Improvement: The combination of algorithms must be continuously improved and optimized in response to input via healthcare settings as well as research discoveries. Looking onward, the algorithm should be updated iteratively using more sources of information, optimizing strategies, and extraction of features approaches to improve its performance and flexibility over time.

Ethical and Regulatory Considerations: Verified permission, protection, along confidentiality of information are legal and moral problems that must be addressed for the combination of algorithms to be developed and implemented in an accountable manner within medical environments. To guarantee that the algorithm has moral and accountable application in healthcare settings, subsequent studies ought to give regulatory compliance and moral standards a top priority.

Through researching these potential avenues for further development, we can expand the combination of the WQLO-EF-GM-ICA algorithm's use to provide early Diabetic prediction, which will ultimately enhance the results of healthcare and give people more control over their health.

References

1. Ahmed, B. M., Ali, M. E., Masud, M. M., & Naznin, M. (2024). Recent trends and techniques of blood glucose level prediction for diabetes control. Smart Health, 100457.

2. Zohair, M., Chandra, R., Tiwari, S., & Agarwal, S. (2024). A model fusion approach for severity prediction of diabetes concerning binary and multiclass classification. International Journal of Information Technology, 16(3), 1955-1965.
3. Arora, A. S. (2024). Automated prediction of diabetes mellitus using infrared thermal foot images: recurrent neural network approach. Biomedical Physics & Engineering Express, 10(2), 025025.
4. Sarkar, P. J., & Pawar, S. (2024). Enhancing Early Detection and Prediction of Diabetes Mellitus in Patients of Indian Origin through Rigorous Machine Learning Techniques with Comprehensive Models Evaluation. International Journal of Intelligent Systems and Applications in Engineering, 12(11s), 634-639.
5. Yu, H., Li, X., Shu, J., Wu, X., Wang, Y., Zhang, C., ... & Li, Z. (2024). Evaluation of salivary glycopatterns based diagnostic models for prediction of diabetic vascular complications. International Journal of Biological Macromolecules, 129763.'
6. Zhang, Z., Zhou, Z., & Li, H. (2024). The role of lipid dysregulation in gestational diabetes mellitus: Early prediction and postpartum prognosis. Journal of Diabetes Investigation, 15(1), 15-25.
7. Parthiban, L., Latchoumi, T. P., Balamurugan, K., Raja, K., & Parthiban, R. (2023). Cognitive Computing for the Internet of Medical Things. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 85-100). Cham: Springer International Publishing.
8. Malarvizhi, K., Sangeetha, M., Deepak, V., Kavikumar, C., & Prakash, P. S. (2024). Prediction of Diabetes Mellitus using Ensemble Methods in Machine Learning. Grenze International Journal of Engineering & Technology (GIJET), 10(1).
9. Ooi, Y. G., Sarvanandan, T., Hee, N. K. Y., Lim, Q. H., Paramasivam, S. S., Ratnasingam, J., ... & Lim, L. L. (2024). Risk Prediction and Management of Chronic Kidney Disease in People Living with Type 2 Diabetes Mellitus. Diabetes & metabolism journal.
10. Yang, W., Zhu, L., He, J., Wu, W., Zhang, Y., Zhuang, B., ... & Lu, M. (2024). Long-term outcomes prediction in diabetic heart failure with preserved ejection fraction by cardiac MRI. European Radiology, 1-13.
11. Althobaiti, T., Althobaiti, S., & Selim, M. M. (2024). An optimized diabetes mellitus detection model for improved prediction of accuracy and clinical decision-making. Alexandria Engineering Journal, 94, 311-324.
12. Ghazal, P., & Mehmood, M. (2024). The use of the Finnish Diabetic Risk Score Scheme (FINDRISC) for the prediction of diabetes type-2 risk among various socio-demographic cohorts of the Pakistani population. Journal of Public Health, 1-8.
13. Mohamed Yousuff, A. R., Zainulabedin Hasan, M., Anand, R., & Rajasekhara Babu, M. (2024). Leveraging deep learning models for continuous glucose monitoring and prediction in diabetes management: towards enhanced blood sugar control. International Journal of System Assurance Engineering and Management, 1-8.
14. Laptev, D. N., Sechko, E. A., Romanenkova, E. M., Eremina, I. A., Bezlepkina, O. B., Peterkova, V. A., & Mokrysheva, N. G. (2024). Clinical prediction model for MODY type diabetes mellitus in children. Diabetes mellitus, 27(1), 33-40.
15. NG, B. A. (2024). En-RfRsK: An ensemble machine learning technique for prognostication of diabetes mellitus. Egyptian Informatics Journal, 25, 100441.
16. Tanaka, M., Akiyama, Y., Mori, K., Hosaka, I., Kato, K., Endo, K., ... & Furuhashi, M. (2024). Predictive modeling for the development of diabetes mellitus using key factors in various machine learning approaches. Diabetes Epidemiology and Management, 13, 100191.

17. El-Sofany, H., El-Seoud, S. A., Karam, O. H., El-Latif, A., Yasser, M., & Taj-Eddin, I. A. (2024). A Proposed Technique Using Machine Learning for the Prediction of Diabetes Disease through a Mobile App. International Journal of Intelligent Systems, 2024.
18. Ghusn, W., Hage, K., Vierkant, R. A., Collazo-Clavell, M. L., Dayyeh, B. K. A., Kellogg, T. A., ... & Ghanem, O. M. (2024). Type-2 diabetes mellitus remission prediction models after Roux-En-Y gastric bypass and sleeve gastrectomy based on disease severity scores. Diabetes Research and Clinical Practice, 208, 111091.
19. Li, J., Zhou, X., Wen, J., Liu, S., & Fan, X. (2024). Establishment and validation of a nomogram clinical prediction model for osteoporosis in senile patients with type 2 diabetes mellitus. Scientific Reports, 14(1), 5343.
20. Latchoumi, T. P., Parthiban, L., Balamurugan, K., Raja, K., Vijayaraj, J., & Parthiban, R. (2023). A Framework for Low Energy Application Devices Using Blockchain-Enabled IoT in WSNs. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 121-132). Cham: Springer International Publishing.
21. Xiong, Y., Liu, Y. M., Hu, J. Q., Zhu, B. Q., Wei, Y. K., Yang, Y., ... & Long, E. W. (2024). A personalized prediction model for urinary tract infections in type 2 diabetes mellitus using machine learning. Frontiers in Pharmacology, 14, 1259596.
22. Detsrisuwan, J., Luewan, S., Sirichotiyakul, S., & Tongsong, T. (2024). Maternal Visceral Fat in Prediction of Gestational Diabetes Mellitus. Journal of Clinical Medicine, 13(2), 493.
23. Sasagawa, Y., Inoue, Y., Futagami, K., Nakamura, T., Maeda, K., Aoki, T., ... & Hoshina, G. (2024). Application of deep neural survival networks to the development of risk prediction models for DM, hypertension, and dyslipidemia. Journal of Hypertension, 42(3), 506-514.
24. Hassan, N. R., Nasr El Dien, E. F., Kishk, R. M., Azab, H., Othman, A. E., Abdelaziz, B. M., & Aboelros, S. A. (2024). Role of Tumor Necrosis Factor Alpha Gene Polymorphism (TNFα-308 G/A) and Prediction of Diabetic Nephropathy. Egyptian Journal of Medical Microbiology, 33(2).
25. Garikapati, P. R., Balamurugan, K., Latchoumi, T. P., & Shankar, G. (2022). A quantitative study of small dataset machining by agglomerative hierarchical cluster and K-medoid. In Emergent Converging Technologies and Biomedical Systems: Select Proceedings of ETBS 2021 (pp. 717-727). Singapore: Springer Singapore.
26. Kopitar, L., Fister Jr, I., & Stiglic, G. (2024). Using Generative AI to Improve the Performance and Interpretability of Rule-Based Diagnosis of Type 2 Diabetes Mellitus. Information, 15(3), 162.
27. Hennebelle, A., Ismail, L., Materwala, H., Al Kaabi, J., Ranjan, P., & Janardhanan, R. (2024). Secure and privacy-preserving automated machine learning operations into end-to-end integrated IoT-edge-artificial intelligence-blockchain monitoring system for diabetes mellitus prediction. Computational and Structural Biotechnology Journal, 23, 212-233.
28. MOHAMED, M. H., KHAFAGY, M. H., MOHAMED, N., KAMEL, M., & SAID, W. (2024). DIABETIC MELLITUS PREDICTION WITH BRFSS DATA SETS. Journal of Theoretical and Applied Information Technology, 102(3).
29. Jain, A., & Singhal, A. (2024). Bio-inspired Approach for Early Diabetes Prediction and Diet Recommendation. SN Computer Science, 5(1), 182.