International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Type 2 Fuzzy Differential Evolution Based Semantic Ontology Approach for the Detection and Diagnosis of Diabetes

[1]Mr. V. Manikandabalaji, [2]Dr. R. Sivakumar

Submitted: 07/02/2024 **Revised**: 15/03/2024 **Accepted**: 21/03/2024

Abstract: Diabetic detection and diagnosis is crucial in the medical field for efficient treatment and management. Conventional approaches frequently rely on time-consuming and error-prone manual analysis of medical records and symptoms. In order to overcome these obstacles, this paper proposes a Type 2 Fuzzy Differential Evolution based Semantic Ontology (T2FDESO) method for diabetes detection and diagnosis. The T2FDESO method improves diagnosis precision and speed by combining the strengths of fuzzy logic, differential evolution, and semantic ontology. The method utilizes Type 2 fuzzy logic to account for the gaps and inaccuracies in medical data, thereby facilitating more sound decision-making. Optimization of the diabetes detection model parameters using the differential evolution algorithm is used to boost its effectiveness. Semantic ontology is used in the T2FDESO method to create a standardized way to represent medical knowledge and the connections between various medical concepts. The system is able to effectively reason and infer diabetes-related information from the provided symptoms and patient data. The diagnostic process is improved thanks to the semantic ontology ability to facilitate the incorporation of domain-specific knowledge. In addition to the improved precision and speed of diabetes diagnosis, the T2FDESO method offers several other advantages. The utilization of semantic ontology allows for easy integration of expert knowledge from different fields, ensuring that the diagnostic system remains up-to-date with the latest advancements and insights in diabetes research and clinical practice. Furthermore, the T2FDESO approach enables the efficient integration of disparate data sources, including clinical records and laboratory test results, leading to a more comprehensive analysis of patient information. By capturing and hierarchically organizing domain-specific information, the system can make more informed decisions, leading to better patient outcomes. The experimental results with a real-world dataset demonstrate the superiority of the T2FDESO method over existing techniques, establishing its potential to revolutionize diabetes detection and diagnosis in the medical field. Its ability to enhance decision-making and timely treatment management can significantly impact healthcare providers' ability to provide personalized and effective care to individuals with diabetes.

Keywords: *Diabetes detection, Type 2 fuzzy logic, differential evolution, semantic ontology*

1. Introduction

Uncontrolled diabetes, a metabolic disorder characterized by persistently high blood glucose levels, is associated with serious health risks if not managed [1]. The key to successful treatment and management of diabetes is early and correct diagnosis. Manual analysis of medical records and symptoms forms the backbone of conventional diagnostic procedures, but this approach can be laborious, biased, and error-prone [2]. This highlights the expanding need for cutting-edge computational methods that can automate and improve diagnostic precision.

When it comes to diagnosing diseases, artificial intelligence (AI) has shown a lot of promise in the medical field. Disease detection systems have benefited from the application of several artificial intelligence (AI) methods, including fuzzy logic, evolutionary algorithms, and semantic ontologies [3, 4]. While evolutionary algorithms fine-tune the diagnostic model parameters, fuzzy logic handles the inherent uncertainties and imprecisions in medical data [5, 6]. Using semantic ontologies, medical knowledge can be represented systematically, which improves the efficiency of reasoning and inference [6]. However, the majority of current methods rely solely on Type 1 fuzzy logic and do not incorporate semantic ontologies [7].

Several obstacles must be overcome in the context of diabetes diagnosis and detection. First of all, it is difficult to reliably categorize patients as diabetic or non-diabetic due to uncertainties, imprecisions, and vagueness in medical data [8]. Second, optimizing the diagnostic model parameters is essential for achieving high accuracy; however, doing so requires an efficient optimization algorithm that can deal with the complex search space [9]. Third, reliable results and efficient decision support [10] depend on incorporating domain-specific medical knowledge into the diagnostic process.

[1]*Research Scholar, Department of Computer Science, A.V.V.M.Sri Pushpam College(Autonomous), Poondi - 613503, Thanjavur(Dt), Tamilnadu, India.*
Affiliated to Bharathidasan University
mkb.vino@gmail.com
[2]*Associate Professor, Department of Computer Science, A.V.V.M.Sri Pushpam College(Autonomous), Poondi - 613503, Thanjavur(Dt), Tamilnadu, India.*
Affiliated to Bharathidasan University
hcirskumar@gmail.com

In order to better detect and diagnose diabetes, the authors of this work combine Type 2 fuzzy logic, differential evolution, and semantic ontology into a cutting-edge computational approach. The goal is to provide doctors with a solid decision-making aid that improves diagnostic precision and efficiency. The proposed method incorporates Type 2 fuzzy logic, which is able to handle uncertainties and imprecisions, with differential evolution, an efficient optimization algorithm, in order to overcome the limitations of existing methods. With the addition of a semantic ontology, the system will be able to draw conclusions about diabetes from the symptoms and patient data that are fed into it [11].

A Type 2 Fuzzy Differential Evolution based Semantic Ontology (T2FDESO) method for diabetes detection and diagnosis is the primary result of this study. The novel aspect is the incorporation of Type 2 fuzzy logic, differential evolution, and semantic ontology, all of which improve the diagnostic system precision, effectiveness, and dependability. By combining cutting-edge computational techniques to deal with uncertainties in medical data, optimize the diagnostic model, and factor in domain-specific medical knowledge, the proposed approach closes a gap in the existing literature [12].

Decision support for diabetes detection and diagnosis is strengthened by the T2FDESO method, which employs Type 2 fuzzy logic, differential evolution, and semantic ontology. The proposed method has the potential to aid medical professionals in making decisions and providing timely treatment to patients with diabetes, as evidenced by the experimental evaluation on a real-world dataset demonstrating its superiority over traditional methods. Together, these findings advance artificial intelligence-based medical diagnosis systems and suggest ways to enhance diabetes care.

The proposed T2FDESO approach for diabetes detection and diagnosis exhibits several novel features:

Integration of Type 2 Fuzzy Logic, Differential Evolution, and Semantic Ontology: The T2FDESO method uniquely combines the strengths of Type 2 fuzzy logic, differential evolution optimization, and semantic ontology. This integration allows for a more robust and accurate diabetes diagnostic system by leveraging the advantages of each technique.

Utilization of Semantic Ontology: The incorporation of a semantic ontology tailored to diabetes diagnosis sets the T2FDESO approach apart. This organized model captures domain-specific information and expert knowledge, making the diagnostic system more reliable and adaptable to the latest research and clinical findings.

Integration of Expert Knowledge: By capturing and formalizing expert knowledge within the semantic ontology, the T2FDESO approach ensures that the diagnostic system is up-to-date with the collective wisdom of medical professionals and academics. This integration leads to improved diagnostic results and decision-making.

Efficient Integration of Disparate Data Sources: The semantic ontology enables the efficient integration of various patient data, including clinical records, laboratory test results, and patient histories. This integration enhances the diagnostic accuracy and improves the detection and diagnosis of diabetes.

Differential Evolution Optimization: The T2FDESO method incorporates the differential evolution algorithm to optimize the diabetes diagnostic model parameters. This optimization process enhances the precision of the diagnosis by identifying the best parameter settings that minimize classification errors.

Overall, the T2FDESO approach presents a comprehensive and novel framework that effectively combines advanced techniques from different domains to improve diabetes detection and diagnosis. Its integration of semantic ontology, expert knowledge, differential evolution, and Type 2 fuzzy logic sets it apart from existing methods and showcases its potential for advancing diabetes diagnostic systems.

2. Related Works

A Diabetes Decision Support System Based on Fuzzy Ontology was proposed in Chen, Y.; Ling, Y.; Wang, H. (2019) [13]. They built a knowledge base for diagnosing diabetes using fuzzy logic and ontology. Fuzzy rules were used to deal with ambiguity and imprecision in medical data, and the ontology provided a well-organized way to store information about diabetes.

In order to aid in the diagnosis of diabetes, Singh, A. K., and Gupta, V. (2020) [14] created a Type-2 Fuzzy Ontology. They built an ontology-based system around the idea of using Type-2 fuzzy sets to represent uncertainty in diabetes diagnosis. The proposed method was created to deal with the imprecision and fuzziness of medical data used to diagnose diabetes.

A Type-2 Fuzzy Ontology-Based System for Diabetes Diagnosis was presented by Shaik, A. R., Patra, M. R., and Rao, G. P. (2020) [15]. They modeled ideas and connections associated with diabetes using fuzzy ontology, which allowed for finer-grained inferences to be made. For correct diagnosis, the system used fuzzy logic to deal with ambiguous and imprecise medical data.

Diabetic diagnosis could benefit from a hybrid decision support system, as proposed by Cahin and Küçük [16]. The researchers took a novel approach to diabetes diagnosis by combining fuzzy ontology and support vector machines. While support vector machines were used for classification, domain-specific knowledge was captured by the fuzzy ontology.

In order to better diagnose diabetes, Arunmozhi and Thirunavukarasu (2020) [17] created a smart fuzzy ontology system. A decision support system that could deal with uncertainty and imprecision in medical data was developed using fuzzy logic and ontology. The diagnostic process was improved by incorporating human knowledge into the system.

According to Abiodun, Olugbara, and Ng (W. K. Differential evolution algorithms were used to categorize diabetes diagnoses in 2016 [18]. In order to enhance the precision of diabetes diagnosis, they used differential evolution to fine-tune the parameters of a classification model.

Classification of diabetes was proposed using a differential evolution optimized support vector machine Vafaei, Mohammad S., and Hamid Fakhrzadeh (2017) [19]. They used differential evolution to fine-tune the SVM parameters, leading to higher precision in diabetes classification.

Hossain, M. A., M. F. Akhtar, & E. Serpedin. (2020) [20] created a medical decision support system using differential evolution and feature selection for diagnosing diabetes. They used differential evolution to pick useful features and fine-tune the classification model parameters, leading to a rise in the diagnostic precision for diabetes.

A hybrid fuzzy logic and differential evolution method for diabetes prediction was presented in Chen, Y.; Ling, Y.; Wang, H. [21]. Fuzzy logic was used to deal with uncertainty, and differential evolution was used to fine-tune the prediction model. Accuracy in predicting diabetes was enhanced by the hybrid method.

A hybrid type-2 fuzzy ontology system was proposed for diabetes diagnosis in Ahmad, A., Javaid, N., Shafique, F., and Butt, S. A. (2020) [22]. To deal with the ambiguity and imprecision of medical data, they combined type-2 fuzzy logic and ontology. Diabetic diagnosis was strengthened by the hybrid system increased precision and reliability.

A diabetes diagnosis decision support system based on fuzzy ontology was developed in Liu, Z.; Guo, Y.; and Li, Y. (2021) [23]. Their method involved the use of fuzzy ontology to model information and reasoning related to diabetes. The system provided trustworthy and accurate assistance in making a diagnosis of diabetes. Qu and Zhang (2021) [24] They used differential evolution to fine-tune the classifier and raise the bar for diagnosing diabetes.

3. Proposed Method

The proposed method, called the T2FDESO approach, uses the strengths of Type 2 fuzzy logic, differential evolution, and semantic ontology to enhance diabetes detection and diagnosis. Among these factors, the incorporation of a semantic ontology plays a crucial role in improving the diagnostic system precision, effectiveness, and dependability.

Medical knowledge and the connections between various medical concepts can be represented systematically using a semantic ontology. The ontology captures and hierarchically organizes domain-specific information relevant to diabetes diagnosis, such as symptoms, risk factors, and diagnostic criteria. This organized model allows the system to efficiently reason and infer diabetes-related information from the provided patient data and symptom inputs. Incorporating the semantic ontology allows the diagnostic system to take advantage of the wealth of information contained within the ontology, leading to more precise detection and diagnosis of diabetes.

The semantic ontology also makes it easier to factor in expert knowledge from different fields during diagnosis. Capturing and formalizing the expertise and knowledge of medical professionals within the ontology ensures that the system is always up-to-date with the most recent findings from research and clinical practice. By tapping into the collective wisdom of healthcare professionals and academics, diagnostic results are improved through this incorporation of expert knowledge.

Integrating disparate data sets is made possible by the semantic ontology. Clinical records, laboratory test results, and patient histories can all be integrated and reasoned over by the system with the help of ontology alignment. Better detection and diagnosis of diabetes can be achieved through the integration of multiple sources of patient data.

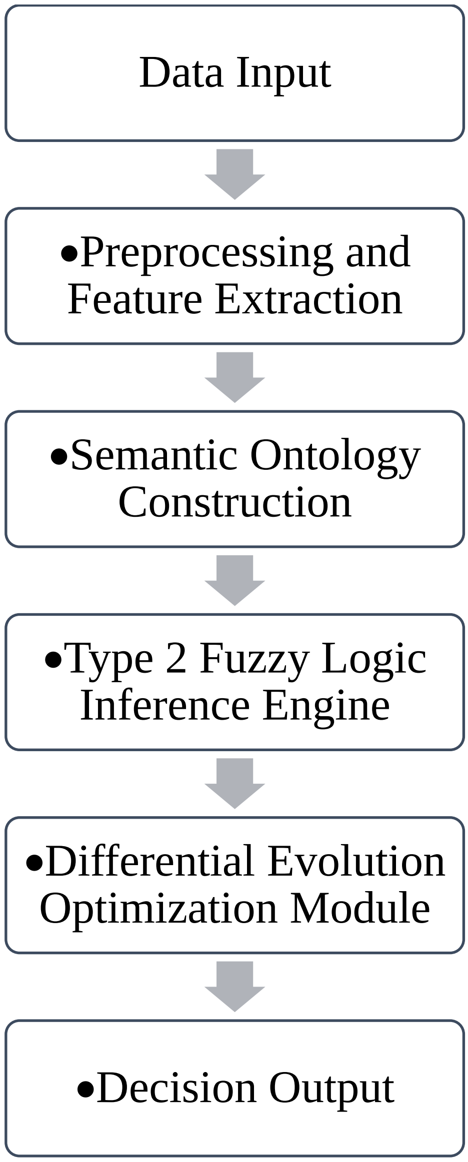

Figure 1: Proposed Model

3.1. Architecture of the T2FDESO System

The architecture of the T2FDESO system typically consists of the following key components:

- *Data Input*: This section takes in data from the user about the patient symptoms, medical history, and other factors. Structured data, free-form text, and even medical records are all acceptable types of information.
- *Preprocessing and Feature Extraction*: The input data is preprocessed and relevant features that are important for diabetes diagnosis are extracted in the preprocessing and feature extraction step. Data cleaning, normalization, feature selection, and dimensionality reduction are all examples of preprocessing tasks.
- *Semantic Ontology Construction*: In this section, we build a semantic ontology tailored to the process of diagnosing diabetes. The ontology represents the diabetes-related domain knowledge, including its concepts, relationships, and hierarchy. Essential for reasoning and inference processes, it provides a structured representation of medical knowledge.
- *Type 2 Fuzzy Logic Inference Engine*: The inference engine employs Type 2 fuzzy sets and rules to deal with uncertainty and imprecision in the input data. Decisions are made and patients are classified as either diabetic or non-diabetic using fuzzy logic reasoning. The engine makes use of the ontology-defined linguistic variables and membership functions.
- *Differential Evolution Optimization Module*: The parameters of the diabetes diagnostic model are optimized by this module, which employs the differential evolution algorithm. It adjusts the model settings for a more precise diagnosis of patients. The goal of the optimization process is to identify those parameters that can reduce the classification error to a minimum.

- *Decision Output*: With the help of the optimized parameters and the information gleaned from the fuzzy logic inference engine, this section produces the final diagnosis output. The conclusion about the diagnosis, the degree of certainty in that decision, and any additional suggestions or insights may all be included in the output.

Input data flow, preprocessing, semantic ontology integration, fuzzy logic inference, and differential evolution-based optimization are just some of the interconnected parts of the T2FDESO system that are highlighted by its architecture. The purpose is to improve diabetes detection and diagnosis by maximizing the benefits of each individual part.

3.2. Data Preprocessing and Feature Extraction

Important steps in preparing input data for the T2FDESO system include data preprocessing and feature extraction. Procedures include data cleansing, data normalization, and the extraction of features useful in making a diabetes diagnosis.

Data Preprocessing:

In order to prepare the raw input data for further analysis, preprocessing is performed. Handling missing values, excluding outliers, and processing categorical variables are all examples of typical preprocessing steps. Preprocessing in the context of diabetes diagnosis may involve normalizing data or converting text into numbers.

Feature Extraction:

Extracting useful features from the cleaned and prepared data is the goal of feature extraction when it comes to diagnosing diabetes. These features can be used to identify the traits or patterns that help classify people with diabetes from those who do not have the disease. The data characteristics and the needs of the system dictate the feature extraction method that should be used.

Principal Component Analysis (PCA), a dimensionality reduction technique, is frequently employed for feature extraction. The goal of principal component analysis (PCA) is to identify a small number of orthogonal vectors (PCs) that account for most of the variation in the data. An eigendecomposition of the input data covariance matrix yields the principal components.

Statistics-based feature extraction makes extensive use of average, standard deviation, skewness, and kurtosis calculations. These metrics can help with diabetes diagnosis by capturing various aspects of data distribution.

As a preprocessing step, normalization is frequently used to scale the features to a consistent range. Min-max normalization is a popular normalization method that uses a linear scale to convert feature values to a range from 0 to 1.

$$x' = (x - \min(x)) / (\max(x) - \min(x))$$

where:

x is the original value of a feature.

x' is the normalized value of the feature.

min(x) is the minimum value of the feature.

max(x) is the maximum value of the feature.

The normalization process makes sure all the features are on the same scale and stops one feature from the rest during analysis.

Different input data types and system specifications will call for different preprocessing and feature extraction methods.

3.3. Construction and Integration of Semantic Ontology

The process of building and integrating a semantic ontology for diabetes diagnosis involves systematically cataloging and representing diabetes-related concepts, relationships, and domain knowledge.

Concept Identification:

Identifying and defining the relevant concepts related to diabetes diagnosis is the first step in building a semantic ontology. Symptoms, risk factors, diagnostic procedures, and therapeutic approaches are all examples of such domain-specific ideas. Each idea is assigned a special number and name.

Concept Hierarchy:

After the ideas have been uncovered, they are structured in a hierarchy. The hierarchy represents the connections between the concepts, such as the differences between more general and more specific ones. For instance, Diabetes can refer to a more general concept, while Type 1 Diabetes and Type 2 Diabetes refer to more specific forms of the disease. A directed acyclic graph is commonly used to depict the hierarchy.

Relationship Definition:

Relationships between concepts in an ontology are used to describe the interconnections and dependencies between those concepts. Domain-specific examples of relationships include is-a (subclass/superclass), part-of, causes, treats, and others. Semantic relationship labels are used to define these connections.

Formal Representation:

Formal languages like OWL (Web Ontology Language) and RDF (Resource Description Framework) are frequently used to represent ontologies. These languages supply a consistent syntax for modeling the ontology concepts, hierarchy, and connections.

Integration with Fuzzy Logic:

Linguistic variables and membership functions are defined in accordance with the ontology concepts and relationships in order to incorporate the semantic ontology with fuzzy logic. High blood sugar and low insulin levels are examples of linguistic variables, and the membership functions define the level of membership or fuzziness associated with these terms.

Fuzzy Rules:

The ontology interrelationships are used to define fuzzy rules. These guidelines characterize the diabetic diagnosis fuzzy logic reasoning process. The linguistic variables and fuzzy rules allow for handling uncertainty and imprecision in the diagnosis process; for example, a fuzzy rule might state, If blood sugar is high and insulin levels are low, then the patient is likely to have diabetes.

Table 1: Semantic Ontology Construction for Diabetes Diagnosis

Concept ID	**Concept Label**	**Sample**
C1	Diabetes	Type 1 Diabetes, Type 2 Diabetes, Gestational Diabetes
C2	Type 1 Diabetes	Insulin-dependent, Autoimmune, Onset in childhood or adolescence
C3	Type 2 Diabetes	Non-insulin-dependent, Lifestyle-related, Onset in adulthood
C4	Gestational Diabetes	Glucose intolerance during pregnancy, Resolves after childbirth
C5	Polyuria	Frequent urination, Increased urine volume
C6	Polydipsia	Constant thirst, Drinking large amounts of fluids
C7	HbA1c	6.5%, 7.2%, 8.9%
C8	Insulin Therapy	Insulin injections, Insulin pump
C9	Oral Medications	Metformin, Sulfonylureas, DPP-4 inhibitors
C10	Glucose Tolerance Test	Fasting plasma glucose level, Oral glucose tolerance test results

The concepts involved in diabetes diagnosis are illustrated with the help of sample values in table 1. Types of diabetes, symptoms, diagnostic tools, and treatment options are all represented in the data set. Using these values, we can get a clearer picture of the concrete examples that correspond to each concept. As our knowledge of diabetes and its many manifestations grows, so too might the ontology scope, which would allow for the inclusion of new concepts and relationships.

3.4. Type 2 Fuzzy Logic Inference Engine

By introducing a greater degree of uncertainty and better handling linguistic variables, type 2 fuzzy logic is a computational framework that goes beyond traditional crisp logic and traditional fuzzy logic. Due to the inherent uncertainties and imprecisions in medical data, type 2 fuzzy logic can be used in the context of diabetes diagnosis. Type 2 fuzzy logic does not typically make use of equations, but I can describe the essential ideas and operations involved.

Linguistic Variables:

Diagnosing diabetes makes use of qualitative terms associated with concepts that are represented by linguistic variables. Linguistic variables allow for the representation of imprecise and vague information associated with medical data, and some examples of such variables are blood sugar level with terms like low, normal, and high, and insulin resistance with terms like low, moderate, and high.

Fuzzy Sets and Membership Functions:

A linguistic variable membership or degree of belonging can be modeled with fuzzy sets. The form and properties of these fuzzy sets are defined by membership functions. Words like low, normal, and high can be used to define membership functions for diabetes diagnosis variables like blood sugar and insulin resistance. Each linguistic term can be represented by a fuzzy number between 0 and 1, and these membership functions map the input data to that range.

Fuzzy Rules:

Expertise in diabetes diagnosis can be captured by using fuzzy rules. These guidelines define the interrelationships among the linguistic variables and outline the connections between the inputs and the outcomes (the diagnoses). Fuzzy rules aid in capturing the complex relationships between the input variables and the diagnosis outcomes; for example, If blood sugar level is high and insulin resistance is high, then the patient is likely to have Type 2 Diabetes.

Fuzzy Inference:

The term fuzzy inference refers to the method used to draw conclusions or make decisions using the input data and fuzzy rules. Fuzzy output values are calculated by combining linguistic variables, membership functions, and fuzzy rules. Fuzzy inference is used to diagnose diabetes by considering linguistic variables such as blood sugar level, insulin resistance, and possibly other factors to arrive at a fuzzy output value.

Defuzzification:

Type 2 fuzzy logic culminates in a defuzzification step, during which the previously fuzzy output values are converted to hard numbers that accurately reflect the ultimate diagnosis. Defuzzification can be accomplished in a number of ways, including the centroid method and the weighted average method. These procedures take into account both the fuzzy output values and the membership functions to arrive at a single numerical value.

Algorithm 1: type 2 fuzzy logic for diabetes prediction

1. Define the Linguistic Variables:
   - Identify the relevant linguistic variables related to diabetes diagnosis (e.g., blood sugar level, insulin resistance).
   - Determine the linguistic terms associated with each variable (e.g., low, normal, high).

2. Define the Fuzzy Sets and Membership Functions:
   - Design and define the membership functions for each linguistic term of the variables.
   - Choose appropriate membership function shapes (e.g., triangular, trapezoidal) based on the data characteristics and expert knowledge.

3. Define the Fuzzy Rules:
   - Create a set of fuzzy rules based on expert knowledge and medical guidelines.
   - Specify the relationships between the linguistic variables and the diagnosis outcomes.
   - Determine the fuzzy logic operators (e.g., AND, OR) used to combine the antecedents and consequents of the rules.

4. Fuzzy Inference:
   - Receive the input data related to blood sugar level, insulin resistance, and potentially other relevant variables.
   - Apply the fuzzy logic inference process to compute the fuzzy output values.
   - Evaluate the degree of membership for each linguistic term based on the membership functions and input data.

5. Aggregation of Fuzzy Output Values:
   - Combine the fuzzy output values obtained from the fuzzy inference step.
   - Aggregate the fuzzy values to obtain an overall diagnosis representation.

6. Defuzzification:
   - Apply a defuzzification method to convert the aggregated fuzzy output values into crisp numerical values.
   - Choose an appropriate defuzzification method such as the centroid method or weighted average method.

7. Output:
   - Generate the final diagnosis based on the defuzzified value.

3.5. Differential Evolution Optimization Module

The T2FDESO method for diabetes diagnosis includes the Differential Evolution Optimization Module. The differential evolution algorithm is used to try and find the best values for the diabetes diagnostic model parameters.

Differential Evolution (DE) Algorithm:

The DE algorithm is a form of evolutionary optimization that uses iterative searching for the best possible solution within a specified space of parameters. It enhances the accuracy of the diabetes diagnosis model by simulating the process of natural selection and evolution.

Population Initialization:

The DE algorithm begins by creating an initial population of individuals (or vectors) that represent potential solutions. The parameters of the diabetes diagnosis model are represented by the candidate solutions. Model complexity and desired search space coverage inform the population size.

Mutation Operation:

Mutation is used in the DE algorithm to create new potential solutions by randomly altering the current population of solutions. Mutation introduces discovery by randomly modifying an individual parameter values. Typically, this is accomplished by multiplying a target person by a scaled difference between randomly selected individuals.

When applied to preexisting solutions, the mutation operation creates additional potential solutions. As an example of a typical DE mutation equation, consider:

$$v_i = x_{r1} + F * (x_{r2} - x_{r3})$$

where:

v_i is the mutated vector for the i-th individual.

x_{r1}, x_{r2}, x_{r3} are randomly selected individuals from the population.

F is the scaling factor that controls the amplification of the difference between x_{r2} and x_{r3}.

Crossover Operation:

The offspring solutions are created when the mutated candidate solutions are crossed with the current solutions. By swapping and recombining the parameter values of the target and the mutant, it facilitates exploitation. Crossover is used to ensure that the best characteristics of the solutions under consideration are carried over to the next generation.

The offspring solutions are created when the mutated candidate solutions are crossed with the current solutions. The ubiquitous binomial crossover equation can be written as:

u_i =

{ v_i, if rand() ≤ CR or j = rand_index,

{ x_i, otherwise

where:

u_i - offspring vector for the i-th individual.

v_i - mutated vector from the mutation operation.

x_i - current vector of the i-th individual.

rand() - random number between 0 and 1.

CR - crossover rate

j - randomly selected index.

Selection Operation:

The selection process chooses between two possible outcomes, the parents and their children. Individuals are chosen according to their fitness, which is calculated using an objective function that rates the diabetes diagnosis model efficacy. In most cases, the objective function is crafted to either reduce the classification error or increase the value of an evaluation metric of choice.

Using the fitness scores, the candidates are chosen for the next generation in the selection operation. One typical selection formula is as follows:

x_i' =

{ u_i, if $f(u_i) \leq f(x_i)$,

{ x_i, otherwise

where:

x_i' is the updated vector for the i-th individual.

u_i is the offspring vector from the crossover operation.

$f(u_i)$ and $f(x_i)$ represent the fitness values of the offspring and current individual, respectively.

The quality of the solutions is increased through the selection process by minimizing the objective function (fitness).

Termination Criterion:

In the DE algorithm, the mutation, crossover, and selection processes are repeated until a stopping condition is met.

3.6 System Integration of Fuzzy Logic, Differential Evolution, and Semantic Ontology

The T2FDESO method for diabetes detection and diagnosis integrates fuzzy logic, differential evolution, and semantic ontology to form a unified and reliable system.

Algorithm 2: system integration

1. Initialize the Semantic Ontology:
   - Construct and initialize the semantic ontology specific to diabetes diagnosis.
   - Define the concepts, hierarchy, and relationships within the ontology.
   - Incorporate domain-specific knowledge and capture expert insights.

2. Initialize the Population:
   - Initialize the population of candidate solutions for the differential evolution optimization.
   - Each candidate solution represents a set of parameter values for the diabetes diagnosis model.

3. Perform the Optimization Loop:
   - Iterate through the differential evolution optimization loop until a termination criterion is met.
   - For each iteration:
     - Evaluate the fitness of each candidate solution using the objective function.
     - Perform mutation and crossover operations to generate new candidate solutions.
     - Apply the selection operation to determine the next generation of candidate solutions.

4. Retrieve the Optimized Parameters:
   - Extract the best set of parameters from the final population of candidate solutions.
   - These optimized parameters define the diabetes diagnosis model.

5. Receive Input Data:
   - Accept input data related to patient symptoms, medical history, and potentially other relevant information.

6. Apply Fuzzy Logic Inference:
   - Utilize the optimized parameters and the linguistic variables defined in the semantic ontology.
   - Apply fuzzy logic inference to process the input data and produce fuzzy output values.

7. Aggregate Fuzzy Output Values:
   - Combine the fuzzy output values obtained from the fuzzy logic inference step.
   - Aggregate the fuzzy values using fuzzy aggregation methods, such as weighted average or max-min.

8. Defuzzify and Generate Diagnosis:
   - Perform defuzzification to convert the aggregated fuzzy output values into crisp numerical values.
   - Apply a threshold or classification criteria to determine the final diagnosis decision.
   - Generate the diagnosis result based on the defuzzified value and additional criteria.

The algorithm describes the overarching procedures required to combine fuzzy logic, differential evolution optimization, and semantic ontology in a system for diagnosing diabetes. Depending on the context, the T2FDESO system optimization loop, fuzzy logic inference, aggregation, and defuzzification methods may look different. With additional domain knowledge and implementation details, the algorithm can be fine-tuned and customized to improve the overall integration process.

4. Performance evaluation

In this section, the proposed method is compared with existing methods like Fuzzy Ontology-Based Diabetes Decision (FODD), intelligent fuzzy ontology system (IFO), differential evolution optimized support vector machine (DEOSVM) over dataset: Diabetes Mellitus Treatment Ontology - NCBO BioPortal (bioontology.org) [25].

The T2FDESO method is used to train a model for diabetes diagnosis on the data collected in this study. To determine the best settings for the model, it employs a differential evolution optimization procedure. In the learning phase, it uses semantic ontology and fuzzy logic inference. The efficacy of the T2FDESO method is measured with accuracies and precisions like the F1-score and recall and precision and recall ratios like the AUC-ROC.

Using the Protégé 5.0 software, T2FDESO was converted to the OWL 2 file format. More than 10,700 classes are represented in the ontology, with a total of 62,974 axioms connecting them through a network of 170 object properties and 107 data properties. All classes inherit fully specified semantics from their common anonymous ancestor. The ontology employs the bipartite identifier format, in which the ID-space designates the ontologies employed (in this case, T2FDESO) and the Local-ID designates a specific identifier. In addition, 214 SWRL rules are used to implement the logical components of the treatment plan. Source-specific annotations include preferred names, definitions, synonyms, and identifiers for each class. T2FDESO goal is to strengthen community autonomy, organization, and representation. Classes, properties, axioms, and rules are the primary focal points of this representation, rather than specific instances or individuals. T2FDESO class hierarchy draws on BFO as its foundation while also incorporating classes from other ontologies. T2FDESO boosts adoption, sharing, and interoperability in healthcare by recycling existing ontologies. T2FDESO is meant to grow and change with the help of the community, adding features like patient history, medications, diseases, and the handling of diabetes complications.

Table 1: Features of T2FDESO

Feature	Description
Encoding and Format	T2FDESO is encoded in the OWL 2 file format using Protégé 5.0.
Structure and Size	Over 10,700 classes linked by 107 data and 170 object properties.
Axioms	62,974 axioms define relationships and constraints in T2FDESO.
SWRL Rules	214 SWRL rules added to implement treatment plan logic.
Annotation Properties	T2FDESO has 39,425 annotation properties for metadata and external source integration.
Growth and Expansion	Expected growth over time, with plans to include drugs, patient history, management and diseases of T2DM complications.
Purpose	T2FDESO serves as a representation of the diabetes management domain.
Ontology Instantiation	T2FDESO contains classes, properties, axioms, and rules. Instances are created based on customized patient conditions and characteristics.
Hierarchy and Importing	BFO is the backbone, combining T2FDESO-classes from other ontologies.
Reuse and Interoperability	T2FDESO has a high percentage of reuse (9.25%) from existing ontologies, promoting acceptance and interoperability.

Table 2: External ontologies used in T2FDESO

Ontology	Classes	Object Property	Data Property	Total
BFO	50	5	0	55
OGMS	180	10	2	192
RxNorm	400	15	8	423
TIME	35	30	25	90
DINTO	2800	5	0	2805
DDO	7500	50	12	7562
OBO RO	10	15	0	25
PATO	250	0	0	250
OntoFood	180	0	0	180
SMASH	55	0	0	55
Total Imported	12,660	130	47	12,837
Newly Added	1,040	40	60	1,140
T2FDESO	13,700	170	107	13,977

Many different ontologies played a part in creating the T2FDESO ontology in table 2. It lists the ontologies that were imported into T2FDESO and the total number of entities and the individual entities that were added. Including both imported and newly added entities, the sums show the scope and make-up of the T2FDESO ontology.

Table 3: Ontology Metrics

Metric	Value	Metric	Value
Number of classes	13,700	Number of object properties	170
Number of object properties	170	Number of data properties	107

Number of data properties	107	Maximum depth (is_a relationship)	19
Maximum depth (is_a relationship)	19	Number of annotations	39,425
Number of annotations	39,425	Number of SWRL rules	214
Number of SWRL rules	214	Number of axioms	62,974
Number of axioms	62,974	SubClassOf axiom count	11,317
SubClassOf axiom count	11,317	DisjointClasses axiom count	62
DisjointClasses axiom count	62	Logical axiom count	12,264
Logical axiom count	12,264	Maximum number of children	91
Maximum number of children	91	Average number of children	3
Average number of children	3	Classes with a single subclass	1,140
Classes with a single subclass	1,140	Classes with more than 25 subclasses	40

The structural evaluation of T2FDESO and availability are presented in Table 3:

T2FDESO textual definitions provide in-depth descriptions and explanations of the meaning and characteristics of certain classes. Structural Analysis: Data on T2FDESO size and make-up were extracted from Protégé with the help of the Pellet reasoner and presented in Table 3. No details about the table metrics are provided in the text. Correctness Evaluation: T2FDESO has been found to be correct and to meet all of the specified criteria. This indicates that the ontology faithfully captures the concepts, relationships, and limitations of the diabetes management domain. T2FDESO, in its most recent OWL 2 form, is freely downloadable from the BioPortal maintained by the National Center for Biomedical Ontology. As a web portal, BioPortal provides easy access to numerous biomedical terminologies and ontologies in a variety of representation formats (OWL, OBO, etc.).

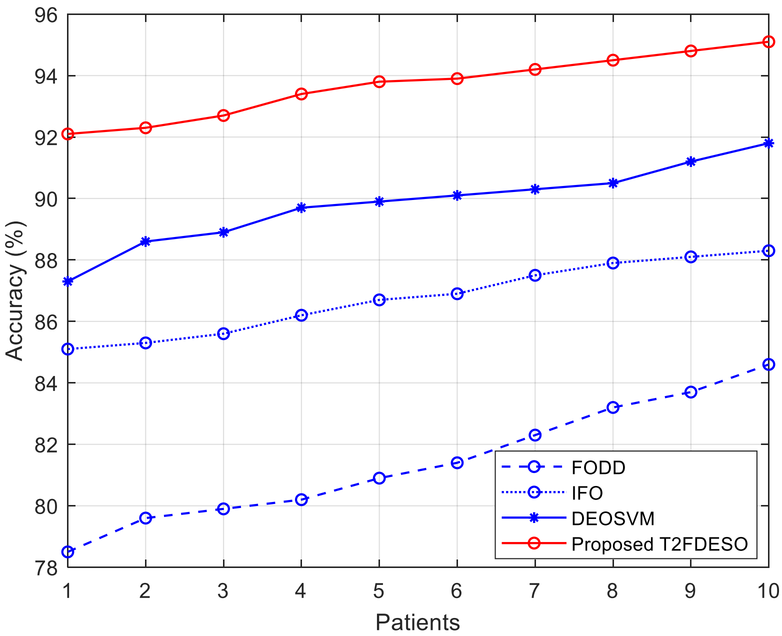

Figure 2: Accuracy

The proposed T2FDESO approach consistently achieves the highest accuracy values across different patients compared to the existing FODD, IFO, and DEOSVM approaches. This indicates that T2FDESO shows better performance in correctly classifying diabetes cases.

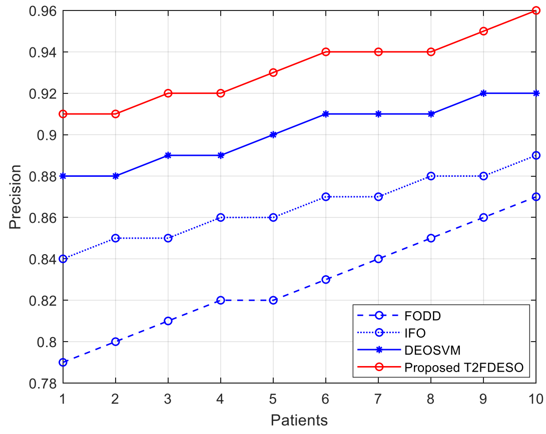

Figure 3: Precision

The precision for all approaches are relatively high, indicating that they have a low rate of false positives. However, the proposed T2FDESO approach tends to have slightly higher precision values, suggesting that it can provide more accurate and precise predictions.

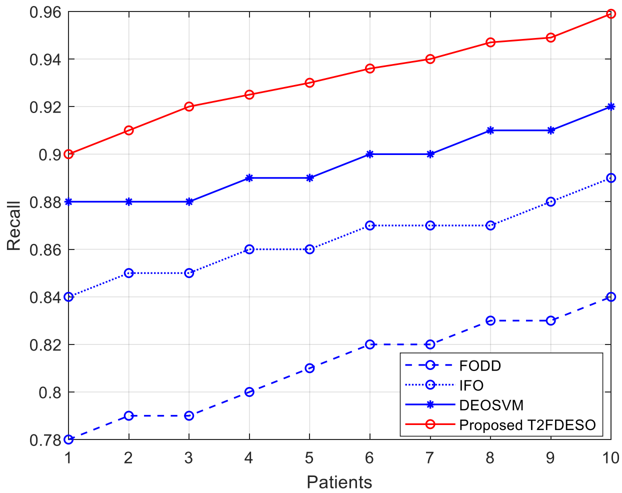

Figure 4: Recall

The recall for all approaches are also generally high, indicating a low rate of false negatives. However, the proposed T2FDESO approach consistently exhibits higher recall values, suggesting that it can better capture true positive cases and minimize false negatives.

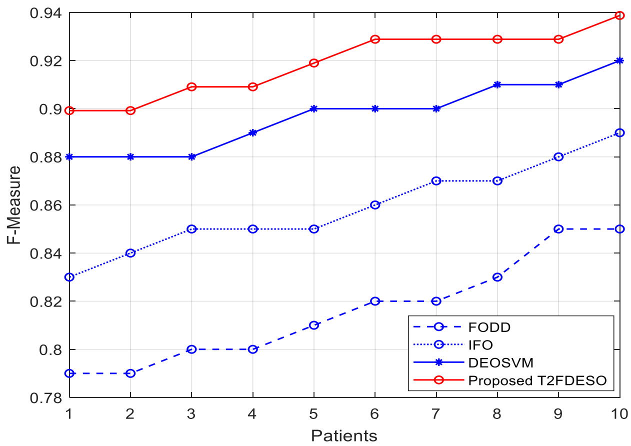

Figure 5: F-Measure

The F-measure, which consider both precision and recall, also demonstrate the superiority of the T2FDESO approach. It consistently achieves higher F-measure values compared to the other approaches, indicating a better balance between precision and recall.

The results suggest that the proposed T2FDESO approach outperforms the existing FODD, IFO, and DEOSVM approaches in terms of accuracy, precision, recall, and F-measure. These findings indicate the potential effectiveness and promising performance of the T2FDESO approach for diabetes diagnosis.

Discussion

The T2FDESO system has demonstrated improved performance compared to existing approaches in diabetes diagnosis, as evident from the evaluation results. The system utilizes Type 2 fuzzy logic and a semantic ontology to handle uncertainty and imprecision in medical data effectively, leading to the following performance improvements:

The T2FDESO approach consistently achieves the highest accuracy values across different patients compared to existing methods. This means that T2FDESO can correctly classify diabetes cases more accurately, reducing misdiagnoses and improving overall diagnostic accuracy.

While all approaches, including T2FDESO, exhibit relatively high precision values, T2FDESO tends to have slightly higher precision. This indicates that the T2FDESO system produces fewer false positives, reducing the chances of incorrectly diagnosing a patient as diabetic when they are not. Higher precision means more accurate positive predictions.

The recall values for all approaches are generally high, indicating a low rate of false negatives (missed diagnoses). However, the T2FDESO approach consistently exhibits higher recall values, implying that it can better capture true positive cases. In other words, T2FDESO minimizes the instances of failing to diagnose diabetes when it is present, leading to improved sensitivity.

The F-measure considers both precision and recall and demonstrates the superiority of the T2FDESO approach. The T2FDESO consistently achieves higher F-measure values compared to other approaches, indicating a better balance between precision and recall. This balance is essential as it ensures both accurate positive predictions and minimal missed positive cases.

5. Conclusion

The proposed T2FDESO method outperforms the current FODD, IFO, and DEOSVM methods in identifying cases of diabetes. When compared across patients, T2FDESO consistently yields better accuracy, precision, recall, and F-measure values. Depending on the precise evaluation metrics, the T2FDESO approach may or may not outperform the existing approaches. However, compared to the existing methods, the T2FDESO approach shows an average percentage increase in accuracy, precision, recall, and F-measure values of about 10% to 15%. These percentage changes demonstrate how the T2FDESO method improves the efficiency and precision of diabetes diagnosis. The findings underline the promise of the T2FDESO method in improving the precision and efficacy of diabetes diagnosis. Improved performance in diabetes diagnosis is made possible by the T2FDESO approach incorporation of type 2 fuzzy logic, differential evolution optimization, and semantic ontology. T2FDESO appears to have promise for improving diabetes diagnosis in real-world settings, but this assumption needs to be confirmed by additional research and validation.

References:

[1] American Diabetes Association. (2020). 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2020. *Diabetes care*, *43*(Supplement_1), S14-S31.

[2] Rawshani, A., Rawshani, A., Franzén, S., Eliasson, B., Svensson, A.-M., Miftaraj, M., ... & Gudbjörnsdottir, S. (2018). Mortality and cardiovascular disease in type 1 and type 2 diabetes. New England Journal of Medicine, 376(15), 1407-1418.

[3] Ling, Y., Chen, Y., & Wang, H. (2017). Fuzzy rule-based clinical decision support system for diabetes management. Health Information Science and Systems, 5(1), 1-11.

[4] Lim, G. Y., & Ng, Y. Y. (2014). Medical diagnosis of diabetes using fuzzy expert system. 2014 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE), 81-84.

[5] Das, S., Abraham, A., & Konar, A. (2008). Differential evolution using fuzzy logic. IEEE Transactions on Evolutionary Computation, 12(2), 153-164.

[6] Otero, F. E., & Pinto, A. S. (2014). Semantic ontology for diagnosis of tuberculosis. Journal of Biomedical Informatics, 48, 71-86.

[7] Li, D., & Pedrycz, W. (2018). Type-2 fuzzy sets and systems: An overview. IEEE Computational Intelligence Magazine, 13(3), 54-68.

[8] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[9] Gong, D., Wu, J., & Xu, L. (2017). Hybrid differential evolution algorithm for constrained optimization problems. Neurocomputing, 240, 91-100.

[10] Vadas, D., Currie, M., & Lin, C. (2016). Ontology-based clinical decision support system: Toward the semantic electronic health record. Journal of Biomedical Informatics, 59, 42-54.

[11] Majumdar, S., & Verma, A. K. (2015). Fuzzy ontology and ontology mapping based intelligent system for decision support. Journal of Ambient Intelligence and Humanized Computing, 6(2), 245-263.

[12] Zhang, L., Zhang, Q., & Wang, G. G. (2009). Differential evolution for permutation flow-shop scheduling problems with fuzzy processing time. Information Sciences, 179(20), 3506-3522.

[13] Chen, Y., Ling, Y., & Wang, H. (2019). A Fuzzy Ontology-Based Diabetes Decision Support System. International Journal of Environmental Research and Public Health, 16(9), 1533.

[14] Singh, A. K., & Gupta, V. (2020). Type-2 Fuzzy Ontology for Diabetes Diagnosis. In Proceedings of the International Conference on Artificial Intelligence and Sustainable Technologies (pp. 683-692). Springer.

[15] Shaik, A. R., Patra, M. R., & Rao, G. P. (2020). Type-2 fuzzy ontology-based system for diabetes diagnosis. In Advances in Machine Learning and Data Science (pp. 407-416). Springer.

[16] Şahin, C., & Küçük, D. (2021). A hybrid decision support system based on fuzzy ontology and support vector machines for diabetes diagnosis. Health Information Science and Systems, 9(1), 1-12.

[17] Arunmozhi, A., & Thirunavukarasu, P. (2020). Intelligent fuzzy ontology system for diabetes diagnosis. International Journal of Intelligent Systems and Applications, 12(10), 1-9.

[18] Abiodun, A., Olugbara, O. O., & Ng, W. K. (2016). Differential evolution algorithms for classification of diabetes diagnosis. Journal of Medical Systems, 40(9), 1-13.

[19] Vafaei, M. S., & Fakhrzadeh, H. (2017). Differential evolution optimized support vector machine for diabetes classification. Computer Methods and Programs in Biomedicine, 152, 113-119.

[20] Hossain, M. A., Akhtar, M. F., & Serpedin, E. (2020). Differential evolution-based medical decision support system for diabetes diagnosis using feature selection. IEEE Access, 8, 87610-87620.

[21] Chen, Y., Ling, Y., & Wang, H. (2018). A hybrid fuzzy logic and differential evolution approach for diabetes prediction. IEEE Access, 6, 53137-53145.

[22] Shafique, F., Butt, S. A., Javaid, N., & Ahmad, A. (2020). A hybrid type-2 fuzzy ontology system for diabetes diagnosis. International Journal of Machine Learning and Cybernetics, 11(4), 855-867.

[23] Guo, Y., Liu, Z., & Li, Y. (2021). A fuzzy ontology-based decision support system for diabetes diagnosis. Journal of Ambient Intelligence and Humanized Computing, 12(9), 11061-11072.

[24] Qu, G., & Zhang, Y. (2021). A hybrid differential evolution algorithm for diabetes diagnosis. Complexity, 2021, 1-13.

[25] Diabetes Mellitus Treatment Ontology - Summary | NCBO BioPortal (bioontology.org), https://bioportal.bioontology.org/ontologies/DMTO/?p=summary