Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Fuzzy logic and machine learning for diabetes risk prediction using modifiable factors

Rabia Khushal *, Ubaida Fatima

Department of Mathematics, NED University of Engineering and Technology Karachi, Karachi, Pakistan

A R T I C L E I N F O

Article history:
Received 28 August 2024
Received in revised form
14 November 2024
Accepted 29 November 2024

Keywords:
Diabetes risk
Fuzzy logic
Modifiable factors
Machine learning
Lifestyle intervention

A B S T R A C T

Diabetes mellitus, a global health concern, includes type 1 diabetes, with an uncontrollable risk, and type 2 diabetes, where risk can be managed through lifestyle modifications. This study examines the impact of modifiable risk factors—diet, physical activity, and body mass index (BMI)—on type 2 diabetes development. Using fuzzy logic, binary variables from a healthcare diabetes dataset were transformed into a fuzzy format, generating three output classes: "no diabetes risk," "possible diabetes risk," and "diabetes risk present." The intermediate class, "possible diabetes risk," serves as an alert for adopting healthier lifestyles to mitigate risk. Machine learning was applied to both the original and fuzzy-transformed datasets. While the original dataset provided binary outputs with moderate accuracy and higher computation times, the fuzzy-transformed dataset yielded more nuanced predictions, reduced computation time, and improved classifier performance. This approach enhances diabetes risk assessment and supports proactive interventions.

1. Introduction

Diabetes mellitus is a chronic metabolic disorder. It is rapidly progressing globally and causes substantial social, economic, and health consequences (Kaul et al., 2023). Insulin is secreted by the pancreas. Insulin plays a significant role in maintaining the blood glucose concentration level. The pancreas also produces glucagon, another hormone that functions in conjunction with insulin. Glucagon's function is to boost blood sugar levels if they fall too low, whereas insulin's function is to reduce blood sugar levels when necessary. By using this mechanism, the body makes sure that blood glucose levels stay within predetermined ranges, allowing the body to continue functioning (Guyton and Hall, 2011).

Diabetes mellitus is a bihormonal disorder that results from combined defects of insulin and glucagon (Knudsen et al., 2019). Diabetes mellitus is classified into two types. Type 1 diabetes mellitus, often called juvenile diabetes, usually develops before the age of 40. It is an insulin-dependent condition where the body produces little or no insulin. Type 2 diabetes mellitus, the more common type, typically occurs in adults and is therefore also known as adult-onset diabetes. It is a non-insulin-dependent condition where the body becomes resistant to using insulin effectively. Common symptoms of diabetes mellitus include excessive thirst, frequent urination, increased hunger, fatigue, blurry vision, nerve damage, and weight gain (Guyton and Hall, 2011; Barrett et al., 2019).

The two main causes of diabetes are obesity and the aging population. Pre-existing obesity and diabetes mellitus increase the risk of developing COVID-19 (Conte et al., 2024). In industrialized countries, the prevalence of diabetes mellitus is 5-6 %, and it will increase in the future (Taxirovich et al., 2024). Individuals suffering from diabetes mellitus are at a greater risk of developing cardiovascular diseases. Moreover, diabetes mellitus is a significant cause of uremia-related blindness and death. Diabetic foot syndrome and lower extremity gangrene account for about 40% of all non-traumatic lower extremity amputations (Husna et al., 2024).

Some risk factors for diabetes can be influenced by behavioral changes, while others, such as age and family history, cannot be altered. The risk factors for Type 1 diabetes are non-modifiable, including age and family history. In contrast, Type 2 diabetes is associated with both modifiable and non-modifiable risk factors. Non-modifiable factors include age, race, family history, and gestational diabetes. Modifiable

* Corresponding Author.
Email Address: rabiakhushal21@gmail.com (R. Khushal)
https://doi.org/10.21833/ijaas.2024.12.025
Corresponding author's ORCID profile:
https://orcid.org/0009-0002-5437-3032

factors include weight, physical activity, and dietary habits. Maintaining a healthy body mass index (BMI) of less than 25 can help reduce the risk of Type 2 diabetes associated with modifiable factors (NIDDK, 2022).

A machine learning algorithm is a collection of rules used by an artificial intelligence system for discovering data patterns and output prediction. Gaussian Naïve Bayes (GNB) follows the Bayes theorem and depicts that each predictor has an equal influence on the outcome prediction. Support vector machines (SVM) construct a hyperplane known as a decision boundary, dividing the data point classes on either side of the plane. Adaptive boosting (AB) is a boosting technique that combines the predictive ability of multiple base estimators. Fuzzy logic is introduced by Zadeh (1965) and deals with the uncertainty of knowledge. In a fuzzy logic system, crisp inputs are transformed into fuzzy values through the fuzzification interface. These fuzzy values are then processed by the inference engine, which uses the fuzzy knowledge base to generate a fuzzy output. Finally, the fuzzy output is converted back into a crisp output through the defuzzification process (Hentout et al., 2023).

This research utilizes fuzzy logic and machine learning to predict diabetes risk based on key modifiable factors, including diet, physical activity, and BMI. The standard approach of applying machine learning to diabetes datasets typically offers limited insights, only indicating whether diabetes risk is present or absent. Additionally, input predictors provide restricted information, focusing solely on whether a factor is followed or not.

To address these limitations, fuzzy logic is incorporated to account for uncertainty and introduce the possibility of intermediate outcomes, represented by a "maybe" factor. The dataset is transformed into a fuzzy format, and machine learning techniques are applied to generate a fuzzy output. This output indicates whether the diabetes risk is present, absent, or possibly present. Integrating fuzzy logic with machine learning optimizes both performance metrics and computation time while enhancing the clarity and accuracy of risk predictions.

The research is organized into several sections: Section 2 provides a literature review, Section 3 outlines the methodology, Section 4 presents the results and discussion, and Section 5 concludes the study.

2. Literature review

The literature review relevant to the research work is, in the study of Khushal and Fatima (2024a), a novel methodology comprised of fuzzy machine learning logic has been proposed, which is used to diagnose the absence, presence, and risk of development of polycystic ovary syndrome (PCOS). Similarly, in another study, fuzzy computing concepts were used to determine whether the individual is living a healthy lifestyle or an unhealthy lifestyle (Khushal and Fatima, 2024b). In the study of Moiz et al. (2024), Bio-Link Strength, a fuzzy membership function, has been introduced, which uses fuzzy set theory concepts for the quantification of interaction strength with continuous values between 0 and 1. In a study by Ali et al. (2024), heart disease is diagnosed using a fuzzy expert system. The if-then rules have been formulated by doing extensive research and by consulting doctors in the relevant field. The proposed model has been tested using the Cleveland dataset, and it shows 98.08 % accuracy. A fuzzy homotopy analysis method was discussed in the study of Khwayyit and Suhhiem (2024) to obtain approximate but analytical solutions to initial value problems that contain fuzzy coefficients. The proposed approach allows the solving of fuzzy differential equations as an endless series containing fuzzy numbers.

For the diagnosis of diabetes at the initial stage, Talukder et al. (2024) used different machine-learning techniques on 4 different diabetes datasets. They have found out that random forest surpasses all other techniques for datasets 1 and 2. Moreover, decision tree and extreme gradient boosting surpass all remaining techniques for datasets 3 and 4. It is concluded that the proposed method can be employed for early diagnosis of the metabolic disorder mentioned. For the prediction of future dengue cases in Malaysia, Mustaffa et al. (2024) employed three forecasting techniques, including the neural network autoregressive technique. A prediction model for the diagnosis of Covid-19 is built using different machine learning techniques. The model also provides information regarding the severity of the illness (Luna-Benoso et al., 2024). Machine learning techniques have also been used to forecast the usage of electricity in homes. The accuracy has been improved due to the extraction of correlated features (Janjua et al., 2024). In another study, a forest-optimized neural network classifier model was proposed, and a slime mold algorithm was used for efficient feature selection. The technique is applied to the dataset of low academic performance in higher education. It is concluded that the proposed method is effective and shows improved results (Begum and Ashok, 2024).

A correlation study by Nurhasanah et al. (2024) was conducted to find out the association between diet and mental health of older people suffering from diabetes, and it is concluded that anxiety in older people suffering from diabetes can hinder the ability to follow a diet plan. In a study by Modak and Jha (2024), diabetes, a global concern, has been diagnosed using different machine-learning techniques. Both foundation machine learning and ensemble methods are commonly utilized for diabetes prediction. The dataset has been extracted from Kaggle. Among all machine learning techniques, the accuracy of boost techniques surpasses all other techniques. Gestational diabetes mellitus is experienced during pregnancy due to the increased levels of glucose. In a study by Hasan et al. (2024), a fuzzy expert system has been used for early diagnosis of gestational diabetes using risk factors. It is concluded that the proposed method can serve as a supporting tool in the healthcare environment. A prediction model for diabetes risk assessment has been developed by integrating lifestyle factors with health indicators (Nandan et al., 2024). This research has made use of a variety of datasets. Machine learning techniques have been applied to construct models. The accuracy of the proposed framework improved with the addition of lifestyle factors and health indicators.

For the detection and prediction of diabetes, a fuzzy logic-based model has been described in Godfrey et al. (2023). The rules for diagnosis of diabetes and its risk factors have been acquired from experts in the field. All the gathered information has been incorporated during the model development. The significance of this model is that it can be deployed on edge-based devices. Hence, no medical experts would be needed. Digital pathology uses a virtual microscope for identification. Digital pathology is linked with different techniques, such as machine learning techniques (Gautam Goswami et al., 2023). Type-2 fuzzy logic-based nominal model for the robotic system is discussed in Alnufaie (2023). Money laundering is a significant global threat; therefore, Masrom et al. (2023) evaluated professional accountants' adherence to anti-money laundry regulations by employing machine learning techniques. In a study by Fatima et al. (2023), a novel clustering coefficient-dependent degree centrality has been introduced to analyze the links of the profitable nodes in a large product network. Modi et al. (2023) have addressed lifestyle diseases that are linked to people's daily activities, including smoking, alcohol intake, physical inactivity, and overeating. Additionally, typical machine learning methods for creating diagnostic models are provided. In a study by Abdullah et al. (2018), a fuzzy expert system was developed to assess the risk of developing diabetes. Data has been acquired from the journal and by interviewing domain experts. They also developed a GUI (Graphical User Interface) to aid doctors who are not experts in fuzzy logic. With the help of this system, the diagnostic time can be reduced since the outcome of the diagnosis can be known in a matter of seconds.

3. Methodology

3.1. Dataset description

The dataset utilized for the implementation of the proposed methodology is the diabetes prediction data. The dataset has been extracted from the Kaggle website. There are 17 input features, and the output variable is the binary variable. Total number of patients is 70692. The data description is depicted in Fig. 1.

Age	Sex	HighChol	CholCheck	BMI	Smoker	HeartDiseaseorAttack	PhysActivity	Fruits
4.0	1.0	0.0	1.0	26.0	0.0	0.0	1.0	1.0
12.0	1.0	1.0	1.0	26.0	1.0	0.0	0.0	1.0
13.0	1.0	0.0	1.0	26.0	0.0	0.0	1.0	1.0
11.0	1.0	1.0	1.0	28.0	1.0	0.0	1.0	1.0
8.0	0.0	0.0	1.0	29.0	1.0	0.0	1.0	1.0
Veggies	HvyAlcoholConsump	GenHlth	MentHlth	PhysHlth	DiffWalk	Stroke	HighBP	Diabetes
1.0	1.0	3.0	5.0	30.0	0.0	0.0	1.0	0.0
1.0	0.0	3.0	0.0	0.0	0.0	0.0	1.0	0.0
1.0	1.0	1.0	0.0	10.0	0.0	0.0	0.0	0.0
1.0	1.0	3.0	1.0	3.0	0.0	0.0	1.0	0.0
1.0	1.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0

Fig. 1: Considered diabetes dataset description

3.2. Data preprocessing and feature selection

The considered dataset has been preprocessed to make sure the data is in an appropriate form before the model training. The dataset description has been checked, and features have been visualized to determine the presence of null values, outliers, and categorical values. For feature selection, basic modifiable risk factors for type 2 diabetes, as discussed in Section 1 have been considered, which are BMI, Physical activity, fruits, and vegetables.

3.3. Fuzzy concepts implementation on a considered dataset

The fuzzy logic is applied to the binary input parameters of the dataset. The only numerical parameter from the modifiable risk factors is the BMI, which is divided into two categories: the healthy range if the BMI is less than 25 and the unhealthy range if the BMI is greater than 25. The division is done in the light of literature. The remaining three binary variables were fuzzified by comparing them with each other. If all variables have the same kind of input value, the same value is assigned to their common output. If the value of two variables is 0 and the value of the remaining variable is 1, 0.25 is assigned to their common output. If the value of two variables is 1 and the value of the remaining variable is 0, 0.75 is assigned to their common output. This transformation in the form of a truth table is represented by Table 1.

Table 1: Truth table of fuzzy data transformation of binary input variables

a	b	c	z
0	0	0	0
0	0	1	0.25
1	1	0	0.75
0	1	0	0.25
0	1	1	0.75
1	0	0	0.25
1	0	1	0.75
1	1	1	1

Mathematically, this fuzzy transformation can be represented as

$$z = \begin{cases} 0 \; iff \; a = 0 \; AND \; b = 0 \; AND \; c = 0 \\ 1 \; iff \; a = 1 \; AND \; b = 1 \; AND \; c = 1 \\ 0.25 \; iff \; either \; a = 0 \; AND \; b = 0 \; AND \; c = 1 \; OR \; a = 1 \; AND \\ b = 0 \; AND \; c = 0 \; OR \; a = 0 \; AND \; b = 1 \; AND \; c = 0 \\ 0.75 \; iff \; either \; a = 0 \; AND \; b = 1 \; AND \; c = 1 \; OR \; a = 1 \\ AND \; b = 1 \; AND \; c = 0 \; OR \; a = 1 \; AND \; b = 0 \; AND \; c = 1 \end{cases} \quad (1)$$

The three binary input variables have been transformed into a fuzzy format by following the above-stated equation and truth table. Their output variable is called ***z*** (healthy lifestyle factors). Previously, the binary variables provided only limited information, which included yes and no. However, by implementing fuzzy logic concepts, they are fuzzified and now also consider their uncertain side. Instead of providing yes and no information, it now provides middle values, which are the factors that may be absent and factors that may be present. Due to this transformation, the output variable of the dataset, which is diabetes risk prediction, is also transformed from binary output to fuzzy output. Hence, now provides three kinds of output, which are diabetes risk is absent (0), may be present (1), and present (2), as shown in Table 2. Thus, if the diabetes risk may be present, the necessary factors that are absent can be followed to remove the risk of diabetes.

Table 2: Depiction of fuzzy data transformation of considered dataset

Physical activity	Fruits	Veggies	Healthy lifestyle factors	BMI	Diabetes risk prediction
0	0	0	0	20	1
1	1	0	0.75	26	2
1	1	1	1	34	1
1	0	1	0.75	24	0
1	0	0	0.25	20	2

3.4. Machine learning of normal and fuzzy transform data

Machine learning techniques are first applied to the non-fuzzy dataset. Machine learning classifiers that have been utilized are GNB, SVM, and AB. The usual metrics of classifiers and computation time have been computed. Afterward, machine learning techniques, GNB, SVM, and AB, were applied again to the fuzzy transformed dataset. As depicted in Fig. 2, the crisp input that is the considered dataset goes into the fuzzification box, where it is transformed into a fuzzy input, as discussed in section 3.2. Subsequently, the transformed dataset goes into an inference engine whereby, using machine learning techniques and rules stored in the rule base, the fuzzy output is produced. The rules are represented in Table 3.

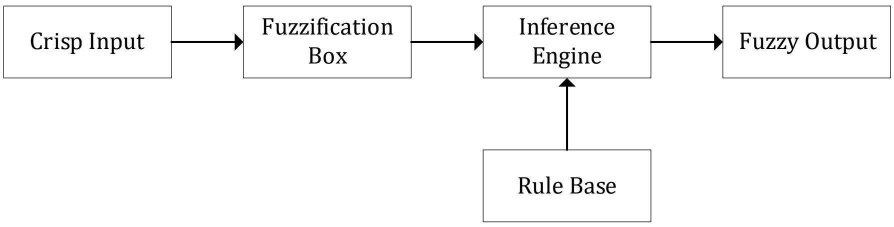

Fig. 2: Visualization of fuzzy logic fusion with machine learning

Table 3: Rules for diagnosis of diabetes risk

#	Rules
1	If BMI is in the healthy range and healthy lifestyle factors are present, then diabetes risk is absent
2	If BMI is in the healthy range and healthy lifestyle factors may be present, then diabetes risk may be absent
3	If BMI is in the healthy range and healthy lifestyle factors may be absent, then diabetes risk may be absent
4	If BMI is in the healthy range and healthy lifestyle factors are absent, then diabetes risk may be absent
5	If BMI is in the unhealthy range and healthy lifestyle factors are present, then diabetes risk may be present
6	If BMI is in the unhealthy range and healthy lifestyle factors may be present, then diabetes risk is present
7	If BMI is in the unhealthy range and healthy lifestyle factors may be absent, then diabetes risk is present
8	If BMI is in the unhealthy range and healthy lifestyle factors may be absent, then diabetes risk is present

4. Results and discussion

The results of machine learning implementation on normal and fuzzy transformed datasets have been discussed in this section. The metrics of classifiers and computation time have been calculated. Accuracy is used to find out the percentage of correct predictions. Precision is the percentage of accurate predictions made overall for each class across all classes in the sample. Recall is the percentage of accurate predictions made from each class's total predictions across all classes in the dataset. The f-1 score is the harmonic mean of precision and recall.

4.1. Results of considered diabetes dataset

The application of machine learning on the normal dataset shows that the average result and accuracy are in the range of 60s, as shown in Table 4. The computation time is also long. However, by incorporating the fuzzy concept, the accuracy of all three machine learning techniques, GNB, SVM, and AB, has been significantly improved. The computational time for all three techniques has been significantly improved, as demonstrated in Table 5. This indicates that the proposed method is effective for optimizing both computation time and classifier performance metrics. Additionally, it offers a more comprehensive approach to diabetes diagnosis, leading to greater accuracy. By considering the uncertainty in the output variable—indicating that diabetes risk may be present—the method provides a valuable alert for patients. This encourages them to adopt healthy lifestyle habits to reduce the risk of diabetes associated with key modifiable factors.

Table 4: Results of the normal dataset

Machine learning Classifier	Diabetes dataset				
	Accuracy	Macro avg precision	Macro avg recall	Macro avg F1-score	Time (Seconds)
GNB	0.62	0.63	0.62	0.62	0.26
SVM	0.64	0.64	0.64	0.63	471.58
AB	0.64	0.64	0.64	0.64	5.73

Avg: Average

Table 5: Results of fuzzy transformed diabetes dataset

Machine learning classifier	Fuzzy adapted diabetes dataset				
	Accuracy	Macro avg precision	Macro avg recall	Macro avg F1-score	Time (Seconds)
GNB	0.92	0.91	0.92	0.91	0.09
SVM	0.98	0.97	0.98	0.97	62.24
AB	0.89	0.91	0.81	0.82	2.84

4.2. Result of first validation dataset (PCOS dataset)

The validation of the technique is performed using the polycystic ovary syndrome dataset (PCOS) (Kottarathil, 2020). Modifiable factors, such as weight gain, fast food, regular physical activity, and BMI, have been considered from the dataset. The binary variables weight gain, regular sleep, and physical activity have been transformed into a single variable, z. Machine learning techniques are applied to the original dataset, and afterward, they are applied to the fuzzy transformed dataset. The machine learning techniques on the fuzzy transformed dataset have shown improved results, and computation time has also been optimized. Moreover, the concept provides a broad vision for diagnosing PCOS. The results of both implementations are shown in Tables 6 and 7.

Table 6: Results of normal PCOS dataset

Machine learning classifier	PCOS dataset				
	Accuracy	Macro avg precision	Macro avg recall	Macro avg F1-score	Time (Seconds)
GNB	0.87	0.90	0.87	0.87	0.05
SVM	0.94	0.95	0.93	0.94	0.54
AB	0.95	0.95	0.95	0.95	124.43

Table 7: Results of fuzzy transformed PCOS dataset

Machine learning classifier	Fuzzy adapted PCOS dataset				
	Accuracy	Macro avg precision	Macro avg recall	Macro avg F1-score	Time (Seconds)
GNB	0.92	0.91	0.92	0.91	0.09
SVM	0.98	0.97	0.98	0.97	62.24
AB	0.89	0.91	0.81	0.82	2.84

4.3 Results of the second validation dataset (Osteoporosis risk prediction dataset)

The proposed technique has also been validated using the osteoporosis risk prediction dataset (Kulkarni, 2024). The six modifiable risk factors considered from a dataset are smoking, alcohol consumption, physical activity, body weight, intake of calcium, and intake of vitamin D. The first three factors have been compared with each other, and their resultant value is termed z_1. Moreover, the remaining three factors have also been compared with each other. Their output is termed as z_2. These two modified variables have been used to find out the absence, may be present, and presence of osteoporosis risk in the individuals. Machine learning techniques were applied to the normal dataset and then to the fuzzy transformed dataset. Their results are represented in Tables 8 and 9. The implementation of machine learning techniques on the fuzzy transformed dataset has optimized all considered metrics of the classifiers along with the computation time.

The results of the research highlight several significant improvements achieved by integrating machine learning with fuzzy logic. Traditional machine learning techniques, as applied in previous studies, were limited in terms of providing normal accuracy and computation times. They did not address the inherent uncertainty in input and output variables and offered limited predictions, merely indicating whether the disease was present or absent. Furthermore, these methods lacked the capacity to provide proactive alerts to individuals regarding necessary lifestyle changes to mitigate risks.

In contrast, the proposed approach, which integrates fuzzy logic with machine learning, has demonstrated optimized accuracy and reduced computation times. By considering the uncertainty of input and output variables, the method broadens the scope of predictions, introducing a middle category that indicates the potential presence of risk. This additional category serves as a valuable warning mechanism, encouraging individuals to adopt preventive measures. Ultimately, this innovative approach provides a more comprehensive and effective method for predicting and mitigating disease risks.

Table 8: Results of normal osteoporosis risk prediction dataset

Machine learning classifier	Osteoporosis dataset				
	Accuracy	Macro avg precision	Macro avg recall	Macro avg F1-score	Time (Seconds)
GNB	0.87	0.91	0.87	0.87	0.3
SVM	0.89	0.89	0.89	0.89	5.4
AB	0.91	0.91	0.91	0.91	7.9

Table 9: Results of fuzzy transformed osteoporosis risk prediction dataset

Machine learning classifier	Fuzzy adapted osteoporosis dataset				
	Accuracy	Macro avg precision	Macro avg recall	Macro avg F1-score	Time (Seconds)
GNB	0.99	0.99	0.99	0.99	0.18
SVM	0.99	0.99	0.99	0.99	0.2
AB	0.99	0.99	0.99	0.99	2.5

5. Conclusion

Diabetes is a significant global health issue. While the risk factors for Type 1 diabetes are non-modifiable, those for Type 2 diabetes include both modifiable and non-modifiable factors. This research focuses on three key modifiable risk factors: physical activity, BMI, and diet. These factors were used to develop a machine learning model to predict whether the risk of diabetes is present or absent. However, real-world problems often involve uncertainty, where symptoms and factors do not provide a clear distinction between risk being present or absent. To address this, a combination of fuzzy logic and machine learning was introduced. Fuzzy concepts were applied to transform input variables, which in turn transformed the output variable. The transformed output provides three possible outcomes: the risk is absent, may be present, or present. The additional "may be present" category serves as a warning for patients to adopt healthier lifestyle habits to reduce the risk of Type 2 diabetes associated with modifiable factors. The proposed methodology was also validated using datasets for PCOS and osteoporosis risk prediction. The results were satisfactory, demonstrating that this method can be effectively applied to various healthcare datasets.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abdullah AA, Fadil NS, and Khairunizam W (2018). Development of fuzzy expert system for diagnosis of diabetes. In the International Conference on Computational Approach in Smart Systems Design and Applications, IEEE, Kuching, Malaysia: 1-8. https://doi.org/10.1109/ICASSDA.2018.8477635

Ali ML, Sadi MS, and Goni MO (2024). Diagnosis of heart diseases: A fuzzy-logic-based approach. PLOS ONE, 19(2): e0293112. https://doi.org/10.1371/journal.pone.0293112 **PMid:38319925 PMCid:PMC10846727**

Alnufaie L (2023). Fuzzy nonsingular fast terminal sliding mode controller for a robotic system. International Journal of Advanced and Applied Sciences, 10(10): 166-173. https://doi.org/10.21833/ijaas.2023.10.019

Barrett KE, Barman SM, and Boitano S (2019). Ganong's review of medical physiology. McGraw-Hill Education, New York, USA.

Begum S and Ashok MV (2024). A novel approach to mitigate academic underachievement in higher education: Feature selection, classifier performance, and interpretability in predicting student performance. International Journal of Advanced and Applied Sciences, 11(5): 140-150. https://doi.org/10.21833/ijaas.2024.05.015

Conte C, Cipponeri E, and Roden M (2024). Diabetes mellitus, energy metabolism, and COVID-19. Endocrine Reviews, 45(2): 281-308. https://doi.org/10.1210/endrev/bnad032 **PMid:37934800 PMCid:PMC10911957**

Fatima U, Hina S, and Wasif M (2023). A novel global clustering coefficient-dependent degree centrality (GCCDC) metric for large network analysis using real-world datasets. Journal of Computational Science, 70: 102008. https://doi.org/10.1016/j.jocs.2023.102008

Gautam Goswami N, Karnad S, Sampathila N, Bairy GM, Chadaga K, and Swathi KS (2023). Current trends of artificial intelligence and applications in digital pathology: A comprehensive review. International Journal of Advanced and Applied Sciences, 10(12): 29-41. https://doi.org/10.21833/ijaas.2023.12.004

Godfrey K, Rushingabigwi G, Habtie AB, and Gatare I (2023). Fuzzy logic-based driven model for detection and prediction of diabetes mellitus type 2. In the IEEE International Conference on Advanced Systems and Emergent Technologies, IEEE, Hammamet, Tunisia: 1-5. https://doi.org/10.1109/IC_ASET58101.2023.10150948

Guyton AC and Hall JE (2011). Textbook of medical physiology. Saunders Elsevier, Philadelphia, USA.

Hasan N, Islam MS, and Hayat MT (2024). A fuzzy logic approach for predicting gestational diabetes mellitus using risk factors. In the 6th International Conference on Electrical Engineering and Information and Communication Technology, IEEE, Dhaka, Bangladesh: 1479-1484. https://doi.org/10.1109/ICEEICT62016.2024.10534451

Hentout A, Maoudj A, and Aouache M (2023). A review of the literature on fuzzy-logic approaches for collision-free path planning of manipulator robots. Artificial Intelligence Review, 56(4): 3369-3444. https://doi.org/10.1007/s10462-022-10257-7

Husna C, Akmal A, Syarif H, and Agustina S (2024). Exploring the connection between diabetes duration, HbA1c levels, and peripheral neuropathy in type 2 diabetes patients. International Journal of Advanced and Applied Sciences, 11(4): 236-240. https://doi.org/10.21833/ijaas.2024.04.025

Janjua JI, Ahmad R, Abbas S, Mohammed AS, Khan MS, Daud A, Abbas T, and Khan MA (2024). Enhancing smart grid electricity prediction with the fusion of intelligent modeling and XAI integration. International Journal of Advanced and Applied Sciences, 11(5): 230-248. https://doi.org/10.21833/ijaas.2024.05.025

Kaul K, Tarr JM, Ahmad SI, Kohner EM, and Chibber R (2013). Introduction to diabetes mellitus. In: Ahmad SI (Ed.), Diabetes: Advances in experimental medicine and biology: 1–11. Volume 771, Springer, New York, USA. https://doi.org/10.1007/978-1-4614-5441-0_1 **PMid:23393665**

Khushal R and Fatima U (2024a). Fuzzy machine learning logic utilization on hormonal imbalance dataset. Computers in Biology and Medicine, 174: 108429. https://doi.org/10.1016/j.compbiomed.2024.108429

Khushal R and Fatima U (2024b). Fuzzy computing in healthcare. In the International Visualization, Informatics and Technology Conference, IEEE, Kuala Lumpur, Malaysia: 78-83. https://doi.org/10.1109/IVIT62102.2024.10692652

Khwayyit RI and Suhhiem MH (2024). Applying the fuzzy homotopy analysis method to solve fuzzy initial value problems with variable coefficients. International Journal of Advanced and Applied Sciences, 11(4): 85-92. https://doi.org/10.21833/ijaas.2024.04.010

Knudsen JG, Hamilton A, Ramracheya R, Tarasov AI, Brereton M, Haythorne E, Chibalina MV, Spegel P, Mulder H, Zhang Q, and Ashcroft FM et al. (2019). Dysregulation of glucagon secretion by hyperglycemia-induced sodium-dependent reduction of ATP production. Cell Metabolism, 29(2): 430-442. https://doi.org/10.1016/j.cmet.2018.10.003 **PMid:30415925 PMCid:PMC6370947**

Kottarathil P (2020). Polycystic ovary syndrome (PCOS). Available online at: https://www.kaggle.com/datasets/prasoonkottarathil/polycystic-ovary-syndrome-pcos

Kulkarni A (2024). Osteoporosis risk prediction. Available online at: https://www.kaggle.com/datasets/amitvkulkarni/lifestyle-factors-influencing-osteoporosis

Luna-Benoso B, Martínez-Perales JC, Cortés-Galicia J, and Morales-Rodríguez US (2024). Detection and risk assessment of COVID-19 through machine learning. International Journal of Advanced and Applied Sciences, 11(1): 207-216. https://doi.org/10.21833/ijaas.2024.01.025

Masrom S, Tarmizi MA, Halid S, Rahman RA, Rahman ASA, and Ibrahim R (2023). Machine learning in predicting anti-money laundering compliance with protection motivation theory among professional accountants. International Journal of Advanced and Applied Sciences, 10(7): 48-53. https://doi.org/10.21833/ijaas.2023.07.007

Modak SKSS and Jha VK (2024). Diabetes prediction model using machine learning techniques. Multimedia Tools and Applications, 83(13): 38523-38549. https://doi.org/10.1007/s11042-023-16745-4

Modi K, Singh I, and Kumar Y (2023). A comprehensive analysis of artificial intelligence techniques for the prediction and prognosis of lifestyle diseases. Archives of Computational Methods in Engineering, 30(8): 4733-4756. https://doi.org/10.1007/s11831-023-09957-2

Moiz A, Fatima U, and Haque MZU (2024). A new framework for pinpointing crucial proteins in protein-protein interaction networks. IEEE Access, 12: 108425-108444. https://doi.org/10.1109/ACCESS.2024.3437215

Mustaffa NA, Zahari SM, Farhana NA, Nasir N, and Azil AH (2024). Forecasting the incidence of dengue fever in Malaysia: A comparative analysis of seasonal ARIMA, dynamic harmonic regression, and neural network models. International Journal of Advanced and Applied Sciences, 11(1): 20-31. https://doi.org/10.21833/ijaas.2024.01.003

Nandan TK, Ramesh G, Reddy GS, ShivaRam G, Anusha M, and Shafi S (2024). Predicting diabetes with integrated health-lifestyle fusion. In the 1st International Conference on Trends in Engineering Systems and Technologies, IEEE, Kochi, India: 1-6. https://doi.org/10.1109/ICTEST60614.2024.10576176

NIDDK (2022). Risk factors for type 2 diabetes. National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, USA.

Nurhasanah N, Juanita J, and Putra A (2024). Anxiety and diet among elderly people suffering from diabetes mellitus in Banda Aceh, Indonesia: A correlational study. International Journal of Advanced and Applied Sciences, 11(1): 123-128. https://doi.org/10.21833/ijaas.2024.01.015

Talukder MA, Islam MM, Uddin MA, Kazi M, Khalid M, Akhter A, and Ali Moni M (2024). Toward reliable diabetes prediction: Innovations in data engineering and machine learning applications. Digital Health. https://doi.org/10.1177/20552076241271867 **PMid:39175924 PMCid:PMC11339751**

Taxirovich DA, Ulugbekovna RN, and Abduxalimova YJ (2024). Status and prospects for the fight against diabetes mellitus. Educational Research in Universal Sciences, 3(1): 4-9.

Zadeh LA (1965). Fuzzy sets. Information and Control, 8(3): 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X