EVALUATING MACHINE LEARNING MODELS FOR PREDICTING THE SEVERITY OF COMPLICATIONS DUE TO TYPE 1DIABETES: INSIGHTS FROM CLINICAL DATA

MRS. ANISHA SATHISH

E-mail: anisat556@gmail.com

Abstract - Predicting major adverse complications in individuals with type 1 diabetes is a critical challenge in healthcare. Traditional approaches, such as statistical methods and clinical risk scores, often face limitations including linearity assumptions and lack of real-time data integration. This study explores the potential of advanced machine learning to improve predictive capabilities by employing four distinct models: K-Nearest Neighbors, Random Forest Classifier, Multilayer Perceptron, and Logistic Regression. These machine learning models offer a novel approach by leveraging advanced algorithms to analyze complex data patterns. The data set was carefully divided into training and testing cohorts to rigorously evaluate model performance. Our findings reveal that the Random Forest Classifier achieved a superior accuracy rate of 92%, outperforming the other models and confirming the hypothesis that advanced machine learning techniques can effectively predict complications due to type 1 diabetes. The implications of these findings are substantial for medical practice. Integrating the high-accuracy Random Forest model into healthcare systems can improve patient outcomes through early intervention and personalized care plans. This research demonstrates the efficacy of machine learning in medical prediction and sets a new standard for diabetes management practices.

Keywords - Type 1 diabetes, machine learning, K-Nearest Neighbors, Random Forest Classifier, Multilayer Perceptron, Logistic Regression

I. INTRODUCTION

Every day, more and more people are getting diagnosed with type 1 diabetes. In the United States, about 3.2 out of every 1000 children under the age of 18 are diagnosed with diabetes [1]. Type 1 diabetes is a condition in which an individual's immune system targets the cells that produce insulin, and the cells of the individual are therefore unable to absorb glucose from the bloodstream. In a person without type 1 diabetes, insulin binds to the receptors of cells, signaling them to absorb glucose from the bloodstream [2]. In individuals with type 1 diabetes, this process doesn't occur, and sugar remains in the bloodstream. This can cause the glucose to bind to the wrong molecules, and can have disastrous consequences, including cardiovascular disease, neuropathy, kidney damage, diabetic retinopathy, glaucoma, foot problems, higher risk for developing other autoimmune diseases, and a higher risk for developing cancer [3].

These complications can be extremely difficult to predict, and can be a major source of stress for those with type 1 diabetes. Managing type 1 is stressful on its own. People with type 1 diabetes need to constantly check their blood sugar levels, maintain a strict routine, and keep track of the food they consume and the exercise they perform [4]. In a certain study participants reported having a stressful event related to type1 diabetes on nearly 60% of days [5]. Added with the risk of developing complications, having type 1 diabetes can often feel too overwhelming to be able to manage [6]. This may lead someone to give up on their health, and in severe cases, can lead to death.

Currently, the main approach to predict the severity of complications due to diabetes is through a doctor [7]. But even doctors can have a difficult time considering all the factors that may contribute to future complications. Using AI allows for a better overview of the problem by considering factors often forgotten by doctors, and can predict a more accurate risk of developing complications.

AI stands for artificial intelligence. Uses for artificial intelligence vary wildly, including learning, language, communication, security, healthcare, and research [8]. AI can be classified into different types, such as deep learning or machine learning [9].

AI models have been proven to be effective in classifying things for many different scenarios. Specifically, researchers have been looking into its benefits in healthcare [10]. Interest in using AI in healthcare has been growing in recent years, and there have been many applications for AI in this particular field. AI is being studied in several contexts regarding type 1 diabetes. From technology used to manage type 1 diabetes to predicting nutritional information, AI has high potential to improve the lives of those with type 1 diabetes. However, AI is still very much in testing. Privacy, as well as the drastic effect that a wrong prediction would have on health, are among the factors that should be taken into consideration when implementing AI in healthcare.

Having a physician nearby to consult when considering the complications due to type 1 diabetes would be very helpful, but may not always be

practical. It can be difficult for many families to reach a physician, due to any number of things, including financial constraint or extensive distance to reach the nearest physician. Using AI could be a practical alternative for those who may not readily have access to a physician.

Of course, if AI were to be implemented in such a way, it is crucial to make sure that it can work effectively and accurately in order to avoid further difficulties for those with type 1 diabetes. In our study, we aim to prove our hypothesis that AI can be effectively utilized to predict the severity of complications due to type 1 diabetes. By using AI models, we can better understand data by analyzing it for underlying patterns that humans may not be able to easily recognize. Hopefully, this can lead us to earlier diagnoses before health conditions become too severe and an overall better outcome for patients.

This paper has seven sections in total. The first section is the introduction. The second section highlights other works in this area. The third section specifies the methods used for this experiment, the process used in creating the model, and the results of the experiment. The fourth section is the results of our research. The fifth section discusses our findings in depth. The sixth and final section is the conclusion section.

II. RELATED WORKS

Regarding type 1 diabetes, there have been many studies that aim to improve its treatment. There have been numerous different approaches, including improvement in diabetes related technology and studying the risk of developing type 1 diabetes in the future. Type 1 diabetes is a condition that consists of many different aspects, including technology to help patients manage diabetes, physiological factors that contribute to the risk of developing type 1 diabetes or complications due to diabetes.

In a certain Finnish study, antibodies such as islet antigen-2 antibodies and GAD antibodies were evaluated for predicting the development of type 1 diabetes in individuals [11]. In type 1 diabetes, most B cells are rendered useless in the production of insulin by the body's immune system. Autoantibodies such as these are known to accompany the degradation of B cells in the development of type 1 diabetes. This study followed 3175 participants under the age of 18 for 27 years, and demonstrated the antibodies to be dynamic. Changes in the antibodies were correlated to a diagnosis of type 1 diabetes. This study shows the importance of antibody screening when considering the development of type 1 diabetes.

One study in this area was conducted by developing a prognostic model to predict complications due to type 1 diabetes [12]. Over a course of seven years, this study followed 1973 individuals with type 1 diabetes. This particular study incorporates five different factors that could be used to predict the likelihood of developing certain complications after a type 1 diabetes diagnosis. Those five factors are age, HbA1c waist/hip ratio, albumin/creatinine ratio, and HDL levels. These factors were incorporated into a Weibull regression model. Weibull regression models tend to be one of the most popular forms of parametric regression models in healthcare. It uses an equation to determine the survival of patients [13]. The C statistic was 0.74, demonstrating good discrimination. This model was validated in three cohorts in the USA and in Finland, and it aligned with or performed better than it was initially, though it had a tendency to underpredict complications in populations that were at high risk. Overall, this model was able to reliably predict high versus low risk patients, allowing for better treatment plans and improved decision making.

Currently, there are two main methods to determine blood glucose for people with diabetes. There is blood glucose monitoring, which consists of taking a pinprick of blood from the tip of the finger, and putting the blood into a device which can read it and determine the blood glucose levels of that individual. A more novel approach is continuous glucose monitoring, which offers patients the ability to constantly monitor the blood glucose levels over an extended period of time, as is said in the name. A study by Guido Freckmann et al aims to compare the methods of obtaining accuracy of blood glucose monitoring (BGM) to continuous glucose monitoring (CGM) [14]. This study has concluded that it is difficult to compare the methods of determining blood glucose accuracy when comparing BGM and CGM, although BGM has a longer history of research and data behind it to support its accuracy. CGMs are relatively new technology, and further advances will have to be made before they are able to be used without BGM as well.

A study by Tara Dohety et al has explored the use of machine learning algorithms to augment CGMs [15]. This is a promising avenue, as it could work around current issues with CGMs, including accuracy of the device, interference, and safety. However, introducing machine learning comes with its own set of issues. Algorithm bias, confidentiality, and security threats are among some of these concerns. These issues must be overcome in order for widespread usage of machine learning models in CGMs.

A more recent study by Annuzzi et al. explores nutritional factors that contribute to blood glucose levels around and after meals [16]. This study emphasizes the nuances brought by nutrition when predicting blood glucose. Using neural networks

across nine different scenarios, this study emphasized how detailed descriptions of nutritional value in food could increase the accuracy of predicting blood glucose levels around mealtimes. It also shows how important it is to personalize data according to patients, and that there isn't one best model for all people with type 1 diabetes.

Although research concerning type 1 diabetes and AI is incredibly varied, and covers several different areas within type 1 diabetes care and management, there are still gaps to be filled in research. Antibody markers, while incredibly helpful in diabetes screening, are not always the most reliable, due to their change over time. Models such as the Weibull Regression Model generally perform well, but tend to underpredict in populations that are at high risk, which is a critical issue when considering complications due to type 1 diabetes. BGM is the traditional method of checking blood sugar values, with CGMs not having the same reliability as BGM due to technical issues, which can result in life-threatening conditions such as hypoglycemia or ketoacidosis. Our study aims to address the gap in research about AI models and complications due to diabetes.

III. DATASET

This research uses data from various hospitals and diagnostic centers in Dhaka, Bangladesh, as cited by S. Asaduzzaman et al. [17]. It was collected through the use of a questionnaire developed by healthcare researchers and professionals in the medical field. It consists of the data for 306 individuals. The dataset includes 21 features like age, sex, area of residence, mother's level of education, HbA1c values, and BMI and one label column named "affected", which means if a person is affected by the complications associated with type 1 diabetes or not. 2 of these features are numeric type features, and the remaining 19 are object type. This data is balanced, with about half the patients predicted to be affected and half the patients predicted to not be affected.

IV. METHODOLOGY

We used Python to create the code. The software libraries used were NumPy and Pandas. The process of evaluating our models started with data preprocessing, which includes the cleaning of the dataset by removing extra features like features that did not seem to have any correlation with developing severe complications due to type 1 diabetes, such as removing height and weight because BMI was available. All categorical data was converted into a numerical form using one hot encoder so that it could be fed into the machine for training purposes. The processed dataset was split into two parts. 80% of it was used to train the machine learning models, and 20% was used to test their accuracies.

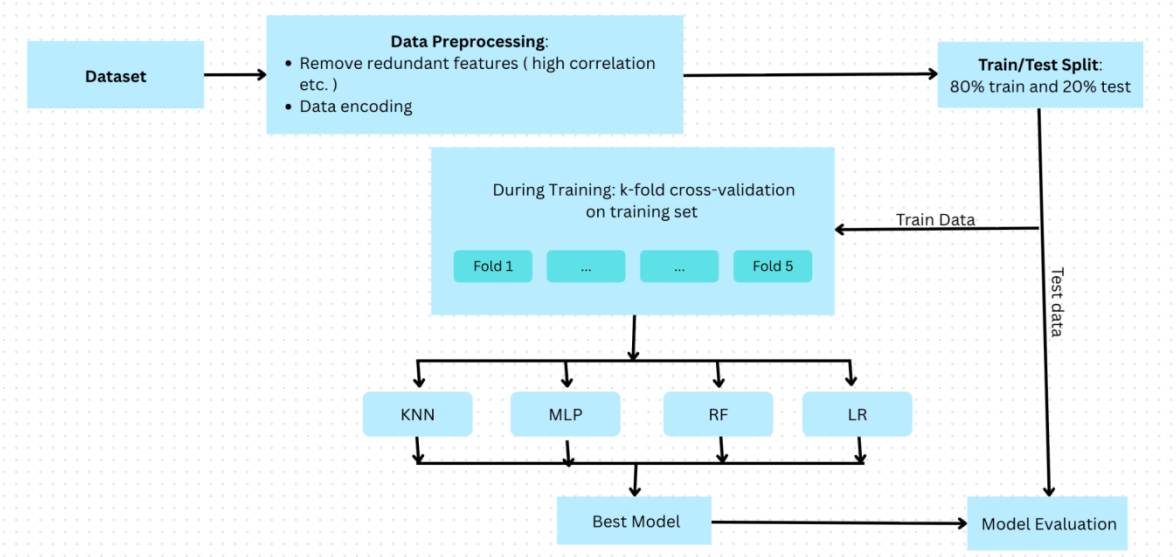


Figure 1: Flowchart of method used

Figure 1 shows how we trained our model. After data preprocessing, model training was next. Four different models were trained, each with different hyperparameter values. These models include K-Nearest Neighbors (KNN), Random Forest Classifier, Multiplayer Perceptron (MLP), and Logistic Regression (LR). Cross validation was a technique we employed to increase the effectiveness and accuracy of our models. The data was divided into two categories: training and testing. The training data was then split again, and some of it was used in cross validation. Cross validation helps a model apply itself to new data instead of relying solely on the data it was trained on. The best model was then chosen and tested. This process can be seen in Figure 1.

K-Nearest Numbers is a nonparametric supervised model that compares given points in order to estimate the value of a new point [18]. This model is most often used for classification and regression purposes. These neighboring points, or "k", are a chosen number of closest data points. Our k values were between 1 and 10, with higher values sometimes not picking up on finer patterns, but also did not incorporate noise.

Random Forest Classifier is a model that incorporates the decisions of several different trees to produce the most accurate result [19]. It is often used for categorization and prediction. It uses trees and depth as a metric to evaluate data. We used 100 trees and a depth of 10 to be able to pick up on complex patterns without overfitting.

Multilayer Perceptron is a popular type of neural network [20]. It progresses from the input layer to the output layer, with hidden layers in between. It goes through a nonlinear function. Its parameter is a learning rate, which helps it balance the error produced from the output.

Logistic Regression is a model used primarily for classification [21]. It is used when evaluating two distinct possibilities for classification, such as true/false or yes/no. It estimates the probability that the output will be in one of these two classifications. The maximum number of iterations determines how many times the model should be used depending on data size.

V. RESULTS

For this study, four different models were explored in order to predict if someone would develop severe complications due to type 1 diabetes. We used K-Nearest Neighbors, Logistic Regression, Random Forest Classifier, and Multi-Layer Perceptron. We experimented with different hyperparameters in order to achieve the highest accuracy with each model.

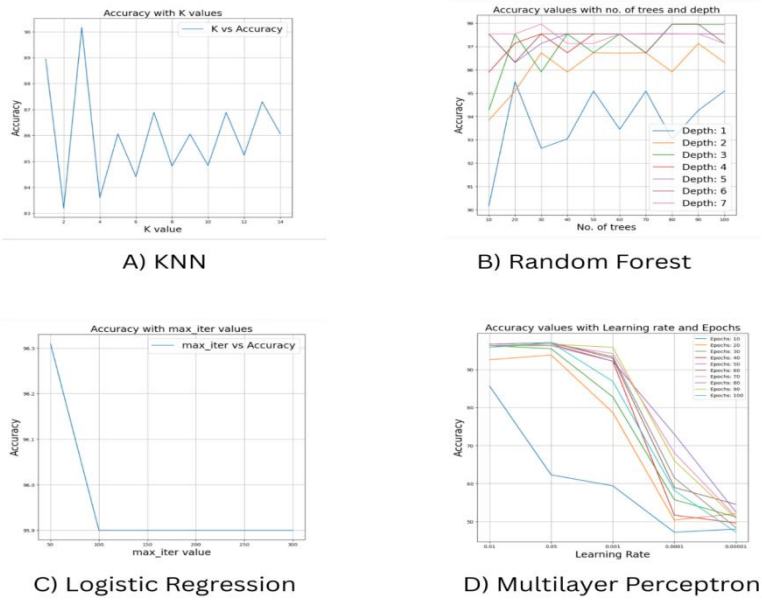


Figure 2: Graphs Comparing Accuracies to Hyperparameter values

The graphs seen in Figure 2 compile the accuracies achieved by different hyperparameters. Adjusting the hyperparameter values even slightly can significantly change the accuracy of the predictions of the models.

Model	Optimized Hyper parameter	Accuracy	Precision	Recall
KNN	k = 3	90.17	90.17	90.17
RF	Depth = 7 Trees = 30	97.96	97.96	97.96
MLP	Epochs = 60 Learning rate = 0.05	97.14	97.14	97.14
LR	Maximum iterations =50	96.31	96.31	96.31

Table 1: Metrics comparison of all models

The first model we used was K-Nearest Neighbors. It had the highest accuracy when "k" was three, meaning that three nearest neighbors were analyzed. Its highest achieved accuracy for this study was 90.17%. The next model we tested was Logistic Regression. It performed best when the number of maximum iterations was set to 50, with an accuracy of 96.31%. The third model we used was Random Forest Classifier. When the maximum depth was 7, and 30 trees were used, it produced an accuracy of 97.96%. This accuracy was the highest achieved by any of the models we used. The final model evaluated was Multilayer Perceptron. When it had 60 epochs and a learning rate of 0.05, it had an accuracy of 97.14. The best model was found to be Random Forest Classifier, with a training accuracy of 97.96% and a testing accuracy of 98.00%. Table 1 shows the highest accuracies of each model, and what hyperparameter value they were achieved at.

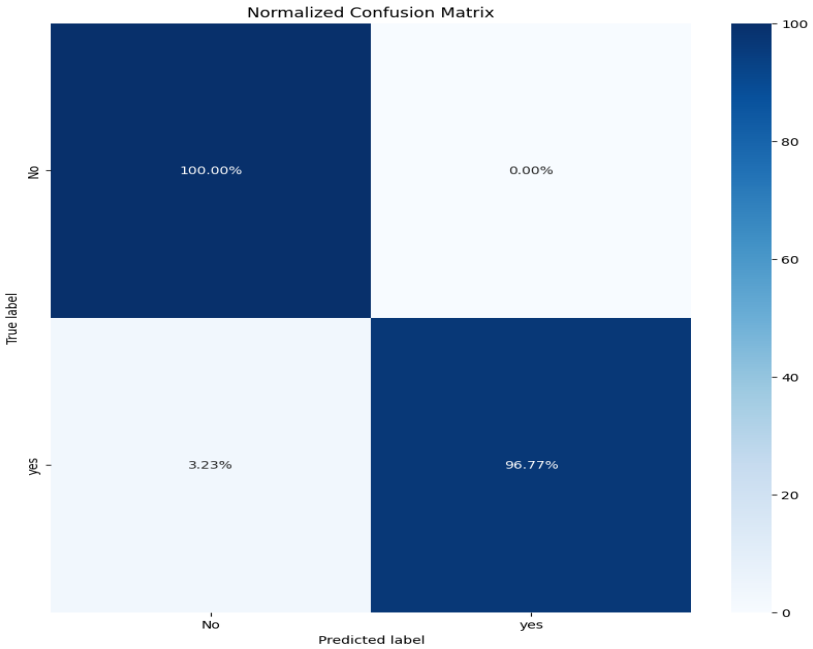


Figure 3: Confusion Matrix for Random Forest Classifier

Figure 3 is the confusion matrix for Random Forest Classifier after it completed its testing. The rows represent the true value of the data, while the columns show what the model predicted. This model did not predict any false positives, and predicted very few false negatives. This shows promise because the amount of times false negatives and false positives were predicted is very small. With further improvement, those values can become even smaller, leading to more accurate predictions of a patient's future.

VI. DISCUSSION

The results of our research shows that machine learning can effectively predict if someone is at high risk for developing complications due to type 1 diabetes or not. Out of the models we evaluated, Random Forest Classifier had the highest accuracy, with a training accuracy of 97.95% and a test accuracy of 98.00%. All of the other models also had an accuracy of 90% or higher. Random Forest Classifier, in particular, is effective at making predictions when given new data, especially when the relationship between the data is complex.

Referring back to Figure 2, different hyperparameters have significant effects on the accuracies achieved by different models. As seen in the graph for KNN, different k values have a significant effect on accuracy. Small K values show low accuracies, likely due to overfitting. As k increases, the values start to become more accurate. However, if k becomes too large, then accuracy becomes lower again, likely due to underfitting.

Random Forest is a model that uses two hyperparameters in order to achieve the best accuracy. Out of all the other models, changing its hyperparameters led to the most variability in accuracy. Different depths impacted the accuracies in very different ways. Lower depths tended to have lower accuracies, but both intermediate and higher depths corresponded to higher accuracies. There were no consistent trends among changing the number of trees used.

The model of Logistic Regression uses one hyperparameter, which is the maximum number of iterations. Lower max_iter values lead to more accurate predictions. Increasing the max_iteration did

not show improvement. In fact, increasing the number of maximum iterations leads to decreased accuracy. This shows that the model may have converged at lower max_iter values, and increasing this value does not necessarily improve performance.
The graphs for Multilayer Perceptron, like Random Forest Classifier, have two hyperparameters. Its hyperparameters are learning rates and epochs. Low learning rates tend to have poor accuracies, but accuracy improves as the learning rate becomes higher. Generally, increasing epochs increases accuracy, but it is important to find the best learning rate for the right epoch.

Each of the models have different aspects that they bring to the data. A model such as KNN compares surrounding data with a new input. Although the accuracy of KNN is high, it is the lowest among the models we used. It shows that a more complex approach might be needed in order to evaluate the data.

Multilayer Perceptron had a higher accuracy than KNN, showing that it is better suited to our purposes than KNN was. Our data has two distinct results that could appear based on what data is inputted.

Random Forest Classifier appears to be best suited to understand the nuances in the data. The process of understanding data is very thorough in Random Forest Classifier. Each tree used comes up with its own prediction, and all of these predictions are compiled together.

The results from using Logistic Regression suggest that our data is best suited to a nonlinear function, and that this model was effective at evaluating each piece of the data to form a comprehensive output.

Our research holds promise for the future, showing that artificial learning models can help with early detection of complications due to type 1 diabetes: a pressing concern for many. With the help of AI, these risks can be found earlier, and preventive measures can be taken. Healthcare providers can work proactively to ensure the best patient outcomes, with patients being alleviated of the stress that occurs when contemplating these complications.

A limitation of this research is that the dataset we used was very concentrated, with all of the data coming from one city. If all of the data is from the same place, then the algorithms may not pick up on the patterns present in other places. In our case, the data was very concentrated in Bangladesh, and our models were trained specifically to Bangladeshi culture and regional aspects. Although it could be helpful to have a more globally trained model, our model that is specialized to Bangladesh can be helpful when evaluating regional patterns, and when understanding the nuances of a specific place or culture.

In the future, we would like to retrain the models using datasets that encompass data from different demographics and regions. We could also use a hybrid of machine learning methods to further improve the efficiency of our models.

VII. CONCLUSION

In this study, we addressed the critical challenge of predicting severe complications in individuals with type 1 diabetes using machine learning models: K-Nearest Neighbors, Random Forest Classifier, Multilayer Perceptron, and Logistic Regression. Among these, the Random Forest Classifier achieved the highest accuracy at 98%, validated through cross-validation techniques, demonstrating robustness in predictive performance.This research contributes to the field by establishing a reliable predictive framework, potentially guiding proactive healthcare strategies for better patient outcomes.

The practical implications are significant, with our AI models offering the prospect of timely interventions in clinical settings to reduce complications and enhance patient quality of life. Future work will focus on incorporating diverse demographic datasets to improve model generalizability and exploring hybrid approaches to enhance predictive accuracy. These steps will refine our models and expand their applicability in medical AI, ensuring more personalized and effective care for type 1 diabetes patients.

REFERENCES

[1] Menke, Andy, et al. "The prevalence of type 1 diabetes in the United States." Epidemiology 24.5 (2013): 773-774.
[2] Atkinson, Mark A., George S. Eisenbarth, and Aaron W. Michels. "Type 1 diabetes." The lancet 383.9911 (2014): 69-82.
[3] Melendez-Ramirez, L. Yvonne, Robert J. Richards, and William T. Cefalu. "Complications of type 1 diabetes." Endocrinology and Metabolism Clinics 39.3 (2010): 625-640.
[4] Smith, Andrew, and Chelsea Harris. "Type 1 diabetes: management strategies." American family physician 98.3 (2018): 154-162.
[5] Kelly, Caitlin S., et al. "Daily experiences of type 1 diabetes stress across adulthood." Diabetic Medicine 39.1 (2022): e14628.
[6] Rechenberg, Kaitlyn, et al. "General and diabetes-specific stress in adolescents with type 1 diabetes." Diabetes research and clinical practice 130 (2017): 1-8.
[7] Skyler, J. S. "Prediction and prevention of type 1 diabetes: progress, problems, and prospects." Clinical Pharmacology & Therapeutics 81.5 (2007): 768-771.
[8] Venkatesh, Viswanath. "Adoption and use of AI tools: a research agenda grounded in UTAUT." Annals of operations research 308.1 (2022): 641-652.
[9] McCormick, Kyle. "AI Technology." From GI to Z: A Generational Guide to Technology (1985).
[10] Rong, Guoguang, et al. "Artificial intelligence in healthcare: review and prediction case studies." Engineering 6.3 (2020): 291-301.

[11] Knip, Mikael, et al. "Prediction of type 1 diabetes in the general population." Diabetes care 33.6 (2010): 1206-1212.
[12] Soedamah-Muthu, Sabita S., et al. "Predicting major outcomes in type 1 diabetes: a model development and validation study." Diabetologia 57.11 (2014): 2304-2314.
[13] Zhang, Zhongheng. "Parametric regression model for survival data: Weibull regression model as an example." Annals of translational medicine 4.24 (2016): 484.
[14] Freckmann, Guido, et al. "Measures of accuracy for continuous glucose monitoring and blood glucose monitoring devices." Journal of diabetes science and technology 13.3 (2019): 575-583.
[15] Doherty, Tara, et al. "Use of continuous glucose monitoring systems in pediatric patients in the perioperative environment: Challenges and machine learning opportunities." World Journal of Clinical Pediatrics 14.4 (2025).
[16] Annuzzi, Giovanni, et al. "Impact of nutritional factors in blood glucose prediction in type 1 diabetes through machine learning." IEEE Access 11 (2023): 17104-17115.
[17] Asaduzzaman S, Masud FA, Bhuiyan T, Ahmed K, Paul BK, Rahman SAMM. Dataset on significant risk factors for Type 1 Diabetes: A Bangladeshi perspective. Data Brief. 2018 Oct 9;21:700-708. doi: 10.1016/j.dib.2018.10.018. PMID: 30666315; PMCID: PMC6205358.
[18] Comment, et al. “K-Nearest Neighbor(KNN) Algorithm.” GeeksforGeeks, 23 Aug. 2025, www.geeksforgeeks.org/machine-learning/k-nearest-neighbours/.
[19] “Random Forest Classifier Using Scikit-Learn.” GeeksforGeeks, GeeksforGeeks, 13 Nov. 2025, www.geeksforgeeks.org/dsa/random-forest-classifier-using-scikit-learn/.
[20] Popescu, Marius-Constantin, et al. "Multilayer perceptron and neural networks." WSEAS Transactions on Circuits and Systems 8.7 (2009): 579-588.
[21] LaValley, Michael P. "Logistic regression." Circulation 117.18 (2008): 2395-2399.

★★★