J. Electrical Systems 20-5s (2024): 1801-1813

[1]K. Pushpavathi
[2]K.L. Shunmuganathan

Diabetic Drug ontology mapping for individual Diabetic Person and predict Insulin Dosage on Daily Basis

Abstract: -Managing diabetes requires a personalized treatment plan based on the patient's health records and drug usage effectively. In this paper, we have proposed accurate prediction of insulin dosage for diabetic patients daily by using ontology mapping with deep learning recurrent neural network models such as long short-term memory (LSTM), Gated recurrent unit(GRU), and Bidirectional long short-term memory(BiLSTM). The patient's information is arranged in a structured format. The structured ontology data provides a meaningful connection between patients' personal information, physical activity, insulin intake, carbohydrate intake and glucose level. The structured data is beneficial to take decisions effectively to recommend the insulin dosage level daily to avoid a sudden decrease in sugar level. The proposed LSTM model outperforms well on insulin prediction compared to GRU and BiLSTM models. The proposed model is very effective for diabetic personal care and improves patient outcomes and quality of life.

Keywords:Ontology Mapping, Recurrent neural network, LSTM, GRU, BiLSTM

I. INTRODUCTION

Diabetes has direct and indirect adverse effects on vascular complications. Major vascular complications are microvascular and macrovascular complications. Macrovascular complications occur due to large blood vessel damage. Microvascular complications occur due to small blood vessel damage. Macrovascular complications usually occur in Type 2 diabetes mellitus(T2DM). T2DM patients are affected by cardiovascular diseases such as heart attack, cardiomyopathy, arrhythmia, cerebra vascular disease, peripheral artery disease [1], myocardial infarction, stroke, and renal failure [2]. Microvascular [3] complications are diabetic retinopathy [4], diabetic neuropathy [4], and diabetic nephropathy [5]. Micro and macrovascular complications are prevented by maintaining good glycaemic control, controlling blood pressure, managing lipid profiles, and lifestyle modifications. High blood glucose levels for over 10 years cause micro and macrovascularcomplications[6]. Controlling high blood glucose levels by monitoring glucose levels regularly, following a healthy diet, exercising regularly, and taking medications as prescribed, such as insulin and oral hyperglycaemic agents [7].

Insulin plays an important role in controlling high blood glucose levels. There are different types of insulin, such as needle-free insulin injector (NFII), continuous insulin infusion (CSII), multiple subcutaneous insulin (MSI), and visfatin gene variation. NFII injects basal insulin to improve fasting glucose variation in T2DM patients [8]. CSII treatment reduces glucose levels faster with a lower dosage of insulin [9]. The vistagin gene modifies insulin creation, and conventional insulin reduces microvascular complications such as retinopathy, neuropathy, and dermopathy in insulin-dependent diabetes miletus(IDDM) [10]. Basal insulin dosage with conventional injection is 18IU, and NFII insulin dosage is 16IU. Adverse effects of insulin include insulinogenic edema, allergic reactions, lipo hypertrophy and lipoatrophy[12], ecchymosis[13], and hypoglycemia [14].

Ontology plays an important role in managing diabetes. An ontology represents the patient's records in the structured document, integrates the data, and supports integral decision-making. The researcher developed diabetes management pathway control (DPC)[15] to manage diabetes in China. The model contains 119 classes, 28 objects,58 data properties, and 81 individuals. The author designed a biocultural ontology for managing diabetes personally in American Indians [16]. American researchers constructed ontology for Native American Indians Nato [17] to manage diabetes individually for Americans. Data quality ontology is constructed for T2DM patients to acquire knowledge from electronic health records (EHR) [18]. The researcher developed an ontology-based support vector classifier model in India to manage diabetes [19]. Overall, ontology provides advantages such as knowledge representation of data in a common way, integration of different data, efficient decision-making, integration of natural languages effectively, and support of personalized diabetic care.

[1]Associate Professor, Department of Computer Science, School of Arts and Science, Vinayaka Mission's Chennai Campus, Vinayaka Mission's Research Foundation (DU), Paiyanoor, Tamil Nādu, India, Email: spushpavathi@rediffmail.com.
[2]Director (Academic) & Professor-CSE dept., Jaya Engineering College, Thiruninravur, Chennai, Email: director@jec.ac.in

Ontology-based dataset (OBD) combines data from multiple heterogeneous data. OBD integrates data between existing global schema and produces vocabulary for user needs [20]. OBD enables external database accessibility through a conceptual domain based on ontology. Conceptual domain-based database accessibility integrates the semantics of data [21]. Ontology-based semantic annotation enhances interoperability and understanding of data[22]. Ontology-based datasets organize the data and labels in a hierarchically structured data format. The data are arranged for specific domains. Creating an ontology-based dataset needs manual annotation for specific classes and attributes because of the variation of domains. The advantages of ontology-based datasets are (i) structured data arranges the data more precisely, complex analysis, and inference. (ii) Capture semantic information from classes, subclasses, and attributes easily. (iii). Easily represents unique characteristics of a specific domain. (iv). Bridges the gap between large datasets of specific domains such as audio, video, and image.

The insulin prediction system is an important factor in controlling hyperglycemia. The author designed a deep neural network model to predict blood glucose concentration after short-acting insulin injection instead of long-term blood glucose monitoring records [23]. The author developed an ensemble machine learning model to predict the inpatient insulin total daily dosage from electronic health records [24]. The author predicted insulin dosage for gestational diabetes women through a logistic regression model [25]. The author designed automated insulin delivery (AID) system to predict glucose variability in individuals using long short-term memory (LSTM) and an autoregressive model [26]. Using the Kalman filter, the author developed an onboard insulin model to predict insulin in real-time [27]. The proposed model is applicable under a variation of different conditions, such as stress and exercise. The author developed a 3D insulin model to predict patient-specific hyperglycaemic control [28]. The machine learning and deep learning automated insulin prediction system predicts insulin based on age, sex, BMI, and other glucose levels.

However, the insulin dosage does not predict accurately. To solve the above problem, in this paper, we have proposed ontology-based insulin predicted system to predict accurate insulin dosage levels daily. Ontology-based insulin prediction gives knowledge about insulin, including insulin type, administration, dosage, and blood glucose level. The type of insulin includes rapid, short, intermediate, and long-acting insulin. Insulin administration is done through subcutaneous injection, insulin pumps, and inhalation. Insulin dosage depends on weight, age, physical activity, carbohydrate intake, and insulin sensitivity. The author designed a diabetes prediction system using semantic web rule and decision tree algorithm to predict diabetes and insulin dosage level based on weight, age, BMI, carbohydrate intake, and glucose level [29]. The author constructed network ontology for diabetic patients in Mexico to derive new knowledge by applying semantic web rules. The ontology network contained several classes such as control plan, clinical Entity, Education level, clinical information and administration, geographic location, and persons [30].

Objective

1. Develop an ontology-based mapping structure to map the diabetic patients' personal details, daily activity, carbohydrate intake, insulin intake and glucose level.
2. Create an ontology-based dataset by extracting datas from ontology.
3. Develop on RNN models such as LSTM, GRU and BiLSTM model to predict insulin dosage of individual patient.
4. Compare prediction accuracy using mean square error(MSE), Root mean square error(RMSE), and mean absolute error(MAE) to predict insulin dosage on real time.

II. METHODOLOGY

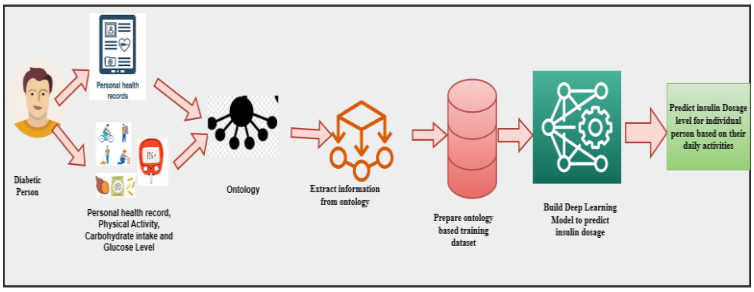


Fig. 1. Architecture of Ontology based insulin prediction System (OIPS)

Figure 1 depicts an ontology-based daily insulin dosage prediction system. The architecture captures each patient's health records and everyday activities, such as glucose levels, carbohydrate intake, and physical activity. The domain-based diabetes ontology is then built using personal health records and daily activities. The ontology is then used to extract training data. Finally, individual insulin dosage is learned using recurrent neural network models such as long short-term memory (lstm), bidirectional lstm, and gated recurrent unit (GRU). The best model is then chosen based on its performance metric for predicting insulin dosage level.

2.1 Creation of Ontology

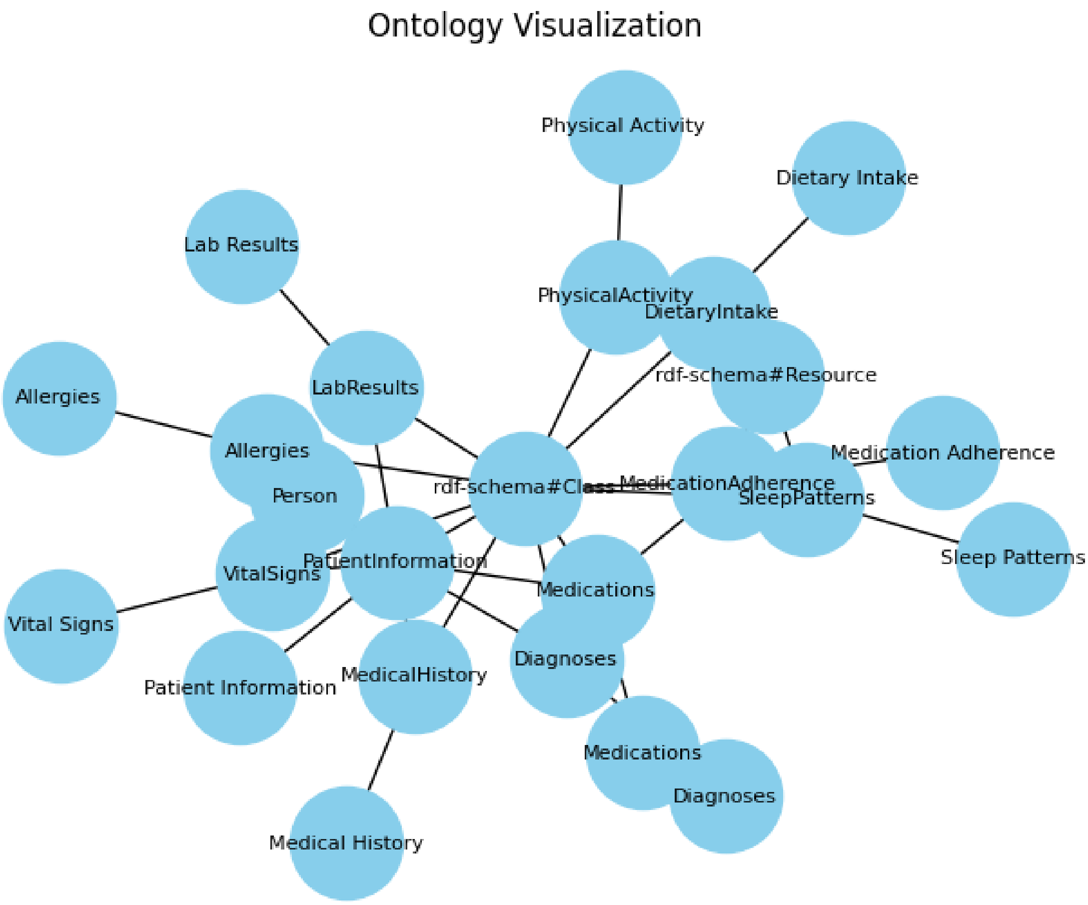


Fig. 2. Sample Ontology of Diabetic Patients Daily Activity and Patients Personal History

Diabetic patients' daily activity includes carbohydrate intake, physical activity such as yoga, walking, running, cycling and other exercises, medication, glucose level and their sleep patterns. We have created an ontology with patient's personal record and their daily activities. Figure 2 shows the sample ontology of Diabetes patients' daily activity and their personal health record. The ontology contains patient meta data such as name, age, date of birth, gender, contact information. Next, Medical History class includes allergies, surgery and illness. Medication class includes medication name, dosage and frequency of dosage usage. Allergies class includes any allergies or sensitive of patient. Lab results class stores various lab results of patient. Vital signs class includes heart rate, blood pressure and temperature. Medical diagnoses class includes symptoms, treatment plan. Daily activities class includes subclasses such as dietary intake, physical activity, sleep patterns and medication adherence. Dietary class includes food consumed, portion sizes and meal time. Sleep pattern includes information about sleep duration and quality. Physical activity class includes type of activity, duration and intensity.

2.2 Generation of ontology-based training samples

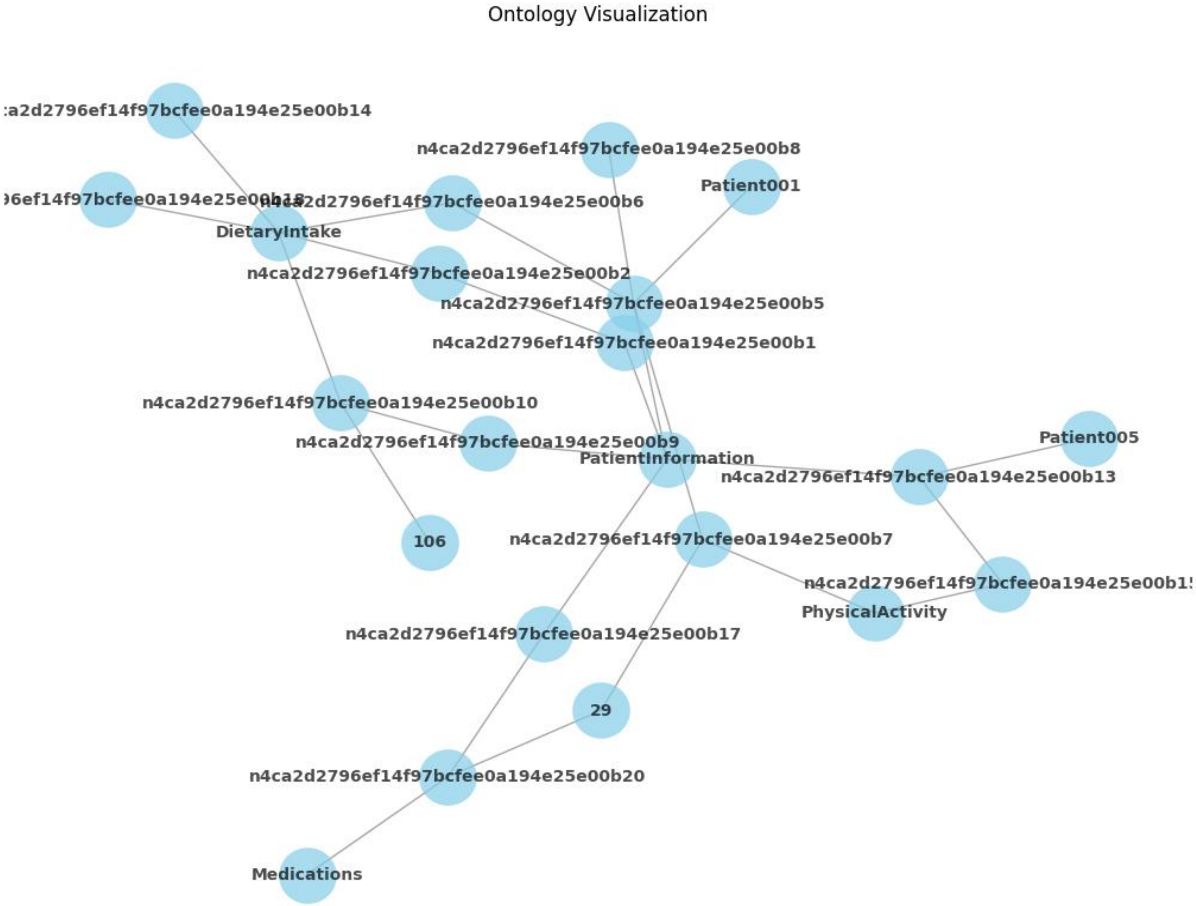


Fig. 3. Ontology based training data generation

The training data based on ontologies is shown in Figure 3. The semantic relationships of domain-specific ontologies are part of the training data based on ontologies; these relationships help to represent the complex and hierarchical nature of the domain knowledge. Machine learning models can more accurately and intelligently make decisions by incorporating ontology-based training data because it helps them understand the context and meaning of the data. Integrating training data from various sources yields thorough training data. The data can perform domain-specific tasks like predicting an individual patient's insulin dosage based on their daily activities and monitoring their blood sugar levels. With the help of this extensive training data, machine learning models can analyse a variety of variables and patterns to produce accurate predictions and suggestions for adjusting insulin dosage.

Additionally, by factoring in factors like lifestyle decisions, medication histories, and genetic predispositions, the integration of training data from various sources improves the accuracy and reliability of the predictions. Food intake, exercise, and patient data are shown in the graph above. The ontology graph stores the following information: patient ID, patient name, carbohydrate intake in grammes, minutes of physical activity, and insulin dosage level. Healthcare professionals can develop a thorough understanding of each patient's particular needs and recommend insulin dosages that are more specifically tailored by analysing data from various sources. Additionally, the ontology graph's inclusion of the patient ID and name ensures that the information is securely linked to each individual, making quick access and updates possible as needed. Interoperability, explicit knowledge, improved findability, accessibility, and reusability are all features of ontology-based training data that use controlled vocabularies.

2.3 Deep learning based Recurrent Neural Network Model

Long short-term memory is a type of recurrent neural network model to predict the value of long-term sequences. The LSTM model stores long term memory in sequence in the cell state. The cell state is placed at the top of the LSTM model. The LSTM Model contains an input gate, an output gate, and a forgotten gate. The input gate takes the previous output and current input as input and produces an output between 0 and 1. The input gate decides whether the information from the input should be stored in the cell state. The output gate takes the previous output and current input as inputs and produces an output between 0 and 1. It determines whether the information in the cell state should be used to compute the final output. The forgotten gate takes the previous output and current input as inputs and produces an output between 0 and 1. It controls how much of the previous cell state should be forgotten or retained in the current cell state. The forgotten gate selects the information that should be forgotten. The output gate selects the output from the cell state. The output gate takes the current input and previous output as inputs and produces an output between 0 and 1. It determines how much of the current cell state should be used to compute the final output. The output gate selects the relevant information to be passed on as the final output. The output is the final state; it selects the output from the cell state and output gate. Figure 4 shows the individual insulin prediction model using LSTM architecture.
The mathematical equations of LSTM Model is shown in equation (1),(2),(3),(4),(5),(6).

Input Gate (IG) = σ (Weight of input [Previous Hidden State, input Sequence] + bias of input (1)

Forget Gate (FG) = σ [Weight of forget [Previous Hidden State, input Sequence] + bias of forget (2)

Output Gate (OG) = σ [Weight of output gate [Hidden State, input Sequence] + bias of output (3)

Candidate Coll State (CCS)
= tanh (Weight of Candidate [Hidden State, Input Sequence] + bias of Candidate Cell state (4)

Updated Cell State (UPS) = FG+CCS+IP+~CCS (5)

Hidden State (HS)=(6+tanh(S) (6)

Fig. 4. Insulin Dosage Prediction Using LSTM Model

2.4 Gated Recurrent neural Network

Gated Recurrent neural network (GRU) is another RNN network. The GRU updates the network's hidden state selectively at each time and controls the information flow of the network. GRU-RNN contains a reset gate and an update gate. The reset gate measures the forgotten information of the previous hidden state, and the update gate updates the number of new input sequences. GRU consists of three states: Update Gate (z), Reset Gate and Current memory gate. Update Gate regulates the amount of past information passed into the future. Reset Gate fixes the amount of past information that needs to be forgotten. The current memory gate (CMG) introduces non-linearity in the input and the input to zero. CMG also acts as a subpart of the reset gate to reduce the effect on the current information passed into the future.GRU contains fewer parameters, is designed to capture relevant longer sequence information and removes past information because of its update and reset gate. GRU network is applicable to predict insulin dosage of Type1 and Type2 diabetes. Figure 5 shows the insulin dosage prediction system using GRU. The mathematical equations of GRU are shown in equations (7), (8),(9), and (10).

Update Gate (UG) = σ (Weight of update Gate * [Hidden State, input Sequence] + bias of input gate (7)

Reset Gate (RG)= σ (Weight of Reset Gate* [Hidden State, Input Sequence] + bias of Reset Gate (8)

Candidate Hidden State (CHS)
=tanh (Weight of hidden State * [Reset Gate * Hidden State, Input Sequence] +bias of hidden State (9)

Updated Hidden State (UNS) = (1- UG) * Hidden State + UG *~ (Hidden State) (10)

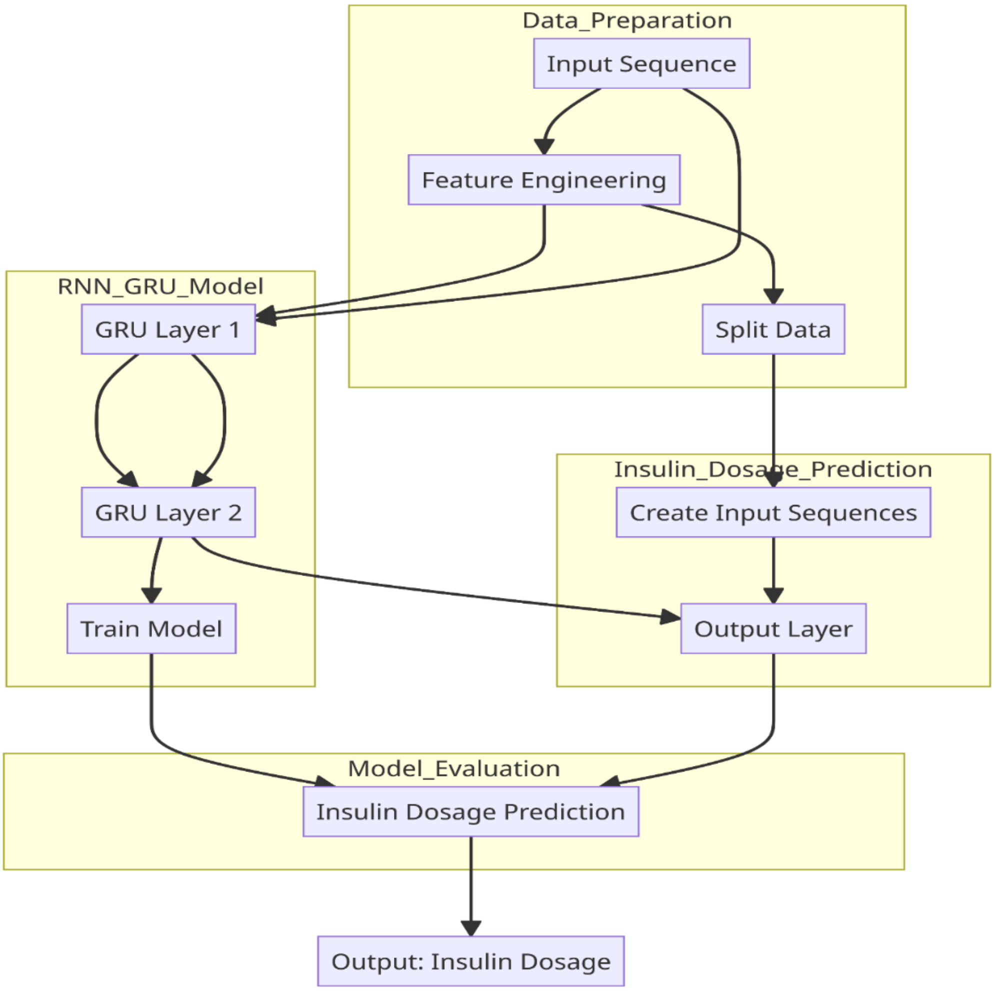


Fig. 5. Insulin dosage prediction Using Gated Recurrent Unit

2.5 Bidirectional LSTM

Bidirectional RNNs contain two independent RNNs. The network accepts both forward and backward information about the input sequence. BiLSTM handles the input in two ways: past to future and future to past. Hence, we can preserve the input information from past and future directions. The inputs depend on the amount of food consumed, physical activity done by the patient, either slow or fast, glucose level measured before and after the food and insulin taken before and after the food. Hence, the BiLSTM model considers past and future data to predict insulin dosage levels. BiLSTM model increases the amount of information in the network and improves the context of the input sequence to predict the output accurately. Figure 6 shows the insulin dosage prediction of the BiLSTM Model. The mathematical equations of the BiLSTM Model is shown in equation (11),(12),(13),(14),(15),(16),(17),(18),(19),(20),(21),(22),(23),(24).

Forward LSTM

Input Gate Forward (IGF)
= σ (Weight of input Forward = [Hidden State Forward, Input Sequence] + bias of input forward (11)

Forget Gate Forward (FGF)
= (Weight of forget forward *[Hidden State Forward, Input Sequence] +bias of forget forward (12)

Output Forward
= σ (Weight of output forward *[Hidden State Forward, Input Sequence] +bias of output gate forward (13)

~Candidate Forward
=tanh (Weight of Candidate Forward *[Hidden State forward, input Sequence] + bias of Candidate Forward (14)

Candidate Forward
= tanh (Weight of Candidate Forward *[Hidden State Forward, Input Sequence] + bias of Candidate (15)

Candidate Forward = forget gate Forward * Candidate Forward + input Gate Foreword + ~ (Candidate Forward) (16)

Hidden State Forward (HSF) = Output Forward * tanh (Candidate forward) (17)

Backward LSTM

Input Gate Backward (IGB)
=σ (Weight of input Backward * [Hidden State Backward, Input Sequence] +bias of input Backward (18)

Forget Gate Backward (FGB)
= (Weight of forget backward *[Hidden State Backward Input Sequence] + bias of forget Backward (19)

Output Backward
= σ (Weight of output Backward *[Hidden State Backward, Input Sequence] +
bias of Output gate Backward (20)

~Candidate Backward
= tanh (Weight of Candidate Backward *[Hidden State Backward, input Sequence]
+bias of Candidate Backward (21)

Candidate Backward
= tanh (Weight of Candidate Backward *[Hidden State Backward, Input Sequence] + bias of Candidate (22)

Candidate Backward
= forget gate Backward * Candidate Backward + input Gate Backward ~ (Candidate Backward) (23)

Hidden State Backward (HSR) = Output Backward * tanh (Candidate backward) (24)

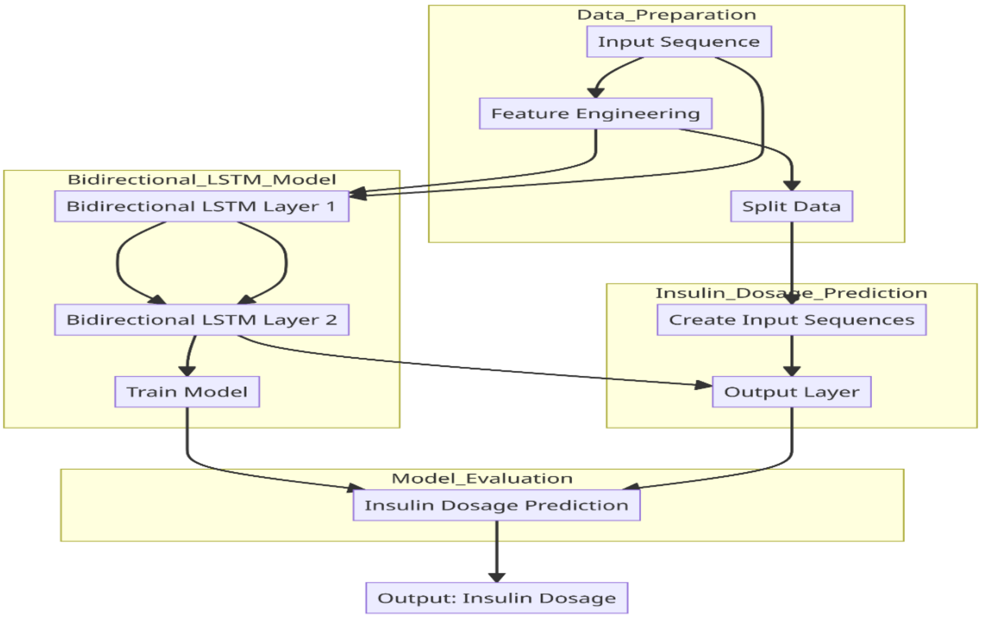


Fig. 6. Insulin dosage prediction system using BiLSTM Model

III. RESULTS AND DISCUSSIONS

3.1 Dataset Description

The proposed model's primary goal is to create a daily dosage insulin prediction system using RNN models such as the LSTM, GRU, and BiLSTM models. These models are known for capturing long-term dependencies in sequential data, making them suitable for predicting insulin dosage requirements. By training the model on historical patient data, it can learn patterns and trends to forecast the insulin dosage required for a given day accurately. Additionally, the proposed model will include features such as meal intake, exercise levels, and glucose readings to improve prediction accuracy. The dataset contains the diabetic's personal health record, physical activity, carbohydrate intake, glucose level, and insulin dosage. This extensive dataset enables the model to capture a comprehensive view of insulin requirements influenced by a person's health and lifestyle factors. By considering multiple variables, the model can provide personalised and precise insulin dosage recommendations, improving diabetes management for each patient. The information is organised in an ontological dataset format. The ontological dataset format structures the data, allowing for efficient analysis and interpretation. This format ensures that all relevant information about the person's health is accurately captured, including their health record, physical activity, carbohydrate intake, glucose level, and insulin dosage.

Furthermore, the ontological dataset format allows for seamless integration with other healthcare systems and accessible information sharing between healthcare providers for collaborative decision-making. The ontological dataset contains semantic relationships between diabetic patients, and the data is subject to the domain. Ontology-based diabetic dataset enables healthcare providers to analyse data more effectively and make informed decisions based on the unique needs of each diabetic patient. Furthermore, the ontological dataset format ensures that the data is structured, making it easier to retrieve and interpret relevant information for research or developing personalised treatment plans.

3.2 Architecture of RNN Models

3.2.1 Long and short-term memory

In this model, we used 100 cells for the encoder and decoder with one dense layer. The encoder and decoder are crucial components in the model as they handle the task of encoding and decoding information. Using 100 cells allows a sufficient capacity to store and process data efficiently. Additionally, including one dense layer helps enhance the model's ability to learn complex patterns and relationships within the data. The proposed model was trained for 100 epochs with a batch size of 32, and the Adam optimizer was used to optimize the model's performance. The Adam optimizer is beneficial as it combines the advantages of both AdaGrad and RMSProp algorithms, allowing for faster convergence and better generalization.

3.2.2 GRU

In this model, we used 100 cells for the encoder and decoder, with one dense layer. The proposed model was trained with 100 epochs with a batch size of 40, and the Adam optimizer was used to optimize the model's performance. The Adam optimizer is beneficial as it combines the advantages of both AdaGrad and RMSProp algorithms, allowing for faster convergence and better generalization. Additionally, a learning rate of 0.001 was used with the loss function set as MSE to further enhance the model's training process. And the Adam optimizer was used with a learning rate of 0.01 and a loss function of MSE.

3.2.3 Bi-LSTM Model

In this model, we used 100 cells for the encoder and decoder, with one dense layer. The Bi-LSTM model was chosen for its ability to capture both forward and backward dependencies in the input sequence. It is beneficial for tasks such as natural language processing. The use of 100 cells allows for a sufficient amount of memory and complexity in the model while including a dense layer helps further refine the output predictions. The proposed model was trained with 100 epochs with a batch size of 32, and the Adam optimizer was used with a learning rate of 0.001 and a loss function of MSE.

3.3 Statistical Measure

The performance of RNN models such as LSTM, Bi-LSTM and GRU model is predicted with the following parameters, they are shown in equation (25), (26) ,(27).

3.3.1 Mean Square Error (MSE)

MSE is used to measure the accuracy of predicted model by measuring the average squared difference between predicted and actual insulin dosage.MSE should be lower one.

MSE= (1/number of diabetic persons) pow (Σ (actual insulin dosage - predicted insulin dosage)) (25)

3.3.2 Root Mean Square Error (RMSE)

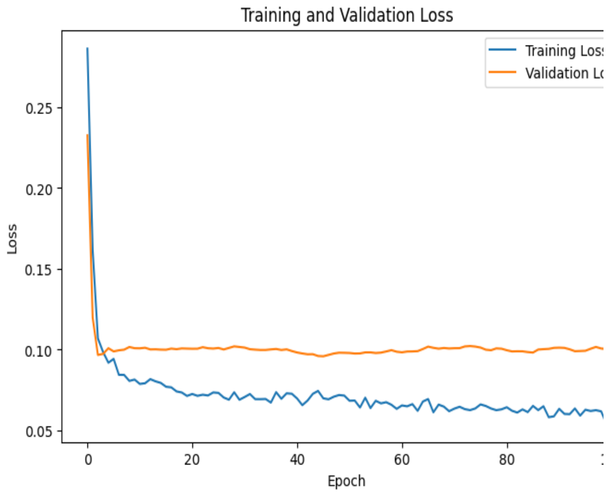
RMSE is measured by taking the square root of MSE.RMSE should be lower one.

$$RMSE=\sqrt{MSE} \tag{26}$$

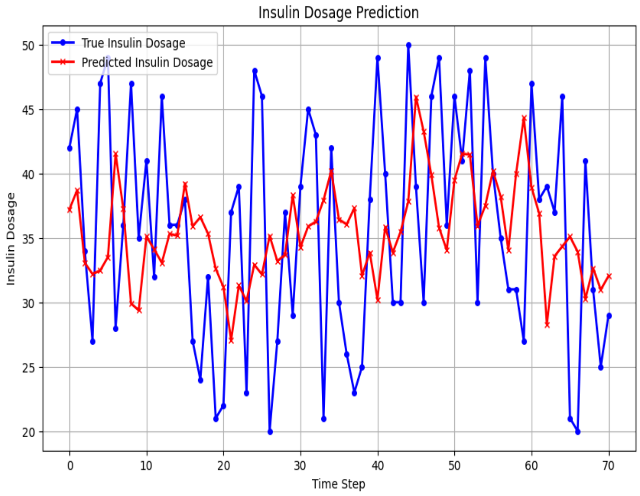
3.3.3 Mean Absolute Error (MAE)

MAE is measured the difference between the actual insulin dosage with the predicted insulin dosage. The MAE value should be lower, then only we can decide our model predicts insulin dosage accurately for diabetes patient correctly.

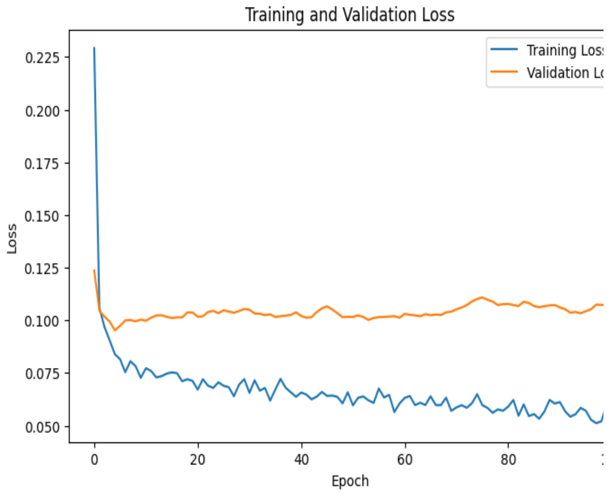
MAE =1/number of Diabetic patient Σ │actual insulin dosage - predicted insulin dosage│ (27)



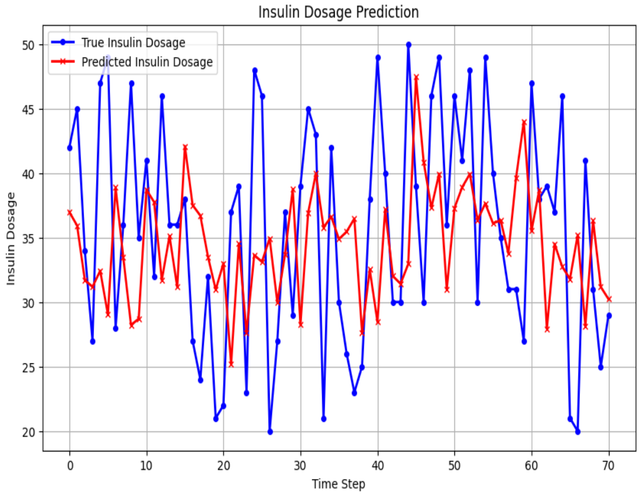
(i). Training and Validation Loss of LSTM



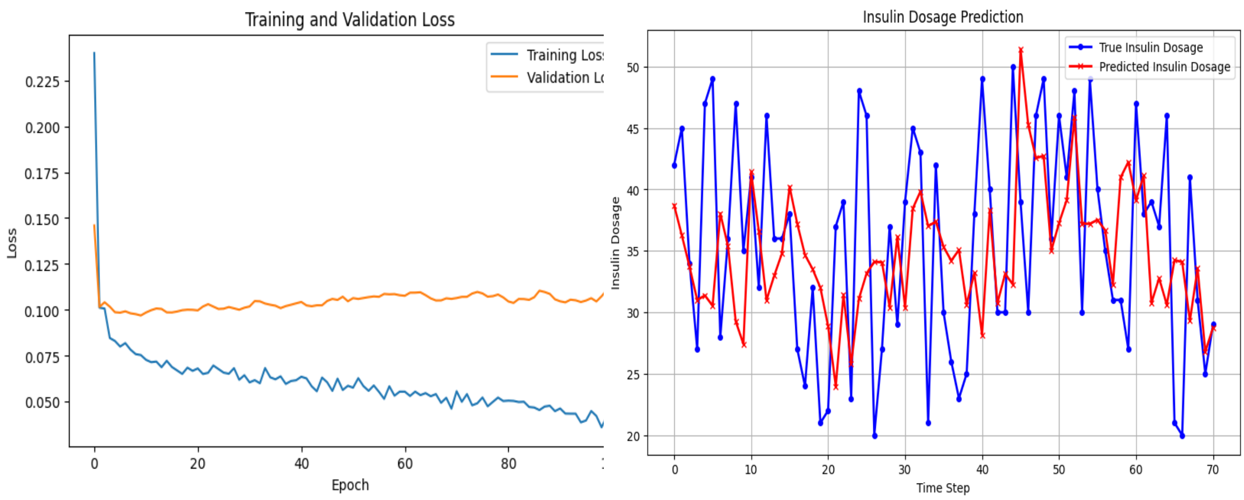
(ii). Insulin Dosage prediction in LSTM



(i). Training and Validation Loss of GRU



(ii). Insulin Dosage Prediction in GRU



(i). Training and Validation Loss of BiLSTM (ii). Insulin dosage Prediction in BiLSTM

Fig. 7: Insulin Dosage Prediction Using RNN Models

Figure 7 shows the training and validation loss curve and insulin dosage prediction curve using the RNN model. The training and validation loss curve in Figure 7 demonstrates the performance of the RNN model during the training process. It provides insights into how well the model learns from the data and whether it is overfitting or underfitting. The insulin dosage prediction curve also showcases how accurately the RNN model can predict insulin dosages based on the given input data. The training loss curve shows the value of the loss during training. The training loss decreases continuously, indicating that the model correctly fits the data. In Figure 7, all LSTM, GRU, and BiLSTM models provide the data accurately. There is no overfitting of diabetic patients' data. The validation loss occurs due to the fitting of new or unknown data. The validation loss also decreases continuously for each training epoch. We have trained our models for 100 generations with a batch size 32 and a dropout of 0.5. The proposed model leaves a small gap between the training and validation curves. The curves indicate that the model generalizes well to unseen data and is not overly reliant on the training data. The small gap between the training and validation curves suggests that the model is balanced with the training data, as both curves consistently decrease. Overall, these results demonstrate the effectiveness of our proposed model in accurately predicting outcomes for diabetic patients. Hence, the model fits the data accurately.

Table 1:Performance of Insulin dosage prediction of RNN Models

Models	**MSE**	**MAE**	**RMSE**
LSTM	86.35	7.8	9.3
GRU	91.79	8.102	9.6
BiLSTM	88.37	7.65	9.4

Table 1 shows the performance of the insulin dosage prediction system of the RNN Model. The MSE is the average error between the actual and predicted values. The higher MSE value means a more significant error between the true and predicted insulin dosage. Here, LSTM produces a lower error (86.35) than the other two models. Hence LSTM expects insulin dosage correctly. Similarly, MAE and RMSE are also lower in LSTM compared to the other two models. Hence, LSTM predicts the insulin dosage correctly compared to the other two models. The GRU model performs slightly worse than the LSTM and BiLSTM models. However, it is essential to note that the difference in performance between the LSTM, BiLSTM, and GRU models is relatively small. While the LSTM and BiLSTM models have lower error metrics, the GRU model still provides reasonably accurate predictions for insulin dosage. Therefore, although more accurate than the LSTM and BiLSTM models, the GRU model can still be considered a viable option for predicting insulin dosage. The BiLSTM model falls between the LSTM and GRU models. Hence The BiLSTM model predicts the insulin dosage moderately. Table 2 shows the prediction of insulin dosage for various levels of basal insulin dosage predicted with high accuracy; error values are comparatively low; the model [31] considers only glycaemic control of sensor data of Type 1 diabetes. The second model [32] considers only the data with 30 minutes interval period. The proposed model

contains the daily dosage prediction based on their physical activity, carbohydrate intake and glucose level. Hence, the error is increased compared to the other two models.

Table 2: Insulin Dosage prediction with other models

Type of Diabetes	Method	RMSE	MAE	MSE
Type 1 Diabetes Age, Sex,weight, height, body mass index, basal insulin dosage [31]	Support Vector machine +glycaemic control with sensor data in between(6 to 8)	0.033	0.001	0.0128
Type1 Diabetes Continuous Glucose monitoring data with 30 minutes interval [32]	LSTM,BiLSTM,Temporal Convolutional Network	34.7	-	-
Proposed Type1 and Type2 Diabetes with Age, sex, Male, carbohydrate intake, Glucose Level and physical activity	LSTM, BiLSTM , GRU	86.35	7.8	9.3

IV. CONCLUSION

In this study, we proposed diabetic drug ontology mapping and insulin dosage prediction for diabetic patients daily. The ontology mapping dataset analyses drug information based on patient data, physical activity, glucose level, insulin intake, and carbohydrate level. Hence, the proposed methodology improves treatment plans and drug management. By utilizing diabetic drug ontology mapping, healthcare professionals can accurately identify the most suitable medications for patients based on their specific needs and conditions. Additionally, the insulin dosage prediction feature assists in determining the optimal amount of insulin required for each patient, ensuring precise and personalized treatment. **Integrating drug ontology mapping with the RNN model accurately predicts the insulin dosage level of individual patients. Among the RNN models, the LSTM model outperforms well in daily insulin dosage predictions. Healthcare providers offer personalized care to diabetic patients and timely insulin dosage recommendations by combining drug ontology mapping and predictive modelling. In the future, we will include more patient data to improve patient outcomes in insulin prediction** and further enhance the accuracy of our predictive models. Additionally, incorporating real-time monitoring devices and patient feedback will allow for more dynamic and precise insulin dosage recommendations, ultimately leading to better management of diabetes and improved patient quality of life.

REFERENCE

[1] Viigimaa M, Sachinidis A, Toumpourleka M, Koutsampasopoulos K, Alliksoo S, Titma T. Macrovascular Complications of Type 2 Diabetes Mellitus. *CurrVascPharmacol*. 2020;18(2):110-116. doi:10.2174/1570161117666190405165151.

[2] Ahmad MS, Alslamah T, Alannaz SM, et al. Prevalence of micro and macro vascular complications and their risk factors in type 2 diabetes in Saudi Arabian population: an analysis from SHIS. *Eur Rev Med Pharmacol Sci*. 2021;25(12):4308-4316. doi:10.26355/eurrev_202106_26137

[3] Vora, A.R. (2021). Microvascular complications of diabetes mellitus. *Indian Journal of Applied Basic Medical Sciences*.

[4] Reichard P, Nilsson BY, Rosenqvist U. The effect of long-term intensified insulin treatment on the development of microvascular complications of diabetes mellitus. *N Engl J Med*. 1993;329(5):304-309. doi:10.1056/NEJM199307293290502

[5] Cheema S, Maisonneuve P, Zirie M, Jayyousi A, Alrouh H, Abraham A, Al-Samraye S, Mahfoud Z, Al-Janahi IM, Ibrahim B, Lowenfels AB, Mamtani R. Risk Factors for Microvascular Complications of Diabetes in a High-Risk Middle East Population. J Diabetes Res. 2018 Jul 2;2018:8964027. doi: 10.1155/2018/8964027. PMID: 30057913; PMCID: PMC6051261.

[6] Hanssen KF. Blood glucose control and microvascular and macrovascular complications in diabetes. *Diabetes*. 1997;46Suppl 2:S101-S103. doi:10.2337/diab.46.2.s101.

[7]. Imran, M., Begum, S., Kandhro, A.H., Ahmed, N., &Qasim, R. (2017). The Management of Glycemic Control in Associated Disorders.

[8]. Sun F, Gao B, Yang A, et al. Needle-free injection of basal insulin improves fasting glucose variability as assessed by continuous glucose monitoring in T2DM: a prospective randomized multicenter open-label crossover study. *Expert Opin Drug Deliv*. 2022;19(12):1725-1734. doi:10.1080/17425247.2022.2147504

[9] Jiang, Li-yan and Li-li Zhenc. "Clinical efficacy observation of insulin pump therapy on newly diagnosed type 2 diabetes." *Central Plains Medical Journal* 39 (2012): 5-6.

[10] Reichard P, Nilsson BY, Rosenqvist U. The effect of long-term intensified insulin treatment on the development of microvascular complications of diabetes mellitus. *N Engl J Med*. 1993;329(5):304-309. doi:10.1056/NEJM199307293290502

[11]. Guarneri AM, Hoffman RP. Non-glycemic Adverse Effects of Insulin. *Curr Diabetes Rev*. 2022;18(2):e012821190877. doi:10.2174/1573399817666210129104420

[12]. Oliverio, S., Chiarabolli, S., &Tignani, R. (2007). Adverse Effects of Insulin Detemir.

[13]. Kondo T. *Nihon Rinsho*. 2012;70Suppl 6:196-199.

[14]. Ukpds. "Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes." (2002).

[15]. Fan Z, Cui L, Ye Y, Li S, Deng N. Telehealth System Based on the Ontology Design of a Diabetes Management Pathway Model in China: Development and Usability Study. JMIR Med Inform. 2022 Dec 19;10(12):e42664. doi: 10.2196/42664. PMID: 36534448; PMCID: PMC9808585.

[16]. Li, Juan and ShadiAlian. "Design and Development of a Biocultural Ontology for Personalized Diabetes Self-Management of American Indians." *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)* (2018): 1-7.

[17]. Pandey, Vikram et al. "Evaluation and Evolution of NAOnto – An Ontology for Personalized Diabetes Management for Native Americans." *2021 7th International Conference on Computer and Communications (ICCC)* (2021): 1635-1641.

[18]. Rahimi, Alireza et al. "Development of a Methodological Approach for Data Quality Ontology in Diabetes Management." *Int. J. E Health Medical Commun.* 5 (2014): 58-77.

[19]. Muthuraman, Thangaraj and GnanambalSankaran. "Rule Based Exception Handling and Priority Modeling For Diabetes Management Using Ontology and SWRL." (2014).

[20]. Kontchakov, R., Rodríguez-Muro, M., Zakharyaschev, M. (2013). Ontology-Based Data Access with Databases: A Short Course. In: Rudolph, S., Gottlob, G., Horrocks, I., van Harmelen, F. (eds) Reasoning Web. Semantic Technologies for Intelligent Data Access.Reasoning Web 2013. Lecture Notes in Computer Science, vol 8067. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39784-4_5

[21]. TadeuszPankowski, Modeling and Querying Data in an Ontology-Based Data Access System,Procedia Computer Science,Volume 192,2021,Pages 497-506,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2021.08.051.

[22]. Kostovska, A., Džeroski, S., Panov, P. (2020). Semantic Description of Data Mining Datasets: An Ontology-Based Annotation Schema. In: Appice, A., Tsoumakas, G., Manolopoulos, Y., Matwin, S. (eds) Discovery Science. DS 2020.Lecture Notes in Computer Science(), vol 12323. Springer, Cham.https://doi.org/10.1007/978-3-030-61527-7_10

[23] Tang B, Yuan Y, Yang J, Qiu L, Zhang S, Shi J. Predicting Blood Glucose Concentration after Short-Acting Insulin Injection Using Discontinuous Injection Records. Sensors (Basel). 2022 Nov 3;22(21):8454. doi: 10.3390/s22218454. PMID: 36366151; PMCID: PMC9653564.

[24] Nguyen M, Jankovic I, Kalesinskas L, Baiocchi M, Chen JH. Machine learning for initial insulin estimation in hospitalized patients. J Am Med Inform Assoc. 2021 Sep 18;28(10):2212-2219. doi: 10.1093/jamia/ocab099. PMID: 34279615; PMCID: PMC8449602.

[25] Eleftheriades, M., Chatzakis, C., Papachatzopoulou, E. *et al.* Prediction of insulin treatment in women with gestational diabetes mellitus. *Nutr.Diabetes* **11**, 30 (2021).https://doi.org/10.1038/s41387-021-00173-0

[26] Zafar A, Lewis DM, Shahid A. Long-Term Glucose Forecasting for Open-Source Automated Insulin Delivery Systems: A Machine Learning Study with Real-World Variability Analysis. *Healthcare*. 2023; 11(6):779. https://doi.org/10.3390/healthcare11060779

[27] Eleonora M. Aiello, Kelilah L. Wolkowicz, Jordan E. Pinsker, EyalDassau, Francis J. Doyle III,A novel model-based estimator for real-time prediction of insulin-on-board,Chemical Engineering Science,Volume 267,2023,118321,ISSN 0009-2509,https://doi.org/10.1016/j.ces.2022.118321.

[28]. Vincent Uyttendaele, Jennifer L. Knopp, Kent W. Stewart, Thomas Desaive, BalázsBenyó, NoémiSzabó-Némedi, Attila Illyés, Geoffrey M. Shaw, J. Geoffrey Chase, A 3D insulin sensitivity prediction model enables more patient-specific prediction and model-based glycaemiccontrol,Biomedical Signal Processing and Control,Volume 46,2018,Pages 192-200,ISSN 1746-8094,https://doi.org/10.1016/j.bspc.2018.05.032

[29]. El Massari, H., Mhammedi, S., Sabouri, Z., Gherabi, N. (2022). Ontology-Based Machine Learning to Predict Diabetes Patients. In: Maleh, Y., Alazab, M., Gherabi, N., Tawalbeh, L., Abd El-Latif, A.A. (eds) Advances in Information, Communication and Cybersecurity. ICI2C 2021. Lecture Notes in Networks and Systems, vol 357. Springer, Cham.https://doi.org/10.1007/978-3-030-91738-8_40.

[30]. Reyes-Peña C, Tovar M, Bravo M, Motz R. An ontology network for Diabetes Mellitus in Mexico.J Biomed Semantics. 2021 Oct 9;12(1):19. doi: 10.1186/s13326-021-00252-2. PMID: 34625104; PMCID: PMC8500829.

[31]. Guzman Gómez GE, BurbanoAgredo LE, Martínez V, BedoyaLeiva OF. Application of Artificial Intelligence Techniques for the Estimation of Basal Insulin in Patients with Type I Diabetes. Int J Endocrinol. 2020 Nov 3;2020:7326073. doi: 10.1155/2020/7326073. PMID: 33204261; PMCID: PMC7655245.

[32]. Bhimireddy, A., Sinha, P., Oluwalade, B., Gichoya, J. W., &Purkayastha, S. (2020, August). Blood glucose level prediction as time-series modeling using sequence-to-sequence neural networks.In CEUR workshop proceedings.