A PERSONALIZED DEEP LEARNING APPROACH FOR BLOOD GLUCOSE PREDICTION IN PEOPLE WITH T1DM

By

Ayesha Parveen

A THESIS

Submitted to the Faculty of the Stevens Institute of Technology
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE – COMPUTER SCIENCE

_____

Ayesha Parveen, Candidate

ADVISORY COMMITTEE

_____

Dr Samantha Kleinberg, Advisor Date

_____

Dr Yue Ning, Reader Date

STEVENS INSTITUTE OF TECHNOLOGY
Castle Point on Hudson
Hoboken, NJ 07030
2021

A PERSONALIZED DEEP LEARNING APPROACH FOR BLOOD GLUCOSE PREDICTION IN PEOPLE WITH T1DM

ABSTRACT:

Managing a chronic disease like Type 1 diabetes (T1D) is both tough and time-consuming. Although new technologies that allow continuous measurement of glucose and delivery of insulin have led to significant improvements, it is widely known that achieving a tight glycaemic control can be a very complex process for some individuals as there are wide variations in their blood glucose signals for various factors like physical activity, age, diet, or their stress levels. Due to these factors which are not easily identifiable and can vary from person to person, having a personalized model for BG prediction is important. Personalized models can capture the factors of a particular patient's lifestyle and utilize that to make a forecast of their future BG.

In this work, we propose a new deep learning method for personalized BG prediction based on an LSTM model. Unlike vanilla LSTM models, this model uses the method of incremental learning to constantly update itself as and when the patient data is available. This method would help us achieve better-personalized predictions over time for patients. The obtained results show a fairly good prediction accuracy. For a prediction horizon of 30 minutes, we get the best average RMSE = 19.152 mg/dl. The model can achieve a reliable glucose forecasting performance. In the future, we plan to include other features like meal, activity, stress to get better personalized BG prediction for patients.

Author: Ayesha Parveen
Advisor: Samantha Kleinberg

Date: December 8, 2021
Department: Computer Science
Degree: Master of Science – Computer Science

Dedication:

To my family, who always supported my dreams…….

Acknowledgements

I would like to express my deepest gratitude to my advisor Dr Kleinberg for her support and constant guidance, my thesis reader, Dr Ning, for her cooperation and valuable suggestions, and my colleagues from HAIL who have been the best of friends and the most fun to work with.

Table of Contents:

Chapter 1:

Introduction

Diabetes is a chronic disease that occurs when either the pancreas is not able to produce enough insulin or when the body cannot effectively use what it produces. In 2014 around 8.5% of the adults older than 18 years of age had diabetes and in 2021 diabetes was the direct cause of death for almost 6.7 million people in the world [38]. Currently, over 573 million people all over the world have diabetes with approx. 34.2 million people in the United States have diabetes while 88 million are pre-diabetic [14]. Although the majority of people suffer from Type 2 diabetes, the rate of people with Type 1 diabetes is also rising. T1 diabetes, also known as juvenile diabetes, is a lifelong chronic disease where the pancreas produces little or no insulin. So, people with T1D need to rely on external insulin delivery to the body to maintain their blood glucose levels [15]. And this is a very tedious and difficult task because for each meal and activity the individual must calculate the right amount of insulin dosage. While calculating the insulin they must consider external factors like stress that affect insulin, because getting the correct amount of dosage is especially important to help maintain glucose levels within a healthy range in the body, as that would prevent many severe secondary complications of diabetes like chronic kidney disease [16] and heart disease and stroke [17][41].

Due to all the challenges for T1D, a lot of modern technologies have been developed like continuous glucose monitors which measure CGM (Continuous Glucose Monitor) every 5 minutes, unlike fingersticks that are regularly used to measure blood glucose before and after a meal [40]. An artificial pancreas is a diabetes management system that helps to track the BG levels using a CGM or continuous glucose monitor and then algorithmically decides the insulin dosage after which delivers the instructions to the insulin pump to deliver the insulin [39]. So, the artificial pancreas system will not only help monitor the glucose levels in the body, but it also helps to involuntarily adjust the insulin delivery to help reduce high glucose levels and minimize any incidence of low glucose levels [42]. But the glucose forecasting needs to be done accurately and precisely so that the adjustments are done correctly to be able to reach the correct BG range. And this accurate forecast is a big challenge since the CGMs (Continuous Glucose Monitor) do not measure the blood

glucose but the glucose from the fluids in the cells, which impedes the process of BG measurement. Furthermore, there can be errors and noise from factors that do not directly affect blood glucose.

With the rapid development in the field of machine learning, it can be used for BG forecasting in the AP (Artificial Pancreas) system. Earlier studies have shown that simple auto-regressive models can be used to predict BG 30 minutes in advance with a root mean square value of 27.7 mg/dl [1]. Now, with the recent developments of deep learning models RNN [2], CNN [6] and LSTM [3] have been used as well with RMSE values as 21.07, 21.72, and 20.80, respectively. However, these studies only considered data that was collected over a brief period with very few patients. So, how these models would handle data collected in real life with there are a lot of errors and noises that are not known (for example external factors like being stressed out for an exam or doing a rigorous workout). But now there is a large-scale-patient generate heath data that can be used to test the ML models for BG forecasting. One such project is the Open Artificial Pancreas, led by people with T1D who wanted to build a system that could cater to their needs, that uses the overnight closed-loop APS technology. As, of July 19, 2021, there are more than 2344 people around the world that are using the OpenAPS, and a subset of these users have made their data available for research. In this work, we focus on the task of forecasting BG in real-world settings by leveraging the data made available from OhioT1DM. We present a BG level prediction method that uses the concept of incremental learning or learning as data is made available. We aim to make a model that will help to personalize the BG prediction particular to the specific patients by having the model learn from the data as in when it is available over time and update the features of the model, instead of just one time in the beginning when it is created. This way, we aim to improve our accuracy over some time.

Chapter 2:

Background

2.1 Time-series analysis:

Time series can be defined as a set of observations on the values that a variable takes over a certain period of time i.e., $X_T = \{x_1, x_2 \ldots\ldots x_T\}$ where T is the length of the time period. For example, daily whether temperatures recorded for 24 hours. Now the values of the variable can be recorded continuously over time called 'Continuous-time series' or it can be recorded at a discrete set of time instances called 'Discrete Time series'. So, developing models for accurate forecasting is an active area of research in fields like stock market [4], forecasting electricity loads [5], estimating glucose values for effective diabetes management [18] and many more applications.

A stationary time series property does not depend on the time at which the series is observed. So, time series with trends and seasonality are not stationary, as these properties would affect the value of the variable at different times. Whereas white noise series is a stationary series as the values would look the same at any point in time. Thus in general, stationary time series does not have any predictable patterns in long term and the time plots show the series to be roughly horizontal with constant variance. Furthermore, a time series can be unimodal or multimodal that is data is either obtained from a single source or n different sources.

Time series forecasting happens when we make scientific predictions based on historical time-stamped data. So, we build models through historical analysis and use them to make observations and derive a future strategic decision-making model. One of the distinctive features of such forecasting is that at the time of predicting the future outcome is completely unavailable and it can only be estimated through careful analysis and evidence-based priors. Thus, time series forecasting is not always an exact prediction, and the likelihood of the forecasting can vary extensively, especially if there are fluctuating variables or factors that are outside of our control. Therefore, the more comprehensive data we have the more accurately we will be able to forecast. Time series analysis usually curtails developing models that can help gain an understanding of the data and its underlying causes. So, the analysis can provide us answers for the reasons "why" we are seeing the outcomes. Time series forecasting has various applications in different industries like weather forecasting, economic forecasting, healthcare forecasting, retail forecasting, financial forecasting, business

forecasting, social studies forecasting and many more. So, any industry practitioner that uses consistent historical data can analyse the data with time series analysis methods and then build a model that can forecast and predict required outcomes.

Two types of forecasting can be done:

- Ex-ante forecast – is the forecast based on the information available at the time of forecast. So, prediction is done based on the data that is known in advance and then forecast is done for the data beyond the point known. One example is forecasting percentage change in the US fuel consumption for the next quarter based on the previous quarters for a year.
- Ex-post forecast – is the forecast that uses information beyond the time at which the forecast is made. For example, ex-post forecasts of consumption might end up using actual observations of the predictors once they have been observed. These forecasts might not seem authentic, but they are especially useful in studying the behaviour of the forecasting models.

Time series forecasting can be handled recursively, where a single time series model is estimated, and each forecast is computed by using previous forecast outcomes. However, direct forecasting can also be done where a separate time series model is estimated for each forecasting horizon and the forecasts are computed only on the observed data. Secondly, the output can be a single future value, or it can be a multi-step output where all the values are estimated. This is described further in chapter 3. Next, we describe the techniques used for time series forecasting.

2.2 Deep Learning Models

When we have sequential or time-series data, typical feed-forward networks that take in a sequence of inputs and map them to an output space using a combination of linear operations and non-linear activations functions cannot be used for learning and making predictions. We need a mechanism that can retain past /historic information and use that to make forecasts for the future [36]. Recurrent neural networks (RNNs) are variants of conventional feed-forward neural networks that can deal with sequential data and can be trained to hold temporal information. Although RNNs can handle sequential data and store it to ‘memorise’ historical information, it has a vanishing gradient problem that prevents the model network from capturing

long-term dependencies in the sequential data. This problem is alleviated by an advanced configuration of RNN called Long-Short Term Memory [2].

Long short-term memory (LSTM):

Long-short term memory networks are a type of RNN that are capable of learning order dependence in sequence prediction problems. The long short-term memory architecture is designed to overcome the existing flaw in RNNs i.e., the problem of vanishing gradients during backpropagation that prevents from capturing the long-term dependencies. LSTMs use three gates namely input, output and forget gate. These three gates have values ranging from 0 to 1, and they are used to update the cell and the hidden states for each time step. So, they can control how much information from the current input should be contributing to the computing of the hidden states and the output. [53]

The LSTM model passes two kinds of states across time steps -- cell state and hidden state. The cell states decide which information should be carried forward from different observations that the network is trained on [52].

One of the first things that happen within an LSTM is the activation function of the forget gate, given by equation 2.1. The activation function takes in the inputs of h_{t-1} and x_t and then outputs a value between 0 to 1 to every number in the cell state C_{t-1} from the previous layer. So, a value of 0 would mean completely forgetting the seen data seen and a value of 1 would mean keeping the seen data.

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f) \quad (2.1)$$

The following steps in the LSTM network are the input gate layer and the tanh layer. The input layer determines how to update the cell state, followed by the tanh layer that creates new candidate value vectors, which could be potentially added to the state. These two layers are combined to create an update for the state given by equations 2.2 and 2.3.

$$i_t = \sigma (W_i . [h_{t-1}, x_t] + b_i) \quad (2.2)$$

$$\tilde{C}_t = \tanh (W_C . [h_{t-1}, x_t] + b_C) \quad (2.3)$$

The previous cell state C_{t-1} is then updated to the new cell state C_t.

We get equation (2.4) for updating the cell, by using the values from equations (2.1), (2.2) and (2.3)

$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$ (2.4)

The last state determines the output which is based on the filtered version of the cell state, i.e., we run a sigmoid layer on the cell state to decide which portions from the state should go to output followed by running the cell state through tanh layer to make our values between the range of –1 and 1.

$o_t = \sigma (W_o . [h_{t-1}, x_t] + b_o)$ (2.5)

$h_t = o_t * \tanh (C_t)$ (2.6)

2.3 Incremental Learning for time series:

Incremental Learning is the type of machine learning where a learner tackles some predictive tasks by learning from a sequence of data one step at a time. The goal of using such a technique is to be able to improve accuracy for the sequence of predictions made by the learner given the knowledge of the correct answers to the previous prediction's tasks. This type of learning can be used on time series data as well where the goal is to improve accuracy/corrections over time.[19][20][21].

Incremental learning has been used in a lot of time series related problems [45] like activity recognition systems [47], activity tracking systems [44][46], recommendation systems [37].

For our problem, we use the incremental method to personalize our model by updating our model features after every prediction. This would help the model personalize the prediction for the subjects and in turn improve the accuracy of predictions for each subject over time.

2.4 Blood Glucose Forecasting

A key issue faced by Type-1 diabetes patients while controlling the amount of insulin delivery is to predict high or low BG levels fairly early so that timely interventions can be made. For example, if we can predict with a reasonably high probability that the blood glucose level for a particular patient will drop to 65mg/dl in

the next hour, cause the patient to go into a state of hypoglycaemia, then the insulin can be reduced by sufficient quantity to prevent any complications.

2.4.1 Open-loop systems

Open-loop systems usually use patients administering insulin delivery methods. The patient has to administer the insulin delivery himself or herself at different times of the day. Usually, the patients inject the insulin formation in the morning that provides them with the basal dose requirement for the day. Additionally, patients need to inject another dose called bolus before each meal intake. The amount of dosage calculation is usually dependent on the glucose level and the amount of carbohydrate intake. Additionally, there is often a delay between injecting insulin into the subcutaneous tissue and it is diffused into the bloodstream which then helps in the regulation of the blood glucose levels in the body. Patients can also replace the insulin injections with external insulin pumps. The pump is usually always attached to the diabetic patient and a basal amount of insulin is supplied throughout the day. Secondly, when the patient wants to modify the insulin delivery because of a meal or exercise, he/she can easily do so from the pump. Besides subcutaneous delivery, recently other open-loop methods have also been proposed like Exuberant ® which allows patients to inhale insulin. Now, the biggest reason for such insulin delivery methods is to increase patient compliance since the patient now no longer has to receive injections or carry an insulin pump with them, but these oral methods have quite a few disadvantages. One of them majorly is that the bioavailability of the insulin is less than that of the subcutaneous infusion. Furthermore, a slow-releasing insulin analog has not yet been developed in the inhaled form, so basal administration is still necessary. Moreover, the rate of absorption of insulin via the lungs can vary from patient to patient, so the doses have to be very carefully determined. Thus, a lot of work is yet to be done to have oral insulin delivery systems.

2.4.2 Closed-loop systems

An alternative approach to the open-loop system is the closed-loop delivery system, which requires minimal involvement for the patient to maintain glucose control. The closed-loop system determines the insulin requirement in real-time, regardless of the situation and then deliver the proper insulin dosage to the patient. Thus, the system will adjust the dosage based on the patient's activities which means that getting the amount of dosage amounts correct based on the level of stress or meal intake is the most crucial aspect of the system. Although this type of method is nice to have since it requires minimal intervention from the patient and ideally all the changes are happening internally so the number of injections required could be reduced significantly, but this system has a lot of shortcomings. Since the best closed-loop delivery system would be to repair the patient's body's natural ability to infuse insulin and for that, pancreas transplantation would be required, which further involves a lot of complications.

So, a lot of natural methods are being explored that would help restore the patient's pancreas ability by planting foreign cells that would help secrete insulin. A lot of research is being done to construct artificial and engineered solutions. Thus ideally, a fully automated artificial pancreas system would have a CGM that would help to record the glucose levels of the patient regularly, there would be a control algorithm that would take in the glucose values and possibly other values like meal, activity, stress and they compute an appropriate insulin infusion rate which would then pass on to the insulin pump.

The FDA in 2016 approved the first commercial artificial pancreas [22], a closed-loop insulin delivery system, which paved a way for intelligent decision support systems for diabetes management. The artificial pancreas has three key components: a CGM which is used to measure glucose, an insulin pump which is used to deliver insulin to lower the blood glucose and a control algorithm that takes in input from the CGM and sends instructions to the pump. In our work, we try to improve the algorithm for the blood glucose prediction for the AP system.

2.5 Related Work

With the approval of the Artificial Pancreas System, numerous AP projects are being developed where all of them aim to make accurate blood glucose predictions [26][29]. The aim is to better the algorithm for BG prediction to be able to make as accurate a glucose prediction as possible for T1DM patients. With the recent development in machine learning techniques, a lot of them are being applied to help improve the accuracy of blood glucose prediction. Like M.W Percival et al. developed a multi-parametric model for predictive control algorithm for insulin delivery in T1Dm patients using clinical parameters [39]. Furthermore, a lot of deep learning techniques [32][43] are being developed with the key focus on trying to improve the accuracy of BG forecasting [28]. Md Fazle Rabby et al. developed a stacked LSTM model to make blood glucose predictions [22]. Another such work done was by Taku Yamagata et al. who designed a model-based reinforcement learning for type-1 diabetes blood glucose control and insulin dose decisions [23]. Harry Rubin-Falcone et al. tried to improve blood glucose forecasting, by building a deep residual time-series forecasting approach where propose a method to forecast in blocks and stages. They were able to achieve an RMSE score of 18.2mg/dl [24]. In addition, deep physiological models are being developed to achieve better accuracy of blood glucose prediction [27][34]. J. Chen et al. developed a dilated recurrent neural network for blood glucose concentration prediction where they were able to achieve an RMSE of 19.04mg/dl [48]. Another similar work by F.Allam et al. used recurrent neural networks for the prediction of blood glucose and their results showed that RNN was better at prediction than a feed-forward neural network model (NNM) for longer prediction horizons [50]. John Martinsson et al. developed an automatic blood glucose prediction with confidence using RNN where the model not only outputs the predicted future BG but also an estimate of the certainty of accuracy [51]. Furthermore, many researchers have tried using deep reinforcement learning for blood glucose forecasting. For instance, Taiyu Zhu et al. developed a dual-hormone control for T1DM using reinforcement learning [35]. Although a lot of work is being done to improve the accuracy of the blood glucose forecasting for the T1DM patients, it is a well-known factor in clinical practices that being able to achieve a tight glycaemic control for certain diabetic patients is an extremely difficult and complex process. Since different patients exhibit varied variations in their BG signals due to numerous factors that play a role in affecting the blood glucose in the person. These factors could be

weather, physical activity, stress, diet or age of the patient, maybe some other diseases or menstrual cycle or pregnancy in women. Due to the variation in these factors that are not easily identifiable, having models that can personalize the BG prediction for an individual is very important. Since personalized models would be able to capture specific lifestyle factors of particular individuals. Due to these reasons, many new models are being developed to help to personalize the BG predictions for the patients. Alessandro Aliberti et al. developed a data-driven patient specialized neural network for BG prediction [29], while Zitao Liu et al. developed a personalized predictive framework for multivariate clinical time series prediction using adaptive model selection [30]. Silvia Oviedo et al. did a review of the ways personalized blood glucose prediction strategies are being developed for T1DM patients [31]. Joon Bok Lee et al. in their work talk about how they use a model-based personalization scheme for an AP system for T1DM [33]. D-Y Kim et al. developed an individual glucose prediction model using RNN where they were able to achieve an RMSE pf 21.5 mg/dl for the patients [49]. The research study conducted by Taiyu Zhu et al. [6] used a convolutional neural network (CNN) model to personalize the forecast of the glucose levels of patients with type 1 diabetes. The model is mainly built by using causal dilated CNN layers and it employs WaveNet algorithms. Upon testing their model on the Ohio dataset, they were able to achieve an RMSE of 21.72 among the six patients. Giacomo et al. [7] were able to employ a personalized bi-directional LSTM for blood glucose prediction in T1DM individuals where they obtained a prediction accuracy of (P_H=30/60) RMSE: 20.20/34.19 mg/dl. Ran Cui et al. [8] applied an attention-based deep learning network for personalized glucose prediction where the model takes in both temporal dependencies and physiological relations among glucose and glucose-related life events like insulin and carbohydrate intake. The model proposed by the authors when tested on the OhioT1DM dataset was able to achieve an average RMSE of17.82 mg/dl for 30 minutes and 28.54 mg/dl for 60 minutes. Taiyu Zhu et al. [9] proposed a dual hormone control algorithm for people with type 1 diabetes that uses deep reinforcement learning. The authors developed a data-driven model using the UVA/Padova Simulator. They first pre-trained a generalized model using long-term exploration in an environment with average T1D subject parameters which are provided by the simulator, and then they adopt an important sampling to train personalized models for each patient. The proposed algorithm was able to reduce glycaemic events by 93% for adults and 83% for adolescents. Ivanoe De Falco et al. [10] also investigated a

neuroevolution approach to personalize the prediction of blood glucose for TIDM patients. They developed a neuroevolution algorithm to model and predict future personalized blood glucose levels.

Machine Learning techniques are slowly being used to personalize the future predictions of the blood glucose levels in T1DM patients. Accurate estimations of these personalized models for the prediction horizons of 30 to 60 min will allow a T1DM patient to take proper actions in advance to avoid hypo and hyperglycaemic episodes. Thus, we focus our efforts and our review of related work on approaches for personalized BG forecasting.

Although the different models are proposed for personalizing the blood glucose prediction in T1DM, what they don't consider is the fact that this prediction should be improving over time as more data of the patient is available. In this paper, we try to address this problem. Inspired by the prior work in personalization for blood glucose forecasting, we design an incremental deep learning model using LSTM that would make personalized BG prediction along with updating itself when new data is available. We refer to the work of Haoran Xu et al. where they propose ways incremental learning can be used in time-series forecasting [18]. Since incremental learning for personalization had not been applied to blood glucose forecasting before, we take inspiration from other time series forecasting problems that use incremental learning like the work done by Po-Han Chang et al. where they propose online and offline incremental learning techniques for personalized blood pressure prediction [19].

Chapter 3:

Methods

3.1 Problem Setup:

In this work, we focus on the task of blood glucose (BG) forecasting. BG forecasting can be modelled in two ways: (1) recursively forecasting where the estimates are used for future forecasts or (2) directly forecasting where the prior estimates do not influence future estimates.

The basic recursive setup is defined as follows:

$$x'(t) = \varphi[(x(t-1), x(t-2), ..., x(t-n))] \quad (3.1)$$

$$x'(t+1) = \varphi[(x'(t), x(t-1), x(t-2), ..., x(t-n))] \quad (3.2)$$

$$x'(t+h) = \varphi[(x'(t+(h-1)), ..., x(t-1), x(t-2), ..., x(t-n))] \quad (3.3)$$

In the above setup, x and x' are the actual and estimated values of the variable being forecasted (so, in our case it is BG), t is the timepoint the prediction is being made for, and φ is the function which is being used to make predictions. P_H is the prediction horizon which indicates how many time points in the future we want to forecast. So, $h = 0,\ldots P_H$. So, for a prediction horizon of 30 minutes with the input values being recorded every 5 minutes, $P_H = 6$ and as shown in equation (3.3) the prediction for the last time point t+h is a function of all prior predictions.

In the direct method, the data is split into chunks of fixed-sized sequences of historical and future points which can then become features and outputs for a classic regression task. With the variation definition being the same as previously described, there is now no dependence between the forecasts:

$$x'(t+h) = \varphi_h[(x(t-1), x(t-2), ..., x(t-n))] \quad (3.4)$$

The basic setup shown here provides single step forecasts which is what we are going to use in our work, i.e. single step BG forecasting. So, the single-step methods estimate $p(x_{t+P_H} | x_{0:t-1})$ with x_{t+P_H} being the glucose value of P_H time instances in the future. This can be achieved in a recursive or a direct way.

3.2 Data

3.2.1 OhioT1DM

The experiments are repeated on the OhioT1DM dataset which is a widely used benchmark dataset for BG prediction [10]. The dataset was released in 2018 and then updated in 2020. The updated dataset in 2020 has eight weeks of data for 12 people with type 1 diabetes. In total this yielded 600 days of data, with 177,000 CGM values(~ 144k training, ~33k testing). The average and the median number of recorded days were 54±3.02 and 56 days per person. Out of the twenty features originally recorded in the dataset, we used CGM values for our experiment.

3.3 Data Pre-processing

For the dataset, we extract the single feature (raw CGM values) for our BG forecasting. Due to a high error rate and noise in the CGM data, we include several pre-processing steps shown in Table 3.3 to remove implausible values and fill in missing ones. Firstly, we exclude any CGM values less than 15mg/dl, as glucose values cannot fall below this level for a person who is still conscious.

Pre-processing step	**Technique**	**Parameters/Constraints**
Windowing	Overlapping. Sliding window	history window = 12 stride=1, $P_H = 6$
Thresholding CGM values	Remove samples with CGM $\leq$ k	k=15mg/dl

Imputing CGM values	Interpolating (training) Extrapolation (test)	Linear and for gaps ≤ 30 min
Standardization	Standard Scaling	Mean=0 and std=1
Filtering (training data only)	Median filter	Window size = 5*

P_H: Prediction horizon

Table 3.3 Pre-processing steps performed on the dataset

Missing values in the training set are imputed using linear interpolation for gaps less than 30 minutes in duration. Interpolation helps to impute a missing value x_t at time t by assuming a linear relationship between the past data points (x_{t-1}), and future data points (x_{t+1}) and fits a straight-line using equation (3.5) [11].

$$x_t = \{(x_{t-1} - x_{t+1})(t - t_{+1})\}/ (t_{-1} - t_{+1}) + x_{t+1} \quad (3.5)$$

where $t_{-1} < t < t_{+1}$. For test data, we use first-order extrapolation to avoid using data from the future. We then use a median filter with a window size of 5 samples [12] to smooth the CGM data and address any discontinuities present in the data, which might cause unnecessary variance in the predictions. The median filtering was only done for the training data and not for the test data to make sure we can test the robustness of the models for real-world use.

For the direct-forecasting task, we use a sliding window to split the data into fixed-sized sequences of the past and future BG values. There were three parameters for the moving window namely the size of the history window i.e., size of the historical data to use for forecasting, prediction offset and horizon i.e., how far ahead in the future and how many values we want to predict, and a stride i.e., number of samples to skip while we are sliding the window ahead. An hour which is 12 samples for 5-minutes intervals of the past

values were used with a prediction horizon of 30 minutes (P_H = 6 samples). We use a stride of size 1 which means that we use overlapping windows to partition the data. In the case of single-step forecasting, a single future glucose value is being estimated.

3.4 Experiments

In this work, our idea of personalization is to update the features of the model after a certain amount of time which we refer to as update_window and then check to see if we can gradually decrease the error rate for each patient. The experiments aimed to see if the model can be able to improve the accuracy of prediction for the subjects and achieve a satisfactory accuracy rate.

3.4.1 Method

The method we use for our work is incremental learning. Here first we define an update window of 1 day. Update window is the period of time we want to predict a value for prediction horizon of 30 minutes and update our model by updating the features of the model. We then train the model again and then test on the next day of data; we repeat the process till all the data of the test subject has been tested upon. We then calculate the average RMSE values over all the days we have and compare them with the non-incremental LSTM model and with other prior work to see the performance of our model.

For example:

- Train data: Data from 1 to n-1 subjects
- Test data: Data from the n^{th} subject (8 weeks of data)
- History window = 12 (the amount of data we want to see)
- Prediction horizon = 30 min (we want to predict into the future)
- Update window = 1 day

So, after training the model on the n-1 subject's data, we then take the test subject data and firstly divide the 8 weeks of data into chunks of days- that is roughly 56 days. Then for the first chunk of 1-day data, we take data points of the history window size, make a prediction of 30 minutes into the future, and then record the error score. We then slide the history window of size 1 and again predict the CGM value 30 minutes into the future. We repeat the process for that day and then record the average RMSE score. After that, we update the model features based on our previous predictions and train the model again. Now with this updated model, we test on the next day's data. We repeat this process till we have covered all 56 days. Finally, we take an average of all the 56 average RMSE scores.

We use Glorot normal initialization [13] to initialize the weight matrix in deep learning models. We use Python libraries such as sci-kit- learn, pyramid, and Keras to run the deep learning model. For our work, we use an LSTM model with a single layer of 32 units and an output layer of 1 unit since we are performing single step forecasting. For the activation function, we use the 'reLu' activation function.

3.4.2 Training

For the univariate setting, the time series consists only of the raw CGM values which are given to the input layer. We divide our dataset into a train set that consisted of data from n-1 subjects and a test set that consisted of data from the n^{th} subject.

3.5 Evaluation:

For our evaluation, we compare the results using statistical metric RMSE:

a) with a Vanilla LSTM model where the model is tested on each patient without any updates with a prediction horizon of 30 min.
b) compare our results with the prior personalization work

Root mean Square Error (RMSE):

The statistic metric used for the evaluation was RMSE, which is given by equation (3.6) where x and $\hat{x}$ are the actual and estimated BG values, respectively.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \qquad (3.6)$$

In the RMSE metric, n is the number of non-missing data points, x_i is the actual observations in the time series, and $\hat{x}_i$ is the estimated values in the time series.

We use RMSE as a standard statistical metric as it has been widely used in prior work to measure the deviation of the estimated BG values from the true BG levels.

For our work, we calculate the average RMSE scores for every day of the 56 days. We then take the average of all the RMSE scores for 56 days and report the best, total average, and worst RMSE for each subject.

Chapter 4:

Results Analysis

Table 4.1 shows the RMSE values for the baseline non-incremental LSTM model. The experiment was repeated 5 times and the average RMSE was recorded for each subject. We report the best, worst, and average RMSE for each of the 12 subjects in the OhioT1DM dataset.

Patient	P540	P544	P552	P559	P563	P567	P570	P584	P588	P591	P596	P575
Average	23.20	24.20	25.90	25.63	19.76	22.92	27.33	27.33	27.86	24.25	25.19	22.73
Best	21.21	32.32	18.45	32.79	18.43	21.18	23.83	23.33	23.89	25.82	24.04	21.60
Worst	27.23	19.24	26.59	23.20	25.06	24.35	30.42	30.42	29.69	21.80	28.33	23.56

Table 4.1: RMSE values for the non-incremental LSTM model

Table 4.2 shows the RMSE values for the incremental LSTM model. In the experiment we set the update window to be 1 day and the average RMSE is recorded for each subject. We report the best, worst, and average RMSE for each of the 12 subjects in the OhioT1DM dataset.

Patient	P540	P544	P552	P559	P563	P567	P570	P584	P588	P591	P596	P575
Average	25.16	24.59	23.68	23.73	22.54	21.67	24.51	24.97	23.52	25.71	23.23	26.48
Best	21.05	18.15	18.17	19.17	17.18	17.29	19.70	20.41	18.59	20.97	18.12	21.03
Worst	30.59	30.28	27.25	26.68	28.01	24.91	28.57	30.35	27.64	31.83	28.75	30.88

Table 4.2: RMSE scores for our incremental LSTM model for each subject

From the results shown in Table 4.2, we see that although our starting RMSE scores are high gradually they become better. This behaviour is expected since we have a cold start problem for the test subject- i.e., we

have not seen any data from this subject and as gradually we get more data, our prediction gets better over time.

Figures 4.1 to 4.12 shows the results of the 12 subjects and how their RMSEs are decreasing over time for each subject. This helps us visualize better, how with our cold start problem for the test subject we see a very high RMSE at the beginning and then slight fluctuations after some of the days' data from the subject is seen by the model, and then gradually the RMSEs start decreasing over the period of 56 days where towards the end we get a satisfactory accuracy.

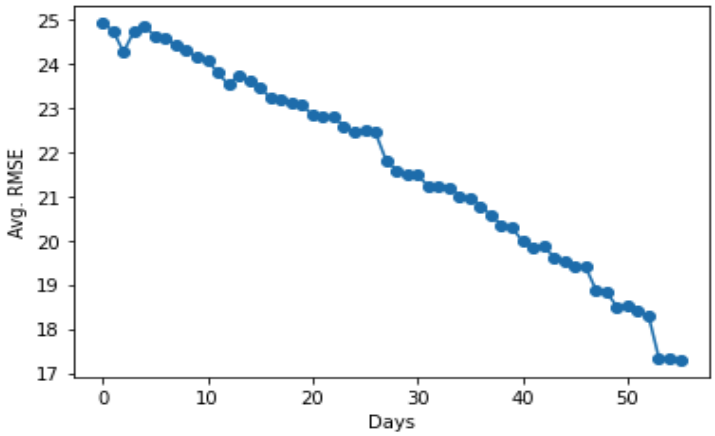

Figure 4.1 Average RMSE scores for 56 days for P567

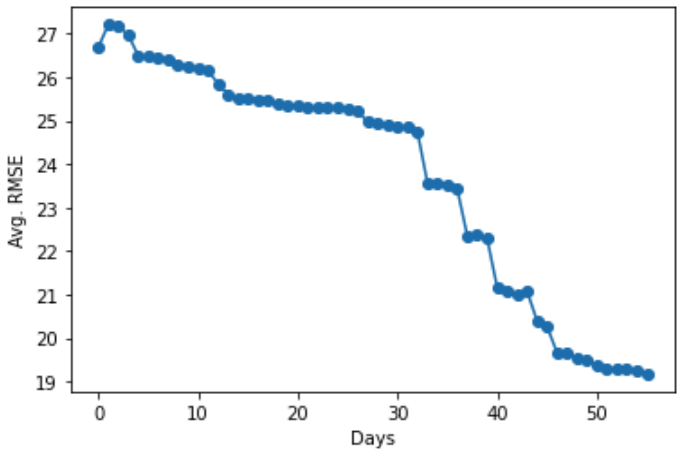

Figure 4.2 Average RMSE scores for 56 days for P559

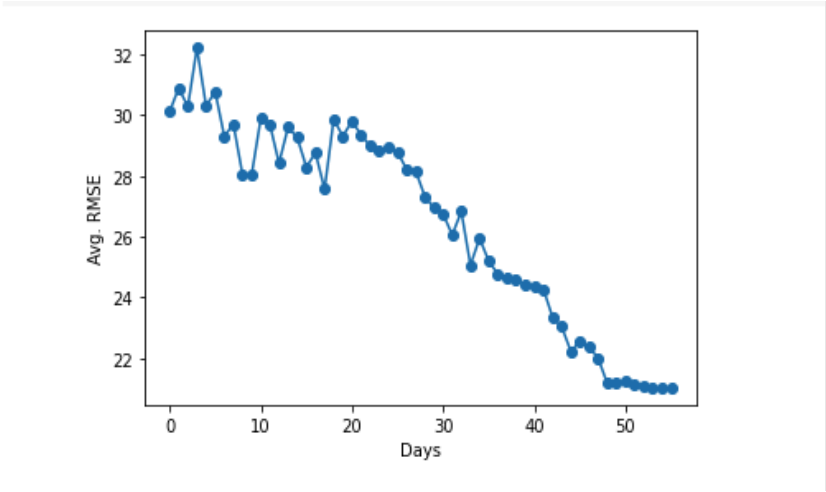

Figure 4.3 Average RMSE scores for 56 days for P575

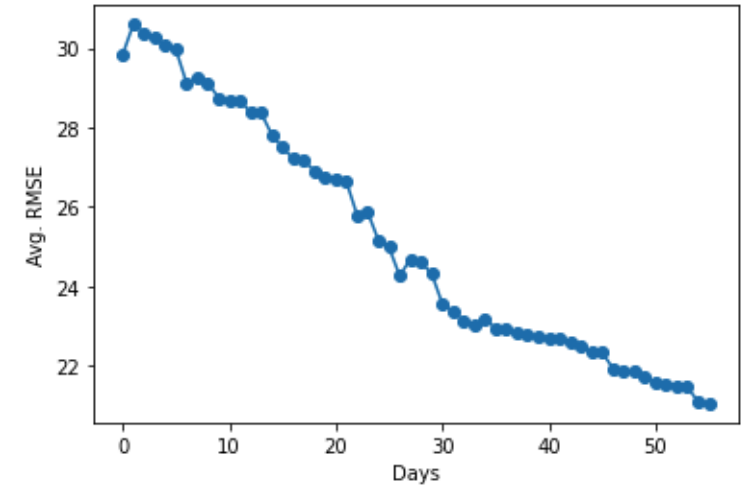

Figure 4.4 Average RMSE scores for 56 days for P540

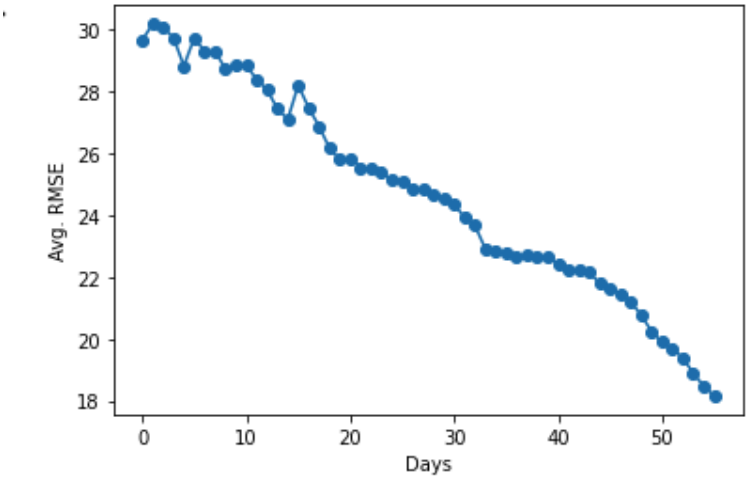

Figure 4.5 Average RMSE scores for 56 days for P544

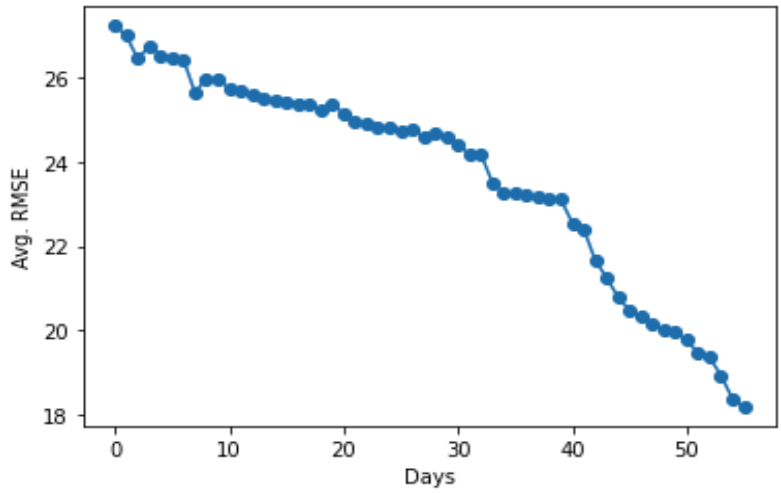

Figure 4.6 Average RMSE scores for 56 days for P552

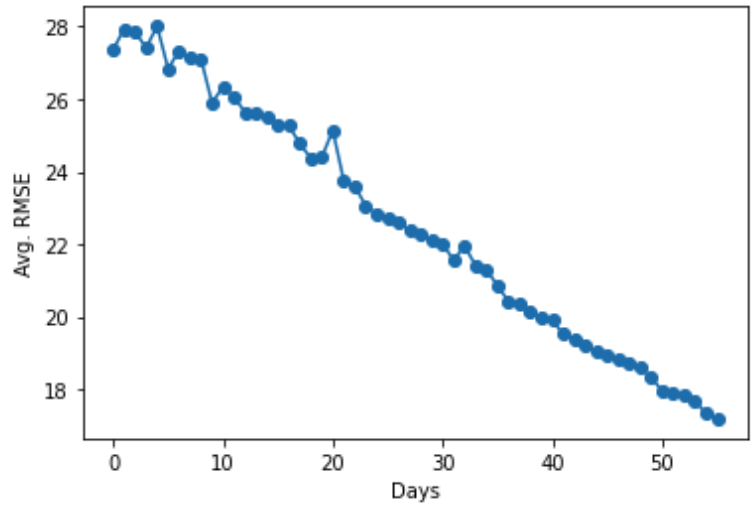

Figure 4.7 Average RMSE scores for 56 days for P563

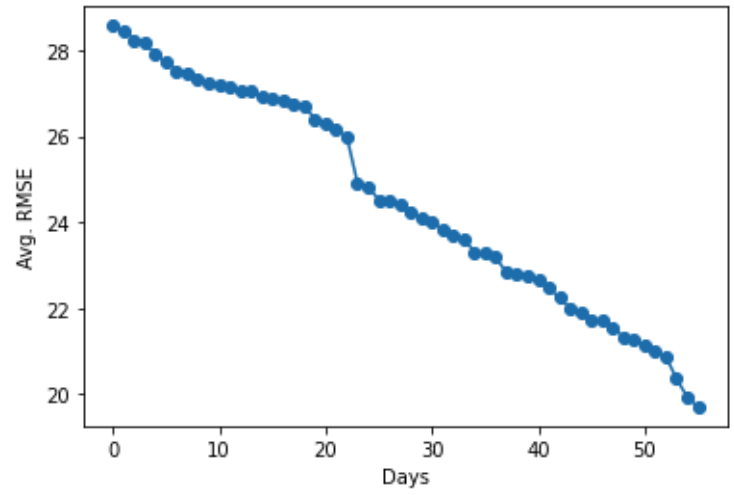

Figure 4.8 Average RMSE scores for 56 days for P570

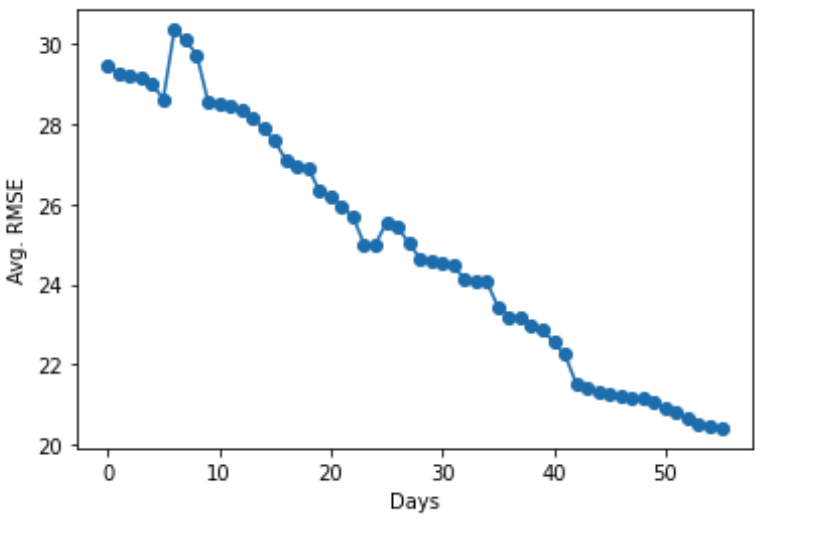

Figure 4.9 Average RMSE scores for 56 days for P584

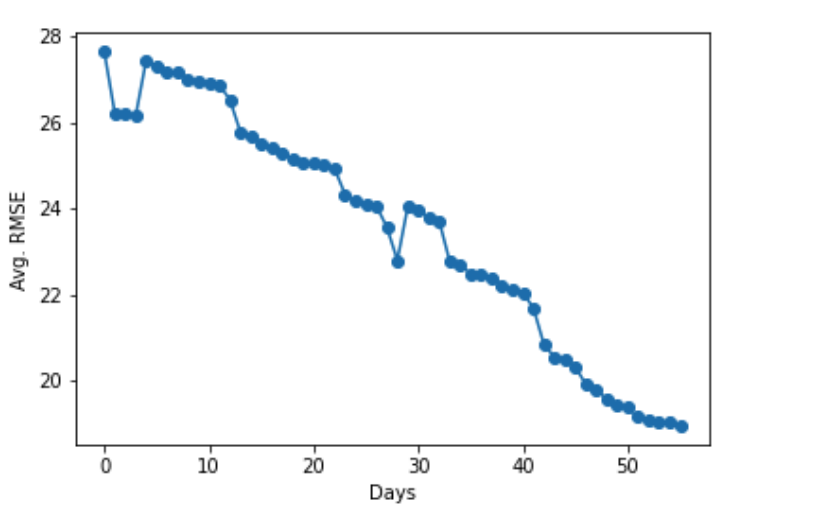

Figure 4.10 Average RMSE scores for 56 days for P588

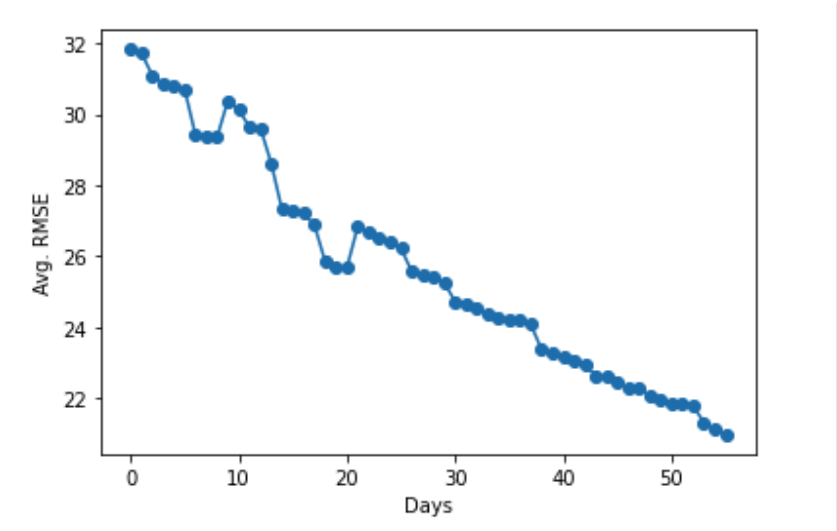

Figure 4.11 Average RMSE scores vs days for P591

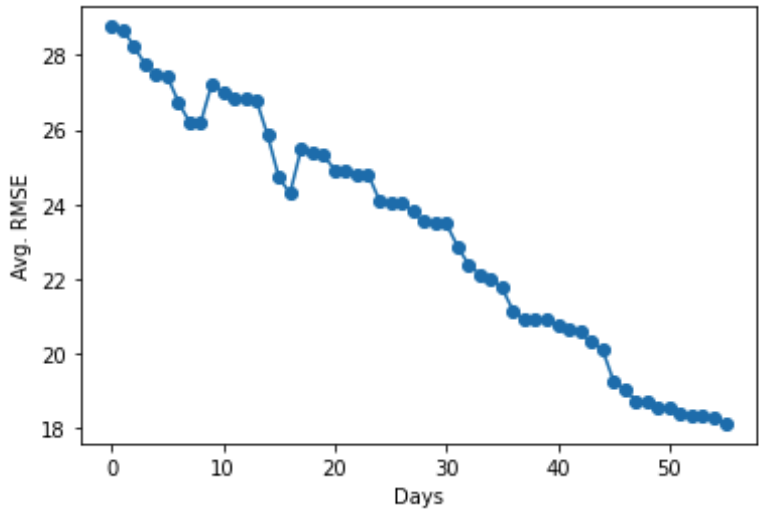

Figure 4.12 Average RMSE scores for 56 days for P596

Compared to our baseline LSTM model (Table 4.1), the incremental LSTM, as shown in Table 4.2, seems to perform better over time where there is a gradual decrease in the RMSEs. This behaviour is found to be consistent for all the subjects. So, we can conclude that updating the model over time as and when more data is available from the patients can be beneficial to improving the accuracy of predictions for them.

We further compare our results with one of the prior works for the personalization of BG forecasting [6]. In this work, the authors use a convolutional neural network (CNN) model to personalise the forecast of the glucose levels of patients with type 1 diabetes.

In comparison to the results achieved by the authors, our model seems to perform better for these six subjects, where our RMSE values get better over time as seen in Table 4.3. Our best RMSE scores are also better than the ones they achieved.

Taiyu Zhu et al. [6]	Patient	P559	P563	P570	P575	P588	P591
	Average	22.48	20.35	18.26	25.65	21.69	24.59
	Best	21.72	20.17	18.03	24.80	21.41	24.22

	Best Average	21.7267 ± 2.5237					
Our incremental LSTM model	Average	23.73	22.54	24.51	26.48	23.52	25.71
	Best	19.17	17.18	19.70	21.03	18.59	20.97
	Best Average	19.44 ± 2.365					

Table 4.3: Comparison of our results with prior work [6]

Thus, our personalized model can improve the accuracy over time for the subjects and seemed to perform significantly better. So, having a personalized model where features are updated incrementally seemed to be better than an offline personalized model.

Chapter 5:

Conclusion

We implement an incremental LSTM personalized model where the goal is to update the features of the model and help to improve the accuracy of blood glucose prediction for the subjects with type 1 diabetes of the OhioT1DM dataset over time.

We aim to replicate the real-life scenario where if a person initially starts using the artificial pancreas system, the predictions won't be accurate since we do not have any data of the patient. After a warm-up period, when we have some data from the patient, the accuracy of the blood glucose predictions should get better over time.

So, we try to build a personalised model that can help us to improve the accuracy of predictions for the subjects, over time. We implement an incremental LSTM model where we update the features of the model after an update window of 1 day and we observe that the accuracy of our model improves gradually which eventually gets better over a period of time. Initially, we get very high RMSE values for the subjects, which is expected behaviour as it is a cold start problem for that subject, but eventually as and when we get more data our model can improve the accuracy of predictions. When compared to a non-incremental LSTM model, our incremental model seems to be doing better for most of the subjects. Also, in comparison to prior works [6] our model can achieve better results where our best average RMSE score is comparatively better. It is interesting to see how the model performs poorly in the initial stage for the subjects, which is consistent behaviour for all the subjects, as seen in figures 4.1- 4.12. Then the values gradually become better towards the end. We do see some spikes in between which can be due to noisy data but overall, the performance of the model gradually improves towards the end, and we can achieve a satisfactory score for all the subjects in the OhioT1DM dataset.

For future work, we plan to use a larger dataset like OAPS with 3 years of data from the subjects, to test if seeing data over a very long period of time can help improve the accuracy or whether we reach a point of

stagnation where the accuracy starts deteriorating again. Although our model performs well in the univariate setting, to make this model more personalized, a multi-variate setting can be used where features other than CGM values like meal intake, bolus, activity can be included. Adding these features should improve the accuracy further while giving us an updated version of the personalized model for the subjects.

Other variations of the model can also be used like using an update window of a week or a month. The stride to shift the history window can also be varied. Another modification that can be included especially for large datasets like OAPS is ‘forgetting of data’ i.e., putting more weightage on current data over the data at the very beginning. Integrating all these variations could potentially help us improve the prediction accuracy in future work.

Bibliography:

1. Ransford Henry Botwey, Elena Daskalaki, Peter Diem, and Stavroula G Mougiakakou. Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4843–4846. IEEE, 2014.
2. Kezhi Li, John Daniels, Chengyuan Liu, Pau Herrero-Vinas, and Pantelis Georgiou. 'Convolutional recurrent neural networks for glucose prediction'. *IEEE Journal of biomedical and health informatics*, 2019.
3. Ali Mohebbi, Alexander R Johansen, Nicklas Hansen, Peter E Christensen, Jens M Tarp, Morten L Jensen, Henrik Bengtsson, and Morten Mørup. 'Short term blood glucose prediction based on continuous glucose monitoring data. *arXiv preprint arXiv:2002.02805*, 2020.
4. Amaury Lendasse, Eric de Bodt, Vincent Wertz, and Michel Verleysen. Non- linear financial time series forecasting-application to the bel 20 stock market index. *European Journal of Economic and Social Systems*, 14(1):81–91, 2000
5. O ̈mer Faruk Ertugrul. 'Forecasting electricity load by a novel recurrent extreme learning machines approach'. *International Journal of Electrical Power & Energy Systems*, 78:429–435, 2016.
6. Taiyu Zhu*, Kezhi Li*, Pau Herrero, Jianwei Chen, Pantelis Georgiou Department of Electronic and Electrical Engineering, Imperial College London,
   'A Deep Learning Algorithm for Personalized Blood Glucose Prediction' London SW5 7AZ, UK
   taiyu.zhu17@imperial.ac.uk, kezhi.li@imperial.ac.uk
7. Giacomo Cappon et al. 'A Neural Network-Based Approach Personalize Insulin Bolus Calculation Using Continuous Glucose Monitoring' *Journal of Diabetes Science and Technology 2018, Vol. 12(2) 265–272*
8. Ran Chui et al. 'Personalized Short-Term Glucose Prediction via Recurrent Self Attention Network', *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS) 10.1109/CBMS52027.2021.00064*

9. Cindy Marling and Razvan Bunescu. 'The OhioT1DM dataset for blood glucose level prediction': Update 2020. *IJCAI*, 2020.
10. Mathieu Lepot, Jean-Baptiste Aubin, and François HLR Clemens. Interpola- on in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment. *Water*, 9(10):796, 2017.
11. Taiyu Zhu, Kezhi Li, Pau Herrero, Jianwei Chen, and Pantelis Georgiou. 'A deep learning algorithm for personalized blood glucose prediction'. In *KHD@ IJCAI*, pages 64–78, 2018
12. Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.
13. International Diabetes Federation. Idf diabetes atlas eighth edition 2021. https://www.idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html, 2021. [Online accessed 09-Dec-2021].
14. Chengyuan Liu, Josep Vehi, Nick Oliver, Pantelis Georgiou, and Pau Herrero. Enhancing blood glucose prediction with meal absorption and physical exercise information. *arXiv preprint arXiv:1901.07467*, 2018
15. Andrew S Levey, Josef Coresh, Ethan Balk, Annamaria T Kausz, Adeera Levin, Michael W Steffes, Ronald J Hogg, Ronald D Perrone, Joseph Lau, and Garabed Eknoyan. National kidney foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Annals of internal medicine*,139(2):137–147, 2003
16. D Mozaffarian, EJ Benjamin, AS Go, DK Arnett, MJ Blaha, M Cushman, SR Das, S de Ferranti, JP Despre´s, HJ Fullerton, et al. Heart disease and stroke statistics-2016 update: a report from the American heart association. *Circulation*, 133(4): e38, 2016.
17. Ian Fox, Lynn Ang, Mamta Jaiswal, Rodica Pop-Busui, and Jenna Wiens. 'Deep multi-output forecasting: Learning to accurately predict blood glucose trajectories. In *Proceedings of the 24th*

ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pages 1387–1395, 2018.

18. Haoran Xu, et al. 'An Online Incremental Learning Algorithm for Time Series', *2015 International Joint Conference on Neural Networks (IJCNN)*
19. Po-Han Chang et al. 'Offline and Online Learning Techniques for Personalized Blood Pressure Prediction and Health Behavior Recommendations', *IEEE Access 10.1109/ACCESS.2019.2939218*
20. Steven C. H. Hoi et al. 'Online Learning: A Comprehensive Survey' – *arXiv:1802.02871 [cs] 2018*
21. FDA. Fda's efforts to advance artificial pancreas device systems. https://www.fda.gov/medical-devices/consumer-products/ artificial-pancreas-device-system, 2018. [Online; accessed 15-Nov- 2021].
22. Md Fazle Rabby et al. 'Stacked LSTM based deep recurrent neural network with Kalman smoothing for blood glucose prediction', *BMC Medical Informatics Decision Making (2021) 21:101* https://doi.org/10.1186/s12911-021-01462-5
23. Taku Yamagata et al. 'Model-Based Reinforcement Learning for Type 1 Diabetes Blood Glucose Control', AAI4H@ECAI 26 June 2020
24. Harry Rubin-Falcone et al. 'Deep Residual Time-Series Forecasting: Application to Blood Glucose Prediction' *KDH @ECAI 2020*
25. Qingnan Sun et al. 'A dual model adaptive basal-bolus advisor based on reinforcement learning', *IEEE Journal of Biomedical and Health Informatics 10.1109/JBHI.2018.2887067 2168-2194, 2168-2208, 2019*
26. F.Stahl et al. 'Diabetes mellitus model and short-term prediction based on blood glucose measurements', *Mathematical Biosciences 10.1016/j.mbs.2008.10.008, 2009*
27. Mario Munoz Orgnaero, 'Deep Physiological Model for Blood Glucose Predictions in T1Dm Patients', *Sensors doi:10.3390/s20143896, 2020*
28. Kezhli Li et al. 'GluNet: A Deep Learning Framework for Accurate Glucose Forecating', *IEEE Journal of Biomedical and Health Informatics DOI: 10.1109/JBHI.2019.2931842, 2020*

29. Alessandro Aliberti et al. ‘Data-Driven Patient Specialized Neural Networks for Blood Glucose Prediction’, *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) 10.1109/ICMEW46912.2020.9105950*

30. Zitao Liu et al. ‘A Personalized Predictive framework for Multivariate Clinical Time Series via Adaptive Model Selection, *Proc ACM Int Conf Inf Knowledge Manag. 2017 November; 2017: 1169–1177. DOI: 10.1145/3132847.3132859.*

31. Silvia Oviedo et al. ‘A review of personalized blood glucose predictions strategies for T1DM patients’, *International Journal for Numerical Methods in Biomedical Engineering DOI: 10.1002/cnm.2833, 2017*

32. Dae - Yeon Kim et al. ‘Developing an Individual Glucose Prediction Model using Recurrent Neural Network’, *Sensors DOI: 10.3390/s20226460, 2020*

33. Joon Bok lee et al. ‘Model-Based Personalization Scheme of an Artificial Pancreas for Type1 Diabetes Applications, *2013 American Control Conference (ACC) DOI: 10.1109/ACC.2013.6580276*

34. Ivan Contreras et al. ‘Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models’, *PLOS ONE DOI: 10.1371/journal.pone.0187754, 2017*

35. Taiyu Zhu et al. ‘Personalized Dual-Hormone Control for Type 1 Diabetes Using Deep Reinforcement Learning, *Explainable AI in Healthcare and Medicine Series Title: Studies in Computational Intelligence DOI: 10.1007/978-3-030-53352-6_5, 2021*

36. Lim, B, Zohren S. 2021 ‘Time-series forecasting with deep learning: a survey’. *Phil. Trans. R. Soc. A* **379**: 20200209. https://doi.org/10.1098/rsta.2020.0209

37. Tigran Ishkhanov et al. ‘Time-based Sequence Model for Personalized and Recommendation System’, *arXiv:2008.11922 [cs, stat], 2020*

38. National Institute of Diabetes and Digestive and Kidney Diseases, National Diabetes Information Clearinghouse, National Institutes of Health. *National Diabetes Statistics, 2011*. Available: http://diabetes.niddk.nih.gov/dm/pubs/statistics/

39. M. W. Percival, Y. Wang, B. Grosman, E. Dassau, H. Zisser, L. Jovanovič, and F. J. Doyle III, "Development of a Multi-Parametric Model Predictive Control Algorithm for Insulin Delivery in Type 1 Diabetes Mellitus Using Clinical Parameters," *J. Process Control,* vol. 21, pp. 391-404, 2011.

40. D. Daneman, "Type 1 diabetes," *Lancet,* vol. 367, pp. 847-58, 2006.

41. O. Franco, E. Steyerberg, F. Hu, J. Mackenbach, and W. Nusselder, "Associations of Diabetes Mellitus with Total Life Expectancy and Life Expectancy with and Without Cardiovascular Disease," *Arch Intern Med,* vol. 167, pp. 1145-1151, 2007.

42. S. A. Weinzimer, G. M. Steil, K. L. Swan, J. Dziura, N. Kurtz, and W. V. Tamborlane, "Fully Automated Closed-Loop Insulin Delivery Versus Semiautomated Hybrid Control in Pediatric Patients with Type 1 Diabetes using an Artificial Pancreas," *Diabetes Care,* vol. 31, pp. 934-9, May 2008.

43. H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A deep learning approach to diabetic blood glucose prediction," *Frontiers in Applied Mathematics and Statistics*, vol. 3, p. 14, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fams.2017.00014

44. Weiss, G.; Lockhart, J. The impact of personalization on smartphone-based activity recognition. In Proceedings of the AAAI Workshop on Activity Context Representation: Techniques and Languages, Toronto, ON, Canada, 22–26 July 2012; pp. 94–104.

45. Gepperth, A.; Hammer, B. Incremental learning algorithms and applications. In Proceedings of the European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, 27–29 April 2016; pp. 357–368.

46. Wang, Z.; Jiang, M.; Hu, Y.; Li, H. An incremental learning method based on probabilistic neural networks and adjustable fuzzy clustering for human activity recognition by using wearable sensors. *IEEE Trans. Inf. Technol. Biomed.* **2012**, *16*, 691–699.

47. Siirtola, P.; Koskimäki, H.; Röning, J. Personalizing human activity recognition models using incremental learning. In Proceedings of the European Symposium on Artificial Neural Networks,

Computational Intelligence and Machine Learning (ESANN), Bruges, Belgium, 24–26 April 2018; pp. 627–632.

48. J. Chen, K. Li, P. Herrero, T. Zhu, and P. Georgiou, 'Dilated recurrent neural network for short-time prediction of glucose concentration.', *KDH,* (2018).
49. D.-Y. Kim, D.-S. Choi, J. Kim, S. W. Chun, H.-W. Gil, N.-J. Cho, A. R. Kang, and J. Woo, "Developing an individual glucose prediction model using recurrent neural network," *Sensors*, vol. 20, no. 22, p. 6460, 2020.
50. F. Allam, Z. Nossai, H. Gomma, I. Ibrahim, and M. Abdelsalam, "A recurrent neural network approach for predicting glucose concentration in type-1 diabetic patients," in *Engineering Applications of Neural Networks*. Springer, 2011, pp. 254–259.
51. J. Martinsson, A. Schliep, B. Eliasson, C. Meijner, S. Persson, and O. Mogren, "Automatic blood glucose prediction with confidence using recurrent neural networks," in *the 3rd International Workshop on Knowledge Discovery in Healthcare Data*. CEUR Workshop Proceedings, 2018, pp. 64–68.
52. Nick McCullum, 'Long Short-Term Memory Networks', https://nickmccullum.com/python-deep-learning/lstms-long-short-term-memory-networks/,
53. Jason Brownie, 'How to develop LSTM Models for Time Series Forecasting', https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/, 2018