# Comparing a Personalised LSTM Model and a Biomedical Model for Blood Glucose Forecasting

Esha Manchanda[1]

[1] *Department of Computer Science, Ashoka University*

## I. INTRODUCTION

Type One Diabetes Mellitus is a condition wherein the pancreatic beta cells do not produce the hormone insulin. To manage the condition, type one diabetics must take exogenous insulin injections/infusions to monitor their blood glucose daily.
Insufficient insulin can cause elevated blood glucose levels which leads to long term complications, while too much insulin can cause hypoglycemia which is immediately catastrophic. Managing type one diabetes consists of delivering "just enough" insulin to regulate blood glucose levels. However, the determination of "just enough" at any given moment is a challenge.
Lately, automated insulin delivery systems (also known as artificial pancreas) are becoming more popular for the management of type one diabetes. Automated insulin delivery systems forecast blood glucose levels a few hours into the future and suggest insulin doses accordingly. Any system, however, can only be helpful in suggesting insulin dosing if its blood glucose forecasting is accurate.

## II. LITERATURE REVIEW

Currently, there are two main approaches to blood glucose forecasting: biomedical models and statistical/machine learning models. Biomedical models of insulin glucose dynamics are based in the physiology of glucoregulatory systems. An example is the Bergman minimal model which uses a multicompartment to describe interaction between active insulin, blood glucose, and endogenous insulin production [1]. The more recent UVA/Padova Type One Diabetes simulator describes blood glucose levels in the postprandial state [2]. Alternatively, various statistical and machine learning models have been proposed for blood glucose forecasting. Shanthi [3] introduced a classical statistical method, i.e. an autoregressive integrated moving average (ARIMA) model-based algorithm to predict blood glucose in 30 to 60 minutes prediction horizons. Daskalaki et al.[4] used a real-time learning recurrent neural network (RNN) fed with both glucose and insulin information. Bunescu et al. [5] proposed the use of support vector regression (SVR) for glucose prediction. However, LSTM Neural Networks outperformed all other methodologies for blood glucose forecasting[6]. Very few works have however incorporated insulin and carbohydrates in their LSTM prediction model, which are the main factors that dictate blood glucose dynamics.

## III. PROJECT DESCRIPTION AND GOALS

The management of type one diabetes is highly personalized — two people can have completely different blood glucose dynamics under identical exogenous circumstances. Biomedical model parameters are set to fit the mean data of a large subject database and the model can not be manipulated in any way to cater to an individual. In this work, I wanted to explore if personalised machine learning models for blood glucose predictions can prove to be more accurate than biomedical models of type one diabetes. To investigate this, I have implemented a personalised LSTM based recurrent neural network and compared its forecasting accuracy with the forecasting accuracy of the UVA/Padova Simulator.

## IV. THE DATASET

### A. Data Specifications

This work studies the data of a single type one diabetic, collected over a 8 month time period. The T1D is using a continuous glucose monitoring system (CGM) with a sampling rate of 5 minutes and an insulin pump throughout daily life. The following quantities are recorded using Apple's HealthKit framework and used as features in the LSTM network:

- instantaneous continuous glucose monitor measurement (mg/dL)
- insulin dose delivered (insulin units),
- estimated carbohydrate ingestion (grams)

It is also worth mentioning that the individual is on DIY Loop [7], an insulin delivery system which automatically adjusts basal insulin values based on CGM measurements. However, this does not impact our model since basal insulin values were not considered.

### B. Cleaning and Preprocessing

Each of the above features were stored in separate files with dissimilar timestamps. The dataset was preprocessed such that each entry was rounded off to the nearest five minute timestamp. Linear interpolation was used to calculate missing blood glucose measurements. Since our model only considered boluses, insulin doses that were under 0.5IU/hr and took over 10 minutes to deliver were dropped. Normalisation of the dataset was skipped since it did not improve prediction accuracy.
The first 50000 data points (approximately 180 days) were used to train and validate the forecasting model, and the subsequent 17000 data points (approximately 60 days) were used to measure prediction accuracy.

### C. Dataset Limitations

Sensor fault is very common in continuous glucose monitors and therefore blood glucose data sets are prone to be inaccurate. Sensor fault can be caused by decay of sensor sensitivity, pressure-induced sensor attenuation (PISA), and interruption in signal transmission, etc.[8]
Also, carbohydrate intake is estimated by type one diabetics and therefore may not always be precise.

## V. METHODOLOGY

### A. LSTM Prediction Model

Recurrent neural networks (RNNs) are particularly used for sequential data such as time series. Vanilla RNNs, however, are plagued by the vanishing gradient problem, which limits their capacity to learn long-term sequences. Long-short term memory architecture resolves this problem by introducing a memory cell and forget gate into the Vanilla RNN network. The memory cell holds/discards information according to the cell's mechanism, and the forget (f), input (i) and output (o) gates control how much the cell state value is preserved, updated with input values, and contributing to the hidden state, respectively[6].
In our experiment, we will be forecasting a future sequence of blood glucose using a current sequence of blood glucose. The problem can be represented as the following:

$$\boldsymbol{BG_{t+\tau} = f(BG_t + e_t)}$$

$BG_t$ is the sequence of blood glucose levels till time t and $e_t$ is the sequence of exogenous events (meals and insulin) till time t.
$BG_{t+\tau}$ represents the prediction horizon $\tau$ relative to the present time $t$. In our experiment, we have used 120 minute glucose, carbohydrate and bolus history to forecast blood glucose values 60 minutes in the future. Therefore, $\tau$ is set to 60 minutes and $t$ is a sequence of blood glucose going from time $t - 120$ minutes to time $t$. The setting of this history and prediction window was done in accordance with the time horizons type one diabetics consider while making treatment decisions.

| | DateTime | BG | Carbs | Bolus |
|---|---|---|---|---|
| 67363 | 2021-10-30 21:35:00 | 162.0 | 0.0 | 0.00 |
| 67364 | 2021-10-30 21:40:00 | 163.0 | 0.0 | 0.00 |
| 67365 | 2021-10-30 21:45:00 | 163.0 | 0.0 | 0.00 |
| 67366 | 2021-10-30 21:50:00 | 169.0 | 0.0 | 0.00 |
| 67367 | 2021-10-30 21:55:00 | 170.0 | 0.0 | 0.00 |
| 67368 | 2021-10-30 22:00:00 | 171.0 | 0.0 | 0.00 |
| 67369 | 2021-10-30 22:05:00 | 168.0 | 0.0 | 0.00 |
| 67370 | 2021-10-30 22:10:00 | 166.0 | 0.0 | 0.00 |
| 67371 | 2021-10-30 22:15:00 | 164.0 | 0.0 | 0.00 |
| 67372 | 2021-10-30 22:20:00 | 165.0 | 0.0 | 0.00 |
| 67373 | 2021-10-30 22:25:00 | 163.0 | 0.0 | 0.00 |
| 67374 | 2021-10-30 22:30:00 | 165.0 | 0.0 | 0.00 |
| 67375 | 2021-10-30 22:35:00 | 166.0 | 0.0 | 0.00 |
| 67376 | 2021-10-30 22:40:00 | 167.0 | 0.0 | 0.00 |
| 67377 | 2021-10-30 22:45:00 | 165.0 | 129.0 | 4.25 |
| 67378 | 2021-10-30 22:50:00 | 164.0 | 0.0 | 0.00 |

FIG. 1. Dataset after cleaning and preprocessing.

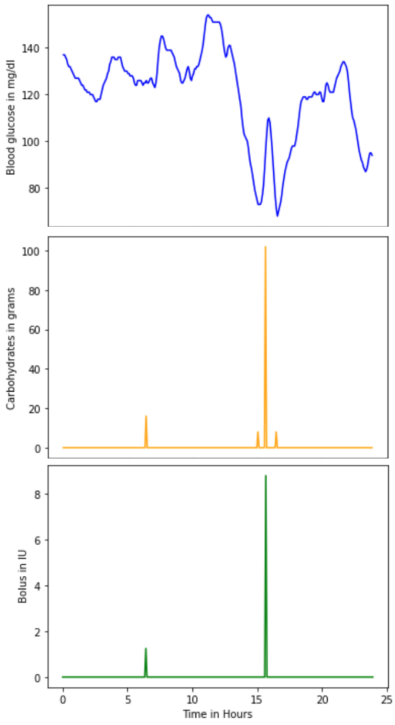


FIG. 2. Data for a single random day: blood glucose as measured by CGM, insulin delivered by pump, user carbohydrate log

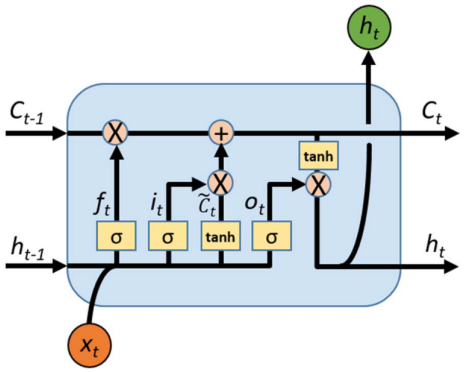


FIG. 3. Structure of LSTM Cell [6]

A rolling forecast scenario or walk-forward model validation is used wherein each time step of the test dataset will be walked one at a time. The model will be used to make a forecast for the time step. As and when the true blood glucose value arrives, it will be made available to the model for the forecast on the next time step.

This prediction model was built using the high-level neural networks API Keras version 2.7.0 in the Python 3.7.12 environment. The neural network consisted of a single LSTM layer with 70 units, and a dense layer with 12 units, i.e. the output layer. The output layer was used to return a 60 minute blood glucose prediction horizon. The number of train epochs was fixed as 20 iterations. The mean squared error objective is minimized using Adam and a learning rate of 0.001 is chosen. Implementation details can be found here.

### B. UVA/Padova Model

The UVA/Padova simulator is a physiological model that describes the interaction of glucose and insulin within different subsystems in the human body and returns the instantaneous state of the variables in those subsystems. The subsystems broadly consist of the glucose subsystem and the insulin subsystem. The model considers different glucose and insulin fluxes represented in Fig 4.

The scipy library in python was used to integrate ODEs and implement this model.

## VI. EXPERIMENTAL RESULTS

We used root-mean-squared error (RMSE) to calculate the difference between the CGM values and predicted BG values. It is given by

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\Big(y_i - f(x_i)\Big)^2}$$

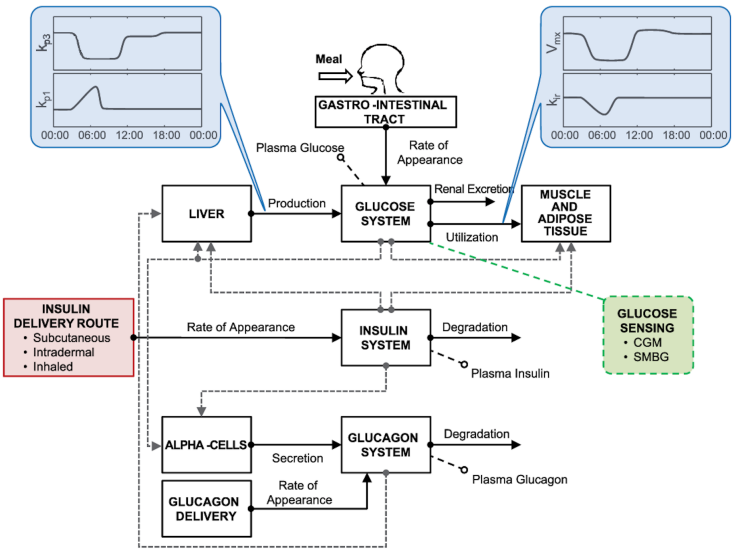


FIG. 4. Schematic diagram of all processes and subsystems in the UVA/Padova Simulator [9]

where $f(x_i)$ denotes the model's prediction, $y_i$ denotes the CGM reading and n is the number of datapoints. The results for the LSTM model can be found in table 1.

| LSTM Model | RMSE in mg/dl |
|---|---|
| Train Set | 11.9214 |
| Validation Set | 13.4585 |
| Test Set | 13.7520 |

Table 1. Root Mean Squared Error on the LSTM Model.

Next we compared the accuracy of the LSTM model and the UVA/Padova model. Since the UVA/Padova model is designed to simulate the postprandial state, we only considered test data points with carbohydrate intake. The results can be found in Table 2.

| Model | RMSE in mg/dl |
|---|---|
| LSTM | 17.7921 |
| UVA/Padova | 33.4669 |

Table 2. Root Mean Squared Error in the postprandial state.

Root-mean-squared error, however, may not be the best measure of prediction accuracy. You may have a lower RMSE by getting 100% accuracy for some readings and substantial error for other readings, which could be catastrophic for treatment decisions. If such an algorithm is deployed in the real world, the the user interface will consist of a 60 minute graph forecast. Therefore, we plotted random 60 minute graphs for the LSTM and UVA/Padova model and compared their accuracy with the CGM graphs found in Fig. 5 and 6.

Like we discussed above, ultimately all that matters in blood glucose predictions is not their absolute value but whether they can help make the right treatment decisions. Therefore, it makes sense to measure their clinical accuracy. To quantify the clinical accuracy of the considered predictors, we used the

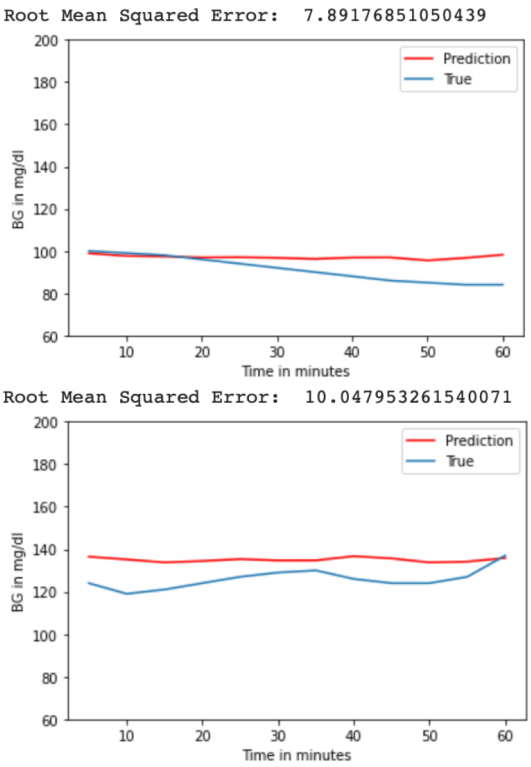


FIG. 5. 60 minute LSTM Forecasts.

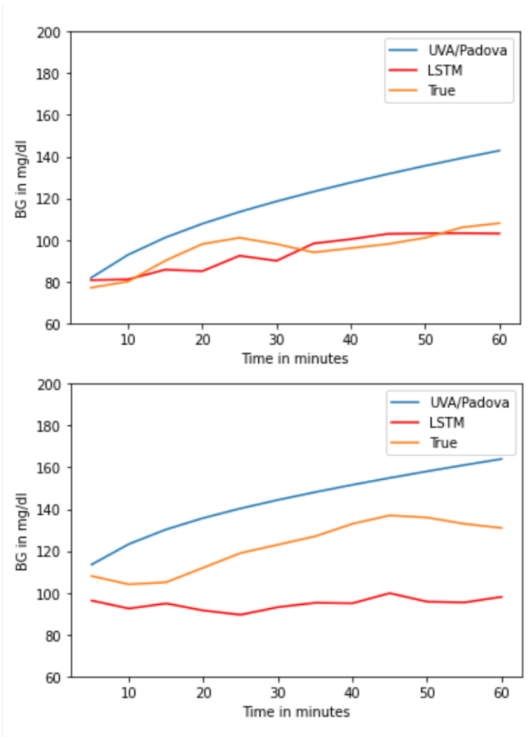


FIG. 6. LSTM & UVA/Padova prediction comparison in the postprandial state.

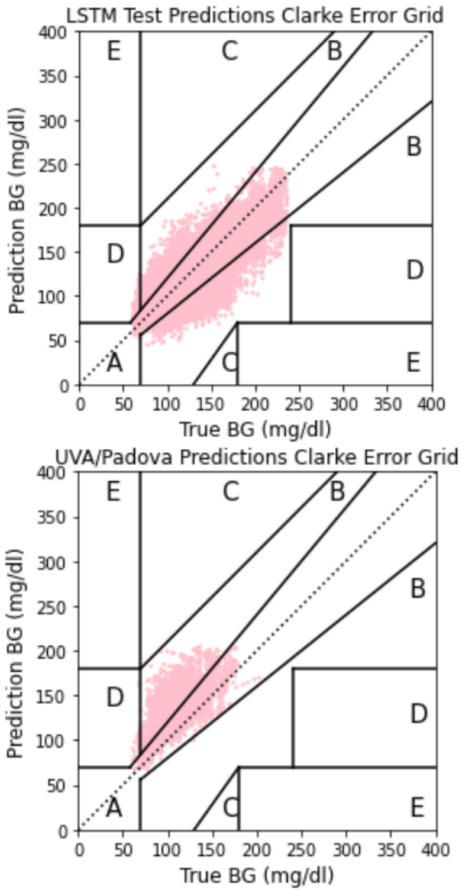


FIG. 7. Clarke Error Grid Analysis.

Clarke Error-Grid Analysis[10]. The Clarke Error-Grid Analysis is designed especially for glucose predictors. It classifies each prediction into one of the following 5 zones:

- Zone A are those prediction values within 20% of the sensor reading.
- Zone B contains points that are outside of 20% but would not lead to inappropriate treatment.
- Zone C are those points leading to unnecessary treatment,
- Zone D are those points indicating a potentially dangerous failure to detect hypoglycemia or hyperglycemia, and
- Zone E are those points that would confuse treatment of hypoglycemia for hyperglycemia and vice versa.

For the LSTM model, approximately 92% of the readings lie in zone A, 7% in zone B and less than 1% in zone D. None of the readings lie in zone C and E.
For the UVA/Padova model, approximately 45% of the readings lie in zone A, 53% of the readings lie in zone B and less than 1% in zone D. Here too, none of the readings lie in zone C and E.

## VII. CONCLUSIONS

In line with previous studies, we find the LSTM blood glucose prediction model to be remarkably accurate. Moreover, our LSTM model outperformed other LSTM models for a 60 minute prediction model [6]. This could be because of three reasons. First, our model uses a personalised approach and caters to a specific individual. Second, we have considered carbohydrates and bolus insulin as features in our model along with CGM readings. Third, the T1D dataset maintains great glycemic control. The T1D's routine seems to stay similar each day and blood glucose values normally lie in a narrow range, making it easier to predict values.

The UVA/Padova model, on the other hand, has a RMSE twice that of the LSTM model (in the postprandial state). This is because the simulator is built using fixed parameters and does not account for fluctuations in insulin sensitivity and meal absorption rates between different individuals.

Moreover, there are other factors that influence blood glucose levels such as time of day which our LSTM model may be able to catch but the UVA/Padova model does not incorporate. Biomedical simulators are developed to be accurate in high controlled clinical settings. Real-world blood glucose data, however, does not follow smooth curves and is noisy.

## VIII. FUTURE DIRECTIONS

While the LSTM model is accurate in making predictions on the test dataset, unlike the UVA/Padova simulator, it is not able to formulate a positive correlation between insulin-blood glucose and carbohydrates-blood glucose. This makes sense because insulin might not always decrease blood glucose and carbohydrates might not always increase blood glucose in our dataset (as a result of multiple external factors). To develop a clear correlation between insulin-blood glucose and carbohydrates-blood glucose, we can build a model which strikes a balance between purely statistical and purely physiological approaches. Such a fused model could inherit the realistic inductive biases from the physiological model and the flexibility and predictive power of modern machine learning sequence methods.[11]

Moreover, different kinds of food follow different carbohydrate absorption kinetics. The time that carbohydrates take to enter the bloodstream depends on the food's glycaemic index(a measure of how quickly a food causes our blood glucose levels to rise). Consuming 50 grams of carbohydrates from a slice of pizza and from a can of juice will elicit contrasting glycemic responses. Therefore, it is also important to develop a way to quantify carbohydrate absorption rates for different foods which can be incorporated in blood glucose prediction models.

## IX. REFERENCES


[1] Richard N Bergman. Minimal model: perspective from 2005. Hormone Research in Paediatrics, 64(Suppl. 3):8–15, 2005.

[2] Chiara Dalla Man, Francesco Micheletto, Dayu Lv, Marc Breton, Boris Kovatchev, and Claudio Cobelli. The uva/padova type 1 diabetes simulator: new features. Journal of diabetes science and technology, 8(1):26–34, 2014.

[3] S. Shanthi, "Advanced Engineering Informatics A novel approach for the prediction of glucose concentration in type 1 diabetes ahead in time through ARIMA and differential evolution," vol. 38, pp. 4182–4186, 2011.

[4] E. Daskalaki, A. Prountzou, P. Diem, and S. G. Mougiakakou, "Real-Time Adaptive Models for the Personalized Prediction of Glycemic Profile in Type 1 Diabetes Patients," Diabetes Technol. Ther., vol. 14, no. 2, pp. 168–174, 2012.

[5] R. Bunescu, N. Struble, C. Marling, J. Shubrook, and F. Schwartz, "Blood Glucose Level Prediction Using Physiological Models and Support Vector Regression," 2013 12th Int. Conf. Mach. Learn. Appl., pp. 135–140, 2013.

[6] Q. Sun, M. V. Jankovic, L. Bally and S. G. Mougiakakou, "Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network," 2018 14th Symposium on Neural Networks and Applications (NEUREL), 2018, pp. 1-5.

[7] https://loopkit.github.io/loopdocs/

[8] Z. Mahmoudi et al., "Application of the continuous-discrete extended Kalman filter for fault detection in continuous glucose monitors for type 1 diabetes," 2016 European Control Conference (ECC), 2016, pp. 714-719.

[9] Visentin, R., Campos-Náñez, E., Schiavon, M., Lv, D., Vettoretti, M., Breton, M., Kovatchev, B. P., Dalla Man, C., Cobelli, C. (2018). The UVA/Padova Type 1 Diabetes Simulator Goes From Single Meal to Single Day. Journal of Diabetes Science and Technology, 12(2), 273–281.

[10] Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL: Evaluating clinical accuracy of systems for self-monitoring of blood glucose.

[11] Miller, A.C., Foti, N.J. amp; Fox, E.. (2020). Learning Insulin-Glucose Dynamics in the Wild. Proceedings of the 5th Machine Learning for Healthcare Conference, in Proceedings of Machine Learning Research 126:172-197.