Glucose Levels Forecasting using Online Machine Learning

Bruno Carmo
57418
bm.carmo@campus.fct.unl.pt

Sahil Kumar
57449
ss.kumar@campus.fct.unl.pt

Problem

According to the Centers for Disease Control and Prevention of the United States (CDC 2023), diabetes is a chronic (long-lasting) health condition that affects how the human body turns food into energy. The human body cracks down most of the eaten food into sugar (glucose) and releases it into the bloodstream. In order to use glucose as energy for the body, the pancreas releases insulin to regulate the glucose levels in the bloodstream. With diabetes, the body does not make enough insulin or cannot use it as well as it should. As such, glucose can accumulate in the bloodstream and over time, that can cause serious health problems.

The WHO, indicates that the number of people with diabetes rose from 108 million in 1980 to 422 million in 2014 (WHO 2023) and in 2019, diabetes was the 9th highest cause of death globally with a 70% increase in deaths since 2000 (WHO 2020).

Therefore, it is essential to effectively control glucose levels in individuals with diabetes. However, measures of glucose levels taken by monitoring devices reflect the glucose level at the recent past and the moment, which is not useful to prevent any glucose anomaly that might happen in the near future. As such, by using data collected by a continuous glucose monitoring (CGM) device, we try to forecast the glucose level of an individual in the next 30 minutes after the measurement taken by the CGM.

State of the Art

In relation to our problem, there are a lot of studies that focus on forecasting glucose levels, from which we refer the most interesting ones.

In (Della Cioppa et al. 2023), the authors propose a Grammatical Evolution algorithm to induce personalized and interpretable forecasting glucose models assessed with the Prediction-Error Grid Analysis metric to satisfy both clinical and numerical requirements of the estimated predictions.

The authors of (Cappon et al. 2023) advance with a blood glucose prediction algorithm based on a personalized physiological model inspired by the UVA/Padova T1D Simulator. Also, they present a comparison of white-box and advanced black-box personalized prediction techniques.

In (Faccioli et al. 0), the authors put forward a glucose-specific mean square error (gMSE) cost function that is used in autoregressive integrated moving average with exogenous input (ARIMAX) models to show that real-time hypoglycemia forecasting can be improved.

The authors of (De Paoli et al. 2021) devoloped a Jump Neural Network in order to overcome the issue of predicting blood glucose values during physical activity. The authors developed and tested three learning configurations: offline training, online training, and online training with reinforcement.

Our work differs from most of the state of the art presented in this section because we focus more on simple online machine learning models.

Dataset

The dataset we used for this project consists of time series data, a collection of observations taken at successive equally spaced points in time. The observations collected are measurements of blood sugar in mg/dL of a person with diabetes, taken at a 5 minute interval during 12 days, from the 17th of October of 2018 to the 29th of October of 2018. Unfortunately, there are 301 null values during these 12 days, but given the scale of the data we decided that the dataset is appropriate for our problem. This dataset has 2 columns that contain the timestamp of the measurement and the glucose amount in mg/dL and is available here.

Exploratory Data Analysis

Although the recommended glucose levels change throught the day, depending if a person is fasting or just ate, a general way to check if a person with diabetes has a "good" level of sugar in their blood is to check if their glucose level is in between 70 and 180 mg/dL. With this in mind we can visualise the frequency distrubution of the glucose levels of subject during the 12 days of the experiments through a histogram in figure 1.

As we can see our test subject mostly stays within the recommended values but there are many occurrences of excess or deficit in glucose levels.

We can visualise the evolution of glucose levels throughout the experiment by using a color coded line graph as in figure 2.

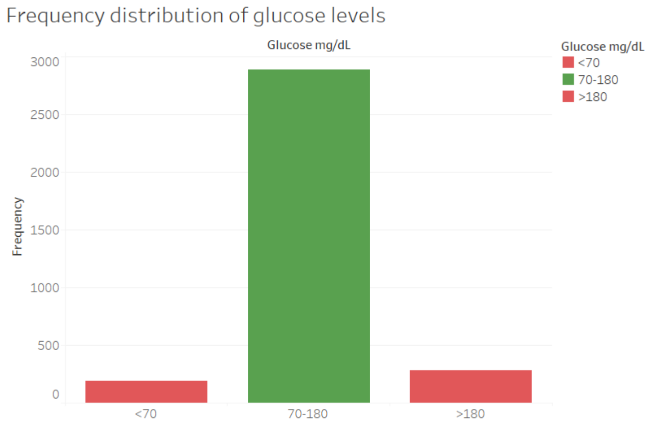

Figure 1: Frequency distribution of glucose levels

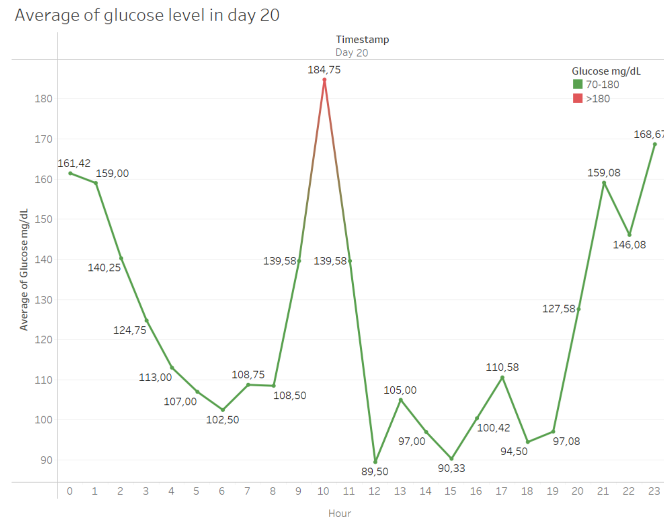

Figure 3: Glucose levels during the experiment

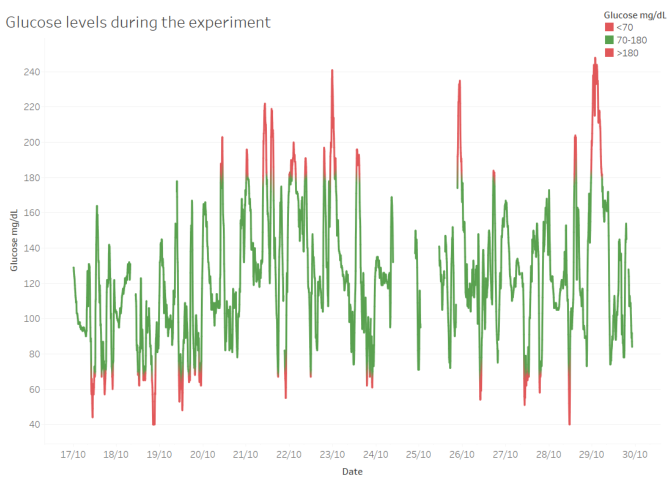

Figure 2: Glucose levels during the experiment

The plot reinforces the idea of the previous plot that most of the time, the patient has the glucose levels within the recommended values. However by analysing this plot, we can see that most of the days the patient enters in hyperglycemia at some interval of the day. The days where this does not occur, the patient enters in hypoglycemia.

Using the visualisation in figure 3 we can analyse the levels of glucose in a more detailed fashion. Here we can observe the average levels of glucose per hour during the twentieth day of October, using a a color coded line graph.

This plot reveals that during the 20Th the subject mostly stays within the recommended glucose values, although at 10 o'clock the subject enters in a state of hyperglycemia, possibly because of a late breakfast.

Breaking down even further the analysis during the day in figure 4 we can observe the average glucose values on the 13Th and 14Th hours of the twentieth first of October.

With this plot we can deduct that the glucose level starts to increase in the the 13Th hour reaching a state of hyperglycemia for most of these two hours. These high levels of glucose in the patient's blood could be related to the fact that these two hours are usual lunch hours.

To better understand our dataset in the context of machine learning in data streams we can do a Drift Detection analysis. Drift detection is the process of detecting changes in data over time. It is used to identify when the data being used by a machine learning model has changed significantly from the data that was used to train the model. This is important because if the data being used by a model has changed significantly, the model's predictions may no longer be accurate. In our problem this is a very important topic to assess because it could very well affect the well being of the subjects.

There are several methods that can be used to adapt to drift detection. One such method is called Adaptive Windowing (ADWIN), which is a distribution-based method that uses an adaptive sliding window to detect concept drift based on changes in data distribution.

By using the River API (Montiel et al. 2021) we can take advantage of the ADWIN method for concept drift detection, to visualize the possible occurrences of Concept Drift in the first 1000 examples from our data in figure 5. The plot demonstrates that there are various occurrences of Concept Drift in the first 1000 examples. This could be solved by using Concept Drift adaptation strategies such as simple retraining, ensemble retraining or model adjusting.

Research Question

Thus, according to the previous sections our research question is the following: Can we accurately, effectively and efficiently forecast glucose levels for the next 30 minutes based on previous glucose levels?

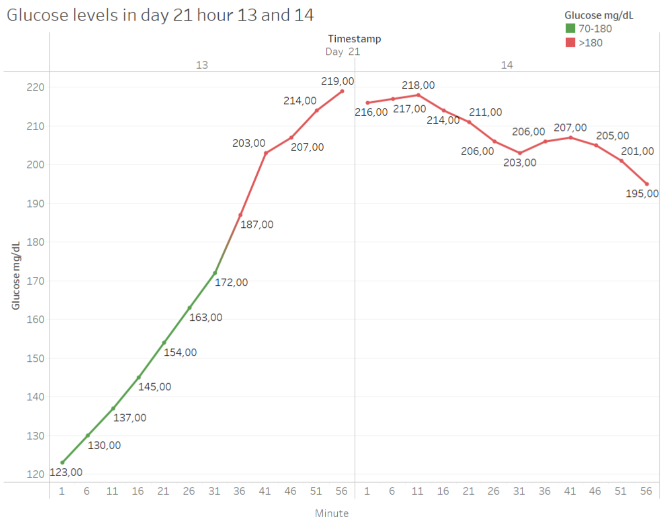

Figure 4: Glucose levels during the experiment

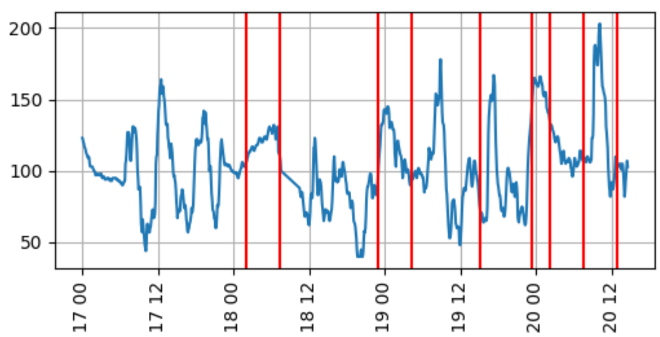

Figure 5: ADWIN drift analysis of the dataset in the first 1000 examples

Methodology

Taking into account that our work focuses on a streaming context, where new data points are constantly arriving, we used the River API to adapt our dataset to this context and to train machine learning models according to this paradigm. Therefore, taking advantage of the API, before any data point is fed to the models, we standardized it, such that, the glucose values got re-scaled to a range of values where the mean and the standard deviation are 0 and 1, respectively.

Additionally, in order to get better performance and reduce the error of the models, we added a new calculated feature: The average of the previous glucose levels per hours.

Considering our research question, our problem consists in a regression problem where we try to predict a real value. Therefore, we used four models from the River API that focus on regression: Linear Regression, Bagging Regressor, ARF Regressor and the MLP Regressor.

The Linear Regression model, as its name states, consists in doing linear regression on the streaming dataset. To compare results, we used a Linear Regression model with the Stochastic Gradient Descent optimizer and another Linear Regression model with the Adam optimizer.

The Bagging Regressor consists in an ensemble model for regressors, combining N independent regressor models. As such, to perform regression on the streaming dataset, the ensemble outputs the average of the outputs from each of the N models. In order to learn, the N models are updated for each data point according to a weight that is sampled from a Poisson distribution of parameter 1.

The Adaptive Random Forest (ARF) Regressor is a random forest model for regressors, combining N independent decision trees, drift detectors for each of the trees and adaptive operators. This regressor permits parallel computation, which leads in a decrease in prediction and learning time.

The Multi-layer Perception (MLP) Regressor is a model that uses neural networks to perform regression. Even though the River API states that there is still work in progress concerning this model, we decided to use it because it could give interesting results. The architecture we used for the model is depicted in figure 6.

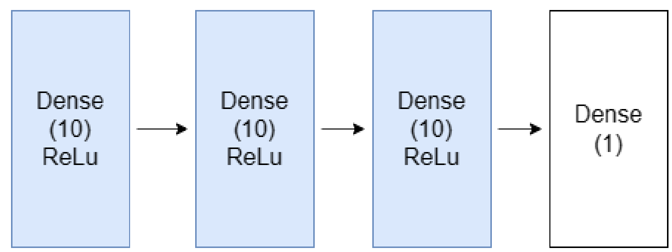

Figure 6: MLP architecture

We used this simple architecture with less hidden layers and less neurons because of the streaming context where prediction speed and learning speed is highly important.

Results

We obtained our results by performing various experiences according the previous section.

For the Linear Regression model, using the Stochastic Gradient Descent optimizer, after running many experiments, we arrived at the conclusion that the best learning rate was 0.01 and noticed that the Mean Average Error would increase exponentially with the increase of learning rates. In contrast, using the Adam optimizer, learning rate that achieved the best results was 0.1. One thing to note is that the Mean Average Error would decrease, with both optimizers, as the data samples were processed.

In order to have the best performance using the The Bagging Regressor we use 10 Linear Regression models with the Adam optimizer using a 0.1 learning rate. As in the last models we could observe that the Mean Average Error would decrease as the samples were processed.

Regarding the Adaptive Random Forest (ARF) Regressor, we achieved the best results using 10 models with the ADWIN drift detector using a delta equal to 0.005. Here we noticed that the Mean Average Error would flutuate between lower and higher values as the samples were processed by the model.

Moving on to the Multi-layer Perception (MLP) Regressor we achieved the best results, using the aforementioned network, and an Adam optimizer using a learning rate of 0.01. Here the Mean Average Error would decrease as the samples were processed by the model.

Model	MAE
Linear(SGD)	17.42
Linear(ADAM)	17.70
Bagging(Linear)	17.69
(ARF)	16.54
(MLP)	31.66

As we can observe in the table the Linear Model using different optimizers performed similarly. The Bagging Regressor, combining multiple linear models using the ADAM activation, performed slightly worse than the singular counterpart. The best performing model for our dataset was the Adaptive Random Forest (ARF) Regressor having a Mean Average Error of 16.54, and the model with the most error was the Multi-layer Perception (MLP) Regressor with 31.66 Mean Average Error.

Conclusion

We used four different online machine learning models from the the River API to answer our research question. Each of the models had a different type, such as, linear regression, ensemble methods, random forests and neural networks.

However, none of the models could get a low Mean Average Score. Taking into consideration, that is of extremely big importance to get a good accurate forecast of the glucose levels, none of the models could be used to accurately predict the glucose levels for the next 30 minutes.

Unfortunately, we conclude that, taking into account the problem described here, using the River API, and solely based on previous glucose levels, we cannot accurately, effectively and efficiently forecast the glucose levels for the next 30 minutes.

Probably, if we had more features, we could lower the Mean Average Error, but that stays for future work.

References

[Cappon et al. 2023] Cappon, G.; Prendin, F.; Facchinetti, A.; Sparacino, G.; and Del Favero, S. 2023. Individualized models for glucose prediction in type 1 diabetes: Comparing black-box approaches to a physiological white-box one. *IEEE Transactions on Biomedical Engineering*.

[CDC 2023] CDC. 2023. What is diabetes?

[De Paoli et al. 2021] De Paoli, B.; D'Antoni, F.; Merone, M.; Pieralice, S.; Piemonte, V.; and Pozzilli, P. 2021. Blood glucose level forecasting on type-1-diabetes subjects during physical activity: a comparative analysis of different learning techniques. *Bioengineering* 8(6):72.

[Della Cioppa et al. 2023] Della Cioppa, A.; De Falco, I.; Koutny, T.; Scafuri, U.; Ubl, M.; and Tarantino, E. 2023. Reducing high-risk glucose forecasting errors by evolving interpretable models for type 1 diabetes. *Applied Soft Computing* 134:110012.

[Faccioli et al. 0] Faccioli, S.; Prendin, F.; Facchinetti, A.; Sparacino, G.; and Favero, S. D. 0. Combined use of glucose-specific model identification and alarm strategy based on prediction-funnel to improve online forecasting of hypoglycemic events. *Journal of Diabetes Science and Technology* 0(0):19322968221093665. PMID: 35611461.

[Montiel et al. 2021] Montiel, J.; Halford, M.; Mastelini, S. M.; Bolmier, G.; Sourty, R.; Vaysse, R.; Zouitine, A.; Gomes, H. M.; Read, J.; Abdessalem, T.; et al. 2021. River: machine learning for streaming data in python.

[WHO 2020] WHO. 2020. The top 10 causes of death.

[WHO 2023] WHO. 2023. Diabetes.