Technology and Diabetes Management: Meat Detection

December 2022

Contents

1 Introduction

1.1 Diabetes recap

The body does not produce enough insulin or can't use it as well as it should. When there isn't enough insulin or cells stop responding to insulin, too much blood sugar stays in the bloodstream. Over time, that can cause serious health problems, such as heart disease, vision loss, and kidney disease. There are 2 main types of diabetes.

Type 1 diabetes is thought to be caused by an autoimmune reaction (the body attacks itself by mistake). This reaction stops the body from making insulin. Approximately 5-10% of the people who have diabetes have type 1. Symptoms of type 1 diabetes often develop quickly. It is usually diagnosed in children, teens, and young adults. If one has type 1 diabetes, one will need to take insulin every day to survive. Currently, no one knows how to prevent type 1 diabetes.

With type 2 diabetes, the body does not use insulin well and can't keep blood sugar at normal levels. About 90-95% of people with diabetes have type 2. It develops over many years and is usually diagnosed in adults (but more and more in children, teens, and young adults). One may not notice any symptoms, so it is important to get the blood sugar tested if one is at risk. Type 2 diabetes can be prevented or delayed with healthy lifestyle changes, such as:

- Losing weight
- Eating healthy food
- Being active

1.2 Current research

For important fields in diabetes, such as insulin dose prediction and diabetes prediction, machine learning is widely applied. Maintaining the insulin level without human interaction, such as meal announcements by the users, are still quite a challenge. An important challenge to overcome is incorrect information. That incorrect information could jeopardize the control system. Data could be incorrect due to the fact that the user has to estimate the amount he took in. Or the user may even forget to put it in entirely or at put the information into the system after a delayed time span. Elimination of this need to enter manual information can be achieved through the integration of automatic meal detection modules.

1.3 Project aim & basic approach

A method will have to be be created to detect a meal if the patient forgot to manual entry it. Using machine learning for this task is an overkill, although a good exercise.

The clinical importance of automatic meal detection is to simplify the patient's life and fool-proof the data to ultimately help the patient.

Our approach was and is to develop a simple model fast, adjust the hyperparameters and hope for the best. Then switch to more advanced models, if we are satisfied.

2 Related Work

In this section related work is summarized.

2.1 Reinforcement Learning-Based Adaptive Insulin Advisor for Individuals with Type 1 Diabetes Patients under Multiple Daily Injections Therapy.

The paper talks about an improvement to an already existing basal-bolus advisor (ABBA). The advisor was developed to help patient that are under insulin therapy with multiple daily injections. Some in silicio experiments were conducted using the DMMS.R simulator, likely similar data than the one used in this report. They used a combined approach of self-monitoring through blood glucose and insulin injection devices, such as an insulin pen. The DMMS.R dataset was used to validate that said approach outperform the conventional method. It does indead increase the time spent within the target range, as well as reducing the risk of hyperglyacimic and hypoglycaemic events.

2.2 Deep Physiological Model for Blood Glucose Prediction in T1DM Patients

For Type 1 Diabetes patient it is important to have accurate estimation for the near future. Control algorithms in glycose regulating systems are heavily dependend on those accurate predictions. There have been countless studies to provide better predictions for those glucose levels, categorizable into two major families: Those that are based on tuning glucose physiological-metabolic models and those that are based on machine learning. State of the art is discussed in this paper and in the end a hybrid-model is proposed using LSTM

2.3 A model for early prediction of diabetes.

Diabetes as a chronic and common disease is important to the authors. Prediction of diabetes can lead to improved treatment. Therefore techniques, such as data mining are widely used for prediction at early stages. There findings suggest a very strong association of diabites with the body mass index (BMI) an with glucose level.

2.4 A Deep Learning Framework for Automatic Meal Detection and Estimation in Artificial Pancreas Systems

State of the ard articial pancreas systems are hybrid closed-loops. They require manual meal announcement to control efficiently. This, as one can imagine, poses a challenge to users, because they have to stay engaged. The papers Authors propose moving to fully automated closed-looop glucose control via an algorithm that is based on deep learning.

2.5 Detection of a meal using continuous glucose monitoring: implications for an artificial beta-cell

A novel meal detection algorithm was developed to be used as a part of an artificial *beta*-cell that uses CGM. Multiple meal events using record of multiple children were used. Breakfast insulin was withheld for 1h to obtain glucose levels every 10 minutes after the meal during a pilot study. A reader monitored CGM. Based on that data one could be build a automated insulin dosing apparatus, that will responde to meals.

3 Methods

The data was loaded and unnecessary information thrown away. Kept were only patient name, minutes, CGM, and bolus data. Then the data was put into time windows, and binarized to 0/1 depending if bolus data !=0 per window.

3.1 Data and Experiment setup

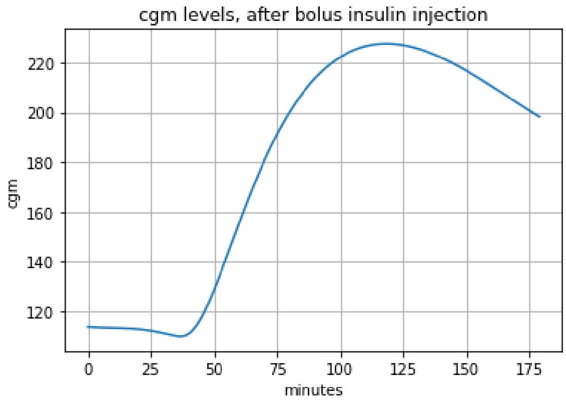
Simulated data for 30 individuals with type 1 diabetes has been provided. For each individual data for 1 month is provided. The data includes continuous glycose monitoring (CGM) measurements and insulin just as in a real-word scenario would be provided. The measurements are taken in a 1 minute interval.

3.2 Creating datasets pre-processing

To generate enough Data for the Model to learn from, multiple examples of each feature in a time series can be created. For simplicity let's say that even values in a series of numbers want to be detected. And the goal would be to detect an even number in a series of three. If the original dataset is [1,2,3,4,5,6], we would only have two sample [1,2,3], [4,5,6]. However if we vary the starting point we get [1,2,3],[2,3,4],[4,5,6]. So that's three samples instead of two, which is an increase by 50 % of the amount. Now this might not seem like much, however if the fraction $\frac{length(dataset)}{length(series)}$ is lower, than the amount of data for training will highly increase.

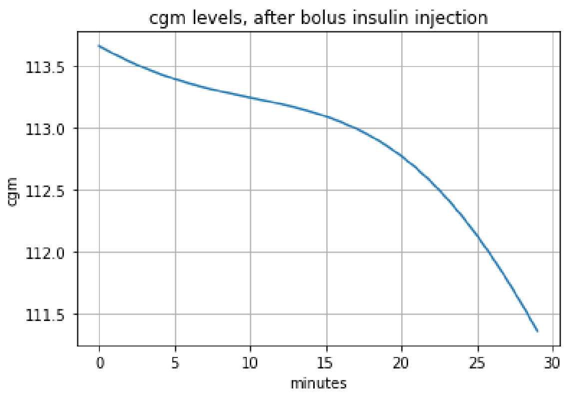
Now that the the data was handled for our model, the question became, how to create the labels corresponding to the data. The goal was to determine whether or not the patient, has eaten food, based on his glucose level.
As an Indication for that the insulin Bolus was taken. A diabetic patient should inject insulin 15 Minutes before taking the meal, because it can take around (15-30 Minutes) for the Insulin to act [1]. When eating, it also takes around 15 Minutes for blood sugar levels to rise (depending on the meal of course, for poor glucose it may rise faster). Then blood sugar slowly increase until it hits a peak in about 2 hours and then decrease back to baseline. [2]. Therefore it makes sense to take the insulin bolus as an indicator for when the patient is having a meal. Since the patient should take the insulin injection about 15 minutes before the meal. an offset of 15 minutes is added to detect the time when patients starts the meal. From this point on it takes about another 15 minutes for blood sugar levels to rise. Therefore the series length was selected to be 30 minutes, to make sure that the rising of blood sugar level is captured. In this time window the effects of lowering blood sugar due to insulin and heightening blood sugar due to meal intake, are both presents. This combination creates a pattern which indicates, that the person has taken a meal.

The Illustration below shows the blood sugar levels fro 3 hours after an insulin injection.

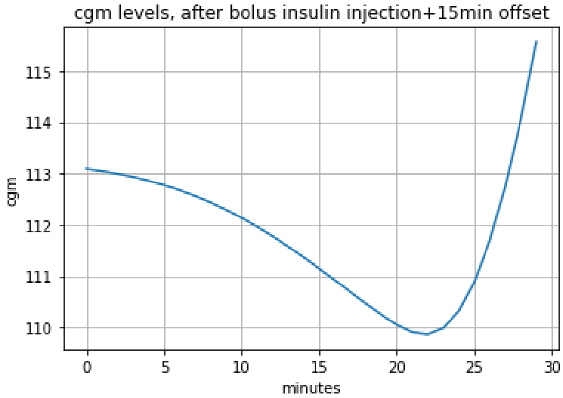


Here the behaviour of the blood sugar levels after, insulin intake can be observed very well. For the first few minutes, we can see the blood sugar staying at base level, while slowly starting to decrease, that's when the insulin starts working.

We can clearly observe when the meal was taken, indicated by local minima in the curve. So this is the pattern, which indicates food intake, that we want our model to discover. However a time-span of 3 hours (180min) would be too long to track. So if we decrease the series length to 30 minutes, the following pattern appears.



We can see, that a window of 30 minutes on might seem too short, because the local minima is not detected. It is for that reason, that, as mention before, an offset of 15minutes from the injection point on is added.



Now with an offset, the behaviour of the blood sugar level when a meal is taken, can be observed clearly, within a short time window. This is the pattern that the model aims to discover.
To improve the performance further, the data was then normalized with the z-score.

Such short series lengths, compared to the overall time span of measuring (1Month), inevitably creates imbalanced Data. To compensate, the values, where a meal was detected, were over-sampled using the tensorflow keras imblearn library. This created a more balanced dataset to train, thus decreasing the necessary training time to achieve sufficient results.

3.3 LSTM model

The short-term memory model (LSTM) is yet another type of articifical neural networks. The name comes from short-term memory and the model is a subvariant of recurrent artificial neural networks (RNNs).

LSTM networks are well suited for time-based data, according to literature. This is why it was used in this report. Also, in difference to feed-forward neural networks in has feedback connections, that allows for more than just image processing. Common task may be handwriting recognition or speech recognition.

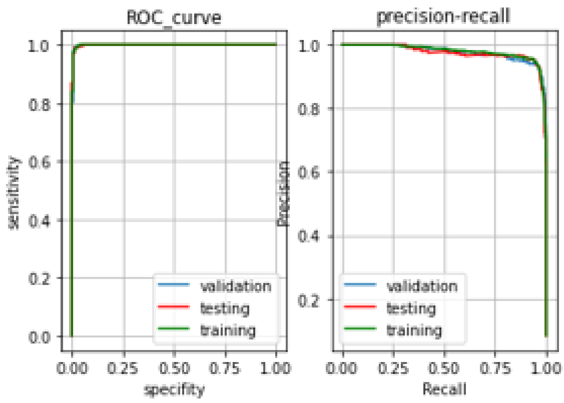
The LSTM contains a memory-cell, known as ”cell state”. It could be visualised as a conveyor belt through which information just flows, unchanged. The gates then add ord remove from the cell state. In difference to a general RNN, the LSTM makes only small modifications on the data.

The model was optimized not only for accuracy but also for precision, recall and f1 score. This was done so to ensure, that sufficient results for precision and recall are reached.

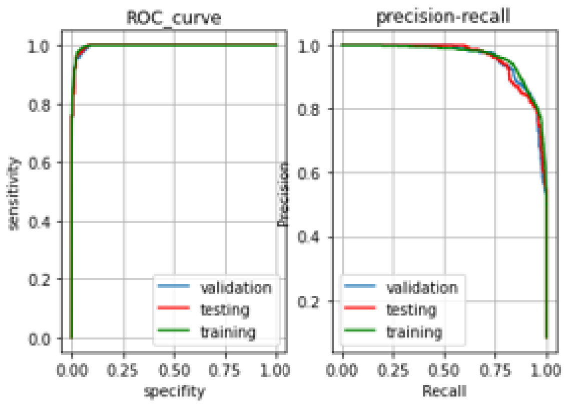
4 Results

The parameter optimization for the LSTM model and the training, greatly improved the quality of the results. Same as the z-score normalization and the minority oversampling. As well as not only training for accuracy. Together with choosing the right offset of 15 from the insulin bolus and the series length of 30 Minutes, the values of accuracy, sensitivity and specificity are in 90% range for, children, adolescents and adults. The values of precision and recall are in the 80% range. There is little difference between the results for the training, validation and testing sets.

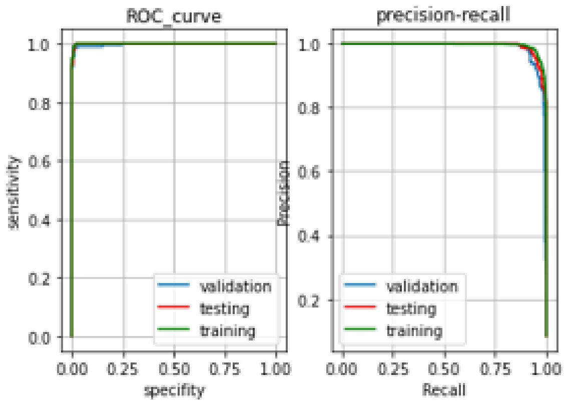
The illustration below shows the ROC curve and the precision recall curve, for a child with only 10 epochs of training.



Next the same training for one of the adolescents.

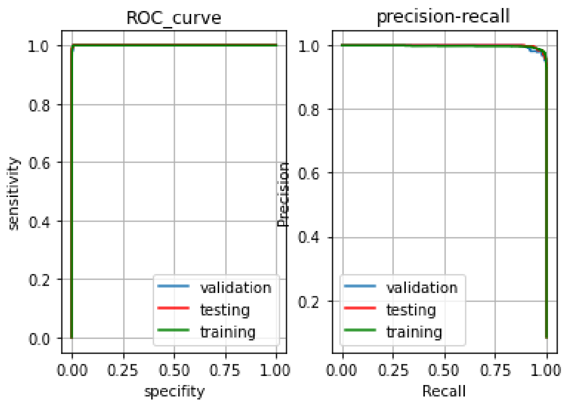


And the result for one one of the adults:



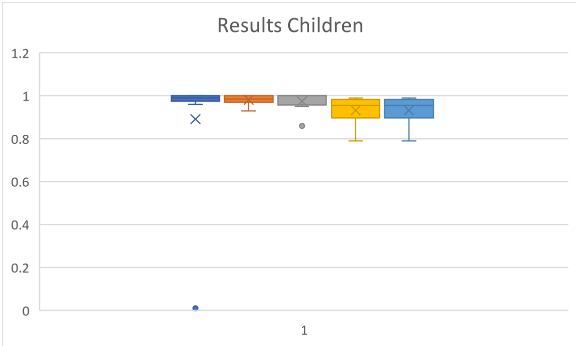
We can see that only training with 5 epochs, already leads to good results. Adolescent 5 however seems to be a more difficult case, but the number of epochs can easily be increased.

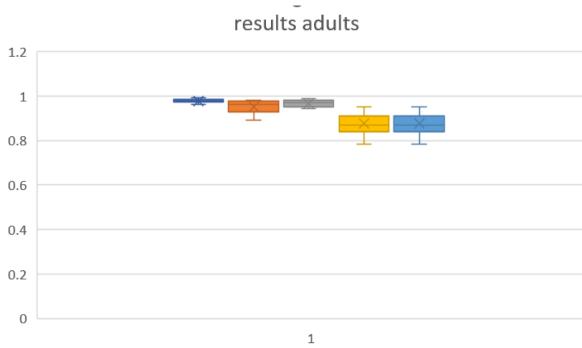
If they are for example increased to 30, the results for adolescent 5 look like this:

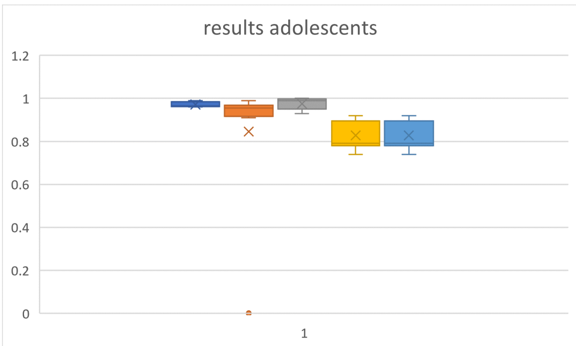


The scores for accuracy, specificity and sensitivity are are at 1 (rounded for 2 digits). While the scores for precision and recall are each at 0.98.

The following tables show the overall results for each group, children, adults and adolescents. All the results remain in the upper 90% range, excepts for a few outliers in the upper 80% to lower 90% range. Interestingly, the precision and recall scores are always the same, this will be further commented on at the end of the discussion part.







5 Discussion

The main challenge of this assignment was how to handle the datasets. For a long time no offset was taken from the data. Different methods were implemented to improve the results with series lengths of 15 and 30 Minutes. With a series length of 15 Minutes, even very long training, with lots of epochs, returned poor values for precision and recall. Meaning there were a lot of false positive and false negatives. With a series length of 30 minutes sufficiently good precision and recall values could be obtained, after long training times, which were around 80%. However adding an offset to to the time window, increased the quality and decreased the training time dramatically.

The results seems almost too good to be true. And the pattern described in the section data handling, could sometimes be observed even without an offset. Furthermore, sometimes it wano not observed int the data offset it was not observed. Nonetheless the meal detection performs significantly better with offset than without. And since the result are done on the testing and validation set, these high values can't be caused by overfitting.

So this begs the question, why the pattern can be learned so much better with offset, even when it's seemingly the same than without and vise versa in many cases.

Our assumptions is that there might be pattern which the model picks up on, that when as human don't pick up on immediately. For example this could the curvatures of the curve and the differences in the amplitude. In the end it is always unknown what exactly goes on in the model, but all the results suggest, that there is no performance error, since we evaluate for many different metrics. And on a practical note, since this is medical data, which could be used to determine eventual medication.

One other questions that remains, is why the precision for all samples are the same. As we can see in the precision-recall curves. it was not, that the same curve was calculated twice. The problem was, that the micro average instead of the macro average was calculated. Unfortunately this was only detected at the end of the project, however some sample tests, showes that the difference in precision and recall is only in the range of 2% points. And the overall scores for precision and recall remain in the 80-90% range.

6 Attachements

- https://github.com/bodging/DiabetsTechMeatDetection

References

[1] https://www.aafp.org/pubs/afp/issues/1999/0801/p649.html

[2] https://www.boost.com/blog/low-blood-sugar