TFG University of Granada

Mario García Jiménez

February 2023

Chapter 1

Acknowledgements

Contents

Chapter 2

Introduction

2.1 Thesis Goal

The objective of this thesis is to develop a Machine Learning model that allows predictions about blood glucose levels in patients with type 1 diabetes using data from real patients.

2.2 Context and Motivation

2.3 The challenge of Predicting Blood Glucose

The nature of blood glucose (BG) data makes prediction a tough task. BG values are the result of complex interactions between various physiological, hormonal and metabolic processes that are not yet fully understood. They are influenced by several factors such as age, weight, genre, diet, physical activity, illness, menstrual period, stress and medications.

Blood glucose levels can change rapidly and unpredictably in response to these factors. The changes can vary greatly between different patients and even within the same patient over time. Imperceptible differences in the routine can result in significantly different outcomes [**?**]. For these reasons, we can consider BG data series as highly non-linear.

2.4 Context

2.5 Comparison data between patients

Speak about the big difference in data between different patients.
Se puede poner mejor en la parte del dataset

Chapter 3

State Of Art

Todo: Add here a small introduction

3.1 Search Methods

Three different approaches were taken to gain an overview of the current state of the art:

3.1.1 Reviews

Three reviews were considered [2, 26, 32]. They provide a broad view of the state of the art, comparing several scientific publications based on used techniques, input datasets, objectives, and results. They also include a summary of the most interesting publications and comparison matrices. [32] was published in 2021 and [2, 26] in 2022, so the publications up to 2021 and the first part of 2022 are covered.

3.1.2 Personalized Query

The following query was used to cover the papers published in 2022 and 2023:

((TI= ((blood glucose))) AND (TS=((predict OR forecast) AND (neural network OR deep OR machine learning OR nonlinear) NOT (type 2 diabetes OR T2D or ECG))) AND (PY==("2023" OR "2022"))) AND (DT==("ARTICLE"))

On 6th March, Web of Science returned 12 results. Of these, [6, 28, 30, 14, 27] were rejected because they are clearly not relevant to the topic. At first glance, [24, 10] might seem to be relevant to the topic. However, they were also rejected after reading the abstract and introduction. [11] focuses on blood glucose prediction, but the dataset includes only patients with type 2 diabetes, so this paper was declined too. The rest of the results (4 papers) have been added to the matrix.

3.1.3 Publication References

As part of the comprehensive analysis of the state of the art, certain works cited by selected publications were also reviewed. This category includes works considered relevant by the author, previous works, complementary studies, contributed code, etc. This process helped to ensure the rigor and validity of the articles, as well as to understand the motivation behind specific choices made by the authors.

3.2 Blood Glucose Prediction Models

Several techniques have been tried over the years. In [7], Bunescu proposed a physiological model that uses mathematical equations to attempt to model the trend of blood glucose. Equation-based solutions can give acceptable results, but their formulation is cognitively demanding and not scalable: When new physiological parameters become available, the model has to be redesigned. The model outperformed physician predictions in a comparative study. This method can also be easily combined with others, as in [8], where it is used to model the effects of carbohydrates and insulin on BG values in combination with grammatical evolution technique.

Statistical approaches have also been widely employed. Traditionally, ARIMA has been the most commonly used technique for time series forecasting. In recent years, variants of ARIMA have been used in BG prediction, such as [29] which uses ARIMA with adaptive orders to try to capture the non-stationary changes of the CGM over time. In [22], ARIMA model was tested on the same dataset as thirty different linear and nonlinear predictive algorithms (including LSTM and feed-forward neural networks) and outperformed them all with results of 22.15 (mg/dl) RMSE error for 30 minutes prediction. [4, 31, 34] demonstrate that ARIMA is not the only statical used technique. Remarkably, Aljamaan [4] proposes a long-term prediction horizon of 6 days, while most of the works use prediction windows of 30-60 minutes, which are considered rigid and challenging to improve.

In the last few years, the most popular approaches have been based on neural networks. In [25], Idrissi suggests a 1-dimensional filter convolutional neural network (CNN) and tests it with 5 different forecasting strategies. The model outperforms an LSTM model also presented by the author, obtaining an RMSE of 8.68 (mg/dl) for a 30-minute prediction horizon. In [13], Li proposed a dilated CNN model. Dilated CNNs can increase the receptive field of the network without increasing the number of parameters, allowing them to capture temporal patterns over a wider range of time steps. In addition, Artificial Neural Networks (ANNs) and Autoregressive Neural Networks (AR-NNs) are examined in [32, 3] with poor results.

However, among the various machine learning models employed in blood glucose prediction, recurrent neural networks (RNN) have gained significant popularity and are considered the leading approach. RNNs are a powerful technique that implements the concept of "memory" to neural networks. The main problem with classical recurrent neural networks (RNN) is the vanishing or exploding gradient problem [5]. For this reason, most of the publications used variations of classical RNN units such as LSTM [17, 3, 18, 1, 16, 33, 19, 20] or dilated RNN [33]. The work of Mirshekarian [17] creates a personalized LSTM model per patient, improving the results of the previous work [7] in the same dataset. On the other hand, Aliberti [3] utilizes the LSTM model to study a large and heterogeneous cohort of patients and then applies it to completely new patients. The main objective of this work is to learn a generalizable prediction model that can be easily extrapolated to real-life applications.
Several works have taken the theoretical base of LSTM and modified it to try to get stronger models that give better results. Examples include [18] with a Memory-Augmented LSTM that allows searching for similar BG trends further back in time, [1] whose LSTM model has as input the past information and the suggested future information of insulin and meals or [20] that combines RNN with Restricted Boltzmann Machines (RBM).

Another approach worth mentioning is ensemble models. Ensemble models try to enhance the weaknesses of each algorithm by combining multiple models to improve the overall performance and accuracy of the predictions. In the literature, there are proposals of ensemble models conformed by LSTM and bidirectional LSTM models (bi-LSTM) [23, 21], CNN and LSTM [12] or autoregressive multi-output model with polynomial forecasting system [9]. The combination strategy of the results can vary widely between ensemble models, [21] explains three approaches and compares them.

3.3 Comparison Between BG Prediction Publications

Comparison between BG prediction publications is not an easy task. Two publications may differ in multiple points such as the input dataset, the measuring device, the type of model, the results metric, etc. Before making a comparison, it is important to bear these points in mind in order to avoid erroneous conclusions.

It is possible to divide the blood glucose prediction datasets into two broad groups:

1. **In-silico data:** Consists of computer-generated data. The most popular generator is the UVA/Padova T1D simulator [15]. Several works exploit this generator to complement the real patients' data [23, 12, 17, 13, 1, 33] or to create all input datasets [8, 4].

2. **Real patients data:** Data obtained from real patients under professional supervision using continuous glucose monitoring (CGM) technology. Not all the cited publications have open-source patient datasets [9, 29, 1], making it difficult to contrast results.

Input information can also include complementary information like carbohydrates values (meals), injected insulin, physiological statistics, medical tests, etc [18].

Models using both, generated and real patient data show that the prediction error is significantly lower for generated data than for real data. This could explain exceptional results in papers with only simulated data as [4], which presents less than 1 (mg/dl) RMSE for 15 and 30 minutes estimation windows. Moreover, in some publications the real patient dataset is extremely short: 3 days of 12 patients in [31] or 4 days of 9 patients in [19]. This casts doubt on whether these results can be extrapolated to larger and more diverse datasets. Finally, different datasets may use different measurement sensors, which can lead to differences in time between measurements, and more difficult to detect differences in the likelihood of gaps or the amount of added noise. This variability between publications explains why choosing the model with the lower error metric is not always the best idea. Therefore, it is crucial to consider all these factors when selecting the proper blood glucose prediction model for our concrete dataset.

Añadir aquí una comparativa de que metodos de error se utilizan (RMSE y la rejilla de Clark) y además meter lo que se ha conseguido en el estado del arte.

Chapter 4

Data Exploration

4.1 Description

110 different patients

4.2 First Data Visualization

The exploration of data has been supported by PowerBi plots. In PowerBi, I have imported the data and added a combined temporal column:
Time = Glucose_measurements_sample[Measurement_date] + Glucose_measurements_sample[Measurement_time]

TODO: Create a boxplot of patients to show the differences

TODO: Crar boxplot que muestren la diferencia de tendencias y de seasonality en los distintos pacientes https://sarit-maitra.medium.com/take-time-series-a-level-up-with-walk-forward-validation-217c33114f68

Chapter 5

Data Preprocessing

The preprocessing techniques employed by various interesting papers were reviewed in order to select the most useful ones for the current dataset (see matrix preprocessing). In summary, the techniques applied in this work consist of four steps:

1. Removing duplicate samples.
2. Removing outliers.
3. Resampling.
4. Normalization??

5.1 Removing Duplicate Samples

A duplicate sample refers to a reading with the same measurement time and BG value as another sample. No duplicate samples were found in the dataset.

5.2 Removing outliers

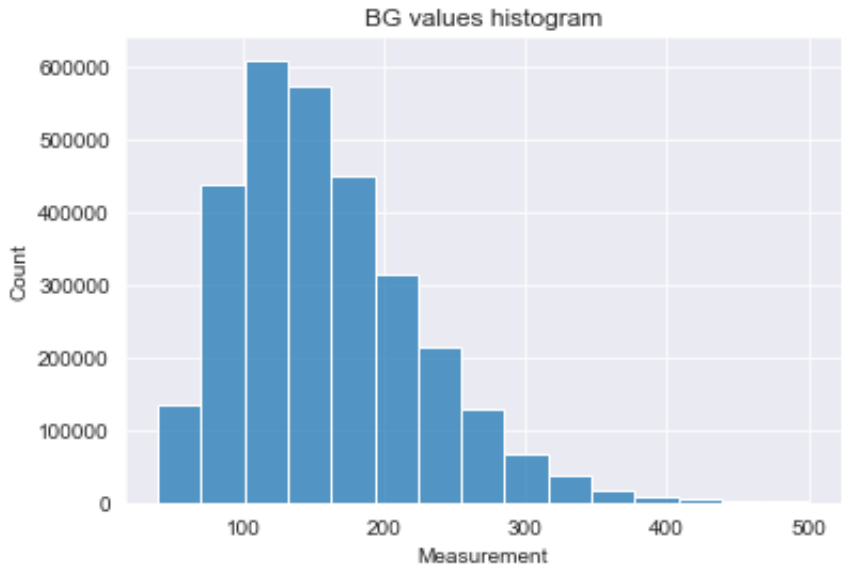
The used sensor measures BG values between 40 and 500 mg/dL, so all measurements in the dataset are within this range. If the BG of a patient were lower than 40 mg/dL or higher than 500 mg/dL, the device would register 40 and 500 mg/dL respectively. These two values are extreme and highly unusual, but possible. Therefore, it is not correct to consider isolated measurements within this range as unrealistic values.

Another strategy for finding outliers is to use relative differences between measurements. Relative differences between samples show how the BG value changes over a known time interval. It would not be feasible to consider a blood glucose reading of 40 mg/dl 15 minutes after a reading of 500 mg/dl for the same patient. Unfortunately, it is not possible to set a global threshold that separates realistic variation from outliers because BG changes are very patient and context-dependent. According to medical opinion (Miguel Quesada), the changes could reach 100 mg/dL over a period of 15 minutes, and even more abrupt changes in extreme cases such as an overdose of the medical prescription to combat hypoglycemia or hyperglycemia. The percentiles of the blood glucose differences have been calculated in order to contrast the medical claim in the dataset. $\mathbf{P}_{99.00}$ does not reach 50 mg/dL. $\mathbf{P}_{99.99}$ must be calculated to reach 80 mg/dL. There are 5,461 consecutive samples with differences greater than 100 mg/dL, which is less than 0.01% of the total dataset. These results show that the medical assumption is consistent with the tendency of the dataset studied.

A different approach would be to remove samples with readings close to the sensor extremes. These are rare readings and most measurements fall within a narrower interval because most of the time patients have healthy values of BG 5.1. The problem with this is that the prediction of BG is particularly important when the patient is inside hypoglycemia (less than 70 mg/dL) or hyperglycemia (more than 270 mg/dL) scenarios, and if this approach were followed, a lot of information from these scenarios would be lost.

None of these strategies were able to find clear outliers. The data exhibits a realistic behavior that, combined with the lack of other information such as meals, exercise, or insulin makes it impossible to judge whether the samples with large BG differences are outliers or

Figure 5.1: Blood Glucose Measurements Histogram for Complete Dataset



just extreme patient situations. Also, removing samples in time series could create gaps that break the continuity of the data. For these reasons, no samples were removed at this point.

5.3 Resampling

The used sensor takes an automatic measurement every 15 minutes. If the time between measurements is shorter, there is more information than needed for a considered interval. The reason could be anything from a sensor error to a patient taking several readings to test the device, to make sure their BG levels are stable, etc. If this time is longer, it receives the name of gap and there is a loss of information. The causes of missing values are always related to complications during data collection, such as sensor failure, both sensors running out of battery, or patients leaving the study and then returning. Whatever the reason, the used prediction models need equal time differences between measurements to give consistent predictions.

After the preprocessing, the dataset should have a sample every 15 minutes with no samples in between. This requires first a downsampling process to convert 15-minute intervals with more than one sample to 15-minute intervals with only one sample, and then an upsampling process to fill the gaps with readings.

5.3.1 Downsampling. Multiple Measurements at the Same Interval

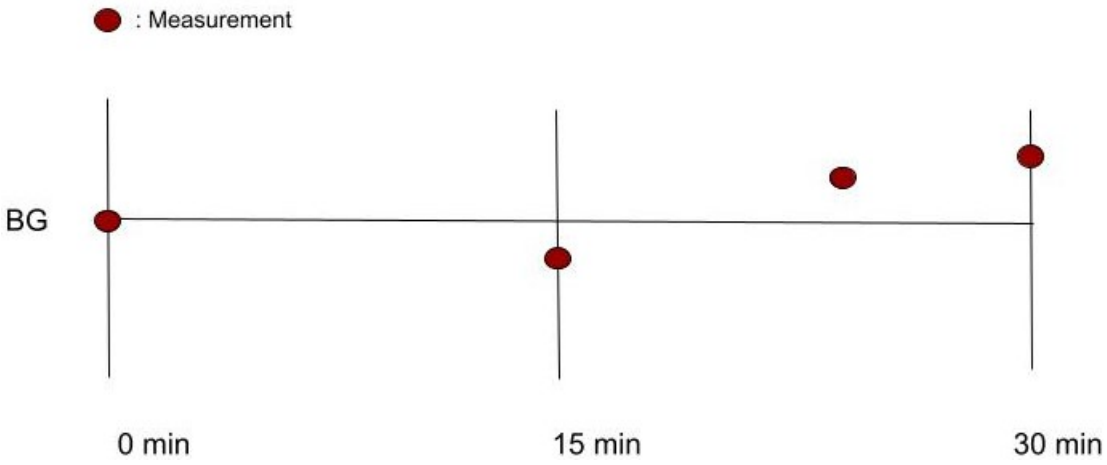
5.3.2 Longer Frequency Problems. Gaps Treatment

In the literature, two main strategies for dealing with gaps were followed: Simulate the missing values or split the dataset into two sub-datasets. The problem with simulating the missing data is that the techniques are not precise enough and the BG value can change rapidly, so for long gaps, the simulated data is likely to have little to do with the missing real data. Also, including generated data in the training data can lead to overfitting. The problem with splitting is that the continuity of the data is lost. The ideal strategy should be to simulate as few samples as possible and to maintain continuity in the data whenever possible.

In [23], the shortest gap needed to split the dataset is 5 days [1] long. This means that a gap of 4 days is fully simulated. This long simulation creates a trend that is totally different from the real one because the error is being added to each simulated measurement. The figure 5.5 shows an example of interpolation for a random time for a random patient of the dataset. The used interpolation method is polynomial with order 2. Results for simulations of 24 and

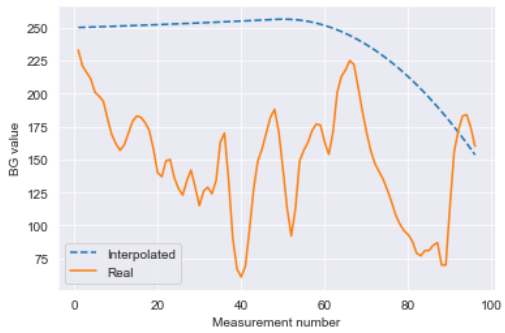
[1] 5 days correspond to 480 measurements in the studied dataset

Figure 5.2: Downsampling example

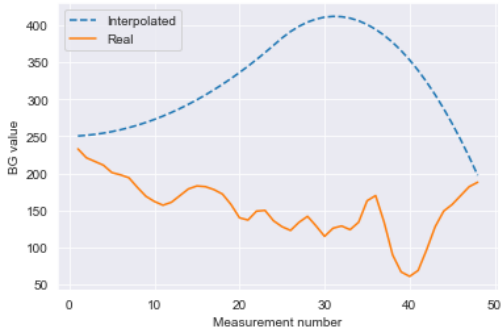


12 hours are poor. The results improve as the interpolation time decreases, obtaining moderate results for simulations of 6 hours (24 measurements) and good results for 3 hours (12 measurements). The same experiment was replicated for another patient with similar results 8.1.

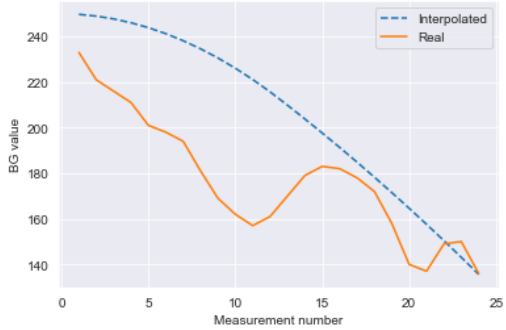
Figure 5.3: BG values interpolation for a random patient *LIB193367*



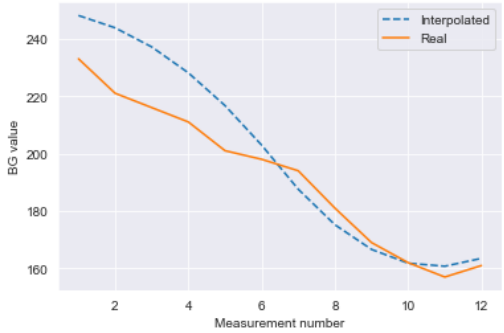
(a) 24 hours interpolation



(b) 12 hours interpolation



(c) 6 hours interpolation



(d) 3 hours interpolation

Probably interpolation is not the best approach for the dataset. The dataset has high-quality data and the inclusion of generated dummy data may lead to inaccurate results. The problem with simulated data in this context is that the BG tendency is highly non-linear and using a function to interpolate the new data would add unrealistic data behaviors. It is true that sometimes the simulation of some data is unavoidable in BG problems, but as a first approach, all missing data is kept as unknown. An upsampling procedure was used to generate samples with unknown measurements corresponding to the missing values.

5.4 Reframing the dataset

At this point, the current dataset shows the BG value for a patient at a specific time. For the 10 first samples:

Time	PatientID	Measurement
2020-06-09 19:30:00	LIB193263	86.00
2020-06-09 19:45:00	LIB193263	85.00
2020-06-09 20:00:00	LIB193263	85.00
2020-06-09 20:15:00	LIB193263	87.00
2020-06-09 20:30:00	LIB193263	88.00
2020-06-09 20:45:00	LIB193263	93.00
2020-06-09 21:00:00	LIB193263	106.00
2020-06-09 21:15:00	LIB193263	134.00
2020-06-09 21:30:00	LIB193263	165.00
2020-06-09 21:45:00	LIB193263	187.00

This time series must be transformed into a supervised learning problem to enable the forecasting of blood glucose values. The simplest way to do this is to predict the value at the next time given the value at the current time. The supervised learning problem with 1-shifted values looks as follows:

Time	PatientID	Measurement	Measurement-1
2020-06-09 19:30:00	LIB193263	86.00	-
2020-06-09 19:45:00	LIB193263	85.00	86.00
2020-06-09 20:00:00	LIB193263	85.00	85.00
2020-06-09 20:15:00	LIB193263	87.00	85.00
2020-06-09 20:30:00	LIB193263	88.00	87.00
2020-06-09 20:45:00	LIB193263	93.00	88.00
2020-06-09 21:00:00	LIB193263	106.00	93.00
2020-06-09 21:15:00	LIB193263	134.00	106.00
2020-06-09 21:30:00	LIB193263	165.00	134.00
2020-06-09 21:45:00	LIB193263	187.00	165.00

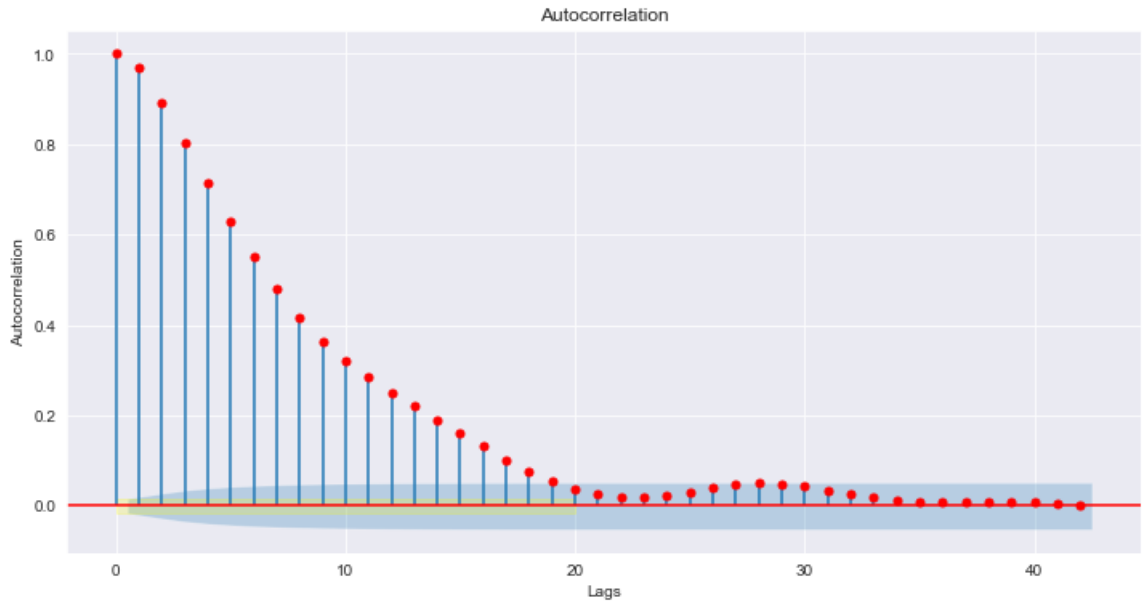
The rolling window size is the number of shifted values used to predict a new value. For each sample, the shifted values act as the input values and the current observation as the output or label. A number of shifted values corresponding to times of 30, 60, 90, and 120 minutes have been widely tested in the literature. This corresponds to 2, 4, 6, and 8 samples respectively for the utilized sensor. A common practice to study the strength and type of relationship between observations and their lags are autocorrelation plots 5.4. The yellow line shows that autocorrelation exists up to the 20-lagged sample. All of the considered rolling window sizes are smaller than 20, so the plot doesn't provide enough information to choose one of the rolling window sizes. In addition, the plot was computed only for one patient subdataset because it is a very expensive process. Considering this, the selection of the rolling window size will be done in the hyperparameter tuning process.

5.5 Missing Values Treatment

A gap is a period of time where data is missing. For this particular dataset, the time between two measurements should be a maximum of 18 minutes (15 minutes plus 3 minutes of assumed delay), but about 20,000 samples have a longer time difference with the lagged sample. In contrast to duplicate values, the causes of missing values are always related to complications during data collection, such as sensor failure, both sensors running out of battery, patients leaving the study and then returning, etc.

There are two main strategies for dealing with gaps: Simulate the missing values or split the dataset into two sub-datasets. The problem with simulating the missing data is that the techniques are not precise enough and the BG value can change rapidly, so for long gaps the

Figure 5.4: Autocorrelation plot for first subdataset of patient *LIB193367*



simulated data is likely to have little to do with the missing real data. Also, including too many generated readings in the training data can lead to overfitting. As for splitting, the problem is that the continuity of the data is lost. The ideal strategy should be to simulate as few samples as possible and to maintain the continuity in the data whenever possible. There are papers that don't treat gaps because they use algorithms that are robust to missing data but this is not the case for the proposed ones.

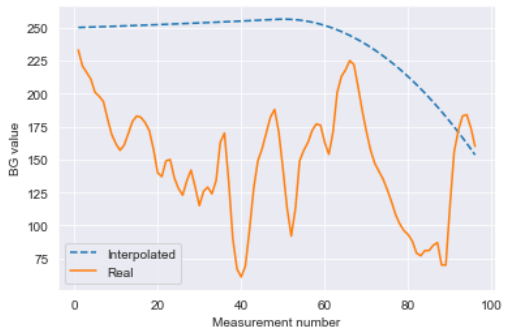
As a first approximation, the state of the art is reviewed.

In [12], the missing values strategy is to interpolate gaps shorter than 1 hour [2] and to split gaps longer than 1 hour into two segments. When the same process is applied to the dataset, longer gaps are so common that several subsamples per patient are obtained **??**. Before this experiment, the larger patient (*LIB193277*) dataset had 114,598 samples. After the experiment, the larger subsample of all patients has 2,603 subsamples. The figure **??** shows the differences between the number of measurements read for each patient and the number of measurements for their larger subsample. This is not desirable because the long and continuous data has been converted to several short datasets.

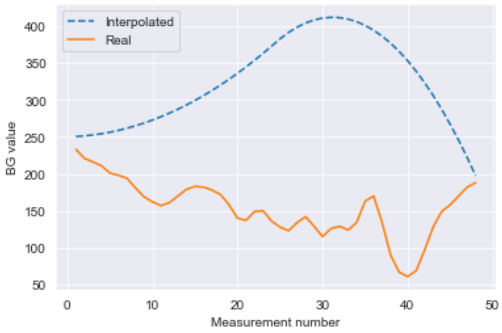
In [23], the shortest gap needed to split the dataset is 5 days [3] long. This means that a gap of 4 days is fully simulated. This long simulation creates a trend that is totally different from the real one because the error is being added to each simulated measurement. The figure 5.5 shows an example of interpolation for a random patient of the dataset. The used interpolation method is polynomial with order 2. Results for simulations of 24 and 12 hours are poor. The results improve as the interpolation time decreases, obtaining moderate results for simulations of 6 hours (24 measurements) and good results for 3 hours (12 measurements). The same experiment has been replicated for another patient and the results are similar 8.1.

[2] 1 hour corresponds to 4 points in the studied dataset.
[3] 5 days correspond to 480 measurements in the studied dataset

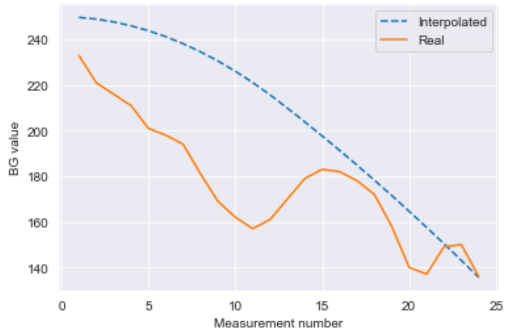
Figure 5.5: BG values interpolation for a random patient *LIB193367*



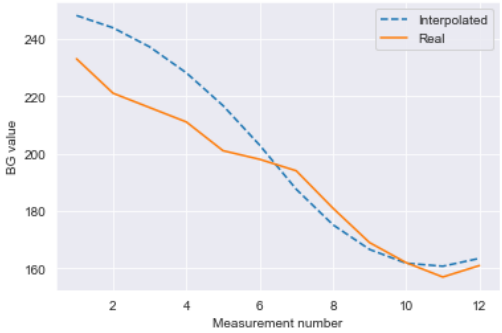
(a) 24 hours interpolation



(b) 12 hours interpolation



(c) 6 hours interpolation



(d) 3 hours interpolation

Chapter 6

Notes from interesting papers

General notes:

0.001 Learning rate and 0.9 decay rate is very used

The split train-test is done randomly (not sequentally), normally 80-20 or 70-30 or 60-20-20

LSTM hyperparameters:
The major parameter of the proposed model consists of the number of memory units inside each LSTM cell. A low number of memory units will make the model unable to learn all the patterns in the signals while a big value could be prone to overfitting.

Test selection: Randomly or sequential
In [17]:

1. Defines a set of hyperparameters for all points prediction about the selection of hyperparameters:
The hyper-parameters of the SVR were tuned separately for each point in [11], by using one week of data before the training week. This tuning procedure is computationally expensive and unfeasible in a real-time setting. Therefore, we used instead the same generic set of hyper-parameters for all points in the dataset, by tuning the SVR on a patient from the development dataset.

2. Cites the universal approximation theorem

3. Says preprocessing elections

In [23]:

1. Experimental selection of hyperparameters:
For the LSTM method, the epoch number for the two rounds of pre-train process was determined by running the experiment with epoch number from 100 to 2000, with increment of 100.

In [3]:

1. LSTM architecture:
A LSTM network consisting of a layer of 30 LSTM units and a single output layer (dense), **with a number of units equal to the future glucose samples that need to be predicted** (i.e. 18, corresponding to a 90 min prediction horizon at a 5 min sampling rate).

2. Experimental selection of number of regressors

3. Explains training elections (minibatch, learning rate...)

4. A lot of information about results analysis

In [33]:

1. Describe a very bassic preprocessing process.

Added time row
A time index (T) channel is added by normalising the one-day duration of 288 samples into a range of [0, 1). For example, 0:00 and 12:00 refer to 0 and 0.5, respectively.

2. A lot of information about results analysis

In [12]:

1. The preprocessing is explained

2. About hyperparameters:
While values of most other parameters are automatically derived via training. At the end of the training, the validation loss is quite close to the training loss, and both loss values are converging.

In [19]:

Two testing scenarios:
One based on randomly split into train-tests sets, another sequential, the test are the 20% last data.

In [16]:

Doesn't use interpolation:
Bc it's not clear how it can affect the results, instead only trains with real data.

In [9]:

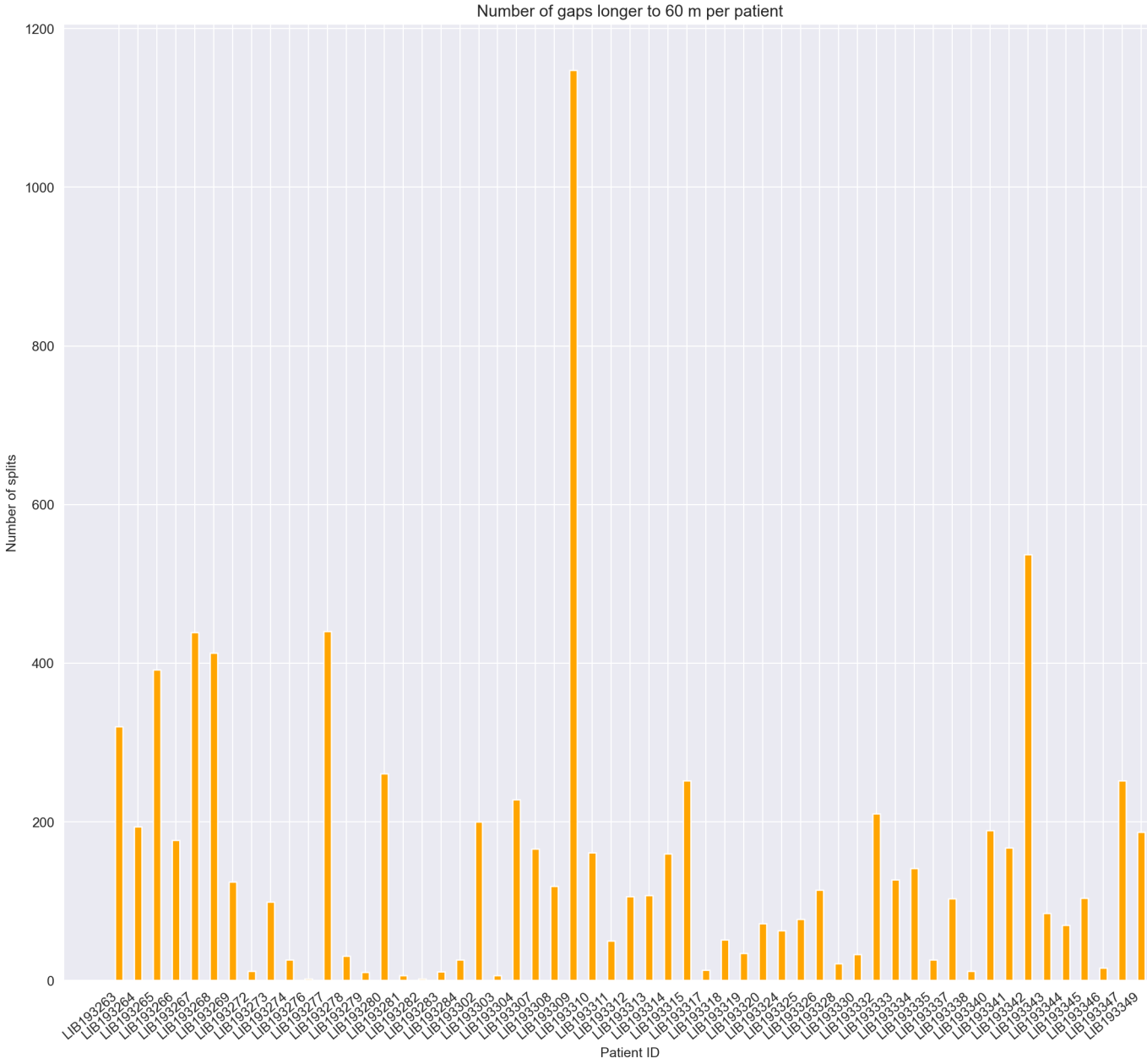
Avaliable code

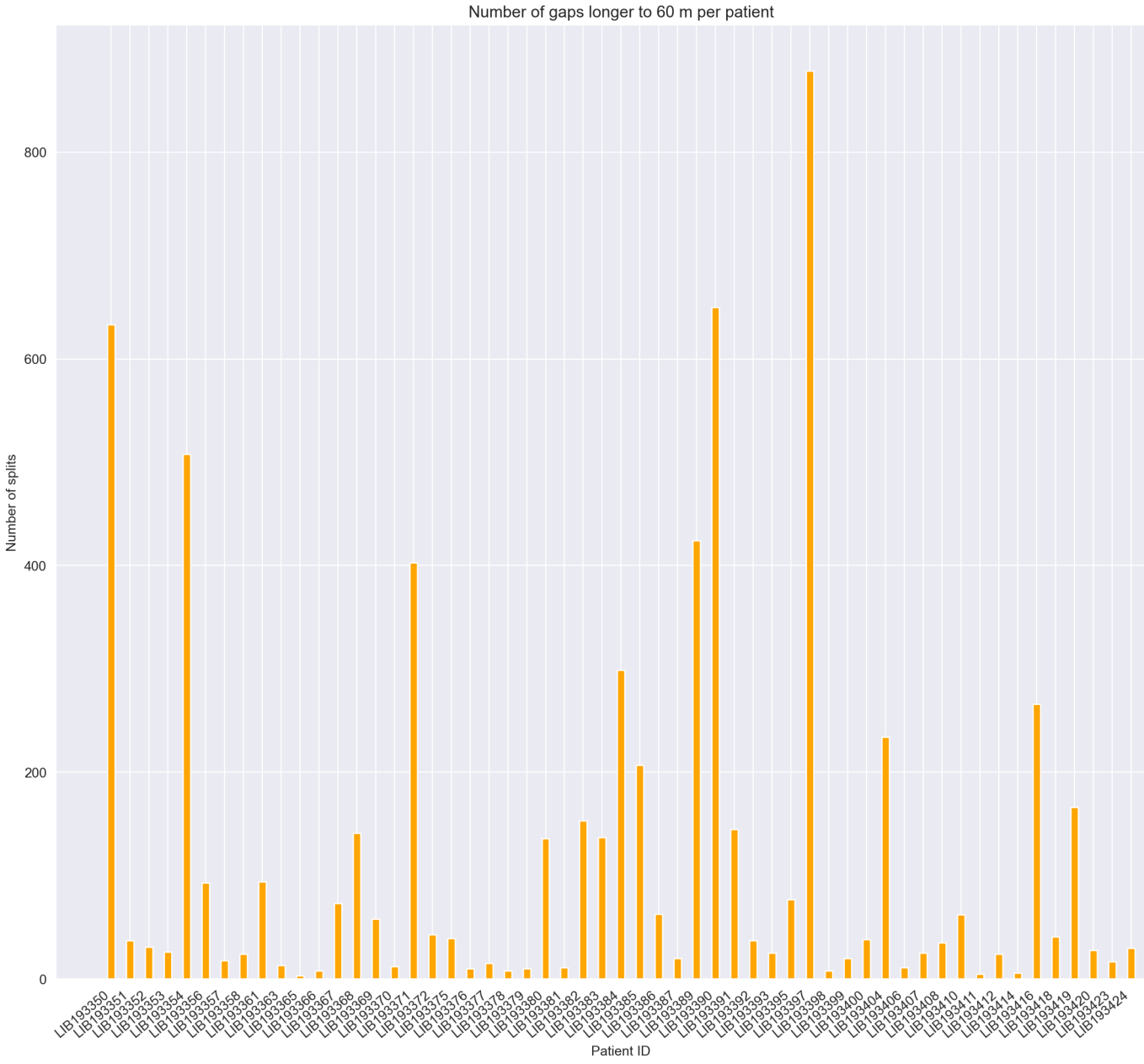
6.1 Chosen works

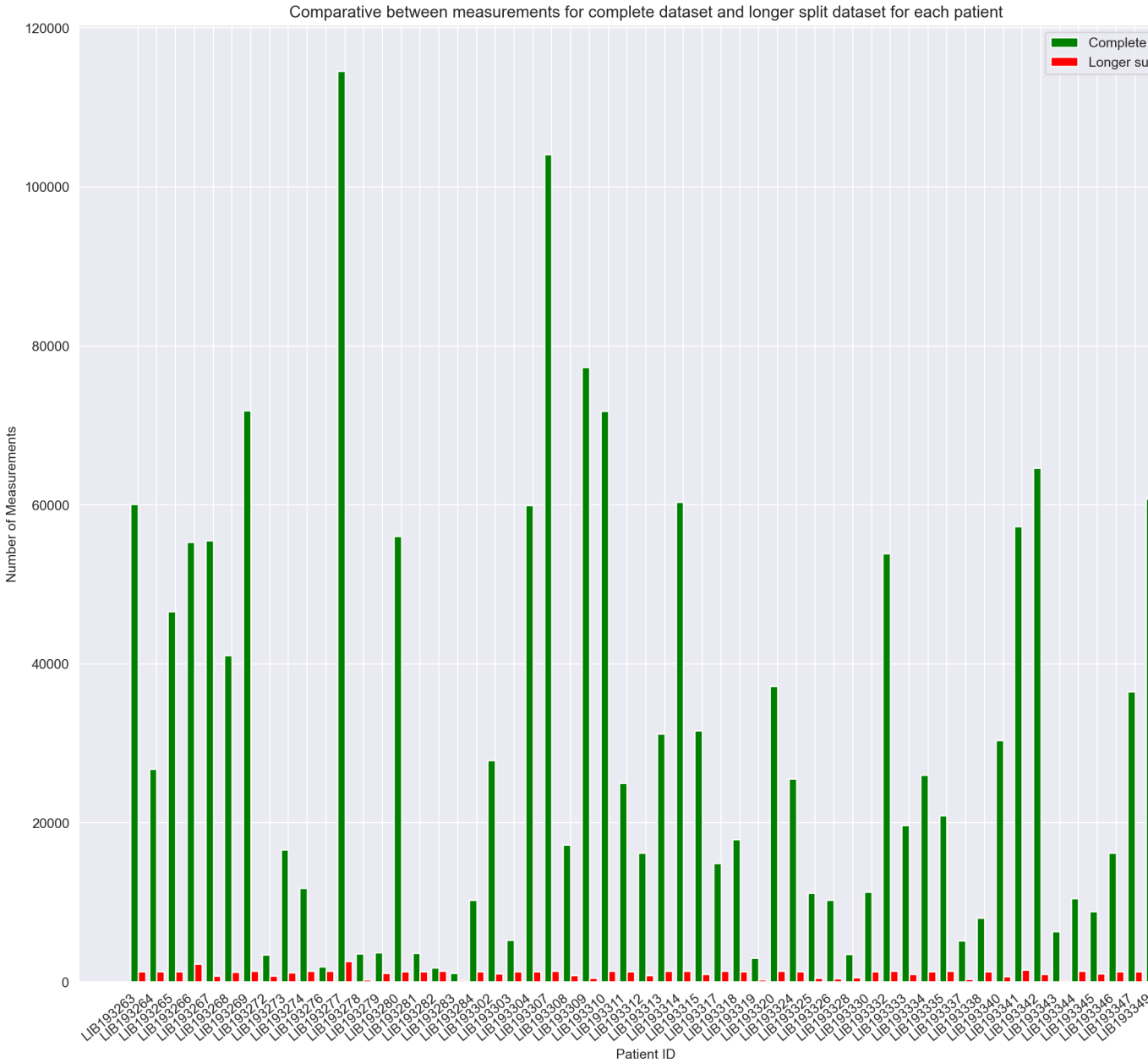
For CNN: [25]
For LSTM: [3]
For ensemble model: [21]

Chapter 7

Appendix

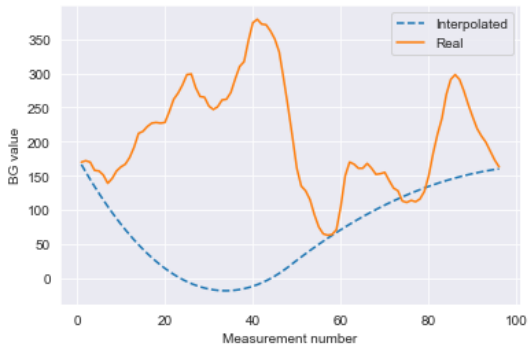



Number of gaps longer to 60 m per patient
Number of splits
0
200
400
600
800
Patient ID
LIB193350
LIB193351
LIB193352
LIB193353
LIB193354
LIB193356
LIB193357
LIB193358
LIB193361
LIB193363
LIB193365
LIB193366
LIB193367
LIB193368
LIB193369
LIB193370
LIB193371
LIB193372
LIB193375
LIB193376
LIB193377
LIB193378
LIB193379
LIB193380
LIB193381
LIB193382
LIB193383
LIB193384
LIB193385
LIB193386
LIB193387
LIB193389
LIB193390
LIB193391
LIB193392
LIB193393
LIB193395
LIB193397
LIB193398
LIB193399
LIB193400
LIB193404
LIB193406
LIB193407
LIB193408
LIB193410
LIB193411
LIB193412
LIB193414
LIB193416
LIB193418
LIB193419
LIB193420
LIB193423
LIB193424

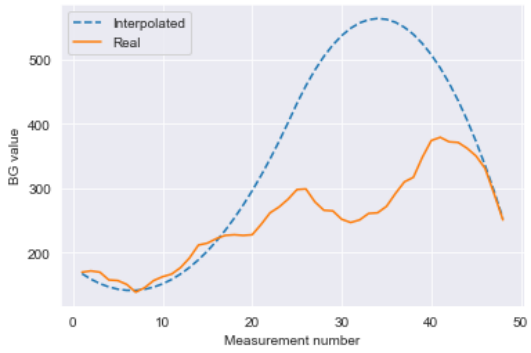

Comparative between measurements for complete dataset and longer split dataset for each patient
Complete
Longer su
Number of Measurements
120000
100000
80000
60000
40000
20000
0
Patient ID

Chapter 8

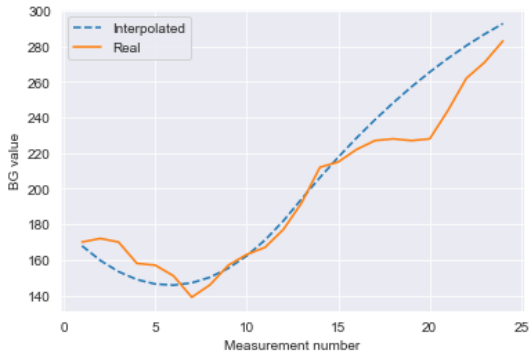
Preprocessing



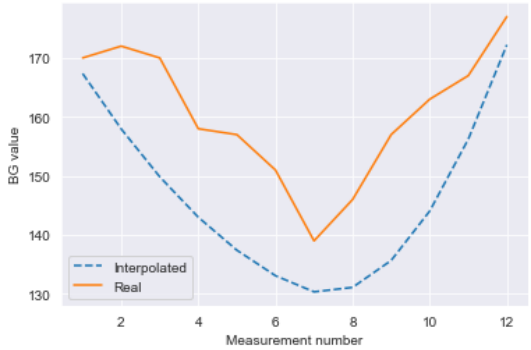
(a) 24 hours interpolation



(b) 12 hours interpolation



(c) 6 hours interpolation



(d) 3 hours interpolation

Figure 8.1: BG values interpolation for a random patient *LIB193307*

Bibliography

[1] Eleonora Maria Aiello, Giuseppe Lisanti, Lalo Magni, Mirto Musci, and Chiara Toffanin. Therapy-driven deep glucose forecasting. *ENGINEERING APPLICATIONS OF ARTIFICIAL INTELLIGENCE*, 87, 1 2020. Leer el estado del arte de este artículo.

[2] Ahmad Yaser Alhaddad, Hussein Aly, Hoda Gad, Abdulaziz Al-Ali, Kishor Kumar Sadasivuni, John-John Cabibihan, and Rayaz A Malik. Sense and learn: Recent advances in wearable sensing and machine learning for blood glucose monitoring and trend-detection. *Frontiers in Bioengineering and Biotechnology*, 10, 2022.

[3] Alessandro Aliberti, Irene Pupillo, Stefano Terna, Enrico Macii, Santa Di Cataldo, Edoardo Patti, and Andrea Acquaviva. A multi-patient data-driven approach to blood of glucose prediction. *IEEE ACCESS*, 7:69311–69325, 2019. Tiene un mini estado del arte¡br/¿¡br/¿Utiliza un dataset público!!!¡br/¿¡br/¿Comprueba testeando cuantos datos utilizar como delay para la predicción.¡br/¿¡br/¿Crea un modelo general que utiliza en ¡b¿datos de pacientes nuevos sin preprocesar¡/b¿¡br/¿¡br/¿Usa un montón de técnicas de regularización, para evitar overfitting... Con nombres raros.¡br/¿¡br/¿Tb muestra el ¡b¿Clarke eror grid¡/b¿.

[4] Ibrahim Aljamaan and Ibraheem Al-Naib. Prediction of blood glucose level using nonlinear system identification approach. *IEEE ACCESS*, 10:1936–1945, 2022.

[5] Y BENGIO, P SIMARD, and P FRASCONI. Learning long-term dependencies with gradient descent is difficult. *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 5:157–166, 3 1994.

[6] Brian Bogue-Jimenez, Xiaolei Huang, Douglas Powell, and Ana Doblas. Selection of noninvasive features in wrist-based wearable sensors to predict blood glucose concentrations using machine learning algorithms. *Sensors*, 22, 5 2022.

[7] Razvan Bunescu, Nigel Struble, Cindy Marling, Jay Shubrook, and Frank Schwartz. Blood glucose level prediction using physiological models and support vector regression. *Proceedings - 2013 12th International Conference on Machine Learning and Applications, ICMLA 2013*, 1:135–140, 2013. Este es el artículo que tiene asociado un artículo que se hizo después de LSTM.

[8] Iván Contreras, Silvia Oviedo, Martina Vettoretti, Roberto Visentin, and Josep Vehí. Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models. *PLOS ONE*, 12:e0187754, 2 2017.

[9] Ian Fox, Lynn Ang, Mamta Jaiswal, Rodica Pop-Busui, and Jenna Wiens. Deep multi-output forecasting learning to accurately predict blood glucose trajectories. pages 1387–1395, 2018. 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), London, ENGLAND, AUG 19-23, 2018.

[10] J j. LI, Z p. Qu, Y w. Wang, and J Guo. Research on multi-parameter fusion non-invasive blood glucose detection method based on machine learning. *EUROPEAN REVIEW FOR MEDICAL AND PHARMACOLOGICAL SCIENCES*, 26:6040–6049, 2022.

[11] Dae-Yeon Kim, Dong-Sik Choi, Ah Reum Kang, Jiyoung Woo, Yechan Han, Sung Wan Chun, and Jaeyun Kim. Intelligent ensemble deep learning system for blood glucose prediction using genetic algorithms. *COMPLEXITY*, 2022, 10 2022.

[12] Kezhi Li, John Daniels, Chengyuan Liu, Pau Herrero, and Pantelis Georgiou. Convolutional recurrent neural networks for glucose prediction. *IEEE Journal of Biomedical and Health Informatics*, 24:603–613, 2 2020. ¡b¿Me gusta el modelo que implementa¡/b¿¡br/¿¡br/¿Usa un método para comparar sus resultados con otros traba-

jos¡br/¿¡br/¿Usa valores relativos en el entrenamiento de la red neuronal en lugar de absolutos.

[13] Kezhi Li, Chengyuan Liu, Taiyu Zhu, Pau Herrero, and Pantelis Georgiou. Glunet: A deep learning framework for accurate glucose forecasting. *IEEE Journal of Biomedical and Health Informatics*, 24:414–423, 2 2020. CNN¡br/¿¡br/¿Realiza el cambio de ¡b¿valores absolutos a relativos¡/b¿¡br/¿¡br/¿Transforma la predicción en un problema de clasificación.

[14] Ming Liu, Ge Xu, Yuejin Zhao, Lingqin Kong, Liquan Dong, Fen Li, and Mei Hui. Diffuse imaging approach for universal noninvasive blood glucose measurements. *Frontiers in Physics*, 10, 3 2022.

[15] Chiara Dalla Man, Francesco Micheletto, Dayu Lv, Marc Breton, Boris Kovatchev, and Claudio Cobelli. The uva/padova type 1 diabetes simulator. *Journal of Diabetes Science and Technology*, 8:26–34, 1 2014.

[16] John Martinsson, Alexander Schliep, Björn Eliasson, and Olof Mogren. Blood glucose prediction with variance estimation using recurrent neural networks. *Journal of Healthcare Informatics Research*, 4:1–18, 2020. Muy largo ¡br/¿¡br/¿Explica la elección de los hiperparámetros¡br/¿¡br/¿Experimentos para ver cuanto ruido queda.

[17] Sadegh Mirshekarian, Razvan Bunescu, Cindy Marling, and Frank Schwartz. Using lstms to learn physiological models of blood glucose behavior. pages 2887–2891. IEEE, 7 2017.

[18] Sadegh Mirshekarian, Hui Shen, Razvan Bunescu, and Cindy Marling. Lstms and neural attention models for blood glucose prediction: Comparative experiments on real and synthetic data. pages 706–712. IEEE, 2019. Se puede utilizar como ejemplo para mostrar las ¡b¿diferencias entre los datos simulados y los reales¡/b¿.

[19] Mario Munoz-Organero. Deep physiological model for blood glucose prediction in t1dm patients. *Sensors*, 20, 2020.

[20] Ahmed R Nasser, Ahmed M Hasan, Amjad J Humaidi, Ahmed Alkhayyat, Laith Alzubaidi, Mohammed A Fadhel, José Santamaría, and Ye Duan. Iot and cloud computing in health-care: A new wearable device and cloud-based deep learning algorithm for monitoring of diabetes. *Electronics*, 10, 2021.

[21] Hoda Nemat, Heydar Khadem, Mohammad R Eissa, Jackie Elliott, and Mohammed Benaissa. Blood glucose level prediction: Advanced deep-ensemble learning approach. *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, 26:2758–2769, 6 2022.

[22] Francesco Prendin, Simone Del Favero, Martina Vettoretti, Giovanni Sparacino, and Andrea Facchinetti. Forecasting of glucose levels and hypoglycemic events: Head-to-head comparison of linear and nonlinear data-driven algorithms based on continuous glucose monitoring data only. *Sensors*, 21, 2021.

[23] Qingnan Sun, Marko V Jankovic, Lia Bally, and Stavroula G Mougiakakou. Predicting blood glucose with an lstm and bi-lstm based deep neural network. 2018. Me gusta mucho cómo elige los hiperparámetros¡br/¿.

[24] Baoyu Tang, Yuyu Yuan, Jincui Yang, Lirong Qiu, Shasha Zhang, and Jinsheng Shi. Predicting blood glucose concentration after short-acting insulin injection using discontinuous injection records. *SENSORS*, 22, 11 2022.

[25] Ali El Idrissi Touria and Idri. Deep learning for blood glucose prediction: Cnn vs lstm. pages 379–393. Springer International Publishing, 2020. Habla de las estrategias utilizadas para el forecasting¡br/¿¡br/¿Avaliable dataset¡br/¿¡br/¿Muestra por qué el multistep les da mejores resultados.

[26] Stella Tsichlaki, Lefteris Koumakis, and Manolis Tsiknakis. Type 1 diabetes hypoglycemia prediction algorithms: Systematic review. *JMIR Diabetes*, 7:e34699, 7 2022.

[27] Ching-Shih Tsou, Christine Liou, Longsheng Cheng, and Hanting Zhou. Quality prediction through machine learning for the inspection and manufacturing process of blood glucose test strips. *COGENT ENGINEERING*, 9, 12 2022.

[28] Qiaoyun Wang, Feifei Pian, Mingxuan Wang, Shuai Song, Zhigang Li, Peng Shan, and Zhenhe Ma. Quantitative analysis of raman spectra for glucose concentration in human blood using gramian angular field and convolutional neural network. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 275, 7 2022.

[29] Jun Yang, Lei Li, Yimeng Shi, and Xiaolei Xie. An arima model with adaptive orders for predicting blood glucose concentrations and hypoglycemia. *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, 23:1251–1260, 5 2019.

[30] Zhenyi Ye, Jie Wang, Hao Hua, Xiangdong Zhou, and Qiliang Li. Precise detection and quantitative prediction of blood glucose level with an electronic nose system. *IEEE Sensors Journal*, 22:12452–12459, 7 2022.

[31] Yongjun Zhang and Guangheng Gao. Optimization and evaluation of an intelligent short-term blood glucose prediction model based on noninvasive monitoring and deep learning techniques. *Journal of Healthcare Engineering*, 2022, 2022.

[32] Taiyu Zhu, Lei Kuang, Kezhi Li, Junming Zeng, Pau Herrero, and Pantelis Georgiou. Blood glucose prediction in type 1 diabetes using deep learning on the edge. IEEE, 2021. IEEE International Symposium on Circuits and Systems (IEEE ISCAS), Daegu, SOUTH KOREA, MAY 22-28, 2021.

[33] Taiyu Zhu, Kezhi Li, Jianwei Chen, Pau Herrero, and Pantelis Georgiou. Dilated recurrent neural networks for glucose forecasting in type 1 diabetes. *JOURNAL OF HEALTHCARE INFORMATICS RESEARCH*, 4:308–324, 9 2020. Explica la elección de hiperparámetros¡br/¿¡br/¿Usa ¡b¿transfer learning¡/b¿ y explica sobre él.

[34] Ting Zhu, Wenbo Wang, and Min Yu. A novel blood glucose time series prediction framework based on a novel signal decomposition method. *CHAOS SOLITONS FRACTALS*, 164, 11 2022.