CISC/CMPE 452

Neural and Genetic Computing

Prediction of Glucose Concentration Patients with Type-1 Diabetes

***Prepared by:***

Nolan Nisbet

Ryan Dick

Evan Cloutier

Aidan Gunda

Jordan Nanos

***December 4, 2016***

# Project Proposal

## Project Value

Type 1 diabetes is a disease characterized by insufficient insulin production resulting in elevated blood glucose levels. Management of type 1 diabetes involves insulin therapy along with close monitoring of blood glucose concentrations. A major challenge type 1 diabetics face is avoiding the occurrence of hypoglycemia and hyperglycemia [1]. There are many factors that impact blood glucose concentration (nutrition, exercise, emotional state, etc.) making it difficult to accurately predict changes [1]. Improvements in the ability to predict glucose highs and lows would be beneficial for diabetes patients, because it would help them to avoid these undesired states.

Conventionally, glucose concentrations have been monitored through the use of home glucose monitors (HGM) with measurements being taken approximately 4 times daily. Recent advances in technology have lead to the development of continuous glucose monitoring systems (CGMS), which are used to measure blood glucose levels on a more frequent basis (approximately every 5 mins.) [1]. This project aims to apply neural computing algorithms to the data collected by a patient’s CGMS in order to learn their individual's characteristics and more accurately predict future blood glucose spikes and lows.

## Data Source

The Jaeb Center for Health Research has published CGMS data collected as part of a study conducted to investigate the efficacy of real-time CGM systems [2]. The data set that will be used contains blinded CGM measurements for a control group that was instructed to continue following a 4-times daily HGM measurement protocol for adjusting insulin doses [2]. Since no nutrition or insulin dosing data is available, the project will aim to predict overnight glucose lows. Limiting the data set to overnight data for individuals using HGM protocols will reduce the likelihood that meals and insulin injections are influencing the observed glucose levels. Targeting overnight prediction still addresses an important issue, as more than 50% of severe hypoglycemic events occur during sleep, mainly due to a lack of monitoring [2].

## Proposed Algorithm

There are two neural network architectures that are most often applied to the challenge of time-series data prediction. One option is a feed forward neural network that takes as input previous glucose concentration measurements to predict future glucose levels at a specific prediction time horizon [1]. A second option is a recurrent neural network, which accepts each measurement sequentially and maintains a notion of the current state in order to make glucose concentration predictions [3]. Both of these algorithms will be implemented and compared for a prediction horizon of 20 minutes in order to determine which architecture is most effective at addressing the proposed problem.

## Validation Criteria

The effectiveness of the neural network solution will be evaluated based on the following three metrics:

1. The mean squared error of the glucose concentration predictions.
2. The percentage of hypoglycemic events (<= 70 mg/dL) and hyperglycemic events (>= 200 mg/dL) that are correctly predicted.
3. The number of false lows and false highs.

The mean squared error will be used to evaluate the overall accuracy of the prediction algorithm and based on previous neural network prediction schemes, an MSE of less than 40 mg/dL will be targeted [1]. The percentage of correctly predicted hypoglycemic and hyperglycemic events provides a measure of the solution's effectiveness at specifically predicting overnight glucose levels that require patient notification. The project goal is to achieve prediction rates greater than 80% for these potentially dangerous conditions. Finally, the quantity of false predictions must be considered as a high frequency of false alarms could invalidate a high prediction accuracy. Evidently, the target value is to achieve zero false high or low notifications.

The above criteria will be applied to both of the proposed neural network designs (FFNN and RNN) to select the most effective option and determine whether it is a reliable tool for guiding treatment decisions.

# Project Implementation

## Implementation Details

### Data Preprocessing

The data used for the prediction of blood glucose levels in Type 1 diabetes patients consisted of glucose measurements and timestamps for RTCGM measurements collected from patients participating in a study conducted by the Jaeb Center for Health Research [2]. In order to achieve accurate glucose concentration prediction based on historical concentration measurements, it was desired to reduce the effect of unknown factors such as insulin injection, meal consumption and exercise. The purpose of the data preprocessing steps was to minimize the effect of external factors and organize the data into a format that could be used to train a neural network.

The first data processing step was to select data collected from the study's control group. This group was preferred because the participants were still making insulin injection decisions based on measurements taken 4 times daily (despite the fact that they were being continuously monitored), thus it was likely that this group would be injecting insulin less frequently than the group making decisions based on the CGM data. From this data set, the measurements were grouped based on patient ID and then organized into 24-hour windows. The code for this initial data extraction step is included in the 'CGMParse.cpp' file.

Within each day's data, only the overnight data was analyzed to determine if it could be included in the training set. Overnight data was selected to minimize the likelihood that the patients would be injecting insulin or consuming food. Patient data was parsed to find 7 consecutive data points separated by 10 minute intervals, and a final data point 20 minutes ahead of the time window to be used as a prediction point. If a valid time window was found, these points were written to either the training set (with probability of 75%) or the test set (with probability of 25%) to be used as a single input into the neural network. This process continued until all of the data was parsed for each patient. The python script used for data processing is included in the 'data\_preprocessor.py'.

Figure 1 contains a flowchart illustrating the data processing steps described above.

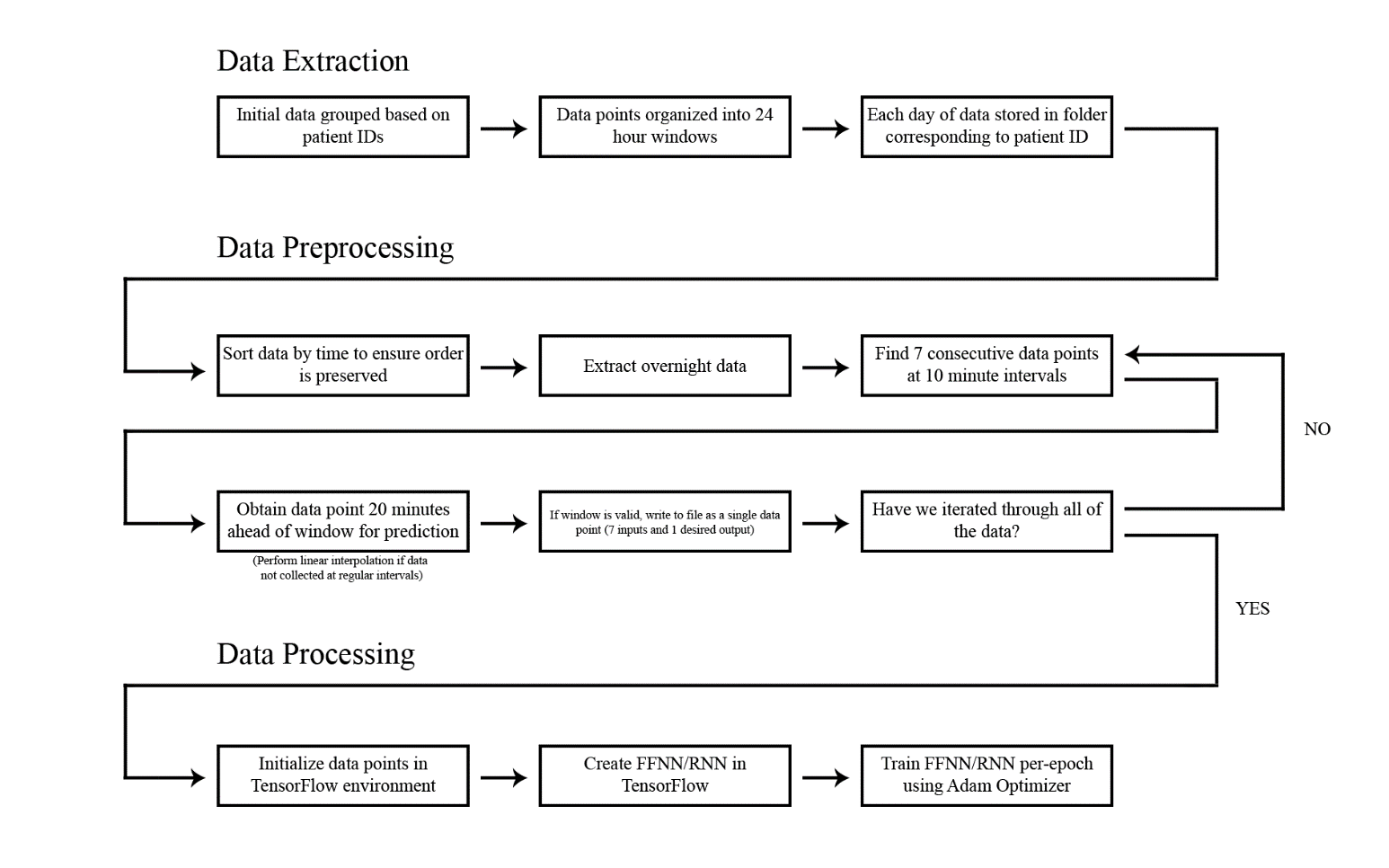


Figure 1: Data preprocessing flow chart.

### Neural Networks

In order to achieve the goal of glucose concentration prediction, two types of neural networks were modelled and compared: a feedforward network and a recurrent network.

The feedforward network followed the tapped delay-line architecture, meaning that there were 7 input nodes to accept each of the last 7 glucose measurements. The network consisted of two hidden layers with 15 nodes in each layer, and a single output neuron. These values were arbitrarily chosen, with no significant improvements resulting when the number of hidden layers or nodes was increased. A linear output function was used in order to achieve the network's purpose: prediction. Since no classification was required, there was no need for a hyperbolic or sigmoidal function; converging output values were not desired. To train the network, the Adam optimization algorithm was used with a mean squared error (MSE) cost function. An extension of gradient descent, Adam is a first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments [4]. A standard gradient descent algorithm was tested for comparison, but it was found that the Adam algorithm produced better results. A per-epoch methodology was used when training the network. The per-epoch method was compared to a per-batch training method, and while it required longer training times it resulted in slightly improved results. The FFNN was implemented using Tensorflow as shown in the 'FeedForwardNN.py' file.

The recurrent neural network model that was tested was composed of a single long short-term memory (LSTM) cell to retain the required values for output prediction – the past 7 glucose measurements. The output produced by the LSTM cell was then multiplied by a weight, with a bias applied, to produce a predicted output value. Several of the same design decisions were made for the RNN as for the feed-forward network. Again, linear output functions were used to achieve prediction, and the Adam optimization algorithm was used to minimize the MSE cost function. The Tensorflow RNN implementation is in the 'RecurrentNN.py' file.

## Results

The two network models were tested with data from three different patients (IDs 149, 151, 174). Both networks were trained for each of the three patients uniquely and then tested on all of the patients. In other words, the networks were tested on patients for which they were not trained. The purpose of these tests were to address the following questions:

1. Which network type is more effective at predicting glucose concentration, RNN or FFNN?
2. Are there significant differences between patients? Is there value in training the network for a specific patient rather than applying a generic prediction algorithm?
3. Are neural networks capable of accurately predicting blood glucose levels?

The following metrics were used to evaluate the performance of each network after training:

1. **MSE**: The average squared difference between the predicted concentration and the actual value at a prediction horizon of 20 minutes.
2. **% of Lows Identified**: Lows are characterized as glucose concentrations below 70 mg/dL. This metric is the percentage of lows for which the network correctly predicted a value below the threshold.
3. **% of Highs Identified**: Highs are characterized as glucose concentrations above 200 mg/dL. This metric is the percentage of highs for which the network correctly predicted a value above the threshold.
4. **Number of False Lows**: A count of the number of predictions below the low threshold, for which the actual concentration was above the threshold. (If the error was less than 8 mg/dL then the false low was not counted)
5. **Number of False Highs**: A count of the number of predictions above the high threshold, for which the actual concentration was below the threshold. (If the error was less than 8 mg/dL then the false high was not counted)

The following tables contain the test results collected after training both the RNN and the FFNN for all three patients.

### Feed Forward Neural Network

Table - Evaluation criteria for FFNN trained on Patient 149 data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Patient** | **MSE** | **% of Lows Identified** | **% of Highs Identified** | **Number of False Lows** | **Number of False Highs** |
| ***149*** | *249* | *79%* | *93%* | *14* | *44* |
| **151** | 100 | 75% | 96% | 8 | 5 |
| **174** | 133 | 77% | 88% | 4 | 11 |

Table - Evaluation criteria for FFNN trained on Patient 151 data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Patient** | **MSE** | **% of Lows Identified** | **% of Highs Identified** | **Number of False Lows** | **Number of False Highs** |
| **149** | 278 | 72% | 93% | 19 | 48 |
| ***151*** | *101* | *82%* | *96%* | *11* | *4* |
| **174** | 154 | 69% | 86% | 5 | 17 |

Table - Evaluation criteria for FFNN trained on Patient 174 data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Patient** | **MSE** | **% of Lows Identified** | **% of Highs Identified** | **Number of False Lows** | **Number of False Highs** |
| **149** | 270 | 84% | 93% | 19 | 49 |
| **151** | 105 | 73% | 97% | 10 | 3 |
| ***174*** | *143* | *77%* | *87%* | *4* | *14* |

### Recurrent Neural Network

Table - Evaluation criteria for RNN trained on Patient 149 data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Patient** | **MSE** | **% of Lows Identified** | **% of Highs Identified** | **Number of False Lows** | **Number of False Highs** |
| **149** | 279 | 75% | 93% | 14 | 46 |
| **151** | 106 | 80% | 97% | 13 | 7 |
| **174** | 161 | 69% | 87% | 5 | 16 |

Table - Evaluation criteria for RNN trained on Patient 151 data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Patient** | **MSE** | **% of Lows Identified** | **% of Highs Identified** | **Number of False Lows** | **Number of False Highs** |
| **149** | 314 | 67% | 92% | 12 | 54 |
| **151** | 105 | 70% | 96% | 9 | 5 |
| **174** | 173 | 69% | 85% | 5 | 17 |

Table - Evaluation criteria for RNN trained on Patient 174 data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Patient** | **MSE** | **% of Lows Identified** | **% of Highs Identified** | **Number of False Lows** | **Number of False Highs** |
| **149** | 408 | 70% | 93% | 16 | 51 |
| **151** | 112 | 65% | 97% | 12 | 9 |
| **174** | 151 | 77% | 89% | 3 | 15 |

## Discussion and Future Improvements

Based on the results of both neural networks, the following are some conclusions that can be made about the data. Firstly, the MSE needs to be significantly lower in order for neural networks to be a reliable prediction mechanism. More data would likely be needed with respect to each patient in order to achieve better performance. A major improvement would likely be observed with the addition of nutritional intake, broken down into carbohydrates (simple and complex), proteins and fats, as an input to the network. Similarly, activity duration and type (aerobic or anaerobic) would allow for a better understanding of nighttime low events. Even factors like sleep quality and duration as well as general stress level can have a noticeable impact on glucose levels and would likely prove valuable in decreasing the prediction MSE. Access to more data about each patient could make neural networks a more feasible tool for blood sugar prediction.

Next, it was observed that the feedforward and recurrent networks produced very similar results, although the feedforward network provided slightly more consistent training compared to the recurrent network. It was observed qualitatively that when the feedforward network was trained multiple times on the same patient, the resultant MSE values were very similar. Contrarily, if this same process was done on the recurrent network, the MSE values fluctuated more from one training cycle to the next.

The number of false lows and false highs were higher than anticipated. This is highly undesirable, as it could lead to misinformed treatment decisions being made. Future work could address this issue in two ways. First, the addition of a larger proportion of high and low events in the training data set would likely yield improved results. Currently, the dataset is composed primarily of 'normal' glucose measurements, thus the networks achieved their optimal prediction accuracy in this range. New high and low data points could be generated artificially by simply adding noise to existing data points. A second possible solution to the issue of false lows and highs is the development of a cost function that penalizes 'false alarms' more aggressively than a standard MSE cost function. The ability to create custom cost functions was limited in the software library used for the neural network implementation (TensorFlow); however, future work could investigate the effectiveness of this option.

It was observed that the networks performed equally well regardless of which patient they were trained for. In other words, the network trained for patient 149 was equally capable of predicting glucose for patient 151 as the network trained specifically for patient 151. This observation suggests that a static algorithm may be equally well suited to solving this problem as a neural network approach. Future work could investigate this observation further by grouping individuals based on their demographics (gender, age, etc.) to determine whether any potential trends exist that were overlooked in the initial study.

## Applications

This project focused solely on the prediction of overnight blood glucose levels in patients with type-1 diabetes. This is a valuable application; however, it does not represent the full scope of benefit that predictive neural networks can offer these patients. If future work is successfully able to eliminate the current system limitations discussed in Section 2.3, then neural networks have the potential to become a standard around-the-clock diabetes management tool. If a sufficiently high level of accuracy is achieved, real-time predictive neural networks could replace current standard diabetes management protocols and greatly reduce the risks experienced by individuals with diabetes on a daily basis.

# References

[1] Pappada, S. M., B. D. Cameron, and P. M. Rosman. "Development of a Neural Network for Prediction of Glucose Concentration in Type 1 Diabetes Patients." Journal of Diabetes Science and Technology 2.5 (2008): 792-801. Web

[2] Beck, Roy W. "A Randomized Clinical Trial to Assess the Efficacy of Real-Time Continuous Glucose Monitoring in the Management of Type 1 Diabetes". Jaeb Center for Health Research (JCHR) - Diabetes Research Studies (2009). <http://diabetes.jaeb.org/RT_CGMRCTProtocol.aspx>. Accessed 8 Oct. 2016.

[3] Mougiakakou, Stavroula G., Aikaterini Prountzou, Dimitra Iliopoulou, Konstantina S. Nikita, Andriani Vazeou, and Christos S. Bartsocas. "Neural Network Based Glucose - Insulin Metabolism Models for Children with Type 1 Diabetes." 2006 International Conference of the IEEE Engineering in Medicine and Biology Society (2006): n. pag. Web.

[2] J. L. B. Diederik P. Kingma, "Adam: A Method for Stochastic Optimization," ICLR 2015, 2015.