HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 4 (2024)

Comparative Analysis of LSTM and GRU Neural Networks in Predicting Hyperglycemic and Hypoglycemic Events

Anish Kumar

Phoenix High School, USA

ABSTRACT

The research aims to compare the effectiveness of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks in predicting hyperglycemic and hypoglycemic events in diabetes patients using wearable sensor data, to identify the more effective model for near-term blood glucose forecasting. Effective prediction of blood glucose levels is crucial for Type 1 diabetes management, as extreme glucose levels can cause fatal hyperglycemic and hypoglycemic incidents. Choosing between LSTM and GRU models can significantly impact predictive accuracy and patient outcomes. The OpenD1NAMO public dataset was used as a training set to compare the efficacies of LSTM and GRU models in predicting near-term glucose levels. Artificial features such as lagged features and rolling averaged lagged features were created to facilitate glucose forecasting comparisons. Various permutations of hyperparameters were then compared to find the most effective set for the provided sensor data. GRU and LSTM models trained on a portion of the preprocessed dataset were compiled and the average of the minimum MSEs per epoch were compared. On average, the LSTM model was 53.5% more accurate than the GRU model for the dataset provided, forecasting blood glucose values with a 15-minute interval. The minimum average MSE of the LSTM model was 2.043 $mmol^2/L^2$ while the minimum average MSE of the GRU model was 3.137 $mmol^2/L^2$. Given the models' performances on the dataset, it suggests that LSTM models will likely be more accurate than GRU models when predicting near-term blood glucose levels using live wearable sensor data.

Introduction

Effective management of diabetes requires accurate prediction of blood glucose levels to prevent potentially life-threatening hyperglycemic and hypoglycemic events. This research seeks to compare the efficacy of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks in predicting hyperglycemic and hypoglycemic incidents using time series data from diabetes patients. The research question addresses which model—LSTM or GRU—is more effective in near-term blood glucose forecasting, based on continuous wearable sensor data.

Diabetes mellitus, a metabolic disorder that can result in abnormal blood glucose levels, can result in many health complications. Individuals with Type 1 diabetes often have trouble regulating their blood glucose, which can result in hyperglycemic and hypoglycemic episodes. Various factors can play a role in the timing of these episodes and an individual's blood glucose levels, including carbohydrate intake, sleep, exercise, etc.

One critical aspect of management is predicting blood glucose levels to anticipate and prevent extreme fluctuations. LSTM and GRU networks are popular choices for time series forecasting due to their ability to capture temporal dependencies effectively[1]. Understanding which model performs better in this context is essential for enhancing predictive accuracy and, consequently, improving patient outcomes.

The research utilizes the OpenD1NAMO public dataset, containing time series wearable sensor data relevant to blood glucose levels in diabetes patients[2]. This dataset provides an opportunity to explore the performance of LSTM and GRU models in predicting near-term glucose levels, which is crucial for the real-time management of Type 1 diabetes.

Glucose levels were predicted using supervised regression, which aims to predict continuous numerical values representing blood glucose levels. The data primarily consists of numerical features extracted from wearable sensor readings, capturing physiological parameters correlated with blood glucose fluctuations. The few categorical features were not used in the final models.

The significance of this research lies in its potential to inform clinical practice and wearable technology development. By identifying the superior model for near-term blood glucose forecasting, healthcare providers can make more informed decisions regarding diabetes management strategies. Advancements in wearable sensor technology coupled with accurate predictive models hold promise for allowing individuals with Type 1 diabetes to proactively manage their condition and improve their quality of life.

Methods

Several articles and approaches have been explored in blood glucose prediction and diabetes management, particularly the influence of diet on glucose levels, and cardiological and respiratory factors. Research from the University of Florida's Diabetes Institute underscores the intricate interplay between dietary factors, activity, and blood glucose regulation[4]. Carbohydrates emerge as a major determinant of blood glucose levels, with proteins and fats also exerting some influence, albeit to a lesser extent. Additionally, minute-by-minute blood sugar fluctuations are influenced by diet and exercise, emphasizing the need for a balanced approach to diabetes management.

Trepanowski et al. (2017) investigated the effects of alternate-day fasting on weight loss, weight maintenance, and cardioprotection in metabolically healthy obese adults[5]. The randomized clinical trial showed promising outcomes and suggested that alternate day fasting could be a viable strategy for weight management and cardiovascular health improvement.

Sutton et al. (2018) delved into early time-restricted feeding and its impact on insulin sensitivity, blood pressure, and oxidative stress in men with prediabetes[6]. Despite no significant weight loss, early time-restricted feeding showed improvements in various metabolic parameters, highlighting its potential as an intervention for metabolic health enhancement.

Moreover, insights from Harvard Health underscore the importance of maintaining normal blood sugar levels and the role of factors like exercise, carbohydrate intake, and sleep in regulating glucose levels[7]. The risks associated with hypoglycemia are outlined, emphasizing the need for balanced lifestyle choices to prevent adverse outcomes.

These approaches collectively contribute to the understanding of blood glucose regulation and diabetes management. While alternate-day fasting and early time-restricted feeding offer novel dietary interventions, traditional strategies emphasizing balanced carbohydrate intake, exercise, and sleep remain fundamental. Incorporating insights from these studies alongside glucose regulation using LSTM and GRU neural networks for near-term blood glucose prediction can provide a comprehensive approach to Type 1 diabetes management.

The dataset utilized in this study is sourced from the Open D1NAMO dataset, a comprehensive resource designed for research on non-invasive Type 1 diabetes management[2]. This open-access dataset offers trusted data and is widely utilized in various research endeavors within the scientific community. The Open D1NAMO dataset contains two distinct subsets: one that was acquired from 20 healthy people, and a second one acquired from 9 patients with Type 1 diabetes.

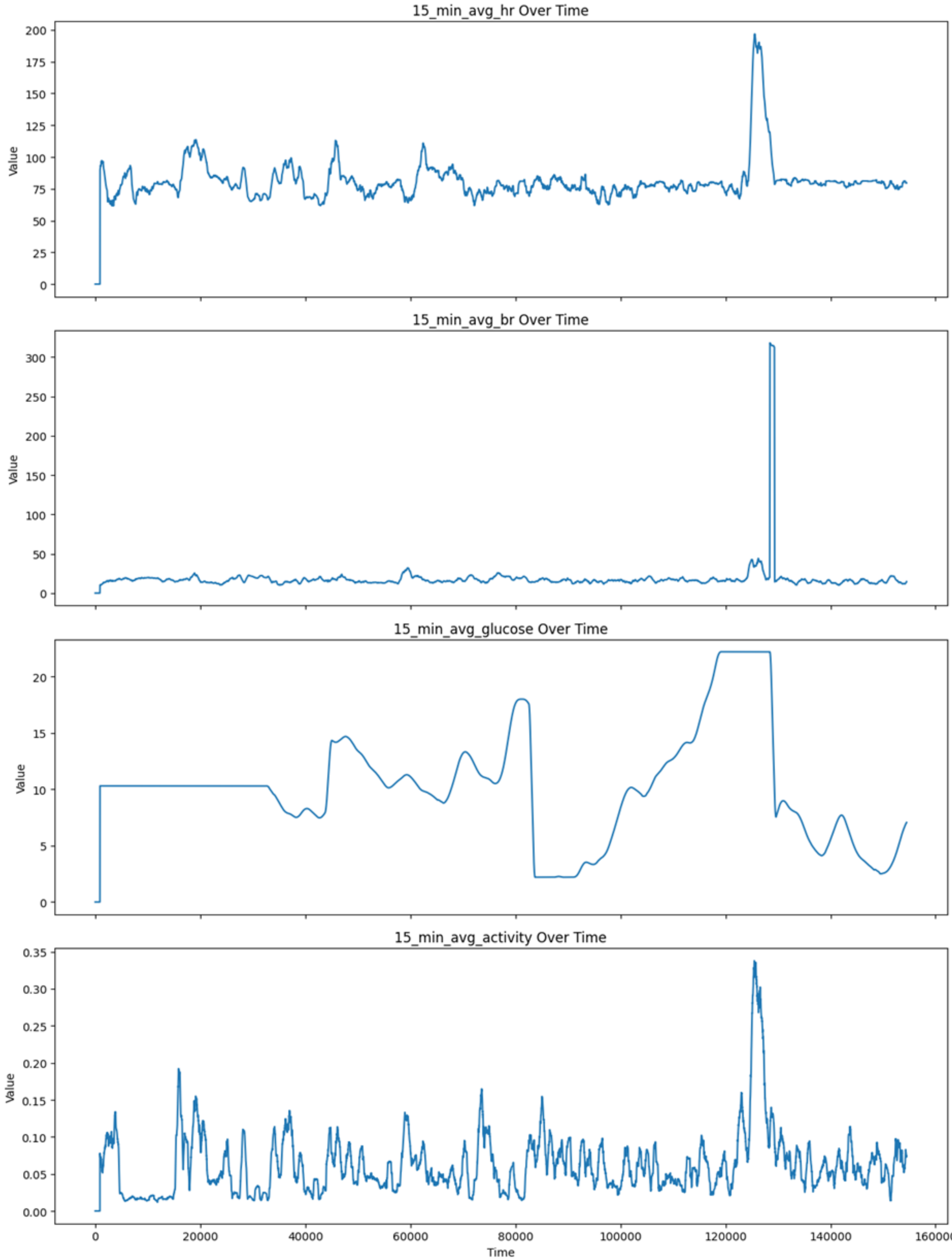


Figure 1. Timeseries plots of all relevant features

Only data from the second subset was used for the purposes of this research, focusing primarily on glucose detection for diabetes patients. The same wearable device was used for both subsets, and the age, gender, height, and weight of each patient were recorded. Both subsets contain ECG signals, breathing signals, accelerometer outputs, glucose measurements, caloric intake, food pictures, and annotations by a dietitian. The dataset contains approximately 450 hours of total continuous data for all participants, with an average of approximately 50 hours of data recorded per participant. An average of 934.9 blood glucose recordings were taken per diabetes patient in the second subset.

The Open D1NAMO dataset primarily contains numerical measurements for ECG signals, breathing signals, accelerometer outputs, blood glucose measurements, and caloric intake. Images and categorical data were included in various food pictures and dietician meal annotations, respectively. Both non-numerical features were not included in the final processed dataset. Each of the numerical features is relevant to the research question, due to their correlation with near-term blood glucose levels[8]. The dataset was preprocessed by extracting individual groups of features with identical time formats into separate data frames. All relevant features (heart rate, breathing rate, activity, glucose levels, long/short-term insulin intakes, meal calories, and time) were merged into a single data frame after converting all datetime columns into an identical format. All null values were interpolated.

Several artificial features were created, both to display the structure of the dataset and to allow for blood glucose forecasting. Features for the time passed since the last insulin injection and last caloric intake were also added, to aid in displaying a correlation between insulin and caloric intake and blood glucose levels. Due to spikes and a high variance in heart rate and breathing measured per second, the rolling average of those features was added to the data frame, to smooth out the data. To train the model to forecast near-term blood glucose levels, every relevant feature was copied and lagged by 15 minutes, which correlated current numerical features to blood glucose values 15 minutes ahead in the time series. Finally, glucose values above and below the thresholds for hyperglycemia and hypoglycemia were marked as such.

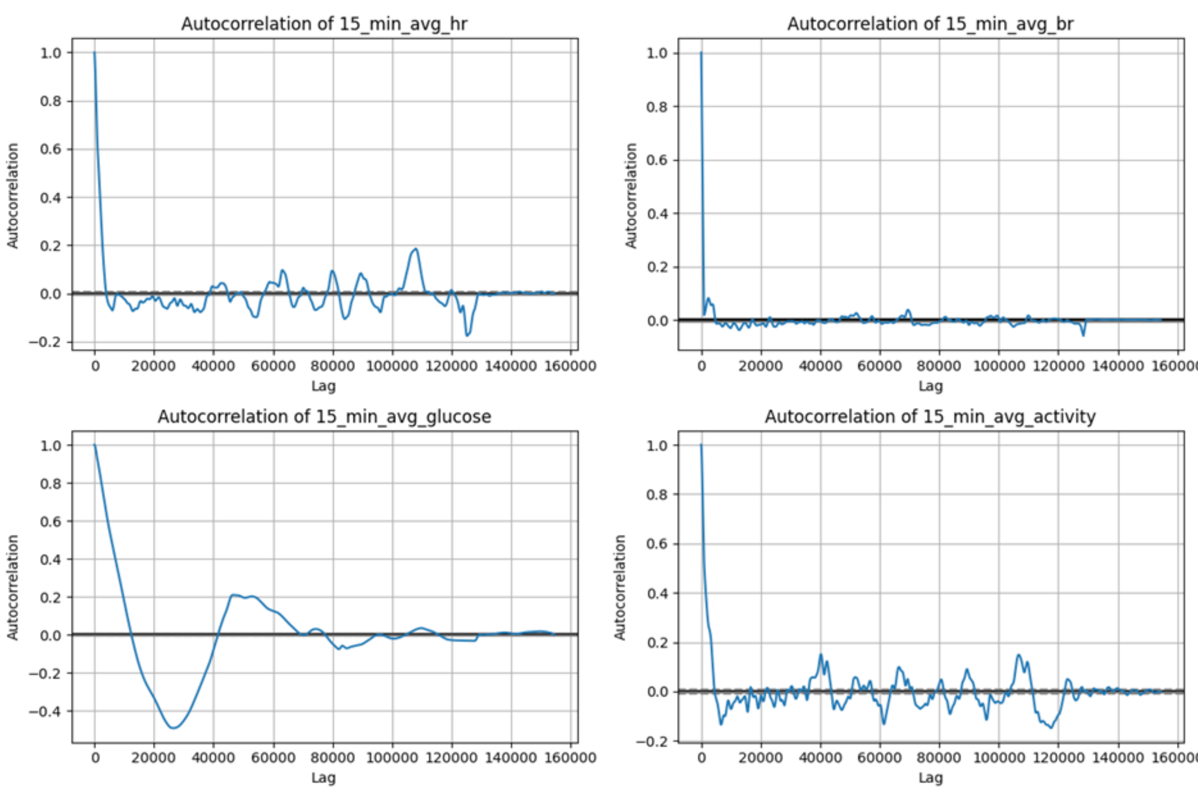


Figure 2. Autocorrelation plots of all relevant features

To address the research problem of predicting hyperglycemic and hypoglycemic events in diabetes patients using time series data, this study employed Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks. LSTM and GRU are popular choices for time series forecasting tasks due to their ability to effectively capture temporal dependencies.

LSTM is a type of recurrent neural network (RNN) designed to overcome the vanishing gradient problem in traditional RNNs. It introduces memory cells and gating mechanisms to control the flow of information through the network over time. The key components of an LSTM cell include the input gate, forget gate, output gate, and memory cell. These gates regulate the flow of information into and out of the memory cell, allowing the network to retain important information over long sequences of data.

GRU is a variant of the LSTM architecture that simplifies the gating mechanism while achieving comparable performance. The structure combines the forget and input gates into a single update gate and merges the cell and hidden states. This simplification reduces the number of parameters in the model and makes it more computationally efficient than LSTM models[1].

The LSTM model architecture consisted of a single LSTM layer with a single unit, intended to capture temporal dependencies, followed by a dense output layer with one unit to return the final blood glucose level prediction. The GRU model was structured similarly, with a single GRU layer containing one unit and a dense output layer with one unit for the final glucose output. Both models were compiled using the Adam optimizer and the mean squared error loss function to minimize large outliers in prediction errors.

The models were trained on the reshaped training data. The input data was reshaped to fit the expected input format for LSTM and GRU layers, with multiple dimensions (samples, time steps, features). For this study, the time step was set to 1. Both models were trained for 25 epochs with a batch size of 32, due to memory limitations. During each epoch, the models learned by iterating over the training data and adjusting their parameters to minimize the loss function. The MSE was monitored throughout the training process to track the models' performance and ensure proper convergence, and the number of epochs was chosen to limit excessive overfitting after testing varied amounts of iterations.

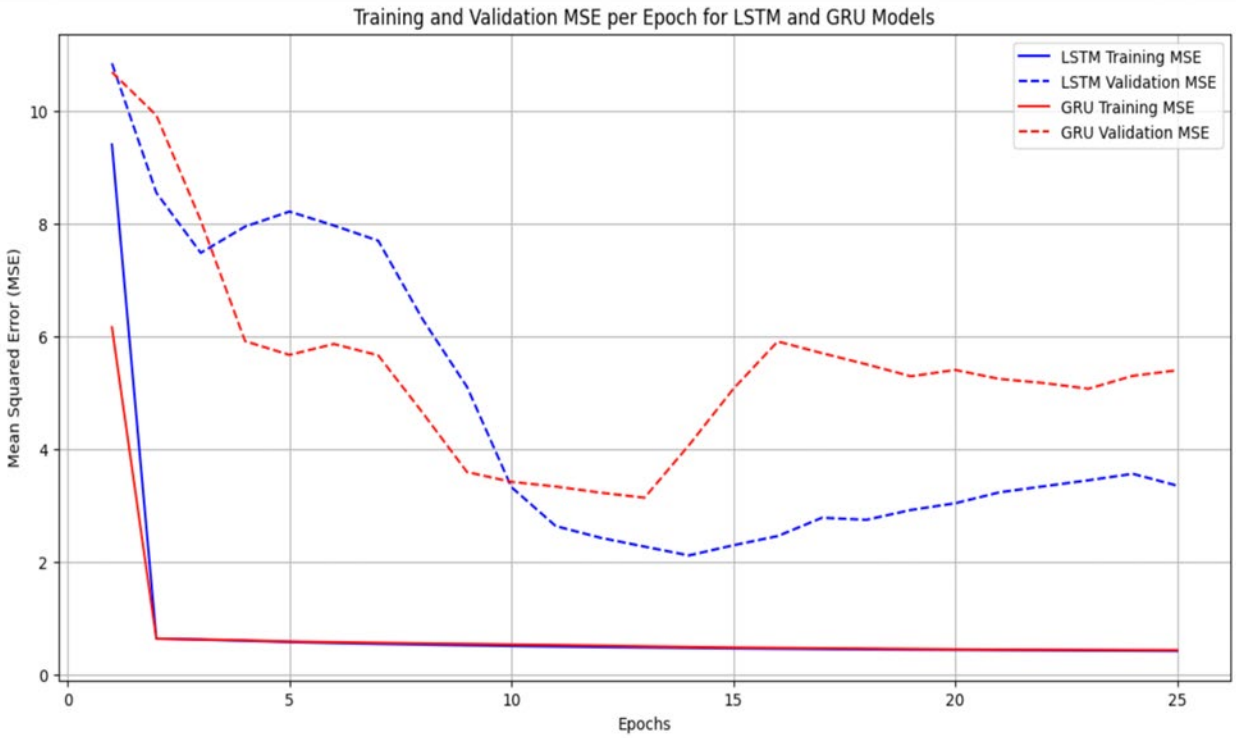
The trained models were evaluated on the testing dataset, then used to make predictions on the standardized and reshaped testing data. The mean squared error (MSE) was calculated to quantify the models' accuracies in predicting near-term blood glucose levels, and this step was repeated for 10 iterations for each model.

Results

The primary metric used to evaluate model performance was their minimum average Mean Squared Error (MSE) over all 10 iterations. The MSE provides a measure of the average squared difference between the predicted and actual blood glucose levels, with lower values indicating a better model performance.

For the GRU model, the average MSE was 3.137 $mmol^2/L^2$, while the LSTM model achieved a significantly lower average MSE of 2.043 $mmol^2/L^2$. This substantial difference indicates that the LSTM model provided more accurate blood glucose level predictions than the GRU model.

HIGH SCHOOL EDITION
Journal of Student Research



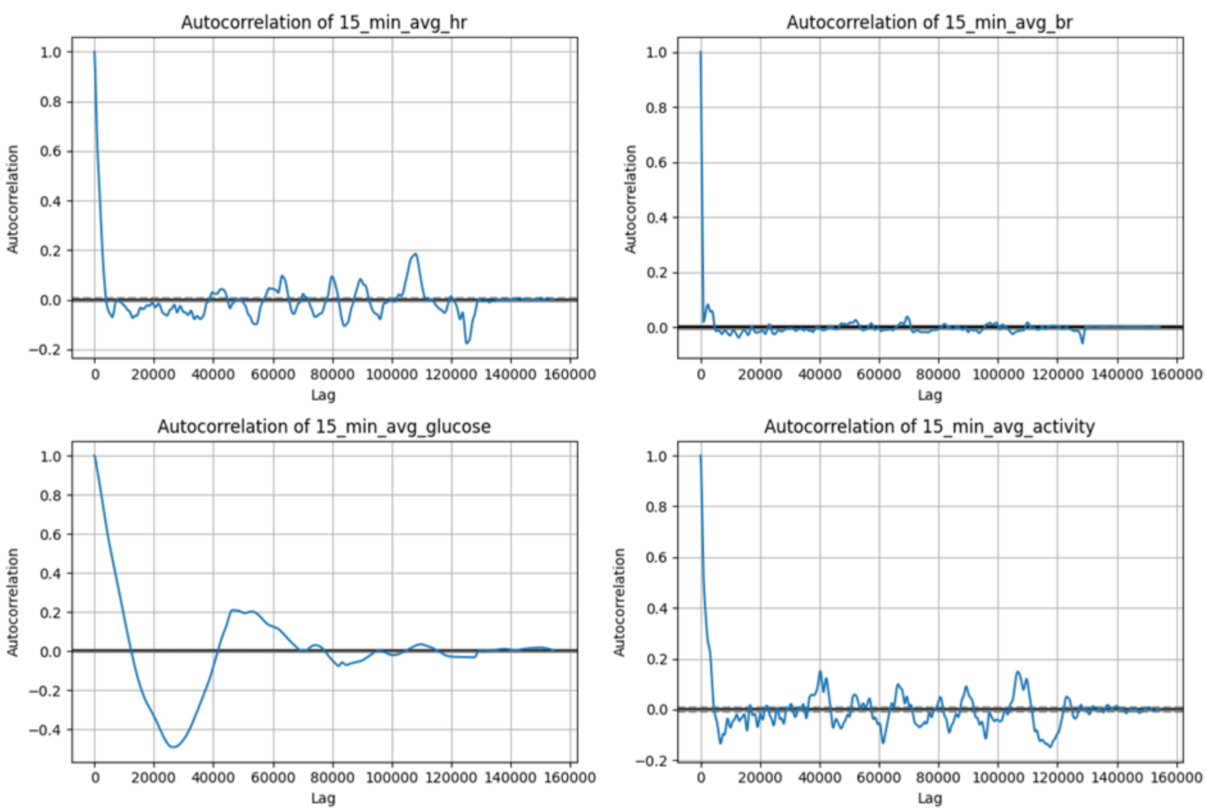


Figure 3. Plot of minimum average MSEs of both LSTM and GRU models on training and validation data

Discussion

While both LSTM and GRU models may be used in forecasting based on analyzing time series data, the results indicate that LSTM neural networks are more effective than GRU models in forecasting blood glucose levels, potentially leading to more accurate alert systems for diabetics. This can especially prove to be better at predicting trends that will help diabetics plan for and manage their blood glucose levels more effectively.

Despite the results obtained through this comparison of LSTM and GRU neural networks for predicting blood glucose levels, several limitations must be acknowledged. The OpenD1NAMO public dataset, although comprehensive, includes data from only nine patients with Type 1 diabetes[2]. This limited sample size may not fully represent the broader diabetic population, potentially leading to overfitting or larger error bounds where the model performs well on the training data but poorly on unseen data. The dataset might lack sufficient variability in lifestyle, dietary habits, and medical conditions, which could restrict the generalizability of the results for other populations.

Conclusion

This research aimed to evaluate LSTM and GRU neural networks in their ability to predict near-term blood glucose levels in diabetic patients. The study does not conclude that LSTM models are strictly more accurate than GRU models for live patient care, yet the conclusions from this study can be used as a precedent for further research. Both model structures were chosen for their ability to capture relationships in time-series data. The models' performances were evaluated on the OpenD1NAMO dataset, containing continuous physiological data including blood glucose levels, heart rate, and breathing rate. The dataset was preprocessed, accounting for missing values, scaling features, and creating lagged and rolling average features. The processed data was split into a training (80%) dataset and a testing (20%) dataset, then reshaped and inputted into both models. The LSTM and GRU neural networks were created using the Python Keras library, and their hyperparameters were tuned using grid-search testing. The LSTM model achieved a lower minimum average MSE of 2.043 $mmol^2/L^2$, compared to the GRU model's minimum average MSE of 3.137 $mmol^2/L^2$, which indicates a greater accuracy. To further increase prediction accuracy, future improvements include exploring different model types, more in-depth hyperparameter tuning, and model ensembling. This study demonstrates that single-layer LSTM models are more effective than single-layer GRU models in forecasting near-term glucose levels in Type 1 diabetic patients, given real-time physiological data.

Limitations

The selection of features and the creation of engineered features were based on assumptions about the relevance of certain parameters to blood glucose levels. While features such as heart rate, breathing rate, and average glucose levels are logical choices based on prior research, other significant factors may not be included in this study. The creation of lagged and rolling average features was intended to capture near-term effects, yet other temporal representations might better capture the characteristics of the data.

Although grid search and cross-validation were used for hyperparameter tuning, the search space and computational resources available were limited. As a result, the chosen hyperparameters may not be the optimal configuration for the models, and a more extensive search or alternative optimization techniques could improve performance for both models[3]. The tuning process was constrained by the available computational power, which limited the usage of more complex model architectures.

Furthermore, computational efficiency must be considered as well. Although GRU models are generally more computationally efficient than LSTM models, both require significant computational resources for training and inference[1]. In real-world applications, especially those involving continuous monitoring and real-time predictions, computational efficiency is a crucial characteristic. The investigation did not explore optimization techniques that could enhance the models' efficiency without compromising accuracy.

Acknowledgments

This research was supported by Inspirit AI. The author thanks their research mentor, Sebastian Bedoya, who provided insight and expertise that greatly assisted the research process.

References

1. R. Cahuantzi, X. Chen, and S. Güttel. "A Comparison of LSTM and GRU Networks for Learning Symbolic Sequences." *arXiv.Org*, 4 Jan. 2023, arxiv.org/abs/2107.02248.
2. F. Dubosson, et al. "The Open D1namo Dataset: A Multi-Modal Dataset for Research on Non-Invasive Type 1 Diabetes Management – DOAJ." *Informatics in Medicine Unlocked*, Elsevier, 1 Jan. 2018, doaj.org/article/babcdeb1b6c34747bb4b6db6a6be3bf.
3. H. Khadem, A. Razavi, and M. Baniahmad. "Blood Glucose Level Time Series Forecasting: Nested Deep Ensemble Learning Lag Fusion." *Bioengineering (Basel, Switzerland)*, U.S. National Library of Medicine, 19 Apr. 2023, www.ncbi.nlm.nih.gov/pmc/articles/PMC10135844/.
4. University of Florida Diabetes Institute. "Dietary factors, activity, and blood glucose regulation." *UF Monogram*, diabetes.ufl.edu/.
5. J.F. Trepanowski, C.M. Kroeger, A. Barnosky, et al. "Effect of Alternate-Day Fasting on Weight Loss, Weight Maintenance, and Cardioprotection among Metabolically Healthy Obese Adults: A Randomized Clinical Trial." *JAMA Internal Medicine*, U.S. National Library of Medicine, 1 July 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5680777/.
6. E.F. Sutton, R.A. Beyl, K.S. Early, et al. "Early Time-Restricted Feeding Improves Insulin Sensitivity, Blood Pressure, and Oxidative Stress Even without Weight Loss in Men with Prediabetes." *Cell Metabolism*, U.S. National Library of Medicine, pubmed.ncbi.nlm.nih.gov/29754952/. Accessed 04 Aug. 2024.
7. "Handling Hypoglycemia." *Harvard Health*, 30 Sept. 2016, www.health.harvard.edu/diseases-and-conditions/handling-hypoglycemia.
8. W.P.T.M. van Doorn, et al. "Machine Learning-Based Glucose Prediction with Use of Continuous Glucose and Physical Activity Monitoring Data: The Maastricht Study." *PloS One*, U.S. National Library of Medicine, 24 June 2021, www.ncbi.nlm.nih.gov/pmc/articles/PMC8224858/.