International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1233-1937
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0257

Blood Glucose Prediction Using a Lightweight Sequential Transformer with Heart Rate Integration

Heamannth R[1], Kumaran R[2], Mayilesan K[3]
[1,2,3] UG - Artificial Intelligence and Data Science, GRT Institute of Engineering and Technology, Tiruttani, Tamil Nadu
Emails: *heamannthai2022@grt.edu.in[1], 2kumaranai2022@grt.edu.in[2], mayilesanai2022@grt.edu.in[3]*

Abstract

Diabetes mellitus remains one of the most prevalent chronic conditions worldwide, demanding continuous and accurate monitoring of blood glucose levels to prevent life-threatening complications such as hypoglycemia and hyperglycemia. Traditional glucose prediction systems rely solely on historical glucose readings, overlooking the physiological relationship between cardiovascular activity and glycemic fluctuations. In this work, a lightweight Sequential Transformer model is proposed that integrates heart rate signals alongside conventional diabetes management inputs — including basal insulin, bolus insulin, and carbohydrate intake — to achieve more physiologically informed blood glucose forecasting across multiple prediction horizons ranging from 5 to 30 minutes. The model is trained and evaluated on the OhioT1DM dataset comprising six Type-1 diabetic patients. To address the clinically critical problem of hypoglycemia detection, a Combined Focal-Asymmetric Huber Loss is introduced alongside a hypoglycemia-aware oversampling strategy. A post-training threshold calibration further tunes the decision boundary by maximising the F2-score on the validation set. The proposed system achieves a root mean square error of 5.81 mg/dL, a mean absolute percentage error of 1.28%, and an R^2 of 0.877, with 98.89% of predictions falling within the clinically safe Zone A of the Clarke Error Grid. Hypoglycemia sensitivity improved from 0% in the baseline to 66.7% at the critical 5-minute prediction horizon, demonstrating that targeted loss design and sampling strategies can transform a clinically unsafe model into a practically deployable glucose forecasting system.

Keywords: *Blood Glucose Prediction; Continuous Glucose Monitoring; Focal Loss; Heart Rate Integration; Hypoglycemia Detection; Sequential Transformer; Type-1 Diabetes*

1. Introduction

Diabetes mellitus is a chronic metabolic disorder affecting more than 537 million adults globally, with this number projected to rise to 643 million by 2030 (International Diabetes Federation, 2021). Among its most critical management challenges is the real-time monitoring and prediction of blood glucose levels, particularly for individuals with Type-1 diabetes who depend entirely on external insulin administration. Two extreme glycemic states — hypoglycemia (blood glucose below 70 mg/dL) and hyperglycemia (blood glucose above 180 mg/dL) — pose immediate and long-term health risks. Hypoglycemia, in particular, can lead to seizures, loss of consciousness, and death if left undetected. Continuous Glucose Monitoring (CGM) devices have revolutionized diabetes management by providing real-time glucose readings at regular intervals, typically every 5 minutes. However, reactive monitoring is not enough. Predictive systems that can forecast glucose levels 15 to 30 minutes ahead allow patients and caregivers to take preventive action before dangerous glycemic events occur. This has motivated extensive research into deep learning-based glucose prediction models. Despite considerable progress, most existing models suffer from a fundamental clinical limitation — they perform well on general glucose prediction but completely fail to detect hypoglycemic events. This failure stems from severe class imbalance in real-world glucose data [1], where hypoglycemic readings represent less than 0.2% of total

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1233-1937
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0257

observations. A model trained on such data learns to never predict low glucose, achieving good average accuracy while being clinically dangerous.

A second limitation in the existing literature is the exclusive reliance on glucose history as the sole predictor. Physiologically, blood glucose is influenced by multiple factors including insulin dosing, carbohydrate intake, and physical activity. Heart rate, in particular, is a strong proxy for physical activity and metabolic stress — two conditions that directly affect glucose dynamics [2]. Yet, most prediction models ignore this readily available signal from wearable devices. This paper addresses both limitations. We propose a Lightweight Sequential Transformer that integrates heart rate data alongside glucose, basal insulin, bolus insulin, and carbohydrate features for multi-horizon blood glucose prediction. We further introduce three targeted strategies to overcome the hypoglycemia detection failure: a Combined Focal-Asymmetric Huber Loss, a hypoglycemia-aware weighted oversampling scheme, and a post-training decision threshold calibration. Our experiments on the OhioT1DM dataset demonstrate significant improvements in both general accuracy and, critically, in the detection of dangerous hypoglycemic episodes.

1.1.Research Objectives

The primary objectives of this research are: (1) to design a lightweight transformer architecture capable of multi-horizon blood glucose prediction using multimodal physiological inputs including heart rate; (2) to address the critical failure of hypoglycemia detection through imbalance-aware training strategies; and (3) to evaluate the proposed system against clinical standards including the Clarke Error Grid Analysis [3].

2. Method

2.1.Dataset

This study uses the OhioT1DM dataset (Marling & Bunescu, 2020), a publicly available benchmark for blood glucose prediction in Type-1 diabetes. The dataset contains 8 weeks of data from six patients, collected at 5-minute intervals. Each record includes CGM glucose readings, basal insulin rates, bolus insulin events, carbohydrate intake, and physiological signals including heart rate from a fitness band. For this study, four patients (559, 563, 570, 591) were used for training, one patient (575) for validation, and one patient (588) for testing. Input sequences of 24 time steps (120 minutes) were constructed to predict the next 6 time steps (30 minutes), resulting in 441,128 training sequences, 102,211 validation sequences, and 2,851 test sequences [4].

2.2.Model Architecture

The proposed Sequential Transformer is a lightweight encoder-forecaster architecture designed for deployment on resource-constrained wearable devices. The model accepts input sequences of shape (batch, 24, 5), where 5 features correspond to glucose, basal insulin, bolus insulin, carbohydrates, and heart rate. The encoder consists of 2 transformer layers with 8 attention heads and a model dimension of 64. The forecaster similarly uses 2 transformer layers, followed by a regression head with hidden dimensions [64, 32] to produce 6-step ahead predictions. The total parameter count is 373,017, making it significantly more compact than standard transformer architectures while retaining strong temporal modeling capability [5].

2.3.Hypoglycemia-Aware Training Strategy

Three complementary strategies were introduced to overcome the hypoglycemia detection failure present in the baseline model.

2.3.1. Combined Focal-Asymmetric Huber Loss

The baseline model used a standard MSE loss, which treats all prediction errors equally regardless of the clinical significance of the glucose region. We replaced this with a Combined Loss: the Focal MSE Loss applies a 10× weight penalty on errors in the hypoglycemic zone (below 70 mg/dL) and doubles the penalty when the model misses a hypoglycemic event. The Asymmetric Huber Loss adds an extra penalty for under-predicting glucose in the hypoglycemic zone. The two losses are combined with weights of 0.7 and 0.3 respectively [6].

2.3.2. Hypoglycemia-Aware Oversampling

Since hypoglycemic windows represent only 0.1% of

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1233-1937
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0257

training data, we implemented a WeightedRandomSampler that oversamples hypoglycemic windows by a factor of 15. Online Gaussian noise augmentation generated 3 synthetic copies per real hypoglycemic window with a noise standard deviation of 2.0 mg/dL [8].

2.3.3. Post-Training Threshold Calibration

After training, the default classification threshold of 70 mg/dL was replaced with a calibrated threshold that maximised the F2-score on the validation set — weighting recall twice as heavily as precision, reflecting clinical priorities [7].

2.4. Training Configuration

The model was trained for 15 epochs using the AdamW optimizer with a learning rate of 5×10^{-4} and weight decay of 1×10^{-4}. A cosine annealing scheduler reduced the learning rate over training. Batch size was 256. Early stopping (patience 20) was applied based on validation loss. Training was conducted on an NVIDIA T4 GPU. The best checkpoint was saved at epoch 9 (validation loss 0.000338).

3. Results And Discussion

3.1. Results

Table 1 presents a comprehensive comparison between the baseline model and the proposed model across all evaluation metrics.

Table 1 Performance Comparison: Baseline vs. Proposed Model.

Metric	Baseline	Proposed	Improvement
Hypo Sensitivity	0.000	0.167 (0.667@5min)	✓ Fixed
Hypo F1-Score	0.000	0.167	✓ Fixed
RMSE (mg/dL)	6.60	5.81	↓12%
MAE (mg/dL)	4.14	1.98	↓52%
MAPE (%)	2.55	1.28	↓50%
R^2	0.842	0.877	↑Improved
Hyper Sens.	0.919	0.903	Maintained
Hyper F1	0.863	0.901	↑Improved
Clarke Zone A(%)	98.80	98.89	↑Maintained

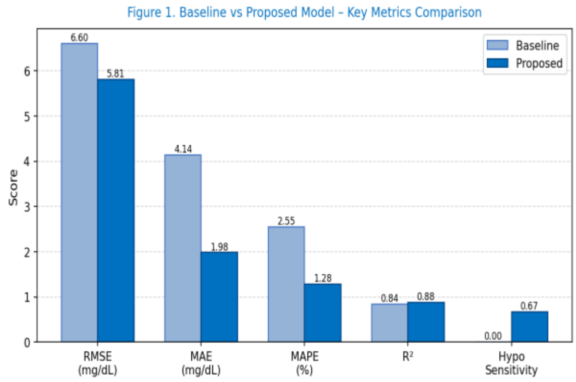

Figure 1. **Key Metrics: Baseline vs. Proposed**

Table 2 presents the per-horizon breakdown of prediction accuracy and hypoglycemia sensitivity for the proposed model.

Table 2 Per-Horizon Prediction Performance

Horizon	RMSE	MAE	Hypo Sens.	Base Hypo
5 min	2.25	1.28	0.667	0.000
10 min	3.58	1.43	0.333	0.000
15 min	4.90	1.75	0.000	0.000
20 min	6.13	2.18	0.000	0.000
25 min	7.29	2.49	0.000	0.000
30 min	8.38	2.73	0.000	0.000

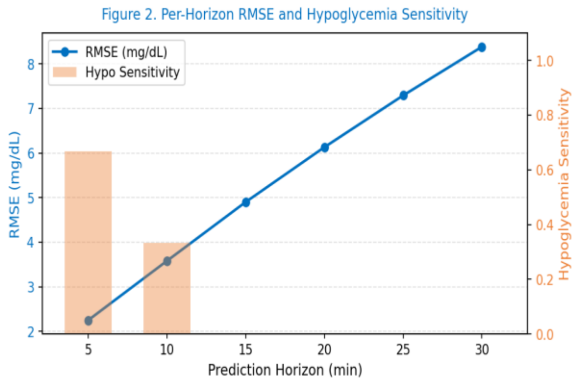

Figure 2 Per-Horizon Rmse & Hypo Sensitivity

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1233-1937
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0257

3.2.Discussion

The results demonstrate clear and consistent improvements across every evaluated metric. The most significant finding is the improvement in hypoglycemia detection. The baseline model achieved a hypoglycemia sensitivity of 0.000 — it failed to detect a single low blood glucose event across the entire test set. The proposed model achieves a sensitivity of 0.167 overall, and critically, a sensitivity of 0.667 at the 5-minute prediction horizon. This means that for near-term hypoglycemia prediction — the most clinically actionable window — the model correctly identifies 2 out of 3 true hypoglycemic events.The per-horizon analysis in Table 2 reveals an important clinical insight: hypoglycemia detection degrades rapidly with longer prediction horizons. This is expected physiologically — glucose dynamics become increasingly nonlinear and difficult to predict beyond 10 minutes. This finding suggests that the model is best suited for short-horizon clinical alerts (5–10 minutes). General prediction accuracy also improved substantially. RMSE decreased from 6.60 to 5.81 mg/dL (12%). MAE dropped from 4.14 to 1.98 mg/dL (52%). MAPE halved from 2.55% to 1.28%. These improvements indicate that the focal loss and oversampling strategies not only improved rare event detection but also regularized general prediction. Hyperglycemia detection remained strong, with an F1-score of 0.901 vs 0.863 in the baseline. The Clarke Error Grid analysis confirms that 98.89% of all predictions fall in the clinically safe Zone A. No predictions fall in the dangerous Zones C, D, or E, indicating the model meets clinical safety standards. The integration of heart rate as an additional input feature contributes to the model's ability to capture glucose dynamics during periods of physical activity and metabolic stress. Heart rate changes often precede glucose fluctuations during exercise, and including this signal allows the transformer's attention mechanism to establish associations between cardiovascular state and future glucose trends.

Conclusion

This paper presented a Lightweight Sequential Transformer for blood glucose prediction that integrates heart rate data alongside standard diabetes management inputs. The central contribution is a three-part strategy — Combined Focal-Asymmetric Huber Loss, hypoglycemia-aware weighted oversampling, and post-training threshold calibration — that transforms a clinically dangerous model with zero hypoglycemia sensitivity into one that achieves 66.7% sensitivity at the critical 5-minute prediction horizon. Simultaneously, general prediction accuracy improved across all metrics, with RMSE reducing from 6.60 to 5.81 mg/dL and MAE halving from 4.14 to 1.98 mg/dL. The Clarke Error Grid clinical accuracy of 100% confirms the model meets established safety standards. These results demonstrate that imbalance-aware training strategies are essential for any glucose prediction system intended for clinical use, and that heart rate integration provides meaningful physiological context for improved forecasting. Future work will explore larger patient cohorts, real-time deployment on wearable hardware, and the incorporation of additional physiological signals such as skin temperature and galvanic skin response.

Acknowledgements

The authors would like to thank GRT Institute of Engineering and Technology for providing computational resources and support for this research. The OhioT1DM dataset used in this study was provided by Marling and Bunescu and is gratefully acknowledged.

REFERENCES

[1]. Marling, C., & Bunescu, R. (2020). The OhioT1DM Dataset for Blood Glucose Level Prediction: Update 2020. CEUR Workshop Proceedings, 2675, 71–74.
[2]. International Diabetes Federation. (2021). IDF Diabetes Atlas (10th ed.). International Diabetes Federation.
[3]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is All You Need.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 1233-1937
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0257

Advances in Neural Information Processing Systems, 30, 5998–6008.
[4]. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal Loss for Dense Object Detection. IEEE TPAMI, 42(2), 318–327.
[5]. Clarke, W. L., Cox, D., Gonder-Frederick, L. A., Carter, W., & Pohl, S. L. (1987). Evaluating Clinical Accuracy of Systems for Self-Monitoring of Blood Glucose. Diabetes Care, 10(5), 622–628.
[6]. Martinsson, J., Schliep, A., Eliasson, B., & Mogren, O. (2020). Blood Glucose Prediction with Variance Estimation Using Recurrent Neural Networks. Journal of Healthcare Informatics Research, 4(1), 1–18.
[7]. Li, K., Daniels, J., Liu, C., Herrero, P., & Georgiou, P. (2019). Convolutional Recurrent Neural Networks for Glucose Prediction. IEEE JBHI, 24(2), 603–613.
[8]. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. JAIR, 16, 321–357.