The University of Hong Kong
Department of Computer Science
Artificial Intelligence of Things Lab

Predicting Adverse Events for Patients with Type-1 Diabetes Via Self-Supervised Learning

Xinzhe Zheng Sijie Ji Chenshu Wu
{xzzheng, sijieji, chenshu}@cs.hku.hk

1 Introduction

Predicting blood glucose levels is fundamental for precise primary care of type-1 diabetes (T1D) patients. However, it is challenging to predict glucose levels accurately, not to mention the early alarm of adverse events (**hyperglycemia** and **hypoglycemia**), namely the *minority class*. In this paper, we propose **BG-BERT**, a novel self-supervised learning framework for blood glucose level prediction.

2 Contributions

1. **Contextual Modeling**: **BG-BERT** effectively models contextual information in blood glucose monitoring data, capturing trends and the influence of observed data on future readings.
2. **Addressing Limited Data**: **BG-BERT** addresses the limited availability of blood glucose data within adverse events by employing data augmentation techniques and a bias-free training process.
3. **Improved Performance**: **BG-BERT** outperforms existing models in predicting blood glucose levels within adverse events, achieving higher accuracy.
4. **Open-source Framework**: **BG-BERT** is an open-source framework available on GitHub: `https://github.com/aiot-lab/BG-BERT`.

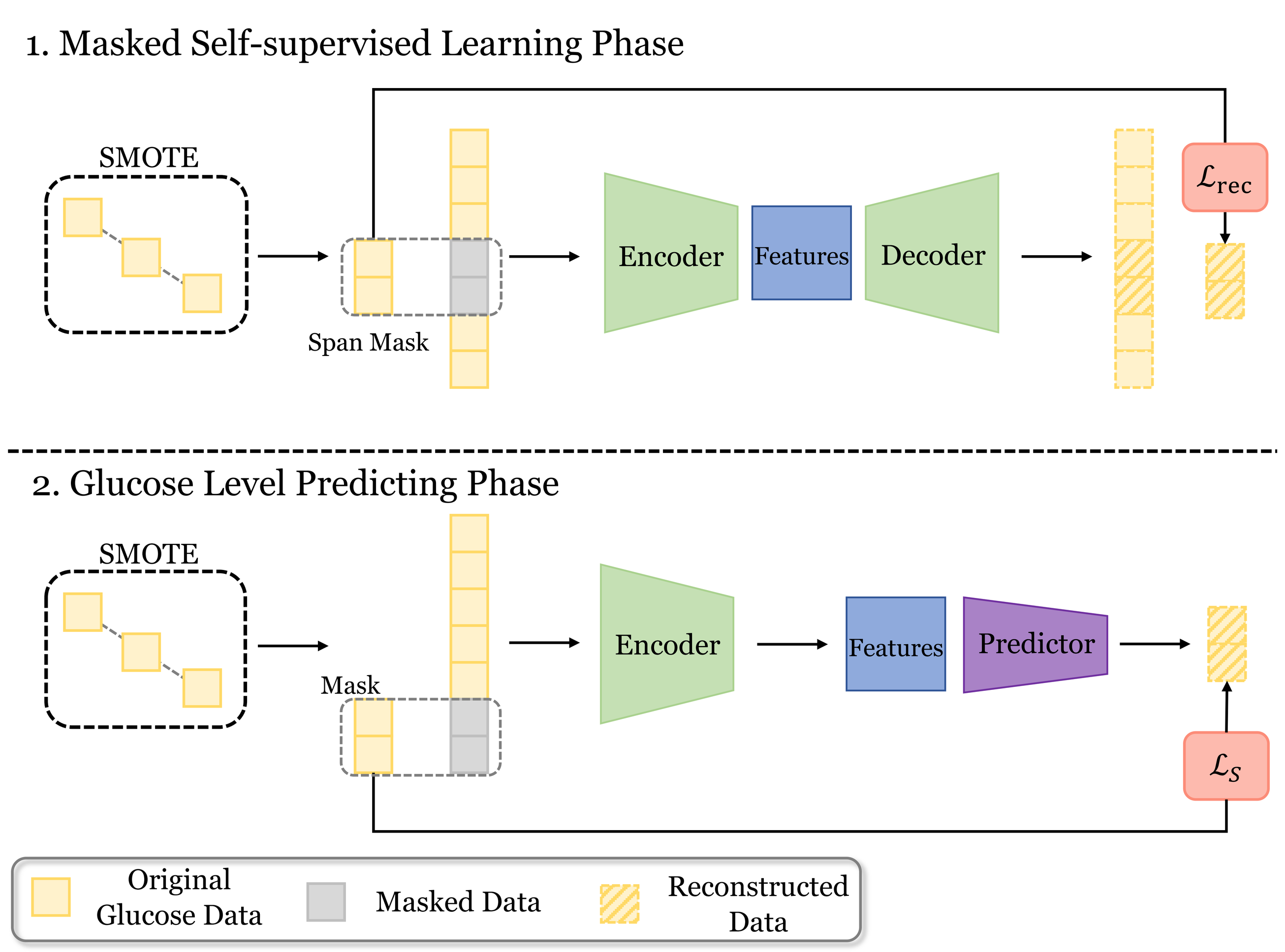

Figure 1: System Workflow

3 Methodology

- **Masked Self-supervised Learning Phase**: Automatically learn rich contextual information from glucose readings.
- **Glucose Level Prediction Phase**: Forecast glucose levels based on learned representations.

In addition, to address the imbalanced data issue within the range of hyperglycemia and hypoglycemia. We incorporate SMOTE data augmentation and a customized shrinkage loss function.

3.1 SMOTE Data Augmentation

the synthetic instances are generated by combining a specific sample from adverse events with one of its k nearest neighbors. The calculation for generating the synthetic object is as follows:

$$\mathbf{X}_s = \mathbf{X}_i + r \cdot (\mathbf{X}_b - \mathbf{X}_i), \quad (1)$$

where $\mathbf{X}_s$ is the synthetic instance, $\mathbf{X}_i$ is one sample from adverse events, and $\mathbf{X}_b$ denotes the neighbor sample of $\mathbf{X}_i$.

3.2 Shrinkage Loss

We amplify large loss values and diminish small loss values during back-propagation, implementing the concept of focal loss for regression tasks. The formula of the shrinkage loss is given as:

$$\mathcal{L}_s = \frac{||\hat{\mathbf{g}} - \mathbf{g}||^2}{1 + \exp(a \cdot (c - ||\hat{\mathbf{g}} - \mathbf{g}||^2))}, \quad (2)$$

where $\hat{\mathbf{g}}$ is the golden standard, $\mathbf{g}$ is the predictions. a and c are hyper-parameters of shrinkage loss.

4 Results

We evaluate **BG-BERT** on two benchmark datasets: OhioT1DM and Diatrend, with two SOTA baseline models: DRTF and MT-NB-L (supervised-learning).

Horizon	30 mins							
Dataset	OhioT1DM				Diatrend			
Metric	RMSE (mg/dL)	TG (mins)	Sen Hype (%)	Sen Hypo (%)	RMSE (mg/dL)	TG (mins)	Sen Hype (%)	Sen Hypo (%)
DRTF	18.21	15.54	80.67	53.58	15.23	15.07	80.28	39.12
MT-NB-L	21.50	14.74	61.36	34.05	19.80	13.88	75.16	39.89
w/o aug	14.38	16.19	81.49	64.41	15.01	16.27	79.53	56.69
w/o Ls	**13.92**	15.72	81.16	70.75	15.13	16.16	80.75	57.42
BG-BERT	14.02	**16.56**	**82.54**	**73.24**	**14.85**	**16.47**	**81.34**	**62.27**

Table 1: Evaluation Results (30mins). TG: temporal gain, which indicates the amount of average time gained for early detection of a potential adverse event; Sen: sensitivity.

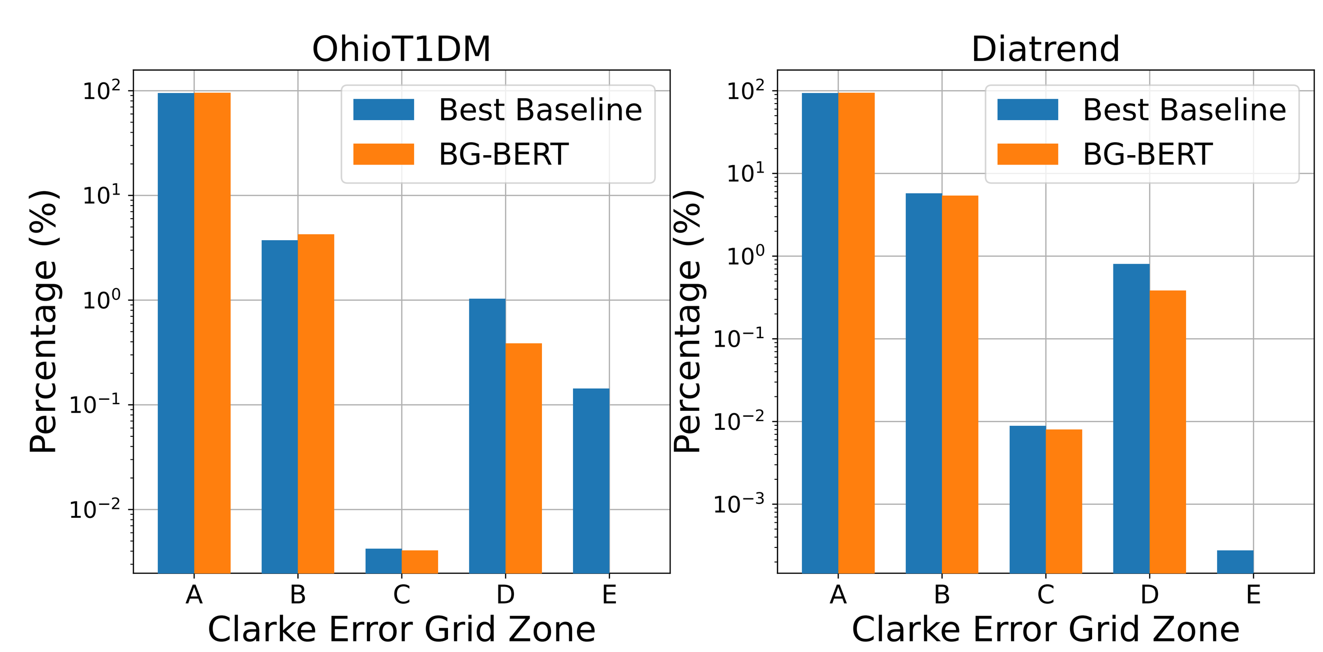

Figure 2: Clarke error grid analysis. (A: medically accurate result, B: medically acceptable, C: unnecessary treatment, D: failure to detect a dangerous condition, E: mistaking adverse events.)

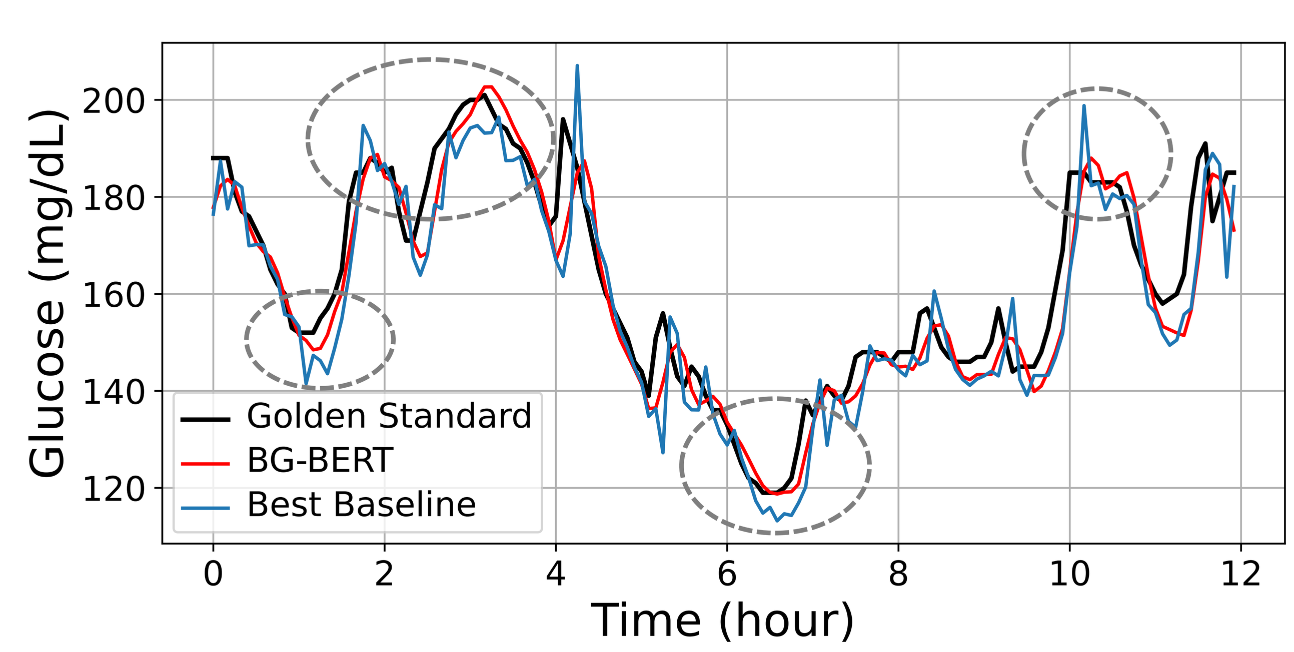

Figure 3: Visualization of half-day glucose prediction. The gray circles highlight the better performance on turning points.