The Ice Cream Experiment – A Real-Life Experiment of Machine Learning-Driven Blood Glucose Prediction

Miriam K. Wolff, Norwegian University of Science of Technology, Department of ICT and Natural Sciences

Accurate blood glucose (BG) prediction is crucial in diabetes management, potentially aiding insulin decisions and automated dosing. While machine learning models claim robustness using traditional metrics, their real-world performance is underexplored [1]. This study compares two established models, Ridge Regressor and Gradient Boosting Regressor, which had demonstrated superior performance in a prior benchmarking study [2]. The prediction model training and test stages are described in Fig. 1. The models are trained on real-life data from a type 1 diabetes patient. We evaluate their predictive accuracy using Root Mean Squared Error (RMSE) and assess performance in extreme scenarios—ice cream consumption without insulin and insulin-induced hypoglycemia.

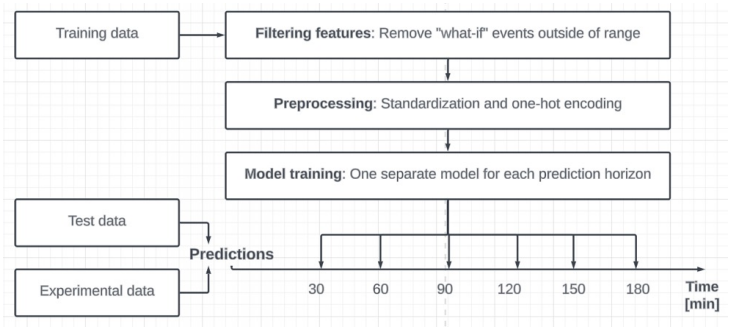

Fig. 1. Prediction model training and test stages.

Two BG prediction models were evaluated using real data from a diabetic patient. Inputs included BG values, insulin, and carbohydrate intake with time lags. RMSE and plots assessed predictive accuracy. Extreme scenarios were tested: ice cream-induced hyperglycemia and insulin-induced hypoglycemia, evaluating model predictions against actual BG outcomes.

Interestingly, the Gradient Boosting Regressor predicted almost the same value in all scenarios. Despite this, it still had the lower RMSE. As observed in Table 1, the RMSE values for a 30-minute prediction horizon were 1.13 mmol/L for Ridge Regressor and 1.04 mmol/L for Gradient Boosting Regressor, with negligible deviation from benchmarks [2]. In Scenario 1, presented in Fig. 2, models failed to predict extreme hyperglycemia. Followingly, in Scenario 2, the models did not anticipate insulin-induced hypoglycemia, see Fig. 3. This study, limited to n=1, highlights the common challenge of models struggling to capture infrequent yet vital real-world scenarios.

This study's novelty lies in incorporating outlier data from real-life experiments and assessing BG prediction models in diverse contexts. Both models exhibited consistent performance with benchmarking. However, they failed in extreme scenarios, highlighting the importance of combining traditional metrics with practical assessments. Future research should further assess prediction models in real-life applications, addressing the gap between theoretical assessments and practical use in diabetes management.

Table 1. RMSE across models and prediction horizons.

	RMSE [mmol/L] for prediction horizon [min]					
Model	**30**	**60**	**90**	**120**	**150**	**180**
Ridge Regressor	1.13	1.66	1.99	2.27	2.46	2.58
Gradient Boosting	1.04	1.59	1.97	2.24	2.41	2.51

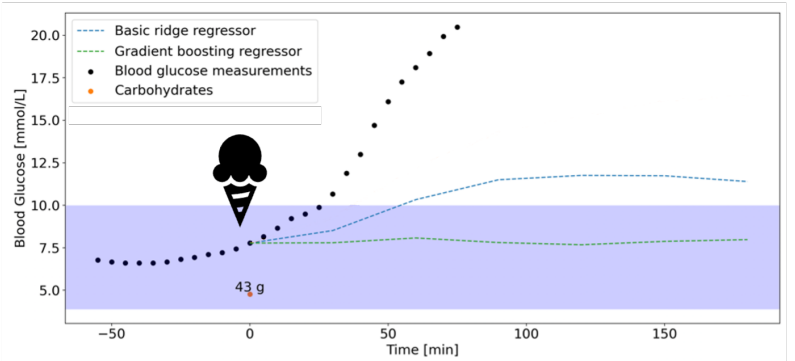

Fig. 2. BG predictions and outcome after ice cream consumption without insulin injection.

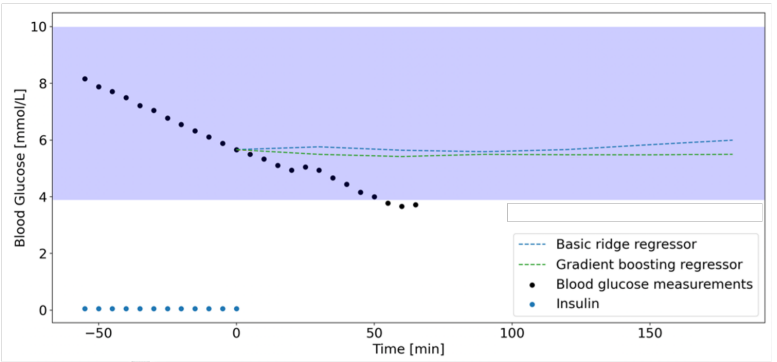

Fig. 3. BG predictions and outcome after invoking hypoglycemia.

[1] O. Diouri *mfl.*, «Hypoglycaemia detection and prediction techniques: A systematic review on the latest developments», *Diabetes Metab. Res. Rev.*, bd. 37, nr. 7, s. e3449, 2021, doi: 10.1002/dmrr.3449.

[2] J. Xie og Q. Wang, «Benchmarking Machine Learning Algorithms on Blood Glucose Prediction for Type I Diabetes in Comparison With Classical Time-Series Models», *IEEE Trans. Biomed. Eng.*, bd. 67, nr. 11, s. 3101–3124, nov. 2020, doi: 10.1109/TBME.2020.2975959.