EXPLAINABLE PREDICTION OF GLYCATED HEMOGLOBIN RESPONSE BEFORE AND AFTER DRUG INITIATION WITH ITE IN FINNISH TYPE 2 DIABETES PATIENTS USING EVIDENCE-BASED MACHINE LEARNING

Gunjan Chandra[1], Anusha Ihalapathirana[1], Piia Lavikainen[2], Pekka Siirtola[1], Satu Tamminen[1], Janne Martikainen[2], and Juha Röning[1]

[1] Biomimetics and Intelligent Systems Group, P.O. BOX 4500, FI-90014, University of Oulu, Oulu, Finland

[2] School of Pharmacy, University of Eastern Finland, Kuopio, Finland

Introduction

This study predicts HbA1c changes in type 2 diabetes patients using machine learning. Demographic and medical data are analyzed with regression algorithms to find the most accurate model, evaluated using metrics like R Square and RMSE. Results support machine learning for personalized glucose management. Combining demographics, lifestyle, and clinical data improves prediction accuracy, and adding post-drug initiation values enhances performance. The second study develops a treatment selection model for DPP-4 vs. SGLT2 inhibitors, categorizing patients into concordant (aligned with prediction) and discordant groups. The study measures outcome improvement (HbA1C after 12 months) in the concordant vs. discordant group by mean change from baseline.

Methodologies

Data were obtained as electronic health records (EHR) from the Siun sote to identify a population-based cohort of patients diagnosed with T2D at the end of 2012. Data provided us with information on utilising both primary health care and specialised care for all patients living in the North Karelia region. For predicting treatment response in diabetes patients, we employed two machine learning models: The first model utilized baseline values before drug initiation, while the second model incorporated follow-up HbA1c values after drug initiation to predict long-term glycated hemoglobin responses. These models also included the expected HbA1c change as a predictor during training, obtained from the randomized controlled trial (RCT) data for various drugs. Regression analysis was performed to create fitted lines and visualize the relationship between predictors and treatment response. We also generated confidence and prediction intervals to assess the uncertainty associated with model predictions, offering clinicians a range of plausible outcomes.

Experiments and results

We employed a regression approach to predict future HbA1c values after initiating blood glucose-lowering medication. Figure 1 illustrates the performance of the regression model, and the drug names are detailed in Table 1. Figure 1 comprises two sub-figures: the left one leverages baseline values before drug initiation, and the right one incorporates follow-up values. Both models demonstrated strong performance, with the second model achieving an impressive R2 score of 0.82. In Figure 2, we present SHAP summary plots for the MLPRegressor models using follow-up HbA1c values post-drug initiation. Figure 3, we provide a SHAP decision plot for additional insights into the model's decision-making process. Finally, Figure 4 illustrates the SGLT2 inhibitor and DPP-4 inhibitor treatment selection model.

Conclusion

The study reveals that the treatment response of patients with T2D can be predicted with high accuracy. As the treatment response is predicted with confidence, different treatment options can be tested for each patient, and the optimal treatment can be selected with much higher certainty.

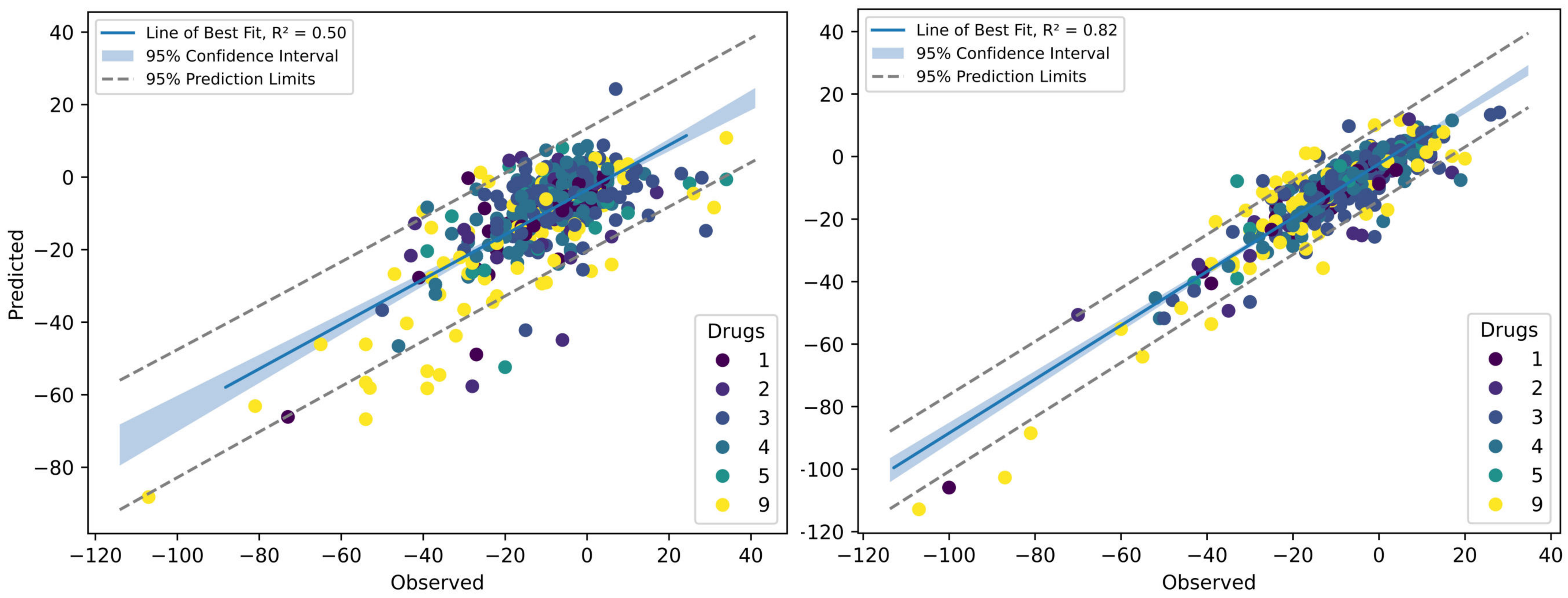

Figure 1. Performance of MLPRegressor Model: Fitted Regression Line for HbA1c Change Before and After Drug Initiation. Base model on the left and model using follow-up HbA1c value after drug initiation on the right.

Table 1. List of drug code and their name.

Code	Name
1	Metformin
2	GLP-1 analogues
3	DPP-4 inhibitors
4	SGLT2 inhibitors
5	Combinations of oral blood glucose lowering drugs
9	Insulin

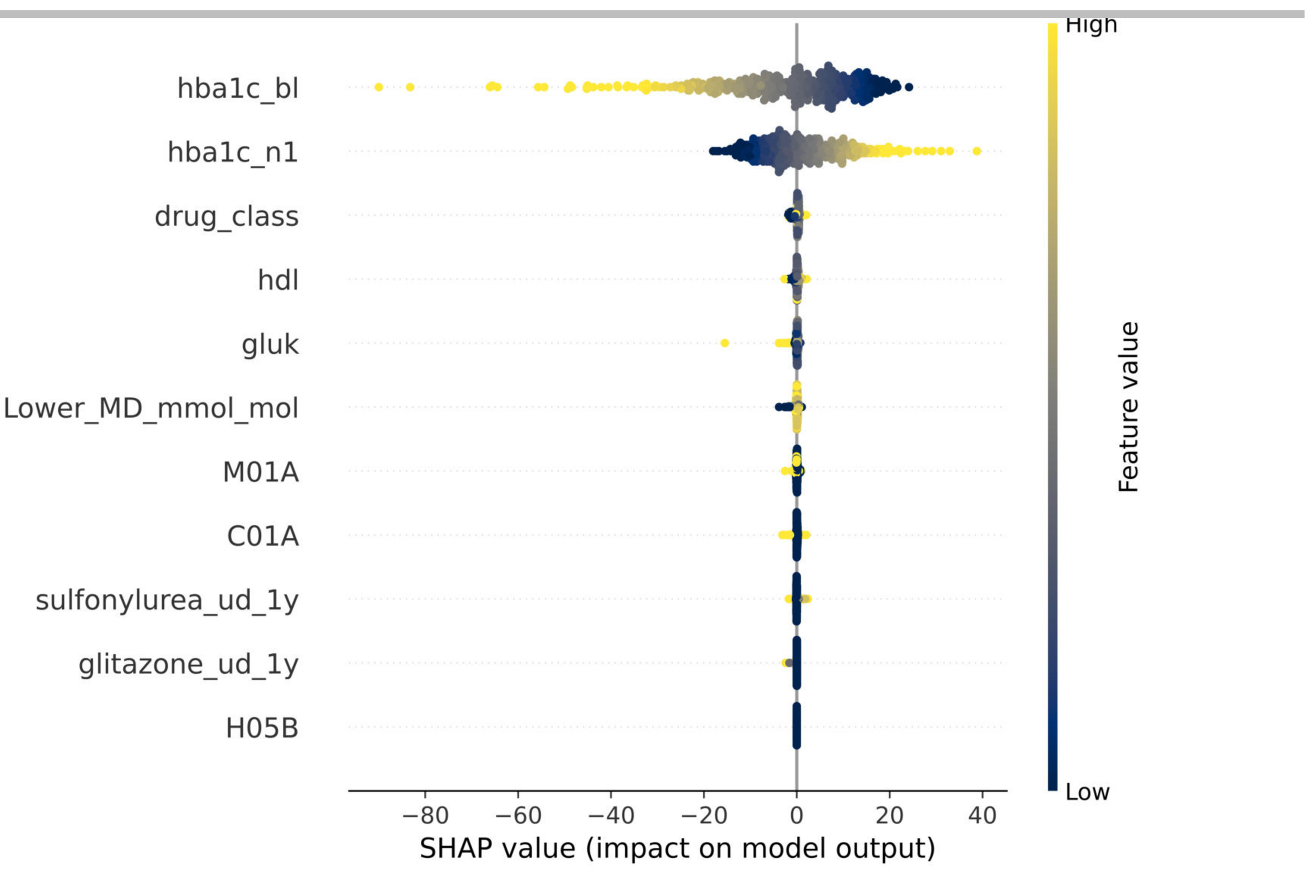

Figure 2. SHAP summary plot of the MLPRegressor models using follow-up HbA1c value after drug initiation.

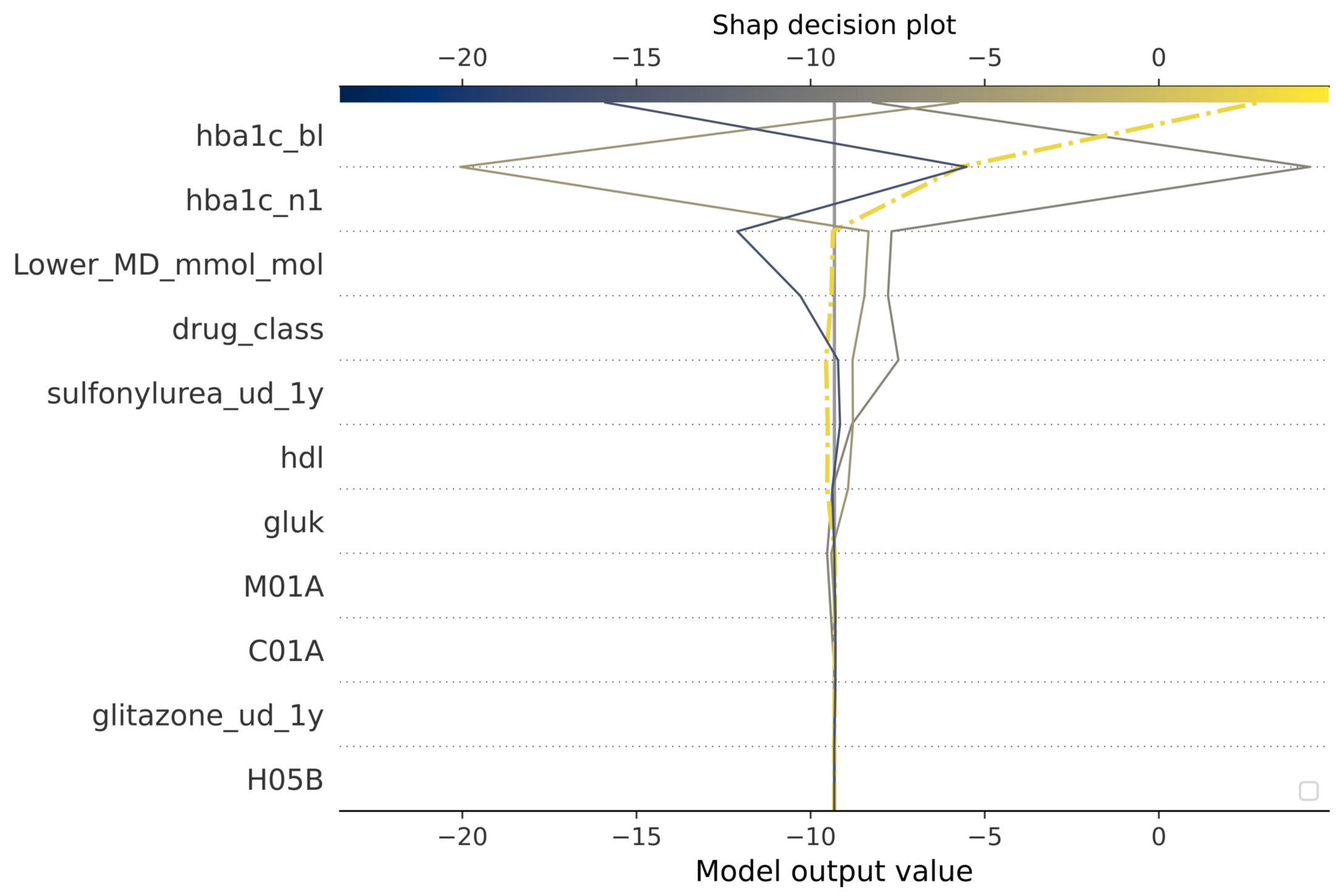

Figure 3. SHAP decision plot of the MLPRegressor models using follow-up HbA1c value after drug initiation.

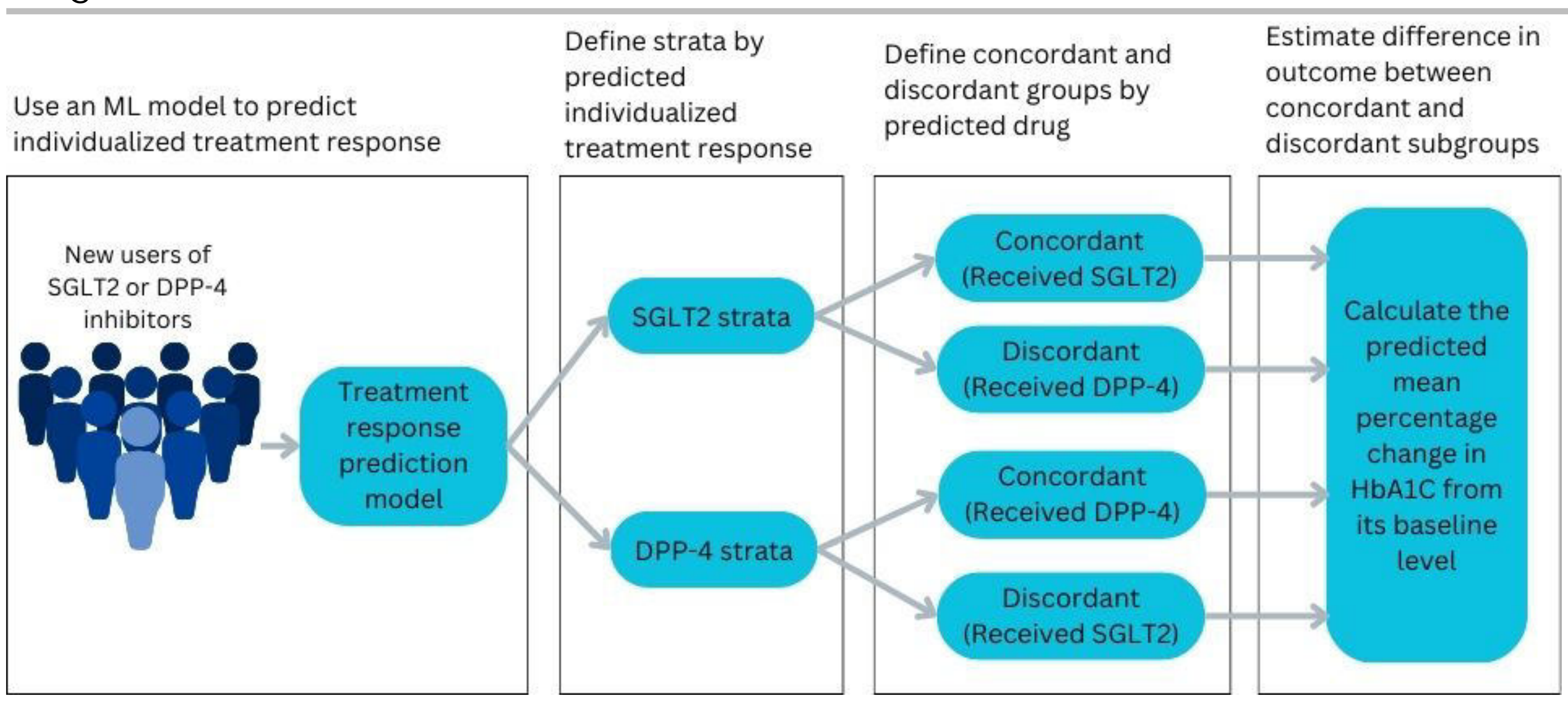

Figure 4. SGLT2 inhibitor and DPP-4 inhibitor treatment selection model.