Explainable Artificial Intelligence to predict clinical outcomes for adults with Type 1 Diabetes

Anusha Ihalapathirana, Pekka Siirtola, Gunjan Chandra, Satu Tamminen and Juha Röning

Biomimetics and Intelligent Systems Group, P.O. BOX 4500, FI-90014, University of Oulu, Oulu, Finland
{anusha.ihalapathirana, pekka.siirtola, gunjan.chandra, satu.tamminen, juha.roning}@oulu.fi

Introduction

Type 1 diabetes patients are prone to life-threatening conditions. Severe hypoglycemia (SH) and diabetic ketoacidosis (DKA) are such conditions that often require urgent hospital care. The objective of this study is to implement an AI-based explainable solution to predict possible SH and DKA events in T1D patients within the next 12 months. The initial models of this study were built with baseline factors identified in prior research. These models were further improved by introducing more features and separating the population by gender. The final models were used to build a decision support system that facilitates precision medicine by prioritizing the high-risk patient group. In addition, it helps to potentially reduce medical expenses through more efficient resource management.

Experiments and Results

In the initial step, the models were developed using the features mentioned in the replication study.

Prediction models with prior study findings

Both the SH and the DKA prediction models were built with 13 baseline features including patients' socioeconomic status, such as annual income, private insurance, and education level.

	Model	Class wise accuracy		Balanced accuracy		Predicted labels		F1 score
		Positive	Negative			Positive (P)	Negative (N)	
SH	LGBM	0.76	0.57	0.66	P	39	12	0.72
					N	692	921	
DKA	AdaBoost	0.70	0.64	0.67	P	35	15	0.64
					N	527	966	

Table 1:Results of the models with baseline factors

Prediction Models with Improvements

The best results for SH prediction were achieved with gender-disaggregated models that use 147 features to predict the outcome. A single prediction model that used all the features in the preprocessed data set achieved the best result in predicting DKA events.

	Model	Class wise accuracy		Balanced accuracy		Predicted labels		F1 score
		Positive	Negative			Positive (P)	Negative (N)	
SH Male	LGBM	1.0	0.72	0.86	P	20	0	0.73
					N	205	530	
SH Female	LGBM	0.85	0.72	0.78	P	23	4	0.72
					N	243	634	
DKA	LGBM	0.88	0.77	0.82	P	44	6	0.77
					N	340	1153	

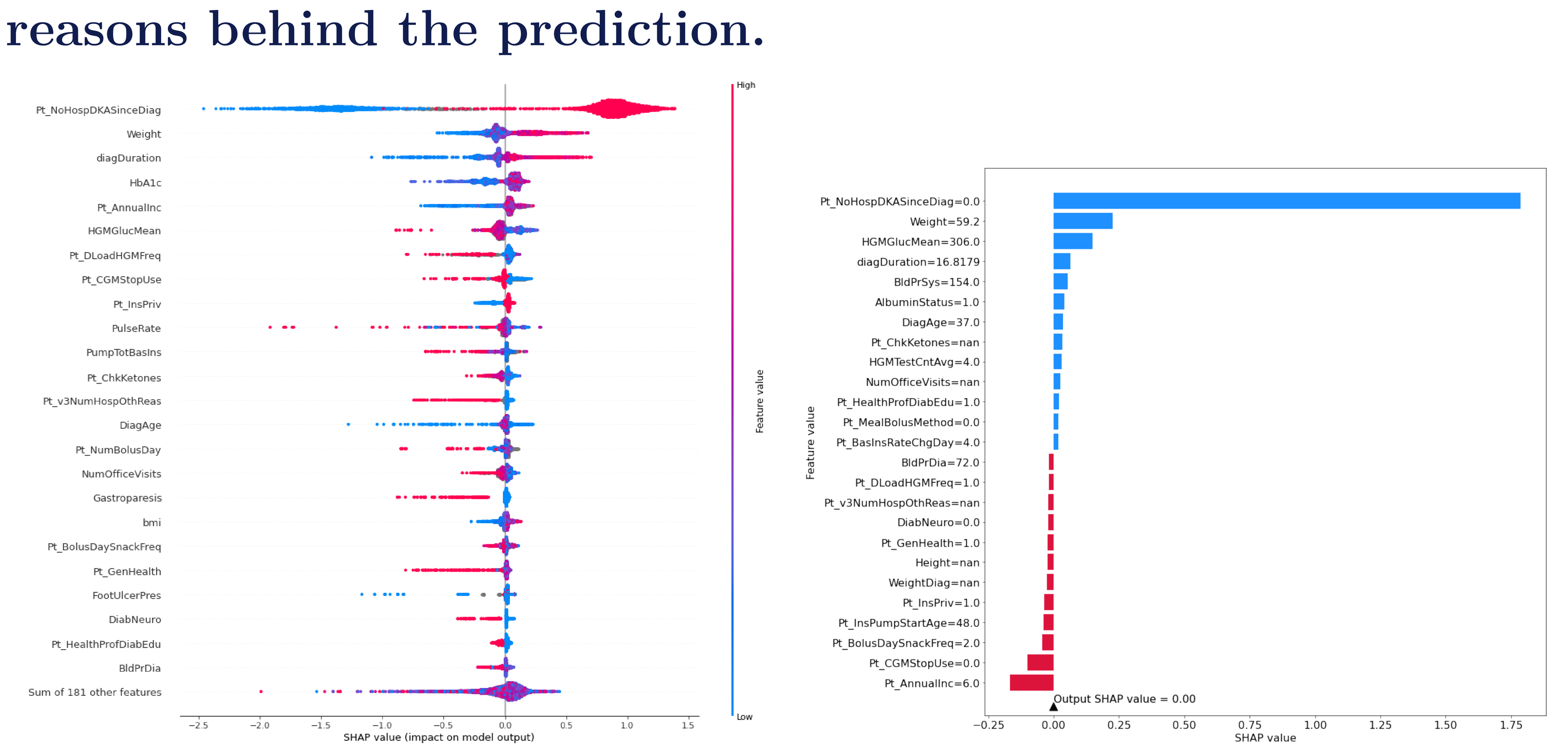
Table 2:Results of the final models

Experimental data set

The experiments of this study are based on the T1D exchange clinic registry open data set that contains 25759 T1D patients data gathered from 67 clinical centers in the United States. This study focuses on 7,155 people in the registry, aged 26 to 93 years and had a T1D duration of at least two years.

DKA Model Interpretation

In this study, we are trying to bridge the interpretability gap in medical machine learning using the SHAP library. Global interpretation is achieved by using the SHAP information dense-summary plot. That gives a view of how features and their values impact on model outcomes across the whole dataset. It is important to have a reason for each individual's outcome separately. Local interpretability is achieved by reviewing the Shapley values of each prediction. It expands the interpretability of the model by predicting possible risks and providing the reasons behind the prediction.



(a) Global interpretation (b) Local Interpretation

Figure 1:Gloabal and Local interpretation of DKA model

Decision Support System

The implemented models were used to build a decision support system for health care personnel. The main objective of this system is to identify patient's risk categories of developing DKA events in the future.

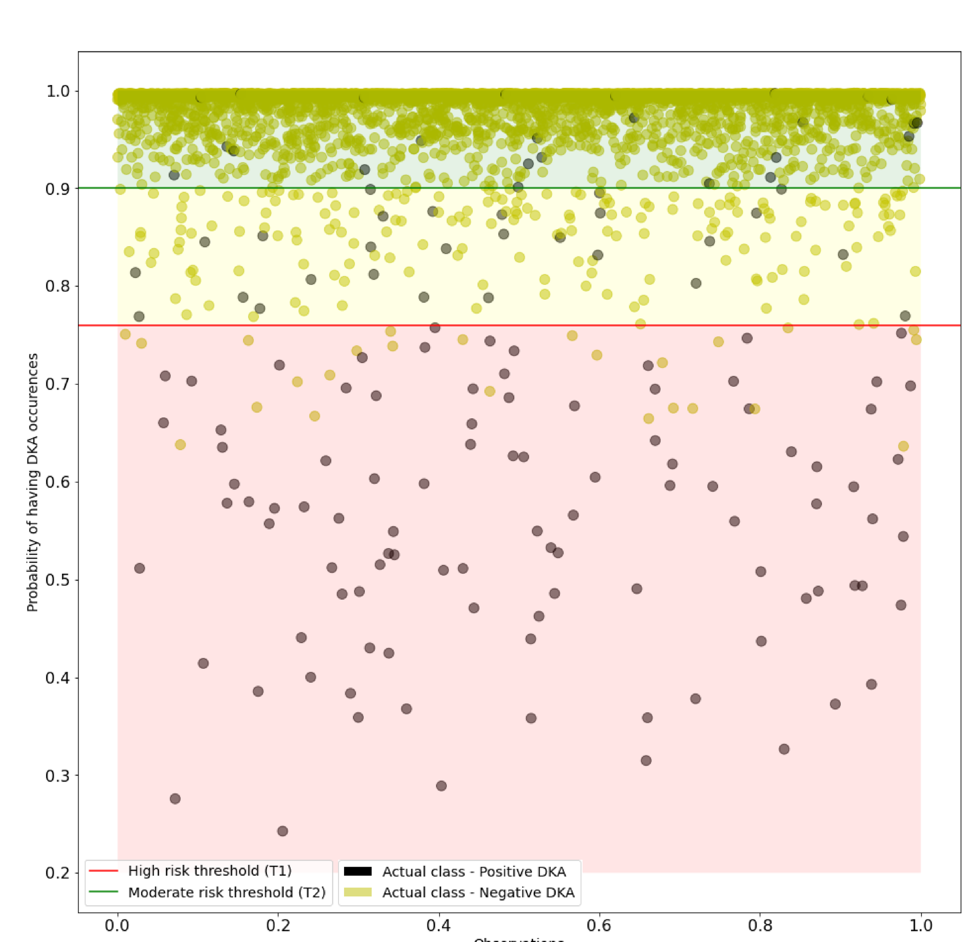


Figure 2:Decision support system of DKA model

Defined thresholds were used to separate patients into risk categories where,

1. Red area - High risk patients
2. Yellow area - Moderate-risk patients
3. Green area - Low risk patients