COMET: Constrained Counterfactual Explanations for Patient Glucose Multivariate Forecasting

Zhendong Wang, Isak Samsten, Ioanna Miliou, and Panagiotis Papapetrou

Zhendong Wang, PhD Student

Stockholm University

Outline

- Background
- Proposed COMET method
- Empirical evaluation
- Conclusions

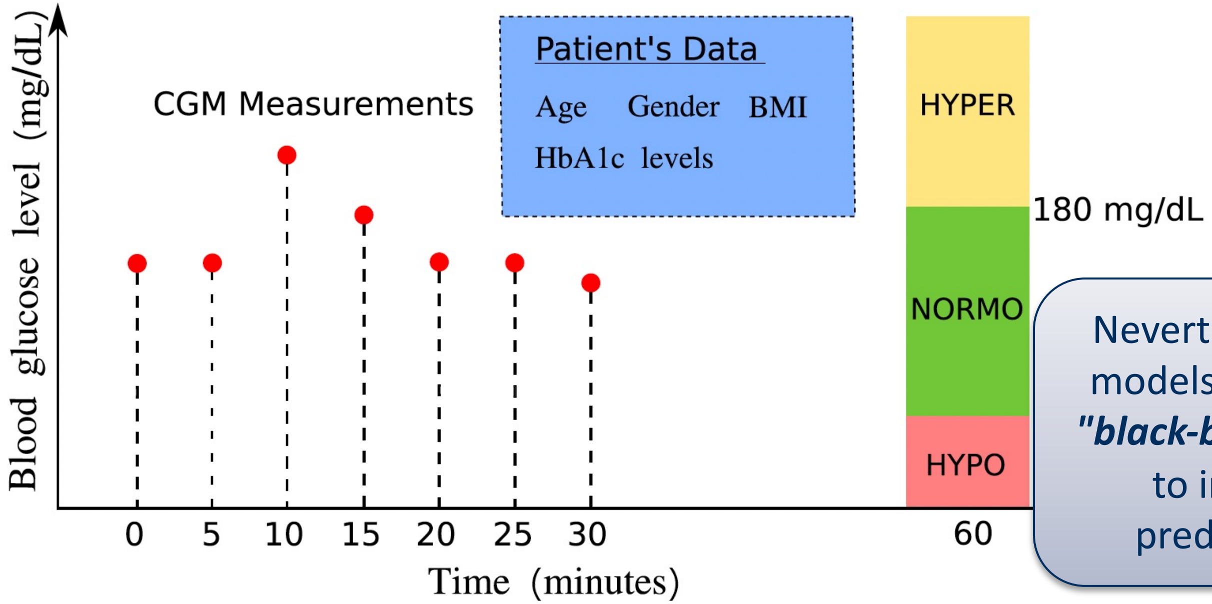
Forecasting in healthcare

- Deep learning (DL) models can assist medical decision--making process in forecasting
  - E.g., Daily patient arrivals and length of stays in the ICU

- For type 1 diabetes mellitus (T1DM) patients, closely tracking glucose levels is a crucial task
  - *RNN-based* models and *transfer learning* have been applied to accurately predict patients' glucose
  - Incorporated additional covariate variables, e.g., insulin dosages and daily carbohydrates

[Deng et al., 2021]
[Cui et al. 2021]

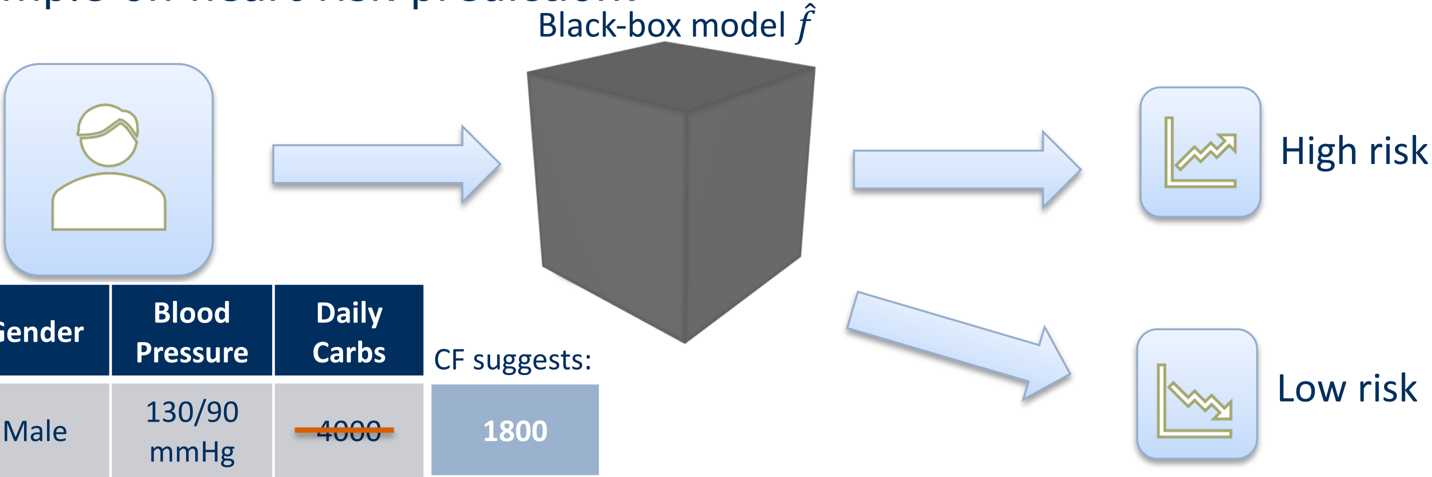
DL forecasting of diabetic blood glucose (30 mins -> 5 mins)



Nevertheless, most DL models are considered ***"black-box"*** and difficult to interpret the prediction results

Counterfactual (CF) explanations

- Show modifications required to change a prediction from an **undesired** (e.g., unhealthy) to a **desired** outcome (e.g., healthy)
- Example on heart risk prediction:



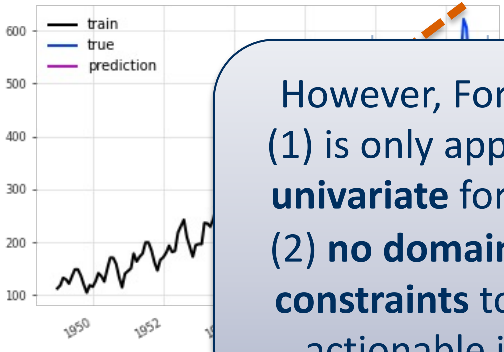
Age	Gender	Blood Pressure	Daily Carbs
43	Male	130/90 mmHg	~~4000~~

CF suggests: **1800**

[Wachter et al., 2017]

Counterfactuals for forecasting

- Recent study proposed to provide counterfactuals for time series forecasting: **ForecastCF**
  - Define desired outcome: *upper* and *lower* bounds that forecasted values are desired to fall into
- Desired bounds for sales forecasting:



Example of Fore

However, ForecastCF:
(1) is only applicable to **univariate** forecasting;
(2) **no domain-specific constraints** to provide actionable insights

 [Wang et al., 2023]

Pseudo-code for proposed COMET

Algorithm 1: COMET counterfactual search

input : Time series $\boldsymbol{X}$, differentiable forecaster $f(\cdot)$, lower and upper bounds $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$, clipping ranges $[\boldsymbol{\rho}, \boldsymbol{\phi}]$, historical value set $\boldsymbol{\mathcal{G}}$, forecast margin weight w, learning rate η, maximum iteration max_iter

output: Counterfactual $\boldsymbol{X}'$ with desired outcome

$\boldsymbol{X}^* \leftarrow \boldsymbol{X}$

$\hat{\boldsymbol{y}}^* \leftarrow f(\boldsymbol{X}^*)$

$\boldsymbol{C} \leftarrow ActivityTemporalConstriant(\boldsymbol{X}^*)$

$loss \leftarrow L(\boldsymbol{X}^*, w, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{X}, \boldsymbol{C})$

$t \leftarrow 0$

while $(\hat{\boldsymbol{y}}^* > \boldsymbol{\beta} \vee \hat{\boldsymbol{y}}^* < \boldsymbol{\alpha}) \wedge (t < max\_iter)$ **do**

- $\boldsymbol{X}^* \leftarrow AdamOptimize(\boldsymbol{X}^*, loss, \eta)$
- $\boldsymbol{X}^* \leftarrow Clip(\boldsymbol{X}^*, \boldsymbol{\rho}, \boldsymbol{\phi})$
- $\hat{\boldsymbol{y}}^* \leftarrow f(\boldsymbol{X}^*)$
- $\boldsymbol{C} \leftarrow HistValueConstraint(\boldsymbol{X}^*, \boldsymbol{\mathcal{G}})$
- $loss \leftarrow L(\boldsymbol{X}^*, w, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{X}, \boldsymbol{C})$
- $t \leftarrow t + 1$

$\boldsymbol{X}' \leftarrow \boldsymbol{X}^*$

return $\boldsymbol{X}'$

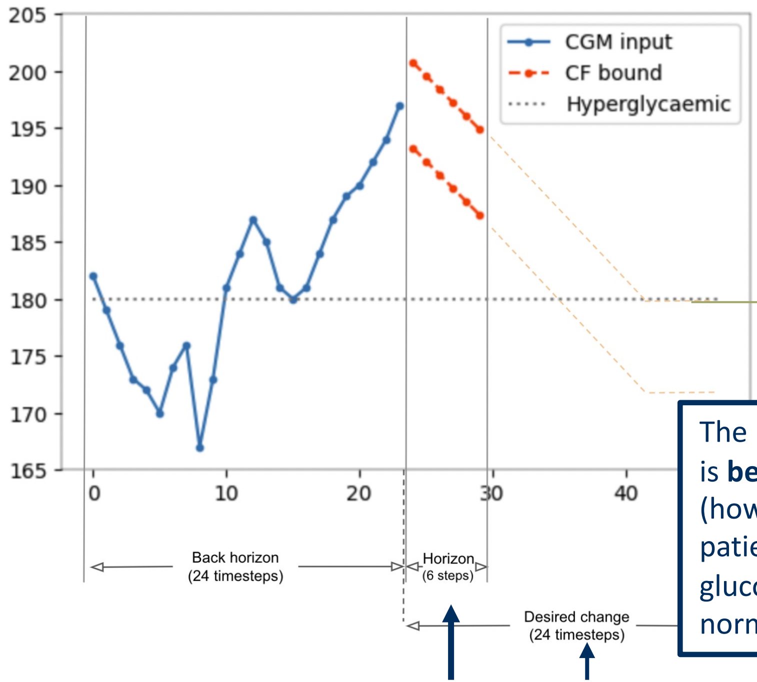
COMET considers three domain-specific constraints for T1DM patients:

1. Activity temporal constraint (-**A**): incorporates constraint **c** to encourage counterfactual changes at specific period with any planned activity (e.g., meal and exercise)

2. Clipping constraint (-**C**): ensures each value at timestep of the CF sample are within the minimum and maximum range

3. Historical values constraint (-**H**): imposes constraints on the CF samples to be closer to the historical values for each patient

Define desired bounds in COMET



The normal glucose range is **between 70 and 180** (however, many T1DM patients often have glucose levels outside the normal range)

…erate constraint …ing a trend to reach the **…ange in 2 hours** (e.g., …teps) and **trim it to the …on** (e.g., 6 steps)

Experimental set-up

- Multivariate forecasting setup:
  - Using 5 clinical features from previous 2 hours to predict CGM measurements in the next 30 mins
  - 2 DL models: GRU and WaveNet
- 2 datasets of diabetes patients
  - *OhioT1DM*: a real-world type 1 diabetes patient dataset
  - *SimGlucose*: simulated from FDA-approved UVA/PADOVA simulator
  - Divide patients in 2 groups: hyper-/hypo-glycemia

- 3 variants of COMET
  - COMET-**C**: included the clipping constraint
  - COMET-**A**: both clipping and activity temporal constraints
  - COMET-**H**: included all three mechanisms (i.e., with historical values constraint)

Empirical evaluation

Stockholm University

Baseline: ForecastCF

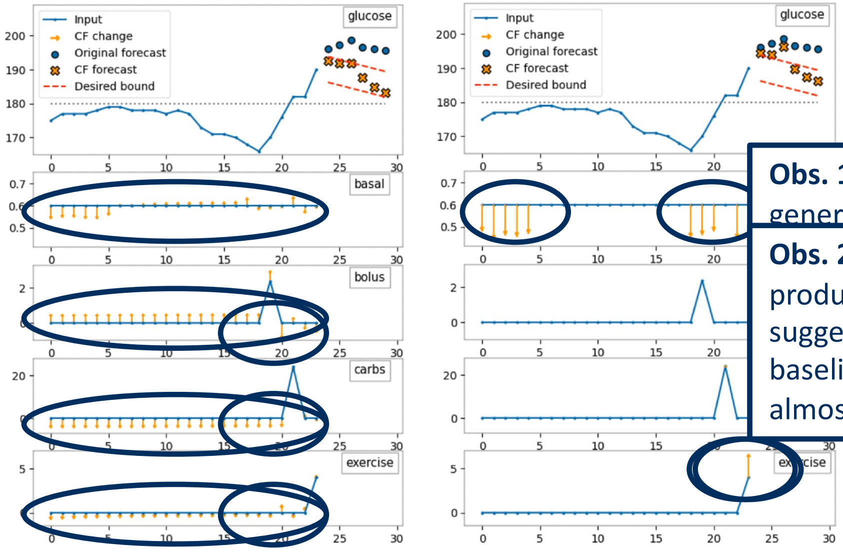
Dataset	Patient group	CF model	GRU *Valid.*	GRU *Proxi.*	GRU *Compa.*	GRU *Step AUC*	WaveNet *Valid.*	WaveNet *Proxi.*	WaveNet *Compa.*	WaveNet *Step AUC*
OhioT1DM	Hyperglycemia	ForecastCF	**0.945**	0.193	0.264	**0.783**	**0.955**	0.159	0.758	**0.790**
		COMET-C	0.867	0.349	0.558	0.706	0.900	0.251	0.844	0.746
		COMET-A	0.675	0.347	0.808	0.526	0.718	0.192	0.935	0.569
		COMET-H	0.585	**0.173**	**0.910**	0.459	0.658	**0.150**	**0.955**	0.521
	Hypoglycemia	ForecastCF	**0.497**	0.284	0.049	**0.188**	0.525	0.162	0.721	**0.196**
		COMET-C	0.490	0.345	0.247	0.175	**0.530**	0.189	0.829	0.171
		COMET-A	0.207	0.396	0.775	0.066	0.238	0.111	0.937	0.074
		COMET-H	0.165	**0.173**	**0.909**	0.046	0.145	**0.042**	**0.973**	0.055
SimGlucose	Hyperglycemia	ForecastCF	**0.984**	**0.056**	0.368	**0.853**	**0.964**	0.117	0.889	**0.823**
		COMET-C	0.858	0.108	0.776	0.758	0.873	0.095	0.935	0.720
		COMET-A	0.732	0.117	0.935	0.624	0.727	0.009	**0.999**	0.457
		COMET-H	0.726	0.116	**0.935**	0.618	0.727	**0.008**	**0.999**	0.457
	Hypoglycemia	ForecastCF	0.115	0.143	0.213	0.041	**0.163**	0.285	0.856	**0.019**
		COMET-C	**0.215**	0.065	0.928	**0.062**	0.094	0.171	0.955	0.009
		COMET-A	0.155	0.031	0.998	0.052	0.103	0.020	**0.999**	0.004
		COMET-H	0.155	**0.031**	**0.998**	0.052	0.103	**0.018**	**0.999**	0.004

Metrics:

- *Validity Ratio* (proportion of valid timesteps)
- *Proximity* (Euclidean distance)
- *Compactness* (unchanged proportion)
- *Stepwise AUC* (consecutive validity)

Qualitative analysis: a T1DM patient

- Forecasting for the next 30 mins (*blue dots*), with upper and lower bounds (*red-dotted lines*)
- Original sample illustrated in *blue*, CF suggestions in *orange arrows*
- Baseline: ForecastCF



(a) Baseline (b) COMET

Obs. 1: Baseline generated out-of-

Obs. 2: COMET produced sparser suggestions while baseline changed almost every step

Conclusions

- Summary of the paper:
  - Proposed COMET generated counterfactual explanations for **multivariate forecasting**, with three **domain-specific constraints** for glucose forecasting
  - Experiments with 2 diabetes datasets outperformed the baseline ForecastCF in proximity and compactness, while maintaining reasonable validity
- Future work:
  - Incorporating clinical experts in assessing counterfactuals' effectiveness and medical relevance
  - Extending into other forecasting applications by involving domain-specific constraints

Github link

SCAN ME

References

1. Wachter, S., Mittelstadt, B., Russell, C.: Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. Technical report, Social Science Research Network (2017)
2. Deng, Y., Lu, L., Aponte, L., Angelidi, A.M., Novak, V., Karniadakis, G.E., Mantzoros, C.S., 2021. Deep transfer learning and data augmentation improve glucose levels prediction in type 2 diabetes patients. npj Digital Medicine, vol 4, pp. 1–13.
3. Cui, R., Hettiarachchi, C., Nolan, C.J., Daskalaki, E., Suominen, H., 2021. Personalised Short-Term Glucose Prediction via Recurrent Self-Attention Network, in: 2021 IEEE International Symposium on Computer-Based Medical Systems (CBMS), pp. 154–159.
4. Wang, Z., Miliou, I., Samsten, I., Papapetrou, P., 2023. Counterfactual Explanations for Time Series Forecasting, in: 2023 IEEE International Conference on Data Mining (ICDM), pp. 1391–1396.

Many thanks
for your attention!