An interpretable deep learning model for time-series electronic health records: Case study of delirium prediction in critical care

Seyedmostafa Sheikhalishahi, Anirban Bhattacharyya, Leo Anthony Celi, Venet Osmani

University of Trento, Italy

Critical Care Services, Mayo Clinic, Jacksonville, FL, USA

Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, USA

Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Information School, University of Sheffield, UK

Journal of Artificial Intelligence In Medicine 2023

BrAIn Lab. Brain and Artificial Intelligence Lab.

SEOULTECH

Author

Mostafa Sheikhalishahi, Ph.D

University of Augsburg
Verified email at med.uni-muenchen.de

Machine learning Explainable AI Bioinformatics Digital health Federated learning

	All	Since 2020
Citations	825	815
h-index	8	8
i10-index	7	7

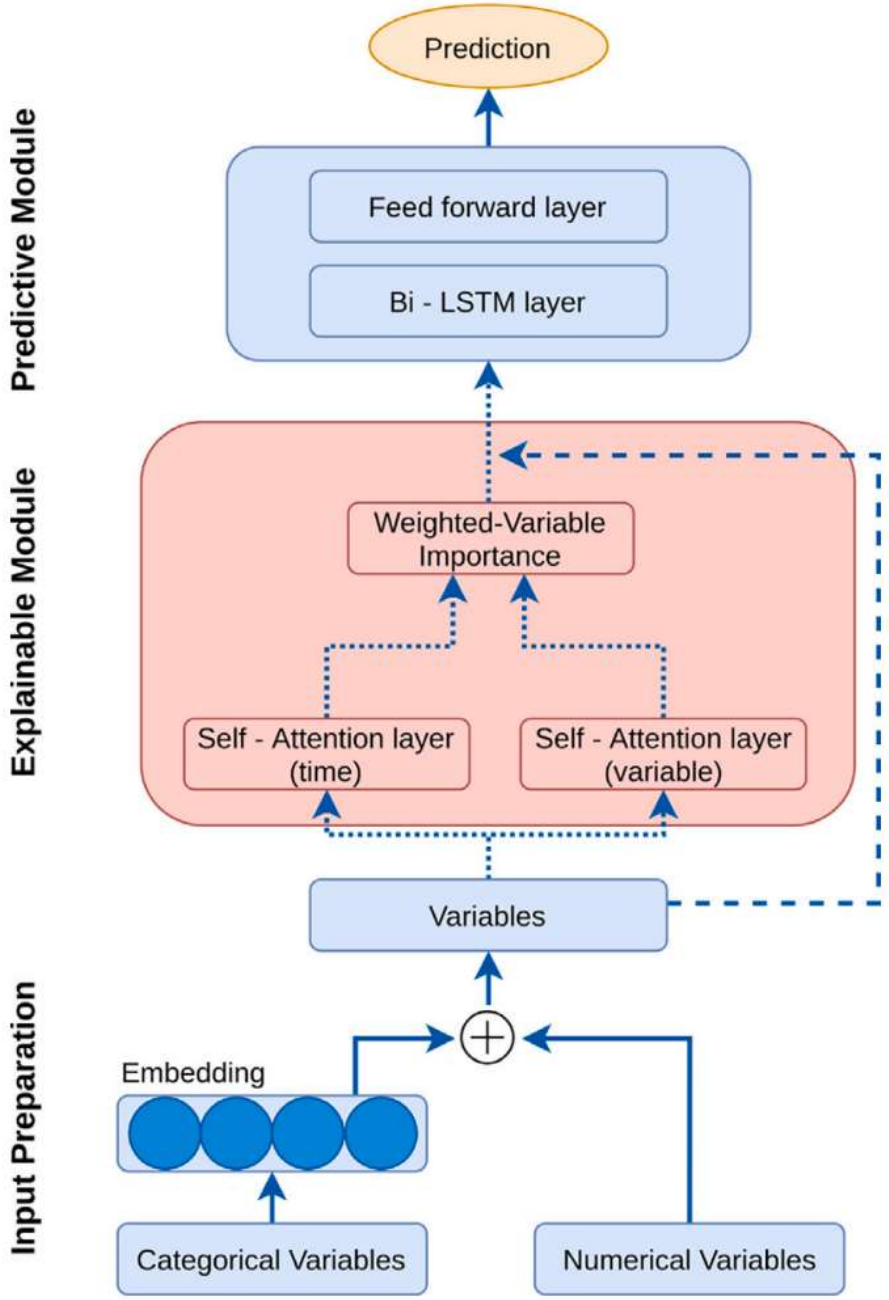
TITLE	CITED BY	YEAR
Natural language processing of clinical notes on chronic diseases: systematic review S Sheikhalishahi, R Miotto, JT Dudley, A Lavelli, F Rinaldi, V Osmani JMIR medical informatics 7 (2), e12239	481	2019
Benchmarking machine learning models on multi-centre eICU critical care dataset S Sheikhalishahi, V Balaraman, V Osmani PloS one 15 (7), e0235424	98 *	2020
Delirium prediction in the ICU: designing a screening tool for preventive interventions A Bhattacharyya, S Sheikhalishahi, H Torbic, W Yeung, T Wang, J Birst, ... JAMIA Open 5 (2), ooac048	15	2022
An interpretable deep learning model for time-series electronic health records: Case study of delirium prediction in critical care S Sheikhalishahi, A Bhattacharyya, LA Celi, V Osmani Artificial Intelligence in Medicine 144, 102659	11	2023
Machine learning applications in Intensive Care Unit S Sheikhalishahi Università degli studi di Trento	1	2022

Introduction

- Conventional machine learning models have been used in the **intensive care unit (ICU)**, which are interpretable but cannot capture non-linear relationship in data.
- It is because the data in ICU is recorded in a **time-series** and conventional ML models do not have an intrinsic ability to deal with time-series inputs.

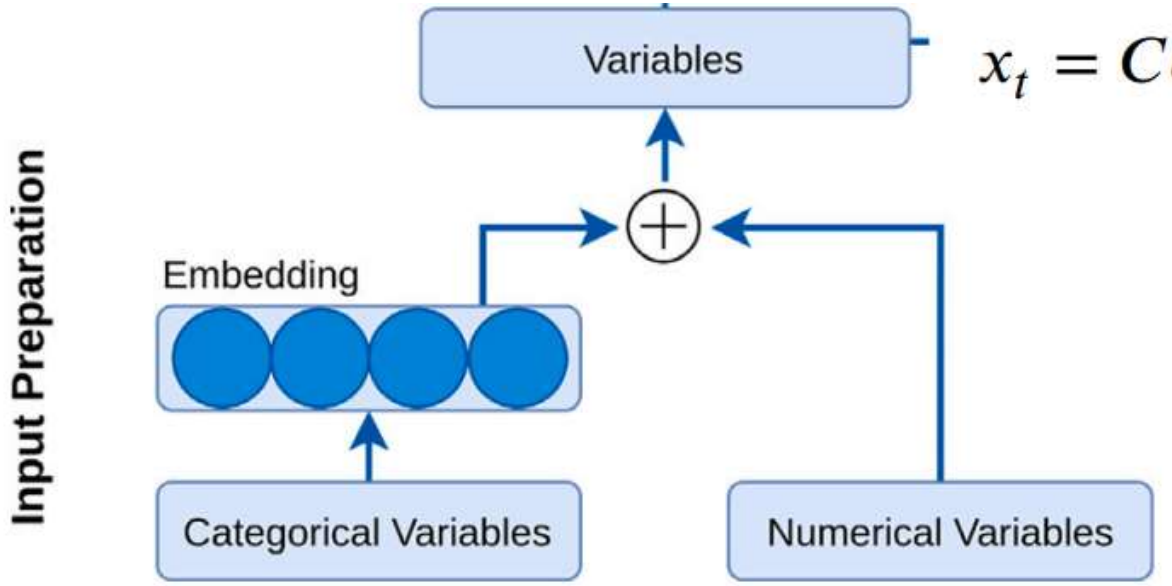
Proposed Method

- Each patient can be viewed as a sequence of medical records (vital signs, laboratory measurements, and demographics) ordered by time, and each record contains a set of clinical variables including **numerical and categorical variables**.
- Categorical variables are represented using **entity embedding**.



BrAIn Lab.
Brain and Artificial Intelligence Lab.

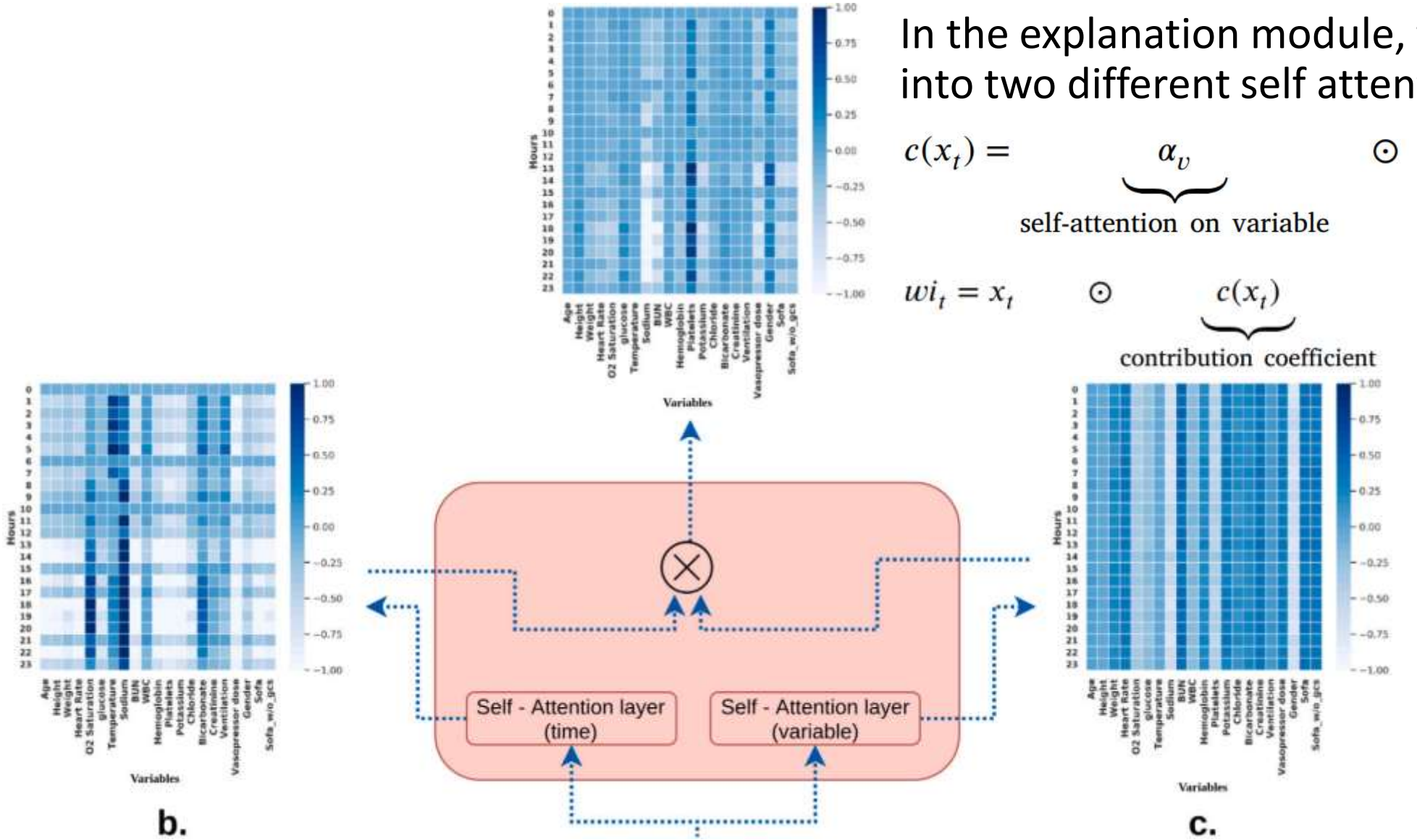
Proposed Method



$$x_t = Concat[Numerical_t, U(Categorical_t)]$$

- $Numerical_t$ stands for the numerical variable, $Categorical_t$ stands for the categorical variable at time t, and U is the embedding matrix.

Proposed Method



In the explanation module, the input x_t is fed into two different self attention layers

$$c(x_t) = \underbrace{\alpha_v}_{\text{self-attention on variable}} \odot \underbrace{\alpha_t}_{\text{self-attention on time}}$$

$$wi_t = x_t \odot \underbrace{c(x_t)}_{\text{contribution coefficient}}$$

Experiment Results

Predictive performance on MIMIC-III dataset.

Model	Observation window 12 h – Prediction window 48 h			
	AUROC% (95% CI)	AUPRC% (95% CI)	Precision% (95% CI)	Recall% (95% CI)
BiLSTM	71.37 (67.99–74.72)	29.81 (27.33–31.88)	28.45 (25.74–31.17)	65.98 (59.82–72.13)
DSA	68.66 (64.99–72.33)	28.58 (23.64–33.20)	26.70 (22.56–30.85)	59.85 (51.93–67.77)
Kaji model	67.56 (64.91–70.22)	27.90 (25.68–30.46)	24.31 (22.01–26.60)	58.07 (52.48–63.66)
Model	Observation window 24 h – Prediction window 12 h			
	AUROC% (95% CI)	AUPRC% (95% CI)	Precision% (95% CI)	Recall% (95% CI)
BiLSTM	81.24 (76.44–86.11)	44.45 (40.35–48.81)	35.03 (30.73–39.33)	71.16 (64.54–77.77)
DSA	80.50 (77.23–83.85)	44.90 (41.28–49.40)	35.58 (33.71–37.44)	68.98 (62.69–75.27)
Kaji model	78.33 (76.11–80.62)	41.69 (38.55–45.15)	32.39 (28.52–36.25)	65.63 (57.78–73.48)

Table 4
Predictive performance on eICU-CRD dataset. Highlighted in bold are results with statistically significant differences.

Model	Observation window 12 h – Prediction window 48 h			
	AUROC% (95% CI)	AUPRC% (95% CI)	Precision% (95% CI)	Recall% (95% CI)
BiLSTM	84.20 (82.52–85.86)	33.24 (29.41–36.54)	**28.37 (27.15–29.58)**	74.44 (70.91–77.98)
DSA	82.51 (80.33–84.67)	31.21 (28.01–33.70)	24.92 (24.22–25.61)	75.99 (72.20–79.78)
Kaji model	81.64 (80.05–83.27)	30.19 (27.41–32.27)	24.60 (23.23–25.97)	75.00 (70.56–79.44)
Model	Observation window 24 h – Prediction window 12 h			
	AUROC% (95% CI)	AUPRC% (95% CI)	Precision% (95% CI)	Recall% (95% CI)
BiLSTM	88.02 (86.31–89.75)	42.69 (38.71–46.1)	**38.19 (36.78–39.6)**	80.39 (76.62–84.16)
DSA	87.10 (85.15–89.03)	42.20 (38.78–45.69)	34.17 (32.88–35.46)	82.89 (78.54–87.25)
Kaji model	85.85 (84.16–87.57)	38.03 (34.64–41.02)	35.82 (34.48–37.15)	76.63 (72.84–80.42)

Conclusion

- This study demonstrated that DSA, can learn importance of variables for predicting delirium where the individual variables underlying these predictions can be explored using attention mechanisms.