Supplementary information

Use of deep learning to develop continuous-risk models for adverse event prediction from electronic health records

Supplementary Table 1: Comparison between selected machine learning models for prediction of inpatient mortality. Where available, the inpatient mortality rate of the dataset is shown. Note MIMIC-III has different rates due to different subsets being used. Triggering time refers to hours after admission when inference in triggered. 95% confidence intervals of performance are shown in brackets where available. GBM, Gradient BoostingMachine; GRU, Gated Recurrent Unit; GRU-D, Gated Recurrent Unit with Delay; FFNN, Feed Forward Neural Network; LSTM, Long Short Term Memory network; RNN, recurrent neural network.

Paper	Dataset (mortality rate)	No. of input features	Primary model	Triggering	Outcome	AUROC [95% CI]	AUPRC [95% CI]
Nakas et al. (2016) [91]	Inselspital Bern (2.4%)	23	Decision trees, FFNN	Static, at admission	In-hospital mortality	0.912	
Johnson et al. (2017) [4]	MIMIC-III	148	GBM	Static, 24h	In-hospital mortality	0.927	
				Random timepoint		0.920	0.665
Aczon et al. (2017) [92]	CHLA paediatric ICU (4.9%)	300	RNN (LSTM)	Static, 12h	In-hospital mortality	0.934	
Che et al. (2018) [93]	MIMIC-III (8.7%)	99	RNN (GRU-D)	Static, 48h	In-hospital mortality	0.853	
Purushotham et al. (2018) [34]	MIMIC-III (10.5%)	20	RNN (FFNN, GRU ensemble)	Static, 24h	In-hospital mortality	0.873	0.477
	MIMIC-III (10.5%)	135		Static, 48h		0.941	0.786
Rajkomar et al. (2018) [20]	Hospital A (2.1%)	Full FHIR embedding	RNN (ensemble LSTM)	Static, 24h	In-hospital mortality	0.95 [0.94-0.96]	
	Hospital B (2.5%)			Static, 24h		0.93 [0.92-0.94]	
Wang et al. (2019) [45]	MIMIC-III	103	RNN (GRU-D)	Static, 24h	In-hospital mortality	0.876	0.532
Caicedo-Torres et al. (2019) [36]	MIMIC-III (9.7%)	22	CNN	Static, 24h	In-hospital mortality	0.822	
				Static, 48h		0.874	
Mayampurath et al. (2019) [94]	U Chicago (2.5%)	156	CNN + recurrent layer	Static, 48h	In-hospital mortality	0.91 [0.90-0.92]	
Harutyunyan et al. (2019) [33]	MIMIC-III (13.2%)	17	RNN (multitask channel-wise LSTM)	Static, 24h	In-hospital mortality	0.870 [0.852-0.887]	0.533 [0.480-0.584]
				Continuous, 1-hourly	Mortality in 24h	0.905 [0.902-0.908]	0.317 [0.307-0.328]
Shickel et al. (2019) [37]	U Florida ICU (10.4%)	14	RNN (GRU)	Static, 24h	In-hospital mortality	0.89 [0.88-0.90]	
				Static, 48h		0.91 [0.90-0.91]	
	MIMIC-III (10.8%)	14		Static, 24h		0.90 [0.89-0.90]	
				Static, 48h		0.91 [0.91-0.92]	
Fritz et al. (2019) [95]	Barnes-Jewish intra-operative (1%)	56	Multi-path CNN	Randomly selected 1h interval	30-day mortality	0.867 [0.835–0.899]	0.095 [0.085-0.109]
Xia et al. (2019) [96]	MIMIC-III (11.7%)	50	RNN (ensemble LSTM)	Continuous, daily	28-day mortality	0.85	0.45
Nielsen et al. (2019) [97]	Danish ICU disease registry (33.4%)	44	FFNN	Static, 24h	In-hospital mortality	0.792	
Brajer et al. (2020) [85]	Duke (3.0%)	195	GBM	Static, at admission	In-hospital mortality	0.87 [0.83-0.89]	0.29 [0.25-0.37]
Hilton et al. (2020) [87]	Cleveland Clinic (1.4%)	171	GBM	Static, 24h	Mortality within 48-72h	0.91	