Appendix A Supplementary Material

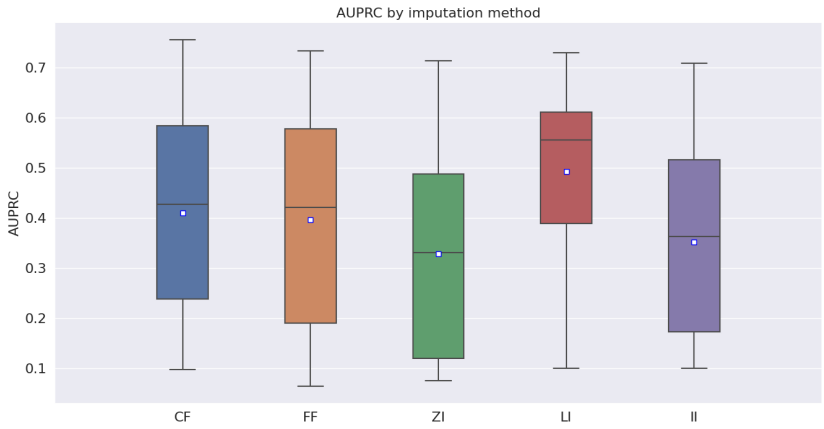

Fig. A1 AUPRC by imputation method. The white square represents the mean value for each method.

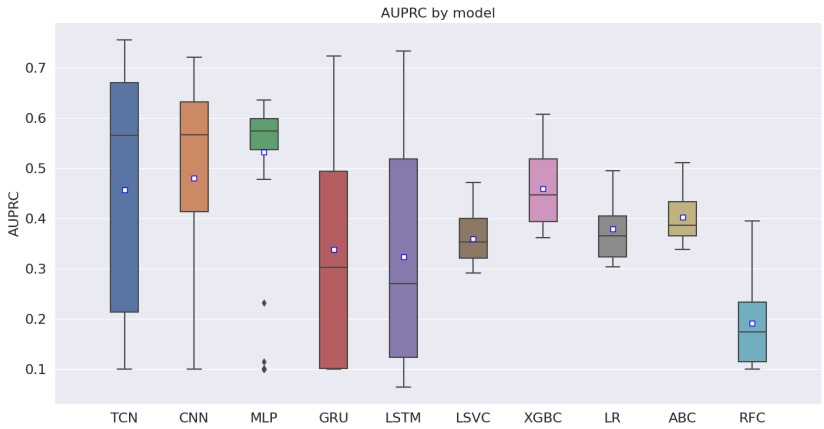

Fig. A2 AUPRC by classification model. The white square represents the mean value for each model.

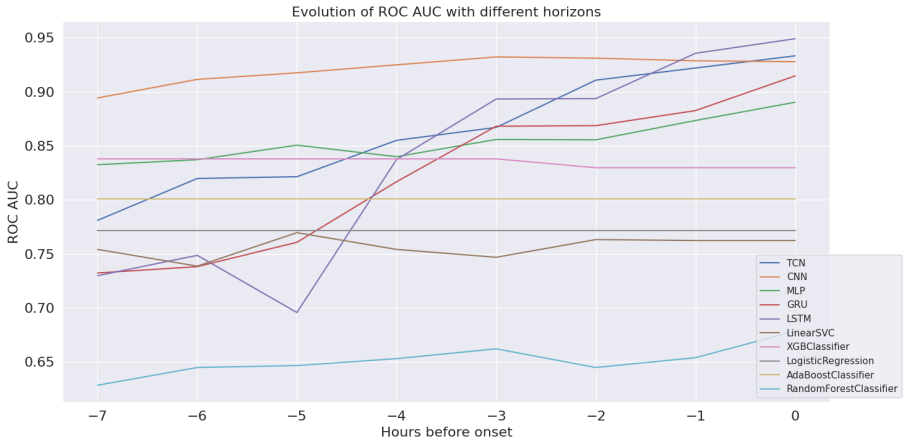

Fig. A3 Predictive performance for the different time horizons. The x-axis indicates the prediction horizon in hours before sepsis onset. The y-axis indicates the AUROC.

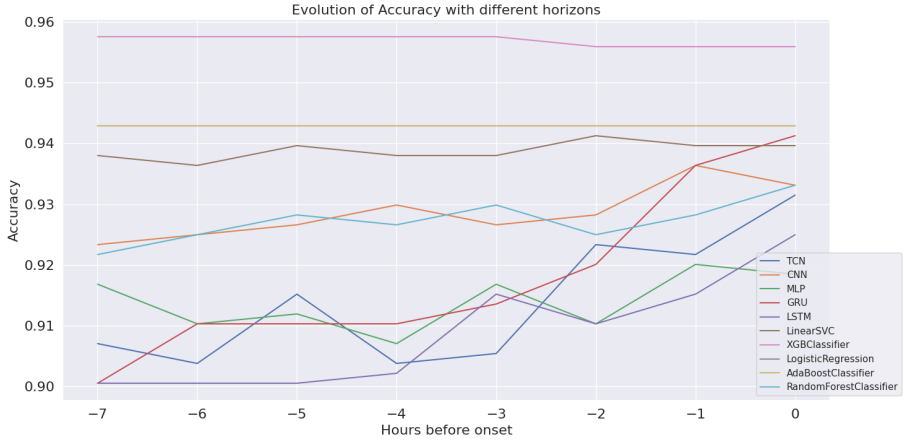

Fig. A4 Predictive performance for the different time horizons. The x-axis indicates the prediction horizon in hours before sepsis onset. The y-axis indicates accuracy.

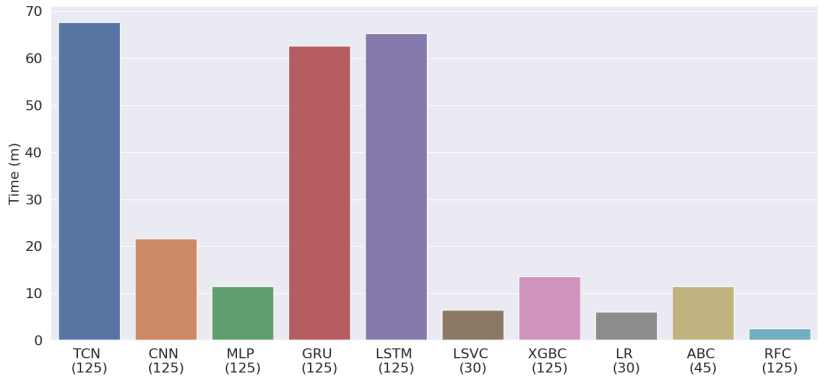

Fig. A5 Total training time, in minutes, for each model, taking into account the parameter optimization process. The total number of parameter combinations tested is also shown for each model.

Type	Measures	
Vital	Systolic Blood Pressure	Heart Rate
	Tidal Volume Set	Total Peep Level
	Diastolic Blood Pressure	SpO2 (Pulse Oximetry)
	Tidal Volume Observed	O2 Flow
	Mean Blood Pressure	Temperature, Celsius
	Tidal Volume Spontaneous	FiO2 (Fraction of Inspired Oxygen)
	Respiratory Rate	Cardiac Output
	Peak Inspiratory Pressure	
Laboratory	AlbuminPotassium	
	Calcium (free)	Blood Urea Nitrogen
	Lactate	Bands (Immature Neutrophils)
	pO2 Blood Gas	White Blood Cells
	Hematocrit	Bicarbonate
	pH Blood Gas	Creatine Kinase
	Hemoglobin	Bilirubin
	pCO2 Blood Gas	Creatine Kinase MB
	Platelet Count	Creatinine
	SO2 Blood Gas	Fibrinogen
	Partial Thromboplastin Time	Chloride
	Glucose	Lactate Dehydrogenase
	Prothrombin Time (Quick)	Sodium
	Troponin T	Magnesium
	INR (Standardized Quick)	

Table A1 Measures selected from MIMIC-III for use. The selection was based on work in [13].

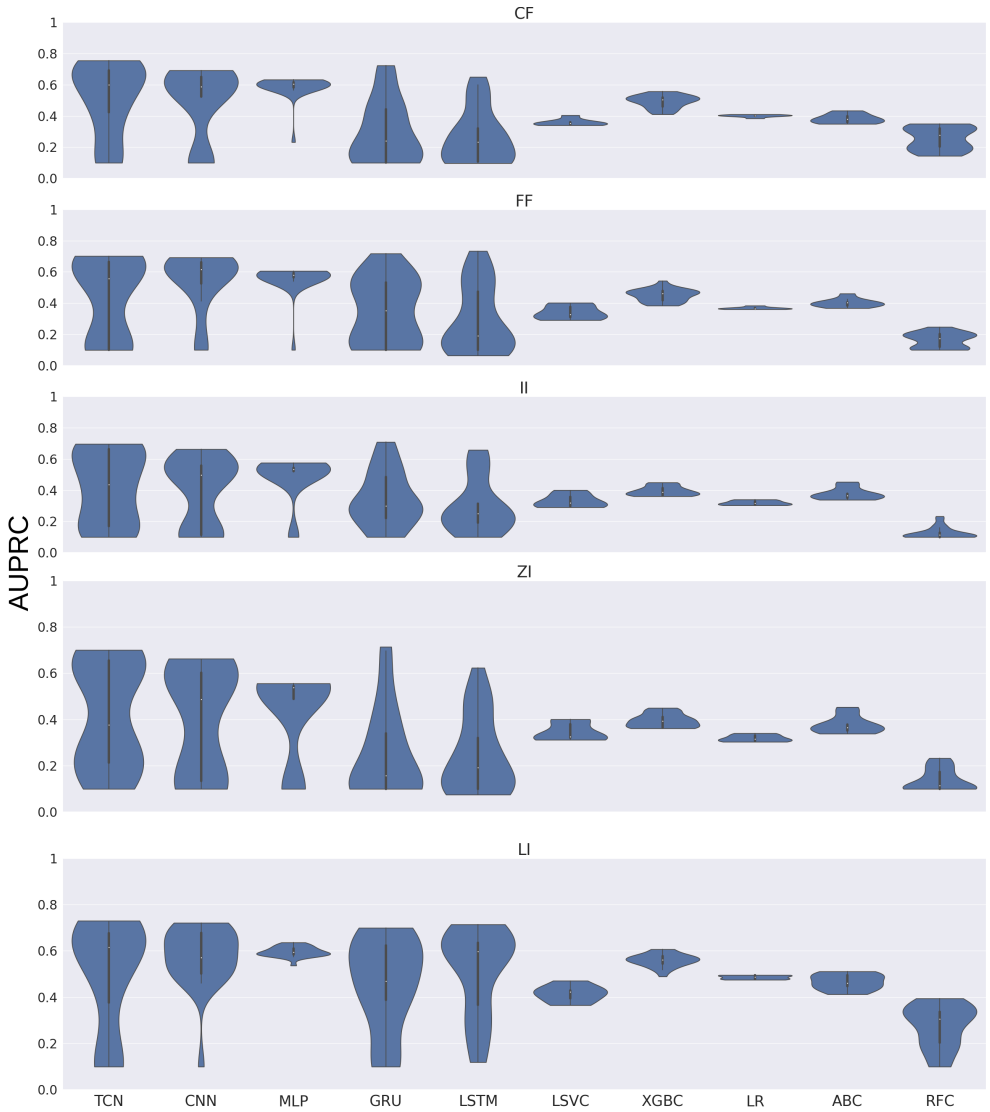

Fig. A6 Distribution of AUPRC measures for each ML model and each imputation method, using density curves.

Imputation method	Model	Count	Mean	Std	Min	25%	50%	75%	Max
	TCN	25	0.5111	0.2325	0.0995	0.4233	0.6007	0.6923	0.7553
	CNN	25	0.4913	0.2302	0.0995	0.524	0.5871	0.6516	0.6923
	MLP	**25**	**0.5872**	**0.0764**	**0.2318**	**0.5817**	**0.6046**	**0.6164**	**0.6329**
	GRU	25	0.2939	0.2022	0.0995	0.0995	0.2391	0.4428	0.7235
Carry forward	LSTM	25	0.2635	0.1736	0.0968	0.1072	0.2335	0.3217	0.65
	LinearSVC	6	0.3593	0.0236	0.34	0.3447	0.3531	0.3608	0.4045
	XGBClassifier	25	0.4923	0.0416	0.4104	0.4604	0.5034	0.5178	0.5573
	LogisticRegression	6	0.4009	0.0089	0.3841	0.4014	0.4019	0.4047	0.4104
	AdaBoostClassifier	9	0.3833	0.0298	0.3487	0.3598	0.3805	0.399	0.4338
	RandomForestClassifier	*25*	*0.2542*	*0.0718*	*0.1438*	*0.203*	*0.2762*	*0.3211*	*0.3497*
	TCN	25	0.4393	0.2548	0.0988	0.0995	0.5581	0.6654	0.7015
	CNN	25	0.5387	0.1902	0.0995	0.5243	0.6155	0.6626	0.6924
	MLP	**25**	**0.5582**	**0.0985**	**0.0995**	**0.5639**	**0.5783**	**0.5945**	**0.6057**
	GRU	25	0.3484	0.2099	0.0995	0.0995	0.3522	0.5321	0.7175
Forward filling	LSTM	25	0.2939	0.214	0.0636	0.1003	0.1901	0.4731	0.7336
	LinearSVC	6	0.3407	0.0452	0.2903	0.3103	0.3274	0.3765	0.4016
	XGBClassifier	25	0.4502	0.0406	0.3834	0.4176	0.4604	0.4777	0.5428
	LogisticRegression	6	0.3676	0.0085	0.3598	0.3627	0.3649	0.3691	0.3834
	AdaBoostClassifier	9	0.4016	0.0277	0.3673	0.3834	0.4019	0.4057	0.4609
	RandomForestClassifier	*25*	*0.1624*	*0.0484*	*0.0995*	*0.1181*	*0.174*	*0.203*	*0.2469*
	TCN	25	0.4193	0.2311	0.0995	0.1696	0.435	0.6641	0.6965
	CNN	25	0.3938	0.2171	0.0995	0.1128	0.4982	0.5593	0.663
	MLP	**25**	**0.4827**	**0.1441**	**0.0995**	**0.521**	**0.5333**	**0.5433**	**0.5757**
	GRU	25	0.3426	0.1785	0.0995	0.2219	0.298	0.4857	0.7089
Indicator imputation	LSTM	25	0.299	0.175	0.0995	0.1925	0.2504	0.3126	0.6581
	LinearSVC	6	0.3312	0.0443	0.2903	0.2971	0.3185	0.357	0.4003
	XGBClassifier	25	0.3958	0.0297	0.361	0.365	0.386	0.4146	0.4493
	LogisticRegression	6	0.3187	0.0151	0.3029	0.3068	0.3153	0.3295	0.3401
	AdaBoostClassifier	9	0.3783	0.0392	0.3381	0.3515	0.365	0.3785	0.4529
	RandomForestClassifier	*25*	*0.1261*	*0.0374*	*0.0995*	*0.0995*	*0.1143*	*0.129*	*0.2324*
	TCN	25	0.4954	0.2375	0.0995	0.3759	0.6155	0.6768	0.7298
	CNN	25	0.5763	0.1316	0.0995	0.5018	0.5705	0.6774	0.7209
	MLP	**25**	**0.5967**	**0.0223**	**0.5363**	**0.5834**	**0.5931**	**0.6096**	**0.636**
	GRU	25	0.456	0.1996	0.0995	0.3872	0.4679	0.6242	0.6992
Linear interpolation	LSTM	25	0.5113	0.1873	0.1187	0.3667	0.5978	0.6355	0.714
	LinearSVC	6	0.4162	0.037	0.365	0.3939	0.4233	0.4289	0.4707
	XGBClassifier	25	0.5602	0.0284	0.4891	0.5448	0.5611	0.5779	0.6067
	LogisticRegression	6	0.486	0.0092	0.4747	0.4786	0.4865	0.4944	0.4953
	AdaBoostClassifier	9	0.4658	0.034	0.4122	0.447	0.4608	0.4953	0.511
	RandomForestClassifier	*25*	*0.2739*	*0.0931*	*0.0995*	*0.2028*	*0.3056*	*0.3375*	*0.3942*
	TCN	25	0.4191	0.2267	0.0995	0.2136	0.3761	0.6544	0.6997
	CNN	25	0.3981	0.2316	0.0995	0.1341	0.487	0.6021	0.6621
	MLP	**25**	**0.4322**	**0.1912**	**0.0995**	**0.4881**	**0.5372**	**0.5441**	**0.5552**
	GRU	25	0.2471	0.1928	0.0995	0.0995	0.1578	0.3397	0.7134
Zero imputation	LSTM	25.0	0.2442	0.1665	0.0747	0.0995	0.1924	0.3196	0.6227
	LinearSVC	6	0.3469	0.0407	0.3116	0.321	0.3257	0.3803	0.4003
	XGBClassifier	25	0.3948	0.0291	0.361	0.365	0.3929	0.4104	0.4493
	LogisticRegression	6	0.3187	0.0151	0.3029	0.3068	0.3153	0.3295	0.3401
	AdaBoostClassifier	9	0.3783	0.0392	0.3381	0.3515	0.365	0.3785	0.4529
	RandomForestClassifier	*25*	*0.1391*	*0.048*	*0.0995*	*0.0995*	*0.1143*	*0.1733*	*0.2324*

Table A2 AUPRC by imputation method. The column count shows the number of experiments for the ML model. The mean, std, and min columns show the average AUPRC, standard deviation, and minimum value, respectively. The next three columns show the quartiles. The max column shows the maximum value. The row for each imputation method with the best mean is in bold, and that with the worst mean is in italics.

Model	Imputation method	Count	Mean	Std	Min	25%	50%	75%	Max
TCN	**carry forward**	**25**	**0.5111**	**0.2325**	**0.0995**	**0.4233**	**0.6007**	**0.6923**	**0.7553**
	forward filling	25	0.4393	0.2548	0.0988	0.0995	0.5581	0.6654	0.7015
	indicator imputation	25	0.4193	0.2311	0.0995	0.1696	0.435	0.6641	0.6965
	linear interpolation	25	0.4954	0.2375	0.0995	0.3759	0.6155	0.6768	0.7298
	zero imputation	*25*	*0.4191*	*0.2267*	*0.0995*	*0.2136*	*0.3761*	*0.6544*	*0.6997*
CNN	carry forward	25	0.4913	0.2302	0.0995	0.524	0.5871	0.6516	0.6923
	forward filling	25	0.5387	0.1902	0.0995	0.5243	0.6155	0.6626	0.6924
	indicator imputation	*25*	*0.3938*	*0.2171*	*0.0995*	*0.1128*	*0.4982*	*0.5593*	*0.663*
	linear interpolation	**25**	**0.5763**	**0.1316**	**0.0995**	**0.5018**	**0.5705**	**0.6774**	**0.7209**
	zero imputation	25	0.3981	0.231	0.0995	0.1341	0.487	0.6021	0.6621
MLP	carry forward	25	0.5872	0.0764	0.2318	0.5817	0.6046	0.6164	0.6329
	forward filling	25	0.5582	0.0985	0.0995	0.5639	0.5783	0.5945	0.6057
	indicator imputation	25	0.4827	0.1441	0.0995	0.521	0.5333	0.5433	0.5757
	linear interpolation	**25**	**0.5967**	**0.0223**	**0.5363**	**0.5834**	**0.5931**	**0.6096**	**0.636**
	zero imputation	*25*	*0.4322*	*0.1912*	*0.0995*	*0.4881*	*0.5372*	*0.5441*	*0.5552*
GRU	carry forward	25	0.2939	0.2022	0.0995	0.0995	0.2391	0.4428	0.7235
	forward filling	25	0.3484	0.2099	0.0995	0.0995	0.3522	0.5321	0.7175
	indicator imputation	25	0.3426	0.1785	0.0995	0.2219	0.298	0.4857	0.7089
	linear interpolation	**25**	**0.456**	**0.1996**	**0.0995**	**0.3872**	**0.4679**	**0.6242**	**0.6992**
	zero imputation	*25*	*0.2471*	*0.1928*	*0.0995*	*0.0995*	*0.1578*	*0.3397*	*0.7134*
LSTM	carry forward	25	0.2635	0.1736	0.0968	0.1072	0.2335	0.3217	0.65
	forward filling	25	0.2939	0.214	0.0636	0.1003	0.1901	0.4731	0.7336
	indicator imputation	25	0.299	0.175	0.0995	0.1925	0.2504	0.3126	0.6581
	linear interpolation	**25**	**0.5113**	**0.1873**	**0.1187**	**0.3667**	**0.5978**	**0.6355**	**0.714**
	zero imputation	*25*	*0.2442*	*0.1665*	*0.0747*	*0.0995*	*0.1924*	*0.3196*	*0.6227*
LinearSVC	carry forward	6	0.3593	0.0236	0.34	0.3447	0.3531	0.3608	0.4045
	forward filling	6	0.3407	0.0452	0.2903	0.3103	0.3274	0.3765	0.4016
	indicator imputation	*6*	*0.3312*	*0.0443*	*0.2903*	*0.2971*	*0.3185*	*0.357*	*0.4003*
	linear interpolation	**6**	**0.4162**	**0.037**	**0.365**	**0.3939**	**0.4233**	**0.4289**	**0.4707**
	zero imputation	6	0.3469	0.0407	0.3116	0.321	0.3257	0.3803	0.4003
XGBClassifier	carry forward	25	0.4923	0.0416	0.4104	0.4604	0.5034	0.5178	0.5573
	forward filling	25	0.4502	0.0406	0.3834	0.4176	0.4604	0.4777	0.5428
	indicator imputation	25	0.3958	0.0297	0.361	0.365	0.386	0.4146	0.4493
	linear interpolation	**25**	**0.5602**	**0.0284**	**0.4891**	**0.5448**	**0.5611**	**0.5779**	**0.6067**
	zero imputation	*25*	*0.3948*	*0.0291*	*0.361*	*0.365*	*0.3929*	*0.4104*	*0.4493*
LogisticRegression	carry forward	6	0.4009	0.0089	0.3841	0.4014	0.4019	0.4047	0.4104
	forward filling	6	0.3676	0.0085	0.3598	0.3627	0.3649	0.3691	0.3834
	indicator imputation	*6*	*0.3187*	*0.0151*	*0.3029*	*0.3068*	*0.3153*	*0.3295*	*0.3401*
	linear interpolation	**6**	**0.486**	**0.0092**	**0.4747**	**0.4786**	**0.4865**	**0.4944**	**0.4953**
	zero imputation	*6*	*0.3187*	*0.0151*	*0.3029*	*0.3068*	*0.3153*	*0.3295*	*0.3401*
AdaBoostClassifier	carry forward	9	0.3833	0.0298	0.3487	0.3598	0.3805	0.399	0.4338
	forward filling	9	0.4016	0.0277	0.3673	0.3834	0.4019	0.4057	0.4609
	indicator imputation	*9*	*0.3783*	*0.0392*	*0.3381*	*0.3515*	*0.365*	*0.3785*	*0.4529*
	linear interpolation	**9**	**0.4658**	**0.034**	**0.4122**	**0.447**	**0.4608**	**0.4953**	**0.511**
	zero imputation	*9*	*0.3783*	*0.0392*	*0.3381*	*0.3515*	*0.365*	*0.3785*	*0.4529*
RandomForestClassifier	carry forward	25	0.2542	0.0718	0.1438	0.203	0.2762	0.3211	0.3497
	forward filling	25	0.1624	0.0484	0.0995	0.1181	0.174	0.203	0.2469
	indicator imputation	*25*	*0.1261*	*0.0374*	*0.0995*	*0.0995*	*0.1143*	*0.129*	*0.2324*
	linear interpolation	**25**	**0.2739**	**0.0931**	**0.0995**	**0.2028**	**0.3056**	**0.3375**	**0.3942**
	zero imputation	25	0.1391	0.048	0.0995	0.0995	0.1143	0.1733	0.2324

Table A3 AUPRC by model. The column count shows the number of experiments for the imputation method. The mean, std, and min columns show the average AUPRC, standard deviation, and minimum value, respectively. The next three columns show the quartiles. The max column shows the maximum value. The row for each imputation method with the best mean is in bold, and that with the worst mean is in italics.

Appendix B Additional Dataset: The PhysioNet Challenge 2019

To validate the models and imputation strategies on another dataset, we used the preprocessed tabular dataset published by PhysioNet in "Early Prediction of Sepsis from Clinical Data: The PhysioNet/Computing in Cardiology Challenge 2019."

B.1 Data Description

A dataset comprising 40,336 patients from two distinct hospitals was made available for this challenge [14]. This dataset contains 40 patient-specific variables, categorized as vital signs, laboratory values, and demographic information. The primary objective of the challenge was to achieve the best possible prediction of sepsis onset within a 6-hour time horizon. This dataset is open access for download without registration. For our experimental analysis, we selected the 48 hours preceding the onset of sepsis for each patient.

B.2 Results and Discussion

This dataset has been extensively used in the challenge and in studies published in other journals, such as [24]. However, results have been reported by using metrics other than the AUPRC. In our experiment, we used the AUPRC metric to facilitate comparisons among future studies using this dataset. Furthermore, according to our main analysis, we performed random splitting, using 80% of the samples for training, 10% for testing, and 10% for validation.

Regarding the results, Table B4 displays the best outcomes achieved for each imputation method, and Figure B7 illustrates the distribution of results obtained during the optimization process. Similarly, Table B5 presents the results of the best values obtained for each model, and Figure B8 depicts the distribution of results obtained by each model during the optimization. These visual representations offer a comprehensive overview of the performance of the proposed methods and the effects of imputation techniques and modeling approaches on the prediction of sepsis onset. Importantly, all these results, in line with the challenge's stated objective, are presented with a temporal prediction horizon of 6 hours. Finally, the best results were obtained with the CNN model, thus demonstrating that deep learning models outperformed other ML models.

Imputation	Model	Time	AUPRC	AUROC	Accuracy
Carry forward	CNN	0.3669	0.7563	*0.9812*	0.9883
Forward filling	GRU	0.7068	0.7745	0.9819	0.989
Indicator imputation	GRU	0.789	*0.749*	0.9886	0.9881
Linear interpolation	CNN	*0.8722*	**0.8215**	**0.9906**	**0.9903**
Zero imputation	CNN	**0.3031**	0.7775	0.9862	*0.9869*

Table B4 Best results by imputation method with the Physionet challenge data. The best results were obtained with the methods for each imputation strategy. For each metric, the best result is in bold, and the worst is in italics.

Model	Imputation	Time	AUPRC	AUROC	Accuracy
TCN	linear interpolation	1.5063	0.8009	0.988	0.9893
CNN	linear interpolation	0.8722	**0.8216**	**0.9906**	0.9903
MLP	linear interpolation	**0.3363**	0.7903	0.9855	0.9891
GRU	linear interpolation	*1.5189*	0.8011	0.9904	0.9888
LSTM	linear interpolation	1.1334	0.7922	0.9904	0.9903
LinearSVC	carry forward	0.3628	0.3667	0.7569	0.9856
XGBClassifier	linear interpolation	0.7142	0.621	0.8808	**0.9916**
LogisticRegression	zero imputation	0.5803	0.374	*0.7511*	0.9861
AdaBoostClassifier	forward filling	2.975	0.5677	0.8858	0.9896
RandomForestClassifier	carry forward	0.3628	*0.3367*	0.7569	*0.9852*

Table B5 Best results obtained by model with Physionet challenge data. For each metric, the best result is in bold, and the worst is in italics.

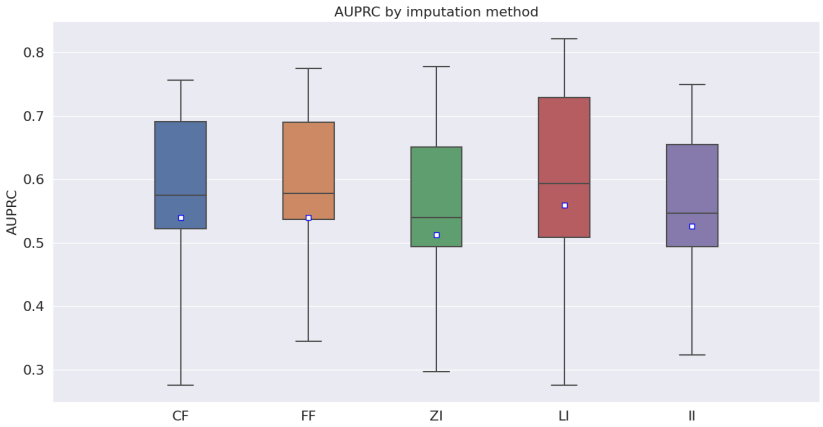

Fig. B7 AUPRC by imputation method with Physionet challenge data. The white square represents the mean value for each of the methods.

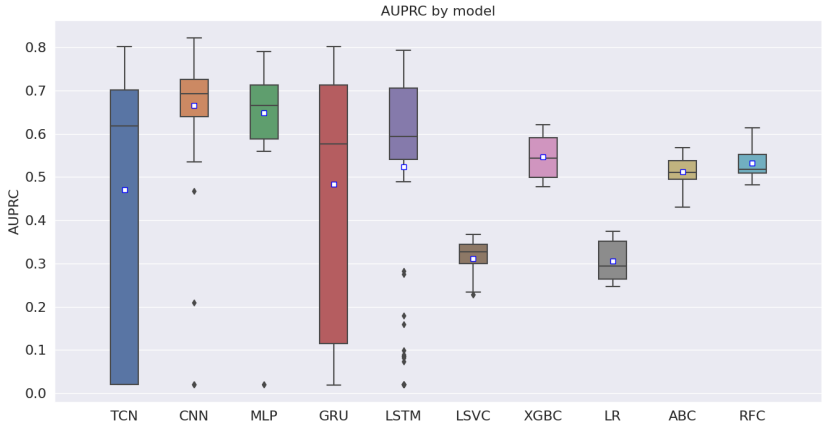

Fig. B8 AUPRC by classification model with Physionet challenge data. The white square represents the mean value for each of the models.