Classification of Irregularly Sampled Clinical Time-Series Data with Convolutional Neural Networks

Michael Wieck-Sosa[1], Diego A. Moreno-Galván[1], Yujuan Fu[1], Shengpu Tang[2], Jenna Wiens[2]

University of Michigan

Introduction

Clinical time series data, such as heart rate, blood pressure and respiration rate, are irregularly sampled because the data is collected by human medical professionals. The hidden structure of these irregularities contains important medical information.

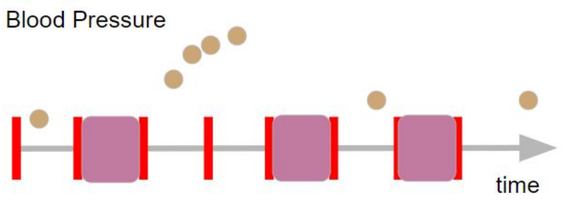


Figure 1. Irregularly Sampled clinical data

Our proposed model is based on the premise that these irregularities are non-random and non-uniform across features, and can be leveraged to make better predictions about a patient's health outcomes. Specifically, in the context of predicting in-hospital mortality within the first 48 hours of an intensive care unit (ICU) admission[1],[2], we propose a novel pipeline to address these irregularities.

Methods

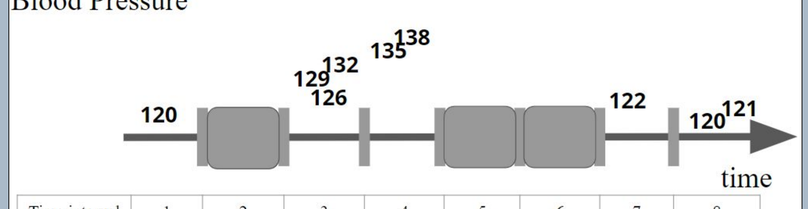
❑ Feature Engineering

We divide the 48 hours of data measurements into time intervals of 1 hour. For each hour timestamp, we encode the structure of the irregularities in the sampling rate with 3 sets of features for each measurement:

- Relative intensity metric
- Missing indicator
- Reliability metric

For example, in the example of Figure 2 below:

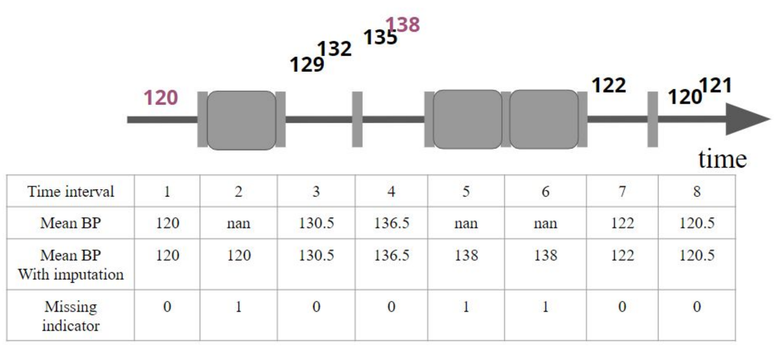
total count=1+3+2+1+2=9
number of time intervals = 8
average count=9 / 8=1.125
⇒ divide every count in each time interval by 1.125.



Time interval	1	2	3	4	5	6	7	8
Count	1	0	3	2	0	0	1	2
Ci/C_avg	0.89	0	2.67	1.78	0	0	0.89	1.78

Figure 2. Relative Intensity Metric

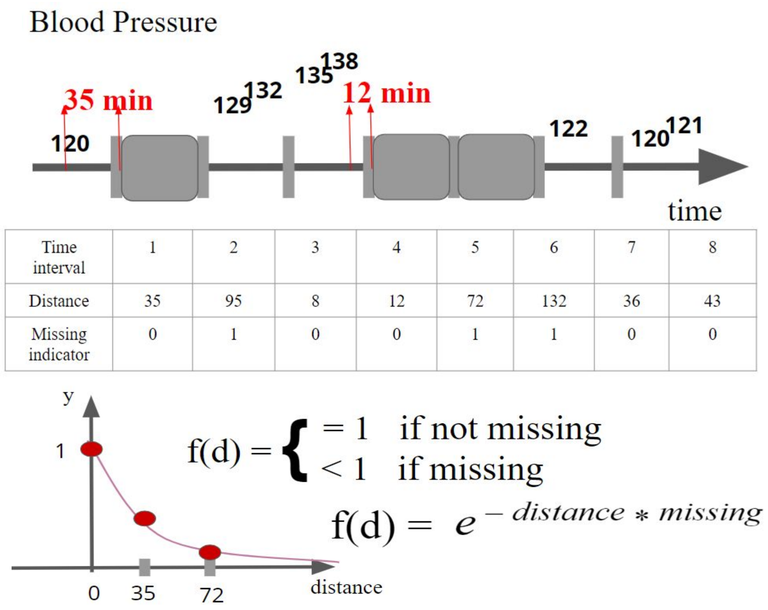
We apply Last Observation Carried Forward (LOCF) to impute missing values, which is common practice in the context of clinical time series data [3], [4].



Time interval	1	2	3	4	5	6	7	8
Mean BP	120	nan	130.5	136.5	nan	nan	122	120.5
Mean BP With imputation	120	120	130.5	136.5	138	138	122	120.5
Missing indicator	0	1	0	0	1	1	0	0

Figure 3.Last Observation Carried Forward

To encode the reliability of LOCF, we record the missing indicator and the time distance to the last measurement in each time interval and employ a missingness indicator. The reliability decreases exponentially with time distance (see Figure 4).



Time interval	1	2	3	4	5	6	7	8
Distance	35	95	8	12	72	132	36	43
Missing indicator	0	1	0	0	1	1	0	0

Figure 4. Reliability Metric

❑ Convolutional Neural Network

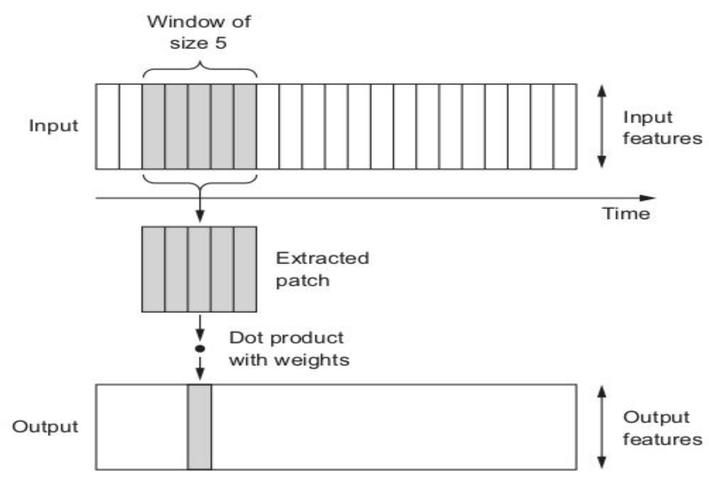


Figure 5. Architecture of Convolutional Neural Network[5]

❑ Proposed Model

Figure 6 below demonstrates the pipeline of our novel convolutional neural network. It incorporates the reliability information to the first convolutional layer.

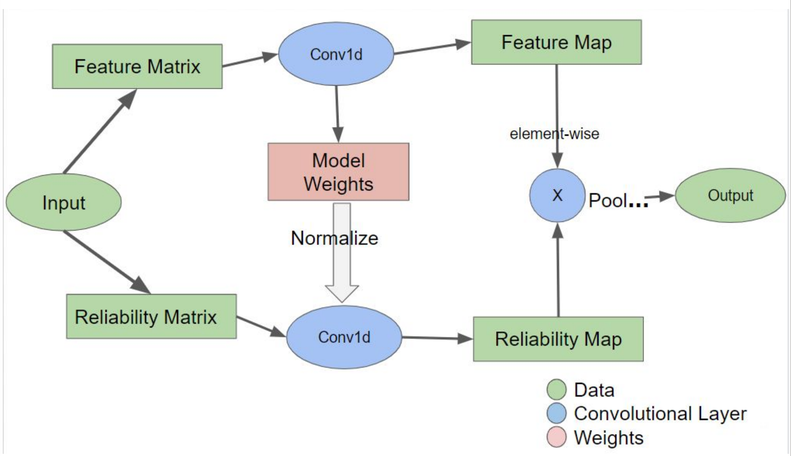


Figure 6. Proposed Model Architecture

For time invariant data, such as height, weight and gender, we concatenate them to the last layer to the neural network. We do the same operation to the flattened missingness indicator and relative intensity of measurements features.

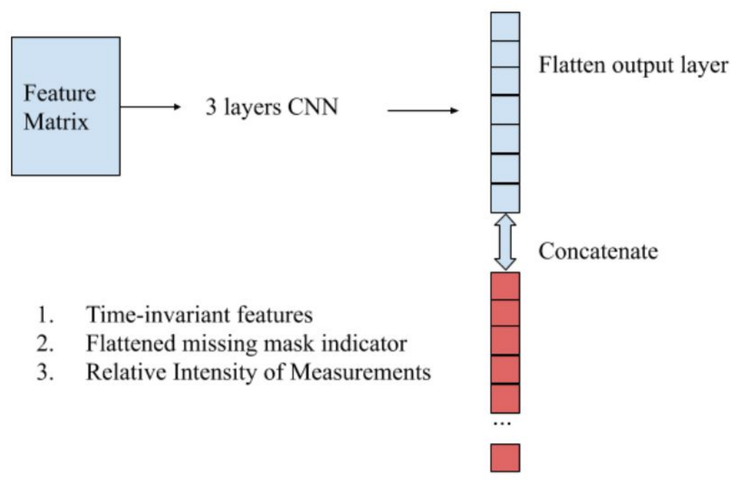


Figure 7. Flattened Layer

Model Evaluation

After hyperparameter selection, we bootstrap the test set 1000 times and get 1000 test score of Area under Receiver Operating Characteristic curve (AUROC).

Model	Median AUROC	AUROC 95% Confidence Interval
Baseline CNN	0.821	[0.804, 0.837]
Proposed CNN	**0.858**	[0.850, 0.866]

Figure 8. Evaluation Performance

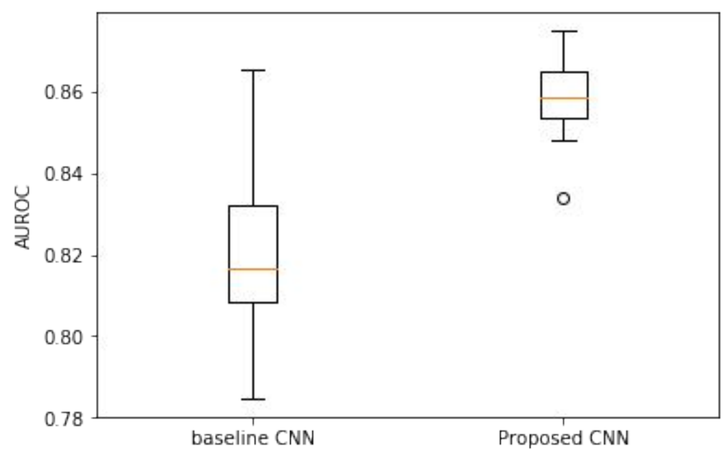


Figure 9. Boxplot of Performance

Figure 8 shows that the proposed model modestly improves prediction performance by 0.037 in AUROC. We run a t-test on the AUROC scores of the proposed and baseline model and got a p-value < 0.005. We conclude that our proposed model has significantly better prediction than the baseline CNN, based on AUROC.

Results

Our proposed pipeline to address sampling irregularities has both clinical and technical relevance:

- ❑ Clinical Relevance: our model can be applied to irregularly sampled clinical time-series data in different contexts
- ❑ Technical Relevance: our model can be generalized to irregularly sampled time-series data in different domains
  - ■ Provided that the irregularities in the sampling rate are non-random and non-uniform across features

Reference

[1] MIMIC database: https://mimic.physionet.org/
[2] PhysioNet CinC Challenge 2012: https://physionet.org/challenge/2012/
[3] Woolley, Stephen B., Alex A. Cardoni, and John W. Goethe. "Last‐observation‐carried‐forward imputation method in clinical efficacy trials: review of 352 antidepressant studies." Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy 29.12 (2009): 1408-1416.
[4] Oh, Jeeheh, Jiaxuan Wang, and Jenna Wiens. "Learning to exploit invariances in clinical time-series data using sequence transformer networks." arXiv preprint arXiv:1808.06725 (2018).
[5] http://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/