Unsupervised Embedding via Locality Preserving Autoencoder for Improved Mortality Prediction in Patients with COVID-19

Anonymous Authors

Abstract

The outbreak of the coronavirus disease 2019 (COVID-19) is putting a huge burden on healthcare workers, and the high fatality rates of COVID-19 have been reported. To alleviate the pressures on the critical care capacity and optimize the allocation of medical resources, the mortality prediction from the patient's records is increasingly becoming a vital factor. This paper proposes the embedding framework to improve the mortality prediction from uneven time series data with missing entries. The proposed embedding framework utilizes the time intervals between patient's records and summarizes them in a format of fixed length vector, by combining deep learning and locality preserving projection. In our experiments, the proposed embedding model shows improved mortality prediction from the blood samples of 485 COVID-19 positive patients in the region of Wuhan, China.

Objective Formulation

Notation

We denotes the records of i-th participant as $\{\mathbf{x}_i, \mathbf{X}_i, \mathbf{m}_i, \mathbf{M}_i\}$, where $\mathbf{x}_i \in \Re^d$, $\mathbf{m}_i \in \{0,1\}^d$ are baseline measurements and masks and $\mathbf{X}_i \in \Re^{d \times n_i}$, $\mathbf{M}_i \in \{0,1\}^{d \times n_i}$ summarizes the measurements and masks of all the available time records. The measurement's time point of i-th participant and j-th record is denoted as t^i_j. The local projection for i-th patient is denoted as $\mathbf{W}_i \in \Re^{d \times r}$.

Objective

The objective to minimize is following:

$$\begin{aligned}
&\mathcal{J}(\mathbf{W}_i, \mathcal{W}_{enc}, \mathcal{W}_{dec}) = \\
&\gamma_1 \sum_{i=1}^{n} \|(\mathrm{dec}(\mathrm{enc}(\mathbf{x}_i, \mathbf{m}_i; \mathcal{W}_{enc}); \mathcal{W}_{dec}) - \mathbf{x}_i) \odot \mathbf{m}_i\|_2^p \\
&+ \gamma_2 \sum_{i=1}^{n} \sum_{\substack{\mathbf{x}^i_j, \mathbf{x}^i_k \in \mathbf{X}_i \\ \mathbf{m}^i_j, \mathbf{m}^i_k \in \mathbf{M}_i}} s^i_{jk} \left\| \mathbf{W}_i^T ((\mathbf{x}^i_j - \mathbf{x}^i_k) \odot \mathbf{m}^i_j \odot \mathbf{m}^i_k) \right\|_2^p \\
&+ \gamma_3 \sum_{i=1}^{n} \left\| \mathrm{enc}(\mathbf{x}_i, \mathbf{m}_i; \mathcal{W}_{enc}) - \mathbf{W}_i^T (\mathbf{x}_i \odot \mathbf{m}_i) \right\|_2^p, \\
&s.t. \quad \mathbf{W}_i^T \mathbf{W}_i = \mathbf{I}
\end{aligned}$$

, where $\mathrm{enc}(\mathbf{x}_i, \mathbf{m}_i; \mathcal{W}_{enc}) \in \Re^r$ is encoded vector from the encoder and $\mathrm{dec} : \Re^r \mapsto \Re^d$ is the decoder. The pairwise similarity coefficient s^i_{jk} is given as the inverse of time interval between j-th record and k-th record, $s^i_{jk} = \frac{1}{\|t^i_j - t^i_k\|}$. The above objective learns a global projection from auto-encoder, and the local projections from the minimization fo Locality Preserving Loss which minimizes the difference between the local projections at the two different time points when the time interval between them is small. We may try below variations:

- Instead of plain auto-encoder, we may use LSTM based auto-encoder as Ball describes in his MS thesis.
- The proposed objective is able to learn the enriched representation of not only baseline but also all the available records. We may learn the enriched representations of all the available records, and then use conventional time series analysis models such as RNN.

Objective 2

The objective without soft constraint between local and global consistency:

$$\begin{aligned}
&\mathcal{J}(\mathcal{W}_{enc}, \mathcal{W}_{dec}) = \\
&\sum_{i=1}^{n} (\gamma_1 \sum_{\substack{\mathbf{x}^i_j \in \mathbf{X}_i, \\ \mathbf{m}^i_j \in \mathbf{M}_i}} \|(\mathrm{dec}(\mathrm{enc}(\mathbf{x}^i_j, \mathbf{m}^i_j; \mathcal{W}_{enc}); \mathcal{W}_{dec}) - \mathbf{x}^i_j) \odot \mathbf{m}^i_j\|_2^p \\
&+ \gamma_2 \sum_{\substack{\mathbf{x}^i_j, \mathbf{x}^i_k \in \mathbf{X}_i, \\ \mathbf{m}^i_j, \mathbf{m}^i_k \in \mathbf{M}_i}} s^i_{jk} \| \mathrm{enc}(\mathbf{x}^i_j, \mathbf{m}^i_j; \mathcal{W}_{enc}) - \mathrm{enc}(\mathbf{x}^i_k, \mathbf{m}^i_k; \mathcal{W}_{enc})\|_2^p)
\end{aligned}$$