

Reiew of the Paper

Student: Yunfei Luo

May 20, 2020

Bidirectional Recurrent Imputation for Time Series
by Wei Cao, Dong Wang, Jian Li, Hao Zhou, Yitan Li, Lei Li
Reference Paper URL: <https://arxiv.org/abs/1805.10572>

Introduction/Main Goal

The paper was published in 2018. The main goal of the paper is to introduce a novel method called Bidirectional Recurrent Imputation for Time Series (BRITS) to handle the missing values in the time-series data, and show that such method is valid and is better than the state-of-art methods.

What's New/Improvement

By treating missing values as the variables in the bidirectional RNNs graph, we could find numerically optimal values for the missing values with their gradients in both forward and backward direction (respect to the direction in the data sequence), so that the imputation could be more accurate. Moreover, while training the model, no assumptions of the data generating process are needed.

Observations

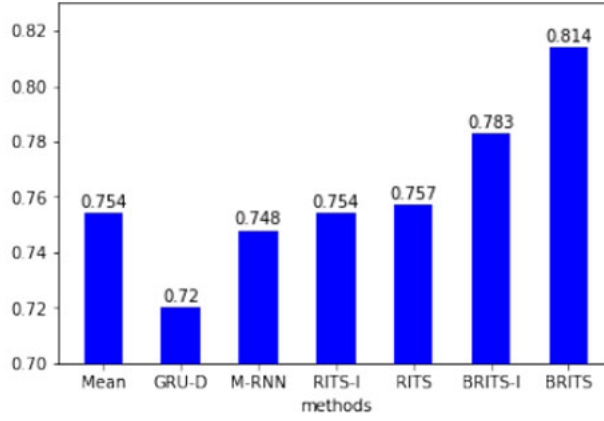
The input is time-series data, and the output is the class value of the target label in the classification/regression tasks. The tasks of missing values imputation and classification/regression are processed simultaneously. That is, when the input value x come into the Recurrent Neural Network (RNN), if it is missing at some point x_t , it will be imputed by both its gradients in the forward and backward directions, then it will be sent for the further computation in the layers in RNN.

Result

The result of heath care classification is shown below (Section 5.3, page 9 in the paper):

Table 2: Performance Comparison for Classification Tasks

Method	Health-care (AUC)	Human Activity (Accuracy)
GRU-D	0.834 ± 0.002	0.940 ± 0.010
M-RNN	0.817 ± 0.003	0.938 ± 0.010
RITS-I	0.821 ± 0.007	0.934 ± 0.008
BRITS-I	0.831 ± 0.003	0.940 ± 0.012
RITS	0.840 ± 0.004	0.968 ± 0.010
BRITS	0.850 ± 0.002	0.969 ± 0.008



The authors also apply the model on different datasets include air quality and localization of human activity. All the case shows that the BRITS method perform better than other state-of-art methods.

Extension/Follow up

The possible follow up would be the scalability of the method, i.e. try these methods on datasets with different size, and see the performance of the BRIST method.