Deep Forecast: Deep Learning-based Spatio-Temporal Forecasting

International Conference on Machine Learning(ICML), Time Series Workshop, 2017

Amir Ghaderi, Borhan M. Sanandaji*, Faezeh Ghaderi

University of Texas Arlington, *Mojio Inc. USA

Introduction

- A spatio-temporal wind speed forecasting algorithm using RNN and LSTM. Deep Learning based Spatio-Temporal Forecasting (DL-STF).
- A graph-based design of spatio-temporal information
  - ➢nodes are data generating entities
  - ➢edges basically model how these nodes are interacting with each other.
- One of the main contributions
  - ➢obtaining forecasts of all nodes of the graph at the same time based on one framework.
- Results on time series data from a collection of wind mills in the north-east of the U.S. show that the proposed DL-based forecasting algorithm significantly improves the short-term forecasts compared to a set of widely-used benchmarks models.

Definitions:

x_i^t : output of the node v_i at time t

s^t : a vector of the output of all nodes, time t

$$s^t = [x_1^t, x_2^t, \dots x_n^t]$$

$\hat{x}_i$: the forecast of output of node i, time t

$\hat{s}$: a vector of all node forecasts at time t

$$\hat{s}^t = [\hat{x}_1^t, \hat{x}_2^t, \dots \hat{x}_n^t]$$

Objective : calculating $\hat{s}^t$using $\{s^k\}, k \in \{t-l, t-l+1, \dots, t-1\}$ when real value for s^k is not available, we use its forecast, $\hat{s}^k$

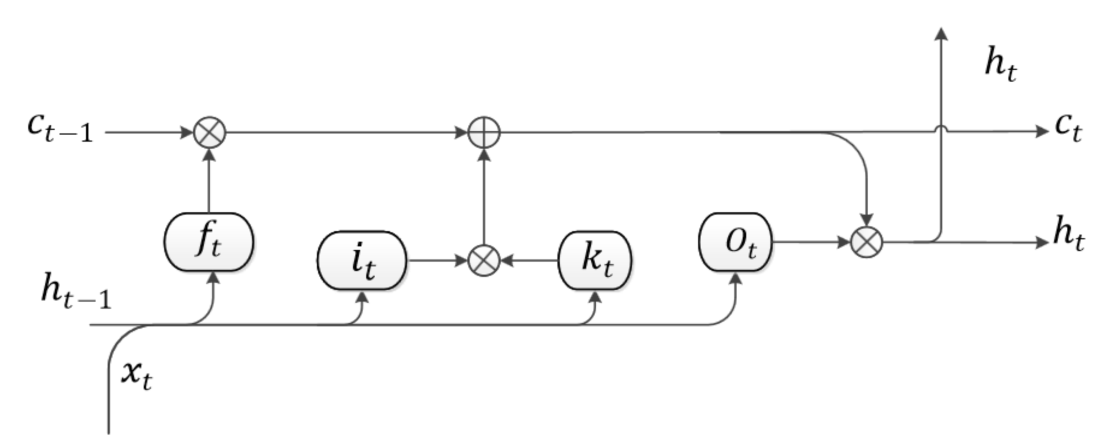
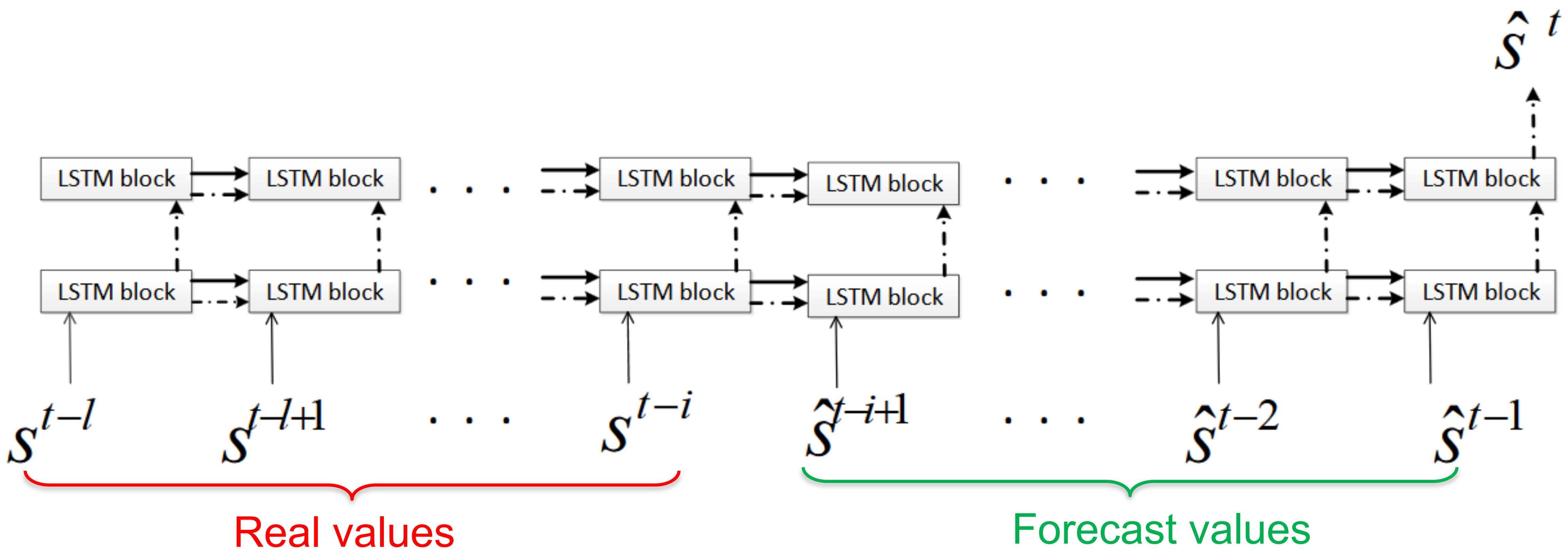


Figure 1. LSTM block at time t



The model M_i trained at time t

Comparison of forecasting error measures

	MAE (m/s)	RMSE (m/s)	NRMSE (%)
one node *	1.99	2.60	15.46
all nodes(ACK) **	1.63	2.19	***13.08***
mean all nodes ***	***1.18***	***1.62***	16.28

Time step models:

First time step : all inputs are real data

Second time step : $l-1$ real data, one forecast

Third time step : $l-2$ real data, two forecast

Different time steps : different kind of inputs so, we need to define a model for each time step

Experiments:

Python , TensorFlow, Keras

Hourly wind speed data from Meteorological Terminal Aviation Routine (METAR) Weather reports

57 stations in east coast U.S.

Test set : time period from January 6, 2014 to February 20, 2014, has the most unsteady wind conditions throughout the year.

The optimizer is MSRPtop which shows good performance for RNN with learning rate 0.001 Activation function is ReLu and data is normalized between [0,1], $h = 6, l = 12$

Results

- Our framework models spatio-temporal data and hidden interactions between nodes
- In spatio-temporal setting we use information of all nodes to forecast one node's output in order to improve the forecasting performance as compared to the case when we only use one node's data (temporal setting)

- * train and test → one node
- ** train → data from all node, test on one node
- *** train and test → all nodes

Results

- Comparison of three common error measures between proposed method and other methods.
- Other methods capable of forecasting ***only one node*** at a time
- Our method can forecast the output of ***all nodes*** at one time
- Our method has smaller error values compared to all other methods
- To the best of our knowledge there is no other method capable of forecasting outputs of all nodes in one framework

Error measures for different methods for one node

Method	MAE (m/s)	RMSE (m/s)	NRMSE (%)
Persistence Forecasting	2.14	2.83	16.86
AR of order 1	2.07	2.76	16.44
AR of order 3	2.07	2.76	16.40
WT-ANN	1.82	2.47	14.68
AN-based ST	1.80	2.30	13.69
LS-based ST	1.72	2.20	13.08
DL-STF	1.63	2.19	***13.08***
DL-STF (All nodes)	**1.18**	**1.62**	16.28

Amir.ghaderi@mavs.uta.edu

Code :https://github.com/amirstar/deep-forecast