

Deep shared representation learning for weather elements forecasting¹

Siamak Mehrkanoon

*Department of Data Science and Knowledge Engineering, Maastricht University,
The Netherlands*

Abstract

This paper introduces a data-driven predictive model based on deep convolutional neural networks (CNN) architecture for wind speed prediction in weather data. The model exploits the spatio-temporal multivariate weather data for learning shared representations and forecasting weather elements for a number of user defined weather stations simultaneously in an end-to-end fashion. The embedded feature learning component of the model as well as coupling the learned features of different input layers have shown to have a significant impact on the prediction task. An experimental setup has been considered based on a high temporal resolution dataset collected from the National Climatic Data Center (NCDC) at five stations located in Denmark. The experiment concerns wind speed prediction at three weather stations located in Denmark for 6 and 12 hours ahead.

1 Introduction

The accuracy and reliability of weather forecasting are of importance for many economic, business and management activities. The use of machine learning techniques to address this data intensive challenge that involves inferences across time and space has recently gained a lot of attentions. In particular, recent years have witnessed the emergence of convolutional neural networks (CNN) as a powerful model for addressing challenging tasks in computer vision. The emerging deep learning techniques together with the availability of massive weather observation data and the advancement of computer technology have motivated researches to explore hidden hierarchical patterns in the weather dataset. Here we employ an upgraded 3d-convolutional neural networks model for learning new feature representations of the given input weather data by exploiting its underlying spatio-temporal multi-modal characteristic. The proposed model uses the historical weather elements from multiple weather stations simultaneously and learns new predictive shared representations.

2 Formulation of the method

The weather datasets naturally follow spatio-temporal structure as each variable (weather element) is recorded in a specific time and location. Let us assume that the number of weather stations is q , and the total number of weather elements (variables) is p . Furthermore, let $y_j^{s_i}(t)$ denotes the measurement corresponding to the j -th weather element of the i -th station at time t . If for instance we set the j -th weather element of the first station at time t as target variable, and also the lag parameter of both input and target signals to d , then one can construct the following regressor vector at time t : $z(t) = [y_1^{s_1}(t-1), \dots, y_p^{s_1}(t-1), \dots, y_1^{s_1}(t-d), \dots, y_p^{s_1}(t-d), \dots, y_1^{s_q}(t-1), \dots, y_p^{s_q}(t-1), \dots, y_1^{s_q}(t-d), \dots, y_p^{s_q}(t-d)]$, which would be a vector of length $p \times q \times d$. Thus the problem is reduced to finding a right mapping from the input vector $z(t)$ to the desired target variable $y_j^{s_1}(t)$ as follows: $y_j^{s_1}(t) = f(z(t))$. In order to exploit the spatio-temporal structure of the input data, we first cast each regressor vector into a tensor with (stations, lags, variables) as (height, width, channel). Here we present a model that learns a bank of

¹Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). The full paper has been published in *Knowledge-Based Systems*, Volume 179, pp. 120-128, September 2019.

three dimensional kernels that are applied on the tensorial data containing the measurements of all the weather stations. It is expected that given enough data, potentially the three dimensional CNN model learns new shared representations of the data as well as existing correlations among channel space. The architecture of the proposed 3d-CNN model in [1], for weather forecasting is depicted in Figure 1. Here, the input data is fed to $(2 \times 2 \times 2)$ -convolution layers with 10 filters followed by ReLU nonlinear activation function. The obtained feature maps are then flattened and the network is followed by fully connected layers with ReLU and linear activation functions respectively.

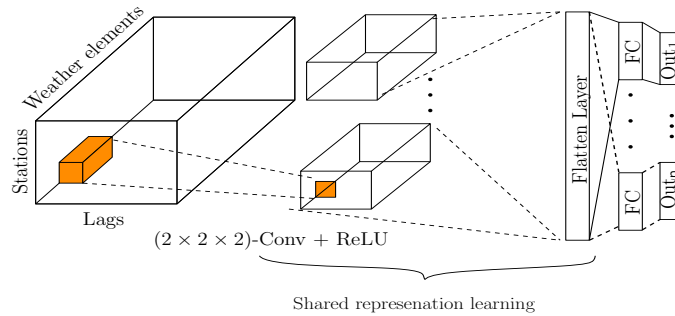


Figure 1: The 3d-CNN architecture proposed in [1] for weather elements forecasting.

3 Experimental results

Wind speed is often considered as one of the most difficult parameters to forecast because its underlying dynamics operates in an intermittent fashion therefore modeling its fluctuation is challenging. Our experiment concerns 6 and 12 hours ahead wind speed prediction for three weather stations located in Denmark. Here the hourly historical data which include four weather elements including temperature, pressure, wind speed and wind direction from 2000-2010 are used. The performance of the proposed 1d-, 2d- and 3d-convolutional neural networks models for wind speed prediction is compared with those of NARX and LSTM networks. The test set consists of the last 10% of the data, while the remaining 90% percent of the data is used for training the models. For this dataset, the sequence length and the number of hidden units in the LSTM cell are set to four days (96 hours) of measurements and 200 respectively. The obtained results are shown in Fig. 2 and tabulated in Table 1.

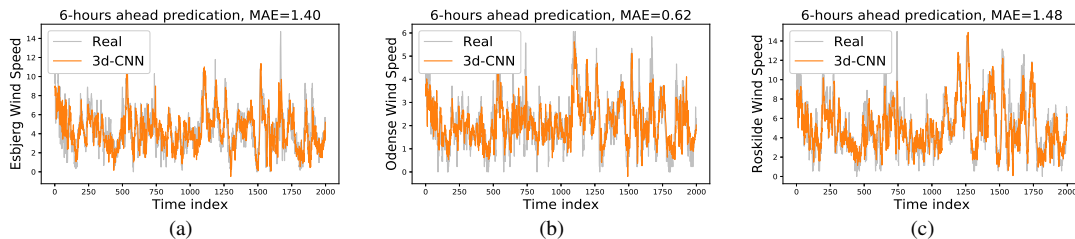


Figure 2: The Obtained 6-hours ahead wind speed forecasts for three stations.

Table 1: The MAEs (mean absolute errors) of the proposed models, the NARX and LSTM models.

Hours ahead	Station	Method				
		3d-CNN [1]	2d-CNN [1]	1d-CNN [1]	NARX	LSTM
6	Esbjerg	<u>1.40</u>	1.42	1.44	1.59	1.54
	Odense	<u>0.62</u>	0.63	0.63	0.68	0.86
	Roskilde	<u>1.48</u>	1.50	1.52	1.56	1.49
12	Esbjerg	<u>1.71</u>	1.75	1.75	1.81	1.77
	Odense	<u>0.79</u>	0.80	0.82	0.86	1.05
	Roskilde	1.84	1.90	1.92	1.96	<u>1.79</u>

Acknowledgments. This work was partially supported by the Postdoctoral Fellowship of the Research Foundation-Flanders (FWO: 12Z1318N). Siamak Mehrkanoon is an assistant professor at the Department of Data Science and Knowledge Engineering, Maastricht University, The Netherlands.

References

- [1] Siamak Mehrkanoon. Deep shared representation learning for weather elements forecasting. *Knowledge-Based Systems*, 179:120–128, 2019.