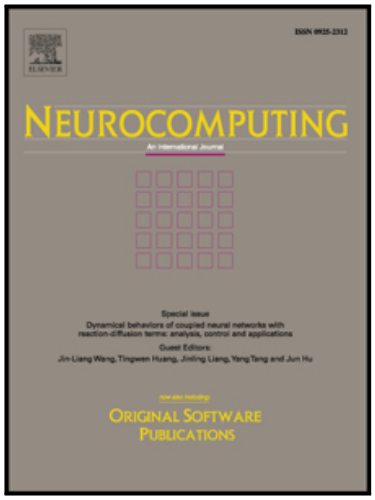
Communicated by Dr. Deng Cheng

Accepted Manuscript



A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction

Junshan Wang, Guojie Song

PII: S0925-2312(18)30785-9
DOI: 10.1016/j.neucom.2018.06.049
Reference: NEUCOM 19727

To appear in: *Neurocomputing*

Received date: 9 June 2017
Revised date: 16 April 2018
Accepted date: 6 June 2018

Please cite this article as: Junshan Wang, Guojie Song, A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction, *Neurocomputing* (2018), doi: 10.1016/j.neucom.2018.06.049

A Deep Spatial-Temporal Ensemble Model for Air Quality Prediction

Junshan Wang[a], Guojie Song[a,*]

[a]*Key Laboratory of Machine Perception, Ministry of Education, Peking University, Beijing, 100871, China.*

Keywords: Air Quality Prediction, Deep Learning, Spatial-Temporal Data

Abstract

Air quality has drawn much attention in the recent years because it seriously affects people's health. Nowadays, monitoring stations in a city can provide real-time air quality, but people also strongly desire air quality prediction, which is a challenging problem as it depends on several complicated factors, such as weather patterns and spatial-temporal dependencies of air quality. In this paper, we design a data-driven approach that utilizes historical air quality and meteorological data to predict air quality in the future. We propose a deep spatial-temporal ensemble(STE) model which is comprised of three components. The first component is an ensemble method with a weather-pattern-based partitioning strategy. It trains multiple individual models and combines them dynamically. The second one is to discover spatial correlation by analyzing Granger causalities among stations and generating spatial data as relative stations and relative areas. The last one is a temporal predictor based on deep LSTM to learn both long-term and short-term dependencies of air quality. We evaluate our model with data from 35 monitoring stations in Beijing, China. The experiments show that each component of our model makes contribution to the improvement in prediction accuracy and the model is superior to baselines.

*gjsong@pku.edu.cn