Rapid prediction of atmospheric pollutants in local region through optimized RBF methods

Jianyang Li, Gang Li, Ming Wang
School of information, Zhenjiang Institute of Technology
School of Management, Hefei University of Technology
E-schooling center, *Zhenjiang Institute of Technology*

Introduction

The total emission control of atmospheric pollutants is to take the local region as a complex system engineering. The aspect of unfixed monitoring points are rapidly deployed for acquiring huge range data of atmospheric pollutants, such as weather balloons or UAV can truly realize the real monitoring of the regional large-area atmospheric environment and solve the problem of insufficient distribution of fixed monitoring points; It can collect environmental parameters of different spatial locations in real time and transmit them through the communication network; Realize the three-dimensional distribution of atmospheric environment, and realize the three-dimensional monitoring of atmospheric environment data through AI algorithms. Through a large amount of data collection, it can guide the production activities of enterprises and the formulation of government policies.

Restricted factors

The emission and accumulation of local air pollutants are subject to the conditions of atmospheric diffusion with the pollutant production level; as well known the meteorological factors are the key that affects atmospheric diffusion including wind direction, wind speed, air temperature, atmospheric temperature, inversion layer and many other factors.

Atmospheric environmental capacity (AEC) refers to the quality of air pollutants that can be carried in a certain area within a period of time, representing the self-purification ability of the atmospheric environment.

Optimized methods

The linear correlation between input and output is analyzed according to Pearson correlation analysis, and the Pearson correlation coefficient r_n (between the input and output variable of the nth input neuron) is calculated through Pearson correlation formula:

$$r_n = \frac{\sum_{k=1}^{K}\left(x_{k,n} - \bar{x}_n\right)\left(y_k - \bar{y}\right)}{\sqrt{\sum_{k=1}^{K}\left(x_{k,n} - \bar{x}_n\right)^2}\sqrt{\sum_{k=1}^{K}\left(y_k - \bar{y}\right)^2}}$$

According the former description, the topology of RBF neural network should be adjusted to construct a cross layer direct connection structure, which is illustrated in the fig 1 -the direct connection between $x_{k,n}$ (the nth input neuron) and y_k (the output neuron) are shown; and set the weight $v_n = r_n$, so the output y_k :

$$y_k = \sum_{n=1}^{N} v_n * x_{k,n} + \sum_{h=1}^{H} w_h * \phi_h\left(x_k\right)$$

Results and outlook

The data of " Abalone" and " Boston house prices "in this experiment are downloaded from UCI repository and their main information are as follows: Abalone- 4177 instances, 8 features; Boston house- 506 instances, 13 features. Set learning rate μ =0.1, maximum number of iteration steps s_{max} =500, 70% of the samples are used for network training, and the rest are used for network performance testing. Experimental parameters are set as shown in table 1.

TABLE I. DATASETS AND EXPERIMENTAL PARAMETERS

	Instan	featu	R	H_{max}	$RMSE_d$
Abal	4177	8	0.6	10	7.5×10-2
Bost	506	13	0.65	10	5×10-2

TABLE II. COMPARISON OF MODEL PERFORMANCE

data	method	training R/ e-2	testing RMSE/ e-2	training cost/s
Abal	RBF	18.360	11.83	255.8
	ERBF	7.416	7.625	4.797
	OURS	7.469	7.777	4.653
Bost	RBF	7.741	13.21	22.76
	ERBF	4.956	8.191	2.815
	OURS	4.704	8.393	2.740

As table 3 shows, there are three RBF-NNs chosen for comparison of our new methods in the experiments.

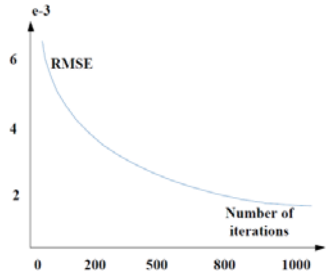

Fig 2. Relationship between RMSE and Iterations

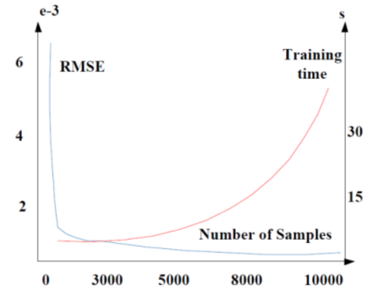

Fig 3. Relationship among RMSE, samples, and time cost

Acknowledgements

Supported by the Natural Industrial Project of Zhenjiang GY2022027, so many people have helped us with this paper that it is hard to pick out the few for special mention, and I would thank for their help with providing many ideas.