CENTRAL ASIAN JOURNAL OF THEORETICAL AND APPLIED SCIENCES

Volume: 04 Issue: 08 | Aug 2023 ISSN: 2660-5317
https://cajotas.centralasianstudies.org

Design and Analysis of Air Pollution Concentration Prediction Models Using Transfer Learning and Recurrent Neural Networks

Mankala Satish
Assistant Professor
Dept of Computer Science Engineering (AI&ML)
CMR Institute of Technology, Kandlakoya, Hyderabad, India

Peddi Niranjan Reddy
Assistant Professor
Dept of Computer Science Engineering (AI&ML)
CMR Institute of Technology, Kandlakoya, Hyderabad, India
.

Received 24th Jun 2023, Accepted 13th Jul 2023, Online 31nd Aug 2023

Abstract: *Since air pollution (AP) poses a serious risk to human health, many people have started paying greater attention to it in recent years. Precise Air Pollution prediction helps individuals schedule their outdoor activities and contributes to human health protection. In this study, recurrent neural networks (RNNs) with long short-term memory (LSTM) were used to predict Macau's future APS concentration. Data on the concentration of APS as well as environmental data have also been used. Additionally, some air quality monitoring stations (AQMSs) in Macau have fewer overall observed data while simultaneously collecting less observed data for specific APS kinds. In order to help AQMSs with less observed data, transfer learning, and pre-trained neural networks have been utilized. The purpose of this study is to show how a collection of neural network algorithms has been utilized for these two pollutant elements. The approach is given considerable thought in this paper, and datasets regarding air and water pollution as well as expected parameters were additionally collected for future development efficiency.*

Keywords: *Transfer learning, Recurrent Neural Network*

I. Introduction

With the increasing development of global industrialization, air pollution has become more difficult. Air pollution could be an important factor in infectious diseases and a decrease in life expectancy, based to studies [1]. Countries across the world have implemented an array of methods to address this global problem, and people have a propensity to change their routines in order to deal with the decreasing quality of the air. Predicting air quality would be an essential and successful means for helping people in their struggle towards such destruction. Long-term exposure to air pollution carried by vehicular traffic can decrease life expectancy. In addition, people who are exposed to vehicle-related AP over an extended period may live shorter lives. In numerous northern Chinese cities, Chen et al. [2] the association between APS and lung cancer patients as well as the connection between NO2, SO2, and PM10 concentrations and lung cancer mortality. The statistical data that they collected shows a positive correlation among the prevalence and death of lung cancer and the level of pollutants in the air in a particular region. The air will travel vertically upward if the atmosphere is unstable, on the contrary together, which will help in the APS's potential to spread to the sky. The amount of APS can be viewed as time series data since LSTM RNNs are effective at predicting time series data. Therefore, LSTM RNN has been used in this paper to be able predict the initially indicated concentration of APS despite an absence of knowledge about the atmospheric dispersion modeling of APS. Additionally, transfer learning has been proposed to be employed in this study to help forecast the AP level in order to achieve decent prediction results considering the lack of observed data.

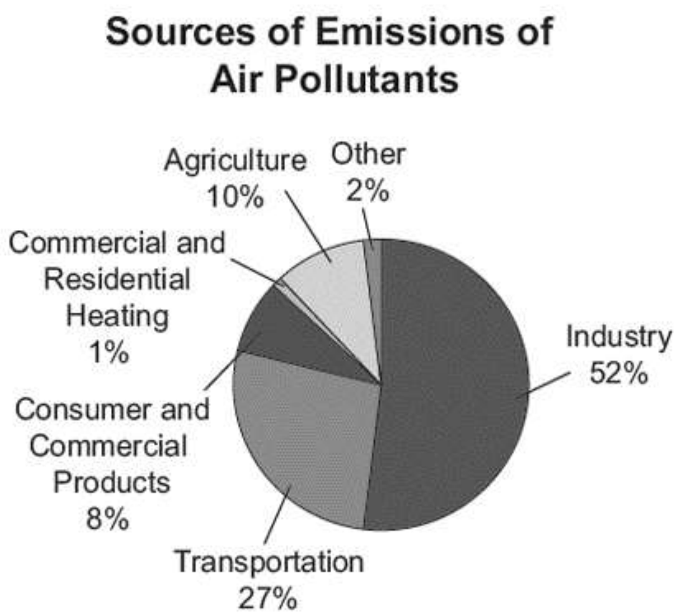

Figure 1: India's sources of pollutants in the air pollutants

Air pollution predictions were achieved through bidirectional LSTM techniques. Recent research on air quality prediction has concentrated on enhancing precision while maintaining the prediction window within 12 hours; some studies have even restricted the window to 1 hour in order to achieve the highest accuracy. Given that combating air pollution takes more time, short-term prediction has minimal practical value despite its exceptional accuracy.

II. Related Work

In contrast with standard methods of machine learning, transfer learning approaches use knowledge accumulated from data in additional domains to facilitate predictive modeling consisting of multiple data patterns in the current domain. Transfer learning approaches seek to establish a framework for making use of previously-acquired knowledge to solve new but similar problems much more quickly and effectively. In order to continue through the timeline in transfer learning, Sinno Jian Pan and Qiang Yang published a survey on the subject in 2009 and thoroughly examined inductive, transductive, and unsupervised transfer learning. By applying a co-training framework, Yu Zheng et al. [3] proposed a semi-supervised learning approach which consists of two separate classifiers: one is a spatial classifier constructed from an artificial neural network, and the other is a temporal classifier developed on a linear-chain Conditional Random Field (CRF). Employed auto-encoder as a trained technique to improve the performance of deep recurrent neural network with the goal to forecast PM2.5 in Japan. The Levenberg-Marquardt approach was used to train artificial neural networks by Asha B. Chelani et al.[4]. to determine the SO2 concentration in three cities in Delhi.

1. Theory-Based Methods

Two theory-based air quality models which have gained popularity in the industry are the community multiscale air quality (CMAQ) modeling system and the comprehensive air quality model with extensions (CAMx). Both were developed using the concept of "one atmosphere," which includes conversion and interactions among multiple kinds of air contaminants, and replicated scales of multipollutants. The majority of airborne allergens continue to be the focus of CMAQ, which also analyzes the general level of air quality over a number of locations. Though being generally applicable, it has an array of intrinsic problems, including errors caused by manually defined parameters and mass conservation being adversely affected by varying meteorological fields[5]. For instance, the particular substance for comprehensive air quality model with extensions (PMCAMx) predicted significantly higher levels of O3 and PM than CMAQ. Considering the fact that theory-based models will perform well given accurate data, their effectiveness is still limited by the models' extensive estimation and inherent little errors. In addition, it makes it tricky for them to gather a lot of data effectively.

III. Methodology

1. Long short-term memory

Long short-term memory is a type of recurrent neural network architecture used in deep learning. RNN solely employs feed forward neural networks to operate, whereas LSTM uses feedback connections. The LSTM system can process complete data sequences as well as to single data points. Due to its capacity to believe time series, this algorithm is frequently used in cases of contamination of the water and air. We increase the constant error carrousel CEC provided by the self-connected, linear unit j by including novel features in order to build an architecture which enables constant error through special, self-connected units without the disadvantages of the naive approach[6]. In order to prevent the memory content stored in j from being disrupted by unrelated inputs, a multiplicative input gate unit is implemented. The addition of a multiplicative output gate unit similarly shields other units from disturbance by stored, currently irrelevant memory contents.

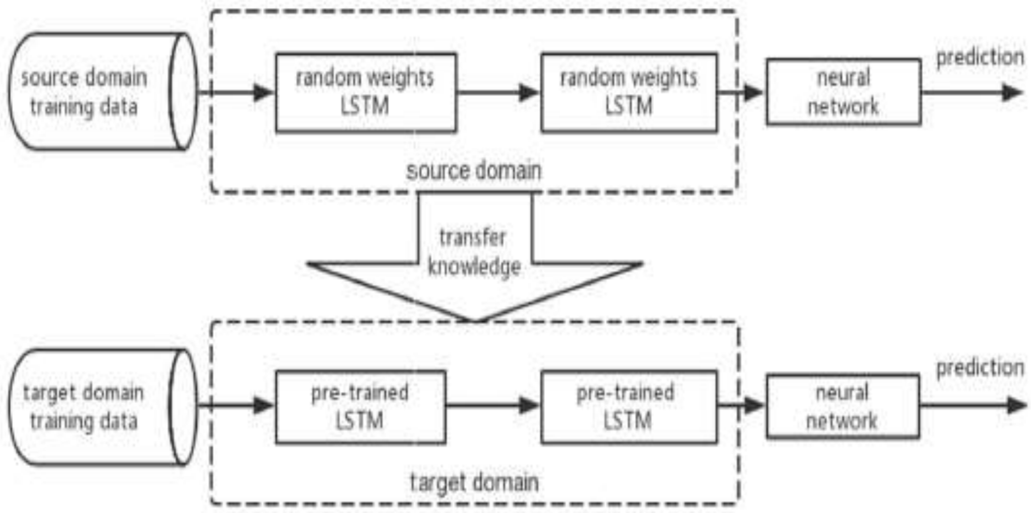

Figure 2: Transfer learning technique.

The signal realizes the storage of data while remaining in a specific time period as it flows through the input gate, output gate, and forget gate in turn. The neural network's LSTM structure shows that the input variable immediately travels horizontally from input to output. In a result, the error in prediction will begin to rise before increasing suddenly in direct proportion with the previous occurrence in the prediction model[7]. The LSTM prediction accumulation error in the forecasting for the regular demand for electricity. The data set used for training for the short-term load forecasting approach usually lasts a day or an entire week.

2. Bi-LSTM

A bidirectional LSTM is recommended here to deal with the accumulative error problem, which is shown in Figure 3. The two layers of LSTM structure which make up the bidirectional LSTM neural network are used to calculate the hidden vector from the front to the back and from the back to the front, accordingly. These two layers regulate the output of the bidirectional LSTM neural network. The standard feed-forward mechanism neural network is unique from the bidirectional LSTM neural network. In a bidirectional LSTM, the internal nodes in each layer have no connection to one another. With the goal to improve the association of single pieces of information in multiple-time series, a directional loop is included in the connection of hidden layers, foregoing data; outcomes are learned and stored in the memory unit. Combining the previous output to the current input produces the neural network's current output. However, due to a lack of delay window width, there currently are going to be gradient disappearance and gradient explosion issues when the time series' input data grows in volume[8].

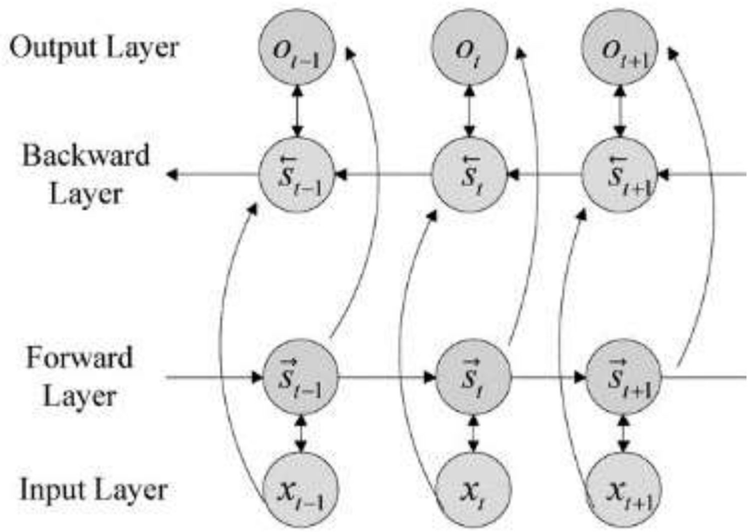

Figure 3: Basic bidirectional LSTM structure

The bidirectional LSTM neural network, which is based on the traditional LSTM model, can take into consideration of the front and back correlation of the load data in time series and enhance the model's performance, for the sequence classification problem.

The forward layer's input data sequence is used as training data during the training phase with one another, and the backward layer serves as the reverse copy of the input data sequence. In order to prevent the forgetting of the order information, the outcomes of bidirectional structure prediction are impacted by the previous input and the subsequent input, enhancing the dependence between the training data [9].

3. Recurrent neural networks

Recurrent neural networks resemble feed-forward neural networks including an addition of edges that span recurrent time steps, providing the model an awareness of time. RNNs may not have cycles among conventional edges, which is comparable to feed-forward networks. Recurrent edges, connecting adjacent time steps, can ultimately generate cycles, including instances of length one that connect a node to itself all through time. A neural network architecture referred to as RNNs can be used to model sequence data. The behavior of RNNs, which are built from feed-forward networks, is comparable to that of human brains. Simply put, recurrent neural networks are better than other algorithms at predicting sequential data[10]. In conventional neural networks, all the inputs and outputs are independent of one another. However, there are instances in which prior words are needed, such as when predicting the next word of a sentence, and it consequently is important to remember the prior words. RNN was developed as a result, which used a Hidden Layer to solve the issue. The Hidden state, which retains particular details about a sequence, is the most essential component of an RNN. RNNs have a memory that allows them to maintain a record of all the algorithm's data. Since it produces the same output by performing the identical action on all inputs or hidden layers, it employs identical parameters for each input.

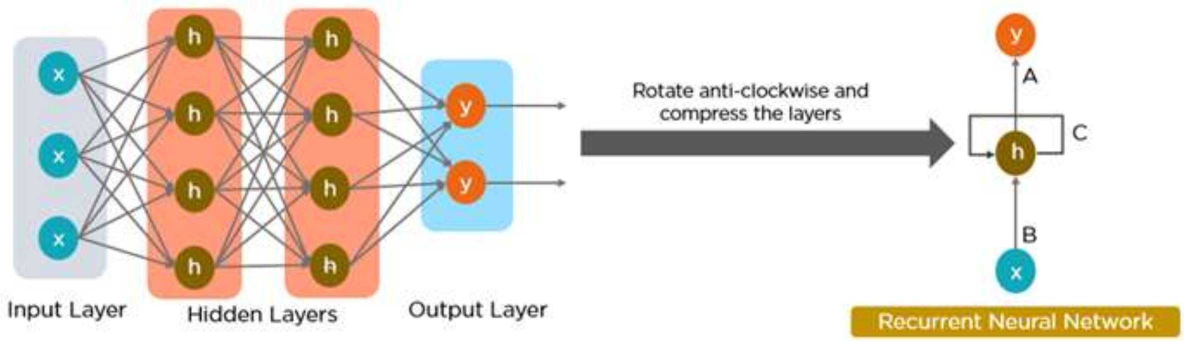

Figure 4: Structure of Recurrent Neural Network

IV. Experimental Results

In order to better understand how LSTM, GRU, and SEQ2SEQ, performed, that represents the expected outcome of the first and third tests. Other models have a similar weakness and periodicity. Table 1 shows the results of the experiment. Segmentation indicates the observed time span may extend up to 48 hours and that the overall time span is 72 hours, whereas the anticipated span is just 24 hours. The results of the experiment were assessed in MSE, for pre-trained and randomly initialized networks utilizing training and validation data. The training results for various networks, including neural networks for various AQMSs that predict PM2.5 based on the PM10 pre-trained neural networks at all AQMSs and neural networks for various AQMSs that predict PM2.5 based on randomly initialized neural networks for various AQMSs, are reported. The neural networks that employed pre-trained techniques and utilized the pre-trained network that is forecasting the same outcomes can be observed[11].

Table 1: Conduct a simulation on the dataset with a time range 48 node

Model	MAE	MAP E	COR
LSTM	2	90.25	0.30
GRU	11.70	77.88	0.38
SEQ2SEQ	15.35	111.4 9	0.35

Predicted values the resulting set of graphs demonstrates that, in general, forecasts were correct when APS concentrations were low. The problem is that outliers like the high concentration of air pollution exist, and RNNs were not intended to consciously deal with outliers during training. In most instances, using pre-trained neural networks were better than random initialized neural networks in terms of Best MSE and the number of epochs necessary to obtain the Best MSE, as can be seen from the other set of charts, which are trends of loss functions using training data and are shown in Figures.

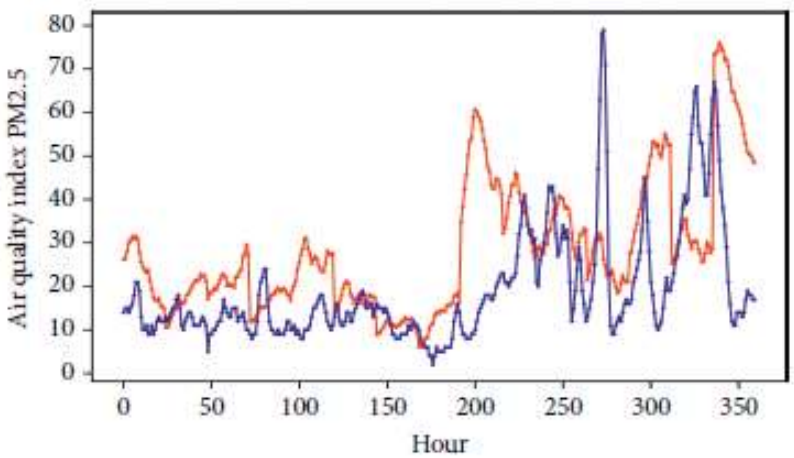

Figure 5: Results from learning employing a 64-node dataset.

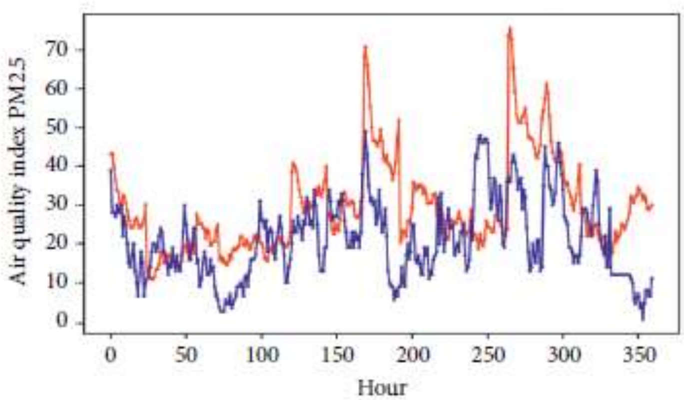

Figure 6: Real and anticipated values in a high-density residential area are contrasted.

V. Conclusion

We included a time component to reflect the "decay" effect of time on forecast because time and air pollution prediction have a significant relationship. A hidden state decoder that may be used to obtain the variation trend in historical and forecasted data was also provided. Moreover, we suggested a window approach to stabilize the total number of concealed states. Pretrained neural networks that are initialized with LSTM RNNs can increase prediction accuracy. Additionally, it is possible to decrease the number of epochs needed to train LSTM RNNs to convergence. The novel approach gives RNNs better beginning states. Predicting the values of the next day's pollutant concentration is the focus of our present research.

References

1. Asha B. Chelani et al, Mohammad M, Al-Sultan A (2020) A new method for prediction of air pollution based on intelligent computation. Soft Computing 24(1):661–680
2. C. Zhang, L. Di, Z. Sun, L. Lin, E. G. Yu, and J. Gaigalas, “Exploring cloud-based web processing service: a case study on the implementation of CMAQ as a service,” Environmental Modelling & Software, vol. 113, pp. 29–41, 2019.
3. Asha B. Chelani et al, C. Gao et al., “Multimodel simulations of a springtime dust storm over northeastern China: implications of an evaluation of four commonly used air quality models (CMAQ

v5.2.1 CAMx v6.50 CHIMERE v2017r4 and WRFChem(v3.9.1)," Geoscientific Model Development, vol. 12, no. 11, pp. 4603–4625, 2019.

4. Y. Bai, Y. Li, B. Zeng, C. Li, and J. Zhang, "Hourly PM2.5 concentration forecast using stacked autoencoder model with emphasis on seasonality," Journal of Cleaner Production, vol. 224, pp. 739–750, 2019.
5. J. Zhu, P. Wu, H. Chen, L. Zhou, and Z. Tao, "A hybrid forecasting approach to air quality time series based on endpoint condition and combined forecasting model," International
6. Journal of Environmental Research and Public Health, vol. 15, no. 9, p. 1941, 2018.
7. M. Đuri´c and D. Vujovi´c, "Short-term forecasting of air pollution index in Belgrade Serbia," Meteorological Applications, vol. 27, p. 5, 2020.
8. A. Glowacz, "Fault diagnosis of electric impact drills using thermal imaging," Measurement, vol. 171, Article ID 108815, 2021.
9. Y. Zhang, B. Chen, G. Pan, and Y. Zhao, "A novel hybrid model based on VMD-WTand PCA-BP-RBF neural network for short-term wind speed forecasting," Energy Conversion and Management, vol. 195, pp. 180–197, 2019.
10. J. Chen, J. Lu, J. C. Avise, J. A. DaMassa, M. J. Kleeman, and A. P. Kaduwela, "Seasonal modeling of PM2.5 in California's san joaquin valley," Atmospheric Environment, vol. 92, pp. 182–190, 2014.
11. D. R. Liu, S. J. Lee, Y. Huang, and C. J. Chiu, "Air pollution forecasting based on attention-based LSTM neural network and ensemble learning," Expert Systems, vol. 37, Article ID e12511, 2020.
12. Swathi, P. (2021). Industry Applications of Augmented Reality and Virtual Reality. Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN) ISSN: 2799-1172, 1(02), 14-18.
13. Swathi, P. (2013). Scope of Financial Management and Functions of Finance. International Journal of Advanced in Management, Technology and Engineering Sciences. ISSN NO-2249-7455, 3, 109-116.
14. Swathi, P. (2022). Implications For Research In Artificial Intelligence. Journal of Electronics, Computer Networking and Applied Mathematics (JECNAM) ISSN: 2799-1156, 2(02), 25-28.
15. Swathi, P. (2022). A Study On The Restrictions Of Deep Learning. Journal Of Artificial Intelligence, Machine Learning and Neural Networks, ISSN-2799-1172, 2(2), 57-61.
16. Bhaskar, N., Ramana, S., & Kumar, G. M. (2023, January). Internet of Things for Green Smart City Application Based on Biotechnology Techniques. In 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF) (pp. 1-7). IEEE.
17. Ramana, S., Bhaskar, N., Murthy, M. R., & Sharma, M. R. (2023). Machine Learning for a Payment Security Evaluation System for Mobile Networks. In Intelligent Communication Technologies and Virtual Mobile Networks (pp. 347-356). Singapore: Springer Nature Singapore.
18. Ramana, S., Bhaskar, N., & Murthy, M. R. (2020). Three Level Gateway protocol for secure M-Commerce Transactions. Solid State Technology, 63(6), 11155-11174.
19. Pothuganti, K. (2021). Long Short-Term Memory (LSTM) algorithm based prediction of stock market exchange. International Journal of Research Publication and Reviews, 2(1), 90-93.
20. Pothuganti, K. (2021). Long Short-Term Memory (LSTM) algorithm based prediction of stock market exchange. International Journal of Research Publication and Reviews, 2(1), 90-93.

