ISSN: 2278-4632
Vol-13, Issue-11, November 2023

UNRAVELING THE IMPACT OF WEATHER CONDITIONS ON AIR QUALITY PREDICTION THROUGH EXPLAINABLE DEEP LEARNING

Dr.B.Indira[1] , Gole Akanksha[2]

[1]Associate Professor, Department of MCA, Chaitanya Bharathi Institute of Technology (A), Gandipet,Hyderabad, Telangana State, India

[2]MCA Student, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad, Telangana State, India

Abstract: Meteorological situations have a robust impact on air quality and may play an critical position in air quality prediction. Air pollution is a major environmental concern affecting human health and climate change. Accurate air quality prediction is crucial for the implementation of effective pollution mitigation strategies. However, air quality prediction is challenging due to the complex relationship between meteorological conditions and air quality. To address the above problem, in this paper, we reveal the influence of weather conditions on air quality prediction by utilizing explainable deep learning. In this paper the information from air pollutant datasets, consisting of PM 2.5, and the meteorological situation datasets measuring the Temperature, humidity, and atmospheric pressure are obtained; the Long Short- Term Memory (LSTM) and Gated Recurrent Unit (GRU) fashions are set up for air quality prediction; the Shapley Additive exPlanation (SHAP) method is employed to analyze the explainability of the air quality prediction models. We discover that the prediction accuracy isn't progressed with only meteorological conditions. When combining meteorological situations with different air pollutants, the prediction accuracy is better than thinking about different air pollutants. In addition, the biggest contribution to air fine prediction is atmospheric pressure, humidity and temperature. The purpose for the unique accuracies of the prediction can also additionally due to the interplay among meteorological situations and different air pollutants.

Index Terms : *Explainable deep learning , air quality prediction , meteorological condition , long short-term memory (LSTM) , gate recurrent unit (GRU)*

1. INTRODUCTION

The continuous acceleration of global urbanization and industrialization has brought environmental problems. One of the serious environmental problems is air quality induced by the development of urbanization and industrialization. Due to the needs of transportation, production, and life, energy production and consumption processes, such as power plants, factories, and automobile exhaust emissions have ultimately led to the continuous deterioration of global air quality. Air pollution can cause various respiratory diseases and may even lead to the occurrence of cancer, which seriously threatens people's lives and health. The main air pollutants include PM2.5, PM10, and SO2, etc. Among them, PM2.5 is a fine particle with a diameter smaller than 2.5 microns. Compared with larger particulate pollutants, PM2.5 particles are more active, meaning that they can easily carry substances that affect human health and the environment, as well as remain in the air for a long time and spread quickly. PM2.5 is one of the most important sources of air pollution. Due to its small particle size, it can enter the nasal cavity and throat of the human body, and then easily cause asthma, bronchial or cardiovascular diseases. Air pollution poses a great threat to people's health. Being in an environment with severe air pollution for a long time may cause various respiratory

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

diseases and even decreased cardiopulmonary function problems. The incidence of various diseases will dramatically increase, which will overdraft people's health, affect people's living and happiness indices, and increase mortality. Air pollution also damages the ecosystem, affects its diversity and stability, and harms the environment.

Frequent air pollution incidents not only cause serious harm to human health but also cause huge economic losses and many social problems. Therefore, based on air pollution parameters, timely scientific analysis, accurate prediction of air quality and effective protection and treatment can help relevant departments and related groups take preventive measures in advance, as well as more reasonably arrange travel. People's health could be ensured, and the occurrence of diseases could be prevented.

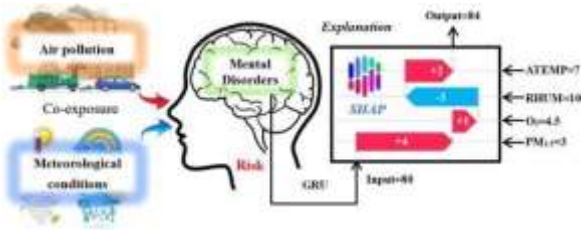

Fig 1 Example Figure

In addition, the prediction of air quality can also provide reliable information for the prevention and control of air pollution. Through further understanding of the influencing factors and changing trends of air pollutants, effective evaluation and prediction of air quality changes are helpful for the control and prevention of air pollution, which would then enable the environment and human health to be better protected. Air quality prediction is also conducive to relevant departments to understand the air quality status, and thus, a valuable theoretical basis can be provided for it. In addition, air pollution prevention and control policies can be formulated according to specific conditions. It also provides constructive opinions and suggestions for decision-makers to take more economical and efficient measures to improve air quality in the future.

2. LITERATURE SURVEY

P. Kumar [3] proposed that growing populations in cities are associated with a major increase in road vehicles and air pollution. The common excessive ranges of city air pollutants had been proven to be of a tremendous danger to town dwellers. However, the effects of very excessive however temporally and spatially limited pollution, and as a consequence exposure, are nonetheless poorly understood. Conventional methods to air best tracking are primarily based totally on networks of static and sparse dimension stations. However, those are prohibitively high priced to seize tempo-spatial heterogeneity and discover pollutants hotspots, that's required for the improvement of strong real-time techniques for publicity control. Current development in growing low-price micro-scale sensing era is substantially converting the traditional technique to permit real-time statistics in a capillary form. But the query stays whether or not there's fee withinside the much less correct information they generate.

E. D. Schraufnagel [4] studied Air pollution poses a great environmental risk to health. Outdoor exceptional particulate count (particulate count with an aerodynamic diameter < 2.5 μm) publicity is the 5th main hazard issue for loss of life

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

withinside the world, accounting for 4.2 million deaths and > 103 million disability-adjusted lifestyles years misplaced in keeping with the Global Burden of Disease Report. Air pollutants can damage acutely, normally manifested through breathing or cardiac symptoms, in addition to chronically, probably affecting each organ withinside the body. It can cause, complicate, or exacerbate many unfavorable fitness conditions. Tissue harm might also additionally end result immediately from pollutant toxicity due to the fact high-quality and ultrafine debris can benefit get right of entry to organs, or not directly thru systemic inflammatory processes.

Y.-F. Xing [5] proposed many researchers paid more attentions to the association between air pollution and respiratory system disease. In the beyond few years, ranges of smog have elevated during China resulting withinside the deterioration of air quality, elevating international concerns. PM2.5 (debris much less than 2. 5 micrometers in diameter) can penetrate deeply into the lung, aggravate and corrode the alveolar wall, and therefore impair lung function. Hence it's miles crucial to research the effect of PM2.5 at the respiration device after which to assist China fight the contemporary air pollutants problems.

X. Qi [8] studied that the air pollution caused by PM2.5, PM10, and O3 is an emerging problem that threatens public health, especially in China's megacities. Meteorological elements have enormous affects at the dilution and diffusion of air pollution which similarly have an effect on the distribution and attention of pollution. In this paper, we examine the relationships among air pollutant concentrations and meteorological situations in Beijing from January 2017 to January 2018. We observe that: the influence of a single meteorological factor on the concentration of pollutants is limited; the temperature-wind velocity aggregate, temperature-strain aggregate, and humidity-wind velocity aggregate are incredibly correlated with the awareness of pollutants, indicating that a variety of meteorological factors combine to affect the concentration of pollutants; and different meteorological factors have different effects on the concentration of the same pollutant, while the same meteorological conditions have different effects on the concentration of different pollutants. Our findings can help in predicting the air great in keeping with meteorological situations even as similarly enhancing the city control performance.

S. Al-Janabi [9] studied detection and treatment of increasing air pollution due to Technological trends constitute a number of the maximum crucial demanding situations going through the sector today. Indeed, there was a sizable growth in stages of environmental pollutants in latest years. The aim of the work presented herein is to design an intelligent predictor for the concentrations of Air pollution over the subsequent 2 days primarily based totally on deep getting to know strategies the use of a recurrent neural network (RNN). The pleasant shape for its operation is then decided the use of a particle swarm optimization (PSO) algorithm. The new predictor primarily based totally on clever computation counting on unsupervised learning, i.e., long short-term memory (LSTM) and optimization (i.e., PSO), is known as the clever air excellent prediction model (SAQPM). Thereafter, the dataset is cut up into education and checking out elements primarily based totally on the 10 cross-validation principle.

3. METHODOLOGY

However, currently, while several deep learning models utilize meteorological conditions for air quality prediction, meteorological conditions are only used as input data, and there is little research work on the influence of meteorological conditions on air quality prediction. In this case, the influence of meteorological conditions on air quality prediction in deep

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

learning models is not yet well understood, such as how it affects air quality prediction. This is because the deep learning model has the common "black box" nature, i.e., the weak explainability. Although it is possible to combine meteorological condition data with air quality data, and then use the deep learning model's powerful fitting advantage for complex data relationships to predict air quality. There are still many difficulties in analyzing the influence of meteorological condition data on air quality prediction and their correlations.

Drawbacks:

1. However, due to the "black-box" nature of deep learning, it is difficult to obtain trustworthy deep learning models when considering meteorological conditions in air quality prediction.
2. the influence of meteorological conditions on air quality prediction in deep learning models is not yet well understood, such as how it affects air quality prediction.

To address the above problems, in this paper, we reveal the impact of meteorological conditions on air quality prediction using explainable deep learning and explain how meteorological conditions affect air quality prediction accordingly. By revealing the influence of meteorological conditions on the prediction of air quality, the accuracy is further improved. Deep learning models for air quality prediction with higher accuracy and credibility can be obtained. Thus, it can be better applied in practice. This can help people plan their travel arrangements reasonably and take corresponding preventive measures on time to protect their health. Through the advanced understanding of the air quality status, corresponding prevention and control measures are adopted to realize timely and effective management.

Benefits:

1. Deep learning models for air quality prediction with higher accuracy and credibility can be obtained
2. This can help people plan their travel arrangements reasonably and take corresponding preventive measures on time to protect their health.

Modules:

- Data exploration: using this module we will load data into system
- Processing: Using the module we will read data for processing
- Splitting data into train & test: using this module data will be divided into train & test
- Model generation: Building the model – LSTM, RNN, GRU, CNN+LSTM, CNN+GRU, ARIMA, RANDOM FOREST,
KNN-SHAP, MLP and voting classifier. Algorithms accuracy calculated
- Prediction: final predicted displayed

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

Algorithms:

4. IMPLEMENTATION

differencing to convert a non-stationary time series into a stationary one, and then predict future values from historical data.

LSTM: LSTM stands for long short-term memory networks, used in the field of Deep Learning. It is a whole lot of recurrent neural networks (RNNs) which are able to mastering long-time period dependencies, mainly in series prediction problems.

RNN: Recurrent neural networks (RNNs) are the state of the art algorithm for sequential data. It is the primary set of rules that recollects its input, because of an internal memory, which makes it ideally suited for system getting to know issues that contain sequential data.

GRU: Gated recurrent units (GRUs) are a gating mechanism in recurrent neural networks. The GRU is like a protracted short-time period memory (LSTM) with a overlook gate, however has fewer parameters than LSTM, because it lacks an output gate.

CNN+LSTM: Long Short-Term Memory(LSTM) and Convolutional Neural Network(CNN).LSTM can efficiently keep the traits of historic data in lengthy textual content sequences, and extract nearby capabilities of textual content through the usage of the shape of CNN.

CNN+GRU: CNN is used for feature extraction, while GRU is used as a fully connected layer. Since COVID- 19 is a novel disease there is limited data publicly available for experiments. The data set used for this study is obtained from two different sources.

ARIMA: ARIMA models are generally denoted as ARIMA (p,d,q) where p is the order of autoregressive model, d is the degree of differencing, and q is the order of moving-average model. ARIMA models use

RANDOM FOREST: A Random Forest Algorithm is a supervised machine learning algorithm which is extremely popular and is used for Classification and Regression problems in Machine Learning.

KNN-SHAP: SHAP is a mathematical method to explain the predictions of machine learning models. It is based on the concepts of game theory and can be used to explain the predictions of any machine learning model by calculating the contribution of each feature to the prediction.

MLP: MLPClassifier stands for Multi-layer Perceptron classifier which in the name itself connects to a Neural Network. Unlike other classification algorithms such as Support Vectors or Naive Bayes, MLPClassifier relies on an underlying Neural Network to perform the task of classification.

Voting classifier: A voting classifier is a machine learning estimator that trains various base models or estimators and predicts on the basis of aggregating the findings of each base estimator. The aggregating criteria can be combined decision of voting for each estimator output.

5. EXPERIMENTAL RESULTS

ISSN: 2278-4632
Vol-12 Issue-11, November 2023

Fig 2 Home page

Fig 3 Registration page

Fig 4 Login page

Fig 5 Main page

Fig 6 Upload input values

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

Fig 7 Prediction result

6. CONCLUSION AND FUTURE SCOPE

In this paper, the essential idea is to interpret the established air quality prediction models and analyze the influence of meteorological conditions on air quality prediction. The results showed that whether only considering meteorological conditions or combining meteorological conditions and other air pollutants for PM2.5 prediction, in both the LSTM and GRU models, the meteorological conditions have a high contribution and importance to air quality prediction, meaning that they are all in the top in terms of contribution. The largest contribution to air quality prediction is made by atmospheric pressure, the second by humidity, and the third by temperature. When meteorological conditions are considered in combination with other air pollutants, the high contribution of meteorological conditions to the prediction facilitates the prediction of air quality and leads to better results. This facilitates the in-depth analysis and understanding of the deep learning models for air quality prediction and improves the trustworthiness of the deep learning models. In the future, we plan to build deep learning models with higher accuracy and trustworthiness for air quality prediction, which can be applied to realistic air quality prediction.

REFERENCES

[1] H. Kan, R. Chen, and S. Tong, ‘‘Ambient air pollution, climate change, and population health in China,’’ Environ. Int., vol. 42, pp. 10–19, Jul. 2012.

[2] H. Zhang, S. Wang, J. Hao, X. Wang, S. Wang, F. Chai, and M. Li, ‘‘Air pollution and control action in Beijing,’’ J. Cleaner Prod., vol. 112, pp. 1519–1527, Jan. 2016.

[3] P. Kumar, L. Morawska, C. Martani, G. Biskos, M. Neophytou, S. Sabatino, M. Bell, L. Norford, and R. Britter, ‘‘The rise of low-cost sensing for managing air pollution in cities,’’ Environ. Int., vol. 75, pp. 199– 205, Feb. 2015.

[4] E. D. Schraufnagel, R. J. Balmes, T. C. Cowl, S. D. Matteis, S.-H. Jung, K. Mortimer, R. Perez-Padilla, B.
M. Rice, H. Riojas-Rodriguez, A. Sood, D. G. Thurston, T. To, A. Vanker, and J. D. Wuebbles, ‘‘Air pollution and noncommunicable diseases a review by the forum of international respiratory societies’ environmental committee, part 2: Air pollution and organ systems,’’ Chest, vol. 155, no. 2, pp. 417–426,
2019.

[5] Y.-F. Xing, Y.-H. Xu, M.-H. Shi, and Y.-X. Lian, ‘‘The impact of PM2.5 on the human respiratory system,’’ J.

Thoracic Disease, vol. 8, no. 1, pp. E69– E74, 2016.

[6] J. J. West, A. Cohen, F. Dentener, B. Brunekreef,
T. Zhu, B. Armstrong, M. L. Bell, M. Brauer, G. Carmichael, D. L. Costa, and D. W. Dockery, ‘‘What we breathe impacts our health: Improving understanding of the link between air pollution and health,’’ Environ. Sci. Technol., vol. 50, no. 10, pp. 4895–4904, 2016.

[7] X. Zhang, X. Zhang, and X. Chen, ‘‘Happiness in the air: How does a dirty sky affect mental health and subjective well-being?’’ J. Environ. Econ. Manage., vol. 85, pp. 81–94, Sep. 2017.

[8] X. Qi, G. Mei, S. Cuomo, C. Liu, and N. Xu, ‘‘Data analysis and mining of the correlations between meteorological conditions and air quality: A case study in Beijing,’’ Internet Things, vol. 14, Jun. 2021, Art. no. 100127.

[9] S. Al-Janabi, M. Mohammad, and A. Al-Sultan, ‘‘A new method for prediction of air pollution based on intelligent computation,’’ Soft Comput., vol. 24, no. 1, pp. 661–680, Jan. 2020.

[10] A. Alimissis, K. Philippopoulos, C. G. Tzanis, and D. Deligiorgi, ‘‘Spatial estimation of urban air pollution with the use of artificial neural network models,’’ Atmos. Environ., vol. 191, pp. 205–213, Oct. 2018.

[11] H. Li, J. Wang, R. Li and H. Lu, "Novel analysis- forecast system based on multi-objective optimization for air quality index", J. Cleaner Prod., vol. 208, pp. 1365-1383, Jan. 2019. Show in Context CrossRef Google Scholar

[12] A. Kumar and P. Goyal, "Forecasting of daily air quality index in Delhi", Sci. Total Environ., vol. 409, no. 24, pp. 5517-5523, Nov. 2011.

[13] W. G. Cobourn, "An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations", Atmos. Environ., vol. 44, no. 25, pp. 3015-3023, Aug. 2010.

[14] T. S. Rajput and N. Sharma, "Multivariate regression analysis of air quality index for Hyderabad city: Forecasting model with hourly frequency", Int. J. Appl. Res., vol. 3, no. 8, pp. 443-447, 2017.

[15] Y. Liu, Q. Zhu, D. Yao and W. Xu, "Forecasting urban air quality via a back-propagation neural network and a selection sample rule", Atmosphere, vol. 6, no. 7, pp. 891-907, Jul. 2015.

[16] S. Xiao, Q. Y. Wang, J. J. Cao, R.-J. Huang, W.
D. Chen, Y. M. Han, et al., "Long-term trends in visibility and impacts of aerosol composition on visibility impairment in Baoji China", Atmos. Res., vol. 149, pp. 88-95, Nov. 2014.

[17] Y. Qi, Q. Li, H. Karimian and D. Liu, "A hybrid model for spatiotemporal forecasting of PM2.5 based on graph convolutional neural network and long short- term memory", Sci. Total Environ., vol. 664, pp. 1-10, May 2019.

[18] C. Wen, S. Liu, X. Yao, L. Peng, X. Li, Y. Hu, et al., "A novel spatiotemporal convolutional long short- term neural

ISSN: 2278-4632
Vol-13, Issue-11, November 2023

network for air pollution prediction", Sci. Total Environ., vol. 654, pp. 1091-1099, Mar. 2019.

[19] J. Schmidhuber, "Deep learning in neural networks: An overview", Neural Netw., vol. 61, pp. 85-117, Jan. 2015.

[20] Z. Ma and G. Mei, "Deep learning for geological hazards analysis: Data models applications and opportunities", Earth-Sci. Rev., vol. 223, Dec. 2021.

[21] S. Du, T. Li, Y. Yang and S.-J. Horng, "Deep air quality forecasting using hybrid deep learning framework", IEEE Trans. Knowl. Data Eng., vol. 33, no. 6, pp. 2412-2424, Jun. 2021.

[22] Y. Zhou, F.-J. Chang, L.-C. Chang, I.-F. Kao and Y.-S. Wang, "Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts", J. Clean Prod., vol. 209, pp. 134- 145, Feb. 2019.

[23] Y. Jiao, Z. Wang and Y. Zhang, "Prediction of air quality index based on LSTM", Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC), pp. 17-20, May 2019.

[24] R. J. Kuo, B. Prasetyo and B. S. Wibowo, "Deep learning-based approach for air quality forecasting by using recurrent neural network with Gaussian process in Taiwan", Proc. IEEE 6th Int. Conf. Ind. Eng. Appl. (ICIEA), pp. 471-474, Apr. 2019.

[25] Y.-S. Chang, H.-T. Chiao, S. Abimannan, Y.-P. Huang, Y.-T. Tsai and K.-M. Lin, "An LSTM-based aggregated model for air pollution forecasting", Atmos. Pollut. Res., vol. 11, no. 8, pp. 1451-1463, Aug. 2020.