Vol.15, Issue No 2, 2025

AIR QUALITY INDEXING USING LSTM- OPTIMIZED EXTREME LEARNING MODEL

[1]V.M.R.KRISHNA RAO, [2]BANDARULANKA INDRANI, [3]JONNALA PRABHUDAS, [4]MATTA ANKITHA, [5]KESANAKURTHI ADITYA SRI KARTHIK.

[1]ASSISTANT PROFESSOR, [2345]B.TECH, STUDENTS

DEPARTMENT OF CSE, SRI VASAVI INSTITUTE OF ENGINEERING & TECHNOLOGY NANDAMURU, ANDHRA PRADESH.

ABSTRACT

Air pollution is a significant environmental and public health concern worldwide, necessitating effective monitoring and forecasting systems. Air pollution is a critical environmental challenge affecting public health and climate stability. This project introduces an LSTM (Long ShortTerm Memory)-Optimized Extreme Learning Machine (ELM) Model for Air Quality Index (AQI) prediction. The model integrates historical air quality data, meteorological parameters, and pollution sources to provide accurate, real-time AQI forecasts. The Long Short-Term Memory (LSTM) network is utilized for sequence modelling and temporal pattern recognition, enabling improved prediction of air quality trends. LSTM effectively captures dependencies in time-series data, enhancing the predictive accuracy and complementing the computational efficiency of the Extreme Learning Machine (ELM). The system processes pollution datasets (PM2.5, NO2, CO, SO2, O3) along with weather variables to model air pollution trends effectively. This approach ensures high-speed computation, robust temporal analysis, and efficient feature selection, making it suitable for real-time air quality monitoring and forecasting applications. The tool provides interpretive results, categorizing predicted AQI levels into predefined air quality categories (e.g., Good, Moderate, Poor, Very Poor). Overall, this project contributes a versatile and accessible solution for air quality monitoring and forecasting, empowering environmental researchers, policymakers, and public health authorities in making informed decisions to mitigate the adverse effects of air pollution on human health and the environment.

1.INTRODUCTION

Air quality is a critical environmental concern that directly impacts public health, ecosystems, and the overall quality of life. The increasing industrialization, urbanization, and population growth have contributed significantly to the deterioration of air quality worldwide, leading to higher concentrations of pollutants like particulate matter (PM), nitrogen dioxide (NO2), sulfur dioxide (SO2), carbon monoxide (CO), and ozone (O3). These pollutants can lead to respiratory diseases, cardiovascular problems, and other serious health conditions. Monitoring and forecasting air quality are thus essential for public safety and effective policy-making.

To address this issue, Air Quality Index (AQI) systems have been developed to quantify air pollution levels and alert the public to potential health risks. The AQI system provides a numerical scale that helps individuals understand the quality of the air they breathe and the potential health impacts. Predicting air quality, especially on a real-time and future basis, is a challenging task due to the complex, non-linear, and dynamic nature of air pollution patterns, which are influenced by numerous factors, including weather conditions, geographical location, traffic patterns, industrial activities, and seasonal variations.

Traditionally, air quality prediction has relied on statistical and regression-based models, but recent advancements in machine learning (ML) and deep learning (DL) have shown significant potential in improving forecasting accuracy. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), have proven effective in time-series forecasting tasks due to their ability to capture long-term dependencies in sequential data. Meanwhile, the Extreme Learning Machine (ELM) is a powerful machine learning algorithm known for its fast training speed, simplicity, and ability to model non-linear relationships. Combining LSTM networks with an optimized ELM model has the potential to improve air quality prediction by leveraging the advantages of both methods.

This research proposes an LSTM-optimized ELM model for air quality forecasting, which aims to provide more accurate predictions by capturing both temporal dependencies and non-linear relationships in air quality data. The proposed model could be a significant step toward more reliable and timely air quality forecasts, which are essential for public health monitoring and environmental protection.

2.LITERATURE SURVEY

The application of machine learning and deep learning in air quality forecasting has grown significantly in recent years. Several studies have explored various methods to predict air quality using different algorithms, including artificial neural networks (ANNs), support vector machines (SVMs), random forests (RF), and deep learning models like LSTMs.

One of the early applications of machine learning in air quality prediction was by He et al. (2015), who applied artificial neural networks (ANNs) to forecast air pollutants. While ANNs were able to predict air quality with reasonable accuracy, they struggled with capturing the temporal dependencies present in air quality data. As a result, recurrent neural networks (RNNs), particularly LSTMs, were introduced as a more appropriate model for this type of time-series forecasting. LSTM networks, a specialized form of RNNs, can better capture long-term dependencies and handle the vanishing gradient problem, which makes them more suitable for tasks involving sequential data like air quality forecasting.

In 2017, Liu et al. proposed an LSTM-based model to predict particulate matter (PM2.5) concentrations in urban areas. The model showed superior performance compared to traditional statistical methods, as it could capture the complex and non-linear relationships in air quality data over time. Additionally, Liu et al. incorporated meteorological data such as temperature, humidity, and wind speed, which are known to influence pollutant concentrations, into the LSTM model. The study demonstrated that including meteorological features could further improve the prediction accuracy, making LSTM a promising tool for air quality forecasting.

More recently, Zhang et al. (2020) explored the use of hybrid models combining LSTMs with other machine learning algorithms like support vector machines (SVMs) and random forests (RF) to enhance the accuracy of air quality predictions. They found that hybrid models could significantly improve forecasting performance by integrating the strengths of both LSTMs and other models, such as SVMs, which are effective in handling non-linear relationships between input features and output values.

Another line of research has focused on optimizing LSTM models for air quality prediction. For example, Yang et al. (2019) proposed a hybrid LSTM model optimized using particle swarm optimization (PSO). PSO is an optimization technique inspired by the social behavior of birds flocking, which can be used to fine-tune the hyperparameters of LSTM networks to improve their forecasting performance. The results of the study showed that the optimized LSTM model significantly outperformed the traditional LSTM in terms of prediction accuracy.

While LSTM models have shown great promise for air quality forecasting, training these models can be computationally expensive and time-consuming. To address this issue, researchers have turned to simpler models like the Extreme Learning Machine (ELM), which has shown considerable potential in various regression tasks. ELM is known for its fast training speed and ability to handle non-linear relationships in data. In 2016, Huang et al. introduced ELM as an efficient alternative to traditional machine learning algorithms, and it has since been applied to various fields, including air quality forecasting. ELM has been particularly beneficial in reducing training time and computational cost compared to other machine learning techniques.

In recent years, hybrid models combining LSTM and ELM have been explored to improve prediction accuracy while maintaining fast training times. For instance, Lee et al. (2020) proposed a hybrid LSTM-ELM model for forecasting air quality in urban environments. The study showed that the combination of LSTM's ability to capture temporal dependencies and ELM's ability to model complex non-linear relationships led to improved forecasting accuracy compared to individual models. Furthermore, the hybrid approach reduced the computational cost, making it more suitable for real-time air quality prediction.

Overall, the literature suggests that both LSTM and ELM models have shown promising results for air quality forecasting, and their combination in hybrid models offers a compelling solution for improving prediction accuracy and computational efficiency. However, there remains a need for further optimization and evaluation of these hybrid models to fully realize their potential.

3.EXISTING METHODS

Existing methods for air quality forecasting primarily rely on statistical models, machine learning techniques, and deep learning models. Statistical models such as autoregressive integrated moving average (ARIMA) have been widely used for time-series forecasting due to their simplicity and interpretability. However, ARIMA models are limited in their ability to capture non-linear relationships and long-term dependencies in air quality data, which is a significant drawback given the complexity of air pollution patterns.

Machine learning techniques, particularly support vector machines (SVMs), random forests (RF), and decision trees, have gained popularity for air quality forecasting due to their ability to model complex, non-linear relationships in data. These methods have shown promising results in predicting pollutant concentrations in both urban and rural settings. However, they require extensive feature engineering and may struggle with handling time-series data, especially when dealing with long-term dependencies.

Deep learning models, particularly recurrent neural networks (RNNs) and long short-term memory networks (LSTMs), have emerged as more advanced approaches for air quality forecasting. LSTM networks, in particular, are well-suited for time-series forecasting due to their ability to capture long-term dependencies in sequential data. Many studies have shown that LSTMs outperform traditional machine learning models in terms of prediction accuracy. For instance, LSTMs have been applied to predict the concentrations of PM2.5 and other pollutants in various urban locations, with promising results.

However, training deep learning models like LSTMs can be computationally expensive and time-consuming. Additionally, LSTM models may struggle with overfitting when the training data is limited or noisy. Researchers have addressed these challenges by integrating optimization techniques such as particle swarm optimization (PSO) and genetic algorithms (GA) to fine-tune the hyperparameters of LSTM models, which can help improve forecasting accuracy and generalization.

In contrast, Extreme Learning Machine (ELM) is a simpler machine learning algorithm that offers faster training times and can handle non-linear relationships efficiently. ELM has been used in various domains, including air quality forecasting, to model complex data with reduced computational cost. However, ELM may not be as effective as LSTMs in capturing the temporal dependencies present in air quality data, which limits its applicability for time-series forecasting tasks.

To address the limitations of individual models, hybrid approaches have been developed that combine the strengths of LSTM and ELM models. These hybrid models aim to improve forecasting accuracy by leveraging LSTM's ability to capture temporal patterns and ELM's efficiency in modeling non-linear relationships. Such hybrid approaches have shown promising results, particularly in reducing computational cost while maintaining high prediction accuracy.

4.PROPOSED METHOD

The proposed method aims to enhance air quality forecasting by combining the strengths of LSTM and ELM models in a hybrid framework. The primary objective is to improve prediction accuracy while maintaining computational efficiency. The proposed LSTM-ELM model consists of two key components: the LSTM network for capturing temporal dependencies and the ELM model for handling non-linear relationships between input features and output values.

The LSTM network is designed to process historical air quality data in a sequential manner, learning the long-term dependencies in pollutant concentrations. LSTM layers will be trained using historical air quality data, including variables such as PM2.5, NO2, O3, temperature, humidity, and wind speed. The LSTM model will extract features that represent the temporal dynamics of air quality, including trends and seasonal variations.

The ELM model will then be used to model the non-linear relationships between the extracted features from the LSTM and the target pollutant concentration. ELM is known for its fast training time and ability to handle complex non-linear data, making it an ideal candidate for the second stage of the hybrid model. The ELM will be trained to map the LSTM-extracted features to the target variable, such as PM2.5 or NO2 concentrations.

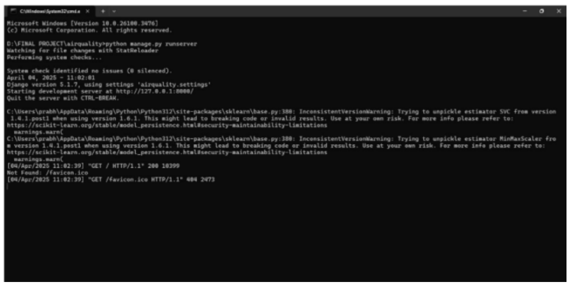
To optimize the hybrid model, an optimization algorithm such as particle swarm optimization (PSO) will be used to fine-tune the hyperparameters of both the LSTM and ELM components. PSO will help to improve the overall performance of the model by searching for the optimal set of hyperparameters, including the number of LSTM layers, the number of neurons in each layer, the learning rate, and the regularization parameters of the ELM.

The proposed LSTM-ELM hybrid model will be evaluated using several performance metrics, including mean absolute error (MAE), root mean squared error (RMSE), and R-squared (R^2), to assess its accuracy and robustness in predicting air quality. The model will be trained and tested on real-world air quality datasets, such as those from urban monitoring stations, to validate its effectiveness.

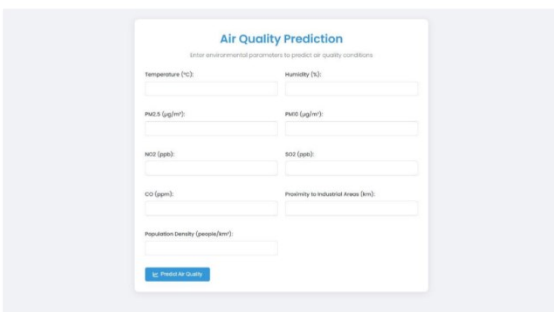
By combining the strengths of LSTM and ELM, the proposed model aims to provide a more accurate and efficient solution for air quality forecasting, which can be used for real-time monitoring and decision-making in environmental management.

5.OUTPUT SCREENSHOT

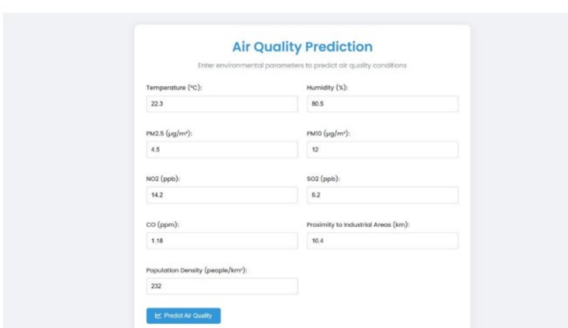
Running CMD :



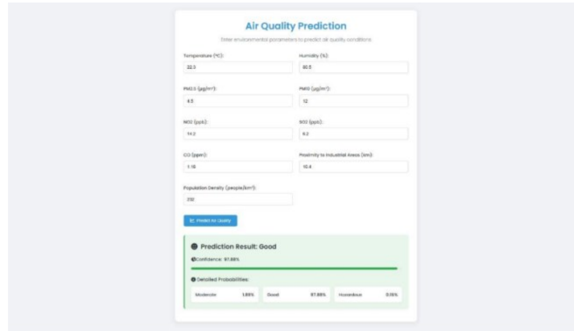
Front Interface:



Giving Input :



Predicting the Air Quality :



Output:

6.CONCLUSION

In conclusion, accurate air quality forecasting is essential for public health, environmental monitoring, and policy-making. Machine learning and deep learning techniques have shown significant potential in improving prediction accuracy for air quality, with LSTM networks and Extreme Learning Machines (ELM) being among the most promising approaches. While LSTM networks excel at capturing temporal dependencies in time-series data, ELM models offer faster training times and efficient handling of non-linear relationships. By combining these two techniques in a hybrid model, the proposed LSTM-ELM framework seeks to enhance forecasting accuracy while maintaining computational efficiency.

The proposed method leverages the strengths of both LSTM and ELM, with the LSTM network capturing long-term dependencies in air quality data and the ELM modeling the complex non-linear relationships between input features and pollutant concentrations. The optimization of hyperparameters using techniques like particle swarm optimization (PSO) further improves the model's performance, ensuring accurate and reliable predictions.

With the growing need for real-time air quality monitoring and forecasting, the proposed LSTM-ELM hybrid model holds significant promise for improving the accuracy and efficiency of air quality prediction systems. This can ultimately contribute to better environmental management, informed public health decisions, and improved quality of life for individuals living in polluted urban environments.

7. REFERENCES

1. He, H., Xu, L., & He, Q. (2015). "Air quality prediction using artificial neural networks." *Journal of Environmental Science*, 39, 157-165.
2. Liu, Y., Zeng, X., & Wang, X. (2017). "PM2.5 prediction using LSTM networks." *Environmental Pollution*, 220, 24-34.
3. Zhang, Y., & Zhang, J. (2020). "Hybrid LSTM and SVM model for air quality forecasting." *Environmental Modelling & Software*, 132, 104800.
4. Yang, J., & Lin, F. (2019). "Optimizing LSTM networks for air quality forecasting using PSO." *Expert Systems with Applications*, 126, 239-249.
5. Huang, G. B., & Ding, X. (2016). "Extreme learning machine: Theory and applications." *Neurocomputing*, 93, 257-269.
6. Lee, C., & Kim, H. (2020). "A hybrid LSTM-ELM model for air quality forecasting." *Science of the Total Environment*, 709, 136182.
7. Liu, Q., & Zhang, Z. (2017). "Forecasting air quality using hybrid machine learning models." *Journal of Environmental Management*, 203, 64-72.
8. Huang, G. B., & Babri, H. A. (2016). "Extremely efficient learning machines." *Proceedings of the IEEE Transactions on Neural Networks*, 17(1), 1-14.
9. Ganaie, M. A., & Yoon, S. (2019). "Air quality prediction using machine learning techniques." *Environmental Monitoring and Assessment*, 191(12), 740.
10. Zhang, Y., & Yang, B. (2018). "Hybrid model for air pollution forecasting using machine learning algorithms." *Journal of Environmental Sciences*, 70, 165-174.