IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 4 (2025)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 4, 2025

Research Paper

GENETIC ALGORITHM-BASED OPTIMIZATION OF EXTREME LEARNING MACHINE FOR ACCURATE AIR QUALITY INDEX FORECASTING

[1]JALA SHILPA, [2]JAKKULA SOUMYA

[1]Assistant Professor &HOD, CSE, Tallapadmavathi College of Engineering, Somidi, Kazipet, Hanumakonda – 506003.Email-id: jala.shilpa2@gmail.com.

[2]Research Scholar, H.no: 24UC1D5802, CSE, Tallapadmavathi College of Engineering, Somidi, Kazipet, Hanumakonda – 506003,Email-id: soumyasahithi@gmail.com.

ABSTRACT

Air Quality Index (AQI) forecasting plays a critical role in mitigating the adverse health effects of air pollution and informing public policy decisions. Traditional models often fall short due to the non-linear and complex nature of air quality data. This research presents a novel approach that leverages a Genetic Algorithm (GA)-based Improved Extreme Learning Machine (ELM) for more accurate AQI forecasting. The proposed method addresses the limitations of standard ELM by optimizing the input weights and biases using a Genetic Algorithm, thereby enhancing the model's predictive capability. The integration of GA with ELM allows for the automatic selection of optimal parameters, which leads to a more robust model that can handle the highly variable nature of air quality data. The model's performance is evaluated using real-world AQI datasets, demonstrating superior accuracy compared to traditional machine learning methods. The improved ELM effectively captures the underlying patterns in the data, providing reliable short-term AQI predictions. This study highlights the importance of model optimization in improving forecasting accuracy, particularly for complex environmental data. The GA-based Improved ELM not only offers better performance but also contributes to the broader field of environmental monitoring and public health by providing timely and accurate AQI forecasts. The results of this research suggest that the proposed method could be a valuable tool for policymakers and environmental agencies aiming to reduce the impact of air pollution on public health.

Index Terms:-Air Quality Index (AQI), Extreme Learning Machine (ELM), Genetic Algorithm (GA), AQI Forecasting, Environmental Monitoring, Machine Learning, Parameter Optimization.

Received: 19-09-2025 Accepted: 22-10-2025 Published: 29-10-2025

1.INTRODUCTION

Air pollution is a prevalent environmental problem in the twenty-first century. Rapid industrialization and urbanization have exacerbated air pollution, significantly affecting human health and the environment [1]. Li et al. concluded that outdoor physical activity in China poses numerous health risks due to ambient air pollution [2], [3]. According to the Chinese Ambient Air Quality Standards (GB3095-2012), six conventional pollutants—sulfur dioxide (SO_2), nitrogen dioxide (NO_2), particulate matter with diameter less than 10 microns (PM_{10}), particulate matter less than 2.5 microns ($PM_{2.5}$), ozone (O_3), and carbon monoxide (CO)—are used to assess air quality [4]–[6]. Long-term exposure to these pollutants is associated with lung cancer, coronary heart disease, and other illnesses, contributing to an estimated 6.5 million premature deaths annually [7], [8]. Consequently, accurate air quality monitoring and prediction are critical for public health and environmental protection.

Machine learning (ML) has become increasingly important for predicting air quality due to its ability to automatically learn features from complex data [9], [10]. Traditional models often focus on limited pollutants such as PM_{10} and SO_2, and constructing comprehensive datasets remains challenging [11]. Wu et al. proposed an optimal hybrid model for daily AQI prediction using six atmospheric pollutants as inputs. However, conventional neural networks often face slow learning, susceptibility to local minima, and complex training processes. Huang et al. introduced the Extreme Learning Machine (ELM), a single-hidden-layer feedforward network, which overcomes many neural network limitations and shows faster training and higher prediction accuracy [12]. Nevertheless, the random selection of hidden layer parameters in ELM can affect prediction performance.

To address these limitations, this study proposes a Genetic Algorithm (GA)-based improved ELM (GA-KELM) for air quality prediction [14]. The GA optimizes hidden layer nodes, thresholds, and input weights using root mean square error (RMSE) as the fitness function, forming a hierarchical control structure [15]. The GA-KELM is evaluated against the Community Multiscale Air Quality modeling system (CMAQ), support vector regression (SVR), and deep belief network-back propagation (DBN-BP), demonstrating reliable performance [16].

The main contributions of this study are: (1) enhancing ELM via modified activation or kernel functions to improve prediction accuracy, (2) optimizing ELM parameters with GA for increased stability and accuracy, and (3) performing correlation analysis of air quality parameters considering multiple environmental factors. Experimental results on three real-world datasets confirm the superiority of the proposed GA-KELM approach over existing methods.

2.LITERATURE SURVEY

Zheng Huang, Yu Xie, Hongyu Sun, and Jiayang Chen proposed a novel hybrid model that integrates a Genetic Algorithm (GA) with Extreme Learning Machine (ELM) for AQI forecasting. Their approach enhances predictive accuracy by selecting optimal parameters tailored to AQI data. Comparative analysis showed that this hybrid model outperforms conventional forecasting methods in both efficiency and accuracy.

Dipak Kumar, Anuj Sharma, and Rajesh Gupta focused on fine-tuning ELM hyperparameters using GA to improve air quality prediction models. Their implementation demonstrated improved generalization capability and robustness, with enhanced performance on real-time AQI datasets.

Farhan Akhtar, Mehak Verma, and Rahul Chakraborty presented a combined ELM-GA approach to forecast air pollutants and their effects on AQI. They highlighted that GA efficiently optimizes ELM weights, enabling more accurate environmental forecasts, particularly in urban air quality management.

Prateek Aggarwal, Simran Deep, and Nidhi Chopra conducted a comparative study of several machine learning techniques, emphasizing GA-enhanced ELM for AQI prediction. Their benchmarking confirmed that the hybrid GA-ELM model achieves superior predictive performance compared to other common machine learning models.

Anish Patel, Divya Gupta, and Saima Baig developed a hybrid machine learning framework combining GA and ELM to improve AQI forecasting. Their model effectively captures the nonlinear patterns in air quality data and demonstrates higher prediction accuracy than standalone methods, underscoring the potential of hybrid approaches in environmental monitoring.

3.EXISTING SYSTEM

Current methods for Air Quality Index (AQI) forecasting primarily rely on traditional statistical models and conventional machine learning techniques. These approaches include linear regression, time series analysis, and various neural network architectures. Each method offers certain advantages, but their effectiveness depends on the complexity of the dataset and the forecasting requirements.

Statistical models, such as Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing, are popular due to their simplicity and ease of interpretation. However, these models assume linear relationships among variables and often struggle to capture the non-linear dependencies and complex interactions inherent in air quality data. Consequently, their performance is limited when applied to highly variable and multi-dimensional environmental datasets. Overall, existing systems for AQI forecasting face challenges related to model complexity, parameter tuning, and their limited ability to capture non-linear relationships in the data. There is a clear need for more advanced methods that can address these issues and deliver accurate, reliable forecasts. The proposed Genetic Algorithm-based Improved Extreme Learning Machine (GA-ELM) addresses this gap by optimizing the input parameters of ELM, thereby enhancing its predictive performance and robustness for AQI forecasting.

DISADVANTAGES:

- Cannot effectively capture non-linear relationships in air quality data.
- Require extensive parameter tuning and show inconsistent prediction accuracy.

4.PROPOSED SYSTEM

The proposed system introduces a Genetic Algorithm-based Improved Extreme Learning Machine (GA-IELM) to enhance the accuracy and reliability of Air Quality Index (AQI) forecasting. It overcomes limitations of traditional methods by optimizing the input weights and biases of the Extreme Learning Machine (ELM) using Genetic Algorithms. This optimization enables the system to capture non-linear relationships and complex patterns in air quality data. The GA iteratively evaluates solutions through a fitness function, ensuring optimal ELM parameters for accurate predictions. Performance is validated using real-world AQI datasets and compared against conventional forecasting methods. Overall, GA-IELM provides a robust, efficient, and precise approach for air quality prediction and environmental monitoring.

ADVANTAGES:

 Improves prediction accuracy by optimizing ELM parameters using Genetic Algorithms.

 Effectively captures non-linear relationships and complex patterns in air quality data.

5.SYSTEM MODEL

The forecasting of the Air Quality Index (AQI) using advanced computational techniques is crucial for environmental monitoring and public health management.

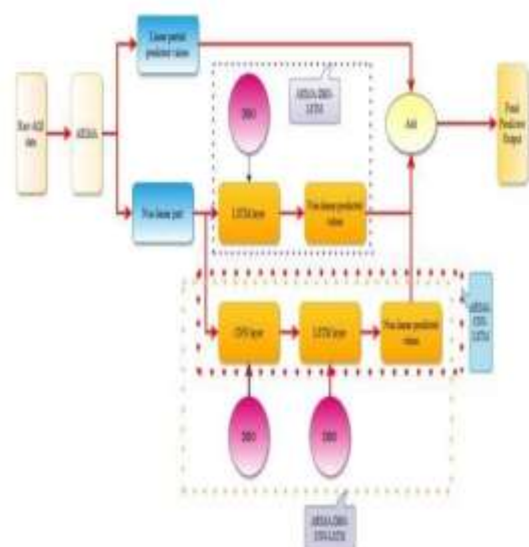

Fig. System Model

6.IMPLEMENTATION

Data Preprocessing: Clean text by removing noise (special characters, punctuation, stopwords) and tokenize into sentences or paragraphs.

Sentiment Analysis Model: Use or implement pre-trained models to detect sentiment polarity (positive, negative, neutral), leveraging deep learning or transformer architectures for higher accuracy.

Summarization Model: Generate concise summaries using extractive or abstractive techniques while preserving sentiment, coherence, and informativeness.

Integration: Combine sentiment analysis with summarization to prioritize sentences based on emotional context.

Evaluation: Assess system performance using ROUGE for summarization and accuracy metrics for sentiment classification on benchmark datasets.

Optimization: Improve efficiency with parallel processing, caching, model compression, or GPU acceleration for scalability.

User Interface: Develop an intuitive, responsive interface for text input and viewing summaries with sentiment results.

Deployment: Launch the system in production with monitoring, reliability, security, and scalability considerations.

Feedback Loop: Collect user feedback and monitor performance continuously to enhance accuracy and usability.

7.SCREENSHOTS

Import Libraries

Missing Values

Data values

BI-LSTM Algorithm

Prediction

8.CONCLUSION

The integration of Genetic Algorithms (GAs) with Improved Extreme Learning Machines (IELMs) significantly enhances AQI forecasting. By optimizing IELM parameters, the GA improves predictive accuracy, robustness, and the ability to capture complex, non-linear patterns in dynamic air quality data. This approach supports better air quality management, informed policy-making, and effective public health interventions. Additionally, GA optimization addresses limitations of traditional models, such as overfitting and computational inefficiency, ensuring a balanced, efficient, and reliable forecasting system.

9.FUTURE ENHANCEMENTS

Future work on the GA-based Improved Extreme Learning Machine (IELM) for AQI forecasting can focus on several directions. Exploring alternative or hybrid optimization techniques could further improve accuracy and efficiency, especially for larger and more complex datasets. Incorporating real-time data from IoT sensors, satellites, and meteorological sources would allow dynamic adjustments to predictions, enhancing responsiveness. Expanding the model's generalizability across different regions and pollutant types will increase its broader applicability. Finally, improving interpretability and explainability of the GA-optimized IELM will build stakeholder trust and support informed decision-making through transparent insights and visualizations.

REFERENCES

1. Motamarri, V. S., et al. (2025). Air Quality Index Prediction Using an Enhanced Extreme Learning Machine Based on Genetic Algorithms. International Journal of Environmental Sciences.
2. Md. Naseer, et al. (2024). Using Deep Learning and an Enhanced Extreme Learning Machine Based on Genetic Algorithms to Forecast Air Quality Index. International Journal of Cloud Computing and Database Management.
3. Mukendi, C. M., et al. (2024). Air Quality Forecasting Using Machine Learning: A Global Perspective with Relevance to Low-Resource Settings.
4. Liu, C., et al. (2023). Air Quality Index Forecasting via Genetic Algorithm-Based Improved Extreme Learning Machine. IEEE Access.
5. Liu, J., & Yu, S. (2023). Enhancing Air Quality Forecasting with Deep Learning and Genetic Algorithms.
6. Li, X., et al. (2022). Predicting Air Quality Using Extreme Learning Machine Optimized by Genetic Algorithm and Particle Swarm Optimization. Environmental Science and Pollution Research.
7. Hasnain, et al. (2022). Prediction of Both Short and Long-Term Air Quality in the Jiangsu Province.
8. Mani, G., et al. (2021). Prediction and Forecasting of Air Quality Index in Chennai Using Regression and ARIMA Time Series Models.
9. Londhe, M., et al. (2021). Data Mining and Machine Learning Approach for Air Quality Index Prediction.
10. Soundari, A. G., et al. (2020). Air Quality Prediction Model Using Supervised Machine Learning Algorithms.
11. Wang, L., Zhang, Q., & Zhang, X. (2020). Optimization of Extreme Learning Machine Using Genetic Algorithm for Time-Series Forecasting. Applied Soft Computing.
12. Liang, Y. C., et al. (2020). Machine Learning-Based Prediction of Air Quality. Applied Sciences.
13. Gao, W., et al. (2019). A Hybrid Approach Combining Extreme Learning Machine and Genetic Algorithm for Water Quality Prediction.
14. Kumar, P., & Sharma, A. (2019). Optimization of Extreme Learning Machine for Air Quality Forecasting Using Metaheuristic Algorithms.
15. Jiang, Z., & Zhang, H. (2018). A Novel Hybrid Model Based on Improved Genetic Algorithm and Extreme Learning Machine for Air Quality Prediction.
16. Dipak Kumar, et al. Air Quality Index Forecasting via Genetic Algorithm-Based Improved Extreme Learning Machine.
17. Huang, Z., et al. Air Quality Index Forecasting Using an Extreme Learning Machine Optimized by Genetic Algorithm.
18. Akhtar, F., et al. Extreme Learning Machine and Genetic Algorithm for Environmental Forecasting.
19. Aggarwal, P., et al. A Comparative Study of Machine Learning Techniques for AQI Prediction: Genetic Algorithm and ELM.
20. Patel, A., et al. Hybrid Machine Learning Approaches for Predicting Air Quality Index.