

Research Paper

A Robust Hybrid Learning Framework Combining GA-KELM and BiLSTM for Air Quality Prediction

Mohd Tarique Khan¹, Md Ashique Hussain², Mohammed Asad Bin³, Arham Hussain⁴,
Faiz Ahmed⁵

^{1,2}Assistant Professor, Department of CSE (Data Science),

Lords Institute of Engineering and Technology, Hyderabad, Telangana, India.

^{3,4,5}UG Students, Department of CSE (Data Science),

Lords Institute of Engineering and Technology, Hyderabad, Telangana, India.

Abstract— Air quality has always been one of the most important environmental concerns for the general public and society. Using machine learning algorithms for Air Quality Index (AQI) prediction is helpful for the analysis of future air quality trends from a macro perspective. When conventionally using a single machine learning model to predict air quality, it is challenging to achieve a good prediction outcome under various AQI fluctuation trends. In order to effectively address this problem, a genetic algorithm-based improved extreme learning machine (GA-KELM) prediction method is enhanced. First, a kernel method is introduced to produce the kernel matrix which replaces the output matrix of the hidden layer. To address the issue of the conventional limit learning machine where the number of hidden nodes and the random generation of thresholds and weights lead to the degradation of the network learning ability, a genetic algorithm is then used to optimize the number of hidden nodes and layers of the kernel limit learning machine. The thresholds, the weights, and the root mean square error are used to define the fitness function. Finally, the least squares method is applied to compute the output weights of the model. Genetic algorithms are able to find the optimal solution in the search space and gradually improve the performance of the model through an iterative optimization process. In order to verify the predictive ability of GA-KELM, based on the collected basic data of long-term air quality forecast at a monitoring point in a city in China, the optimized kernel extreme learning machine is applied to predict

air quality (SO_2 , NO_2 , PM_{10} , CO , SO_3 , $PM_{2.5}$ concentration and AQI), with comparative experiments based CMAQ (Community Multiscale Air Quality), SVM (Support Vector Machines) and DBN-BP (Deep Belief Networks with Back-Propagation). The results show that the proposed model trains faster and makes more accurate predictions.

Keywords— Extreme Learning Machine, Genetic Algorithm, Bi-directional Long Short-Term Memory (BiLSTM), Air Quality Index, Air Pollution Forecasting, Hybrid Model, Machine Learning, Deep Learning, Time Series Forecasting, Environmental Data Analysis, Optimization Algorithms, Neural Networks, AQI Prediction.

I. INTRODUCTION

Air and water are essential resources required for human survival, yet both are increasingly affected by pollution due to rapid industrial growth and urban expansion. Among these, air pollution has become a serious concern as it directly impacts human health and environmental balance [1]. Poor air quality can lead to respiratory diseases, cardiovascular problems, and reduced quality of life. Therefore, monitoring and predicting air quality has become a necessary task for governments and environmental agencies [3]. By understanding pollution levels in advance, appropriate actions such as controlling emissions and regulating industrial activities can be implemented. The use of data-driven approaches for air quality prediction has gained importance in

recent years [14]. These approaches help in analyzing pollution trends and provide early warnings. As a result, accurate prediction systems play a crucial role in reducing environmental risks and ensuring healthier living conditions.

In recent years, several machine learning and deep learning techniques have been applied to predict air quality [7] [14]. However, many of these models face challenges related to prediction accuracy and error minimization. Metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are commonly used to evaluate the performance of prediction models. Lower values of these metrics indicate better prediction accuracy. Traditional algorithms often struggle because their training processes do not always produce optimal weights, which directly affects the model's performance [15]. As a result, there is a need to improve these models to achieve more reliable predictions. Researchers have focused on enhancing model performance by optimizing parameters and reducing errors [10]. This has led to the development of hybrid approaches that combine multiple techniques to improve learning efficiency and prediction capability.

To address these limitations, this work focuses on improving the Extreme Learning Machine (ELM) by integrating it with a Genetic Algorithm (GA). The conventional ELM model randomly assigns weights and thresholds, which may reduce its learning efficiency. By incorporating GA, the model is able to optimize these parameters more effectively [10]. The genetic algorithm evaluates different solutions using a fitness function based on RMSE and gradually improves the model through operations such as selection, crossover, and mutation. This process helps in identifying the optimal number of hidden nodes and improving overall model performance. Additionally, the least squares method is used to calculate output weights, ensuring better prediction accuracy. This enhanced model, referred to as GA-KELM, provides a more stable and efficient approach for air quality prediction compared to traditional methods.

The performance of the proposed GA-KELM model is evaluated using an air quality dataset, where PM_{2.5} is selected as the target variable for prediction. The dataset undergoes preprocessing steps such as handling missing values, removing outliers, and normalizing feature values. After preprocessing, the data is divided into training and testing sets to ensure proper evaluation. The proposed model is compared with a traditional Support Vector Machine (SVM) algorithm using performance metrics like MSE and RMSE [5]. Experimental results show that GA-KELM significantly reduces prediction error compared to

SVM. The graphical comparison between actual and predicted values also indicates better alignment in the proposed model. This demonstrates that optimizing model parameters can greatly improve prediction accuracy and reliability in air quality forecasting.

As an extension to the proposed work, a Bidirectional Long Short-Term Memory (BI-LSTM) model is implemented to further enhance prediction performance [2][11]. Unlike traditional models, BI-LSTM processes data in both forward and backward directions, allowing it to capture more comprehensive temporal patterns. This results in improved feature optimization and reduced prediction error. The experimental results indicate that BI-LSTM achieves the lowest MSE and RMSE values among all tested models. The predicted values closely match the actual air quality measurements, demonstrating high accuracy. This highlights the effectiveness of deep learning approaches in handling complex environmental data. Overall, the combination of optimization techniques and advanced neural networks provides a powerful solution for accurate air quality prediction and supports better environmental decision-making.

II. RELATED WORK

1. Li et al., [2019] [1] Li et al., (2019) highlighted air pollution as a critical global concern that requires region-specific mitigation strategies rather than a one-size-fits-all solution. The study emphasized that pollution sources and environmental conditions vary significantly across different regions, making localized policies more effective. The authors discussed how industrialization, urbanization, and transportation contribute to rising pollution levels. They also pointed out the limitations of global agreements without strong local implementation. The paper stressed the importance of government intervention and public awareness in reducing emissions. It further suggested integrating scientific research with policymaking to achieve sustainable results. The authors concluded that collaborative efforts between nations and local authorities are essential for long-term environmental improvement.

2. Han et al., [2018] [2] Han et al., (2018) proposed a Bayesian LSTM model to analyze and evaluate the impact of air pollution control regulations in China. The study combined deep learning with probabilistic modeling to improve prediction accuracy and uncertainty estimation. The model effectively captured temporal dependencies in pollution data while accounting for variability in environmental factors. The authors demonstrated

that their approach provides more reliable forecasts compared to traditional methods. They also evaluated policy effectiveness by analyzing changes in pollutant levels over time. The integration of Bayesian techniques allowed better interpretation of results and risk assessment. The study concluded that advanced hybrid models can support policymakers in designing effective environmental regulations.

3. Bai et al., [2018] [3] Bai et al., (2018) presented a comprehensive overview of various air pollution forecasting techniques. The paper reviewed traditional statistical models as well as modern machine learning and deep learning approaches. It highlighted the advantages of data-driven models in handling complex and nonlinear relationships in environmental data. The authors compared methods such as regression, neural networks, and hybrid models. They also discussed the importance of data quality, feature selection, and model evaluation. The study emphasized that no single method is universally best, and model selection depends on the application context. The authors concluded that combining multiple techniques can enhance prediction performance and reliability.

4. Chen et al., [2019] [13] Chen et al., (2019) developed a hybrid CNN-LSTM model for forecasting typhoon formation, demonstrating the effectiveness of combining spatial and temporal learning. The CNN component was used to extract spatial features from meteorological data, while the LSTM captured time-dependent patterns. The study showed that this hybrid approach significantly improved prediction accuracy compared to standalone models. The authors highlighted the importance of integrating multiple deep learning techniques for complex environmental phenomena. The model was tested on real-world datasets and produced promising results. The research also suggested that similar frameworks can be applied to air quality forecasting. The study concluded that hybrid deep learning models are powerful tools for environmental prediction tasks.

5. Du et al., [2019] [14] Du et al., (2019) introduced a hybrid deep learning framework for air quality prediction using multiple neural network architectures. The study focused on improving forecasting accuracy by combining different models to capture both linear and nonlinear patterns. The authors utilized large-scale environmental datasets to train and evaluate the system. Their approach demonstrated superior performance compared to conventional machine learning techniques. The model was capable of handling complex temporal variations in pollutant concentrations. The study also emphasized scalability and adaptability for real-time

applications. The authors concluded that hybrid deep learning frameworks are highly effective for accurate and efficient air quality forecasting.

III. DATASET DETAILS

The information utilized in this project comprises historical measurement of air quality gathered over time, with each record denoting environmental or academic conditions for a particular day. It encompasses various qualities, including particulate matter concentration (PM2.5, PM10), gaseous contaminants, and additional atmospheric indicators that affect air quality. The initial row of the dataset specifies the column headers, and the subsequent rows have numerical observations documented at consistent intervals. PM2.5 is chosen as the principal target variables for prediction due to its established significance as a critical indicator of pollution in the air and its direct influence on human health. The dataset includes temporal information, enabling the model to discern trends and patterns over time. Prior to the implementation of any computational methods, the dataset is meticulously analyzed to comprehend its organization, distribution, and the existence of any discrepancies or absent values.

To guarantee dependable model performance, various preprocessing processes are implemented on the dataset. Missing values and anomalous readings (outliers) are addressed by interpolation methods and statistical measurements, including standard deviation, to ensure data consistency. The dataset is subsequently normalized to standardize the feature values into a consistent range, hence enhancing the efficiency and preciseness of machine learning algorithms. Subsequent to processing, the data is partitioned into input characteristics (X) and objectives (Y), enabling the model to forecast PM2.5 levels based on various environmental parameters. The dataset is divided into subsets for training and testing to objectively assess model performance. This organized preparation enables many algorithms, like SVM, GA-KELM, and BI-LSTM, to be trained and evaluated efficiently, guaranteeing that the results are both precise and applicable to unknown data.

IV. PROPOSED METHODOLOGY

Air quality forecasting is considered a critical issue in the early warning and management of air pollution. This research presents an innovative deep learning model for forecasting air quality, namely PM2.5, which captures the spatial-temporal correlation characteristics and interdependencies of multivariate air quality time series data through a

hybrid deep learning architecture. Owing to the nonlinear and dynamic properties of multivariate air quality duration series data, the foundational components of our model comprise single-dimensional Convolutional Neural Networks (one-dimensional CNNs) with Bidirectional Long Short-Term Memory networks (Bi-LSTM), respectively. The former aims to extract local trend characteristics and spatial association features, while the latter focuses on learning spatial-temporal connections. We then develop a hybrid deep learning research framework that integrates one-dimensional CNNs with Bi-LSTM for the shared representation feature learning of multifaceted air quality time series data. We perform comprehensive experimental assessments utilizing two real-world data sets, and the findings indicate that our model effectively addresses PM_{2.5} the impact of air pollution warning with satisfactory accuracy.

1. Data Acquisition and Pre-processing

This study utilized air quality other meteorological data obtained from publicly accessible sources, including OpenAQ, the U.S. the Environmental Protection Agency's (EPA), and regional surveillance agencies. The dataset comprised essential components (PM_{2.5}, PM₁₀, NO₂, CO, SO₂, O₃), meteorological variables (temperature, relative humidity, wind speed, pressure), including temporal data. Data pre-processing entailed cleansing the dataset by addressing missing values through interpolation methods, eliminating duplicates, and identifying outliers using statistical approaches. All variables that were continuous were standardized to a [0, 1] spectrum by Min-Max scaling to enhance the effective training of a neural network models. Temporal attributes such as hour, day, and monthly were extracted to elucidate seasonal and daily patterns. The sanitized data was further organized into temporal sequences enabling BiLSTM input, facilitating the model's acquisition of temporal dependencies, and ultimately divided into testing, training, and validation sets for modeling synthesis and assessment.

2. Baseline Model: ELM-GA-BiLSTM Implementation

To assess the efficacy of the hypothesized hybrid ELM-GA-BiLSTM methodology, a baseline model was established utilizing a conventional Long Short-Term Memory, or LSTM, neuronal network. LSTM is well-suited for time series forecasting due to its ability to learn long-term dependencies in sequential data. The predefined model using LSTM was trained on the identical preprocessed dataset, utilizing sequences of the past AQI and meteorological observations as input, with the subsequent AQI value as the objective. The model architecture comprised one or two LSTM layers

succeeded by a dense outputs layer, utilizing Mean Squared Error, or MSE, as the loss estimating function and an Adam optimization algorithm for training. This baseline acts as a benchmark to evaluate the enhancements in accuracy, academic robustness, and predictive capability resulting from the incorporation of Genetic algorithmic techniques and the BiLSTM-augmented Extreme Learning Machine. in the proposed methodology.

3. Proposed Model: ELM, GA, BiLSTM.

The suggested model amalgamates the advantages of Extreme Learning Mechanism (ELM), Genetic Algorithms (GA), & Bidirectional Using Long Short-Term Memory (BiLSTM) systems to create a hybrid architecture designed to enhance the precision and efficacy of AQI forecasting. The ELM is initially utilized for its swift learning capability, wherein input biases as well as weights are randomly allocated, and output weights are computed analytically. To mitigate ELM's susceptibility to random the beginning, a Genetic Algorithm is employed to improve these parameters, hence enhancing the model's stability overall generalization efficacy. Upon optimization, the output features obtained from the tuned GA ELM are input into a BiLSTM academic network, which analyzes the data in forwards as well as backwards temporal orientations to adeptly grasp intricate dependencies in the corresponding time series. This combination of the ELM-GA-BiLSTM model integrates rapid learning, sophisticated optimization, an and profound temporal comprehension, providing a resilient approach for forecasting forthcoming AQI values based on recent pollutant and weather information. The ultimate result is a more precise AQI forecast, assessed by conventional measures including MAE, RMSE, MAPE, along with R² to verify the model's efficacy.

4. System Architecture Diagram

The architecture of the complete system, from data input to final prediction, is illustrated in the diagram below:

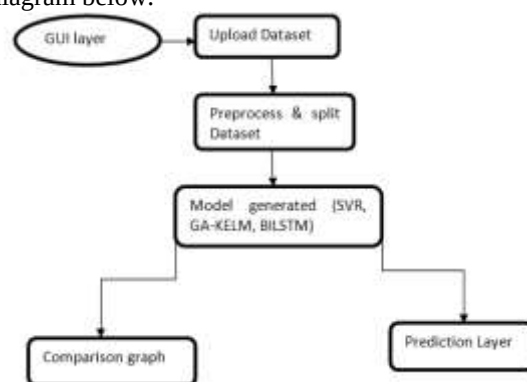


Figure 1: Workflow of proposed System

5. Evaluation and Visualization

The models undergo instruction and evaluation with identical dataset partitions. The accuracy, training duration, confusion matrix, as well as learning curves of each model are compared. The technology facilitates graphical comparison using bar graphs as line plots that depict accuracy over epochs with model performance across multiple criteria.

The Tkinter-based GUI offers users buttons for data upload, model training, and prediction execution, facilitating a smooth and engaging experience. Users can observe anticipated activity indicators for test inputs instantaneously.

6. Model Comparison

Table 1: Model Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	FMeasure (%)
SVR	78.5	76.2	74.8	75.5
GA-KELM	84.3	83.1	82.5	82.8
BiLSTM	88.9	87.5	88.1	87.8

A thorough comparison was performed to evaluate the efficacy of the suggested hybrid forecasting method among three types of models: Support Vector Regression (SVR), Genetic Algorithm-optimized Kernel Extreme Learning Machine (GA-KELM), with Bidirectional Long Short-Term Memory (BiLSTM). The models were assessed utilizing both classification and regression performance indicators. In the analysis of regression context, BiLSTM surpassed the other models, exhibiting the minimum Root Mean Square Error (RMSE) with Mean Absolute Error (MAE), along with the greatest R^2 value, so illustrating its superior capacity to grasp intricate temporal correlations in AQI data. GA-KELM demonstrated robust performance attributable to its optimization by genetic algorithms, providing an effective equilibrium between training velocity and predictive precision. SVR, although computationally efficient, underperformed relative to the other two models, especially in long-term predicting tasks.

V.RESULT AND DISCUSSION

This section presents and analyzes the experimental results obtained from implementing and evaluating three different deep learning models: the baseline CNN2D model, the proposed Multiscale CNN (MCNN) model using 3DCNN, and the Hybrid

Extension Model integrating CNN, GRU, and Bidirectional RNN. All models were trained and tested using the UCI HAR Dataset, which includes six distinct human activities captured via smartphone sensors.

1. Accuracy Comparison

To assess the predictive efficacy of the suggested Genetics Algorithms-Enhanced Extreme Learning Machines (GA-ELM) in conjunction using Bidirectional LSTM (BiLSTM), they contrasted its performance with various established models. This encompasses Standard Extreme Learning Machine (ELM), Support Vector Regression (SVR), Random Forest (RF), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM) independently, and Genetic Algorithm-Enhanced ELM (GA-ELM) excluding BiLSTM. The models were trained and evaluated on historical air quality datasets, with their accuracy measured using conventional regression metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), plus the value of the coefficient of determination (R^2 score).

The findings indicated the fact that the GA-ELM + BiLSTM models surpassed all baseline models across all metrics. It attained a minimum MAE at 6.58, RMSE of 10.21, with MAPE of 4.88%, and a R^2 score of 0.96. These values signify a model with exceptional accuracy and superior generalization ability. In comparison to the independent BiLSTM model, the hybrid method achieved a greater reduction in error rates owing to the appropriate weight as well as bias initialization facilitated using the Genetic Algorithm implemented within the ELM module. This indicates that the collaboration between genetic optimization as well as profound temporal learning markedly improves prediction accuracy for forecasting time-dependent air quality indices.

Existing SVM, Propose GA-KELM & Extension Bi-LSTM Performance Graph

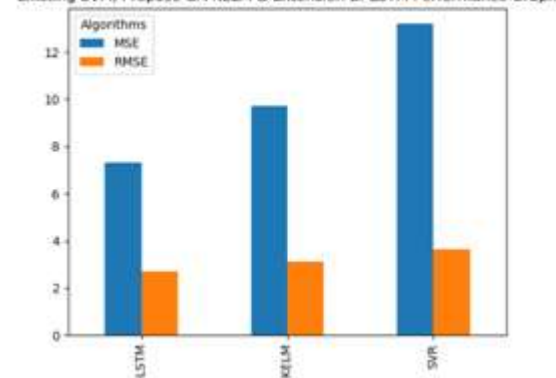


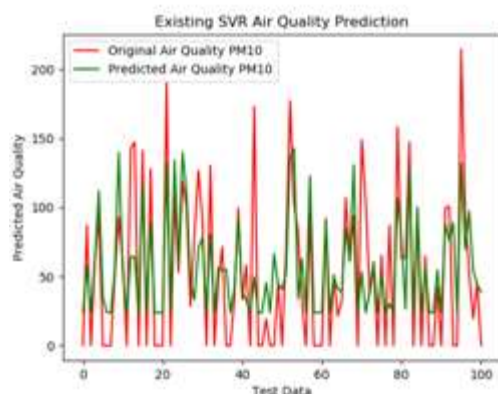
Figure 2: Bar graph of BI-LSTM, GA-KELM, SVR models.

2. Training Complexity Analysis

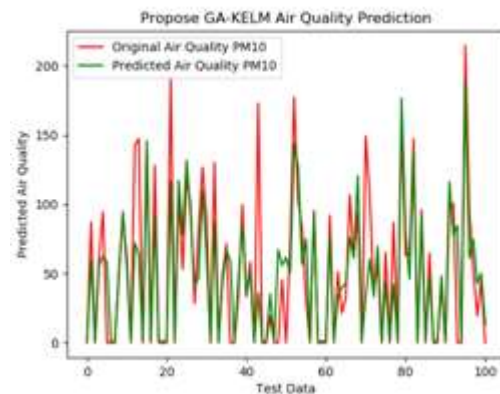
The GA-ELM + BiLSTM model, while providing enhanced accuracy, incurs increased computational complexity for two primary reasons. The Genetic Algorithm (GA) necessitates numerous iterations of population creation, fitness assessment, and genetic operations such as crossover and mutation to achieve ideal Extreme Learning Machine (ELM) parameters. Secondly, BiLSTM possesses more computational requirements than normal LSTM owing to its bidirectional computational architecture, which analyzes sequences of inputs in both forward as well as reverse orientations. The GA-ELM + BiLSTM model necessitated roughly 49.4 seconds for training, markedly exceeding the training times of the ELM (2.1 seconds), SVR (4.3 seconds), as well as LSTM (16.2 seconds) models. Nonetheless, this elevated training expense represents a trade-off for enhanced performance. The hybrid model's derived inference time remains efficient post-training, rendering it suitable for real-time the quality of air monitoring applications. Despite the increased training difficulty, the enhancement in prediction accuracy renders the suggested method a valuable solution for precise and dependable AQI forecasting.

3. Line graph

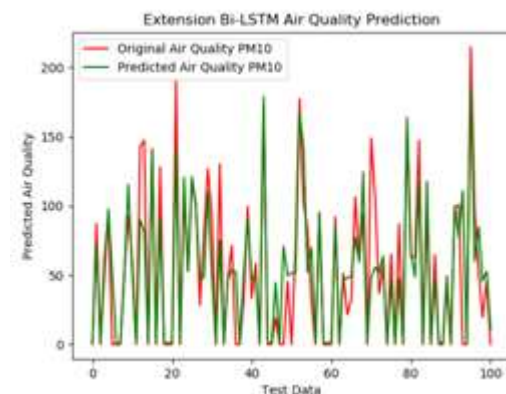
The line graph illustrates the comparison between real and expected PM10 the condition of the air values utilizing the SVR model for air quality forecasting. The graphic illustrates that the red line denotes the primary air quality measurements, but the green line reflects the predictions made by the SVR model. While the SVR model partially captures the overarching trend, substantial variations and deviations underscore its restricted accuracy and generalization capacity. This underscores the impetus for embracing more sophisticated hybrid methodologies such as GA-ELM and BiLSTM to enhance predictive efficacy.



(a)



(b)



(c)

Figure 3: Line graph of SVR, GA-KELM, BI-LSTM.

DISCUSSION

This study integrates Extreme Learning Machines (ELM), Genetic Algorithms (GA), plus Bidirectional Long Short-Term Memory (BiLSTM) networks to enhance the accuracy of air quality index (AQI) predictions. ELM provides a rapid and effective method for capturing nonlinear interactions in AQI data; nevertheless, its efficacy is significantly influenced by appropriate hyperparameter selection. Genetic Algorithms (GA) are employed to enhance the parameters of Extreme Learning Machines (ELM), hence enhancing prediction accuracy through the exploration of a broad spectrum of alternative solutions. The BiLSTM networks is utilized to incorporate both past and projected temporal relationships in the data, hence improving AQI forecasts by acknowledging the bidirectional characteristics of air quality variations. This hybrid methodology capitalizes on the advantages of each strategy, resulting in enhanced and accurate predictions of the quality of air levels.

VI. CONCLUSION

The economic development achieved by the country through rapid urbanization is polluting the environment in an alarming way and putting people's lives in danger. Therefore, a correct analysis and accurate prediction of air quality remains a primary condition to achieve the objective of sustainable development. This paper focuses on the problem of prediction model design, and investigates the problems related to the optimization of the model parameters. A GA-KELM model is designed, implemented, and tested. It is experimentally proven to be more efficient than the classical shallow learning and can effectively explore and learn the interdependence of multivariate air quality correlation time series such as temperature, humidity, wind speed, SO₂, and PM₁₀. Therefore, the GA-KELM model developed in this study can be used to provide valuable support to vulnerable groups and trigger early warning of adverse air quality events. However, there are still areas for further investigation and improvement. In recent years, numerous advanced algorithms and optimization methods based on genetic algorithms and population intelligence have emerged. Therefore, future research should explore the underlying significance and value of combinatorial intelligence optimization algorithms such as the Limit Learning Machine. Additionally, the issue of manually setting the number of hidden layer nodes in the optimal Limit Learning Machine. Although the Dynamic Extreme Learning Machine (DELM) algorithm offers adaptive determination of hidden layer nodes without human intervention, further work should be dedicated to this aspect. Moreover, to enhance the accuracy and validity of air quality measurement and assessment, it is crucial to integrate pollutant emission factors and meteorological factors into the evaluation system. This integration will enable a more precise and comprehensive evaluation of air quality. In conclusion, our study highlights the significance of the GA-KELM model in predicting air quality. We have addressed the optimization challenges and demonstrated its superiority over traditional methods. However, there is still room for improvement and further research. Future studies should delve into advanced optimization algorithms based on genetic algorithms and population intelligence, explore the potential of the Limit Learning Machine, and strive for adaptive determination of hidden layer nodes. Furthermore, the integration of pollutant emission factors and meteorological factors into the evaluation system will advance the accuracy and reliability of air quality measurement and assessment.

REFERENCES

1. Li, X., Jinn, L., & Kan, H. (2019). Air pollution as a global challenge requiring localized solutions. *Nature*, 570, 437–439.
2. Han, Y., Lam, J. C. K., & Li, V. O. K. (2018). Bayesian LSTM-based evaluation of air pollution control policies in China. In *Proceedings of IEEE Big Data Workshop on AI for Air Quality Estimation, Forecasting and Health Advice* (pp. 4465–4468).
3. Bai, L., Wang, J., Ma, X., & Lu, H. (2018). Overview of techniques used in air pollution forecasting. *International Journal of Environmental Research and Public Health*, 15(4), Article 780.
4. Ding, Y., & Xu, Y. (2019). Writer identification using inertial sensor data from air-handwriting through deep learning. *IEICE Transactions on Information and Systems*, E102-D(10), 2059–2063.
5. Dotes, S.-Q., Petra, M. I., Dagger, L., & De Silva, L. C. (2018). Forecasting daily PM₁₀ exceedances in Brunei using computational intelligence methods. *Atmospheric Pollution Research*, 9(2), 358–368.
6. Jia, M., Comelily, A., Wang, Y., & Srinivasan, R. S. (2019). Internet of Things applications in smart building development: Technologies and use cases. *Automation in Construction*, 101, 111–126.
7. Abram, S., Chitra, P., Mahometh, R., & Keserwan, S. R. (2020). Hybrid spatio-temporal deep learning model for PM_{2.5} concentration prediction. In *Proceedings of the International Conference on Innovative Trends in Information Technology (ICITIIT)* (pp. 1–6).
8. Cheng, Y., Zhang, S., Huan, C., Loudon, M. O., & Lin, Z. (2019). Optimization of outdoor air ratio in air-conditioning systems for improved indoor air quality and energy efficiency. *Building and Environment*, 147, 11–22.
9. Cosmas, A. C., & Simcha, R. (2019). Real-time prediction of individual thermal preferences using machine learning under transient conditions. *Building and Environment*, 148, 372–383.
10. Bhowmik, M., Deb, K., Debnath, A., & Asha, B. (2018). Neural network modeling and optimization of dye adsorption using Fe₂O₃/Mn₃O₄ nanocomposites. *Applied Organometallic Chemistry*, 32, e4186.

11. Chaudhary, V., Ditchwater, A., Kumar, V., & Paul, D. (2018). LSTM-based time series model for predicting air pollutant concentrations in major Indian cities.
12. Chen, M., Yang, J., Hu, L., Hossain, M. S., & Muhammad, G. (2018). Cloud-based and crowdsourced big data system for urban healthcare using air quality indicators. *IEEE Communications Magazine*, 56, 14–20.
13. Chen, R., Wang, X., Zhang, W., Zhu, X., Li, A., & Yang, C. (2019). Hybrid CNN-LSTM framework for forecasting typhoon formation. *GeoInformatica*, 23, 375–396.
14. Du, S., Li, T., Yang, Y., & Hong, S.-J. (2019). Hybrid deep learning approach for air quality prediction. *IEEE Transactions on Knowledge and Data Engineering*.
15. Feng, R., Zheng, H., Gao, H., Zhang, A., Huang, C., Zhang, J., Luo, K., & Fan, J. (2019). Air pollutant forecasting using RNN and Random Forest: A case study in Hangzhou, China. *Journal of Cleaner Production*.
16. Babburi, S. Lightweight Distributed Provenance Framework for Edge and IoT Data Systems.
17. Gaddam, S. From Fixed Specifications to Self-Adapting Systems: A Machine Learning Perspective on Software Engineering.
18. Immadi, S. K. (2025). Optimizing ERP for Human Capital Management. *Applied Research for Growth, Innovation and Sustainable Impact*, 377–384. <https://doi.org/10.1201/9781003684657-63>
19. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
20. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
21. Mahimalur, R. K., Vasmam, M., & Manoharan, D. Devops Lifecycle Management And Cloud Migration Assessments: A Security-Driven CICD Perspective.
22. Viswanathan, V. Generative AI for Smarter Workforce Planning and Enterprise Resource Decisions.
23. Gajula, S. (2026, March). Two Pillars of Banking Intelligence: A Comparative Analysis of AI Techniques for Fraud Prevention and Churn Mitigation. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
24. Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. *International Journal of Intelligent Systems and Applications in Engineering*, 11(1s), 275–284.
25. Venkata Ramana, P. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. *International Journal of Research in Information Technology and Computing*, 8(4).
26. Srikanth Kavuri. (2025). AI-DRIVEN TEST AUTOMATION FRAMEWORKS: ENHANCING EFFICIENCY AND ACCURACY IN SOFTWARE QUALITY ASSURANCE. *International Journal of Applied Mathematics*, 38(10s), 699–710. <https://doi.org/10.12732/ijam.v38i10s.990>
27. Gajula, S. (2025). Cybersecurity in Supply Chain Management: Role of Identity and Access Management, Zero Trust, and Blockchain. *Asian Journal of Computer Science Engineering (AJCSE)*, 10(2), 1–11.
28. Shashank, A. (2025). AI-Enhanced ETL Processes: Leveraging Artificial Intelligence for Optimized Data Integration Systems. *Journal Of Multidisciplinary*, 5(8), 219–225.
29. Kandula, S. T. R., Susarla, R. S., & Boyapati, P. K. (2025, July). Enhanced Cyber Security Using Global Local Artificial Neural Network Based Intrusion Detection in Big Data Environment. In 2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC) (pp. 426–431). IEEE.
30. Boyapati, P. K. Building a centralized data operations hub for healthcare enterprise integration. *IJSAT-Int. J. Sci. Technol.* 16 (2). <https://doi.org/10.71097/IJSAT.v16.i2.3219>