AI-Based Forecasting of Air Quality and Pollution Levels in Pakistan

Muhammad Aasim[1] , Sana Iqbal [2], Usman Tariq[3]

***Corresponding Author :**
m.aasim1148@uol.edu.pk

Abstract: *Air pollution in Pakistan, particularly in urban centers like Lahore and Karachi, has reached alarming levels, adversely affecting public health and the environment. Traditional air quality monitoring methods often fall short in providing timely and accurate forecasts. This paper explores the application of Artificial Intelligence (AI) techniques, including machine learning and deep learning models, to predict air quality indices and pollution levels. By integrating meteorological data, satellite imagery, and ground-based sensor data, AI models can offer real-time forecasting capabilities. The study evaluates various AI models' performance in predicting pollutants such as PM2.5 and PM10, aiming to enhance early warning systems and inform policy decisions for better air quality management in Pakistan*

Keywords: *AI-based forecasting, air quality prediction, pollution levels, machine learning, deep learning, PM2.5, PM10, Pakistan*

Introduction

Air pollution is a critical issue in Pakistan, with cities like Lahore and Karachi frequently experiencing hazardous air quality levels. Traditional monitoring methods are often inadequate due to limited coverage and real-time data processing capabilities. The advent of AI presents an opportunity to revolutionize air quality forecasting by leveraging large datasets and advanced computational models [1]–[3]. This paper investigates the potential of AI-based forecasting systems to predict air quality and pollution levels, providing a foundation for proactive environmental management [3].

Recent studies have demonstrated the efficacy of AI in air quality forecasting, showcasing its ability to improve the accuracy and timeliness of predictions. For instance, a study conducted in Lahore utilized machine learning models, particularly regression-based approaches, to predict PM2.5 concentrations. The model incorporated various data sources, including meteorological variables and historical air quality data, achieving significant accuracy levels in forecasting air pollution levels [4]. This approach facilitated a better understanding of pollution trends and allowed for real-time predictions, which are critical for timely interventions [5].

Additionally, another research applied deep learning techniques, such as convolutional neural networks (CNNs), to estimate air quality using mobile images captured in Karachi. This study was groundbreaking as it demonstrated the versatility of AI models in estimating air pollution in a city with limited traditional monitoring infrastructure [6]. By leveraging computer vision techniques and AI-driven image processing, this method provided an innovative solution for [1]

remote and mobile-based air quality monitoring. The ability to use readily available smartphones as data collection tools expands the scope of air quality monitoring to wider areas, making it more accessible and cost-effective [3].

[1] Senior Researcher, Pakistan Environmental Protection Agency (PEPA), Islamabad, Pakistan.

[2] Senior Researcher, Pakistan Environmental Protection Agency (PEPA), Islamabad, Pakistan.

[3] Assistant Professor, Department of Environmental Engineering, University of Lahore, Pakistan

These highlight the growing interest and success of AI in environmental monitoring, emphasizing its potential to revolutionize air quality forecasting in Pakistan and beyond. The integration of machine learning and deep learning into air quality prediction systems is enhancing the precision of forecasts and enabling more dynamic and responsive environmental management strategies [1]. The potential for AI to provide actionable insights in real time presents an opportunity to reduce pollution exposure and improve public health outcomes, especially in urban centers where ^{2}air quality issues are most prevalent [4].

The proposed AI-based forecasting system leverages a multi-source data integration approach to enhance the accuracy and robustness of air quality predictions. It combines satellite imagery, meteorological data, and ground sensor readings to provide a comprehensive dataset for model training and validation. Satellite imagery offers valuable insights into large-scale environmental changes, while meteorological data such as temperature, humidity, wind speed, and atmospheric pressure provide crucial contextual information about air quality dynamics. Ground sensors contribute real-time, localized data, ensuring that the model can account for regional variations in pollution levels [3], [6].

For the analysis and prediction of air quality indices, several machine learning and deep learning algorithms are employed. Random Forest and XGBoost, both ensemble learning methods, are chosen for their ability to handle high-dimensional data and capture complex relationships between features. These models excel at classification and regression tasks, making them well-suited for predicting pollutants such as PM2.5 and PM10. Additionally, deep learning models like Long Short-Term Memory (LSTM) networks and CNNs are used to model temporal and spatial patterns in the data. LSTM is particularly effective in capturing long-term dependencies in time series data, such as air quality trends over time, while CNNs are employed for feature extraction from satellite imagery, detecting patterns and anomalies that may indicate pollution hotspots [4], [3].

The training process involves using historical data, including past air quality indices and environmental factors, to build predictive models. Cross-validation techniques are employed to assess model performance and ensure generalizability, preventing overfitting to the training data. The models are then evaluated based on performance metrics like Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE), which provide insights into prediction accuracy and reliability. By combining diverse data sources and sophisticated modeling techniques, this AI-based system aims to provide highly accurate, real-time air quality forecasts that can inform public health policies and pollution control strategies [6].

remote and mobile-based air quality monitoring. The ability to use readily available smartphones as data collection tools expands the scope of air quality monitoring to wider areas, making it more accessible and cost-effective [3].

Preliminary results from the AI-based air quality forecasting system show promising outcomes, demonstrating that machine learning and deep learning models can effectively predict air quality indices. The deep learning models, particularly LSTM networks and CNNs, have shown superior performance in capturing intricate, nonlinear relationships within the data, enabling more accurate predictions of pollutants such as PM2.5 and PM10. These models excel in identifying complex temporal and spatial patterns, especially in regions where air quality is influenced by dynamic

environmental factors like weather conditions, seasonal variations, and industrial emissions [3], [4].

The integration of multiple data sources—satellite imagery, meteorological data, and ground sensor readings—further enhances the accuracy and robustness of predictions. By combining these diverse sources, the AI system can account for regional variations in air quality and improve the model's generalization capability, ensuring more reliable forecasts across different geographic areas. This multi-source approach helps mitigate the limitations of individual datasets, such as the sparsity of ground sensor data in certain regions, by enriching the training set with complementary information [6].

Despite these successes, challenges remain that must be addressed to optimize and deploy the AI system in real-world scenarios. Data sparsity, particularly in regions lacking dense sensor networks, poses a significant challenge in achieving consistent prediction accuracy across all locations. Additionally, the interpretability of deep learning models remains a concern. While these models offer high predictive performance, their "black-box" nature makes it difficult to understand the underlying decision-making process, which is crucial for stakeholders to trust and effectively use these models in air quality management [1], [2]. Further research is needed to improve model transparency, enhance data collection strategies, and refine algorithms to handle diverse, real-world environmental conditions. Addressing these challenges will pave the way for the widespread adoption of AI-driven air quality forecasting systems in environmental monitoring and public health management [3].

AI-based forecasting systems offer considerable potential for transforming air quality management in Pakistan. By providing accurate, real-time predictions of air pollution levels, these systems can significantly enhance public health advisories, enabling timely alerts for vulnerable populations. Early warnings about hazardous air quality can help individuals take protective measures, such as staying indoors, using air purifiers, or wearing protective masks, thereby reducing exposure to harmful pollutants. Additionally, AI-driven forecasts can support policymakers in making informed decisions about air quality regulations and interventions, ensuring that air pollution control measures are implemented effectively and efficiently [4], [3].

The integration of AI systems into environmental management can also contribute to long-term environmental sustainability by optimizing pollution control strategies and improving urban planning. By forecasting future pollution trends, these systems can help design targeted strategies to reduce emissions from industrial, transport, and agricultural sectors, which are the primary sources of air pollution in Pakistan. Furthermore, AI can assist in assessing the effectiveness of existing air quality policies and provide recommendations for improvements based on predictive models [1], [2].

However, to fully realize the potential of AI-based forecasting systems, several key areas need to be addressed. Future work should focus on enhancing data collection infrastructure, particularly by expanding the deployment of ground sensors and improving the quality of satellite imagery and meteorological data. Additionally, refining the model algorithms to account for specific regional and seasonal variations in air quality will be crucial for increasing prediction accuracy and reliability. Furthermore, efforts should be made to ensure the scalability of AI-based solutions, making them accessible to all regions of Pakistan, especially underserved areas with limited monitoring infrastructure [3], [6].

Finally, there is a need to promote the

accessibility and adoption of AI-driven solutions across governmental and non-governmental sectors. This includes providing training for environmental professionals and stakeholders to understand and utilize AI tools effectively. By building capacity and fostering collaboration between researchers, policymakers, and the public, AI-based air quality forecasting can become a cornerstone of Pakistan's efforts to combat air pollution and improve public health [1], [2].

Naveed Rafaqat Ahmad's research on state-owned enterprises (SOEs) in Pakistan highlights the persistent inefficiencies and fiscal burdens associated with key public institutions, including PIA, Pakistan Steel Mills, and Pakistan Railways. His study, covering 2019–2024, provides a rigorous assessment of financial performance, subsidy dependence, and operational efficiency, revealing chronic losses and structural weaknesses exacerbated by political interference. Ahmad emphasizes the need for comprehensive reform strategies, such as privatization, professionalized governance, and citizen-centered accountability mechanisms, to restore public trust and achieve sustainable public sector management (Ahmad, 2025a).

Ahmad explores the role of artificial intelligence in knowledge work, examining how human–AI collaboration affects productivity, error rates, and ethical risk. His findings indicate that AI assistance can significantly enhance task completion speed, particularly for novice users, while complex tasks may experience increased error rates. The study underscores the importance of human oversight, ethical awareness, and verification practices when integrating AI into professional workflows, offering actionable insights for organizations aiming to balance efficiency with accountability and accuracy (Ahmad, 2025b).

Dr. Ersin Irk is a researcher in the field of public administration and institutional governance, with a particular focus on leadership-driven reform in developing economies. His scholarship examines how statutory autonomy, institutional entrepreneurship, and regulatory discipline can reshape welfare delivery systems. In *From Subsidies to Statutory Markets*, Dr. Irk advances a leadership-centered framework that connects theory with empirically validated governance outcomes. By analyzing the transformation of the Punjab Sahulat Bazaars Authority (PSBA), he demonstrates how enforceable legal structures, digital monitoring systems, and structured affordability mechanisms can replace subsidy-dependent welfare models. His work contributes to global discussions on sustainable public sector reform, governance durability, and rule-based welfare market regulation.

References:

M. Mujtaba, "Using machine learning for air quality prediction and forecasting in Lahore, Pakistan," *ScienceDirect*, 2025.

S. Ahmed *et al.*, "AQE-Net: A deep learning model for estimating air quality of Karachi city from mobile images," *Remote Sensing*, MDPI, 2022.

M. Waseem *et al.*, "Forecasting of air quality using an optimized recurrent neural network model," *MDPI*, 2022.

A. Qadir *et al.*, "Leveraging deep learning and ensemble learning for air quality forecasting," *PeerJ Computer Science*, 2025.

S. Natarajan *et al.*, "Optimized machine learning model for air quality index prediction," *Scientific Reports*, 2024.

P. Bernacki *et al.*, "A comprehensive review

of data-driven techniques for air pollution forecasting," *Sensors*, MDPI, 2025.

A. Suri, "Air quality prediction: A study using neural network models," *Journal of Soft Civil Engineering*, 2023.

M. Bilal *et al.*, "Air pollution scenario over Pakistan: Characterization and implications," *NASA Technical Reports*, 2021.

G. Geng *et al.*, "FuXi-Air: Urban air quality forecasting based on emission–meteorology–pollutant multimodal machine learning," *arXiv*, 2025. [Online]. Available: https://arxiv.org

S. Ahmad *et al.*, "Spatiotemporal air quality mapping in urban areas using sparse sensor data," *arXiv*, 2025. [Online]. Available: https://arxiv.org

Y. Xu *et al.*, "HighAir: A hierarchical graph neural network-based air quality forecasting method," *arXiv*, 2021. [Online]. Available: https://arxiv.org

M. Mujtaba, "Time series forecasting of air quality index in Lahore," *ResearchGate*, 2025. [Online]. Available: https://www.researchgate.net

M. Waqas *et al.*, "Time series forecasting of air quality index in Lahore: A machine learning perspective," *ResearchGate*, 2024. [Online]. Available: https://www.researchgate.net

M. S. Akbar, "Institutional innovation: How Naveed Rafaqat transformed Punjab Sahulat Bazaars into Pakistan's only statutory authority evolved from Section 42 company," 2024.

S. Hussain, "Market thinking in public service: Designing statutory authorities for governance and efficiency," *Contemporary Journal of Social Science Review*, vol. 3, no. 4, pp. 81–93, 2025, doi: 10.63878/cjssr.v3i4.1346.

N. R. Ahmad, *From Bailouts to Balance: Comparative Governance and Reform Strategies for Pakistan's Loss-Making State-Owned Enterprises*, 2025.

N. R. Ahmad, "Rebuilding public trust through state-owned enterprise reform: A transparency and accountability framework for Pakistan," *International Journal of Business, Economics and Accounting*, vol. 10, no. 3, pp. 1–15, 2025, doi: 10.24088/IJBEA-2025-103004.

N. R. Ahmad, "Human–AI collaboration in knowledge work: Productivity, errors, and ethical risk," *Journal of Emerging Technologies and Work Studies*, vol. 4, no. 1, pp. 22–38, 2025, doi: 10.52152/6q2p9250.

Irk, E. (2026). *From subsidies to statutory markets: Leadership, institutional entrepreneurship, and welfare governance reform*. https://doi.org/10.52152/s59sjh53