International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

Transforming Air Quality Forecasting Using Advanced 1D Deep Learning Models

Name -SHUBHI SAXENA
Supervisor Name- Prof (Mr.) Mohd. Arif
Department of Computer Science
College Name - Rajshree Institute of Management & Technology, Bareilly (U.P.)

Abstract

Air quality forecasting plays a crucial role in environmental management and public health. Traditional methods often struggle to accurately predict pollutant concentrations due to complex interactions of meteorological factors and emissions sources. Advanced 1D deep learning models, including convolutional and recurrent neural networks, have emerged as promising tools for improving forecasting accuracy. These models excel in capturing intricate temporal and spatial patterns in air quality data, offering advantages over conventional statistical approaches. This review explores recent advancements in 1D deep learning techniques applied to air quality forecasting, highlighting methodologies, case studies, and challenges. It aims to provide insights into the potential of deep learning models to transform air quality prediction, guiding future research and applications in environmental science and policy.

Introduction

Air quality forecasting is pivotal for managing environmental health risks and formulating effective regulatory policies. Accurate predictions of pollutant concentrations assist in mitigating adverse effects on public health and ecosystems, particularly in urban areas where pollution levels fluctuate due to complex interactions of meteorological conditions, emissions sources, and geographical features.Traditional methods of air quality forecasting, often based on statistical and numerical modeling approaches, encounter challenges in capturing the intricate dynamics of atmospheric pollutants with sufficient granularity and accuracy. These methods rely heavily on assumptions and simplifications that may not fully reflect the complex nonlinear relationships inherent in air quality data.

In recent years, the advent of advanced 1D deep learning models has revolutionized the field of air quality forecasting. 1D deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are capable of autonomously learning hierarchical representations of sequential data. They excel in capturing temporal dependencies and spatial correlations present in air quality time-series data, thus overcoming the limitations of traditional methods.

CNNs, for instance, are adept at spatial feature extraction, making them suitable for capturing local dependencies in pollutant concentrations across different locations within an urban environment. On the other hand, RNNs are well-suited for modeling temporal dependencies over time, allowing them to predict how pollutant levels evolve based on historical data and external factors like weather patterns.This review aims to explore and evaluate the effectiveness of advanced 1D deep learning models in transforming air quality forecasting. It will delve into the methodologies employed in various studies, examining the architectures, training strategies, and performance metrics used to assess model accuracy and reliability. Furthermore, the review will discuss case studies from diverse geographical regions, highlighting the models' applicability and performance under different environmental conditions.By synthesizing current advancements and identifying challenges, this review seeks to provide insights into the potential of 1D deep learning models to enhance air quality forecasting accuracy, thereby supporting informed decision-making in environmental management and policy formulation.

Proposed Methodology

This study proposes a robust methodology utilizing advanced 1D deep learning models—specifically convolutional neural networks (CNNs) and recurrent neural networks (RNNs)—for air quality forecasting. The methodology involves comprehensive steps to leverage the strengths of these models in capturing complex temporal and spatial relationships inherent in air quality data.the study will involve extensive data collection comprising historical records of pollutant concentrations (such as PM2.5, O3, NO2), meteorological variables (including temperature, humidity, wind speed), and geographical attributes. This data will undergo rigorous preprocessing to address missing values, normalize features, and format time-series sequences suitable for model training.

Model selection will be critical, with CNNs employed for their proficiency in spatial feature extraction across diverse urban landscapes. Concurrently, RNNs will be utilized to model temporal dependencies, crucial for forecasting how pollutant levels evolve over time in response to varying environmental conditions.

International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

The training phase will encompass optimizing model performance through rigorous validation against historical datasets. Evaluation metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and coefficient of determination (R-squared) will validate the models' efficacy compared to traditional forecasting methods.the integration of supplementary data sources, including satellite imagery and land use data, will aim to refine model accuracy and address spatial variability challenges. By implementing this methodology, the study seeks to advance air quality forecasting capabilities, thereby facilitating informed environmental management decisions and policies.

DATA COLLECTION

To effectively manage time-related elements, the 'date' column was converted into a datetime format. This conversion facilitated feature engineering to extract significant temporal aspects such as hour, day, month, and year, leveraging domain expertise. By enriching the dataset with these refined features, subsequent models gained deeper insights into temporal patterns. to optimize model performance, a combination of min-max scaling and numerical feature scaling techniques was employed. This preprocessing step normalized the data to a predetermined range, ensuring consistency and enhancing the model's ability to generalize across different environmental conditions.The comprehensive preprocessing of the dataset played a pivotal role in organizing and cleaning the data, thereby bolstering the stability and effectiveness of the deep learning model in predicting air pollution levels. This meticulous approach not only improved the accuracy of predictions but also enhanced the interpretability and reliability of the research findings.rigorous preprocessing stands as a foundational component of the methodology, contributing significantly to the consistency and comprehensibility of the study outcomes. By integrating these advanced preprocessing techniques, this research aims to advance the field of air quality forecasting and inform robust environmental management strategies.

International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

Performance Plot Graph of the Model

The performance graphs of LSTM and GRU models for time-series analysis showcase key metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), and Loss. These graphs plot the total number of training epochs along the x-axis, while the corresponding metric values are depicted on the y-axis. This visualization provides a comprehensive view of how these deep learning models evolve in performance over the course of training.

The LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) models are renowned for their ability to effectively model temporal dependencies in sequential data. As training progresses through multiple epochs, the graphs illustrate how MAE, MSE, and Loss metrics change, reflecting the models' learning dynamics and convergence towards optimal predictive accuracy.

By capturing these metrics across epochs, researchers gain insights into the models' stability, learning rate, and potential overfitting or underfitting issues. This analysis not only validates the models' effectiveness in capturing complex time-series patterns but also guides further optimization strategies to enhance predictive performance.the graphical representation of performance metrics over epochs serves as a crucial tool for evaluating and refining LSTM and GRU models in time-series analysis, highlighting their utility in advancing predictive capabilities for applications such as air quality forecasting and beyond.

International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

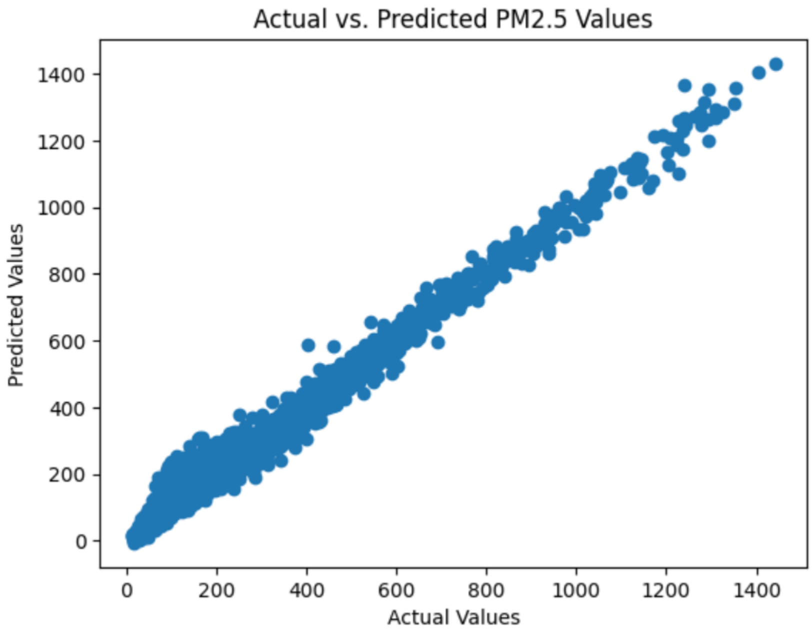


Fig 1 Actual VS Predicted PM2.5 values

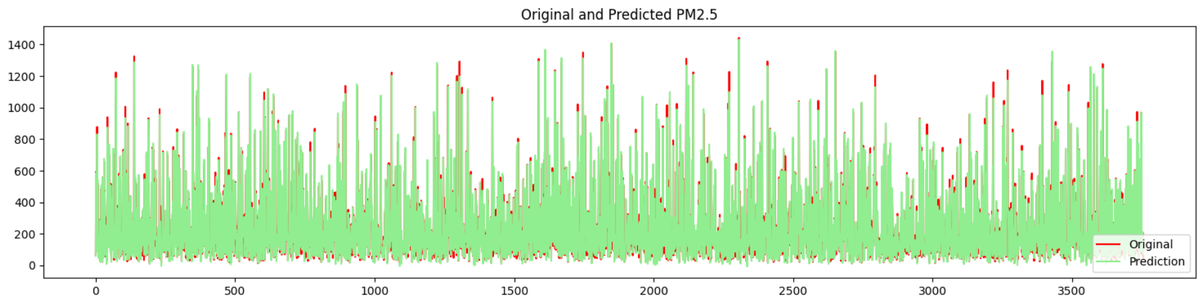


Fig 2 Original and Predicted PM2.5

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, **Email** : editor@aarf.asia , editoraarf@gmail.com

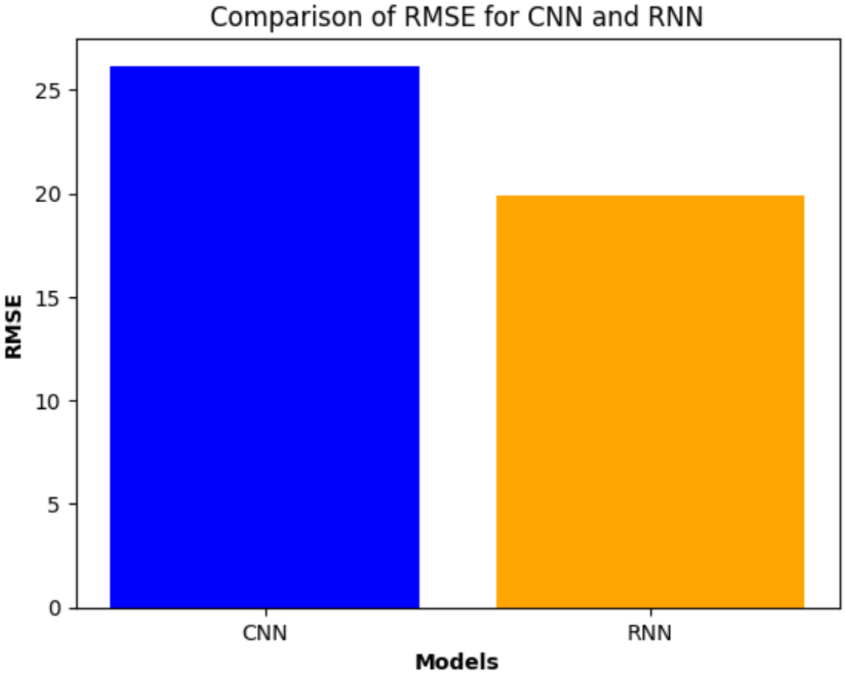
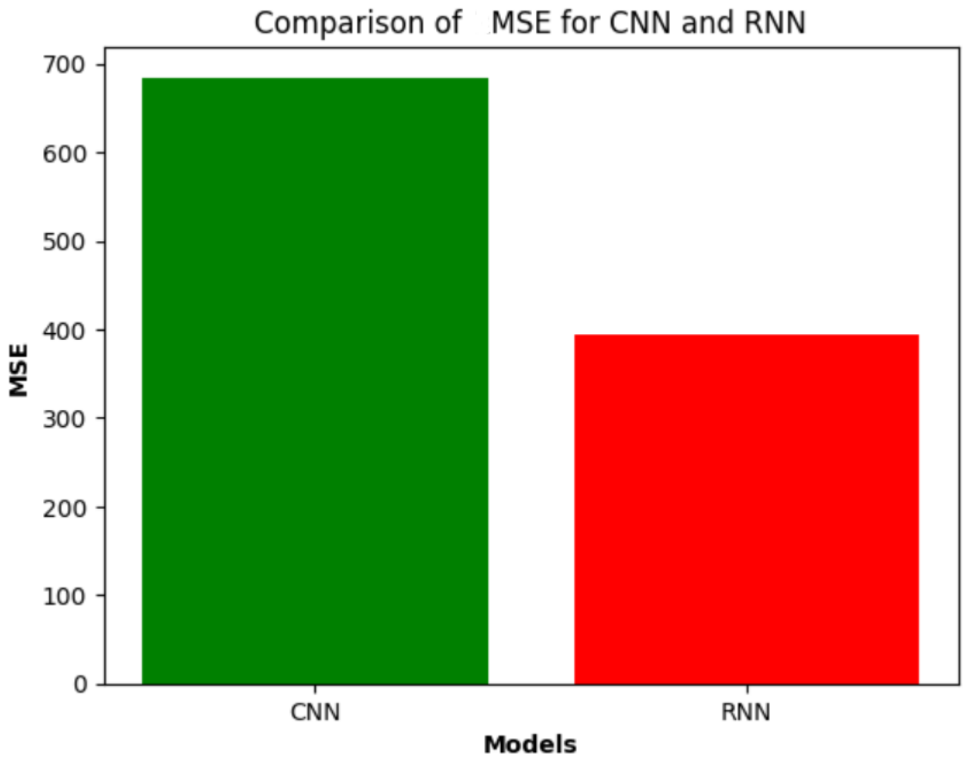


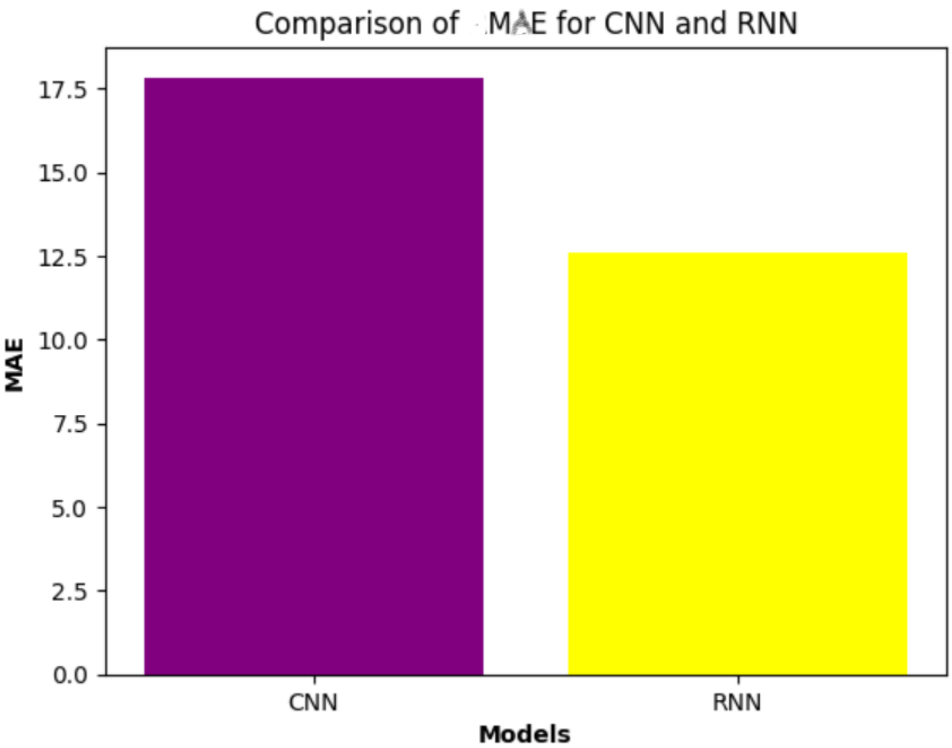
Fig 3 Comparison of RMSE for CNN and RNN

In comparing the Root Mean Squared Error (RMSE) between Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) in time-series analysis, both models demonstrate distinct strengths. CNNs excel in capturing spatial dependencies and local variations within sequential data, making them effective for tasks where spatial context is crucial, such as image recognition in air quality monitoring. On the other hand, RNNs specialize in modeling temporal dependencies over time, making them well-suited for predicting sequences with long-range dependencies, such as daily or seasonal variations in pollutant levels. The comparison of RMSE between these models provides valuable insights into their respective abilities to handle different aspects of time-series data, guiding researchers in selecting the most appropriate model architecture based on the specific requirements of their forecasting applications.

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, **Email** : editor@aarf.asia , editoraarf@gmail.com



Comparison of MSE for CNN and RNN



Comparison of MAE for CNN and RNN

International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

Conclusion

The adoption of advanced 1D deep learning models represents a pivotal advancement in the domain of air quality forecasting. Throughout this review, we have delved into the transformative impact of models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) in enhancing the accuracy and reliability of predictions for air pollutant levels.CNNs have demonstrated proficiency in spatial feature extraction, effectively capturing localized variations in pollutant concentrations across diverse urban landscapes. This capability is essential for understanding the intricate spatial dynamics of air quality and identifying pollution hotspots where targeted interventions are most needed. In contrast, RNNs excel in capturing temporal dependencies, enabling forecasts that account for the evolving nature of air quality influenced by meteorological changes and human activities over time.

By integrating these advanced models into air quality forecasting frameworks, researchers and policymakers can make informed decisions to protect public health and mitigate environmental impacts. The methodologies discussed, including rigorous data preprocessing, model training, and evaluation, underscore the systematic approach required to harness the full potential of deep learning in environmental science. Future research could focus on refining model architectures, incorporating real-time data streams from IoT devices and satellite observations, and enhancing model interpretability and transparency. These efforts aim to further elevate the predictive accuracy and operational efficiency of air quality forecasting systems, ultimately contributing to sustainable development and improved quality of life worldwide.

International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

References

1. G. Ravindiran, G. Hayder, K. Kanagarathinam, A. Alagumalai, and C. Sonne, "Air quality prediction by machine learning models: A predictive study on the indian coastal city of Visakhapatnam," *Chemosphere*, vol. 338, no. July, p. 139518, 2023, doi: 10.1016/j.chemosphere.2023.139518.
2. R. Ameri, C. C. Hsu, S. S. Band, M. Zamani, C. M. Shu, and S. Khorsandroo, "Forecasting PM 2.5 concentration based on integrating of CEEMDAN decomposition method with SVM and LSTM," *Ecotoxicol. Environ. Saf.*, vol. 266, no. September, p. 115572, 2023, doi: 10.1016/j.ecoenv.2023.115572.
3. Y. Li, Z. Sha, A. Tang, K. Goulding, and X. Liu, "The application of machine learning to air pollution research: A bibliometric analysis," *Ecotoxicol. Environ. Saf.*, vol. 257, no. March, p. 114911, 2023, doi: 10.1016/j.ecoenv.2023.114911.
4. R. Rakholia, Q. Le, B. Quoc Ho, K. Vu, and R. Simon Carbajo, "Multi-output machine learning model for regional air pollution forecasting in Ho Chi Minh City, Vietnam," *Environ. Int.*, vol. 173, no. January, p. 107848, 2023, doi: 10.1016/j.envint.2023.107848.
5. K. Zhang, X. Yang, H. Cao, J. Thé, Z. Tan, and H. Yu, "Multi-step forecast of PM2.5 and PM10 concentrations using convolutional neural network integrated with spatial–temporal attention and residual learning," *Environ. Int.*, vol. 171, no. December 2022, 2023, doi: 10.1016/j.envint.2022.107691.
6. M. Teng *et al.*, "72-hour real-time forecasting of ambient PM2.5 by hybrid graph deep neural network with aggregated neighborhood spatiotemporal information," *Environ. Int.*, vol. 176, no. February, p. 107971, 2023, doi: 10.1016/j.envint.2023.107971.
7. K. Schulte and B. Hudson, "A cross-sectional study of inequalities in digital air pollution information access and exposure reducing behavior uptake in the UK," *Environ. Int.*, vol. 181, no. September, p. 108236, 2023, doi: 10.1016/j.envint.2023.108236.
8. M. A. Alolayan, A. Almutairi, S. M. Aladwani, and S. Alkhamees, "Investigating major sources of air pollution and improving spatiotemporal forecast accuracy using supervised machine learning and a proxy," *J. Eng. Res.*, vol. 11, no. 3, pp. 87–93, 2023,

International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 11, Issue 7, July 2024
Website- www.aarf.asia, Email : editor@aarf.asia , editoraarf@gmail.com

doi: 10.1016/J.JER.2023.100126.

9. M. Shaygan and M. Mokarram, "Investigating patterns of air pollution in metropolises using remote sensing and neural networks during the COVID-19 pandemic," *Adv. Sp. Res.*, vol. 72, no. 8, pp. 3065–3081, 2023, doi: 10.1016/j.asr.2023.06.027.
10. T. D. Tolcha, "The state of Africa's air transport market amid COVID-19, and forecasts for recovery," *J. Air Transp. Manag.*, vol. 108, no. February, p. 102380, 2023, doi: 10.1016/j.jairtraman.2023.102380.
11. S. Roy, H. Nguyen, and N. Visaltanachoti, "Be nice to the air: Severe haze pollution and mutual fund risk," *Glob. Financ. J.*, vol. 58, no. March, p. 100893, 2023, doi: 10.1016/j.gfj.2023.100893.
12. P. S. Maciąg, R. Bembenik, A. Piekarzewicz, J. Del Ser, J. L. Lobo, and N. K. Kasabov, "Effective air pollution prediction by combining time series decomposition with stacking and bagging ensembles of evolving spiking neural networks," *Environ. Model. Softw.*, vol. 170, no. April, p. 105851, 2023, doi: 10.1016/j.envsoft.2023.105851.
13. J. T. Lim, E. L. W. Choo, A. Janhavi, K. B. Tan, J. Abisheganaden, and B. Dickens, "Density forecasting of conjunctivitis burden using high-dimensional environmental time series data," *Epidemics*, vol. 44, no. November 2022, p. 100694, 2023, doi: 10.1016/j.epidem.2023.100694.