Air Pollution Detection and Prediction: Leveraging Environmental Physics Using Deep Learning for Enhanced Air Quality Monitoring and Weather Forecasting

Abstract

Air quality is a significant concern impacting public health and economic well-being, as well as a barrier to regional growth and social advancement. The substantial rise in air pollution has led to the worldwide predominance of harmful diseases such as asthma, ischemic heart disease, and pulmonary disease. The government and other relevant organisations can take the required actions to protect the most vulnerable from exposure to toxic air due to their ability to predict air quality. The major aim of the proposed research work is to implement novel deep learning models for air quality monitoring and weather forecasting through air pollution detection and prediction. Initially, a comprehensive dataset encompassing pollution data and meteorological variables (temperature, humidity, wind speed, air pressure, and precipitation) will be gathered from various sources over a specified time. This will be followed by Exploratory Data Analysis (EDA) and data preprocessing. It will address feature normalization, missing values, and outliers. An appropriate feature selection method will be used to identify essential variables, and the dataset will be partitioned into training and testing subsets. A novel deep learning model will be designed for air quality monitoring and weather forecasting via air pollution detection and prediction. The proposed deep learning model, tailored for time-series forecasting, will utilize regularization methods to enhance generalization. The performance of the proposed model will be assessed using accuracy, precision, recall, and F1-score. The model will provide instantaneous predictions to aid environmental policy decisions and public health recommendations.

Keywords: *Environmental Physics, Air Quality Index, Air Pollutants, Weather Forecasting, Air Pollution Detection, Artificial Intelligence, Deep Learning.*

1. Introduction

The world is advancing towards modernization, introducing urbanity to both rural and urban areas, resulting in the highest levels of civilization. Even with advancements in many areas, air pollution remains a major threat to the welfare of society. Air pollution is a global environmental issue that affects people in all aspects of life [1]. Air pollution has become a recent problem as a result of the world's rapid urbanization and industrialization. The principal contributors to air pollution are emissions from vehicles and industries. Air pollution has been shown to hinder lung development in children, impair cognitive function, and contribute to increased mortality from respiratory diseases.

Air pollution results from both human activities and natural occurrences. “Air contaminants” are a group of compounds that alter the physical, chemical, or biological properties of the atmosphere. Contaminants are classified into two categories: primary and secondary. Primary pollutants are compounds derived from automobile emissions, including volcanic ash and carbon monoxide [2]. Solar radiation and heat can transform primary pollutants into secondary pollutants, including ozone (O) and other photochemical pollutants in the lower atmosphere, such as Nitrogen oxides (NO_x) and Sulphur oxides (SO_x).

Governments have invested significant resources in monitoring air pollution to mitigate its detrimental impacts on human health, the global environment and economy. A manual monitoring approach requires considerable human and material resources. Historically, traditional systems for monitoring air pollution have used stationary monitors to track air pollution levels. These monitoring stations are highly reliable and accurate, capable of evaluating a diverse array of pollutants using conventional analytical instruments, such as gas chromatograph-mass spectrometers (GC-MS). The limitations of conventional monitoring devices encompass their size, weight, and expense. Consequently, monitoring stations are distributed sparsely. The status of air pollution is assessed hourly or daily. Consequently, the spatial and temporal resolution of air pollution maps generated by conventional monitoring methods is significantly inadequate. This degree of spatiotemporal resolution is suitable for ambient air monitoring, but it does not reflect personal health concerns and is far too low for the general public to understand their exposure to air pollution.

The Air Quality Index (AQI), ranging from 0 to 500, has historically been considered one of the most dependable measures of air quality. Air pollution levels are assessed based on the AQI of a region. A high value increases the spectrum of contamination, which is harmful to health. An AQI of 40 is considered a safe level of air quality with negligible risk to overall health, whereas an AQI over 300 signifies hazardous air pollution [3]. Air pollutants, including CO , CH , CO, NO , O , and Particulate Matter (PM), are considered the six principal gases contributing to air pollution. This metric estimate relies on atmospheric data and so approximates the presence of airborne pollutants.

There is a substantial correlation between air pollution and meteorological factors. Meteorological parameters greatly influence air quality. Enhancement of air quality monitoring is accomplished by incorporating meteorological variables and significant air contaminants. The meteorological factors are Temperature (Temp), Solar Radiation (SR), Wind Speed (WS), Wind Direction (WD), Relative Humidity (RH), and Barometric Pressure (BP) [4]. Temperature increases can accelerate the pace of chemical reactions that produce secondary pollutants like O . Atmospheric temperature can be affected by sun radiation, influencing the stability and distribution of air contaminants. WS and WD are both critical in the dispersion and transportation of air contaminants. Enhanced wind shear can contribute to the dispersion and attenuation of atmospheric contaminants. The trajectory and endpoint of the contaminants are dictated by the wind direction. An elevation in RH can influence the generation of secondary pollutants, such as PM matter and O . Variations in BP can impact the vertical movement of air masses, consequently influencing the distribution and transport of contaminants in the atmosphere.

Artificial intelligence (AI) and Deep Learning (DL) are essential for air quality monitoring and weather forecasting, facilitating precise detection and prediction of air pollution levels [5]. AI can analyze enormous volumes of real-time data from sensors and satellites and use sophisticated models to spot pollutant trends and predict future air quality. These strategies improve the accuracy of predicting air conditions, pollution dispersion, and meteorological trends. The major aim of the proposed research work is to implement novel DL models for air quality monitoring and weather forecasting through air pollution detection and prediction.

2. Literature Review

Jia Xing and Joshua S. Fu [6] introduced Deep Chemical Transport Models (DeepCTM4D), a technique that emulated Community Multiscale Air Quality model (CMAQ) simulations for four-dimensional air pollution concentrations over space and time by combining DL with a 3D-ResNet and ConvLSTM. This method greatly improved the computational efficiency of atmospheric chemistry modelling. A CONUS domain application proved that DeepCTM4D can precisely replicate CMAQ simulations for six critical species, effectively capturing temporal, spatial, and vertical profile variations. These developments established DeepCTM4D as an effective instrument for operational forecasting systems that enhanced atmospheric chemistry predictions, weather feedback, and overall modelling efficiency and precision.

Valentino Petric *et al.* [7] aimed to enhance the generalization capabilities of Machine Learning (ML) models for predicting hourly air pollution concentrations in situations with restricted access to high-quality data. Various strategies were employed to address this challenge, including the use of the Prophet, Random Forest (RF), and three distinct DL architectures: Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Multilayer Perceptrons (MLP). A hybrid model combining RF and Prophet was also evaluated. Following evaluation, the hybrid model exhibited enhanced generalization abilities, attaining statistically significant advancements in R^2 for hourly concentrations of NO (26 % increase), NO_2 (an 18% enhancement), PM_{10} (fluctuating between an 8% decrease and a 35% increase), and O_3 (with R^2 coefficients between 0.83 and 0.87).

Olivia Bianchi and Herman Parkwood Putro [8] investigated the use of AI methodologies, including ML and DL, in environmental monitoring to enhance the precision of predictions about the impact of climate change and optimize management strategies. The major techniques consisted CNN for land cover categorization and LSTM models for predicting air quality levels. The findings demonstrated that AI significantly enhanced prediction accuracy, with CNN attaining superior performance in land classification and LSTM models providing dependable predictions for air quality changes.

Chengxin Zhang *et al.* [9] used geostationary satellite data to train a Neural Network (NN) model (GeoNet) that could predict the total amount of NO_2 on the ground in eastern China every four hours for the next 24 hours. GeoNet used a spatiotemporal series of satellite NO_2 measurements to explain the complex interrelationships among air quality, meteorology, and emissions across both temporal and spatial dimensions. The simulation results demonstrated that GeoNet effectively predicted diurnal fluctuations and spatial distribution of next-day NO_2 pollution, achieving a coefficient of determination of 0.68 and a Root Mean Square Error (RMSE) of 12.31 $\mu g\ m^{-3}$, considerably exceeding the predictions of conventional air quality models.

Young-Hee Ryu and Seung-Ki Min [10] introduced two methodologies that quantitatively differentiate the impacts of rainfall and pollutant concentrations on wet deposition: one utilised simplified equation to characterize the wet scavenging of pollutants, while the other employed RF models utilizing SHapley Additive explanations. Three-dimensional long-term air quality simulations from 2003 to 2019 served as inputs for both the physics-based and ML models. The outcomes derived from the explainable ML model matched with those from the physics-based methodology.

Raffaele De Palo *et al.* [11] employed eight Quartz-Enhanced Photoacoustic Spectroscopy (QEPAS) sensors to detect eight distinct air contaminants using a uniform acoustic detection module and interchangeable laser sources. Each analyte was addressed utilising identical sensor components and a specific laser source. Both interband cascade lasers and quantum cascade lasers were used to focus on a clear and separate absorption feature from each gas being studied. The wavelengths used ranged from 3.35 to 9.06 μm. The sensor was calibrated using certified concentrations of each gas species within a wet nitrogen matrix. The sensors responded linearly to the target gas concentrations for each gas species, and they could detect gases below their natural abundance levels.

N. Srinivasa Gupta *et al.* [12] evaluated the effectiveness of three prominent data mining models, Support Vector Regression (SVR), Random Forest Regression (RFR), and CatBoost Regression (CR), for accurately forecasting AQI data in several of India's most populous and polluted cities. The synthetic minority oversampling technique (SMOTE) was employed to balance the class data for improved and consistent outcomes. The simulation results demonstrated that RFR and CR yielded favorable outcomes.

A. Samad *et al.* [13] developed ML methodologies for predicting air pollution. This study included five ML methods: ridge regression, SVR, RF, extra trees regression, and extreme gradient boosting (XGB). The pollutants were modelled using meteorological parameters, transportation data, and pollution information obtained from local monitoring stations. The simulation results demonstrated that the pollution data from adjacent stations significantly influences the prediction of pollutant concentrations.

Rajnish Rakholia *et al.* [14] developed a multi-step, multi-output multivariate model for forecasting air quality, incorporating diverse parameters including meteorological conditions, air quality data from urban traffic, residential and industrial zones, urban spatial information, and temporal components for predicting hourly concentrations of NO , SO , O , and CO (1 h to 24 h). The air pollution time series datasets were collected from six healthy air quality monitoring locations. The suggested model's performance was evaluated using real data from healthy air stations and quantified by RMSE, Mean Absolute Percentage Error (MAPE), and correlation indices. The findings indicated that the global air quality forecasting model surpassed existing models designed for the prediction of individual pollutants in Ho Chi Minh City (HCMC).

Cong Cao *et al.* [15] examined the impact of particular climatic variables on air quality by comparing a regression model with K-Means Clustering, Hierarchical Clustering, and RF methods. The predictions of air pollution were analyzed using LSTM and Physics-based Deep Learning (PBDL). The investigation employed ten years of daily traffic, meteorological, and air quality data from three principal cities in Norway. The simulation results revealed that PBDL exhibited greater accuracy in air pollution predictions than LSTM.

Huynh A.D. Nguyen *et al.* [16] developed a DL architecture that demonstrated critical attributes for precise and effective air quality prediction. This study involved real-time data acquisition, automatic data preprocessing, and the ability to train several DL models adapted for diverse forecasting scenarios. Three DL models, LSTM-BNN, CNN-LSTM, and CNN-LSTM-BNN, were created to improve spatio-temporal pattern recognition and reduce prediction uncertainty. This web-based DL framework was developed to improve environmental monitoring and prediction, serving as a vital resource for both authorities and the general public.

Mulomba Mukendi Christian and Hyebong Choi [17] proposed a ML methodology for precise air quality prediction utilising two months of air quality data. The weather, air pollution, and air condition index characteristics of 197 capital cities were taken into account to predict the condition of the air in the following day using the World Weather Repository. The evaluation of multiple ML models showed how well the RF algorithm produced accurate predictions, especially when used for classification rather than regression. This method improved the model's generalisability by 42%, with cross-validation scores of 0.89 for classification and 0.38 for regression.

Ghazi Mauer Idroes *et al.* [18] utilised the CatBoost ML algorithm, recognized for its resilience against overfitting and proficiency in managing missing data, to predict urban air quality based on pollution concentrations. Following preprocessing, 80% of the data was allocated for training and 20% for testing. The simulation results demonstrated that the model exhibited elevated accuracy, precision, and recall. The results of the simulation demonstrated the main contaminants influencing the quality of the air in cities and underlined the necessity of focused measures to lower their concentrations and ensure a healthier and cleaner urban environment.

Xiaohu Wang *et al.* [19] introduced a spatiotemporal hybrid DL model, termed VMD–GAT–BiLSTM, which integrates Variational Mode Decomposition (VMD), Graph Attention Networks (GAT), and Bi-directional Long Short-Term Memory (BiLSTM) for the purpose of air quality forecasting. The suggested approach first used VMD to partition the original $PM_{2.5}$ data into a sequence of relatively stable subsequences, thereby reducing the impact of unknown variables on the model's predictive capabilities. For each subsequence, a GAT was subsequently developed to investigate profound spatial correlations among various monitoring stations. A BiLSTM was employed to capture the temporal characteristics of each decomposed sub-sequence. The prediction outcomes of each decomposed sub-sequence were consolidated and averaged to produce the final air quality predictions. According to the results of the simulation, the suggested model outperformed alternative approaches on both short-term and long-term air quality prediction tasks.

Tshepang Duncan Morapedi and Ibidun Christiana Obagbuwa [20] proposed several ML models to assess air pollution concerning time, cost, and efficiency, enabling diverse scenarios and systems to select the most suitable approach for their requirements. The ML models comprised CatBoost Regressor, Extreme Gradient Boosting Regressor, RF Classifier, Logistic Regression, SVM, K-Nearest Neighbours (KNN), and Decision Tree (DT). The results demonstrated that the proposed hybrid model exhibited greater accuracy than the individual models, confirming its superiority.

3. Research Gaps

The major limitations of the existing study include:

- Small errors grow over time in chaotic nature of the atmosphere, leading to significant biases during long-term forecasts with frequent spatial interactions.
- The quality and availability of environmental data can significantly affect the performance of AI models.
- The forecasting of pollutant concentration is not possible as the data from other monitoring stations is required for prediction.
- Models were previously unable to cover regional forecasts due to a lack of monitoring stations throughout the city.
- Lack of real-time wind direction data prevented analysis or consideration of air pollution from other regions, which affects model performance because each region

has a unique set of potential characteristics that are typically derived from geographic systems.

- Limitations in capturing long-term pollution patterns, resulting in their highly limited performance on air quality forecasting tasks.
- Difficulties in probing complex high-dimensional relationships from massive datasets.
- There is a lack of flexible multiscale frameworks, which prevents the traditional forecasting approaches to solve the curse of dimensionality.
- Time complexity when processing the massive volume of data.
- There are lots of sensor data quality issues which affect the accuracy of air quality evaluation and assessments due to device faults, battery issues, and sensor network communication problems.
- A lack of integrated real-time big data-based air quality evaluation and monitor environments for smart cities to support dynamic air quality evaluation, monitor, and supervision management.
- It is not able to determine the level of PM2.5 after when there is some change in atmospheric condition and it takes into account meteorological condition such as wind speed, temperature.

4. Motivation of the Study

Air pollution is a critical global issue that harms public health, ecosystems, and the environment. Rapid urbanization and industrialization have significantly increased the concentrations of air pollutants, including particulate matter ($PM_{2.5}$ and PM_{10}), NO , SO , CO, and O_3. Long-term exposure to these pollutants is associated with cancer, heart and lung conditions, and negative impacts on vulnerable groups, such as the elderly and children. Traditional systems for monitoring air quality depend on limited, stationary sensors, restricting their ability to detect geographical and temporal shifts in pollution levels. Moreover, these systems frequently lack the predictive ability to accurately anticipate pollution trends, limiting policymaker's ability to execute timely solutions. The integration of environmental physics with DL models offers a potential approach to improve the precision and efficacy of air quality monitoring and prediction. DL algorithms can be used to forecast pollution levels by analyzing complicated environmental patterns. This technique allows for preemptive measures to reduce environmental harm and health hazards. The major aim of this research is to use DL frameworks to overcome the limitations of traditional environmental monitoring techniques and advanced computational intelligence. DL can efficiently model the complex links between atmospheric conditions and pollution dynamics because of its ability to process enormous amounts of data. A more thorough comprehension of pollution dispersion, weather effects, and seasonal variations is made possible by the use of environmental physics concepts. This research work will use DL architectures to create a comprehensive air pollution detection and prediction system. These models will evaluate multi-source data, encompassing real-time sensor readings, meteorological variables, and historical air quality records, to deliver accurate and prompt predictions. In addition to improving public health protection by facilitating early warnings, such a system helps policymakers create evidence-based environmental laws and sustainable urban planning plans. The study will advance theoretical frameworks and practical applications for improving weather forecasting and air quality systems, adding to the expanding field of research on intelligent environmental monitoring.

5. Objectives of the Study

The major objectives of the proposed study include:

- To collect and integrate comprehensive environmental and meteorological data from various sources, encompassing critical pollutants ($PM_{2.5}$, PM , NO , SO , CO, O) and climatic variables (temperature, humidity, wind speed, atmospheric pressure, and precipitation), ensuring temporal synchronization for thorough analysis.
- To execute Exploratory Data Analysis (EDA) and apply advanced preprocessing methods to address outliers, manage missing values, and normalize features.
- To recognize and choose the most critical features affecting air quality and meteorological patterns via rigorous feature selection methodologies.
- To design and optimize novel DL models capable of performing air quality monitoring and weather forecasting.
- To deploy the trained model for real-time air quality and weather forecasting, continuously process incoming data and validate model outputs against ground-truth measurements and external datasets to ensure reliability, thereby supporting data-driven environmental policy decisions and public health recommendations.

6. Scope of the Study

The goal of this study is to generate and apply advanced DL models that are included with environmental physics concepts for the identification and prediction of air pollution.

It includes a thorough examination of principal air pollutants, such as particulate matter ($PM_{2.5}$ and PM_{10}), NO_x, SO_2, CO, and O_3. The method entails the aggregation and analysis of multi-source data from real-time air quality monitoring stations, meteorological databases, satellite imagery, and historical pollution records. The study will investigate the application of advanced DL architectures to model intricate interactions between environmental variables and pollutant concentrations. The study will also investigate how environmental physics, including atmospheric dispersion models, chemical transformations, and weather patterns, might improve the interpretability and accuracy of DL predictions. The geographic scope will concentrate on both urban and rural areas with a range of pollutant profiles to ensure the generalisability and flexibility of the model in various contexts. The research further examines the temporal and spatial dynamics of air pollution and their effects on human health and climate change. The study seeks to incorporate environmental physics into the modelling framework to accurately represent seasonal fluctuations, topography, and meteorological situations that profoundly affect pollution dynamics. This study will evaluate the accuracy, computational efficiency, and generalization potential of several DL models. The study enhances environmental protection and sustainable urban development by offering data-driven insights for politicians and environmental agencies, in addition to its public health implications. The development of precise prediction models can aid in formulating strategies for managing air quality, enhancing traffic rules, and regulating industrial emissions. This discovery has implications for climate change mitigation, as specific air pollutants contribute to global warming. The study offers dependable predictions of pollution trends, facilitating the formulation of adaptive policies that reduce environmental degradation and enhance urban air quality. The study promotes technological progress by investigating innovative DL methodologies and hybrid modelling strategies, thereby enhancing the field of AI in environmental science.

7. Proposed Methodology

The integration of environmental physics concepts and DL models is very significant for weather forecasting, air pollution detection, and prediction because of their complementing abilities to handle complicated environmental data. DL models excel at identifying non-linear correlations, temporal patterns, and spatial dependencies within extensive datasets, making them suitable for analyzing multi-source data from air quality sensors, meteorological records, and satellite imagery. Environmental physics offers a theoretical basis for comprehending pollution dispersion, atmospheric chemistry, and meteorological effects;

hence, it improves the interpretability and precision of these models. These methods can be combined to create advanced systems that can monitor in real time and provide accurate predictions. This allows for early intervention to reduce health concerns, boost urban sustainability, and improve environmental regulations. This integration of DL and environmental physics signifies a significant advancement in predictive modelling by providing accurate and scalable solutions for air quality management and weather prediction. The major aim of the proposed research work is to develop novel DL models for air pollution detection and prediction for enhanced air quality monitoring and weather forecasting. The detailed block diagram of proposed work is shown in Figure 1.

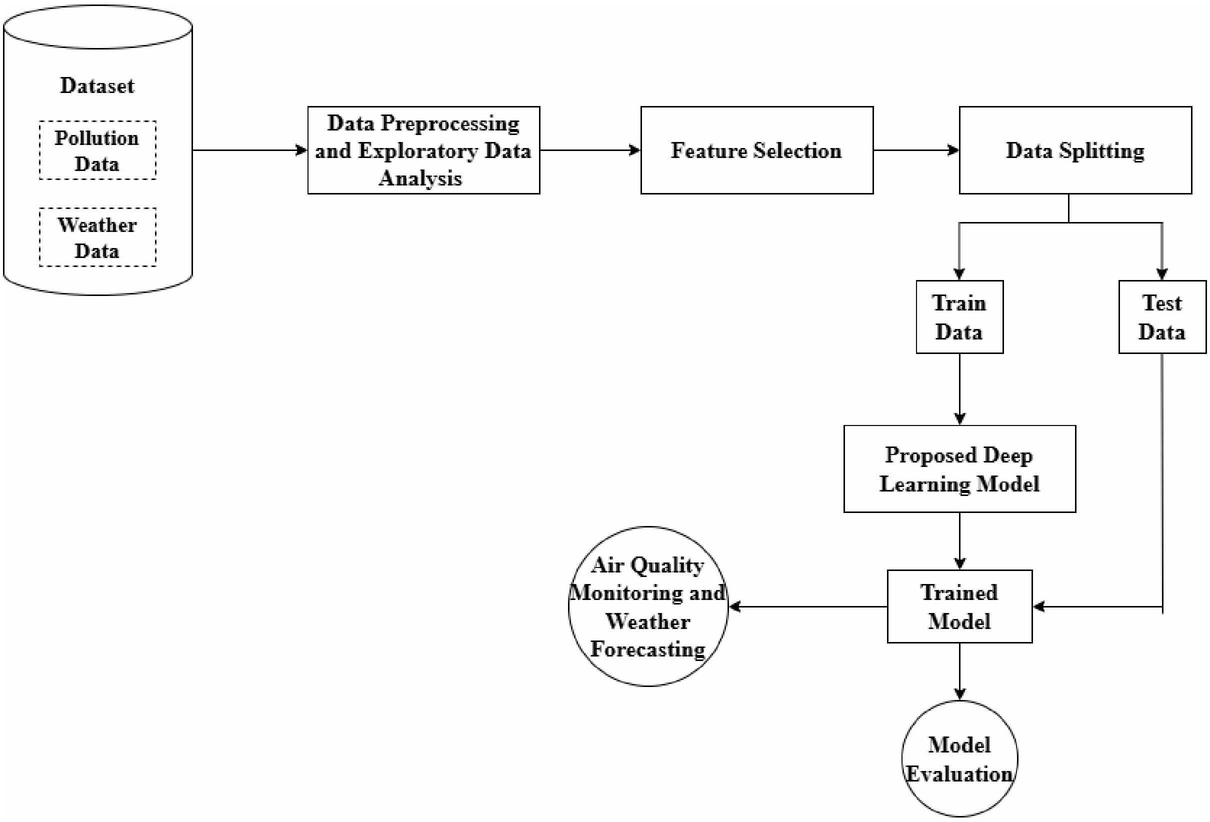

Fig. 1. Block Diagram of Proposed Deep Learning Model for Air Quality Monitoring and Weather Forecasting

In the initial phase, the study will collect an extensive dataset encompassing environmental and weather data from many sources. Pollution data will include measurements of particulate matter ($PM_{2.5}$ and PM_{10}), NO_2, SO_2, CO, and O , whereas weather data will consist of characteristics such as temperature, humidity, wind speed, atmospheric pressure, and precipitation levels. The data will be consolidated over a specified temporal range to ensure consistency between pollution and meteorological variables. Exploratory Data Analysis

(EDA) and extensive data preprocessing will be used to handle outliers, find missing values, and normalise features for increased model efficiency. In order to ensure that the dataset appropriately depicts the physical interactions between pollutants and weather patterns, advanced statistical methods from environmental physics will be used, including the computation of pollutant dispersion and correlation analysis between meteorological conditions and pollutant levels. After the preprocessing stage, feature selection will be carried out to determine which metrics are most important to weather and air quality prediction. The selected features will be split into training and testing data to enable model development and evaluation. The training dataset will be utilised to train the proposed DL model, whereas the testing dataset will assess the model's predictive efficacy, assuring robust generalization to unknown data.

The methodology will focus on developing an innovative DL model optimised for the dual tasks of air quality monitoring and weather forecasting. The train data will be used to train the model, which will use a loss function appropriate for time-series prediction and advanced optimisation techniques. Regularization techniques, such as batch normalization and dropout, will be applied to enhance generalization and avoid overfitting. The model will undergo a thorough evaluation after training to measure its performance on several criteria, such as accuracy, precision, recall, and F1-score. The trained model will subsequently be implemented for real-time air quality monitoring and weather forecasting. It will continually evaluate incoming data to produce precise predictions, supporting environmental policy choices and public health recommendations. Model outputs will be verified with ground-truth data and other datasets to ensure dependability.

8. Implementation Feasibility

The proposed model will be implemented on Google Collaboratory platform with Python language. Google Colaboratory (Colab) is a free, cloud-based platform. It is frequently used for DL tasks due to its free access to GPUs and TPUs, pre-installed frameworks like TensorFlow and PyTorch, and smooth connection with Google Drive for data storage.

References

1. Aguilar-Gomez, S., Dwyer, H., Graff Zivin, J., & Neidell, M. (2022). This is air: The “nonhealth” effects of air pollution. *Annual Review of Resource Economics*, *14*(1), 403-425.

2. Abaje, I. B., Bello, Y., & Ahmad, S. A. (2020). A review of air quality and concentrations of air pollutants in Nigeria. *Journal of Applied Sciences and Environmental Management*, *24*(2), 373-379.
3. *AQI Basics | AirNow.gov*. (n.d.). https://www.airnow.gov/aqi/aqi-basics/
4. Peng, L., Zhao, X., Tao, Y., Mi, S., Huang, J., & Zhang, Q. (2020). The effects of air pollution and meteorological factors on measles cases in Lanzhou, China. *Environmental Science and Pollution Research*, *27*(12), 13524-13533.
5. Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic markets*, *31*(3), 685-695.
6. Xing, J., & Fu, J. S. (2025). Efficient Weather and Air Quality Forecasts Through AI. Artificial Intelligence, Machine Learning, & Robotics in Business, 1(1), 52-61.
7. Petrić, V., Hussain, H., Časni, K., Vuckovic, M., Schopper, A., Andrijić, Ž. U., ... & Lovrić, M. (2024). Ensemble Machine Learning, Deep Learning, and Time Series Forecasting: Improving Prediction Accuracy for Hourly Concentrations of Ambient Air Pollutants. *Aerosol and Air Quality Research*, *24*(12), 230317.
8. Bianchi, O., & Putro, H. P. (2024). Artificial Intelligence in Environmental Monitoring: Predicting and Managing Climate Change Impacts. *International Transactions on Artificial Intelligence*, *3*(1), 85-96.
9. Zhang, C., Niu, X., Wu, H., Ding, Z., Chan, K. L., Kim, J., ... & Liu, C. (2025). Unleashing the potential of geostationary satellite observations in air quality forecasting through artificial intelligence techniques. *Atmospheric Chemistry and Physics*, *25*(2), 759-770.
10. Ryu, Y. H., & Min, S. K. (2024). Leveraging physics-based and explainable machine learning approaches to quantify the relative contributions of rain and air pollutants to wet deposition. *Science of the Total Environment*, *931*, 172980.
11. De Palo, R., Elefante, A., Biagi, G., Paciolla, F., Weih, R., Villada, V., ... & Patimisco, P. (2023). Quartz enhanced photoacoustic sensors for detection of eight air pollutants. *Advanced Photonics Research*, *4*(6), 2200353.
12. Gupta, N. S., Mohta, Y., Heda, K., Armaan, R., Valarmathi, B., & Arulkumaran, G. (2023). Prediction of air quality index using machine learning techniques: a comparative analysis. *Journal of Environmental and Public Health*, *2023*(1), 4916267.

13. Samad, A., Garuda, S., Vogt, U., & Yang, B. (2023). Air pollution prediction using machine learning techniques–an approach to replace existing monitoring stations with virtual monitoring stations. *Atmospheric Environment*, *310*, 119987.
14. Rakholia, R., Le, Q., Ho, B. Q., Vu, K., & Carbajo, R. S. (2023). Multi-output machine learning model for regional air pollution forecasting in Ho Chi Minh City, Vietnam. *Environment international*, *173*, 107848.
15. Cao, C., Debnath, R., & Alvarez, R. M. (2024). Physics-based deep learning reveals rising heating demand heightens air pollution in norwegian cities. *arXiv preprint arXiv:2405.04716*.
16. Nguyen, H. A., Le, H. T., Barthelemy, X., Azzi, M., Duc, H., Jiang, N., ... & Ha, Q. P. (2024). Deep-learning based visualization tool for air pollution forecast. *IEEE Software*.
17. Christian, M. M., & Choi, H. (2024). Air Quality Forecasting Using Machine Learning: A Global perspective with Relevance to Low-Resource Settings. *arXiv preprint arXiv:2401.04369*.
18. Idroes, G. M., Noviandy, T. R., Maulana, A., Zahriah, Z., Suhendrayatna, S., Suhartono, E., ... & Abd Rahman, S. (2023). Urban air quality classification using machine learning approach to enhance environmental monitoring. *Leuser Journal of Environmental Studies*, *1*(2), 62-68.
19. Wang, X., Zhang, S., Chen, Y., He, L., Ren, Y., Zhang, Z., ... & Zhang, S. (2024). Air quality forecasting using a spatiotemporal hybrid deep learning model based on VMD–GAT–BiLSTM. *Scientific Reports*, *14*(1), 17841.
20. Morapedi, T. D., & Obagbuwa, I. C. (2023). Air pollution particulate matter (PM2. 5) prediction in South African cities using machine learning techniques. *Frontiers in Artificial Intelligence*, *6*, 1230087.