Abstract:

Global concern has been raised by the high concentration of particulate matter ($PM_{2.5}$) in ambient air. $PM_{2.5}$ pollution is the suspended fine particles in ambient air that have a smaller diameter of 2.5 micrometer. These particles can come from different sources such as vehicular emission, industrial product, various construction activities, and natural sources like wildfires. $PM_{2.5}$ pollution is a major concern for public health, as these particles can lead to a range of health issues, particularly for people with pre-existing respiratory or cardiovascular conditions. India has reported higher concentration value of $PM_{2.5}$ in all its major cities. Therefore, research is going on all over the world to develop efficient forecasting tools for $PM_{2.5}$concentration prediction.

A subset of artificial intelligence and machine learning called Deep Learning has shown promise in addressing various environmental and pollution-related challenges, including $PM_{2.5}$ pollution prediction, monitoring, and mitigation. Here are a few ways where deep learning is widely used for dealing with $PM_{2.5}$ pollution:

$PM_{2.5}$ Prediction and Forecasting: Deep learning models can be trained on historical environmental data (such as meteorological variables, emission data, and geographical features) to predict and forecast $PM_{2.5}$ concentrations. Deep learning models give accurate results by identifying relationships and patterns in the input data.

Air Quality Monitoring: Deep learning can be used to study in situ data of air pollution monitoring stations, satellite imagery, and other sources to estimate real-time $PM_{2.5}$concentrations across different locations. This helps in identifying pollution hotspots and understanding the spatial distribution of pollution.

Health Impact Assessment: Deep learning techniques can correlate $PM_{2.5}$ concentrations with health outcomes, enabling researchers to better understand the specific health risks associated with different pollution levels. This information can assist in developing targeted public health interventions.

Source Apportionment: Deep learning can help identify the sources of $PM_{2.5}$ pollution by analyzing various data sources, such as chemical composition data and meteorological conditions. This information is crucial for implementing effective pollution control strategies.

Early Warning Systems: Deep learning models can be integrated into early warning systems that alert authorities and the public when $PM_{2.5}$ concentrations are expected to reach hazardous levels. This allows people to take preventive measures and reduce exposure.

Air Quality Index (AQI) Estimation: AQI is a standardized index used to communicate air quality levels to the public. Deep learning models can be trained to estimate AQI values based on various pollutants, including $PM_{2.5}$, providing a comprehensive assessment of air quality.

Policy and Decision Support: Deep learning can assist policymakers by providing insights into the effectiveness of pollution control measures and helping them make informed decisions to improve air quality.

The goal of the current research work is to create a hybrid deep learning model that can predict ambient $PM_{2.5}$ concentrations in various Indian cities with the least amount of error. The proposed model was built with the concept of sequence to sequence prediction methodology. The encoder-decoder-based hybrid model comprises of Bidirectional Long Short Term memory (BLSTM), Convolutional LSTM and Three Dimensional Convolutional Neural Network (3DCNN).

The model was applied in 26 cities covering 13 agroclimatic zones of India. In the model, previous 24 hour data was used as an input to predict the next 8 hour ahead values. The performance of the model output was also analyzed with the values of signal to noise ratio (SNR). We have found a distinct relation among model output and SNR values. Model performance degrades as the SNR values increases. The model shows little variations in terms of performance across different cities and found to be a stable one. The model could be used for long term prediction of ambient $PM_{2.5}$ concentration. A study on different encoder decoder based models was also carried out to find the best performance model.

Keywords: Encoder-decoder, Hybrid deep learning model, LSTM, Convolutional LSTM, , 3D convolution neural network, Bidirectional LSTM, SNR