International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 11 | November - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Deep Learning-Based Air Pollution Prediction in Smart Cities

Shalima Shajahan[1], Fousia M Shamsudeen[2]

[1,2] Department of Masters in Computer Applications, TKM College Of Engineering

---***---

Abstract - One of the most significant issues affecting urban living is air pollution, a form of environmental hazard. Particles having a diameter of under 2.5 µm (PM2.5) is one of the air contaminants that causes a variety of ailments. Therefore, to be able to shield populace using the potentially harmful effects of airborne toxins in advance, it is essential to precisely anticipate PM2.5 concentrations. It is now feasible to anticipate air quality in smart cities using deep learning techniques and their use in practically all facets of life. Using a deep learning solution, CNN Bi-LSTM, with a geographical time-based data, and data from the past on contaminant, weather, PM2.5 concentration at nearby stations, it is possible to predict the hourly prediction of air quality in India. Additionally, compared the performance differences across algorithms for deep learnings like LSTM, GRU, Bi-GRU, Bi-LSTM, CNN LSTM, and the proposed CNN-Bi-LSTM model. The experiment's findings indicate that the mean absolute error of CNN Bi-LSTM and the RMSE both experience drops of 0.20224 and 0.33256, respectively. CNN Bi-LSTM gives the greatest results for air pollution prediction when compared to BI-LSTM and LSTM, and its training time is cut by 12.5 and 50.9 cent, respectively. The primary conclusions of this study based on performance assessment and result comparison: the proposed model can successfully capture the spatial and time aspects of the data using CNN and Bi-LSTM, with more precisely and has stability

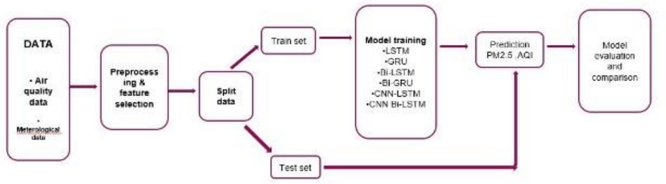
***Key Words*:** Air Pollution, Particulate Matter (PM2.5), Neural Network, Deep Learning**,** CNN Bi-LSTM

1.INTRODUCTION

PM2.5 is hazardous for the environment around the globe. Long term exposure by individuals to these tiny particles which can easily pass through deep into respiratory organs can cause a lot of impairments which can lead to mortality in most cases. Children and elderly citizens, sick people who are already victims of ailments like COPD, asthma, or other lung and heart disorders, pregnant ladies are easily at risk when the level of this concentration is high and predominates for a longer time. The citizens, especially from urban cities, has to track the air quality index to know the severity of PM2.5 and take cautious steps to protect themselves from the pollutant when it reaches the unhealthy level and it is important for any country to update, record and report the concentration level throughout the day at frequent intervals to assess the severity and alert the citizens. Air quality and weather have a close connection especially meteorological factors like air pressure, humidity, temperature, cloud coverage, wind speed, wind direction, and precipitation have huge influence while analyzing the air quality prediction. Evaluating accurately the harmful pollutants by using intelligent and technologically advanced deep learning algorithms which identify trends and patterns more accurately from complex datasets is a revolution and boom of today's world and can help any governing or regulatory bodies to regulate, monitor and account the pollutant levels. Therefore, obtaining air quality data is very important to take preventive measures, but the task of obtaining and analyzing this data remains a difficult task. To analyze the data with high efficiency and accuracy, production methods and techniques must be used to reach the desired results. The goal of this paper is to review articles related to air pollution prediction in smart cities using deep learning techniques. Deep learning as a machine learning methodology is considered a successful tool to predict air pollution as it has the capability to solve problems of big data. The primary objective of this paper is to develop a CNN Bi-LSTM model for Air pollution prediction and compare the performances of the several popular deep learning methods in the air pollution forecasting problem and finding the most accurately predicted model In the rest of the paper, section two contains a review of some research papers, section three contains basic definition of deep learning, sections four and five are discussion and conclusion.

2. METHODOLOGY

The input variables for this study comprise meteorological information, nearby stations, and pollution components throughout various time periods. After preparing the data by normalizing and standardizing the data, selecting features based on the connection between the features and PM2.5 concentration. Correlations between space and time are intricate and extensive. In our study, the model is built using past data from the target station and nearby stations that have been combined with additional characteristics. According to the findings, the suggested combination outperforms the competition in terms of PM2.5 and AQI prediction accuracy as well as effectiveness in extracting spatiotemporal variables. The spatial-temporal properties of the data are retrieved using the suggested model. Integrating the benefits of the CNN model's convolutional neural network, which is efficient at filtering.The system architectural design is used to abstract the overall outline of the software system and the relationships, constraints, and boundaries between components. It is an important tool as it provides an overall view of the physical deployment of the software system. This study uses several deep learning models for prediction of air quality data. The figure 1 below shows system architecture. The proposed system consist of three major phases: • Data preprocessing • Training and testing the data set • Evaluation of the model



Workflow for predicting pm2.5 and AQI concentration

2.1 DATASET AND METHODS

2.1.1.DATA COLLECTION: From several air quality monitoring sites in Delhi, we gathered hourly data from June 2020 to November 2021.Dataset chosen in this article consists of 12237 instances and 17 attributes. In this dataset, the air quality and pollutant concentrations at 47 locations are shown.. The presented method was implemented on two datasets collected from multiple resources. The statistics contain PM10, PM2.5, NO2, CO, O3 & SO2 as air pollutants. The data set also includes meteorological information from the same time period, such as weather conditions ,wind speed, moisture, temperature and pressure. The measurement for air pollution is g/m3. The most common weather kinds are bright, snowy, overcast, light and heavy rain, and blowing sand. Temperature is measured in degrees Celsius. Hectopascals is the name of the unit of atm pressure (hPa). Water vapor proportion in the air is referred to as humidity in percentage. Meters per second (m/s) is the measurement of wind speed. This data set mainly crawled from Meteostat and from OpenAQ The Weather's Record Keeper Meteostat is one of the largest vendors of open weather and climate data. The platform provides access to long-term time series of thousands of weather stations. *OpenAQ* is a non-profit organization empowering community around the globe to clean their air by harmonizing, sharing, and using *open* air quality data and also it is a real-time, interactive air quality monitoring website that enables visitors to find out about air quality monitoring and levels in their local area. Both data are collected based on the location ID, date and time.

2.1.2 PREPROCESSING

Filling The Missing Values: The 12237 entries in this collection have many attributes in each station. Date, PM2.5 and PM10 concentrations, Sulfur Dioxide, Nitrogen Dioxide, Carbon Monoxide, Ozone, Relative Humidity, Temp, Air Pressure, Cumulated Airspeed, Cumulated Hour of Snowfall, and Cumulated Hours of Rainfall make up the data. However, the equipment used to monitor the air quality and the weather will result in data leakage owing to equipment failure for various unavoidable causes. Data mining may be affected to some extent by the presence of such missing variables. The most often used methods for replacing missing field values in time-independent (non-chronological) data are the mean and median values. This is not the case, though, when it comes to time series. Many imputation approaches are used to tackle problems with incomplete data. According to a research, the optimum approach for estimating daily monitoring information for PM10 for each and every percent of generated incomplete data is linear interpolation.. The equation for linear interpolation is as follows:.

$$\mathbf{y - y_1 = \frac{(y_2 - y_1)\times(x - x_1)}{x_2 - x_1}} \quad \textbf{(1)}$$

The Data Set's Normalization: Since their measurement units differ, the value ranges of the various air quality data components also varied significantly. To remove the effects of various numerical ranges, the data should be normalised. Additionally, normalising the data collection can increase prediction accuracy and speed up training. Below is a display of the original sample data's normalising procedure. Where X is the actual data, min denotes the original data's least value, and max denotes the actual data's highest value. The distribution of each element in the processed data set will be between [0, 1]. The training data set and test data set are separated from the normalised data set in an 8:1 ratio.

$$\mathbf{x = \frac{x - min}{max - min}} \quad \textbf{(2)}$$

Selection of features: The process of feature selection, which may be carried out in a variety of methods, is crucial to machine learning applications. To determine the relationship between the dependent and independent variables, the majority of prior research relied on a mathematical relationship. Developing a connection between the goal real output and those attributes decreases the complexity of training and when multiple characteristics are added to the network for training, performance is improved. The Pearson correlation is the method that is most frequently used to establish the link between two variables. Its coefficient may be calculated using the following equation.

$$\gamma = \frac{\sum_{i=1}^{n}(x_1 - \pi)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - x)^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \textbf{(3)}$$

Air-quality attribute: We can identify a variety of contaminants in the atmosphere, and when their concentrations rise, the air's quality suffers. We examined the connections between the air quality characteristics and discovered a strong relationship between PM2.5, PM10, and CO. The 24-hour exposure's levels of health concern (i.e., hazardous, mild, acceptable, etc.) are determined by the AQI scale. The US-EPA 2016 standard is represented by the Air Quality Index scale in Table S.1. According to Wu et al. (2020), the equation below is used to translate concentrations into AQI:

$$\frac{I = (I_h - I_l)\times(C - C_l) + I_l}{C_h - C_l} \quad \textbf{(4)}$$

Metrological feature: Air quality is impacted by weather variables such as air pressure, temp, and humidity levels. For instance, heavy winds speed will lower PM2.5 concentrations, high humidity will usually make air pollution worse, and high air pressure will usually make the air quality better. Therefore, for the purpose of forecasting air quality, meteorological factors are crucial.

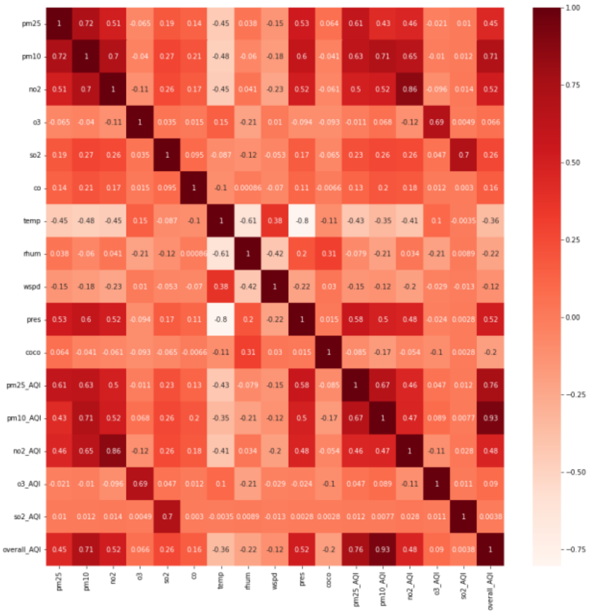


Fig 2: Spacial temporal correlation analysis

Time series data construction: In this experiment, the time of the input data is taken as a sequence, and the input data are segmented into two dimensions and constructed into three dimensions using a time series construction approach. This function takes in a sequence of data-points gathered at equal intervals, along with time series parameters such as length of the sequences/windows, spacing between two sequence/windows, etc., to produce batches of timeseries inputs and targets. Here use batch size as 4 and sequence length 12. In this experiment, training, validation, and test sets of the built time series data are created. For training, validation, and evaluation, the prediction model uses the first 47 data points from each layer as input and the AQI value from the 48th layer as output.

2.1.3 TRAINING AND TESTING THE MODEL

The time series data undergoes training with several deep learning models such us long shot term memory, , bidirectional gated recurrent unit, gated recurrent unit, bidirectional long short term memory, convolutional LSTM and convolutional bidirectional LSTM. All models are trained separately, and these models are saved. The proposed models are tested with testing data and predict PM2.5, AQI and evaluate the predicted vs actual values with evaluation metrics. The accuracy of proposed CNN BI-LSTM is much higher than other models in testing.

Two hidden layers are utilised in the GRU, LSTM, BI-GRU, and BI-LSTM networks, with two hundred units in the layer one and hundred on next . Network's final layers output is connected to a denc layer with just one end neuron. There is a 40% dropouts applied in between the layers. For the hidden layers , the LeakyReLU is used as activation function.

Each network in CNN-LSTM and CNN-Bi-LSTM contains a 1D CNN, which has a convolutional layer with 128 feature detectors sequentially and a convolution window length of 2 with causal padding. Afterwards, there is a MaxPooling layer of one dimensional. This is related to recurrent network employs a dense layer with one output neuron and a 50% dropout rate. Adam was a learning rate optimizer that we utilised (0,001)

CNN Bi-LSTM:A CNN layer and a Bi-LSTM layer are both components of the hybrid CNN-Bi-LSTM model. It can use CNN to extract useful spatial information from all PM2.5-related parameters. Bi-LSTM is used to identify temporal characteristics in two directions of the hidden layer and resolve gradient disappearance and explosion issues in the context of time series. The suggested deep learning network primarily included embedding, 1D convolution, pooling, bilateral LTSM, dropout, flatten, and fully connected layers. Two noteworthy aspects of the convolutional neural network (CNN), a feed-forward network, are local connection and parameter sharing. Local connection is the state of partial connectedness between two adjacent layers. The filter is used by CNN to implement parameter sharing (also called convolution kernel). The filter moves over the data and becomes tangled up with every piece of the data. Here, 1D convolution was utilized. Kernel mobility in a 1D CNN is unidirectional. Data for a 1D CNN's input and output are in two dimensions used often with time-series data.

The optimum option when the input matrix is a 2-D or 3-D tensor, such as an image or a video, is to employ the conventional CNN design. An updated variation Bi-long short term momory called Conv bidirectional LSTM has been developed since the framework of Bi-LSTM is better suited for 1-D Data. This architecture keeps a cell with a convolutional actions and an input feature parametres in the final layer as opposed to a simple one - dimensional vectors. At each gate of a traditional LSTM, a convolutional operations takes the role of matrix multiplication. Conv Bi-LSTM framework combines CNN and Bi-LSTM Network features, so to speak. Combining CNN and LSTM layers one block at a time is an additional strategy for handling spatiotemporal data. Convolutional Bidirectional LSTM is the name of this design Convolutional layers extract key characteristics from the input data in the first portion of this model, and the outputs fattened in one dimensional can be utilised as information for the model's second component (long short term memory). Atlast, information must be moulded into the actual form of input information before travelling through final concealed layer.A different type of neural net framework than multiple layer perception is the recurrent neural network. The RNN shares weights and is especially well suited for sequence analysis tasks like linguistic translation and semantic interpretation. The previous hidden state, which was calculated by where was an activation function and was a bias, as well as the current input were both factors affecting the hidden state at the time interval. Where was a bias, an activation function, it was used to compute the output at the time interval. A catastrophic problem with the RNN, known as vanishing gradient, existed for lengthy sequences. The LSTM is one of the better methods for addressing the vanishing gradient.

Two convolutional layers make up the one-dimensional (1D) CNN found in the CNN Bi-LSTM framework. A series of filter (or kernels) make up the convolution layers, and its parameters must be learnt throughout training. The filters are often smaller in size than the input images. The activation map is produced by each filter's convolution with the image. It is followed by a maximum pooling layer after background subtraction. The feature maps' sizes are decreased by pooling layers. As a result, it decreases the number of elements to learn and the amount of network processing required. A flatten layer is used to combine all of the 2-Dimensional arrays from the pooled feature maps into a single, lengthy continuous linear vector.Two hidden layers with opposing orientations are linked to the same outcome via bidirectional RNN (BRNN). This kind of generative deep learning allows the output layer to concurrently receive data from previous (backwards) and future (ahead) states. then a thick layer made up of only one output neuron comes next. Each neurons in the dense layer of a neural network receives the input from all the neurons in its preceding layer due to the layer's dense connections. The neuron in the dense layer of a deep-connected neural network receives input from all the neurons in the layer below it.

3.RESULT AND DISCUSSIONS

The prediction of PM2.5 concentrations was simulated in this work using multiple deep-learning methods. The historical observation PM2.5 data are compared with the calculated PM2.5 from long shot term memory, gated recurrent unit, bidirectional long short term memory, bidirectional gated recurrent unit, convolutional long short term memory and convolutional bidirectional long short term memory evaluated in one lag day in this part.

Each network makes an attempt to anticipate the outcomes as precisely as possible. In this study, we used Mean Squared Error as a cost function for all networks. For neural networks, an epoch is the number of iteration steps; during each epoch, the network simulates the streamflow time series once. Recurrent networks allow for the arbitrary selection of neurons or network layers. All of the recurrent network models in our study have the same architecture in order to compare them with one another.

Two hidden layers are utilised in the GRU, LSTM, BI-GRU, and BI-LSTM networks, with two hundred units in the layer one and hundred on next . One output neuron-only dense layer is linked to the output of the network's final layer. There is a 40% dropouts applied in between the layers. For the hidden layers , the LeakyReLU is used as activation function.Each network in CNN-LSTM and CNN-Bi-LSTM contains a 1D CNN, which has a convolutional layer with 128 feature detectors sequentially and a convolution window length of 2 with causal padding. Afterwards, there is a MaxPooling layer of one dimensional. This is related to recurrent network ,it employs a thick layer with an output data neuron and applies a drop out of 50% between layers.. Adam was a learning rate optimizer that we utilised (0,001)

Evaluation Index Of The Models

In this article, three measures are used to assess the model's efficiency: mean absolute error , mean squared error , and coefficient of determination .

MAE: The true prediction error scenario can be better described by the arithmetic mean of the absolute variances among the reality value and the predicted data of all samples, known as Mean Absolute Error.The following is the calculating formula:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (5)$$

RMSE: The square-root of the mean square of all mistakes is known as RMSE (Root Mean Square Error). It could accurately reflect the forecast error's accuracy. The following is the calculating formula:

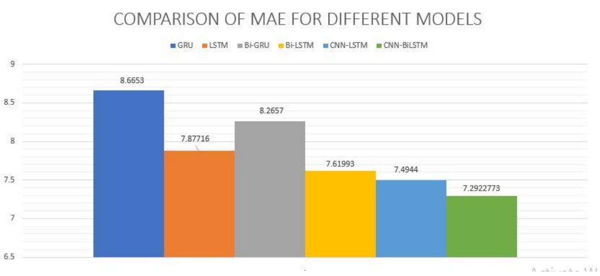
$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}(y_i - \hat{y}_i)^2} \qquad (6)$$

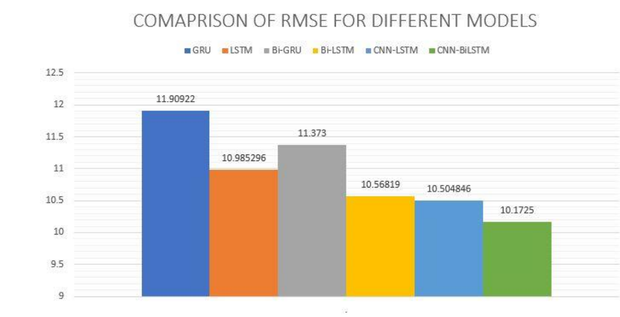
MSE: An estimator's mean squared error calculates the sum of the squared difference between the predicted values and the actual value (MSE). MSE, which corresponds to the expected value of the squared error loss, is a risk function. The following is the calculating formula:

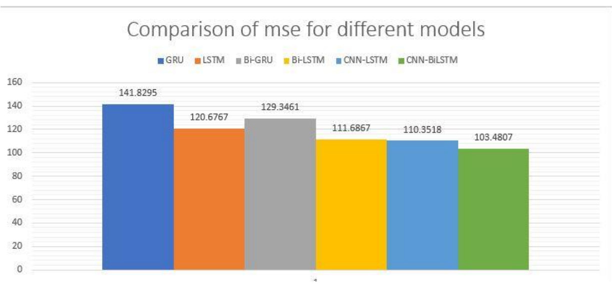
$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y - \hat{y}_i)^2 \qquad (7)$$

3.1 TRAINING AND VALIDATION RESULTS

Model Comparison : The MAE, RMSE, and MSE for six models are shown in Figures 3,4 and 5. These numbers fall within the range of PM2.5 concentrations that are expected and actual.







The MAE, RMSE, and MSE values produced by the model prediction models for the concentration of PM2.5 in the air are listed in Table 1. The CNN BILSTM models' RMSE values were the lowest of all the models. The CNN Bi-LSTM fared better than other models in comparison to the identical settings, according to the results.

The MAE,MSE and RMSE of BI-LSTM are reduced when LSTM is contrasted with Bi-LSTM. Therefore, BI-LSTM was superior than LSTM: MAE decreased from 7.87716 to 7.61993, MSE from 120.6767 to 111.68669, and RMSE decreased from 10.9853 to 10.5682.

• Based on assessment criteria, we may conclude that CNN LSTM outperforms BI-LSTM since MAE lowers from 7.61993 to 7.49447, MSE drops from 111.68669 to 110.3518, and RMSE drops from 10.5682 to 10.5049.

• The suggested model in this research has the least MAE, RMSE, and MSE values when compared to CNN LSTM and CNN BI-LSTM. the MSE falls from 110.3518 to 103.4807, the MAE from 7.4945 to 7.2923, and the RMSE from 10.50485 to 10.1725.

Overall, Table 1's data demonstrate that among the six models, CNN Bi-LSTM performs the best. The MAE is 7.2923 and the RMSE is 10.1725, which is the least of compared models and has the best predicting respectively.

The comparison between the predicted findings and the actual values for PM2.5 data is presented below based on our suggested prediction approach and another comparable model. The numbers show that the suggested strategy is able to produce prediction results that are much closer to the actual values than other models.

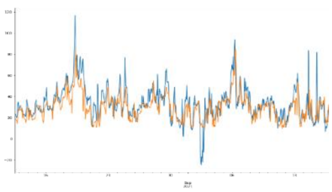


Fig 6: comparison between predicted and actual values in lstm

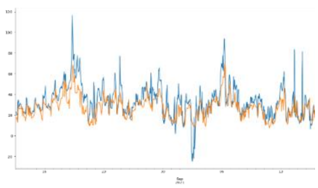


Fig 7:comparison between predicted and actual values in gru

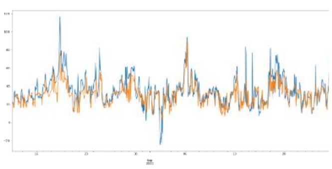


fig 8: comparison between predicted and actual values in bi-lstm

fig 9: comparison between predicted and actual values in bi-gru

MODELS	MAE	MSE	RMSE
GRU	**8.66531**	**141.8295**	**11.90922**
LSTM	**7.87716**	**120.6767**	**10.985296**
BI-GRU	**8.26573**	**129.3461**	**11.3730**
BI-LSTM	**7.61993**	**111.68669**	**10.56819**
CNN LSTM	**7.49447**	**110.35179**	**10.504846**
CNN BILSTM	**7.29228**	**103.480705**	**10.1725**

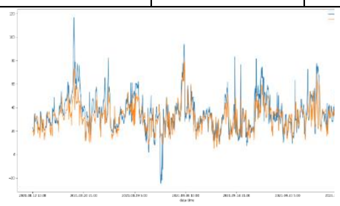


fig 10: comparison between predicted and actual values in cnn lstm

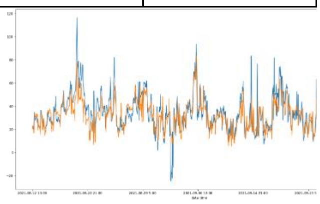


fig 11: comparison between predicted and actual values in cnn bi-lstm

Model Convergence.

Checking if CNN-Bi-LSTM often converges during training is important once the model has been constructed and its parameters have been established. In this experiment, the 1-0 loss function is used to demonstrate its convergence when all parameters are the same and the epoch is 400. The convergences of the CNN-Bi-LSTM in this experiment are depicted in Fig. below. CNN-Bi-convergence LSTM's speed in this experiment is quicker than CNN-due LSTM's to CNN-BI-loss LSTM's function being smaller than CNN- LSTM's.

MODELS	LOSS	VALIDATION LOSS	TIME PER EPOCH (S)
CNN BI-LSTM	**0.1430**	**0.1092**	**36 s**
CNN LSTM	**0.1551**	**0.1228**	**46 s**
LSTM	**0.1853**	**0.1260**	**42 s**
GRU	**0.1808**	**0.1267**	**42 s**
BI-LSTM	**0.1501**	**0.1123**	**80 s**
BI-GRU	**0.1598**	**0.1170**	**82 s**

Table 2: Training loss and validation loss in deep learning models

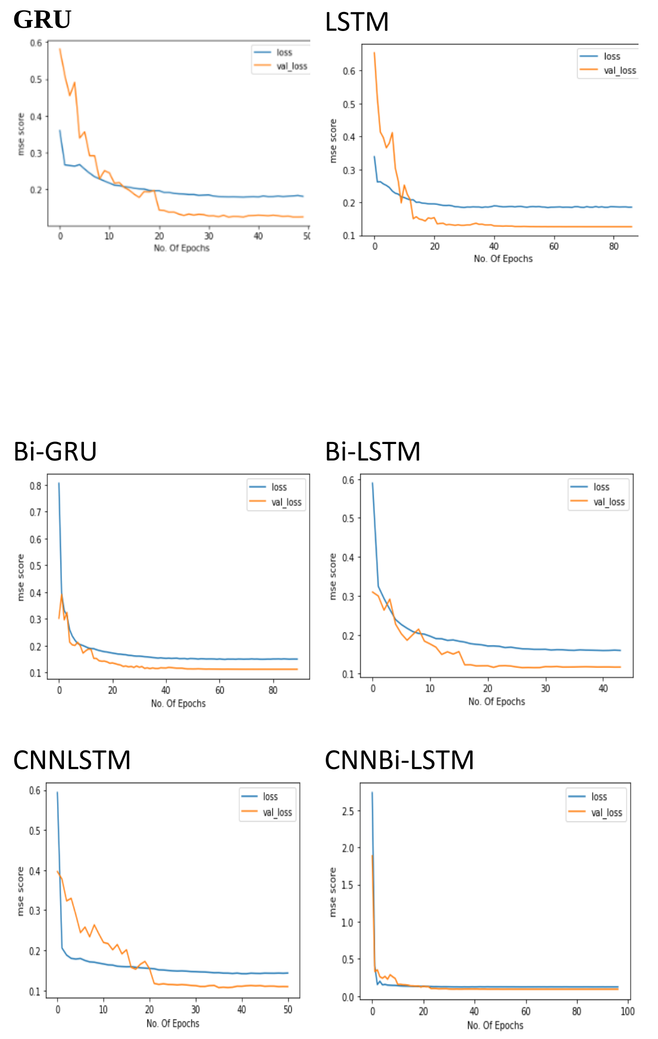


Fig 12: Graphical representation for Training loss and validation loss in models

As shown in Figure 12 it can be seen that when the epoch is almost 10, the loss value of the CNN-BiLSTM is 0.1430 which is the smallest among models and the training time taken is 36 seconds per epochs for CNN BI-LSTM which is fastest while comparing other models.

4. CONCLUSION

The suggested approach, which uses historical and meteorological data to predict air pollution, produces notable results. In this study, a CNN-BI-LSTM-based air pollution prediction model is presented. The CNN BI-LSTM is the deep learning model with the greatest overall assessment of prediction outcomes when compared to Gated recurrent unit, long short term memory, Bidirectional LSTM, CNN LSTM, and Bidirectional GRU. First, a correlation analysis was performed on the stations' historical data. A characteristic with a stronger correlation coefficient with the PM2.5, meteorological information, and correlation with other stations was chosen after experimental comparison. Second, with the hybrid model that was presented, and employed convolutional neural network to successfully take out the spatial features of and the inner characteristics between various variables; concurrently, The temporal characteristics and a more precise and reliable prediction effect were obtained using BI-LSTM. The training time of CNN BI-LSTM is decreased by 12.5 and 50.9 percent when compared to BI-LSTM and LSTM, respectively, and CNN BI-LSTM has the best air quality prediction results. The experiment's findings indicate that the CNN Bi-MAE LSTM's and RMSE both reduce by 0.20224 and 0.33256, respectively. The key conclusions of this project, as determined through performance review and outcomes comparison, are as follows: This model has great accuracy and stability and can efficiently recover the spatially and temporally aspects of the data using CNN and Bi-LSTM.

The focus of future work will be on optimizing the architectures and additional deep learning models. Various problems and study directions may be taken into consideration for future research. Future studies should take into account traffic factors when forecasting air pollution since these elements have an effect on the growth in air pollution concentration both directly and indirectly. Future predictions might be more accurate if the missing characteristics are handled and the anomalies in the pollution dataset are predicted.

ACKNOWLEDGEMENT

First and foremost, I thank GOD almighty and my parents for the success of this project. I owe sincere gratitude and heart full thanks to everyone who shared their precious time and knowledge for the successful completion of my project. With a profound sense of gratitude, I would like to express my heartfelt thanks to Dr. Fousia.M.Shamsudeen,for providing me with best facilities and motivated me throughout the work of my project. I owe my thanks to my friends and all others who have directly or indirectly helped me in the successful completion of this project.

REFERENCES

1. Garg S, Jindal H. *Evaluation of time series forecasting models for estimation of pm2.5 levels in air*, 2021 6th international conference for convergence in technology (I2CT). IEEE; 2021
2. Xayasouk T, Lee H, Lee G. *Air pollution prediction using long short-term memory (lstm) and deep autoencoder (dae) models*. Sustainability, 2020
3. Jeya S, Sankari L. *Air pollution prediction by deep learning model.*, 2020 4th international conference on intelligent computing and control systems (ICICCS) IEEE; 2020
4. Zhao J, Deng F, Cai Y, Chen J. *Long short-term memory-fully connected (lstm-fc) neural network for pm2.5 concentration prediction*. Chemosphere. 2019
5. Rijal N, Gutta RT, Cao T, Lin J, Bo Q, Zhang J. *Ensemble of deep neural networks for estimating particulate matter from images*, 2018 IEEE 3rd international conference on image, vision and computing (ICIVC). IEEE; 2018.
6. Zhang L, Li D, Guo Q, *Deep learning from spatio-temporal data using orthogonal regularizaion residual cnn for air prediction*. IEEE Access, 2020
7. Li J, Jin M, Li H. *Exploring spatial infuence of remotely sensed pm2.5 concentration using a developed deep convolutional neural network model*. Int J Environ Res Public Health, 2019
8. Ma, J. , *Improving air quality `prediction accuracy at larger temporal resolutions using deep learning and transfer learning techniques",* Atmospheric Environment, 214(July), 2019

9. Al-Janabi, S., Mohammad, M., & Al-Sultan, A. *A new method for prediction of air pollution based on intelligent computation*, Soft Computing, 2019
10. Ibrahim Kok,Suat Ozemir, *A deep learning model for air quality prediction in smart cities*" 2017 IEEE International Conference on Big Data (Big Data),2017
11. Chang, Y. S. et al. *An LSTM-based aggregated model for air pollution forecasting, Atmospheric Pollution Research*, 11(8), .2020
12. Bui, T. C., Le, V. D. and Cha, S. K.et.al. *.A Deep Learning Approach for Forecasting Air Pollution in South Korea Using LSTM*, Environmanental science,Arvix Learning ,2018
13. Chang, Y. S. et al. (2020) *An LSTM-based aggregated model for air pollution forecasting*, Atmospheric Pollution Research, .apr.2020.
14. Le, V. D., Bui, T. C. and Cha, S. K. *Spatiotemporal deep learning model for citywide air pollution interpolation and prediction*,,Proceedings - 2020 IEEE International Conference on Big Data and Smart Computing, BigComp 2020
15. Sharma, E. *Deep Air Quality Forecasts: Suspended Particulate Matter Modeling with Convolutional Neural and Long Short-Term Memory Network*s, IEEE Access, 2020
16. Abdellatif Bekkar , Badr Hssina *Air-pollution prediction using deep learning approach*,Journal of Big Data, Springer, 2021

.