A Novel Deep Learning Approach to Predict Air Quality Index

Chapter · December 2020

DOI: 10.1007/978-981-33-4673-4_29

4 authors, including:

Emam Hossain
University of Maryland, Baltimore County

Mohammad Shahadat Hossain
University of Chittagong

Karl Andersson
Luleå University of Technology

A Novel Deep Learning Approach to Predict Air Quality Index

Emam Hossain[1], Mohd Arafath Uddin Shariff[2], Mohammad Shahadat Hossain[3], and Karl Andersson[4]

[1] Department of Computer Science and Engineering, University of Chittagong, Chattogram, Bangladesh
ehfahad01@gmail.com

[2] Department of Computer Science and Engineering, Port City International University, Chattogram, Bangladesh
mhnarfth@gmail.com

[3] Department of Computer Science and Engineering, University of Chittagong, Chattogram, Bangladesh
hossain_ms@cu.ac.bd

[4] Pervasive and Mobile Computing Laboratory, Luleå University of Technology, S-931 87 Skellefteå, Sweden
karl.andersson@ltu.se

Abstract. In accordance with the World Health Organization's instruction, the air quality in Bangladesh is considered perilous. A productive and precise air quality index (AQI) is a must and one of the obligatory conditions for helping the society to be viable in lieu of the consequences of air contamination. If we know the index of air quality in advance then it would ne a great help saving our health from air contamination. This study introduces a Air Quality Index prediction model is hybrid for two mostly polluted cities in Bangladesh: Dhaka and Chattogram. Gated Recurrent Unit (GRU), Long Short Term Memory (LSTM) are the two robust variation of Recurrent Neural Network (RNN). This model combines these two together. We have used GRU as first hidden layer and LSTM as the second hidden layer of the model, followed by two dense layers. After collecting and processing the data, the model was trained on 80% of the data and then validated against the remaining data. We have evaluated the performance of the model considering MSE, RMSE and MAE to see how much error does the model produce. Results reflect that our model can follow the actual AQI trends for both citites. At last, we've juxtaposed the performance of our proposed hybrid model against a standalone GRU model and a standalone LSTM model. Results also show that combining these two models improves the overall model's performance.

Keywords: air quality index · AQI prediction · time series analysis · deep learning · hybrid neural network

1 Introduction

In today's world, one of the crucial issues is air pollution. The pollution of aerosphere is caused by detrimental chemical substances and also biological substances. In 2008, metropolitan air quality and besides indoor air quality are two of the worst contaminated places in the world according to the World's Worst Polluted Places by Blacksmith Institute. Air pollution can cause some serious problem which can be permanent in terms of health. It causes harm to crops, forests, wildlife, lakes, sea, ocean. It involves in the formation of rain including acid rain, which damages trees, rivers, soils, and animals. It affects commercial and agricultural production. Besides it also affects human beings physical state [28]. The discharge of contaminants can cause chronic respiratory and lungs difficulty [5]. The public is so tactful about air quality, because of its consequences which results various illness, sickness and even loss of life [6] [15] [13] [26]. Some environmental outcome of this are eutrophication, fog , and worldwide weather change. Air contamination is also conducting significant economic losses [31]. For all of these reason the general public is too delicate to the upcoming swing of air quality because of its huge effect on human physical state [25]. Keeping track of Air quality is a common method to alert public about the potential air pollution [3] [42].

Air quality index (AQI) is an universal index that describes about the amount of the contaminanrs in the air namely PM2.5, O3, PM10, CO, NO2, SO2 and is calculated with the help of surrounding air quality level [48]. Without that, Air Quality Index is also an important scale for showing that how the human physical state is related to it [35]. So, obtaining information about real-time air quality is surely can help to control the contamination as well as saving the health of the human from polluted air [34] [16]. Because of the nature of unpredictability and entanglement, It is hard to acquire precise Air Quality Index Forecasting results [48].

In this research, our objective is to evolve a model for forecasting daily Air Quality Index data for two of the biggest cities in Bangladesh: the capital, Dhaka and the port city, Chattogram. The Air Quality data of Bangladesh has not been analysed and predicted yet. We know that Air pollution causes a lot of problems. If the Air Quality data can be predicted thoroughly then it could help to control these air pollution related problems. Ww used daily AQI data of these two cities and applied two of the most powerful Recurrent Neural Networks (RNN) which are

Gated Recurrent Unit (GRU) and another is Long Short Term Memory (LSTM). GRU and LSTM are considered to be the most powerful deep learning models for time series analysis [19]. The principal target of this study is to use combined power of GRU and LSTM models to predict daily AQI value in Dhaka and Chattogram. Although many researches have been trying to predict AQI value, but none of the researchers utilized the combined power of LSTM and GRU so far. Therefore, we propose a hybrid model putting together two of the most powerful time series analyzers (GRU and LSTM) to predict daily AQI data.

2 State of the Art

For getting renowned and relevant research works, we have filtered out all the papers using a keyword based search on renowned publishers including *Elsevier*, *Springer*, *ACM Digital Library (DL)*, and *Taylor & Francis (T&F)*. For finding the recent state-of-the-art, we have considered only those article which were published since 2019. We have searched using "AQI Prediction" and "AQI Forecasting" these two keywords. For "AQI Prediction" keyword, we found 154, 7, 0 and 63 research articles (224 in total) since 2019 in Elsevier, Springer, ACM DL and T&F respectively. For "AQI Forecasting", these numbers are 134, 3, 1 and 27 (165 in total) for the corresponding publishers in the stipulated time. But from these 389 articles, only 15 of them have directly predict AQI value. The remaining articles don't actually predict AQI value, but the above mentioned keywords appear somewhere in the article. In this section, we have reviewed these 15 articles which actually predict AQI value.

Chen et al. [7] proposed a deep multi-task learning (MTL) based metropolitan AQI modeling method called "PANDA" to solve two task together which are AQI estimation and AQI prediction. They used CNN for extracting spatial information related to AQI modeling and then exploited Recurrent Neural Network and LSTM to study the pattern of the pursuant data. They compared it previously published prediction methods namely U-Air, Semi-EP, ADAIN, ARMA, and FFA. The MAP and MAE values of PANDA are 0.833, 9.50 respectively. "PANDA" outsails all the previous air quality estimation methods.

Kavita Ahuja and N. N. Jani [2] proposed a continual time series model which is called "Air Quality Data Model". It was developed based on supervised learning to help in the assessment of air quality in metropolitan areas. However, they didn't compare the performance of their model against other models. Moreover, while defining their algorithm, they didn't mention the linear model anywhere.

Gu et al. [12] had taken 365 sets of air contaminant data (PM2.5, AQI, SO2, CO, NO2, O3, PM10) Between the following time frame : 1 Jan, 2018 to 31 Dec, 2018 from Shenzhen. Two important algorithm were used to optimize "Support Vector Machine" model. Those are : 1. Improved SAPSO algorithm and 2. PSO algorithm. Later the model was constructed by using BP Neural Network. The values of average relative error were 34.35%, 21.00%, 23.74%, 25.56%, 32.18%, 31.99% respectively for SO2, NO2, CO, O3, PM10, PM2.5 for 15 iterations of the BP algorithm.

Xu et al. [45] had proposed a hybrid model where 7×24,000 sets of air pollutant data (PM2.5, AQI, SO2, CO, NO2, O3, PM10) Between the following time frame : May 15, 2014 to 6 Feb 6, 2017 from Kunming, Xiamen, Shenzhen and Lasa city. For preprocessing the data and also for reshaping the data "VMD" and "LASSO" were proposed. And besides for reducing the shape and to extract all the features from the data "SAE" was used. And Later after training the "DESN" the result of the prediction was acquired. For MAE, MAPE, RMSE evaluation, the results of Xiamen were 1.5428, 3.1656, 1.9153. For Shenzhen it was 3.5246, 7.3008, 4.4029, for Kunming it was 3.4548, 6.1634, 4.2431 and for Lasa it was 1.2297, 2.3123, 1.6020 respectively.

Shishegaran et al. [36] had developed model to forecast the air quality index in Tehran which was a linear and also a non-linear statistical model. 1. ARIMA, 2. PCR, 3. Combined ARIMA and PCR, 4. Combined ARIMA and GEP these four models were used to forecast the daily AQI. From these four models four equation were obtained to predict daily AQI for each seasons in 2016. The best model among them was non linear ensemble method. MAPE, RMSE (2.870 - 8.125), NMSE (0.012 and 0.51), the coefficient of determination were the evaluation metrics.

Qunli Wu and Huaxing Lin [44] proposed a novel hybrid model SD-SE-LSTM-BA-LSSVM in which SD technique was employed for cutting down the nonlinear features. The MAE, RMSE, MAPE were used to measure the performance of the model. For Beijing it was (8.8920, 6.6885, 0.0877) and for Guilin it was (4.4396, 3.8036, 0.0880) respectively. The proposed model obtains better result compared to other models.

Jiang et al. [22] proposed a hybrid learning method combining WPD-MELM-MSK-MELM. "WPD" does the decomposition of the AQI. "IPIO" optimizes the weights of the "ELM" initially. After that 'MSK" clustering were used. Finally, "MELM" was utilized to predict the AQI. The evaluation metrics used are MAE, NRMSE, RMSE, DS. Their model performs better for high-frequency model.

Weixin Zhai, Chengqi Cheng [47] investigated a "LSTM" approach for predicting AQI. The data was collected from Beijing between the following time frame : 1 Jan 2015 to 31 Dec, 2016. RMSE, MAE, were used as evaluation metrics. This LSTM model outsails other MLR models. Results show that this model is effective for forecasting air quality for a short time period.

Qunli Wu, Huaxing Lin [43] proposed a hybrid AQI forecasting model combining VMD-SE-LSTM. For decomposing the original AQI series, VMD was used and SE was employed for data processing. Finally, LSTM was applied to train the model.The evaluation metrics of this model are MAE, RMSE, MAPE. For Beijing it was 9.38, 5.66, 7.73% and for Baoding, it was 15.10, 11.97, 9.09% respectively. This proposed model can prejudice AQI quickly and conveniently and also superior over the model to some extent for daily urban AQI forecasting.

Wang et al. [41] proposed a hybrid model HI-EMD-SCA-ELM. The Hampel Identifier, VMD decomposition was used for detecting, after that correcting weird values and decomposing the actual AQI dataset. "SCA" was used to optimized the "ELM' model. Finally, all modes values were integrated to get the final values. MAE, RMSE, MAPE, U1 and U2, Pearson's correlation coefficient(r) were used to measure the performance of the model. Results showed that Shenzhen model was better.

Li et al. [29] proposed a new model to predict AQI which was based on MIO Algorithm. To predict the reconstructed series, three models namely HCA (ARIMA model based), HCME and HCFL are used. The hourly average AQI of the Shijiazhuang, Zhengzhou, and Guangzhou cities between the time frame 1 Aug, 2017 to 31 Oct, 2017 were used in this paper. The MAPE values of three datasets are 3.75%, 3.33% and 2.32% respectively.Results showed that this models outperforms other moodels in terms of accuracy.

Wang et al. [40] designed a model based on the L1 norm for monitoring and analysing the air contaminant. The data from Baoding, Tianjin, and Shijiazhuang, China were collected for this model. Using sparse regularization a total analysis was performed on the prediction of three models for getting the final result. By the help of fuzzy framework the prediction results are turned into air quality measurement.

Jiang et al. [23] combined ICEEMDAN-ICA-BPNN for hourly pollutants prediction, involving O3, PM2.5, SO2, CO, and NO2 where the Gamma, Weibull, Rayleighand Lognormal distribution functions were used to model contaminant series for characteristics estimation. The evaluation metrics used in this model are FB, MdAPE, RMSE, MAE, DA and R2.

Liu et al. [30] compared various ML and feature selection methods. The data of Beijing was collected from 2016 to 2017 . MLR, RFR, BPNN and SVR algorithm were used.And they were trained on 10 fold cross validation. Evaluation metrics were used in this model are R,MAE and RMSE. The results were 0.991, 7.089, 9.268 for MLR, 0.998, 1.337, 3.1861 for RFR, 0.996, 4.8286, 6.3371 for BPNN, 0.991, 7.357, 9.599 for SVR respectively. The experimental results follows this way RFR > SVR/BPNN > MLR

Jumaah et al. [24] developed an Air Quality Index forecasting algorithm based on temperature, humidity, wind speed, and precipitation.The data was collected from Kuala Lumpur between the following time frame : Jun, 2018 to Aug, 2018. OLS process was used for the 3 months data individually. The Accuracy was measured (97, 99, and 97%) respectively. The results came out very good and a good fit to the "Ordinary Least Square" model.

3 Overview of LSTM and GRU

For having the ability of memorizing long term dependencies, RNN is the first choice for language models [38]. However, with time lags increasing while unfolding the RNN into very deep feed neural network, gradients of RNN may vanish. As this vanishing gradients problems is a huge problem, therefore, for solving this problem LSTM [17] and GRU [8] was introduced with forget gate units whose job was to give the memory cell capability to decide when the information needs to forget.So, that's how it determines optimal time delays. The following two subsections briefly discuss about the structure of LSTM and GRU.

3.1 LSTM

For building the language based models, "LSTM" was introduced initially back in 1997. Because this Long Short Term Memory (LSTM) has superb power to remember long term allegiance. Four types of gate are available in classical LSTM. They are - 1. Input Gate 2. Input Modulation Gate, 3. Forget Gate 4. Output Gate. The structure of a typical LSTM cell is shown in the figure 1 [11].

Input gate's job is to process all the new data which comes from the outside world. The output of the input modulation gate goes to the memory cell. In the following iteration, the forget gate decides about the information which to keep and which not. By doing this, it picks the optimal delays for the input data sequence. The results which is calculated goes as input into the output gate. The output gate generates the outcome of the Long Short Term Memory cell. Generallly, A softmax layer is stacked on top of the output layer of the LSTM in language models. But in our model a dense layer is stacked on top of the output layer of the LSTM cell. In this procedure,X

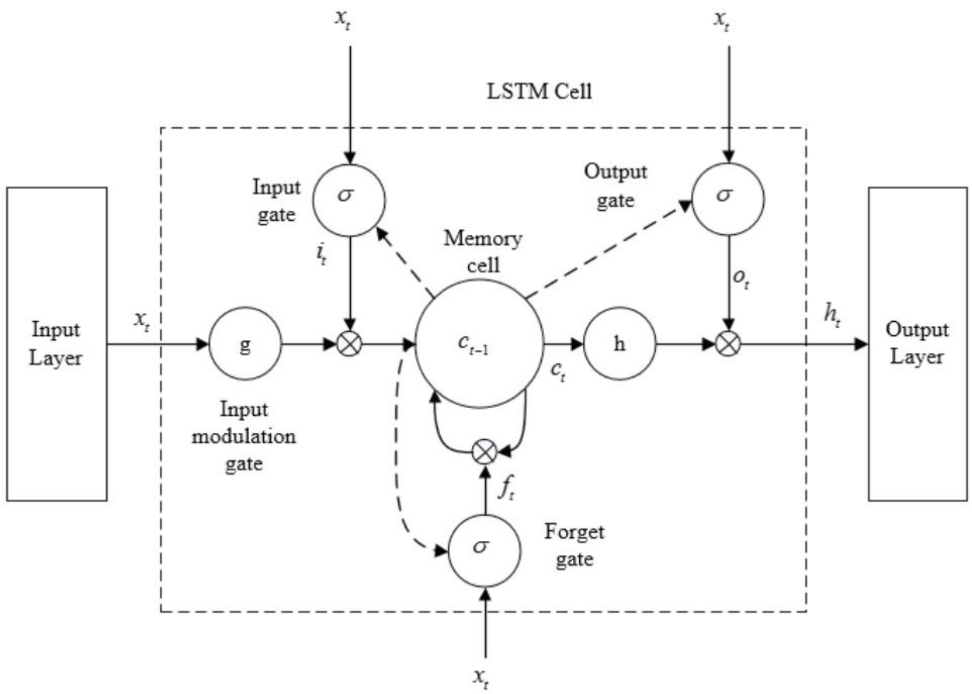

Fig. 1. Structure of a typical LSTM cell

= (x_1,x_2,...,x_n) is the input time series, H = (h_1,h_2,...,h_n) is hidden state of memory cells, Y = (y_1,y_2,...,y_n) is output time series. Hidden state of memory cell is calculated using this formula:

$$h_t = H(W_{hx}x_t + W_{hh}h_{t-1} + b_h) \tag{1}$$

$$p_t = W_{hy}y_{t-1} + b_y \tag{2}$$

Here, weight matrices is denoted as "W" and bias vectors is denoted as "b". And by using the following formulas the hidden state of the memory cells are computed.

$$i_t = \sigma(W_{ix}x_t + W_{hh}h_{t-1} + W_{ic}c_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma(W_{fx}x_t + W_{hh}h_{t-1} + W_{fc}c_{t-1} + b_f) \tag{4}$$

$$c_t = f_t * c_{t-1} + i_t * g(W_{cx}x_t + W_{hh}h_{t-1} + W_{cc}c_{t-1} + b_c) \tag{5}$$

$$o_t = \sigma(W_{ox}x_t + W_{hh}h_{t-1} + W_{oc}c_{t-1} + b_o) \tag{6}$$

$$h_t = o_t * h(c_t) \tag{7}$$

Where the standard sigmoid function is denoted by σ and it is specified in equation (8), the extends of the sigmoid function is denoted by "g". And the range of the function changes to [-2, 2], and [-1, 1].

$$\sigma(x) = \frac{1}{1 + e^x} \tag{8}$$

The square loss function is utilized for objective function according to the following formula:

$$e = \sum_{i=1}^{n}(y_t - p_t)^2 \tag{9}$$

Where the real output is denoted by "y" and the predicted AQI value is denoted by "p". Adam optimizer is applied for doing Back propagation through time (BPTT) in order to avoid local minima and to minimize training errors [27]. Neural networks tend to overfitting. A huge amount of Regularization method is introduced to reduce the overfitting problem. An efficient method to train neural networks named dropouts was proposed in order to gain better features of images in 2012 [37]. For having the recurrent property dropout technique was ineffective in RNN till 2014 [46].

3.2 GRU

In 2014, Cho et al. introduced a much simpler version of RNN which is called GRU [8]. The shape of a typical GRU cell is illustrated in the following figure 2 [11].

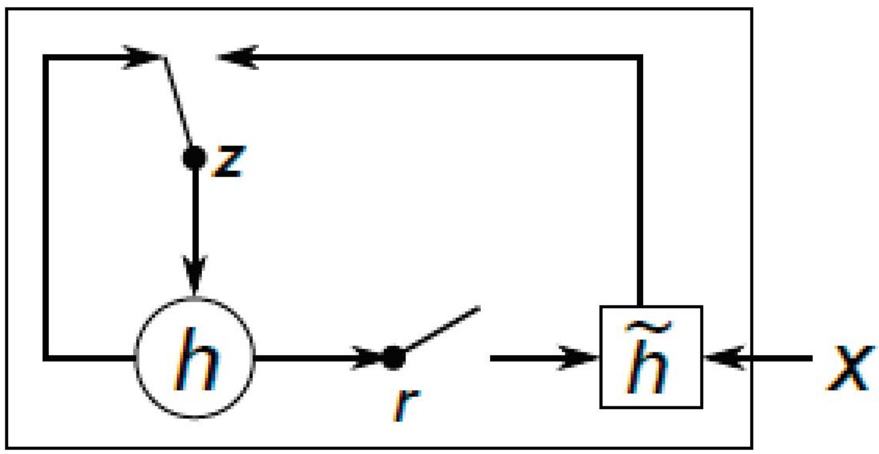

Fig. 2. Structure of GRU cell

There are two gates in a typical GRU cell. The first one is the reset gate (r) and the other one is the update gate (z). Using the value of the output of the hidden state at time t-1 and also the input time series value at the timestamp t, the output of the hidden state is calculated at time t which is basically similar to the LSTM cell. Hidden state output at time t is computed using the hidden state of time t-1 and the input time series value at time t which is similar to LSTM cell. Also, the working procedure of reset gate of the GRU cell matches to the forget gate of a Long Short Term Memory cell.

4 Experimental Setup

In this section, we discuss about the methodology of our experimental model. In the following three subsections, we discuss about the data collection process, model design and how we evaluated the model.

4.1 Data Collection and Preprocessing

Each year, air pollution costs the life of 15,000 Bangladeshis [32]. Dhaka, the capital of Bangladesh, is one of the most polluted cities in the world. A lot of large industries as well as the large amount of cars exist in the city which cause serious air pollution and needs to be taken care of. Besides, Chattogram, second most polluted city in Bangladesh and the port city of the country, has been the main transit for import and export duties of the country. We have collected daily AQI data for these two cities from the website of Bangladesh Ministry of Environment, Forest and Climate Change [10]. For Dhaka, data from June 1, 2017 to July 25, 2020 (1151 observations) are collected and for Chattogram, data collected from June 1, 2017 to March 15, 2020 (1019 observations). After collecting the data, missing values were handled by using the previous 20-days mean value as it produces better results [14]. Then we have normalized the AQI values within the range [0,1] using MinMaxScaler class of python's sklearn package.

4.2 Model Designing

We have built our layers: first layer is GRU layer with 64 neurons, second layer is LSTM layer with 256 neurons, and the remaining two layers are dense layers with 64 and 1 neurons respectively. We have trained our model using 80% of the total collected data and the remaining data are used for testing. Figure 3 shows the internal architecture of the proposed model.

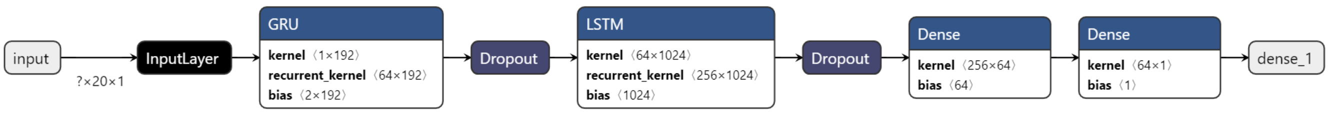

Fig. 3. Proposed Model Architecture

We have used 20 days loopback value in our model. At first, these previous 20 days AQI values is given as input at the GRU layer. GRU neurons collect the input data and pass them along to the LSTM layer. A weight is calculated in this process. Similarly, LSTM passes the collected values to the dense layer and another weight is generated. Finally, this dense layer passes the values into the output layer where the actual output is calculated. The model predicted output is compared against the actual output. This process is repeated for 200 epochs while the weighted values are updated to reduce the value of the cost function (Mean Squared Error). Finally, updated weighted values are used to predict daily AQI value for Dhaka and Chattogram.

4.3 Model Evaluation Criteria

Evaluating the performance of the model is crucial since it is used to measure how the model is behaving. In this work, we have used MAE (Mean Absolute Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error) to validate the performance of our proposed model. Equations 10-12 show the equations for these errors.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2 \tag{10}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_{ACT(i)} - Y_{PRED(i)})^2 \tag{11}$$

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(Y_{PRED(i)} - Y_{ACT(i)})^2}{n}} \tag{12}$$

MAE calculates the overall level of forecasting errors whereas MSE and RMSE put more weight on large error as they square the errors before averaging. The smaller the value of these metrics are, the better the model's prediction.

5 Result & Discussion

For both Dhaka and Chattogram, we have trained our model using 64-256-64-1 formation of the layers and was run for 200 epochs. We also compared our hybrid model against a standalone GRU model and a standalone LSTM model to see whether combining these models improve the performance of the model. The next two subsections discuss about the results of the model for Dhaka and Chattogram respectively. After that, we compare the performance of the proposed model against the two baseline models.

5.1 Dhaka

For Dhaka, we have tested our model against 230 samples which is 20% of our collected data. The model was trained on using the remaining data. Figure 4 shows the actual vs predicted AQI value curve for Dhaka city. X-axis indicates the number validation samples and Y-axis shows the normalized AQI values for Dhaka city. Actual and predicted AQI values are identified by yellow and blue colors respectively.

As we can see from the figure, our model closely predicts the trend of the AQI values over the entire timeframe. The model achieves comparitively smaller error: MSE, RMSE and MAE values for the model are 0.00514, 0.07172 and 0.053440 respectively.

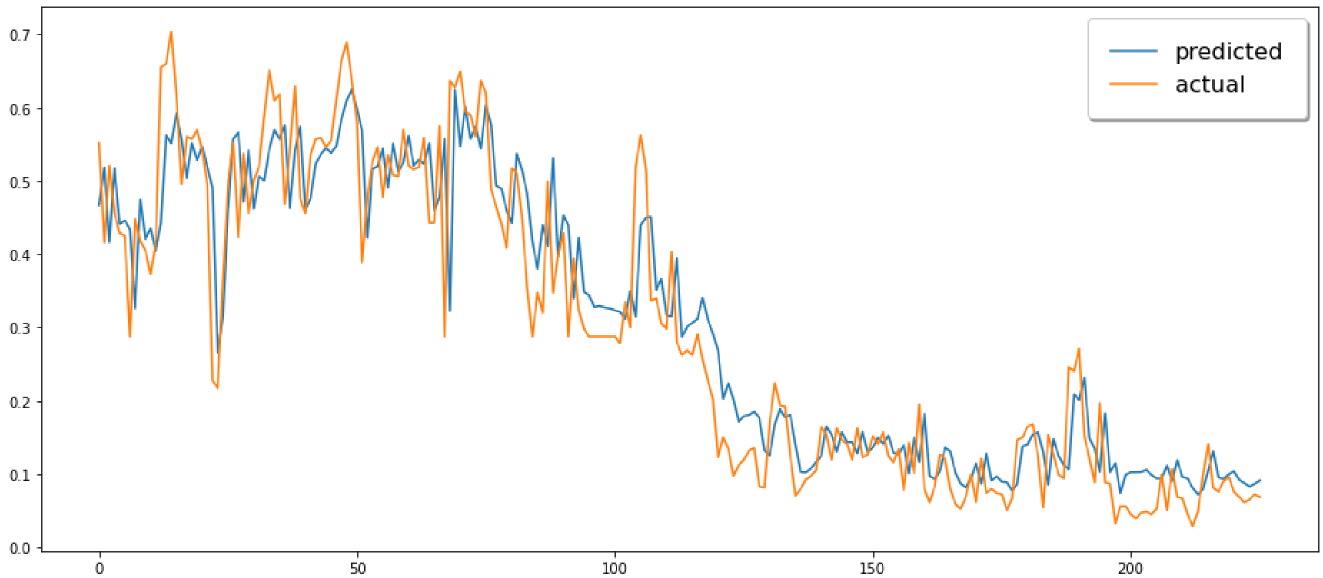

Fig. 4. Actual vs predicted value curve for Dhaka

5.2 Chattogram

In case of Chattogram, we also used 20% of the data for testing and 80% for the training. The model was tested on 204 samples and the predicted values are shown in figure 5. Similarly, X-axis and Y-axis represent the number of samples and normalized AQI values of Chattogram, respectively. Actual and predicted values are shown by yellow and blue colors as before.

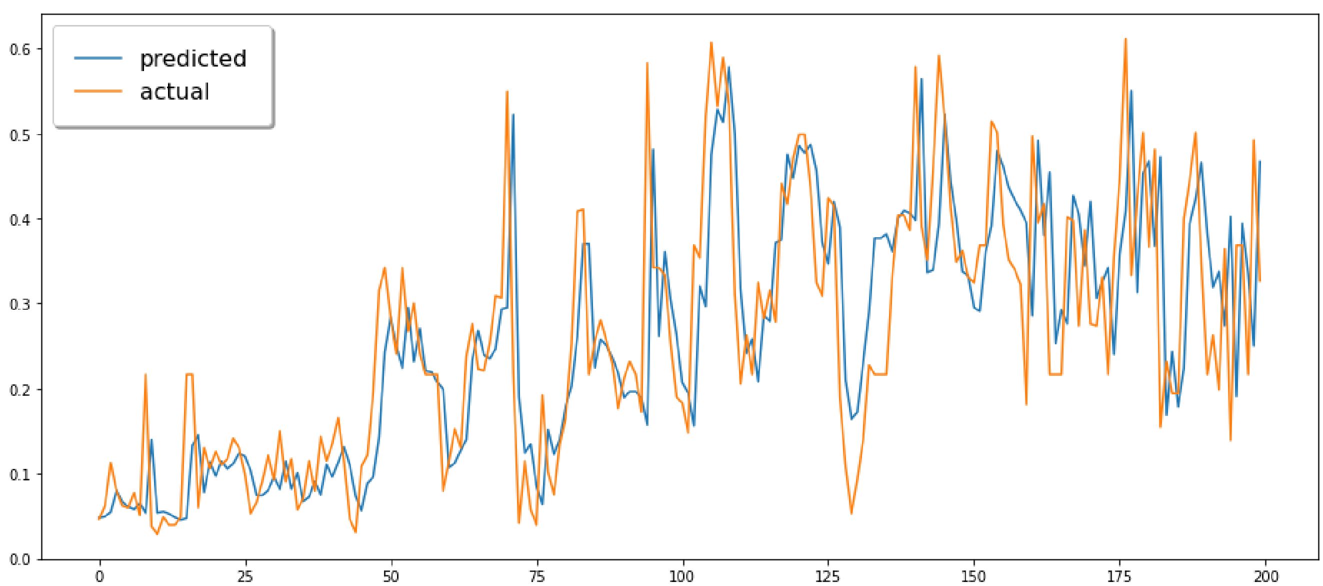

Fig. 5. Actual vs predicted value curve for Chattogram

Although there are lots of sharp trend changes in the testing samples, our model can successfully predict those changes which can be seen in figure 5. The MSE, RMSE and MAE values of this model are 0.01037,0.10184 and 0.072816 respectively.

5.3 Performance Comparison

It is important to compare the performance of a model against other baseline models two justify how good a model is. As we have mentioned before, we tried to demonstrate the combined power of GRU and LSTM models. In this section, we compare our proposed model against a stanalone GRU model and a standalone LSTM model to see whether combining these two models improve the performance or not.

We have used the same model parameters as our proposed model in these two baseline model. Tables 1 and 2 show the the performance comparison of the models in terms of MSE, RMSE and MAE for Dhaka and Chattogram respectively.

Table 1. Performance Comparison for Dhaka

Models	MSE	RMSE	MAE
Proposed Model	**0.00514**	**0.07172**	**0.053440**
LSTM	0.00849	0.09215	0.071245
GRU	0.01086	0.10419	0.063486

Table 2. Performance Comparison for Chattogram

Models	MSE	RMSE	MAE
Proposed Model	**0.01037**	**0.10184**	**0.072816**
LSTM	0.01423	0.11931	0.089882
GRU	0.01230	0.11092	0.083422

It is clear from the tables that our proposed model produces less MSE, RMSE and MAE for both Dhaka and Chattogram cities. This clearly indicates that GRU and LSTM improves the predictability of the model when they are combined together.

6 Conclusion & Future Work

Air pollution poses a great threat not only to human health but also to the environment and economy. This problem is exacerbated in Bangladesh because of the huge population and fast industrial development. Air Quality Index (AQI) shows how much polluted the air of a particular place. In this research, we have proposed a novel hybrid model combining GRU and LSTM to predict daily AQI value in the two most largely polluted cities of Bangladesh: Dhaka and Chattogram. We have collected the daily AQI data for Dhaka and Chattogram from the official website of Bangladesh Ministry of Environment, Forest and Climate Change within the timeframe 2017-2020. We have our model with four hidden layers: GRU layer, LSTM layer, and two dense layers. After handling the missing values, we have passed the dataset into input layer. These values are passed through GRU, LSTM, dense layers and finally to the output layer after 200 epochs. We have validated the performance of the model for 20% of the data and also showed the actual and predicted patterns. We validated the performance against MSE, RMSE and MAE metrics. Finally, we have compared our hybrid model against these two standalone models (GRU and LSTM) to justify our idea of combining. The results show that proposed hybrid model produces less errors than the two standalone models and therefore, proves its superiority. In future, we will apply our proposed model in all the eight divisional cities in Bangladesh and will employ other machine learning and deep learning approaches [1] [39] [18] [33] [4] [21] [20] [9] to handle the prediction in a more natural way.

References

1. Ahmed, T.U., Shahadat Hossain, M., Alam, M.J., Andersson, K.: An integrated cnn-rnn framework to assess road crack. In: 2019 22nd International Conference on Computer and Information Technology (ICCIT). pp. 1–6 (2019)
2. Ahuja, K., Jani, N.: Air quality prediction data-model formulation for urban areas. In: International Conference on Computer Networks and Communication Technologies. pp. 111–118. Springer (2019)

3. Al-Ali, A., Zualkernan, I., Aloul, F.: A mobile gprs-sensors array for air pollution monitoring. IEEE Sensors Journal **10**(10), 1666–1671 (2010)
4. Andersson, K., Hossain, M.S.: Smart risk assessment systems using belief-rule-based dss and wsn technologies. In: 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE). pp. 1–5. IEEE (2014)
5. Błaszczyk, E., Rogula-Kozłowska, W., Klejnowski, K., Kubiesa, P., Fulara, I., Mielżyńska-Švach, D.: Indoor air quality in urban and rural kindergartens: short-term studies in silesia, poland. Air Quality, Atmosphere & Health **10**(10), 1207–1220 (2017)
6. Brunekreef, B.: Health effects of air pollution observed in cohort studies in europe. Journal of exposure science & environmental epidemiology **17**(2), S61–S65 (2007)
7. Chen, L., Ding, Y., Lyu, D., Liu, X., Long, H.: Deep multi-task learning based urban air quality index modelling. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **3**(1), 1–17 (2019)
8. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 (2014)
9. Chowdhury, R.R., Hossain, M.S., ul Islam, R., Andersson, K., Hossain, S.: Bangla handwritten character recognition using convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). pp. 318–323. IEEE (2019)
10. Environment, C.A.S.: Aqi archives. `http://case.doe.gov.bd/index.php?option=com_content&view=category&id=8&Itemid=32`, accessed: 20.08.2020
11. Fu, R., Zhang, Z., Li, L.: Using lstm and gru neural network methods for traffic flow prediction. In: 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC). pp. 324–328. IEEE (2016)
12. Gu, K., Zhou, Y., Sun, H., Zhao, L., Liu, S.: Prediction of air quality in shenzhen based on neural network algorithm. Neural Computing and Applications pp. 1–14 (2019)
13. Gupta, D., Hossain, E., Hossain, M.S., Andersson, K., Hossain, S.: A digital personal assistant using bangla voice command recognition and face detection. In: 2019 IEEE International Conference on Robotics, Automation, Artificial-intelligence and Internet-of-Things (RAAICON). pp. 116–121. IEEE (2019)
14. Harrell, F.E.: Missing data. In: Regression modeling strategies, pp. 45–61. Springer (2015)
15. He, B.J., Ding, L., Prasad, D.: Enhancing urban ventilation performance through the development of precinct ventilation zones: A case study based on the greater sydney, australia. Sustainable Cities and Society **47**, 101472 (2019)
16. He, B.J., Yang, L., Ye, M.: Strategies for creating good wind environment around chinese residences. Sustainable Cities and Society **10**, 174–183 (2014)
17. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997)
18. Hossain, M.S., Andersson, K., Naznin, S.: A belief rule based expert system to diagnose measles under uncertainty. In: World Congress in Computer Science, Computer Engineering, and Applied Computing (WORLDCOMP'15): The 2015 International Conference on Health Informatics and Medical Systems 27/07/2015-30/07/2015. pp. 17–23. CSREA Press (2015)
19. Islam, M., Hossain, E., Rahman, A., Hossain, M.S., Andersson, K., et al.: A review on recent advancements in forex currency prediction. Algorithms **13**(8), 186 (2020)
20. Islam, M.Z., Hossain, M.S., ul Islam, R., Andersson, K.: Static hand gesture recognition using convolutional neural network with data augmentation. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). pp. 324–329. IEEE (2019)
21. Jamil, M.N., Hossain, M.S., ul Islam, R., Andersson, K.: A belief rule based expert system for evaluating technological innovation capability of high-tech firms under uncertainty. In: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). pp. 330–335. IEEE (2019)
22. Jiang, F., He, J., Tian, T.: A clustering-based ensemble approach with improved pigeon-inspired optimization and extreme learning machine for air quality prediction. Applied Soft Computing **85**, 105827 (2019)
23. Jiang, P., Li, C., Li, R., Yang, H.: An innovative hybrid air pollution early-warning system based on pollutants forecasting and extenics evaluation. Knowledge-Based Systems **164**, 174–192 (2019)
24. Jumaah, H.J., Ameen, M.H., Kalantar, B., Rizeei, H.M., Jumaah, S.J.: Air quality index prediction using idw geostatistical technique and ols-based gis technique in kuala lumpur, malaysia. Geomatics, Natural Hazards and Risk **10**(1), 2185–2199 (2019)
25. Khaniabadi, Y.O., Daryanoosh, S.M., Hopke, P.K., Ferrante, M., De Marco, A., Sicard, P., Conti, G.O., Goudarzi, G., Basiri, H., Mohammadi, M.J., et al.: Acute myocardial infarction and copd attributed to ambient so2 in iran. Environmental research **156**, 683–687 (2017)
26. Khaniabadi, Y.O., Hopke, P.K., Goudarzi, G., Daryanoosh, S.M., Jourvand, M., Basiri, H.: Cardiopulmonary mortality and copd attributed to ambient ozone. Environmental research **152**, 336–341 (2017)
27. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
28. Kurt, A., Oktay, A.B.: Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks. Expert Systems with Applications **37**(12), 7986–7992 (2010)

29. Li, H., Wang, J., Yang, H.: A novel dynamic ensemble air quality index forecasting system. Atmospheric Pollution Research (2020)
30. Liu, B., Shi, C., Li, J., Li, Y., Lang, J., Gu, R.: Comparison of different machine learning methods to forecast air quality index. In: International Conference on Frontier Computing. pp. 235–245. Springer (2018)
31. Liu, H., Chen, C.: Prediction of outdoor pm2. 5 concentrations based on a three-stage hybrid neural network model. Atmospheric Pollution Research **11**(3), 469–481 (2020)
32. Mahmood, S.A.I.: Air pollution kills 15,000 bangladeshis each year: The role of public administration and governments integrity. Journal of Public Administration and Policy Research **3**(5), 129–140 (2011)
33. Mahmud, T., Hossain, M.S.: An evidential reasoning-based decision support system to support house hunting. International Journal of Computer Applications **57**(21), 51–58 (2012)
34. Ni, X., Huang, H., Du, W.: Relevance analysis and short-term prediction of pm2. 5 concentrations in beijing based on multi-source data. Atmospheric environment **150**, 146–161 (2017)
35. Ribeiro, M.C., Pinho, P., Branquinho, C., Llop, E., Pereira, M.J.: Geostatistical uncertainty of assessing air quality using high-spatial-resolution lichen data: A health study in the urban area of sines, portugal. Science of the Total Environment **562**, 740–750 (2016)
36. Shishegaran, A., Saeedi, M., Kumar, A., Ghiasinejad, H.: Prediction of air quality in tehran by developing the nonlinear ensemble model. Journal of Cleaner Production p. 120825 (2020)
37. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research **15**(1), 1929–1958 (2014)
38. Tran, K., Bisazza, A., Monz, C.: Recurrent memory networks for language modeling. arXiv preprint arXiv:1601.01272 (2016)
39. Uddin Ahmed, T., Jamil, M.N., Hossain, M.S., Andersson, K., Hossain, M.S.: An integrated real-time deep learning and belief rule base intelligent system to assess facial expression under uncertainty. In: 9th International Conference on Informatics, Electronics & Vision (ICIEV). IEEE Computer Society (2020)
40. Wang, B., Jiang, Q., Jiang, P.: A combined forecasting structure based on the l1 norm: Application to the air quality. Journal of environmental management **246**, 299–313 (2019)
41. Wang, J., Du, P., Hao, Y., Ma, X., Niu, T., Yang, W.: An innovative hybrid model based on outlier detection and correction algorithm and heuristic intelligent optimization algorithm for daily air quality index forecasting. Journal of environmental management **255**, 109855 (2020)
42. Wang, Y.d., Fu, X.k., Jiang, W., Wang, T., Tsou, M.H., Ye, X.y.: Inferring urban air quality based on social media. Computers, Environment and Urban Systems **66**, 110–116 (2017)
43. Wu, Q., Lin, H.: Daily urban air quality index forecasting based on variational mode decomposition, sample entropy and lstm neural network. Sustainable Cities and Society **50**, 101657 (2019)
44. Wu, Q., Lin, H.: A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors. Science of The Total Environment **683**, 808–821 (2019)
45. Xu, Y., Liu, H., Duan, Z.: A novel hybrid model for multi-step daily aqi forecasting driven by air pollution big data. Air Quality, Atmosphere & Health **13**(2), 197–207 (2020)
46. Zaremba, W., Sutskever, I., Vinyals, O.: Recurrent neural network regularization. arXiv preprint arXiv:1409.2329 (2014)
47. Zhai, W., Cheng, C.: A long short-term memory approach to predicting air quality based on social media data. Atmospheric Environment p. 117411 (2020)
48. Zhu, S., Lian, X., Liu, H., Hu, J., Wang, Y., Che, J.: Daily air quality index forecasting with hybrid models: A case in china. Environmental pollution **231**, 1232–1244 (2017)