Research article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

Bidirectional Long Short Term Memory based Recurrent Neural Networks for Air Quality Prediction: Case of Visakhapatnam

Srinivasa Rao Kurapati[1*], Lavanya Devi G[2] and Ramesh Neelapu[3]

[1*]Research Scholar, Department of Computer Science and System Engineering, Andhra University College of Engineering; sri.kurapati@gmail.com
[2]Assistant Professor, Department of Computer Science and System Engineering, Andhra University College of Engineering; lavanyadevig@yahoo.co.in
[3]Research Scholar, Department of Computer Science and System Engineering, Andhra University College of Engineering; rameshauvsp@gmail.com

ABSTRACT

Accurate forecasting of the air quality has become a challenging task in today's scenario. There is an increasing concern on air quality ambience studies to identify and extract patterns for estimating and predicting pollutants' concentrations for a specific geographical area. With the current advances in computation, innovative modeling approaches for effective prediction of air quality have been initiated. The proposed study puts forward Bidirectional Long Short Term Memory (BI-LSTM) which considers both forward and backward dependencies to predict air quality of the city Visakhapatnam.BI-LSTM model is accurate in predicting concentration levels of pollutants by extracting temporal patterns in the past data of the pollutants. Experimental results show that proposed BI-LSTM model achieves higher prediction accuracy when compared with baseline models. Furthermore, this model may be enriched by convolutional recurrent neural networks.

KEYWORDS: air quality; air pollutants' concentrations, prediction, bidirectional long short term memory (BI-LSTM), temporal patterns, convolutional recurrent neural networks

***Corresponding author**

Srinivasa Rao Kurapati

Research Scholar,

Department of Computer Science and System Engineering,

Andhra University College of Engineering,

Visakhapatnam - 530003, A.P, INDIA.

Email: sri.kurapati@gmail.com

INTRODUCTION

Air quality is the state of the air around us. Good air quality refers to clean, clear, pure air. Clean air is essential to maintain the delicate balance of life on this planet. Ambient air quality refers to the quality of outdoor air in our surrounding environment. It is typically measured near ground level, away from direct sources of pollution. Poor air quality endangers humans, animals, plants and the whole environment. It imbalances the ecological equilibrium of the planet which may further leads to a great disaster. Air Quality Index (AQI) is an index for reporting daily air quality. It is the indicator of level of pollution in an area. Smaller the values of AQI indicate good air quality. The calculation of AQI considers major air pollutants such as $PM_{2.5}$, PM_{10}, NO, NO_2, NO_X, SO_2, CO, O_3, NH_3, Benzene, Toluene and Xylene.

Maintenance of ambient air quality is a challenging task for most of the developed and developing Nations. Hence study of air quality has become an active research problem for researchers around the globe. Air quality forecasting techniques are being rapidly advanced as the demand for measuring pollution increases.

The structure of the paper is described as follows: Section II presents the literature review. Section III summarizes the theoretical background of the proposed novel framework for predicting air quality forecasting. Implementation and performance analysis of the proposed system is presents in section IV. The last section deals conclusion and future directions.

LITERATURE REVIEW

Air quality prediction is typically measured as one of the most challenging issues among time series forecasts[1] due to its instable features. How to accurately predict air quality is still an open question with respect to the environmental and meteorological organizations of modern society. During the past decades, machine learning models, such as Artificial Neural Networks (ANNs) [2] and the Support Vector Regression (SVR) [3], have been widely used to predict air quality and gain high predictive accuracy. In the literature, however, a recent trend in the machine learning and pattern recognition communities considers that a deep nonlinear topology should be applied to time series prediction.

Recently, deep learning has proven to be very successful in many areas such as traffic flow forecasting[4], natural language processing responsibilities (emotion classification from noisy speech)[5]. These researches show that deep learning models have a superior or comparable performance with state-of-the-art methods in many fields.

Deep learning model can learn the deep features of air quality data without prior knowledge. For air quality forecasting, the deep learning method has drawn a lot of research interest[6,7,8]. For instance, Athira et al. presented a novel deep learning based air quality prediction method, which used RNN model to learn long term dependencies of air quality features [9].

However, to the best of our knowledge, only a little research has been conducted to apply bidirectional LSTM (BI-LSTM) into air quality prediction. Therefore, it is highly desired to develop a BI-LSTM framework to model air quality evolution.

This research paper aims at developing a data-driven air quality prediction paradigm which is based on the bidirectional mechanism of the LSTMs.

PROPOSED BI-LSTM MODEL

In this paper, we propose a novel design for RNNs called a bidirectional LSTM, to deal with issue of learning long term dependencies. The proposed model given a temporal sequence data of pollutant concentration values and metrological parameters of a particular location, the model will capture the dependencies in the data and predicts the next hour pollutant concentrations. The intuition is to derive inferences from sequential features to better represent data. The model is trained by providing hourly concentration values of pollutants to uncover hidden patterns across the past temporal values of the pollutants. Figure 1 represents the proposed model hypothesis.

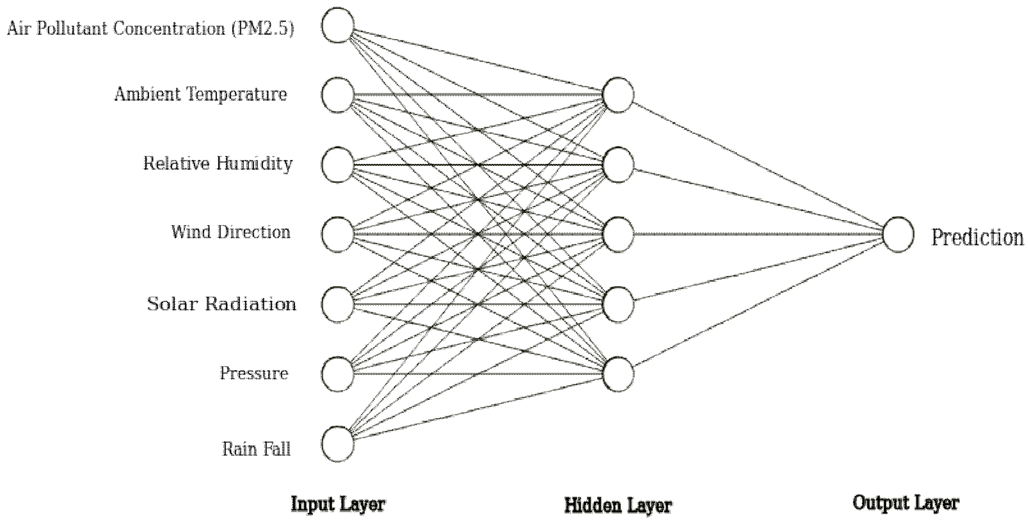

Figure 1. Model Hypothesis

Air Quality forecasting using Bidirectional LSTMs (BI-LSTM)

The idea of BI-LSTM comes from bidirectional RNN (BI-RNN), which processes sequence data in both forward and backward directions with two separate hidden layers. BI-LSTM connects the two hidden layers to the same output layer[10,11,12]. The forward layer output sequence, h_t^f is iteratively calculated using inputs in a positive sequence from time T-n to T-1, while the backward layer output sequence, h_t^b is calculated using the reversed inputs from time T-n to T-1. Both the forward and backward layer outputs are calculated by using the standard LSTM updating equations[13,14]. The BI-LSTM layer generates an output vector, y_T in which each element is calculated by using the following equation:

$$y_T = f_l\,(h_t^f,\ h_t^b) \qquad (1)$$

Where

h_t^f is the forward layer,

h_t^b is the backward layer,

f_l function is used to combine the two output sequences.

The output is a amalgamation produced by the output function f_r of both backward and forward layers.

like to the LSTM layer, the final output of a BI-LSTM layer can be represented by a vector, $y_T = [y_{T-n},\ldots\ y_{T-1}]$, is the predicted air pollutant concentration for the next time iteration when taking air quality prediction as an example.

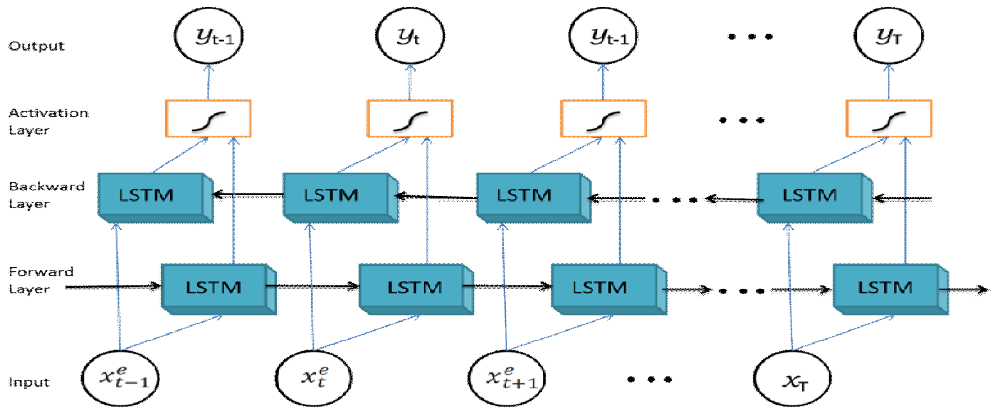

Figure 2. Air Quality Forecasting using Bidirectional LSTM

Proposed BI-LSTM can be trained using all available input information in the past and future of a specific time frame. It was shown that since the network concentrates on minimizing the objective function for both time directions simultaneously, the problem of merging outputs an optimal delay become insignificant as all future and past information around the currently evaluated time point is theoretically available and does not depend on a predefined delay parameter. The architecture of BI-LSTM working with opposing directions can be combined to achieve the goal is shown in figure 2.

IMPLEMENTATION AND PERFORMANCE ANALYSIS

The implementation of the proposed BI-LSTM model is investigated on a bench mark data set of Visakhapatnam air quality data obtained from the Central Pollution Control Board (CPCB) repository available for public use accessed at https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing.

After data preprocessing, dataset partitioning into train and test sets, model is trained with training data and parameters of the model are estimated. The model has three layers i.e. input, recurrent and output layer. Activation function for the neurons is tanh. Loss functions considered for evaluation is mae. Initial values for learning rate and dropout are set to 0.001 and 0.5. hyper parameters of the model such as dropout, momentum and learning rate are optimized by running the experiment 10 times. We implemented all code (models) using Python, Theano, Keras and Scikit-learn frameworks [15] and executed.

The results obtained from proposed framework are compared for correctness in terms of Root Mean Squared Error (RMSE), MSE (Mean Squared Error), MAE (Mean Absolute Error) and Coefficient of Determination (R^2)[8,16].

The experimental results in Table 1, Table 2 and Table 3 are shows that RNN, LSTM, GRU and Bidirectional LSTM models attain higher prediction accuracy when compared to SVR models. Mat plot of compared models is depicted in Figure 3.

Table: 1 Performance comparison of RNN, LSTM and GRU

Sno	Pollutant	RNN MODEL				LSTM MODEL				GRU MODEL			
		RMSE	**MSE**	**MAE**	R^2	**RMSE**	**MSE**	**MAE**	R^2	**RMSE**	**MSE**	**MAE**	R^2
1	PM2.5	18.556	344.334	9.729	0.730	**17.897**	**320.316**	**8.367**	**0.749**	18.328	335.907	9.088	0.736
2	PM10	**28.864**	**833.151**	**23.559**	**0.774**	31.526	993.885	19.129	0.771	31.648	1001.602	19.219	0.769
3	NO	20.574	423.301	9.247	0.223	19.916	396.652	6.700	0.272	20.248	409.992	6.674	0.247
4	NO_2	11.114	123.525	7.327	0.796	10.267	105.402	6.422	0.826	10.394	108.039	6.574	0.822
5	NO_X	19.447	378.198	10.129	0.503	**17.364**	**301.503**	**7.753**	**0.604**	17.886	319.915	8.020	0.580
6	NH_3	8.644	74.723	3.236	-0.012	8.592	73.816	3.628	0.000	8.329	69.368	2.687	0.060
7	SO_2	75.375	5681.383	22.060	-0.580	72.130	5202.686	20.712	-0.447	73.416	5389.960	26.140	-0.499
8	CO	0.464	0.216	0.293	0.430	0.472	0.222	0.302	0.412	0.422	0.178	0.213	0.530
9	OZONE	29.704	882.344	17.572	0.415	28.994	840.660	17.001	0.443	**28.223**	**796.548**	**14.157**	**0.472**
10	BENZENE	1.804	3.254	1.252	0.631	1.597	2.551	1.058	0.711	**1.567**	**2.457**	**0.997**	**0.722**
11	TOLUENE	3.584	12.845	2.264	0.698	3.243	10.520	1.932	0.753	3.303	10.912	2.042	0.744
12	XYLENE	**1.334**	**1.781**	**0.845**	**0.682**	1.417	2.007	0.970	0.641	1.477	2.181	1.049	0.610

Table: 2 Performance of Bidirectional LSTM

Sno	Pollutant	RNN MODEL				LSTM MODEL				GRU MODEL			
		RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2
1	PM2.5	18.105	327.777	8.719	0.743	18.034	325.224	8.630	0.745	18.002	324.062	8.555	0.746
2	PM10	31.356	983.217	19.205	0.774	31.334	981.806	19.141	0.774	31.513	993.049	19.333	0.771
3	NO	19.720	388.871	6.669	0.286	19.760	390.450	6.688	0.283	**19.853**	**394.144**	**6.593**	**0.276**
4	NO2	10.262	105.302	6.323	0.826	10.299	106.071	6.419	0.825	**10.258**	**105.223**	**6.372**	**0.826**
5	NOX	17.545	307.813	7.854	0.596	17.787	316.377	8.109	0.584	17.441	304.186	7.958	0.600
6	NH3	8.551	73.114	3.546	0.010	8.613	74.189	3.511	-0.005	8.467	71.684	3.192	0.029
7	SO2	70.929	5030.896	20.996	-0.399	71.942	5175.595	20.928	-0.439	69.911	4887.528	20.811	-0.359
8	CO	0.462	0.214	0.291	0.435	0.464	0.215	0.290	0.431	0.480	0.230	0.316	0.391
9	OZONE	29.199	852.561	17.965	0.435	28.463	810.122	15.494	0.463	28.403	806.705	14.273	0.465
10	BENZENE	1.586	2.514	1.033	0.715	1.591	2.531	1.041	0.713	1.588	2.522	**1.035**	0.714
11	TOLUENE	3.235	10.463	1.920	0.754	3.243	10.514	1.951	0.753	**3.242**	**10.511**	**1.937**	**0.753**
12	XYLENE	1.509	2.276	1.091	0.593	1.434	2.058	0.991	0.632	1.449	2.100	1.021	0.625

Table: 3 Performance comparison of SVR baseline models

Sno	Pollutant	RNN MODEL				LSTM MODEL				GRU MODEL			
		RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2
1	PM2.5	25.722	661.604	19.179	0.326	30.009	900.540	24.554	0.082	19.922	396.868	13.913	0.596
2	PM10	34.036	1158.426	26.711	0.664	32.478	1054.808	24.709	0.694	36.698	1346.710	29.092	0.609
3	NO	36.019	1297.380	34.157	-1.460	35.257	1243.038	33.322	-1.357	39.966	1597.269	37.912	-2.029
4	NO_2	12.599	158.739	9.963	0.781	12.334	152.115	9.533	0.790	13.632	185.832	10.621	0.743
5	NO_X	26.345	694.038	23.584	0.097	25.996	675.784	23.349	0.120	27.344	747.714	23.870	0.027
6	NH_3	19.852	394.095	18.636	-5.007	19.846	393.878	18.582	-5.004	19.625	385.137	18.334	-4.871
7	SO_2	**66.424**	**4412.194**	**53.811**	**-0.648**	66.958	4483.404	54.623	-0.675	66.940	4481.013	54.070	-0.674
8	CO	1.421	2.020	1.357	-3.265	1.321	1.746	1.255	-2.686	1.633	2.666	1.520	-4.628
9	OZONE	47.476	2253.975	40.646	-0.468	47.313	2238.561	40.331	-0.458	49.576	2457.733	42.895	-0.601
10	BENZENE	2.672	7.142	2.333	0.355	2.582	6.669	2.218	0.398	2.745	7.537	2.384	0.320
11	TOLUENE	4.468	19.962	3.635	0.535	4.517	20.405	3.679	0.525	4.346	18.887	3.391	0.560
12	XYLENE	3.822	14.611	3.461	0.269	3.689	13.609	3.246	0.319	3.476	12.083	2.961	0.395

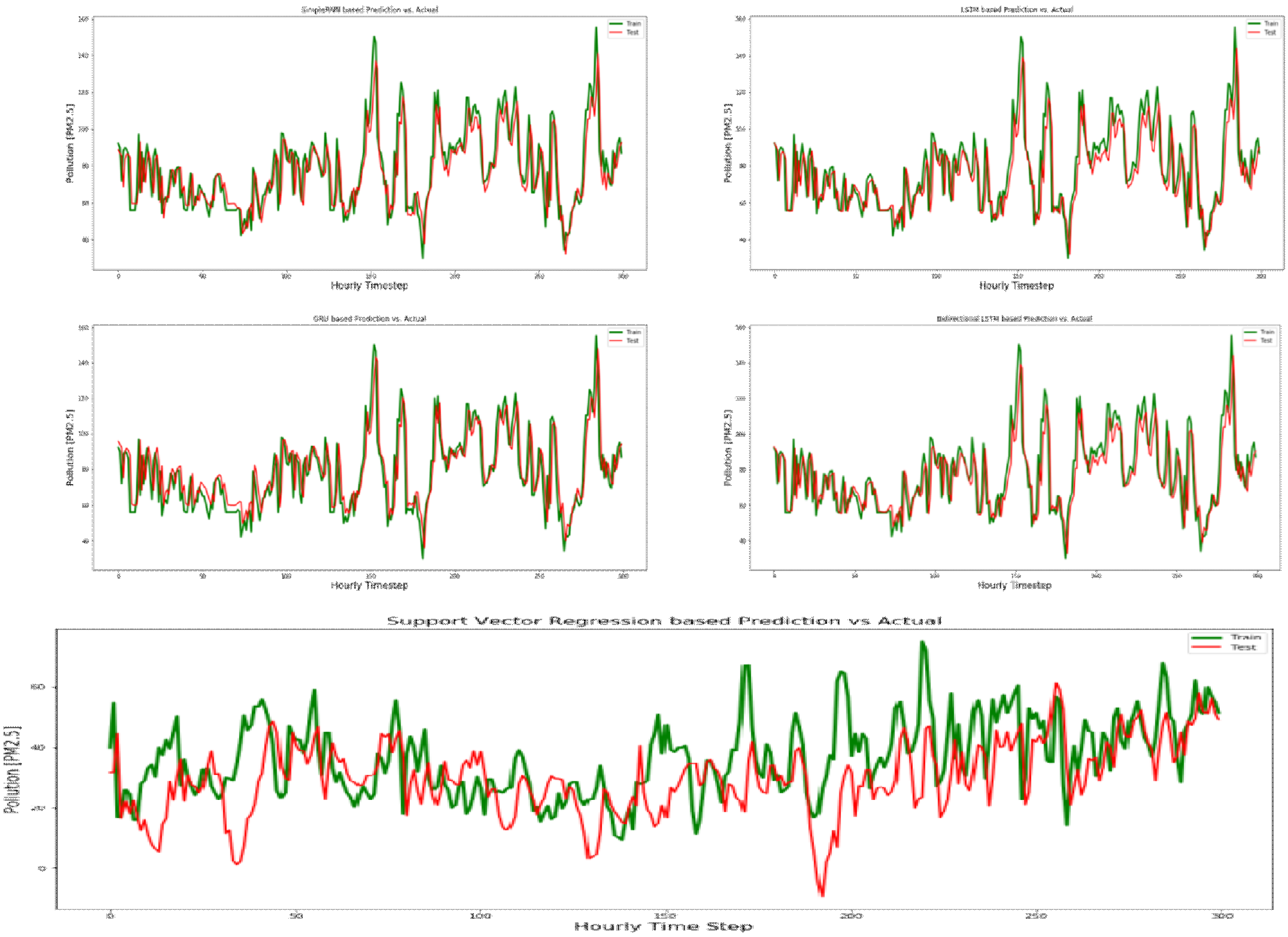

Figure 3. RNN, LSTM, GRU, and Bidirectional (RNN-LSTM-GRU) Models vs. SVR model Prediction performance graph

A quick view of the proposed model's performance is given in Figure 4.

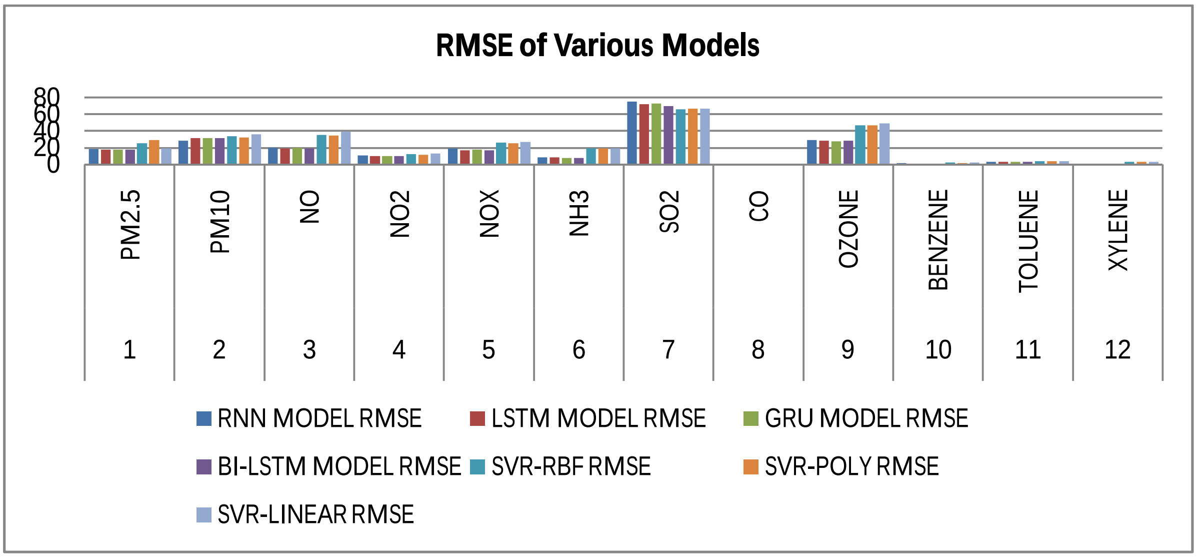

Figure 4. Performance graph of RMSE of all models

The performance is estimated in terms of accuracy, the results were found to be satisfactory. Empirical findings from the conducted experimentations' supports hypothesis of this research paper that is shown in Table 1, Table 2 and Table 3. Figure 4 summarizes the performance of the proposed BI-LSTM and other three models considered on the air quality data set. The proposed BI-LSTM architecture with various loss functions from adversarial network is illustrated here. Some of the interesting findings from the empirical evaluation carried out are listed here. As expected, the proposed approach using BI-LSTM delivered impressive results with considered metrics. Our experiments reveal that the proposed model significantly outperforms the conventional approaches on Visakhapatnam air quality data set. It is clear from figure 4 that the proposed BI-LSTM architecture outperforms the other baseline architectures. In addition, to the final modeling performance, in figure 3, we plotted the learning curves of some models. Here we qualitatively evaluate the BI-LSTM, RNN, GRU, and SVR models trained earlier by predicting air quality. Our experiments show that proposed models are able to achieve notable improvements over baselines models.

CONCLUSIONS

In this paper, proposed BI-LSTM model specific to air quality prediction in Visakhapatnam have been studied and their methodology and significance was investigated. The correctness of the model is checked by comparing the results produced by the RNN, LSTM, GRU and SVR models. The performance of the proposed BI-LSTM model was examined in terms of accuracy using Visakhapatnam air quality data set and the results have been promising. The results were consistent over different 12 pollutants and clearly demonstrated that BI-LSTM architecture is helpful when the models are trained on complicated sequences that involve long term dependencies. BI-LSTM was able to outperform the previously reported best on air quality prediction modeling. This suggests that BI-LSTM is also scalable. In our future studies, we aim to investigate the performance of proposed framework for real-time air quality prediction on smart phones.

REFERENCES

1. Box G.E.P, Jenkins G.M, "Time series analysis: forecasting and control", J.Operational Res.Soc. 1976; 22: 199-201.
2. Gianluigi de Gennaro , Livia Trizio, Alessia Di Gilio, Jorge Pey, Noemi Pérez, Michael Cusack, Andrés Alastuey and Xavier Querol, "Neural network model for the prediction of PM10 daily concentrations in two sites in the Western Mediterranean", Science of the Total Environment, 2013; 875–88.

3. Nieto P.G, Combarro E.F, Del Coz Diaz J.J, “A SVM- based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): a case study”, Appl.Math.Comput, 2013; 8923-8937.
4. Honglei Ren, You Song, Jingwen Wang, Yucheng Hu and Jinzhi Lei, “A Deep Learning Approach to the City wide Traffic Accident Risk Prediction”, https://arxiv.org/pdf/1710.09543.15, April, 2018.
5. R Collobert, J Weston, “A Unified architecture for natural language processing: Deep neural networks with multitask learning”, In Proceedings of the 25th International Conference on Machine Learning, 2008; 5-9.
6. Xiang Li et al., “Deep learning architecture for air quality predictions”, Springer, 2016.
7. Siuli Roy, Anurag D and Somprakash Bandyopadhyay, “Testbed Implementation of a Pollution Monitoring System Using Wireless Sensor Network for the Protection of Public Spaces”, Management Information Systems, Indian Institute of Management Calcutta.
8. Zhongang Qi et al., "Deep Air Learning: Interpolation, Prediction, and Feature Analysis of Fine-grained Air Quality", In CoRR abs/1711.00939,arXiv: 1711.00939. URL: http://arxiv.org/abs/1711.00939,2, Nov 2017.
9. Athira V, Geetha P, Vinayakumar R, Soman K P, “DeepAirNet:Applying Recurrent Networks for Air Quality Prediction”, Elsevier, Procedia computer science, 2018; 132:1394-1403.
10. Alex Graves and Jurgen Schmidhuber, “Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures”, https://doi.org/10.1016/j.neunet.2005.06.042
11. Alex Graves and Jurgen Schmidhuber, “Framewise Phoneme Classification with Bidirectional LSTM Networks”, Proceedings. IEEE International Joint Conference on Neural Networks, 2005.
12. Zhiyong Cui and Yinhai Wang, “Deep Stacked Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction”,https://arxiv.org/abs/1801.02143.
13. K. Srinivasa Rao, Dr. G. Lavanya Devi, N. Ramesh, “Air Quality Prediction in Visakhapatnam with LSTM based Recurrent Neural Networks”, International Journal of Intelligent Systems and Applications (IJISA), 2019; 11(2): 18-24,. DOI: 10.5815/ijisa.2019.02.03
14. K. Srinivasa Rao, Dr. G. Lavanya Devi, N. Ramesh, “Air Quality Prediction in Visakhapatnam using Deep Learning techniques: RNN, LSTM and GRU”, International

Journal of Research in Electronics and Computer Engineering (IJRECE), July - September 2018; 6(3): 2101-2105.

15. Rami Al-Rfou et al., “Theano: A Python framework for fast computation of mathematical expressions”, ar Xiv preprint ar Xiv: 1605.02688, 2016.
16. Ibrahim Kok et al., “A deep learning model for air quality prediction in smart cities”, IEEE International Conference on Big Data, 2017.