A COMPARATIVE ANALYSIS of DEEP LEARNING MODELS FOR AIR QUALITY INDEX PREDICTION

Subhradip Roy[1], Sauvik Roy[2], Manoj Sharma[3], Vishal Mishra[4], Ocean Rashtrapal[5], Pawan Kumar Mall[6]

School of Computer Science and Engineering, Lovely Professional University, Phagwara, India

subhrastien@gmail.com[1], sauviksirkar@gmail.com[2], manoj81017sharma@gmail.com[3], vishalmishra1190@gmail.com[4], oceangajbhiye1999@gmail.com[5], pawankumar.mall@gmail.com[6]

Abstract

The influence of air pollution on human health is measured by the Air Quality Index (AQI). The AQI is a standardized scale that offers a monetary value to represent the degree of air quality in a specific location. The concentration of five main air pollutants—ground-level ozone, particle pollution, carbon monoxide, sulphur dioxide, and nitrogen dioxide—is used to compute the AQI. Prediction of the air quality index (AQI) can offer useful information that can assist people safeguard their health, schedule outside activities, and take actions to enhance air quality, simplifying and improving their lives. The aim of our research work is provide a comparative analysis of different artificial intelligence (AI) in prediction of air quality index (AQI). The AI model involved in our study are LSTM, BiLSTM, BiLSTMConv1D, and GRU. This research will help humans as well other livelihood, as significant number of livelihood experience illnesses brought on by air pollution each year, but only a small percentage of those endure fatalities. Overall based on our research proposal, we can conclude GRU out performs all other model and achieved the minimum values for Mean Squared Error (MSE)= 0.0006, Mean Absolute Error (MAE) = 0.01116 and Root Mean Squared Error(RMSE) = 0.0246.

Keywords: Air pollution, Human Health, Comparative Analysis, Air Quality Index (AQI), Artificial Intelligence (AI), LSTM, BiLSTMConv1D, GRU

1. Introduction

One of the main issues facing civilization today is air pollution. A significant number of people experience illnesses brought on by air pollution each year, but only a small percentage of those endure fatalities. According to data provided by WHO, nearly all people on earth live in excess of the safe WHO guideline limit. Due to the use of unclean fuels like stove fire, coal burning, biomass, etc., about 2.4 billion people are exposed to extremely high amounts of household air pollution. An estimated 6.7 million premature deaths per year are attributed to air pollution, with 3.2 million of those deaths being attributed to domestic air pollution. Children under the age of five account for a startling 237,000 of all premature fatalities caused by household air pollution. Premature fatalities are brought on by a number of illnesses that are brought on by air pollution. Here are the ailments that are most prevalent:

a) Ischemic heart disease
b) Stroke
c) Lower respiratory infection
d) Chronic obstructive pulmonary disease
e) Lung cancer

In addition, air pollution is a key sign of the severe climate change brought on by global warming. High temperatures, harsh winters, heat waves, and rising water levels are some examples relative concerns along the same line.

World organizations must develop strong policies, action plans, and regulations to combat the exponential growth of air pollution (and eventually address global warming). To do this, these organizations require a reliable method that will enable them to take necessary action and predict the air quality accurately or nearly accurately [1].

ARIMA and ARMA (Expensive in terms of computation. Weaker long-term forecast success. Seasonal time series cannot be used. It's harder to describe than exponential smoothing.)
GWR (problems of multi-collinearity and the approaches to calculating goodness of fit statistics.)
MLR (non-linear nature of the data)

Apart from these Statistical models we also got some CTM (chemical transport model)

Community Multiscale Air Quality (CMAQ) model
Weather Research and Forecasting model coupled to Chemistry (WRFChem)
Nested Air Quality Prediction Modeling System (NAQPMS)

The CTM suffer from uncertainties, inaccurate data, in-consideration of less known meteorological elements, connection between different scales over different regions is difficult to simulate.

To solve this problem most of researchers turn towards deep learning [2] [3] [4]. The idea of "Deep Learning" has been a hot subject for debate and study in the artificial intelligence community since 2012. It is found that deep learning is able to solve various problems in the domain of prediction and forecasting which were earlier limited by the amount of data and the nature of complexity it can handle. But with the advancement of technologies and invention of powerful and small CPUs and GPUs along with easy access to cheap storage and computation power (thanks to cloud), deep learning is able to complete multiple projects which were earlier limited by both economic and computational constraint.

In this study, to solve the problem in hand we will work with five models namely LSTM, BiLSTM, BiLSTMConv1D, GRU and SVR then find the best among them. Finally, the dataset we would be working with is Beijing Multi-Site Air-Quality Data Set.

2. Related Work

Since 2012 researchers have been extensively using the power of deep learning to solve multiple problems in the world that was earlier restrained by the capability of the concerned system. It has found its application in many domains throughout the world such as medical sector for example In this[5] Review writing author discusses how patient heterogeneity makes it difficult to diagnose neurodegenerative diseases and produce diagnostic tools, and how machine learning can fix such problems and offer new therapies. Another example could be this study[3] which uses data from Lung Image Database Consortium (LIDC) database and three models namely Deep Belief Networks, CNN and Stacked Denoising Autoencoder. Further the results are compared with existing computer aided diagnosis (CADx) system. Then again the field of medical science isn't restricted to diagnosis only we also have the pharmaceutical sector and machine learning has quite a contribution here as well for instance in[16]; the author discusses various applications of ML a data driven approach in the sector of drug discovery and its development in various stages, it also discusses what are the various challenges. In another study[17] the author suggests a system based on block chain along and machine learning to manage the drug supply chain in order to keep proper tab on the drugs to avoid counterfeiting and provide recommendation to the customers respectively.

Not only in medical we have its use in sector of finance and banking for example here[4] author makes use of complicated, non-linear, and intrinsically dynamic data to make predictions about stock market groups. It makes use of a variety of deep learning models, including random forests, bagging, adaptive boosting, decision trees, extreme gradient boosting gradient boosting, as well as LSTM, ANN and RNN. In an another study[15] author uses logistic regression models on various attributes of a customer to determine what are the chances of defaulting and grant a loan on the basis of that.

Speaking about application of machine learning in field of environmental studies our own study is an example to it but there are a lot more than that say in this study [8] the research compares changes in the area a water body covers in order to gain insight into the effects of global warming. For the investigation, it computes changes of the water-covered region using time-lapsed photos utilizing versions of the Detectron2 instance segmentation architectures.

AQI prediction requires us to work on time series data. In machine learning when we have time series data we turn towards RNN or LSTM and their various advancements and tweaks to solve the problem in hand. There are quite a few numbers of works based on time series data for example In [9] the author describes the application of machine learning for weather forecasting by the National Center for Atmospheric Research. It lists several of the company's AI-based achievements, such as the Dynamic Integrated Forecasting (DICast®) System, one of the initial automated forecasting systems. Sticking to the nature a popular example of time series data is rainfall data, in [19] the author talks about effects of heavy rainfall in a countries such as India which is majorly based on agriculture and reviews major machine learning techniques for its accurate prediction. Another example of time series data could be parking data on various times, in the following study[10] the author implements ARIMA and LSTM models working on data of parking spots usage to predict parking space usage alterations [21][22][23][24].

As we have earlier discuss there is a credible amount of work done in terms of AQI prediction using various models and hybrid composition of it, providing a varying degree of success. In[11], the author employs LSTM to use huge amount of data from number of IoT-based smart city apps to forecast the air quality of smart city. In[15] author proposes VMD-BiLSTM hybrid model for prediction of changes in PM2.5 over various cities of China. The raw PM2.5 complex data is divided in multiple number of sub signal components using variational mode decomposition. Then, each sub signal component is predicted separately by implementing a bidirectional LSTM. In[10] author puts forward SVR-RBF model that is Support Regression Vector along with radial bias function as kernel to forecast level of pollutants and particulate and predict AQI for California region. In[16] the author suggest a (1D-CNNs) BiLSTM where, bi-directional LSTM networks are utilized to understand the spatial-temporal dependencies of time-stamped multivariate air quality data, while 1D-CNNs are employed to carve out local trends and spatial correlation signatures. This model is used to forecast PM2.5 air pollution [25][26][27].

3. Dataset and pre processing

For our study we are using is Beijing Multi-Site Air-Quality Data Set [20]. It contains hourly statistics on air pollutants from 12 stations monitoring air quality under national supervision. Beijing Municipal Environmental Monitoring Center provides the stats on the condition of the air. Each air-quality site's monitoring data is compared against the location of the closest weather monitoring station controlled by the China Meteorological Administration. The duration being March 1st, 2013, to February 28th, 2017.

Originally the dataset consists of 420768 entries, post processing we are left with 401611. The attributes of the data consist of six pollutants (PM2.5, PM10, SO_2, NO_2, CO, O_3), atmospheric attributes (pressure, temperature, dew point temperature, precipitation, wind velocity and direction), date attributes (year, month, day, hour) and finally name of station [28][29][30].

The ratio for training set : validation set is 80:20. Here we are using validation set as test set as well.

The steps involved in pre-processing are as follows:

Step 1: Dropped all Null entries
Step 2: New columns are added (PM10_24hr_avg, PM2.5_24hr_avg, SO2_24hr_avg, NO2_24hr_avg, CO_8hr_max, O3_8hr_max)
Step 3: Created sub-index columns for all pollutants from columns generated in the last step using pre-specified formulas.
Step 4: Finally using the sub-index values and applying required formula we achieve AQI_calculated, i.e. the column consisting of AQI values . [31][32][33].

4. Models Used

For this study our dataset is nonlinear, heterogeneous, multidimensional and sequential data. So, keeping all this in mind we will be using LSTM, BiLSTM, BiLSTMConv1D AND GRU. All The stated models are advancement and variations of Recurrent Neural Network architecture which is used for handling sequential data. Each of them are explained in the upcoming segments [34][35][36].

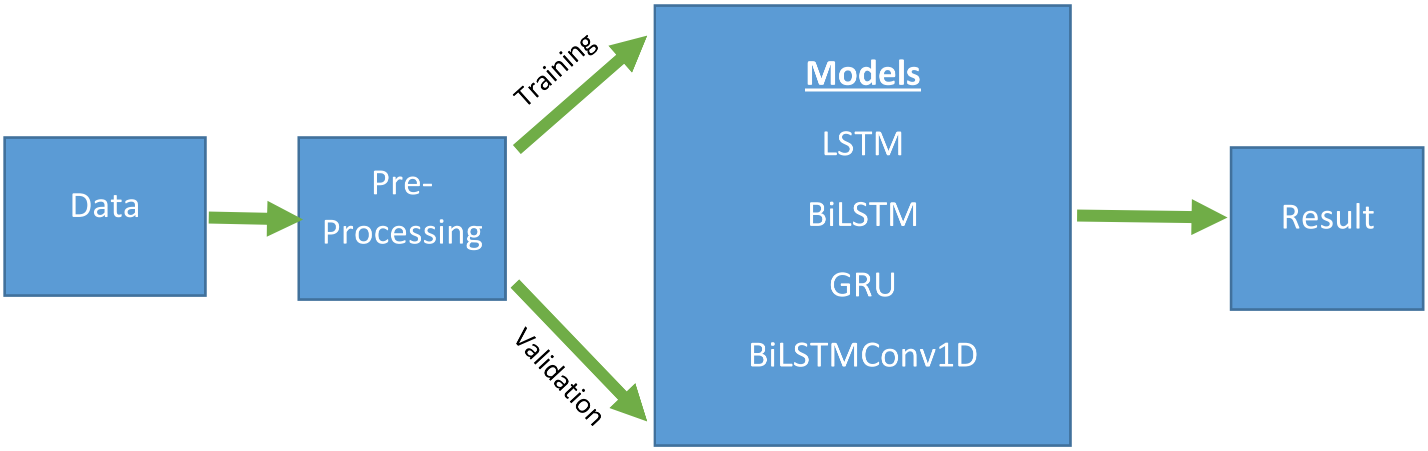


Fig. 1 (Rundown, of the whole process for the study)

4.1 LSTM

Stands for Long Short Term Memory. It is an advancement to the to the classic Recurrent Neural Network. It is used to solve the RNN's problem of gradient vanishing. This model can maintain relevant information from far back in the sequence which RNN is unable to when length increases. Along with the capability of retaining very old data it can also forget irrelevant information and use only a certain part of current input (as per relevance) to maintain the gradient. It also keeps track of two type of data Long term (cell state) and Short term (hidden state) [37][38][39].

To accomplish this a simple neuron in replaced with a memory block consisting of three gates:

a. <u>Forget gate:</u>
   This gate is used to determine what part of the cell state (long term memory of the architecture) will be retained based on current input data and hidden data from the previous hidden state [40][41].
b. <u>Input Gate:</u>
   It is used to determine what new information would be added to the cell state or the long term memory on the basis of current input and previous hidden sate[42][43].
c. <u>Update gate:</u>
   This gate determines the new hidden state based of new cell state, previous hidden state and current input data.

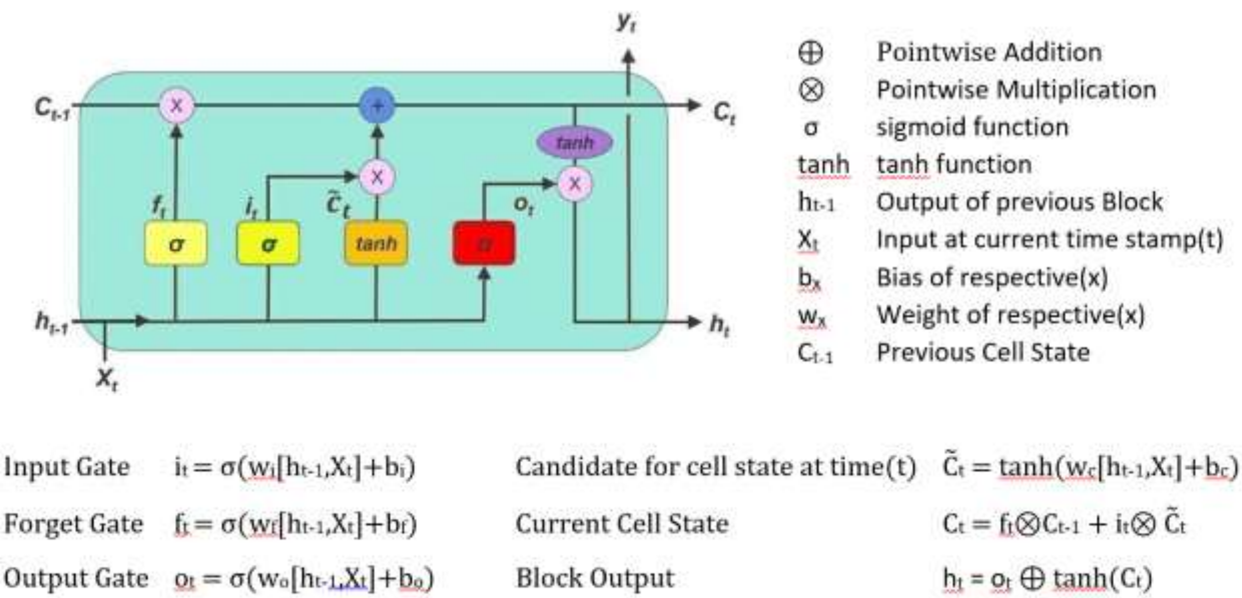


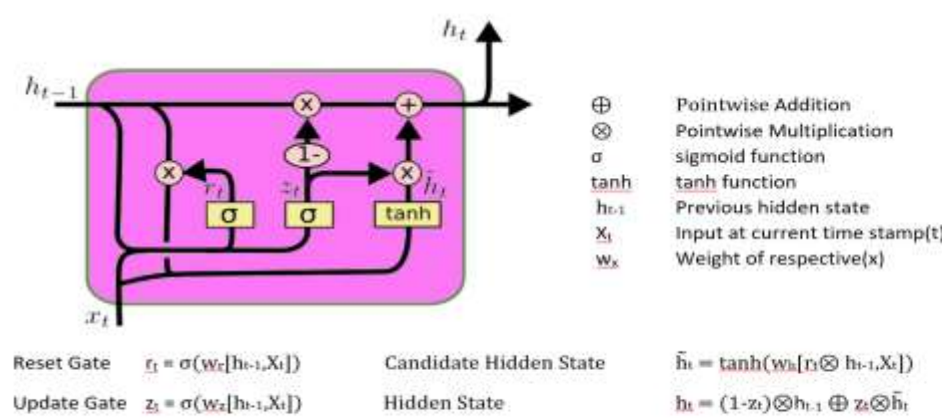
Fig. 2 (LSTM cell unit)

4.2 GRU (Gated Recurrent Unit)

Similar to LSTM it is also an advancement to RNN. This was also developed to solve the vanishing gradient problem. To solve the problem of classic RNN model it incorporates two gates:

a. <u>Update gate:</u>
   It is used to determine what amount of past data from previous steps should be passed along to the next output.
b. <u>Reset Gate:</u>
   This gate is used to determine what part of the previous information has to be rejected for the upcoming part.

Fig. 3 (GRU Cell Unit)



4.3 BiLSTM

It can be simply defined as two RNNs put together, one working in a forward direction whereas the another one working in the backward manner. Here the sequential data is once put in the normal order and once in the opposite order. This feature of this model allows it to take in account of both backward and forward information while prediction instead of just forward direction like in LSTMs, allowing it to be more precise.

The diagram of the same is provided below [15]:

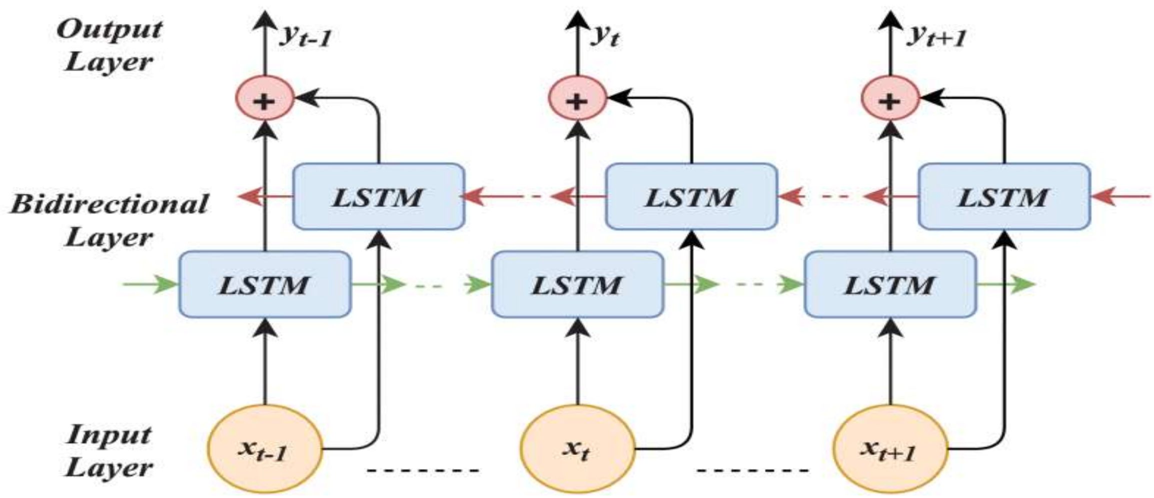


Fig. 4 (BiLSTM, a setup made up of two RNNs)

4.4 BiLSTMConv1D

This is a hybrid BiLSTM which works with a 1D convolutional neural network. In general, when we say CNN we mean 2D CNN which is very successful in image detection, feature and edge detection etc. rather it is sometimes stated as the prime architecture used to create computer vision.

When we talk about 1D CNN we are talking about an architecture that is really good with time series data. It has been used in this domain for quite a time now. This model is effective in capturing both local and global temporal patterns in time series data. Conv1D fetches the local trend along with spatial correlation signatures whereas the BiLSTM finds spatial-temporal dependencies [44].

5. Experimental Setup

The following table provides various attributes about the models we used for AQI prediction.

Table 1 (The list of hyper-parameters and their values used in the setup)

	LSTM	BiLSTM	BiLSTMConv1D	GRU
Total number of	3	5	8	9

layers				
Number of nodes in dense layer	50,1	1	1	1
Number of nodes in input layer	128	100	100	50
Epochs	50 (callback at 26)	50 (callback at 24)	50(callback at 26)	50
Batch Size	64	64	64	64
Learning rate	0.01	0.001	0.001	0.01
Dropout layer(%)	NA	0.4, 0.4	0.2, 0.2	0.2, 0.2, 0.2, 0.2
Activation function	$\tan(h)$	<Not Specified>	LeakyReLU(alpha=0.3)	$\tan(h)$
Loss function	MSE	MSLE	MSLE	MSE
Optimizer	Adam	Nadam	Nadam	SGD(decay=1e-7, momentum=0.9, nesterov=False)
Total params	79,205	332,201	150,273	53,901
Trainable params	79,205	322,201	150,273	53,901
Non-trainable params	0	0	0	0

The Conv1D hyper-parameters for the model BiLSTMConv1D are as follows:
Conv1D filters = 64, kernel size = 3 and MaxPooling1D – pool_size = 2.

6. Result

To understand and compare results let us check the values for MSE, MAE, RMSE, MSLE [21] [22] [23] [24] and a custom error function RMSLE (Root Mean Squared Logarithmic Error) values of various models side by side.

$$MAE = (1/n) * \sum|y - \hat{y}| \quad (1)$$

$$MSE = (1/n) * \sum(y - \hat{y})^2 \quad (2)$$

$$RMSE = \sqrt{1/n * \sum(y - \hat{y})^2} \quad (3)$$

$$MSLE = 1/n * \Sigma_i(\log(y_i+1) - \log(\hat{y}_i+1))^2 \quad (4)$$

$$RMSLE = \sqrt{\left(1/n * \Sigma_i(log(y_i + 1) - log(\hat{y}_i + 1))^2\right)} \quad (5)$$

Table 2 (Performance chart of all models in terms of MSE, RMSE, MSLE, RMSLE and MAE)

	MSE	MAE	RMSE	RMSLE(custom)	MSLE
LSTM	0.02184	0.10525	0.14777	NA	0.01455
BiLSTM	0.02193	0.10538	0.14808	NA	0.01467
1D-CNN-BiLSTM	0.02183	0.10545	0.14775	0.12086	NA
GRU	0.0006	0.01116	0.0246	NA	NA

The data from the previous table can be presented in a graph format to visually compare the data and understand the margin. The graphs are as follows:

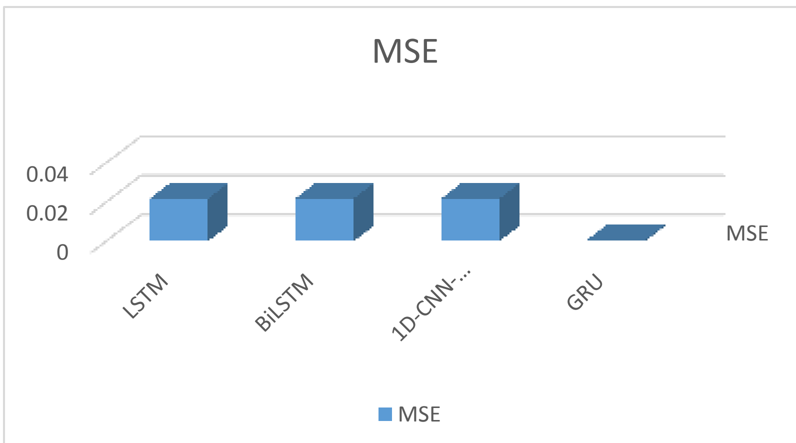


Fig. 5 (Bar graph, showing MSE values for all Model)

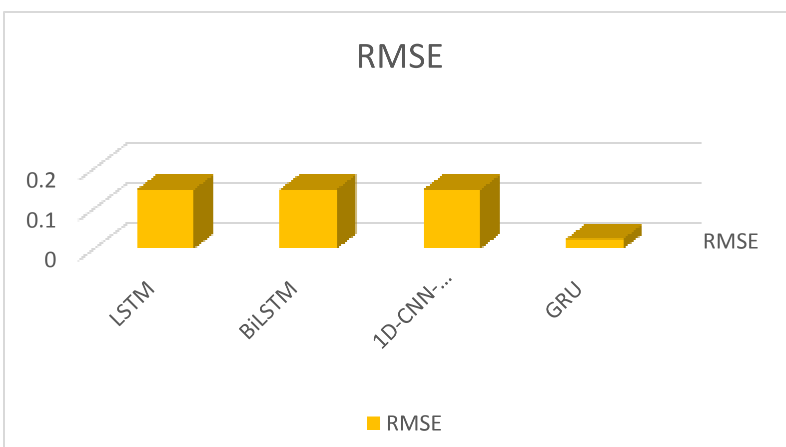


Fig. 6 (Bar graph, showing RMSE values for all Model)

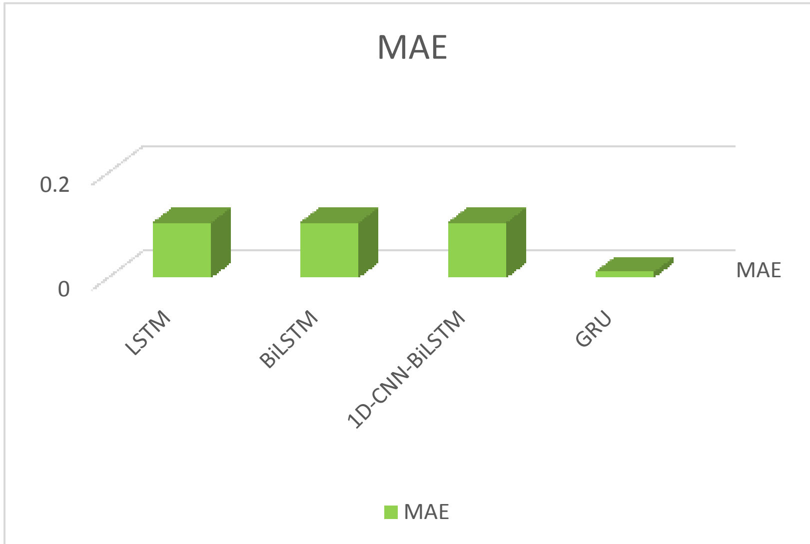


Fig. 7 (Bar graph, showing MAE values for all Model)

Next, we have Actual Vs Prediction graph, Loss function graph and MSLE graph for all the models. For all the experiments we have an epochs value of 50 with different callback values. Other metrics are model specific and are mentioned specifically for each of them (also mentioned in the table previously)

a. LSTM
   Batch size of 64, with an Adam Optimizer, tan(h) activation function and MSE loss function.

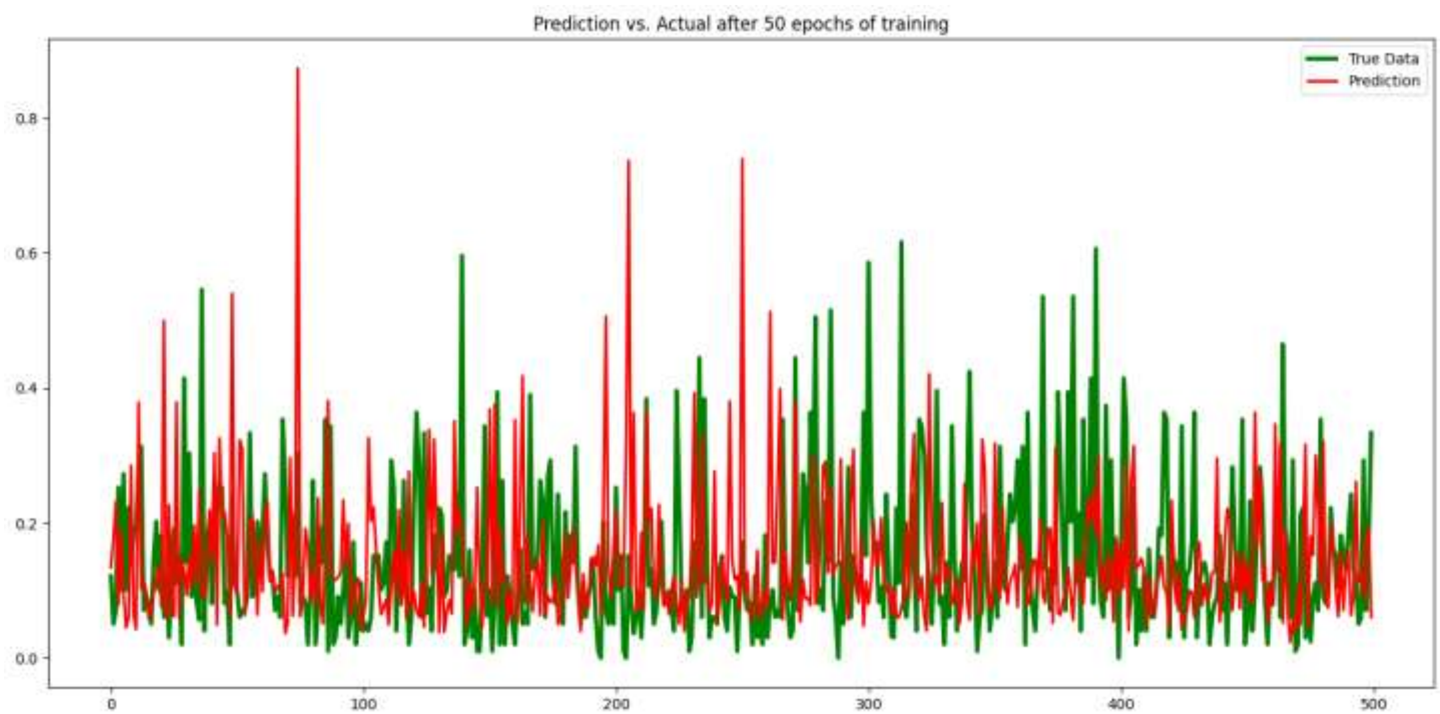


Fig. 8 (Actual vs Prediction graph for LSTM)

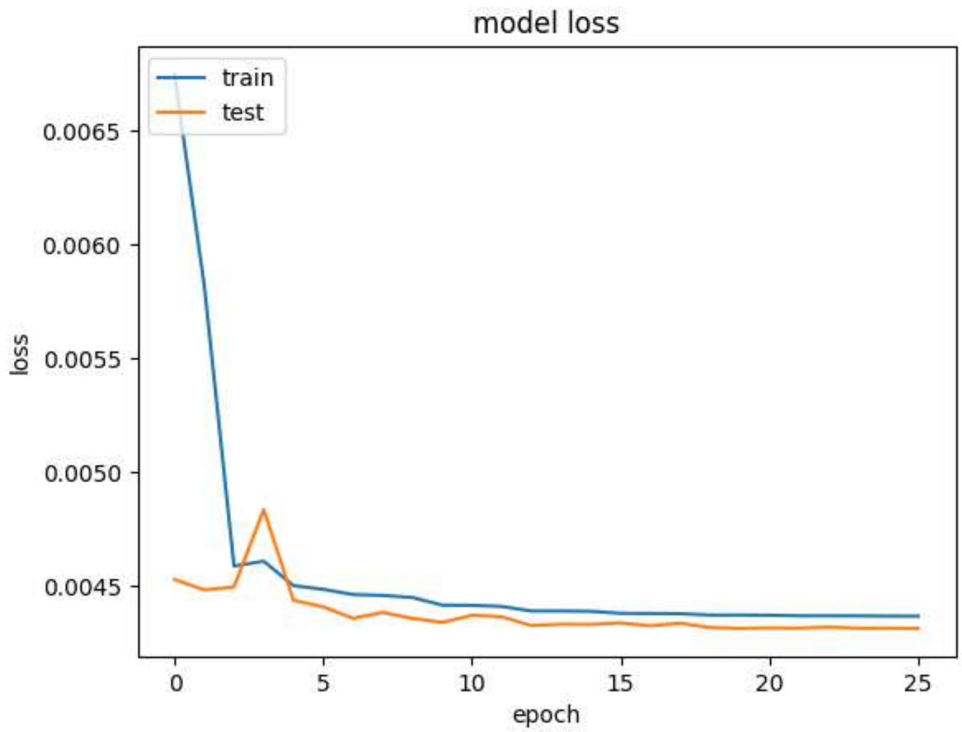


Fig. 9 (Loss Graph for LSTM)

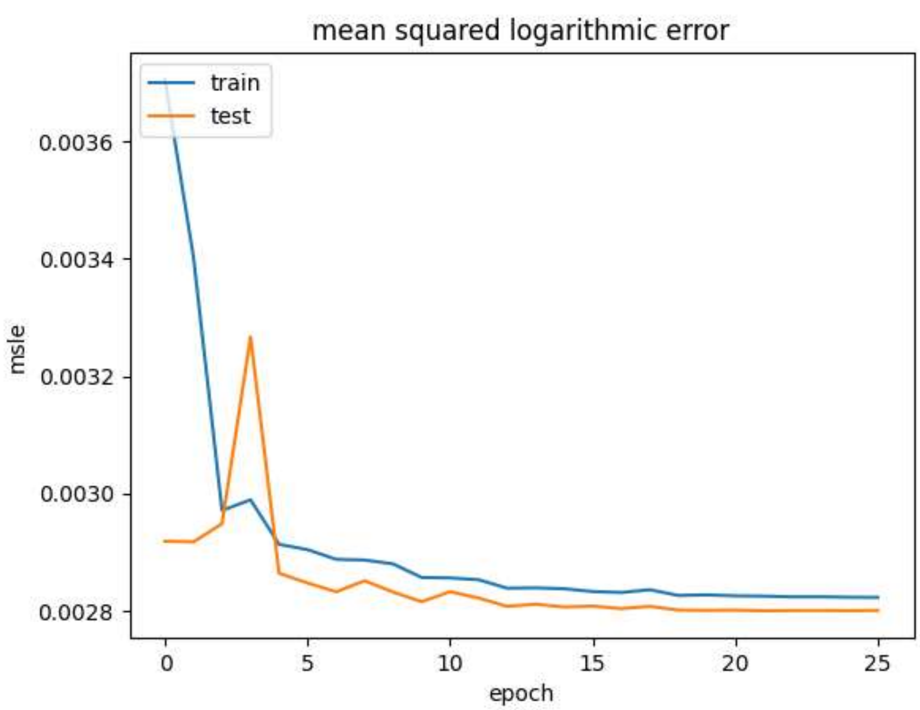


Fig. 10 (MSLE Graph for LSTM)

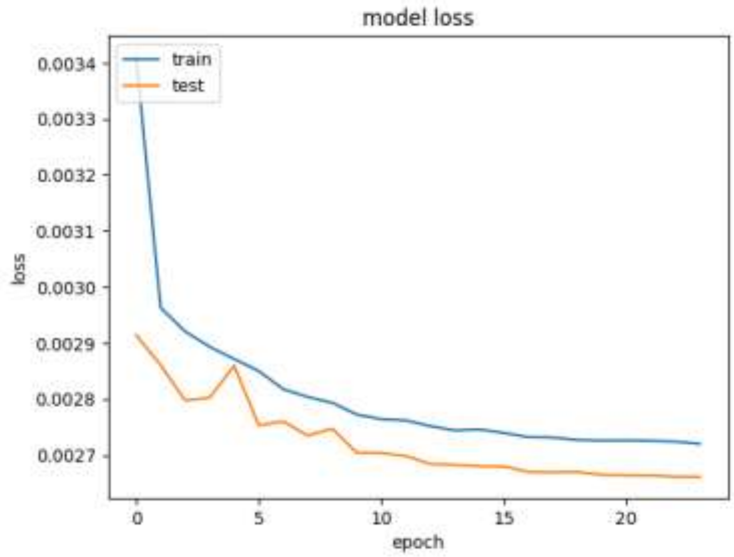


Fig. 11 (Actual vs Prediction graph for BiLSTM)

b. BiLSTM

Batch size of 64, with an Nadam Optimizer, learning rate of 0.001 and MSLE Loss function.

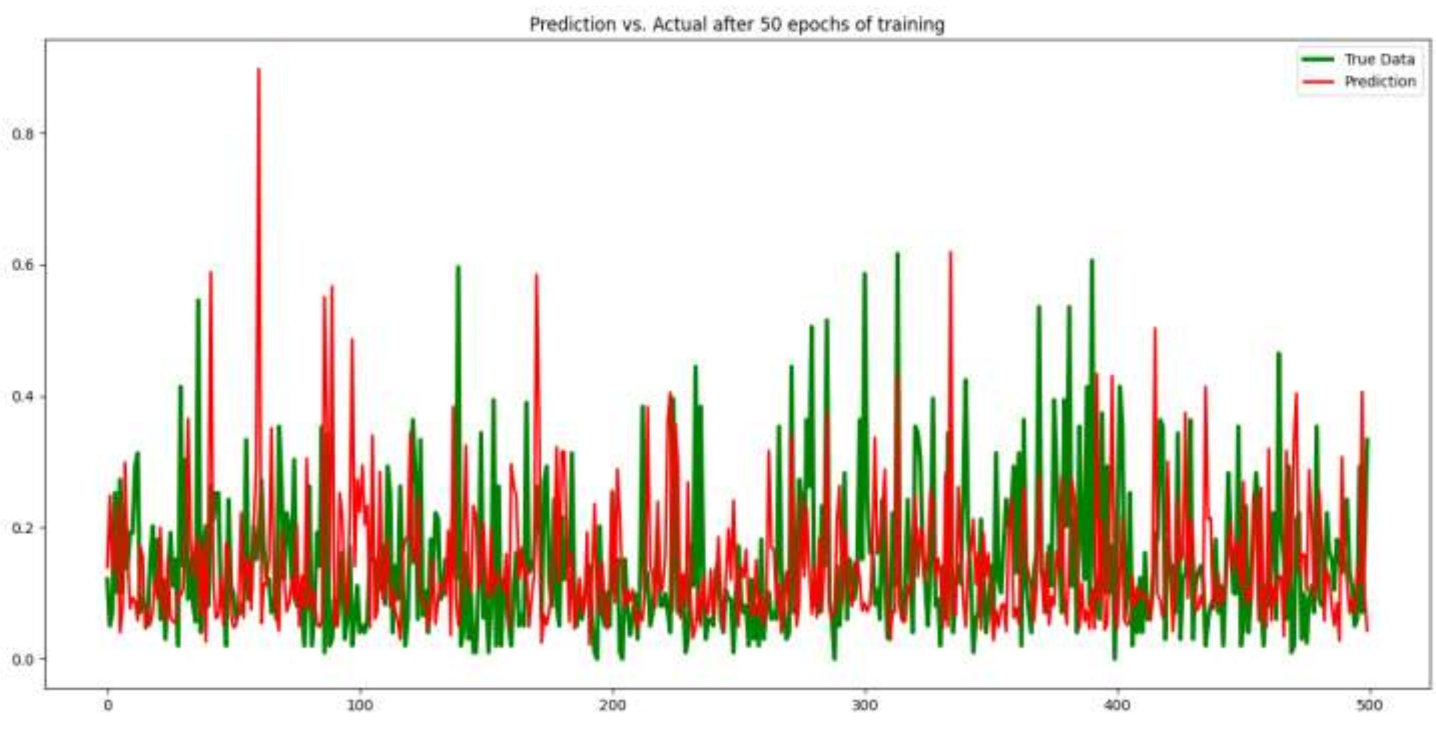


Fig. 12 (Loss Graph for BiLSTM)

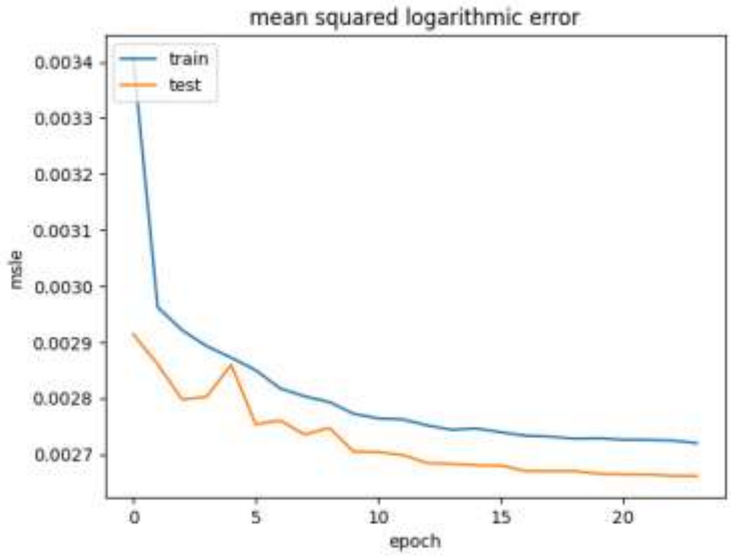


Fig. 13 (MSLE Graph for BiLSTM)

c. BiLSTMConv1D
Batch size of 64, with an Nadam Optimizer, learning rate of 0.001, LeakyReLU(alpha=0.7) activation function and MLSE loss function.

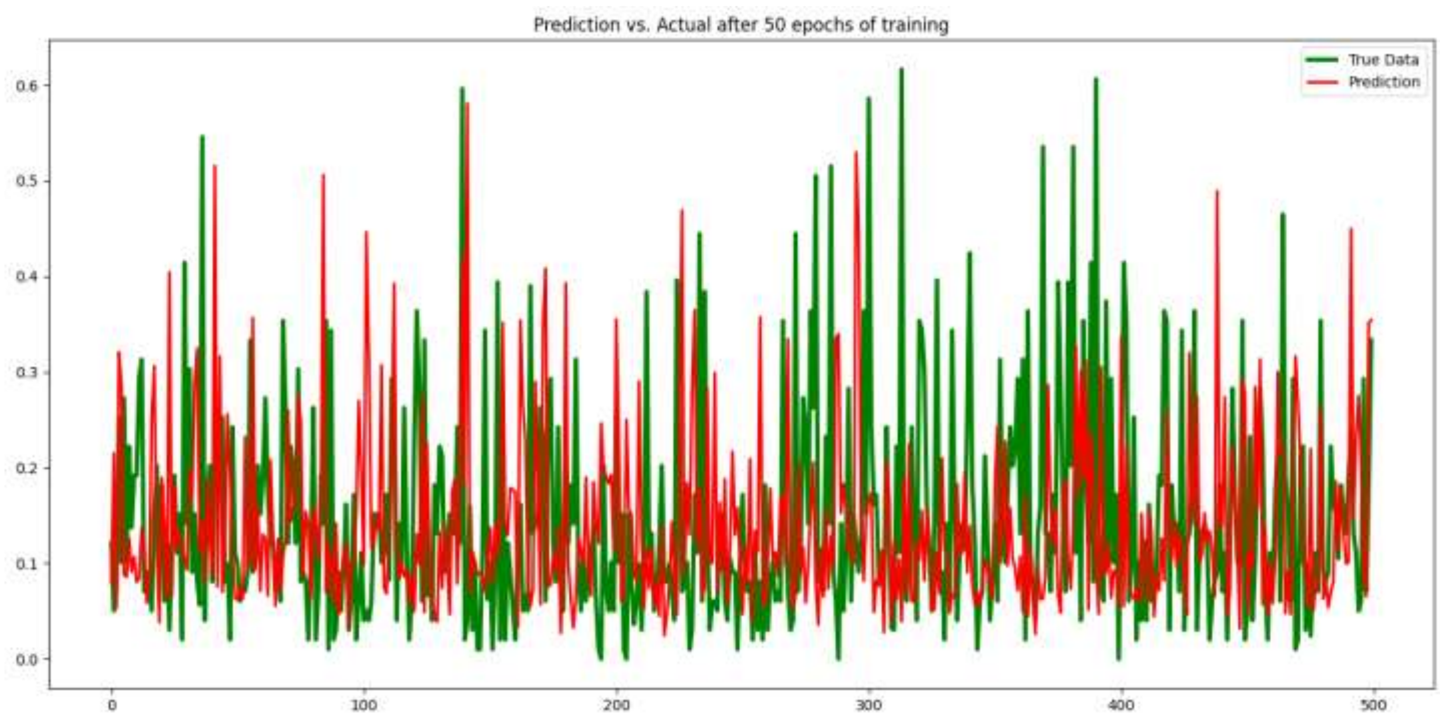
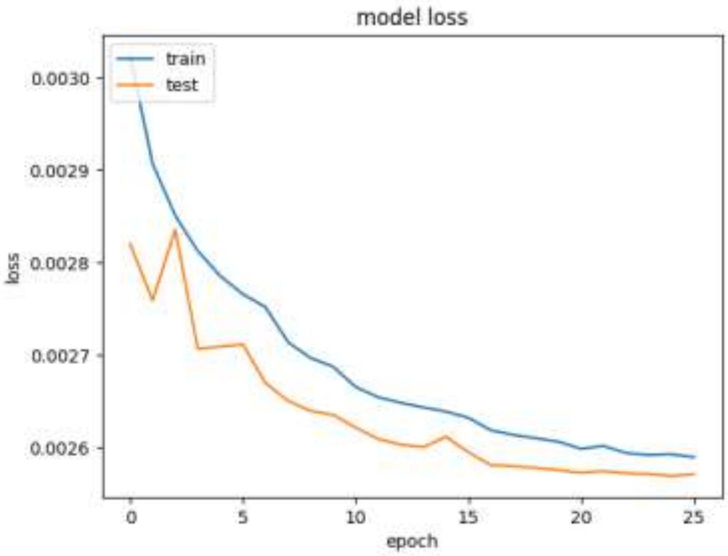


Fig. 14(Actual vs Prediction graph for BiLSTMConv1D)



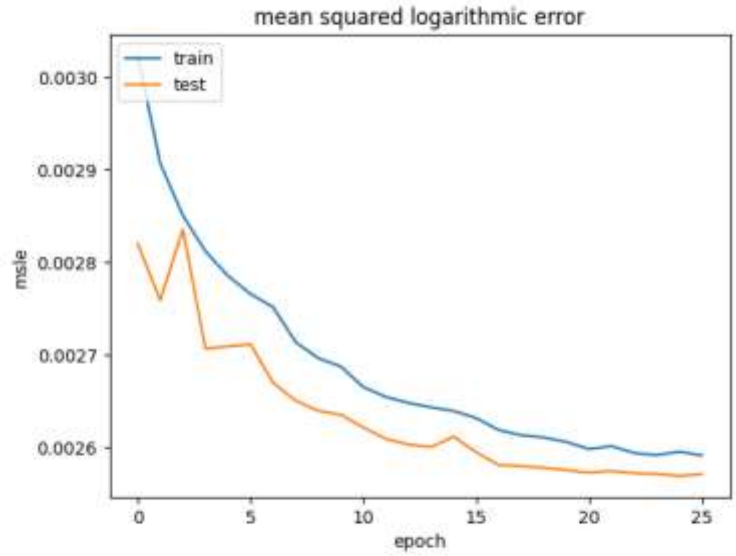


Fig. 15 (Loss Graph for BiLSTMConv1D) ***Fig. 16 (MSLE Graph for BiLSTMConv1D)***

d. GRU

Batch size of 64, with an SGD(lr=0.01, decay=1e-7, momentum=0.9, nesterov=False) Optimizer, learning rate of 0.01, tan(h) activation function and MSE loss function.

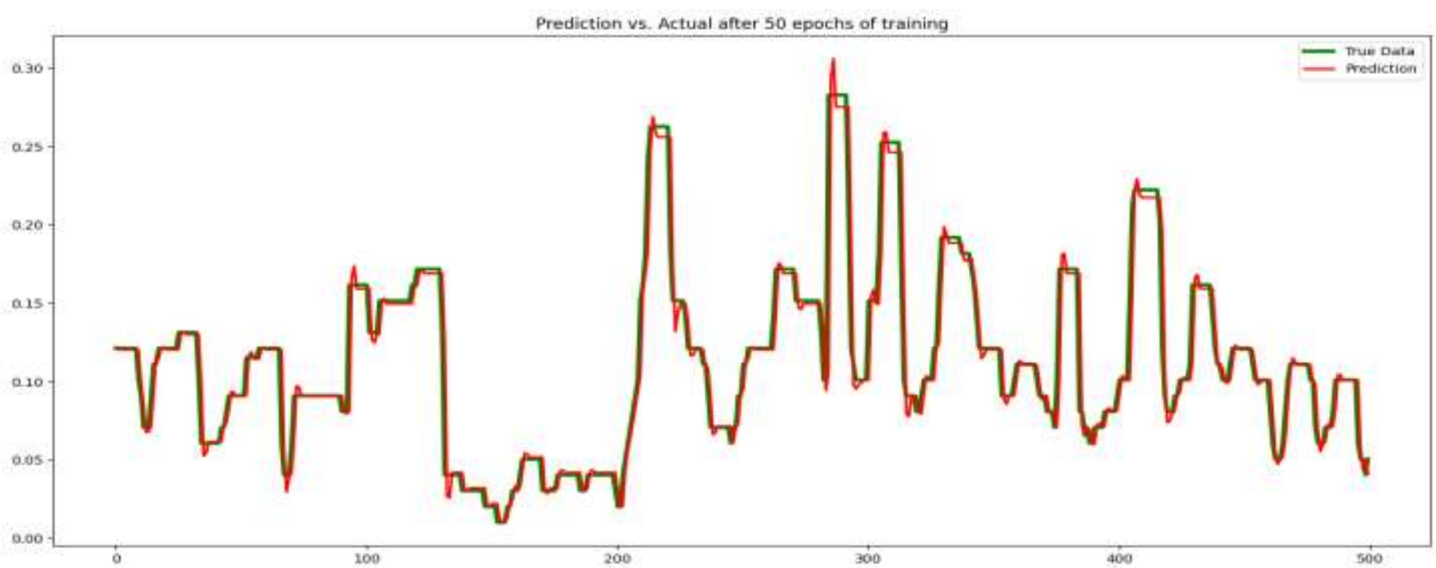


Fig. 17(Actual vs Prediction Graph for GRU)

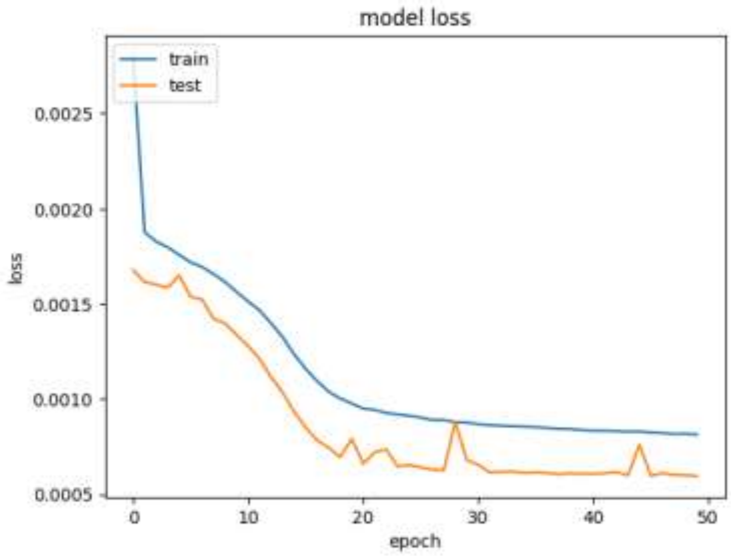


Fig. 18(Loss graph for GRU)

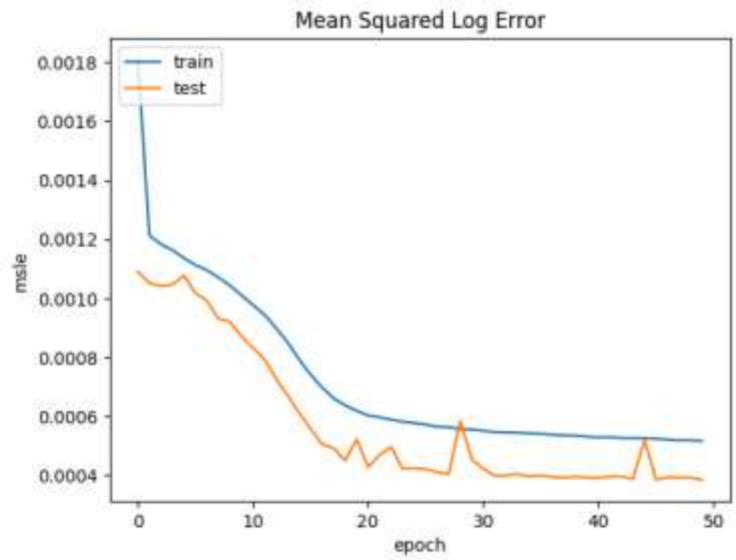


Fig. 19(MSLE graph for GRU)

LSTM is used for solving regression problems and since here we are using LSTM and its other derivatives we cannot use accuracy as a metric to compare the performance of the various models. As previously stated here we will be using MSE, RMSE, MAE, MSLE and a custom function named RMSLE. RMSLE can be provided as:

```
def rmsle_custom(y_true, y_pred):
msle = tf.keras.losses.MeanSquaredLogarithmicError()
return K.sqrt(msle(y_true, y_pred))
```

where K is tensorflow.keras.backend

First up if we compare the actual vs prediction graphs we can easily jump to the conclusion that GRU is completely outperforming all the other models given how close are the actual and prediction lines to coinciding. But for the other three models we can't be so positive about it. Next, let's compare the values of various errors that we initially mentioned. Herein again the GRU way outperformed the other three models presenting the lowest values among all i.e., MSE: 0.0006, MAE: 0.01116, RMSE: 0.0246. It also has the largest difference from the closest value from other models. If we keep GRU aside and compare the other three models we find that in terms of MSE, BiLSTMConv1D performs the best with a score of MSE: 0.2183. In terms of MAE, LSTM performs the best having score of MAE: 0.10525. In terms of RMSE, BiLSTMConv1D performs the best having score of RMSE: 0.14775. So, based on the above individual inspection we may say that BiLSTMConv1D is performing the best in absence of GRU having the lowest score in two parameters MSE and RMSE. Lastly if we analyze the MSLE values that we only have for LSTM and BILSTM we find LSTM performing slightly better than BiLSTM with a score of MSLE: 0.01455

7. Conclusion

Millions of people suffer from ailments caused by air pollution each year, making it a serious problem for contemporary civilization. A range of diseases, including ischemic heart disease, stroke, lower respiratory infections, chronic obstructive pulmonary disease, and lung cancer, are blamed for the premature deaths brought on by air pollution. International organizations need an accurate way to forecast air quality so they can take the required precautions to stop pollution’s exponential expansion. This study compared the performance of various Deep Learning models, including the LSTM, BiLSTM, BiLSTMConv1D, and GRU, in predicting the Air Quality Index (AQI). Our results indicate that the GRU model on average outperforms LSTM, BiLSTM and BiLSTMConv1D in AQI prediction with the optimal MSE, MAE, RMSE and Loss. This finding suggests that deep learning, specifically the GRU model, may be a promising approach for predicting AQI. Further research is needed to validate these findings and explore the potential applications of this model in real-world scenarios. Overall, this study demonstrates the potential of deep learning in predicting environmental variables, which can be valuable for monitoring air quality and making informed decisions for public health and safety. The findings of this study can also contribute to the development of more accurate and efficient environmental prediction models.

References

[1] Liao, Q., Zhu, M., Wu, L. *et al*. Deep Learning for Air Quality Forecasts: a Review. *Curr Pollution Rep* **6**, 399–409 (2020). https://doi.org/10.1007/s40726-020-00159-z

[2] P. K. Mall and P. K. Singh, “Explainable Deep Learning approach for Shoulder Abnormality Detection in X-Rays Dataset.,” Int. J. Next-Generation Comput., vol. 13, no. 3, 2022.

[3] P. K. Mall and P. K. Singh, “BoostNet: a method to enhance the performance of deep learning model on musculoskeletal radiographs X-ray images,” Int. J. Syst. Assur. Eng. Manag., pp. 1–15, 2022.

[4] P. K. Mall, P. K. Singh, and D. Yadav, “GLCM based feature extraction and medical X-RAY image classification using machine learning techniques,” in 2019 IEEE Conference on Information and Communication Technology, 2019, pp. 1–6.

[5] Myszczynska, M.A., Ojamies, P.N., Lacoste, A.M.B. *et al.* Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat Rev Neurol* **16**, 440–456 (2020). https://doi.org/10.1038/s41582-020-0377-8

[6] Wenqing Sun, Bin Zheng, and Wei Qian "Computer aided lung cancer diagnosis with deep learning algorithms", Proc. SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis, 97850Z (24 March 2016); https://doi.org/10.1117/12.2216307

[7] Nabipour, M.; Nayyeri, P.; Jabani, H.; Mosavi, A.; Salwana, E.; S., S. Deep Learning for Stock Market Prediction. Entropy 2020, 22, 840. https://doi.org/10.3390/e22080840

[8] Chatterjee, R., Chatterjee, A. & Islam, S.H. Deep learning techniques for observing the impact of the global warming from satellite images of water-bodies. Multimed Tools Appl 81, 6115–6130 (2022). https://doi.org/10.1007/s11042-021-11811-1

[9] S. E. Haupt, J. Cowie, S. Linden, T. McCandless, B. Kosovic and S. Alessandrini, "Machine Learning for Applied Weather Prediction," 2018 IEEE 14th International Conference on e-Science (e-Science), Amsterdam, Netherlands, 2018, pp. 276-277, doi: 10.1109/eScience.2018.00047

[10] E. Avşar , Y. C. Anar and A. Ö. Polat , "Parking lot occupancy prediction using long short-term memory and statistical methods", MANAS Journal of Engineering, vol. 10, no. 1, pp. 35-41, Jun. 2022, doi:10.51354/mjen.986631

[11] İ. Kök, M. U. Şimşek and S. Özdemir, "A deep learning model for air quality prediction in smart cities," 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 2017, pp. 1983-1990, doi: 10.1109/BigData.2017.8258144.

[12] Zhang, Z., Zeng, Y. & Yan, K. A hybrid deep learning technology for PM2.5 air quality forecasting. Environ Sci Pollut Res 28, 39409–39422 (2021). https://doi.org/10.1007/s11356-021-12657-8

[13] Mauro Castelli, Fabiana Martins Clemente, Aleš Popovič, Sara Silva, Leonardo Vanneschi, "A Machine Learning Approach to Predict Air Quality in California", Complexity, vol. 2020, Article ID 8049504, 23 pages, 2020. https://doi.org/10.1155/2020/8049504

[14] S. Du, T. Li, Y. Yang and S. -J. Horng, "Deep Air Quality Forecasting Using Hybrid Deep Learning Framework," in IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 6, pp. 2412-2424, 1 June 2021, doi: 10.1109/TKDE.2019.2954510.

[15] Ihianle, Isibor & Nwajana, Augustine & Ebenuwa, Solomon & Otuka, Richard & Owa, Kayode & Orisatoki, Mobolaji. (2020). A Deep Learning Approach for Human Activities Recognition From Multimodal Sensing Devices. IEEE Access. 8. 179028-179038. 10.1109/ACCESS.2020.3027979.

[16] Vamathevan, J., Clark, D., Czodrowski, P. et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18, 463–477 (2019). https://doi.org/10.1038/s41573-019-0024-5

[17] Abbas, K.; Afaq, M.; Ahmed Khan, T.; Song, W.-C. A Blockchain and Machine Learning-Based Drug Supply Chain Management and Recommendation System for Smart Pharmaceutical Industry. Electronics 2020, 9, 852. https://doi.org/10.3390/electronics9050852

[18] M. A. Sheikh, A. K. Goel and T. Kumar, "An Approach for Prediction of Loan Approval using Machine Learning Algorithm," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 490-494, doi: 10.1109/ICESC48915.2020.9155614.

[19] Parmar, Aakash & Mistree, Kinjal & Sompura, Mithila. (2017). Machine Learning Techniques For Rainfall Prediction: A Review.

[20] Cini, A., Marisca, I., & Alippi, C. (2021). Filling the g_ap_s: Multivariate time series imputation by graph neural networks. *arXiv preprint arXiv:2108.00298*.

[21] p Mall and P. Singh, “Credence-Net: a semi-supervised deep learning approach for medical images,” Int. J. Nanotechnol., vol. 20, 2022.

[22] Irfan, D., Tang, X., Narayan, V., Mall, P. K., Srivastava, S., & Saravanan, V. (2022). Prediction of Quality Food Sale in Mart Using the AI-Based TOR Method. Journal of Food Quality, 2022.

[23] Mall, P. K., Narayan, V., Pramanik, S., Srivastava, S., Faiz, M., Sriramulu, S., & Kumar, M. N. (2023). FuzzyNet-Based Modelling Smart Traffic System in Smart Cities Using Deep Learning Models. In Handbook of Research on Data-Driven Mathematical Modeling in Smart Cities (pp. 76-95). IGI Global.

[24] Mall, P. K., Yadav, R. K., Rai, A. K., Narayan, V., & Srivastava, S. (2023). Early Warning Signs Of Parkinson’s Disease Prediction Using Machine Learning Technique. Journal of Pharmaceutical Negative Results, 2607-2615.

[25] Narayan, Vipul, et al. "Deep Learning Approaches for Human Gait Recognition: A Review." 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE, 2023.

[26] Narayan, Vipul, et al. "FuzzyNet: Medical Image Classification based on GLCM Texture Feature." 2023 International Conference on Artificial Intelligence and Smart Communication (AISC). IEEE, 2023.

[27] Mall, Pawan Kumar, et al. "FuzzyNet-Based Modelling Smart Traffic System in Smart Cities Using Deep Learning Models." Handbook of Research on Data-Driven Mathematical Modeling in Smart Cities. IGI Global, 2023. 76-95.

[28] Pramanik, Sabyasachi, et al. "A novel approach using steganography and cryptography in business intelligence." Integration Challenges for Analytics, Business Intelligence, and Data Mining. IGI Global, 2021. 192-217.

[29] Irfan, Daniyal, et al. "Prediction of Quality Food Sale in Mart Using the AI-Based TOR Method." Journal of Food Quality 2022 (2022).

[30] Narayan, Vipul, et al. "To Implement a Web Page using Thread in Java." (2017).

[31] Narayan, Vipul, and A. K. Daniel. "Novel protocol for detection and optimization of overlapping coverage in wireless sensor networks." Int. J. Eng. Adv. Technol 8 (2019).

[32] Narayan, Vipul, and A. K. Daniel. "FBCHS: Fuzzy Based Cluster Head Selection Protocol to Enhance Network Lifetime of WSN." ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal 11.3 (2022): 285-307.

[33] Awasthi, Shashank, et al. "A Comparative Study of Various CAPTCHA Methods for Securing Web Pages." 2019 International Conference on Automation, Computational and Technology Management (ICACTM). IEEE, 2019.

[34] Narayan, Vipul, A. K. Daniel, and Ashok Kumar Rai. "Energy efficient two tier cluster based protocol for wireless sensor network." 2020 international conference on electrical and electronics engineering (ICE3). IEEE, 2020.

[35] Narayan, Vipul, et al. "E-Commerce recommendation method based on collaborative filtering technology." International Journal of Current Engineering and Technology 7.3 (2017): 974-982.

[36] Narayan, Vipul, and A. K. Daniel. "Design consideration and issues in wireless sensor network deployment." (2020): 101-109.

[37] Choudhary, Shubham, et al. "Fuzzy approach-based stable energy-efficient AODV routing protocol in mobile ad hoc networks." Software Defined Networking for Ad Hoc Networks. Cham: Springer International Publishing, 2022. 125-139.

[38] Narayan, Vipul, and A. K. Daniel. "Energy Efficient Protocol for Lifetime Prediction of Wireless Sensor Network using Multivariate Polynomial Regression Model." Journal of Scientific & Industrial Research 81.12 (2022): 1297-1309.

[39] Srivastava, Swapnita, and P. K. Singh. "Proof of Optimality based on Greedy Algorithm for Offline Cache Replacement Algorithm." International Journal of Next-Generation Computing 13.3 (2022).

[40] Smiti, Puja, Swapnita Srivastava, and Nitin Rakesh. "Video and audio streaming issues in multimedia application." 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2018.

[41] Srivastava, Swapnita, and P. K. Singh. "HCIP: Hybrid Short Long History Table-based Cache Instruction Prefetcher." International Journal of Next-Generation Computing 13.3 (2022).

[42] Srivastava, Swapnita, and Shilpi Sharma. "Analysis of cyber related issues by implementing data mining Algorithm." 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2019.

[43] Smriti, Puja, Swapnita Srivastava, and Saurabh Singh. "Keyboard invariant biometric authentication." 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT). IEEE, 2018.

[44] Pentyala, S., Liu, M., & Dreyer, M. (2019). Multi-task networks with universe, group, and task feature learning. arXiv preprint arXiv:1907.01791.