*Proceedings of the 9th World Congress on New Technologies (NewTech'23)*
*Brunel University, London, United Kingdom - August 09-11, 2023*
*Paper No. ICEPR 143*
*DOI: 10.11159/icepr23.143*

# A Multi-Layer Perceptron Approach for Estimating Daily Surface $NO_2$ In Thiruvananthapuram City

**Sherin Babu[1,2], Binu Thomas[1,3]**
[1]School of Computer Sciences, Mahatma Gandhi University, Kottayam, Kerala, India
[2]Department of Computer Science, Assumption College Autonomous, Changanassery, Kottayam, Kerala, India
sherinbabu@assumptioncollege.edu.in
[3]Department of Computer Applications, Marian College, Kuttikkanam, Idukki, Kerala, India
binu.thomas@mariancollege.org

## Extended Abstract

In this study, a machine-learning framework based on a multi-layer perceptron (MLP) algorithm is applied to estimate the daily values of air pollutant $NO_2$ in Thiruvananthapuram city of Kerala, India. The risk of human respiratory tract infections rises when exposed to high amounts of $NO_2$ [1]. Due to urbanization and its consequences, the air quality in the study region is getting deteriorated [2]. As a result, there is a pressing need for research and estimation of air pollutants like $NO_2$ in Thiruvananthapuram city.

MLP is a supervised neural network model that is frequently used and it gains experience by learning to simulate the correlation between a set of input-output pairs [3]. This paper proposes a four-layer (i.e. one input, two hidden and one output) multi-layer perceptron neural network model for predicting the daily surface $NO_2$ values. Two year daily data (January 2018 to December 2019) is collected from Central Pollution Control Board, Government of India. The study utilizes 8 air pollutant parameters ($PM_{10}$, $PM_{2.5}$, $SO_2$, NO, $NO_x$, $NH_3$, CO and Ozone) and 7 meteorological parameters (wind speed, wind direction, air temperature, solar radiance, relative humidity, atmospheric pressure and rainfall) in the model development. Due to instrumental errors, certain data are missing and such missing daily data records are excluded from the dataset. Six hundred and eight one daily records are available in the dataset, in which 544 (80%) samples and 137 (20%) samples are divided as training data and testing data, respectively. Additionally, investigations are conducted to evaluate the effects of a limited number of input parameters into the model by reducing the size of the features with Reduced Feature Elimination (RFE) method [4]. By changing the layer size in the neural network design, eight distinct experiments are performed.

The performances of the developed MLP models are evaluated using three performance metrics - coefficient of determination ($R^2$), root mean square error (RMSE), and mean absolute error (MAE) [5]. The best model performance obtained is having $R^2 = 0.968$, MAE = 0.462 μg/m3, and RMSE = 1.028 μg/m3 using the neural network size (3-64-64-1), incorporating the elimination of unimportant features using RFE (included 3 input features – NO, $NO_x$ and $SO_2$ only; excluding all other input features). The results indicate that the models with the reduced feature set performed better than the models with the full feature set. The analysis also demonstrates that the prediction accuracy increases with the number of hidden layers in the neural network.

## References

[1] Hesterberg, Thomas W., William B. Bunn, Roger O. McClellan, Ali K. Hamade, Christopher M. Long, and Peter A. Valberg. "Critical review of the human data on short-term nitrogen dioxide (NO2) exposures: evidence for NO2 no-effect levels." *Critical reviews in toxicology* 39, no. 9 (2009): 743-781.

[2] Kumar, Amogh P., Kshama AV, Paul Lazarus, Salma Muslim, and Santha AM. "An analysis of air pollution in kerala." *Journal of Industrial Pollution Control* 34, no. 1 (2018): 1917-1921.

[3] Fan, Xinghua, Li Wang, and Shasha Li. "Predicting chaotic coal prices using a multi-layer perceptron network model." *Resources Policy* 50 (2016): 86-92.

[4] Yan, Ke, and David Zhang. "Feature selection and analysis on correlated gas sensor data with recursive feature elimination." *Sensors and Actuators B: Chemical* 212 (2015): 353-363.

[5] Şahin, Fatma, Gültekin Işik, Gökhan Şahin, and Mehmet Kazım Kara. "Estimation of PM10 levels using feed forward neural networks in Igdir, Turkey." *Urban Climate* 34 (2020): 100721.