Air Quality Prediction Using Machine Learning

RM Fernando[#], WMKS Ilmini, and DU Vidanagama

Faculty of Computing, General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka

[#]35-CS-18-0001@kdu.ac.lk

Abstract - The main basis of human survival is Air. The Air Quality Index is the value that qualitatively describes the condition of air quality. The greater the Air Quality Index, the more threatening risk to human health and environment. In Sri Lanka, poor air quality is a huge concern, especially in cities like Colombo and Kandy. Accurate Air Quality prediction will minimize health issues that can occur due to air pollution. This research has attempted to identify the best-suited machine learning algorithm-based approach to predict accurate air quality based on PM2.5 concentration in Colombo. In order to identify the most influenced air pollution concentrations for the air quality prediction purpose, correlation analysis was conducted. In this research, PM2.5 was predicted in Colombo city using 4 related air pollution concentrations including SO2 concentration, NO2 concentration, PM2.5 concentration & PM10 concentration. In order to get higher prediction accuracy, the gathered dataset was pre-processed by prediction beforehand. The prediction model trained and tested using machine learning algorithms such as KNN, Multiple Linear Regression, Support Vector Machines, and Random Forest. Multiple Regression was identified as the most suited prediction model which was able to gain 94% higher accuracy.

***Keywords:* air quality, concentration, correlations, machine learning, pollution**

I. INTRODUCTION

Air Quality is a massive luxury to have & unfortunately most people who live on the earth won't have that luxury. According to the World Health Organization, most people live in poor air quality areas. In recent years, Colombo air quality getting poor & poor year-wise. Most times in rush hours Colombo air quality exceeds level four category of the air quality index. Unfortunately, most people who live in the Colombo area don't aware about what is the air quality around them (Mahanta et al., 2019).

In this Research paper, we explore predicting air quality using algorithms of machine learning. Machine learning algorithms are implemented & evaluated for accuracy. The research presented to improve air quality index prediction methods & air quality knowledge in Colombo. The air quality index value in Colombo is average at an unhealthy level, especially in rush hours. With reliable predictions of air quality index levels, people can get precautionary measures, such as minimizing outdoor activities, to minimize the consequences of air pollution. This introduction chapter provides an overview of the air quality index & the motivation for this project, followed by background details about air quality in Colombo & the research goal (Castelli et al., 2020). The availability of clean air to breathe has been of very importance & luxury to have. As air is important for all the living beings on the planet, it is human responsibility to protect the air quality. Industrialization & the increase of vehicles has led the earth into air pollution & an environmental curse. Air pollution basically refers to the high contamination of the air by large quantities of very harmful chemicals, gas & dust substances. Mostly air pollution occurs due to the use of energy & emissions from production, where emissions from vehicles & industries are main contributors to the cause. Air pollution is a threat to human survival due to the impact on human health & the environment. Mostly Urban cities like Colombo normally have the worst air pollution compared to the rural areas due to large human actions within a small area of land. Clear correlations between ambient air pollution & effects on human health have been identified, which includes both long-term & short-term effects on human health issues as well as their living environment. Increase in heart diseases & reduced lung functioning, direct impact on public with asthma & many other types of pneumonia & once air inhaled, the particular matters like PM2.5 & PM10 may very hard to be self-purified by the human immune system (Zhong, Yu and Zhu, 2019).

Air pollution can be considered as one of the most critical factors that affect human survival. Every year air pollution is responsible for millions of deaths worldwide. Not only for human health but also for the environment, air pollution brings negative effects including acid rain, global warming, etc. According to many public perception studies, the main problem is the lack of awareness about the air pollution causes & effects. Therefore, with the increase of air pollution, it has become an important endeavor to predict & aware the people about the effect of air pollution levels on human health & the environment (Nandasena, Wickremasinghe and Sathiakumar, 2010). Predicting the air quality index can avoid the worst effects on human health issues & the environment. Find out the best way to predict air quality index from using various techniques and methods, training the data set to get the most accurate prediction necessary to the success of this research. The main basis of human survival is Air. The air quality index or short firmly AQI is the value that describes qualitatively the condition of air quality. The greater the air quality index the more threatening risk to human health & the environment. The key factors that mainly cause the AQI are NO2, SO2, O3, CO, PM2.5, & PM10. In this research, a past air pollution concentration dataset has gathered from the Central Environment Authority which consists of hourly concentrations of different air pollution parameters & weather parameters such as Solar Radiation, Relative Humidity, Average Temperature, Wind direction, Wind speed, O3, CO, NO2, SO2, PM2.5 & PM10. A number of data preprocessing methods have used to ensure the accuracy of the predicted outcome. A Cross-Validation has done for the preprocessed dataset, by partitioning the data set as 80% for model training & 20% for testing. Several machine learning algorithms have been used as a prediction model such as Multiple Linear Regression, Support Vector Machine, K Nearest Neighbors & Random Forest. Based on the accuracy & performance, the most suitable model for air quality prediction is identified.

This research paper is ordered as follows. The second part consists of the Literature Review for this research. The methodology used in this research & Results has been identified as demonstrating in the third & fourth parts respectively. Finally, the last part consists of the conclusion of this research.

II. LITERATURE REVIEW

There are many air quality index predicting systems available. Few of those provide a common indication of human health diseases according to the level of PM2.5 particles in the air. With the lack of an exact framework within the area of research air quality index prediction, with various problem identifications, locations, & datasets in the research studies, a literature review is conducted to identify an overview of the air quality literature. The research area varies in different techniques & methods, but also, the available datasets are more often different due to the traffic, climate, & environment of the selected geographical area together with the selected air pollutants to predict. For some cities, the poor air quality due to mainly PM-related air pollutants causes, while in other cities the poor air quality might mainly come from NOx, SOx, or COx. Because of these limitations, the literature review is an attempt to get a good understanding of the research scope & find relevant research that doing the same task as in this research. This chapter defines a set of the latest relevant air quality studies for this thesis. Here is a review about exiting air quality index predicting systems. Which method is more suitable to predict air pollution is also an important factor. The deep learning method for air quality prediction is one approach most commonly used in existing systems (Xayasouk and Lee, 2018). As an example, Korean air pollution prediction systems use the Stacked Autoencoders prediction model for training & learning datasets. The predicted output shows the overall performance of the air quality prediction using Machine Learning algorithms. Machine learning approach is the most popular technique when predicting the air pollution. Machine learning techniques able to train a model using big data and algorithms (Iskandaryan, Ramos and Trilles, 2020). Data trained by using regression models and regularizations like nuclear norm regularization and standard Frobenius regularization. One of the experiments had already shown that consecutive regularization and parameter formulation achieve far better performance than existing regularizations and standard regression models. Another technique that can be used to predict air pollution is using a neural network. Accurate predictions of Air quality index is possible with the simple neural network and further modifications of the model achievable using different experimental setups and different input parameters (Sampath, 2019).

Another research paper, machine learning prediction model for AQ prediction for urban cities. In this paper, the Author mentions Air pollution vastly remains a huge challenge for people & governments all around the globe. Air pollution can cause noticeable effects to human health as well as on the environment resulting in global warming, acid rain, skin cancer, and heart problems to the public. This research study addresses the issues of predicting Air Quality, with the focus to minimize air pollution in cities before air pollution gets impacts human health and the environment, using two Machine Learning algorithms, SVM, & Neural Networks(NN). The Machine Learning (ML) model is supposed to predict the AQI. Predicted results will show an increase in the Air Quality prediction outcomes accuracy & recommend that the machine learning approach can be suitable in predicting other city's air quality as well (Bellinger, 2017). This research paper proposed a system using a Machine Learning approach for AQI prediction for big cities. The machine learning model is evaluated with the New Delhi Air pollution data mainly due to the fact that poorness of New Delhi’s air quality. Using the Support Vector Machines and Neural Networks, Air Quality Index is predicted accurately by using two machine learning models with 91.62% higher accuracy for the Neural Networks model & 97.3% accuracy for the SVM model. Six of the SVM algorithm functions were identified to predict Air Quality Index accuracy, & finally, it was identified that the “Gaussian Support Vector Machines” gives the highest accuracy value of 97.3%(Mahalingam et al., 2019).

According to the research work of Timothy & Dela Cruz for predicting the Air Quality Index requires decisive & perfect readings & more complex calculations, therefore it is not recommended portable predicting devices. The main aim of the research is to identify another way of characterizing & monitoring to obtain solutions to minimize the effects of poor air quality. Five predictive machine learning models have developed, K-nearest neighbors, support vector machine, random forest, neural network, and Naïve Bayesian classifier. Results of the paper clearly mention that the research team obtains accuracies of 97.78%, 98.67%, 94.22%, 99.56%, and 98.67% for the five machine learning models respectively, clearly having the model of a neural network(NN) be the perfect accuracy model (Amado and Dela Cruz, 2018)

Table 1. Summary of the Literature Review

Author	**Application**	**Technique**	**Remark**
Sara Silva & others	Air quality prediction for smart cities	•Support vector regression	•Predict PM 2.5 levels variability. •Model is suitable for predict hourly air pollution. •Obtain an accuracy of 94.1%
Usha Mahalingam & others	Air Quality prediction	•Neural Networks •Support vector machine	•Accuracy of 91.62% for neural network •Accuracy of 97.3% for support vector machine
Min Lee & others	Air pollution prediction	•Deep Learning	•predict against PM 2.5, PM 10 particulars. •Accuracy based on PM 10 is very low. •Accuracy based on PM 2.5 is very high.
Timothy M.	Air quality	• Naïve Bayesian	•Highest accuracy was obtained

Amado & others	monitoring models development	classifier • KNN • SVM • Neural network •Random forest	through Neural Networks. •Sometimes Neural Networks leads to slower responses.
Chen Zhao & others	Air Quality Index Prediction	•Linear regression	•Prediction based on one-year data of PM10, PM2.5, etc. •There is a deviation between predicted results and actual data.
Esmail Ahmadi	Air pollution prediction	•Data Mining •Decision Tree	•Used Clementine software for data clustering. •Data sample include climate data of 53 years
Niraj Tailor & others	Predict Air Quality in Urban cities Using Regression techniques for analysis	•Linear regression •Neural Networks •Lasso regression •Elastic Net regression •Ridge regression •Extra Trees •XGBoost •Decision Forest •Boosted tree • KNN	•84.68% accuracy for Linear regression •82.52% accuracy for Neural networks •84.77% accuracy for Lasso regression •84.772% accuracy for ElasticNet •84.89% accuracy for Decision Forest •85.31% accuracy for Extra Trees •83.89% accuracy for Boosted Tree •84.56% accuracy for XGBoost •69.48% accuracy for KNN •84.68% accuracy for Ridge regression •Prediction based on Weather & AQI datasets.
Colin Bellinger & others	A systematic review of Machine Learning & data mining for Air Pollution	•Machine Learning Algorithms •Data Mining •Big Data	•Refer 400 research papers & reduce to 47 after the inclusion/exclusion criteria's •Divided papers into 3 categories •End of the survey that highest accuracy levels always obtain in Machine Learning Algorithms.

According to the literature review, a few drawbacks exist in the available air quality prediction approaches, such as theproblems that can occur when collecting datasets. Inaccuracy & amount of null values have affected the predicted output low accuracies of the existing air quality prediction systems in Sri Lanka. Another factor that affects accuracy reductions in the data preprocessing. Considering all these factors, most of the existing systems in Sri Lanka have been failed to obtain a correct prediction. Considering these problems with foreign countries they are able to achieve these drawbacks & gain high accuracies. Also, air quality prediction systems based on machine learning have faced the issue of selecting the most suitable algorithm. Most used machine learning algorithms are not the appropriate algorithms for the target research area.

III. METHODOLOGY

The proposed structure of predicting the air quality index includes sequence of steps. Those sequence of steps include Gather data, pre-process the dataset, analysis the dataset, use suitable machine learning algorithms & finally findout the best suitable machine learning algorithm & analys the result.

A. Data Gathering & Pre-processing

In Data Gathering & Pre-processing phase for the first part, past dataset of air pollution parameters hourly concentration data in Colombo is gathered from Central Environment Authority & National Building research organization. Dataset Consists of hourly concentration of air pollution parameters like PM2.5, PM10, NO2, SO2, CO & humidity from 2019 January to 2021 February. To get higher accuracy & ensure the quality of the predicted values, the gathered dataset is pre-processed using preprocess technique.

B. Data Analysis

In order to find out the correlations between air pollution parameters & identify the distribution of the dataset & identify the nature of the dataset correlation matrices & distribution graphs are used. For the data analysis R studio IDE is used. Using correlation matrices & distribution graphs can be used to identify the most affectable parameter for PM2.5.

C. Testing

Under the cross validation technique, the Train Test Split method is the most common method which is used for the already pre-process data, by splitting dataset into two sets as 80% for training the prediction model & 20% for the testing the predicted results.

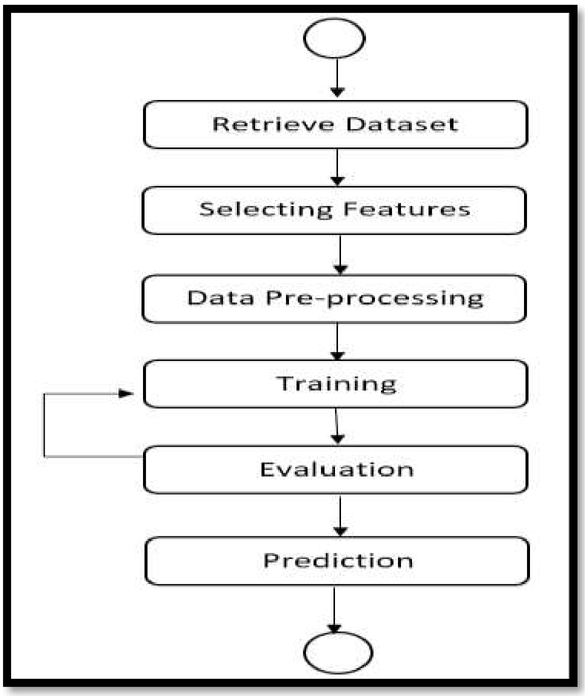


Figure 1. Overall Architecture

D. Training Model

- Multiple Linear Regression
- K Nearest Neighbors
- Random Forest
- Support Vector Machine

Machine Learning algorithms are used to train the dataset. For each of these cases default parameters are used. For the implementation, python based scikit learn, Pandas libraries are used & pycharm IDE also used.

E. Model Evaluation

After finishing the model training phase, model is used to predict PM2.5 value based on pre-processed dataset. Based on the accuracy most suitable machine learning algorithm is selected.

IV. RESULTS

In this research study, the gathered dataset includes nearly 15000 data records & 12 Air pollution concentrations & weather attributes, such as Solar Radiation, Relative Humidity, Average Temperature, Wind direction, Wind speed, O3, CO, NO2, SO2, PM2.5 & PM10. According to the Figure 2 correlation matrix chart, the correlation matrix has computed using through R, between PM2.5 & PM10 have the highest correlation among than each & other. According to the correlation matrix PM10, SO2, NO2 & CO has the highest correlation with the PM2.5 with compared to other air pollution & weather attributes. Therefore, to train the prediction model we considered PM2.5, PM10, NO2, SO2 & CO parameters.

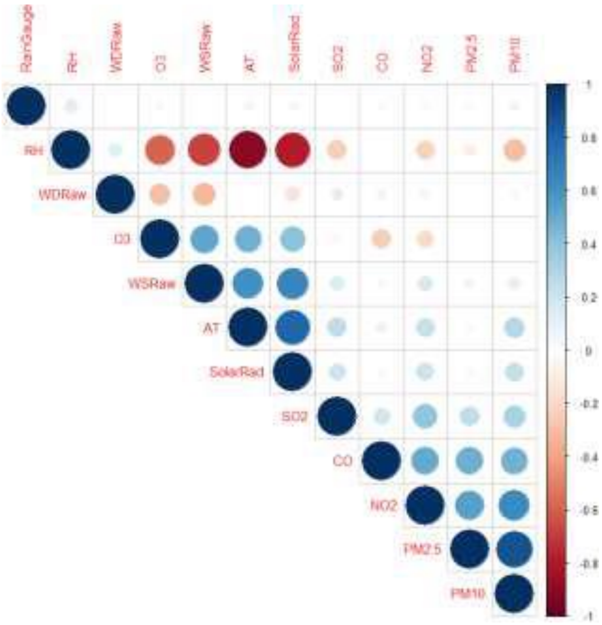
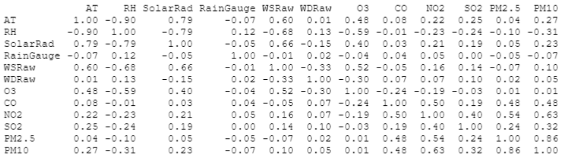


Figure 2. Correlogram

Out of all those Air pollution concentrations & weather parameters, four air pollution parameters were selected from the gathered dataset after a complete correlation analysis. They are NO2 concentration, SO2 concentration, PM2.5 concentration & PM10 concentration.



```
                AT    RH SolarRad RainGauge WSRaw WDRaw    O3    CO   NO2   SO2 PM2.5  PM10
AT            1.00 -0.90     0.79     -0.07  0.60  0.01  0.48  0.08  0.22  0.25  0.04  0.27
RH           -0.90  1.00    -0.79      0.12 -0.68  0.13 -0.59 -0.01 -0.23 -0.24 -0.10 -0.31
SolarRad      0.79 -0.79     1.00     -0.05  0.66 -0.15  0.40  0.03  0.21  0.19  0.05  0.23
RainGauge    -0.07  0.12    -0.05      1.00 -0.01  0.02 -0.04  0.04  0.05  0.00 -0.05 -0.07
WSRaw         0.60 -0.68     0.66     -0.01  1.00 -0.33  0.52 -0.05  0.16  0.14 -0.07  0.10
WDRaw         0.01  0.13    -0.15      0.02 -0.33  1.00 -0.30  0.07  0.07  0.10  0.02  0.05
O3            0.48 -0.59     0.40     -0.04  0.52 -0.30  1.00 -0.24 -0.19 -0.03  0.01  0.01
CO            0.08 -0.01     0.03      0.04 -0.05  0.07 -0.24  1.00  0.50  0.19  0.48  0.48
NO2           0.22 -0.23     0.21      0.05  0.16  0.07 -0.19  0.50  1.00  0.40  0.54  0.63
SO2           0.25 -0.24     0.19      0.00  0.14  0.10 -0.03  0.19  0.40  1.00  0.24  0.32
PM2.5         0.04 -0.10     0.05     -0.05 -0.07  0.02  0.01  0.48  0.54  0.24  1.00  0.86
PM10          0.27 -0.31     0.23     -0.07  0.10  0.05  0.01  0.48  0.63  0.32  0.86  1.00
```

Figure 3. Correlation Matrix Char

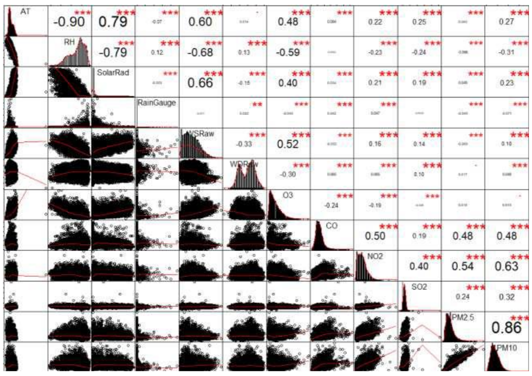


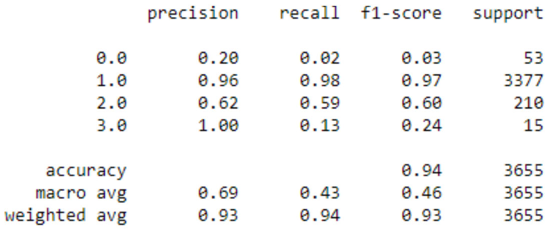
Figure 4. Correlation Matrix

As represented in Figure 3 and Figure 4, correlations between PM2.5 & some weather & air pollution concentration parameters are weak other than PM10, NO2, SO2 & CO. To get a higher correlation value, computed the correlations among PM2.5 & multiple air pollution concentration as represented in Figure 5. The correlation among PM2.5 & the combination of PM10, CO, NO2, SO2 is 0.8644 which is an excellent value.

Table 2. Summary of Multiple Correlation

	PM2.5
PM10 + NO2	0.8635649
PM10 + NO2 + CO	0.8623043
PM10 + NO2 + CO + SO2	0.8644263
PM10 + NO2 + SO2	0.8642185

A. Multiple Regression



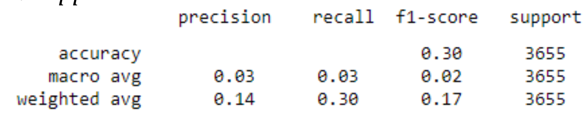
```
              precision    recall  f1-score   support

         0.0       0.20      0.02      0.03        53
         1.0       0.96      0.98      0.97      3377
         2.0       0.62      0.59      0.60       210
         3.0       1.00      0.13      0.24        15

    accuracy                           0.94      3655
   macro avg       0.69      0.43      0.46      3655
weighted avg       0.93      0.94      0.93      3655
```

Figure 5. Linear Regression Classification

According to the represented matrix in Figure 6 & the

classification chart in figure x, the predicted accuracy of the output is 94% which is a great accuracy. The predicted accuracy of the regression model is always influenced by the nature of the gathered dataset, for this dataset & prediction process multiple regression is suited nicely.

B. Support Vector Machines

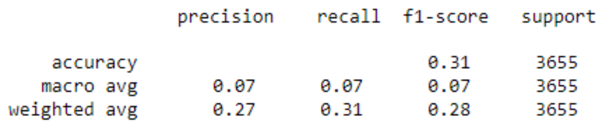


```
              precision    recall  f1-score   support

    accuracy                           0.30      3655
   macro avg       0.03      0.03      0.02      3655
weighted avg       0.14      0.30      0.17      3655
```

Figure 6. SVM Classification

As represented in the classification chart in Figure 7 the Support Vector Machine model has achieved 30% accuracy. This SVM model has got a very low accuracy when compared with the regression model. The reason behind achieving much low accuracy is Support Vector Machines handle inputs with polynomial features, for this prediction work SVM might not be suited.

C.Random Forest

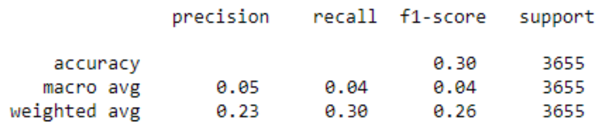


```
              precision    recall  f1-score   support

    accuracy                           0.31      3655
   macro avg       0.07      0.07      0.07      3655
weighted avg       0.27      0.31      0.28      3655
```

Figure 7. Random Forest Classification

Random Forest algorithm was also considered for this approach. Random Forest supervised learning algorithm can be used as both classification problems and regression problems as well as Random Forest is not difficult in calculating the relative importance of every feature that consists of the prediction. As represented in the report in Figure 8. Random Forest model has achieved 31% accuracy. This 31% accuracy is also very low when compared with the Multiple Regression model accuracy.

D.KNN



```
              precision    recall  f1-score   support

    accuracy                           0.30      3655
   macro avg       0.05      0.04      0.04      3655
weighted avg       0.23      0.30      0.26      3655
```

Figure 8. KNN Classification

KNN is another model which has used to predict air quality index in many research studies. As represented in the report in Figure 9, the KNN prediction model has achieved 30% accuracy when k=5. Since some air pollution parameters are very weak, it is very difficult to gain a higher accuracy from this prediction model.

V. DISCUSSION

In this research, we have achieved 30% accuracy for KNN & SVM models, 31% accuracy in Random Forest & 94% accuracy for the Regression model. According to Table1, many existing foreign studies obtain higher accuracies in using KNN, SVM & Random Forest algorithms. The main reason that has affected to the low accuracies obtained by these three machine learning algorithms is the incompleteness of training dataset. Missing values and noisy features that exist with the dataset affect the accuracy of the results. Different data pre-processing techniques have been followed in order to increase the quality of the training dataset.

Since the Multiple Regression model gives the highest overall best accuracy compared to the SVM, Random Forest & KNN models, the Multiple Regression model can be identified as the best-suited model for this prediction process.

Table 3. Summary table of the Model Evaluation

Model	**Accuracy**
Multiple Regression	94%
Support Vector Machines	30%
Random Forest	31%
KNN	30%

VI. CONCLUSION

The main basis of human survival depends on Air. The air quality index or short firmly AQI is the value that describes qualitatively the condition of air quality. The greater the air quality index the more threatening risk to human health & the environment. Air Pollution always caused by due to human actions. In Sri Lanka, poor air quality is a huge concern especially in cities like Colombo & Kandy. Accurate Air Quality prediction will minimize the health issues that can occur due to air pollution. This research has attempted to identify the best-suited machine learning algorithm-based approach to predict accurate air quality based on PM2.5 concentrations in Colombo. In order to identify the most influenced air pollution concentrations for the air quality prediction purpose. In this research, PM2.5 was predicted in Colombo city using 4 related air pollution concentrations including SO2 concentration, NO2 concentration, PM2.5 concentration & PM10 concentration. In order to get higher prediction accuracy, the gathered dataset was pre-processed by prediction beforehand. For the prediction model, cross-validated data according to 80 to 20. The prediction model trained & tested using machine learning algorithms such as KNN, Multiple Linear Regression, Support Vector Machines, & Random Forest. For the model evaluation, Multiple Regression was identified as the most suited prediction model which was able to gain 94% of higher accuracy.

VII. FUTURE WORK

For future work, expected to gather more datasets from air quality monitoring stations in Sri Lanka & apply more suitable pre-processing methods for the dataset. Since some machine learning models have low accuracy levels, the research team plans to build a deep learning prediction model for this approach on the prediction of air quality.

REFERENCES

Amado, T. M. and Dela Cruz, J. C. (2018) 'Development of
Machine Learning-based Predictive Models for Air Quality
Monitoring and Characterization', in TENCON 2018 - 2018 IEEE Region 10 Conference. TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South): IEEE, pp. 0668–0672. doi: 10.1109/TENCON.2018.8650518.

Bellinger, C. (2017) 'A systematic review of data mining and
machine learning for air pollution epidemiology', p. 19.

Castelli, M. et al. (2020) 'A Machine Learning Approach to Predict Air Quality in California', Complexity, 2020, pp. 1–23. doi: 10.1155/2020/8049504.

Iskandaryan, D., Ramos, F. and Trilles, S. (2020) 'Air Quality
Prediction in Smart Cities Using Machine Learning Technologies based on Sensor Data: A Review', Applied Sciences, 10(7), p. 2401. doi: 10.3390/app10072401.

Mahalingam, U. et al. (2019) 'A Machine Learning Model for Air Quality Prediction for Smart Cities', in 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET). 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India: IEEE, pp. 452–457. doi: 10.1109/WiSPNET45539.2019.9032734.

Mahanta, S. et al. (2019) 'Urban Air Quality Prediction Using
Regression Analysis', in TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON). TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India: IEEE, pp. 1118–1123. doi:10.1109/TENCON.2019.8929517.

Nandasena, Y. L. S., Wickremasinghe, A. R. and Sathiakumar, N. (2010) 'RAesierarpchoarltilculetion and health in Sri Lanka: a review of epidemiologic studies', p. 14. Sampath, S. (2019) 'Air Quality Analysis &amp; Prediction'. doi: 10.13140/RG.2.2.14624.23040.

Xayasouk, T. and Lee, H. (2018) 'AIR POLLUTION PREDICTION SYSTEM USING DEEP LEARNING', in. AIR POLLUTION 2018, Naples, Italy, pp. 71–79. doi: 10.2495/AIR180071.

Xi, X. et al. (2015) 'A comprehensive evaluation of air pollution prediction improvement by a machine learning method', in 2015 IEEE International Conference on Service Operations And Logistics, And Informatics (SOLI). 2015 IEEE International Conference on Service Operations And Logistics, And Informatics (SOLI), Yassmine Hammamet, Tunisia: IEEE, pp. 176–181. doi:10.1109/SOLI.2015.7367615.

Zhong, S., Yu, Z. and Zhu, W. (2019) 'Study of the Effects of Air Pollutants on Human Health Based on Baidu Indices of Disease Symptoms and Air Quality Monitoring Data in Beijing, China', International Journal of Environmental Research and Public Health, 16(6), p. 1014. doi: 10.3390/ijerph16061014.

ACKONWLEDGEMENT

The corresponding author would like to thank supervisors for their commitment & guidance throughout this research study.

AUTHOR BIOGRAPHIES

Mr. RM Fernando is a Computer Science undergraduate at the Department of Computer Science, Faculty of Computing, KDU.

Ms. WMKS Ilmini is a Lecturer at the Department of Computer Science, Faculty of Computing, KDU.

Ms. DU Vidanagama is a Lecturer at the Department of Computer Science, Faculty of Computing, KDU.