International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 1, November-December 2020 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Air Quality Prediction using Seaborn and TensorFlow

Rahul Kumar Sharma[1], Kuldeep Baban Vayadande[2], Rahul Ranjan[1]

[1]Master of Computer Application, [2]Assistant Professor,
[1,2]Jain Deemed-to-be University, Bengaluru, Karnataka, India

ABSTRACT

Air quality is considered as a vital issue in the current world and is the underlying driver of sicknesses identified with respiratory organ, skin malignant growth, corrosive downpour and a worldwide temperature alteration. Anticipating air quality has been the consistent test with the developing industrialization, vehicles out and about, deforestation and different variables. Air contamination has been the issue of the entire world. In this paper, we propose to foresee the air nature of a specific spot, with the information gathered in past and take preventive measure to stop the disaster. We will utilize Spearman's Correlation as information used to foresee air quality is non-straight and monotonic. Spearman's Correlation coefficient (rs) can invigorate us of the connection between highlights of information. The coefficient esteem is obliged to - 1 =<rs =< 1.

KEYWORDS: *Air Quality, Spearman's Correlation, Python, Google Colab, TensorFlow, Seaborn*

How to cite this paper: Rahul Kumar Sharma | Kuldeep Baban Vayadande | Rahul Ranjan "Air Quality Prediction using Seaborn and TensorFlow" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-1, December 2020, pp.479-481, URL: www.ijtsrd.com/papers/ijtsrd37975.pdf

IJTSRD37975

INTRODUCTION

With the fast development of industrialization and the improvement of urban areas, the nature of air is diminishing quickly. Air Pollution in metropolitan zones has become an explanation behind fear and critical concern of making a city. Anticipating the nature of air is significant and is viewed as a major question in ecological insurance. Numerous urban areas have begun gathering information of air with the goal that they can watch out for the nature of air.

This paper addresses foreseeing Air Quality utilizing Machine Learning. In this paper, we will utilize libraries like Seaborn, TensorFlow and Keras. We will construct a model utilizing the Spearman Correlation.

Spearman Correlation is used while predicting the quality of air as data is nonlinear and monotonic. Each pollution has its own index and scale. Some of the major pollutants are NO2, SO2, RSPM, SPM. Each pollutant affects the human body in a different way. If the index is too high it may create a major problem related to health. Fine Material (PM2.5) is a great concern for human health. PM2.5 alludes to the molecule that has a width under 2.5 micrometres.

Literature Review- In AI, we can choose any calculation or model dependent on our information and issue which will give us a superior understanding of the dataset. Each time we run our model, it gives us diverse worth dependent on our test size and preparing the size of information. Lately, a few specialists have put forth attempts on air contamination event and air quality estimating [1] [2].

Air Quality Standard:

1. Primary Standard: It shields us from unfavourable wellbeing issues.[3]
2. Secondary Standard: It shields us from horticultural harm and harm to buildings.[4]

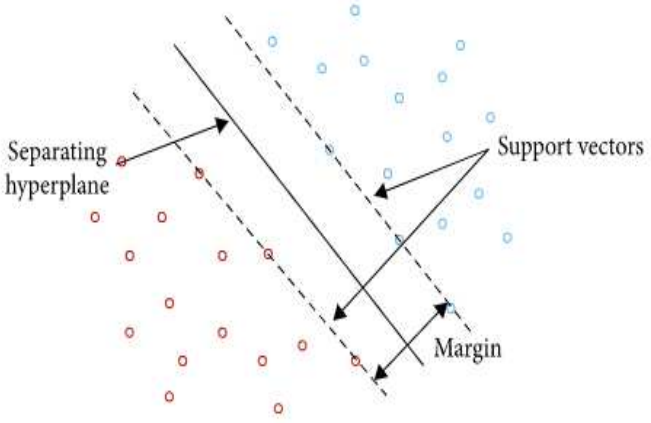
Values	AQI(Range)	PM10	PM2.5	NO2	O3	CO	SO2	NH3	Pb
1	Good (0-5)	0-50	0-30	0-40	0-50	0-1.0	0-40	0-200	0-0.5
2	Satisfactory (51-100)	51-100	31-60	41-80	51-100	1.1-2.0	41-80	201-400	0.5-1.0
3	Moderate (101-200)	101-250	61-90	81-180	101-168	2.1-10	81-380	401-800	1.1-2.0
4	Poor(201-300)	251-350	91-120	181-280	169-208	10-17	381-800	801-1200	2.1-3.0
5	Very poor(301-400)	351-430	121-250	281-400	209-748	17-34	801-1600	1200-1800	3.1-3.5
6	Severe (401-500)	430+	250+	400+	748+	34+	1600+	1800+	3.5+

Standard AQI Range

AQI is a definable measure that is utilized to create a report of the nature of air and its various constituents concerning the climate and human wellbeing [5] [6]. The proposes of an ongoing AQI Monitoring framework is that it estimates various boundaries of air like Particulate Matter (PM 2.5 and PM10), O3, NO2, SO2, CO, CO2, Humidity, Temperature and Air Pressure in the climate. These boundaries are estimated on an hourly premise to get an exact outcome. The framework will be prepared from the information gave to it

and it will be tried in a similar climate with the goal that the outcome is more exact. Information will be gathered through sensors and another medium, it will be cleaned and be handled to get the outcome.

Support Vector Machine is utilized for grouping issues [7]. The primary goal to search for least separation between classes. Focuses lying on classes limits are known as help vector and space between them is called hyperplanes.

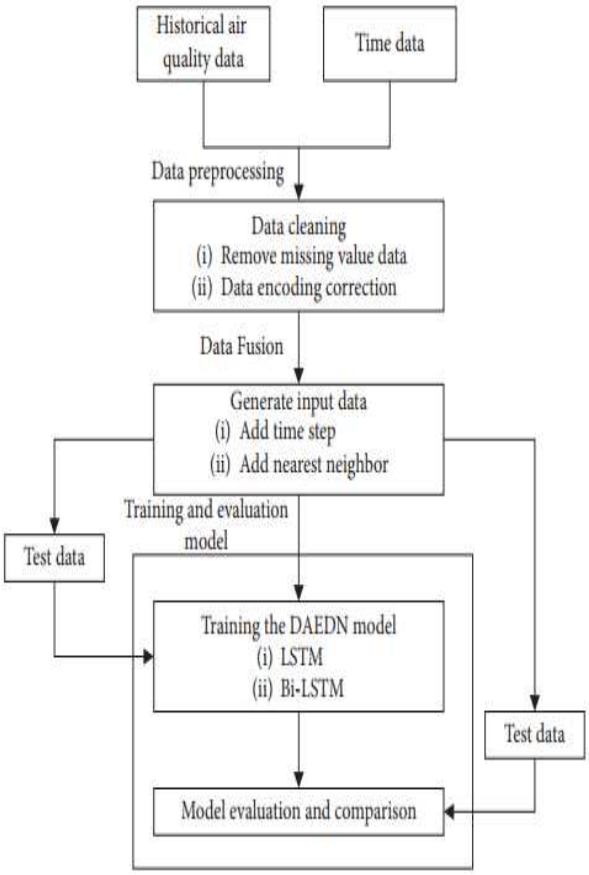


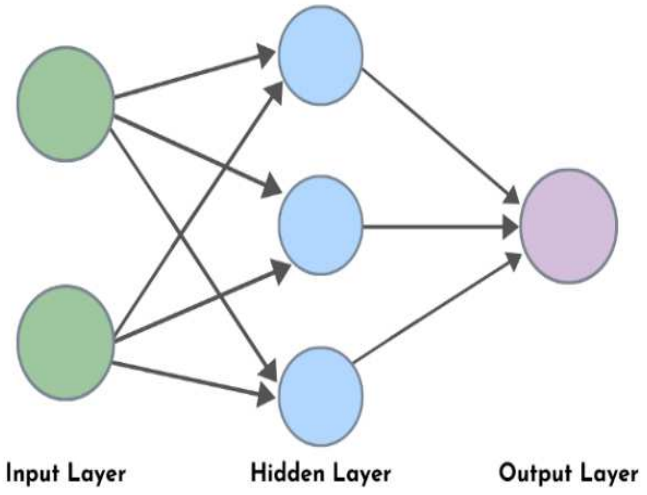
Artificial Neural Network is used by the researcher to solve problems by machine learning. It is a natural model which thinks like the human brain then computes the information and gives us the result. Our human mind has billions of neurons and is connected to each other and communicate over the electrochemical sign. It works like a human mind to solve complex issues in the machine learning part. [8]

Proposed System- The Spearman rank-request relationship coefficient is a nonparametric proportion of the quality and bearing of affiliation that exists between two factors estimated on at any rate an ordinal scale. It is meant by the image rs (ρ). Spearman's connection evaluates monotonic connections (if straight). In the event that there are no rehashed information esteems, an ideal Spearman relationship of +1 or −1 happens when every one of the factors is an ideal droning capacity of the other. The Spearman connection between the two factors will be high when perceptions have a comparable (or indistinguishable for a relationship of 1) position between the two factors and low when perceptions have a disparate rank between the two factors.

We take a dataset and clean it with the goal that it is justifiable. The dataset contains various boundaries for anticipating air quality. Boundaries are Date, Time, CO (mg/m^3), Tin oxide, NMHC, Benzene, Titania, NOx (ppb), Tungsten oxide, NO2 (miniature mg/m^3), True HA, 'Tungsten oxide, nom NO2 ','Indium oxide, Temperature, Relative Humidity, Absolute Humidity. When we know the boundaries, the invalid qualities should be eliminated so it doesn't influence the outcome. Presently we train the information, whenever preparing is done test can be performed.

Architecture:

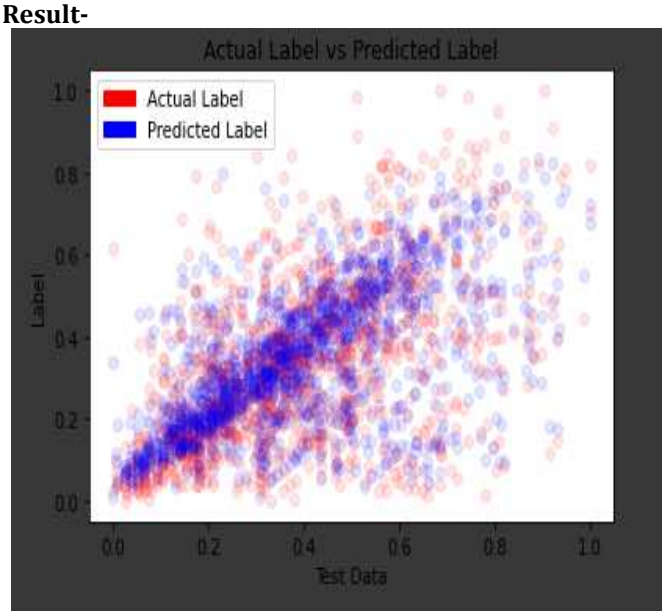




A simple depiction of a Dense layered neural network.

Once the data is normalised and split, we will be creating a model. The model will be a three-layered Sequential model with TensorFlow. All the three layers will be Dense Layers. Dense Layers means that all the nodes in the neural network are connected.

Result-



#calculating R^2 #importing library for R^2 value from sklearn. metrics import r2_score r2 = r2_score(test_label, predictions)

By taking a gander at the above picture, we can see the model has anticipated the marks well. We discover the estimation of R^2 ~0.701, which implies our R coefficient esteem is ~0.837 so the information has corresponded now we can utilize this model to make an expectation on other information and get the outcome.

Conclusion

Machine Learning and Artificial Intelligence assume a significant part in Healthcare, Banking, Stocks, Cyber Security, Weather Forecast, etc. We as a client gather information, clean it, train it, test it and make a forecast. The exactness of the outcome relies upon the nature of the information.

Foreseeing air quality is an intense work because of dynamic nature, expanding vehicle and expanding industrialization.

I would be dealing with Spearman's Correlation. Spearman's Correlation is utilized to make air quality expectations since the information is frequently non-straight and monotonic.

Spearman's connection coefficient (rs) discloses to us the relationship quality between the boundaries of the information. This coefficient is restricted to - 1 =<rs =< 1, and a worth closer to positive or negative 1 demonstrates a more grounded relationship.

A worth more noteworthy than 0.6 is viewed as solid. In this model, we will zero in on connections more prominent than or equivalent to 0.8, which means the relationship is solid.

To get some valuable data from this, SciPy has a helpful inherent

Spearman's Correlation work that will reveal to us both the rs (rho) esteem and the p-estimation of the looked at the information.

Acknowledgement-

I want to communicate my genuine inclination and commitment to Dr MN Nachappa and Asst. Prof: Kuldeep BabanVayadande and venture facilitators for their viable steerage and steady motivations all through my examination work. Their convenient bearing, total co-activity and moment perception have made my work productivity.

Reference-

[1] Zhang Y., Bouquet, M., Mallet, V., Seigneur, C., and Baklanov, A., "Real-time air quality forecasting, Part I: History, techniques, and current status, "Atmospheric Environment, vol. 60, pp. 632-655, 2012

[2] Zhang, Y., Bouquet, M., Mallet, V., Seigneur, C., and Baklanov, A., "Real-time air quality forecasting, Part II: State of the science, current research needs, and future prospects," Atmospheric Environment, vol. 60, pp. 656-676, 2012.

[3] Air Quality Prediction: Big Data and Machine Learning Approaches Gaganjot Kaur Kang, Jerry Zeyu Gao, Sen Chiao, Shengqiang Lu, and Gang Xie

[4] Q. Feng, "Improving Neural Network Prediction Accuracy for PM10 Individual Air Quality Index Pollution Levels", *Environmental Engineering Science*, vol. 30, no. 12, 725–732, 2013.

[5] S. Kumar, A. Jasuja, Air quality monitoring system based on IoT using Raspberry Pi. In Computing, Communication and Automation (ICCCA), 2017 International Conference on 2017 May 5 (pp. 1341–1346). IEEE.

[6] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[7] C. M. Bishop, "Neural Networks for Pattern Recognition", Oxford, 1995.