International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

# Survey Paper on Deep Learning Model based Air Pollution Forecasting Prediction

**[1]Shagufta Perween, [2]Dr. Sushil Kumar**
M. Tech. Scholar, Department of Computer Science and Engineering, Lakshmi Narain College of Technology, Bhopal, India[1]
Professor, Department of Computer Science and Engineering, Lakshmi Narain College of Technology, Bhopal, India[2]

**Abstract:-** Air pollution has become a critical environmental and health issue worldwide, necessitating accurate forecasting models to mitigate its impact. Traditional statistical and physical models have limitations in capturing complex spatiotemporal dependencies of air pollutants. Recently, deep learning models have demonstrated superior performance in air pollution forecasting by leveraging large datasets and capturing intricate patterns. This paper presents a comprehensive survey of various deep learning models applied to air pollution prediction, discussing their architectures, strengths, challenges, and potential future directions. So, the traditional computational intelligence models are not adequate to predict the weather accurately. Hence, deep learning-based techniques are employed to process massive datasets that can learn and make predictions more effectively based on past data. The effective implementation of deep learning in various domains has motivated its use in weather forecasting and is a significant development for the weather industry. The deep learning architectures like Recurrent Neural Networks, and Long Short-Term Memory Networks are proved to be reliable models for weather forecasting tasks.
Keywords: - Air Quality Index (AQI), Machine Learning, Pollutant, Relative Humidity

## I. INTRODUCTION

The process of predicting the atmospheric conditions for a specific location using several meteorological data is known as weather forecasting. Data on the current condition of the atmosphere is gathered in order to make weather forecasts. Meteorologists and scholars have found it difficult to make accurate weather forecasts. Every aspect of life, including agriculture, tourism, the airport system, the mining industry, and power generation, depends on weather data. Knowing the weather in advance aids farmers in making the decisions they need to increase crop production. Continuous meteorological data is necessary for airport or naval systems to detect abrupt changes in the weather. For wind farms to regulate the operation of wind turbines during wind power generation, accurate wind speed prediction is essential. For mining companies to continuously monitor the Earth's crust, accurate meteorological data is essential. Daily, weekly, monthly, or annual weather forecasting becomes crucial since it may more accurately reflect the climatic trend and offer timely and effective environmental information for micromanagement decisions [1, 2]. Because of the rapid growth in the amount of weather data and the development of climate observation methods like satellite meteorological observation, weather forecasting has now entered the Big Data age. Accurate weather prediction is beyond the capabilities of conventional computational intelligence models. Furthermore, more accurate and efficient weather forecasting and climate prediction are possible with the development of deep learning techniques and suitable data visualization approaches [3].

A serious environmental problem that impacts ecosystems, human health, and the climate is air pollution. Proper air quality forecasts can help the public and authorities take preventative action. Non-linearity and high-dimensional interdependence in air pollution data are common problems for traditional approaches, such as statistical regression models and numerical simulations. By identifying intricate patterns in both past and current data, deep learning offers the potential to improve prediction accuracy [4, 5].

## II. AIR POLLUTION

Currently, air pollution is recognized as a significant public pathological condition, responsible for an increasing range of health impacts which are well documented from the outcomes of extensive studies in several areas of the world. While there's little doubt that fast urbanization means we tend to are currently exposed to unhealthy concentrations and additional numerous ambient air pollutants. X-ray radiation imaging studies on the bodies of ancient mummies have detected proof of respiratory illness, emphysema, respiratory organ lump and arteriosclerosis, while autopsies have represented in depth carbon deposits within the respiratory organ [6].

This, in turn, has led to a speculative link to the daily inhalation of smoke in confined areas from fuels used for heat, cooking, and lighting. Air pollution is the release of natural or human made harmful gases and particles into the environment. Air pollution has a higher impact on society and threatens humanity's ability to survive. There was a substantial rise in the use of coal in factories and homes during the urbanization and industrial revolution. These outcomes in smog which caused dismalness and mortality in the stale climatic conditions. During the 1952 Great London smog, the tragedy of losing 4000 life in heavy pollutions highlights the connection between air pollution and human health. In this way air pollution is a developing issue in the urban locales around the globe. Air pollution is a composite of either natural or human activity gasses or particles released into the atmosphere. The number of particles released will be more harmful than the tolerable amount. There are two types of pollutant sources. Natural Sources: Natural pollutants are harmful substances which are emitted by natural phenomenon. SO2, CO2, NO2, CO, and Sulphate are few natural pollutants discharged due to eruptions of volcanoes and forest wildfires [7, 8].

**Man created (Anthropogenic) sources:** The primary causes of air pollution are man-made sources, such as the burning of fuel, emissions from mechanical generation forms, and transportation discharges. Numerous pollutants, including as hydrogen, oxygen, nitrogen, sulfur, metal composites, and particulate matter, are spread by man-made sources. Interest in

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

vitality has increased significantly across the planet due to the expanding global population and the increasing economy. The widespread and massive usage of fossil fuels has also resulted in an increase in ecological issues that have been considered due to their immediate influence on the earth and human well-being. In many parts of the world, air pollution is a major problem. There are two main concerns: the effects on human health, such as cardiovascular disorders, and the effects on the environment, such as corrosive rain, environmental change, and global temperature changes. The bio-geo-chemical cycle refers to the processes that move and change matter in the four categories that make up the Earth's atmosphere, hydrosphere, lithosphere, and biosphere. The planet's operations are governed by these cycles [9]. The nonparticulate receptive radiation from the sun is represented by the planet. As a result, the planet's surroundings are updated and dispersed geographically from the moment of its creation. The existence of life on Earth is attributed to the planet's atmosphere, which in turn influences life on Earth due to the swift changes in these cycles. To forecast the atmospheric conditions in various locations, the Department of Science and Technology uses a number of regulatory agencies. Air quality modelling and monitoring system play a major role in all elements of pollution management and air quality, wherever forecast can play a major role. Air quality predictions provide the general public with air quality data informing them of preventive measures to avoid or limit their exposure to unhealthy levels of pollution [10, 11].

**Implications of Air Pollution to the Quality of Living: -** Air pollution is a substantial threat to health worldwide. As indicated by 2015 Global Burden of Disease exposure to external pollution is that the world's fifth major death risk problem, accounting although air pollution is a universal problem, it is probable to cause the biggest damage in sensitive people exposed to harmful pollutants. People with chronic diseases (especially cardiorespiratory diseases), very little social support, and lack of medical facilities are most at risk from pollution. In 2015, 4.2 million fatalities and losses of 103.1 million life-years adjusted for disability. The most worrying sign is that the incidence of chronic obstructive pulmonary disease and lung cancer is likely to be higher in older populations (aged >55 years) than younger populations (aged also liable for the depletion of the ozone layer that causes Ultraviolet rays to penetrate the Earth and acid rain that has adverse effects on trees and wildlife. Hence, regulation of air quality and its forecasting has become an important task for both developed and developing nations. As a result of increasing man-made developments, growth of pollutants concentration into the atmosphere has become inevitable and thus, depreciating air quality. Air pollutants are classified into two categories – primary pollutants and secondary pollutants. Pollutants which are generated through the process, for instance, ash from a volcanic eruption are referred to as primary pollutants. Examples – Carbon Monoxide (CO), Sulphur Dioxide (SO2). Secondary pollutants are a result of the direct or indirect reaction of primary pollutants. A prominent example of secondary pollutant includes ground level Ozone (O3). Among the six criteria pollutants, Particulate Matter 2.5 (PM2.5) is considered as one of the most pernicious (Pandey et al., 2013). With only 2.5 microns in diameter in size and

International Journal of Research and Technology (IJRT)

International Open-Access, Peer-Reviewed, Refereed, Online Journal

ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529

| An ISO 9001:2015 Certified Journal |

light advocating them, these tiny particles tend to stay for an extended period of time in the atmosphere and cause harmful effects inside filters of the nose and throat. Growing lung cancer mortality and 4-8 percent increase in the cardiopulmonary threat is directly associated with the upgradation of each 10- µg/m3 long-term average PM2.5. A major source of Nitrogen Dioxide (NO2) formation is emissions of power plants and automobiles which in turn leads to the formation of ground-level O3 and fine particle pollution. The mixture of O3 and NO2 is considered as a major threat to children and people suffering from lung diseases like chronic bronchitis, asthma, and emphysema. Prolonged exposure to a certain concentration level of O3 can also result in detrimental effects on plants, crop yield, flora, and fauna. Natural and anthropogenic emission sources of CO include forest fires, animal metabolism, IC engines and burning of carbon enriched fuels. It directly leads to the condition of subnormal oxygenation of the arterial blood and augmentation of greenhouse gases.

## III. AIR QUALITY INDEX

An AQI is a measurable statistic used to record and report uniformly on the air quality of different constituents in terms of human health. The AQI is a daily reporting measure for air quality. It tells you how clean the air we breathe is. The air quality index, or AQI, serves as a benchmark. People can learn more about their living conditions thanks to this AQI. This helps allocate funds to the air pollution control boards and provides the essential information about the pollution caused by the industries' emissions. It can also serve as a feedback factor to adjust the emissions from such industries. By ranking the various sites according to their pollution levels, this AQI assists us in identifying the most contaminated areas and the frequency of possible risks. Predicting air quality and pollution management actions is made possible by AQI, which assists in identifying changes in air quality over a specified time period. The significant amounts of air contaminants used to calculate this AQI. Depending on the major contaminants present, the AQI is calculated differently in each country. India's air quality index, as announced by the Central Pollution Control Board (CPCB) (http://www.cpcb.nic.in) The pollutants SO2, O3, CO, PM (including PM10 and PM2.5), NO2, Lead (Pb), and Ammonia (NH3) make up the AQI.

| AQI Category | $PM_{10}$ 24 Hr µg/m³ | $PM_{2.5}$ 24 Hr µg/m³ | $NO_2$ 24 Hr µg/m³ | $O_3$ 8 Hr µg/m³ | CO 8 Hr (mg /m³ ) | $SO_2$ 24 Hr µg/m³ |
|---|---|---|---|---|---|---|
| Good(0-50) | 0-50 | 0-30 | 0-40 | 0-50 | 0-0.1 | 0-40 |
| Satisfactory(51-100) | 51-100 | 31-60 | 41-80 | 51-100 | 1.1-2 | 41-80 |
| Moderate(101-200) | 101-250 | 61-90 | 81-180 | 101-168 | 2.1-10 | 81-380 |
| Poor(210-300) | 251-350 | 91-120 | 181-280 | 169-208 | 10-17 | 381-800 |
| Very Poor(310-400) | 351-430 | 121-250 | 281-400 | 209-748 | 17-34 | 801-1600 |
| Severe(401-500) | 430+ | 250+ | 400+ | 748+ | 34+ | 1600+ |

**Challenges and Limitations**

1. The pollutant types and levels will vary from one location to another. The sources of pollutants are also different like Natural and manmade.

2. Each pollutant will have an adverse effect on human beings.

3. The monitoring stations will give measurements of various pollutants. More commonly the data available is 24 Hrs average.

4. All the pollutants have to consider calculating AQI for a particular location. Most of the pollutants will vary with time. Handling a huge amount of data will be a tough task.

5. More research has been done on predicting the individual forecasting of pollutants but not on the AQI.

6. The information is supplied straight without scrutiny from the analysers for realtime AQI, so it may not be for a statutory purpose.

7. Monitoring and subsequent AQI dissemination involves various steps including the operation of sensors and analysers, their calibration, local server data acquisition, transmission via the Internet to a central database, etc. Due to multiple technical and operational elements such as lengthy power cuts and maintenance issues, monitoring station functioning may also be impacted. Given these constraints, some interruption in the continuous flow and dissemination of information may occur. However, in the event of breakdowns, immediate action is taken to restore the system to operation within a reasonable period of time.

**Need for Air Quality Forecasting**

Because of restricted resources and practical execution, an alternative approach to tracking air quality is needed to estimate roughly the temporal and spatial distribution of pollutants. Air Quality models are used to indicate air quality standards. It is the least expensive techniques. Regulatory officials can use this modeling as monitoring instruments to evaluate the impact of emissions on ambient air quality. This can also be used to decrease the emissions required to meet the requirements. 9 Air Quality models are generally mathematical descriptions of pollutant transport, diffusion, and chemical reactions from the sources of pollutants (Duc and Azzi, 2009). They accommodate one or a lot of mathematical formulae that include parameters that have an effect on concentrations of pollutants at varying distances downwind of emission sources. Typically, they care for sets of pollutant input data that characterize the emissions, meteorology, and topography of a section and turn out outputs that describe that basis's required air quality. Based on the vital input variable treated the models can be simple or advanced. Advanced models are essentially suited for photochemical air pollution, dispersion in complex terrain, and long-range transport of pollutants. Simpler models are suited for the prediction of particulate matter pollutants of downwind sources. Air Quality prediction has a higher degree of uncertainties than another forecast, as the forecast must diagnose in addition to standard meteorological variables. The forecasting models reduce the uncertainties by "anchoring" the forecast with prior information available and by 'adjusting' the model with additional information like input variables (past measurement values, pollutant concentrations) and meteorological parameters. All the statistical forecasting methods must be calibrated to the effect of large-scale emissions. CO, fine particulate matter with an aerodynamic diameter of less than 10 µm (PM10) and less than 25 µm (PM2.5) are the most prevalent predicted pollutants.

## IV. LITERATURE REVIEW

**Ghufran Isam Drewil et al. [1],** air contamination is a main source of wellbeing concerns and environmental change, one of mankind's generally perilous issues. This issue has been exacerbated by an excess of autos, modern result contamination, transportation fuel utilization, and energy age. Thus, air contamination determining has become indispensable. Because of the huge sum and assortment of information procured via air contamination observing stations, air contamination determining has turned into a well-known theme, especially while applying profound learning models of LSTM. The capacity of these models to learn long haul conditions in air contamination informational collections them separated. Be that as it may, LSTM models utilizing numerous other factual and AI approaches may not offer sufficient forecast results because of loud information and inappropriate hyper parameter settings. Subsequently, to characterize the contamination levels for a gathering of toxins, an ideal portrayal of the LSTM is required. To resolve the issue of perceiving the best hyper boundaries for the LSTM model, In this paper, we propose a model considering the GA computation as well as the LSTM profound learning calculation.

**Jannah Mohammad et al. [2],** have air quality checking framework depends on the RFM design of AI strategies to foresee the convergence of air poisons in metropolitan regions. This paper assesses air contamination information from various adjoining nations into grouped values. In information, gathering fundamental data is likewise depicted by analyzing information. Getting precise expectations requires changing the different area upsides of the bunching calculations utilized in the k-implies calculation. K-means must be utilized when information objects reach out to Euclidean terms. Here, the exhibition of the dataset is seen by K-Means bunching utilizing the elbow strategy with the ideal number. In the recreation, an exact degree of understanding has been brought to the air contamination information and with more prominent accuracy, this issue utilizes scientific devices with Google Co-lab and Python. The group values contrasted and these various sorts of ideal worth conversation will help in arranging the execution of the anticipated AI model.

**Harshit Srivastava et al. [3],** air contamination incorporates air contamination brought about by hurtful gases, buildups, smoke, and so on. Contaminated air creates critical issues for the economical presence of plants, living beings and people, including normal life. This article centers around air contamination expectation utilizing AI (ML) strategies and breaking down its presentation. Different relapse and grouping models, for example, Backing Vector Machine (SVM), Arbitrary Timberland Order, Calculated Relapse, Straight Relapse and Irregular Backwoods Relapse are utilized to advance the impurities air contamination for more exact determining. The presentation of the ML models was assessed utilizing the informational collection of State Contamination Control Board (SPCB), Odisha. The presentation of the relapse models was assessed utilizing root mean square mistake (RMSE) and mean outright blunder (MAE). It overwhelms irregular woodland relapses with RMSE and MAE of 2.63 and 3.32, separately. For grouping models, Irregular Backwoods Classifier leads with 93.5% precision. The model's successful presentation in anticipating air toxins can assist with alarming general society towards more secure ways of life. Air contamination has turned into a significant issue on the planet. The world is dirtied because of the outflow of harmful gases up high like CO2, SO2, NO2 and CO. These harmful gases disintegrate in the air and are capricious. Thus, there is a requirement for an instrument to test air quality. Air contamination can be checked utilizing web gadgets like IoT. Web of Things (IoT) gadgets can gather information and in light of that information can dissect it to foresee regardless of whether the air quality is great. Consequently, the air nature of a specific region can be observed with IOT based sensors and gadgets utilizing Arduino/Raspberry Pi. The point of this exploration is to comprehend data about ecological factors and furthermore empower simple coordination into some other kind of Web based (IoT) design that permits the utilization of sensors equipped for capacity to gather sensor data connected with savvy city natural estimations, to give information on data connected with ecological contamination. **Amar Catovic et al. [4],** a list of air quality reports is called an air quality profile (AQI). It estimates the effect of air pollution on an individual's health over a short period of time. AQI will likely inform the general public about the negative health effects of air pollution in the surrounding area. Air pollution levels in India's urban areas have substantially increased. There are several ways to develop a numerical formula to determine air quality recordings. Many studies have linked openness to air pollution and adverse health effects. The information mining process is one of the most attractive ways to predict and analyze AQI. The objective of this article is to investigate the best AQI prediction method to support environmental control. The best technique can be improved to find the most ideal arrangement. Therefore, the development of this document includes top-down exploration and expansion of new strategies, such as Demolition, to ensure the most ideal answer to the air quality problem. Another important goal is to display and display accurate estimates related to our work in an informative and enlightening manner, thereby making applicable assessments and helping useful for future scientists. In the proposed study, three unmistakable strategies - Supportive Vector Recurrence (SVR), Facultative Forest Recurrence (RFR) and CatBoost Recurrence (CR) - were used to determine the AQI. After comparing the impact of imbalanced data sets, it was found that uneven swab recurrence gave

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

the lowest root mean square error (RMSE) estimate for Bangalore (0.5674), Kolkata (0.1403) and Hyderabad (0.3826), as well as higher contrast accuracy.

**B D Parameshachari et al. [5],** have SVR and CatBoost recurrence in Kolkata (90.9700%) and Hyderabad (78.3672%), while CatBoost recurrence gave the smallest RMSE estimate in New Delhi (0.2792) and notable accuracy highest achieved in New Delhi (79.8622%) and Bangalore (68.6860 %). Regarding the data set after accounting for the (destroyed) technical minority sampling procedure, it is noted that discretionary forest land reuse yielded the lowest RMSE value in Kolkata (0.0988) and in Hyderabad (0.0628) and achieved higher accuracy for Kolkata (93.7438%). Furthermore, Hyderabad (97.6080%) contrasts with Relapse SVR and CatBoost, while Relapse CatBoost provides the most superior accuracy. This clearly shows that datasets to which the Destructive calculation was applied yielded higher accuracy. The puzzling thing about this article is that the best recurrence models were selected through extensive exploration to test their accuracy. Additionally, unlike most related repositories, information manipulation is done through Destruction. Additionally, all runs are reported as plots and measurements, clearly displaying the differences between results and helping.

**Sunori Sandeep Kapil Ghai et al. [6],** monitoring compliance with environmental standards is often performed intermittently due to high costs. However, intermittent monitoring may encourage polluters to demonstrate compliance during monitoring but increase polluting activities at other times. This article documents the strategic response to a cyclical six-day monitoring program as part of large-scale federal air quality regulation. Use one Independently measures atmospheric pollution levels using satellites, I show that air quality is subject to this periodic monitoring it's much worse on unsupervised days. Larger effects were observed in regions (1) failure to comply with air quality standards and may therefore face more stringent sanctions; (2) highly qualified concentration of certain industries, such as wood product manufacturing, and (3) are intersected by highway. Consistent with this latter finding, I show that local governments are more likely to issue Quality notices urge people to reduce car use and outdoor activities on these days Pollution monitoring is planned. This evidence shows that local authorities are coordinating prevention measures. Finally, I document some of the consequences of these strategic responses: Higher pollution levels on unmonitored days lead to lower school test scores and higher pollution levels. violent crime today.

**Chakradhar Reddy et al. [7],** along with rapid urbanization, many non-industrialized countries face serious air pollution problems. The need to predict future air quality is increasingly important in improving government strategies and providing independent guidance to citizens. In this article, we predict air quality for taking into account air quality information, climate information and estimated weather information. Based on information from the field of air pollution, we propose a deep brain organization (DNN)-based approach (named Profound Air), which integrates the spatial change part and the total organization fit.

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

deep diffusion. Considering the spatial relationship of atmospheric toxicants, the main part converts spatially incomplete information on air quality into a reasonable contribution to the reconstruction of pollution sources. The following organization applies distributed brain design to interweave heterogeneous urban information while capturing factors that influence air quality, such as climatic conditions. We have included Profound Air in our air pollution expectations framework, regularly providing accurate air quality estimates for more than 300 Chinese urban communities. Test results from information from Chinese cities over three years show the benefits of Profound Air beyond the 10 core techniques. Unlike the previous web-based approach to the air pollution expectation framework, we achieved relative accuracy improvements of 2.4%, 12.2%, and 63.2, respectively. % by predicting sudden, current and long-term changes.

## V. METHODOLOGY

Machine Learning (ML) is a data handling framework, which are built and executed to show the human cerebrum. The main object of the ML analysis is to develop a process of computing device for modeling the brain to execute various processes of computing tasks at a faster rate than the traditional systems. ML executes various tasks such pattern matching and classification, optimization function and data clustering. These errands are exceptionally troublesome for conventional PCs, which are quicker in algorithmic procedure of registering undertakings and exact number juggling tasks. ML gangs' substantial number of exceedingly interconnected preparing components called hubs or unit or neuron, which often work in parallel and are arranged in standard models [8]. Every neuron is associated with the other by an association interface. Every association interface is related to weights, which contain data about the info flag. This data is required by neuron net to take care of a specific issue. ML, s aggregate conduct is portrayed by their capacity to learn, review and sum up preparing examples or information like that of human mind.

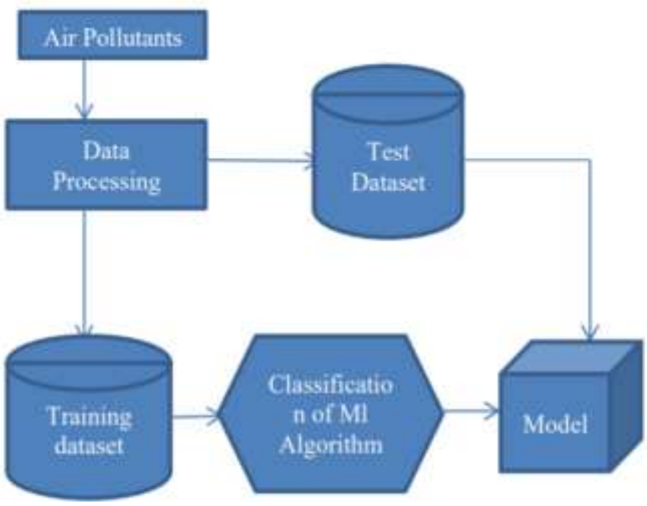


Fig 1: The proposed model's architecture

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

## Deep Learning

Deep learning is a subset of machine learning techniques focused on classification tasks and evolutionary algorithms [14]. There are three types of learning: supervised learning, semi-supervised and unsupervised. Deep-learning architectures incorporating deep learning models, fully connected networks, recurrent neural networks, and artificial neural networks were used in fields involving machine learning, artificial intelligence, computer vision, data analysis, realized, social media site filtering, computational linguistics, computational biology, drug design, information retrieval, and clear overview, among others [15]. Knowledge acquisition and decentralized organizational infrastructure in biological systems influenced artificial neural networks (ANNs). ANNs vary from the human brain in several ways. Neural networks are constant and symbolic, whereas most functioning entities' biological brains are dynamic and analog.

Deep learning gets its name from the fact that it employs many layers in the network. Early research demonstrated that a linear perception cannot be used as a universal classifier but that a network with a non-polynomial input layer and one unrestrained width hidden layer may [59]. Deep learning is a more recent variant involving many layers of bounded size, allowing for functional application and optimization while maintaining theoretical subjectivity under mild conditions. For the sake of performance, trainability, and intelligibility, deep learning structures are also allowed to be diverse and veer away widely from scientifically informed connectionist models, hence the "organized" portion [16].

Most new deep learning techniques focus on machine learning, especially convolutional neural networks (CNNs). They may also include propositional formulas or latent variables structured layer-wise in deep generative models like deep belief networks and deep Boltzmann machines. Each level of deep learning learns to turn the data it receives into a slightly more abstract and composite representation. The raw input in an image recognition program could be a matrix of pixels; the first representative layer could abstract the pixels and encode edges; the second layer could compose and encode edge arrangements; the third layer could encode a nose and eyes, and the fourth layer could recognize that the image contains a face. Importantly, a deep learning algorithm may figure out which features belong to which level on its own.

The term "deep learning" refers to the number of layers that the data is transformed through. Deep learning systems, in particular, have a significant credit assignment path (CAP) depth [5]. The CAP is the input-to-output transition chain. CAPs are used to define possible causal relationships between input and output. The depth of the CAPs in a feedforward neural network is equal to the network's depth plus the number of hidden layers plus one. The CAP depth in recurrent neural networks, where a signal can propagate through a layer multiple times, is theoretically unlimited. Although no generally agreed-upon depth level separates shallow and deep learning, most researchers agree that deep learning needs a CAP depth greater than 2. In the sense that it can imitate any function, CAP of depth two is a universal approximate [7].

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

More layers, on the other hand, do not improve the network's ability to approximate functions. Extra layers aid in learning the features effectively because deep models can extract better features than shallow models. Deep convolutional layers can construct deep learning architectures in CNN. The DL can aid in the deconstruction of these abstractions and the identification of which features improve results. Deep learning methods eliminate feature engineering for supervised learning tasks by converting data into compact feature vectors analogous to factor loading and generating layered structures that reduce redundancy. Unsupervised learning tasks may benefit from deep learning algorithms. This is a significant advantage since unlabeled data is more plentiful than labeled data. ANN and deep belief networks are the two basic neural networks that work like unsupervised learning approach. There are a few different types of deep learning algorithms, which are mentioned below [8].

## VI. CONCLUSION

Monitoring is an activity of measuring an area's ambient air pollution concentrations. The measurement shows the air quality that we breathe. Long term monitoring of information is particularly helpful as it enables us to extract patterns that help policies to control air pollution. These trends include spatial pollution gaps and temporal differences. Thus, while monitoring air pollution does not decrease air pollution by itself, it provides us hints as to where the pollution comes from and what its level is.

Machine Learning (ML) will be suitable in this context. ML’s ability in performing task-specific functions is accomplished through the utilization of statistical models and algorithms by exploiting the data patterns and the inferences of the same. It is seen as a subset of artificial intelligence. ML algorithms are used for the prediction with mathematical models built based on the “training data” from the samples, which are not programmed for a definite task. The ML algorithm can produce precise predictor features from complicated, high-dimensional datasets where statistical and mathematical techniques are susceptible to inaccuracies and are hard to apply because of their fundamental assumptions.

## REFERENCES

[1] Ghufran Isam Drewil and Riyadh Jabbar Al-Bahadili, “Air pollution prediction using LSTM deep learning and metaheuristics algorithms”, Elsevier Science Direct, Vol. 24, pp. 01-07, 2022.

[2] Jannah Mohammad and Mohammod Abul Kashem, “Air Pollution Comparison RFM Model Using Machine Learning Approach”, IEEE 7th International conference for Convergence in Technology (I2CT), IEEE 2022.

[3] Harshit Srivastava, Goutam Kumar Sahoo, Santos Kumar Das and Poonam Singh, “Performance Analysis of Machine Learning Models for Air Pollution Prediction”, International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), IEEE 2022.

[4] Amar Catovic, Esad Kadusic, Christoph Ruland, Natasa Zivic and Narcisa Hadzajlic, “Air pollution prediction and warning system using IoT and machine learning”, International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), IEEE 2022.

[5] B D Parameshachari, G M Siddesh, V. Sridhar, M Latha, Khalid Nazim Abdul Sattar and G Manjula, “Prediction and Analysis of Air Quality Index using Machine Learning Algorithms”, IEEE International Conference on Data Science and Information System (ICDSIS), IEEE 2022.

[6] Sunori Sandeep Kapil Ghai Amit Mittal Manoj Chandra Lohani Mehul Manu and Pradeep Juneja "Prediction of Air Pollutant PM 10 using Various SVM Models" 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) pp. 1541-1546 2022.

[7] Chakradhar Reddy K, Nagarjuna Reddy K, Brahmaji Prasad K and P.Selvi Rajendran, “The Prediction of Quality of the Air Using Supervised Learning”, 6th International Conference on Communication and Electronics Systems (ICCES), IEEE 2021.

[8] Eric Yongchen Zou "Unwatched pollution: The effect of intermittent monitoring on air quality" American Economic Review vol. 111, no. 7, pp. 2101-26 2021.

[9] K. Veljanovska and A Dimoski "Deep Distributed Fusion Network for Air Quality Prediction" Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 19–23 August 2021.

[10] I.H. Sarker "Machine Learning: Algorithms Real-World Applications and Research Directions" SN COMPUT SCI., vol. 2, pp. 160, 2021.

[11] S. Sunori P. Negi S. Maurya P. Juneja A. Rana and Bhawana, “K-Means Clustering of Ambient Air Quality Data of Uttarakhand India during Lockdown Period of Covid-19 Pandemic”, 6th International Conference on Inventive Computation Technologies (ICICT) 2021.

[12] Singh Jayant Kumar and Amit Kumar Goel "Prediction of Air Pollution by using Machine Learning Algorithm", 7th International Conference on Advanced Computing and Communication Systems (ICACCS) vol. 1, pp. 1345-1349, IEEE 2021.

[13] Jovan Kalajdjieski, Eftim Zdravevski , Roberto Corizzo, Petre Lameski, Slobodan Kalajdziski, Ivan Miguel Pires, Nuno M. Garcia and Vladimir Trajkovik, “Air Pollution Prediction with Multi-Modal Data and Deep Neural Networks”, Remote Sensing, 2020.

[14] Zhang, Y.; Guo, L.; Wang, Z.; Yu, Y.; Liu, X.; Xu, F. Intelligent Ship Detection in Remote Sensing Images Based on Multi-Layer Convolutional Feature Fusion. Remote Sens., vol. 12, pp. 3316, 2020.

[15] J. Shafi and A. Waheed "K-Means Clustering Analysing Abrupt Changes in Air Quality", 4th International Conference on Electronics Communication and Aerospace Technology (ICECA), 2020.