



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 7, July 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Air Pollution Forecasting Prediction using Deep Learning Techniques: A Review

¹Sachin Bhargava, ²Prof. Goldy Saini, ³Dr. Sneha Soni

M. Tech. Scholar, Department of Computer Science and Engineering, SIRTE, Bhopal, India ¹

Assistant Professor, Department of Computer Science and Engineering, SIRTE, Bhopal, India ²

Associate Professor, Department of Computer Science and Engineering, SIRTE, Bhopal, India ³

ABSTRACT:- Air is one of the most important components of life on earth. Quality of air in an area significantly contributes to the health of the inhabitants in that particular area. Air pollution poses a severe risk to life and the environment, and if this air pollution continues, there are many severe threats to the atmosphere that can cause the earth's temperature to rise, which can, in turn, destroy life on earth. Air pollution leads to global warming, greenhouse effect, respiratory problems such as bronchitis, lung cancer, and many other dreadful diseases. Accurate weather forecasting has proven to be a challenging task for meteorologists and researchers. Weather information is essential in every facet of life like agriculture, tourism, airport system, mining industry, and power generation. Weather forecasting has now entered the era of Big Data due to the advancement of climate observing systems like satellite meteorological observation and also because of the fast boom in the volume of weather data. 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

Weather forecasting is the practice of predicting the state of the atmosphere for a given location based on different weather parameters. Weather forecasts are made by gathering data about the current state of the atmosphere. Accurate weather forecasting has proven to be a challenging task for meteorologists and researchers. Weather information is essential in every facet of life like agriculture, tourism, airport system, mining industry, and power generation. In agriculture, prior information about weather helps farmers to take necessary decisions to improve their crop yields. Airport or naval systems require continuous weather data to know if there is a sudden change in climatic conditions. Accurate wind speed prediction is critical for wind farms to control the working of wind turbine during wind power generation. Mining industries require precise weather information to monitor the Earth's crust continually. Weather forecasting on a daily, weekly, monthly, or yearly basis becomes essential as it can better reflect the changing trend of climate and also provide timely and efficient environmental information for the decisions at the micro-management level (Salman et al., 2015). Weather forecasting has now entered the era of Big Data due to the advancement of climate observing systems like satellite meteorological observation and also because of the fast boom in the volume of weather data. Traditional computational intelligence models are not adequate to predict the weather accurately. Moreover, with the advancement of deep learning techniques and appropriate data visualization methods, weather forecasting and climate prediction can be made more effectively and accurately.

Hence it is reasonable to use deep learning approaches to process massive datasets that can learn and make predictions more effectively based on past data. Deep learning techniques use layers of neural networks to identify and extract meaningful patterns from the datasets. A neural network with deep architectures can extract high-level abstract features of Big Data accurately (Gheisari et al., 2017). 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. Forecasting models can be categorized as temperature prediction models, wind speed prediction models, rainfall prediction models, dew point prediction models, and so on depending upon the parameter to be predicted. Each prediction model forecasts a particular parameter based on previous observations.

Temperature forecasting models aim at predicting the minimum, maximum, or average temperature of a location, based on various weather parameters. The maximum temperature forecast depends on sunshine during day time, whereas the minimum temperature forecast relies on cloud conditions at night. Temperature prediction plays a vital role in agriculture. Prior information about weather helps farmers to take necessary decisions to enrich their crop cultivation. The extreme temperature will cause severe damages to plants and animals, and it is always a great concern for farmers. The temperature should be optimal for the growth of plants. Low temperatures may cause injury to crops. So, the prediction based on temperature is essential in agriculture.

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 be 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 (Zweifel et al., 2009 ; Thompson et al., 2013), while autopsies have represented in depth carbon deposits within the respiratory organ (Zimmerman et al., 1971). 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 (Kelly and Fussel, 2015). Air pollution is the release of natural or human made harmful gases and particles into the environment (Kemp et al., 2011). 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 (Ferreira et al., 2003). 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 (Kemp et al., 2011). 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.

Man created (Anthropogenic) sources: Man created sources, for example, fuel burning, releases from mechanical generation forms, transportation discharges are the fundamental wellsprings of air contamination. Many pollutants are transferred by man-made sources, including hydrogen, oxygen, nitrogen, sulphur, metal composites and particulate matter. With the growing complete population and the developing economy of the scene, the planet's interest in vitality has grown dramatically. The enormous use of fossil vitality universally has also led to a rise in ecological problems that have been taken into account because of their imminent impact on human well-being and the earth (Kemp et al., 2011; Song Y et al., 2015). Air pollution is a key problem in many areas of the globe, with two critical issues: the impact on human well-being, for instance, cardiovascular diseases, and the impact on the earth, for instance, corrosive rain, environmental change, and global temperature change. (Wang J et al., 2016). The pathways for transporting and transforming matter within four categorical fields that create planet Earth (biosphere, lithosphere, hydrosphere, and atmosphere) square measure known as the Bio-geo-chemical cycle (Niharika and Rao, 2014). These cycles regulate the planet's functioning. The planet is the sun's nonparticulate receptive radiation. Thus, from the moment of its birth, the matter surrounding the planet is renovated and geographically distributed. With the rapid changes in these cycles, the planet's atmosphere is affected to blame for the existence of life on Earth that affects life on Earth reciprocally. The Department of Science and Technology employs several regulatory bodies to predict the state of the atmosphere for different places. 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.

Implications of Air Pollution to the Quality of Living: - Air pollution is a substantial threat to health worldwide (Burnett R, 2018; Cohen AJ et al., 2017; Forbes, 2016). As indicated by 2015 Global Burden of Disease (Balakrishnan et al., 2019) 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 (Zhu S et al., 2017). The related studies by Zhang S et al. (2016) and Fougère B et al., (2018) have shown that the related to air pollution for chronic obstructive pulmonary disease and lung cancer varies with age,

and these results are biologically possible. 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 Ultra Violet rays to penetrate the Earth and acid rain that has adverse effects on trees and wildlife (Gaganjot Kaur Kang et al., 2018). 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 (PM_{2.5}) is considered as one of the most pernicious (Pandey et al., 2013). With only 2.5 microns in diameter in size and 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 threat of cardiopulmonary is directly associated with the upgradation of each 10- µg/m³ long-term average PM_{2.5} (Pope III et al., 2002). 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, emphysema (Jiang et al., 2016). 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 measured metric used to record and report evenly on the air quality of various constituents in terms of human health (Feng and Yang, 2012). AQI is a daily air quality reporting index. It informs you how well we are breathing clean air. The AQI is being used as a standard for the quality of air. This AQI gives information to the people about the environment in which they are living in (Mamta and Basin, 2010). This gives the necessary information about the pollution caused due to the emissions from the industries and can act as a feedback factor so as to modify the emissions from those industries and also helps in allocation of funds to the air pollution control boards. This AQI helps us to rank the different locations in order of the pollution they have, thus highlighting the areas which are more polluted and the frequency of potential hazards. AQI helps to determine the changes in air quality that have happened over a given period of time, allowing for the prediction of air quality and control measures for pollution. This AQI has been calculated by significant levels of air pollutants. The calculation of AQI differs from nation to nation, based on the significant pollutants involved. India AQI, as per Central Pollution Control Board (CPCB) notified (<http://www.cpcb.nic.in>) AQI is constituted by SO₂, O₃, CO, PM (included PM₁₀ and PM_{2.5}), and NO₂, Lead (Pb), Ammonia (NH₃), pollutants.

| AQI Category | PM ₁₀ 24 Hr µg/m ³ | PM _{2.5} 24 Hr µg/m ³ | NO ₂ 24 Hr µg/m ³ | O ₃ 8 Hr µg/m ³ | CO 8 Hr (mg /m ³) | SO ₂ 24 Hr µg/m ³ |
|----------------------|--|---|---|---|--|---|
| 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 pollutants 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 the measurements of various pollutants. More commonly the data available are 24 Hrs average.



4. All the pollutants have to consider for 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 (Ryan, 2016). 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 (PM_{2.5}) are the most prevalent predicted pollutants.

IV. METHODOLOGY

Machine Learning (ML) is the data handling frameworks, 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 process of computing tasks at a faster rate than the traditional systems. ML execute 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 more often than not 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 with 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.

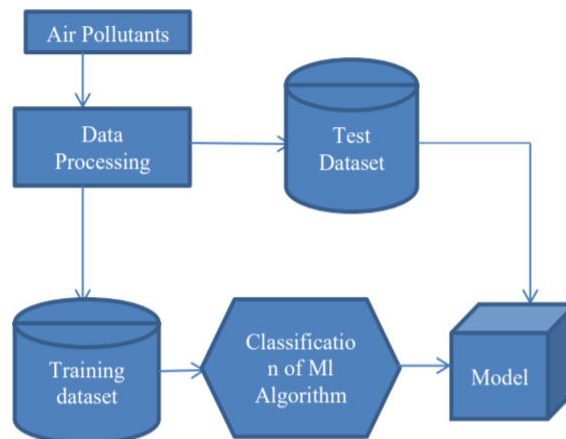


Fig 1: The proposed model's architecture

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. In particular, 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 perceptron 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].

The majority of 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].

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 network that works like unsupervised learning approach. There are a few different types of deep learning algorithms, which are mentioned below [8].

V. CONCLUSION

Monitoring is an activity of measuring an area's ambient air pollution concentrations. The measurement shows the air quality status 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] 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.
- [2] 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.
- [3] 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.* 2020, 12, 3316.
- [4] Bai, L., Wang, J., Ma, X., and Lu, H. (2018), 'Air pollution forecasts: An overview', *International journal of environmental research and public health* 15(4), 780.
- [5] Balakrishnan, K., Dey, S., Gupta, T., Dhaliwal, R. S., Brauer, M., Cohen, A. J. and Sabde, Y. (2019), 'The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: the Global Burden of Disease Study 2017'. *The Lancet Planetary Health* 3(1), e26-e39.
- [6] Cai, S., Wang, Y., Zhao, B., Wang, S., Chang, X., and Hao, J. (2017), 'The impact of the "air pollution prevention and control action plan" on PM_{2.5} concentrations in JingJin-Ji region during 2012–2020', *Science of the Total Environment* 580, 197-209.
- [7] Cascio, W. E., and Long, T. C. (2018), 'Ambient Air Quality and Cardiovascular Health Translation of Environmental Research for Public Health and Clinical Care', *North Carolina medical journal* 79(5), 306-312.
- [8] Huang, M., Zhang, T., Wang, J., & Zhu, L. (2015, September). A new air quality forecasting model using data mining and artificial neural network. In 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS) (pp. 259-262).
- [9] Huang, M., Zhang, T., Wang, J., and Zhu, L. (2015), A new air quality forecasting model using data mining and artificial neural network. In '2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS)', pp. 259-262.
- [10] Kang, G. K., Gao, J. Z., Chiao, S., Lu, S., and Xie, G. (2018), 'Air quality prediction: Big data and machine learning approaches', *International Journal of Environmental Science and Development*, 9(1), 8-16.
- [11] Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. (2002), 'An efficient k-means clustering algorithm: Analysis and implementation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (7), 881-892.
- [12] Kelly, F. J., & Fussell, J. C. (2015). Air pollution and public health: emerging hazards and improved understanding of risk. *Environmental geochemistry and health*, 37(4), 631-649.
- [13] Kemp, A. C., Horton, B. P., Donnelly, J. P., Mann, M. E., Vermeer, M., and Rahmstorf, S. (2011), 'Climate related sea-level variations over the past two millennia', *Proceedings of the National Academy of Sciences*, 108(27), 11017-11022.



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 **6381 907 438** **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details