CHAPTER 1

INTRODUCTION

The World Health Organization (WHO) states that breathing in air pollution increases the risk of developing heart disease, lung cancer, and respiratory infections, among other serious illnesses. It has numerous harmful effects on human health, primarily to the respiratory and cardiovascular systems and can even cause an early death. Environmental deterioration, which exacerbates ecosystems, fuels climate change, and acidifies water bodies, is another significant consequence of air pollution. Particulates, ozone, sulfur dioxide, and nitrogen dioxide are the main contributors to air pollution. It is estimated that indoor and outdoor air pollution kills 3.3 million people worldwide each year. Moreover, air pollution exacerbates environmental issues that are harmful to crops, such as acid rain, climate change, ozone layer depletion and ecosystem degradation. When dangerous or excessive amounts of certain substances are present in the atmosphere, it can harm human health, harm the environment, and exacerbate climate change. Particulate matter (PM), nitrogen oxides (NOx), carbon monoxide (CO), sulfur dioxide (SO2), volatile organic compounds (VOCs), and ground-level ozone (O3) are the primary pollutants.

1.1 DEEP LEARNING

Artificial neural networks are used in deep learning, a subset of machine learning, to identify intricate patterns and relationships in data. The availability of large datasets and improvements in processing power have contributed to its recent surge in popularity. These neural networks, which can learn from enormous volumes of data, are modeled after the morphology and operation of biological neurons in the human brain. Figure 1.1 shows how the layers of deep learning are organized.

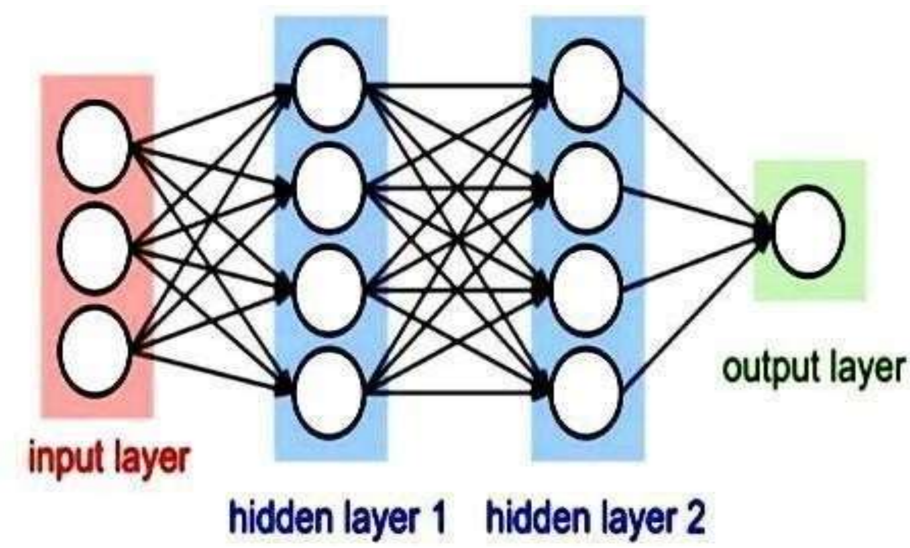

Figure 1.1 Deep learning Layer

Input Layer: The initial layer that takes in input data, like text or images.

Hidden Layers: Between the input and output layers are one or more hidden layers where complex patterns and representations are learned.

Output Layer: The last layer uses the learned representations from the hidden layers to generate the model's predictions or outputs.

1.2 TYPES OF DEEP LEARNING TECHNIQUES

- Artificial Neural Network(ANN)
- Feed Forward Neural Networks(FFNN)
- Convolutional Neural Network(CNN)
- Recurrent Neural Network(RNN)
- Long Short-Term Memory(LSTM)
- Autoencoders (AE)
- Transfer Learning(TL)

Artificial Neural Network:

The arrangement of neurons and connections within an Artificial Neural Network (ANN) determines the network's structure and information. With hundreds of neurons, or processing units, carrying out brain-like tasks, artificial neural networks (ANNs) perform similarly to the human brain. These units consist of input and output units; input units take in data and produce output according to a weighting scheme. Learning rules are used to improve ANN outputs, and mistakes are backpropagated to yield better outcomes. Figure 1.2 illustrates the architecture of an Artificial Neural Network (Bre, Gimenez, and Fachinotti 2018).

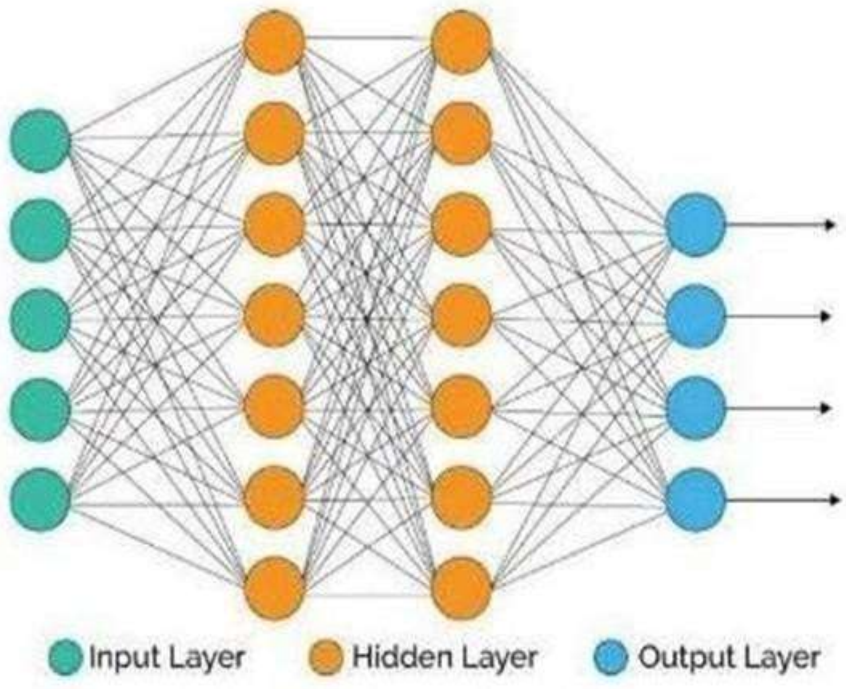

Figure 1.2 Architecture of Artificial Neural Network

Feed forward Neural Networks (FFNN)

Feed Forward Neural Networks (FNN) are the simplest type of neural network architecture. They comprise an input layer, hidden layers, and an output layer. The input layer receives the input data, and each neuron in the hidden layers (also referred to as nodes or units) performs computations on this input data. The output layer then performs a final computation based on the inputs it receives from the neurons in the preceding layer, producing the network's output.

Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are a particular kind of network designed for deep learning algorithms; it is mainly used for pixel-based tasks such as image recognition. CNNs are the most efficient neural networks in deep learning, despite not being the only ones, for tasks like object recognition and identification.

One type of deep learning neural network that is especially well-suited for processing structured arrays of input, like photographs, is the convolutional neural network (CNN). State-of-the-art methods are image processing and other computer vision applications have been attained by CNNs. Additionally, they have demonstrated success in tasks involving natural language processing, like text categorization.

Recurrent Neural Network (RNN)

Applications like Apple's Siri and Google's Voice Search use Recurrent Neural Networks (RNNs), which are algorithms built for sequential data. They are perfect for machine learning problems involving sequential data because they are the first algorithm to remember their inputs through internal memory. The output of a prior computation step is used as the input for the current computation in an RNN. An RNN's interconnected layers, which imitate the composition and functions of the human brain, allow the network to efficiently interpret enormous volumes of data.

Long Short-Term Memory (LSTM)

In artificial intelligence and deep learning, an Artificial Neural Network (ANN) type known as Long Short-Term Memory (LSTM) is employed. In contrast to feed forward neural networks, LSTMs use feedback connections to improve their performance. Because they resolve the vanishing gradient problem that Recurrent Neural Networks (RNNs) commonly encounter, they are helpful for tasks requiring long-term memory.

Autoencoder (AE)

For unsupervised learning, an artificial neural network called an Autoencoder (AE) is employed. Learning effective data representations or encodings is its main goal, which is typically accomplished by lowering the dimensionality of the input data.

- Encoder: This network component creates a lower-dimensional representation of the input data by compressing it. By converting the input data into a compressed or encoded format, it extracts the most crucial elements from the input.
- Decoder: Using the encoded representation, the decoder recreates the original input data. In order to preserve the most important information, the reconstruction should ideally resemble the original input as much as feasible.
- Loss Function: To evaluate an autoencoder's performance, a loss function that measures the difference between the input and the reconstructed output is frequently utilized. For continuous data, the Mean Squared Error (MSE) and for binary data, the binary cross-entropy are common loss functions.

Transfer Learning

The process of fine-tuning a pre-trained model (source model), which has been trained on a larger dataset for a task, on a smaller dataset for a related task (target task) is known as transfer learning. This approach is particularly useful when the target dataset is small and insufficient to train a high performing model from scratch. By adapting a previously trained model for the new task using a smaller dataset, transfer learning makes use of the data from the first task to enhance performance on the second task. Transferring learning has several benefits.

- Data Efficiency: Because the model makes use of information from the larger source dataset, it is particularly helpful when there is a shortage of labeled data for the target task.
- Computational Efficiency: Fine-tuning on a smaller dataset for a particular task usually requires less computational resources than pre-training on a large dataset.
- Enhancement of Performance: When the source and target tasks are related, transfer learning frequently yields superior performance when compared to training a model from scratch.

India's national air quality monitoring network, established in 1987, expanded notably from 2015 to 2019, with a significant rise in both continuous and manual monitoring stations.

However, the network's spatial coverage remains limited, focusing on urban areas, particularly large cities, while rural regions lack monitoring stations. Although the data offer a general view of urban air quality, the limited spatial and temporal coverage presents challenges for long-term trend analysis.

The Global Burden of Disease Study of 2017 analyzed in a report by The Lancet indicated that 76.8% of Indians are exposed to higher ambient particulate matter over 40 μg/m3, which is significantly above the national limit recommenced by national guidelines on ambient air pollution.

India's high levels of fine particulate matter (PM10 and PM2.5), along with a growing awareness of its health risks, led to the National Clean Air Program (NCAP) in 2019. The NCAP set a goal to reduce PM10 and PM2.5 levels by 20–30% by 2024, using 2017 as a reference. While the monitoring network has grown, comprehensive studies analyzing air quality across India remain limited, with most research focused on larger cities like Delhi.

According to the WHO, India has 14 out of the 15 most polluted cities in the world in terms of PM 2.5 concentrations. An examination of daily AQI data from 2015 to 2019 shows that Ahmedabad (484), Delhi (267), Patna (252), Gurugram (234), and Lucknow (224) had the highest average AQI levels, with Ahmedabad categorized as severe (https://earth.org/). In May 2014 the World Health Organization announced New Delhi as the most polluted city in the world. In November 2016, the Great smog of Delhi was an environmental event which saw New Delhi and adjoining areas in a dense blanket of smog, which was the worst in 17 years.

Particulate pollution is particularly high in northern India, with PM10 and PM2.5 levels exceeding national standards by 150% and 100%, respectively. Southern India also recorded elevated levels but at comparatively lower exceedances (50% for PM10 and 40% for PM2.5). Comparisons of PM2.5 data from CPCB monitoring stations and satellite measurements show strong alignment overall, though discrepancies were noted in the western desert region before 2018.

Annual averages for SO_2, NO_2, and ozone (O_3) generally adhered to residential NAAQS standards across India. However, northern India recorded pollutant levels that were 10–

130% higher than in the south, with the exception of SO_2, which showed similar concentrations nationwide. Although there was year-to-year variability, no significant trends were identified for most pollutants from 2015 to 2019, aside from a slight winter decrease in PM10 and PM2.5, particularly in northern and central regions. Despite improvements in monitoring infrastructure, additional stations are still needed in smaller cities and rural areas to achieve comprehensive air quality coverage across India.

1.3 AIR POLLUTION

When dangerous or excessive amounts of certain substances are present in the air we breathe, it can negatively impact human health, the environment and the climate. This is referred to as air pollution. These materials, sometimes referred to as pollutants, can originate from a variety of sources, such as natural disasters like volcanic eruptions and wildfires, as well as industrial processes, vehicle emissions, and agricultural methods. They are two types of pollutants:

Primary Pollutants

- Primary pollutants are emitted directly into the atmosphere from sources like vehicles, industrial activities, and natural events. Unlike secondary pollutants, which are created through chemical reactions in the atmosphere, primary pollutants are released straight into the air. Common examples of primary pollutants include:
- **Particulate Matter (PM):** These minuscule particles are directly released into the atmosphere by a variety of sources, including industrial processes, construction sites, automobile exhaust, and wildfires. The size of PM can vary, but PM10 or PM2.5 are especially dangerous because they can enter the bloodstream and penetrate deeply into the lungs, leading to issues with the heart and lungs.
- **Nitrogen Oxides (NOx**): When materials burn, mostly in power plants, automobiles, and industrial settings, these gases are produced. By causing respiratory system inflammation, aggravating asthma, and increasing the risk of respiratory infections, NOx can directly harm human health. A secondary pollutant called ground-level ozone is also formed by it.
- **Sulfur Dioxide (SO_2):** The main source of this gas is the burning of fossil fuels that

contain sulfur, such as coal and oil, in power plants and other industrial facilities. In addition to aggravating pre-existing respiratory disorders like asthma, SO2 can lead to respiratory problems and produce acid rain, which is detrimental to infrastructure and ecosystems.

- **Carbon Monoxide (CO):** The incomplete burning of fossil fuels in automobiles and other combustion sources results in the production of this colorless, odorless gas. The body's capacity to carry oxygen can be hampered by CO, which can result in symptoms like headaches, nausea, and dizziness, as well as, in extreme situations, death.
- **Volatile Organic Compounds (VOCs):** These organic compounds are released at room temperature and vaporize when exposed to solvents, industrial processes, and vehicle exhaust, among other sources. VOCs are linked to a number of harmful health effects, such as irritation of the eyes, nose, and throat, headaches, nausea, and damage to the liver, kidneys, and central nervous system. They also play a role in the creation of ground-level ozone and fine particulate matter.

Secondary Pollutants

Secondary pollutants are not directly emitted into the atmosphere but form through chemical reactions between primary pollutants and other atmospheric compounds. These reactions typically occur in the presence of sunlight and other atmospheric conditions. Secondary pollutants can significantly impact human health, the environment, and air quality. Common examples of secondary pollutants include:

- **Ground-Level Ozone (O_3):** produced when photochemical reactions between nitrogen oxides (NOx) and volatile organic compounds (VOCs) occur in the presence of sunlight. Ground-level ozone, which can cause respiratory problems like coughing, sore throats, chest pain, and shortness of breath, is one of the primary causes of smog. Long-term exposure to ozone can also exacerbate asthma, reduce lung function, and increase susceptibility to respiratory infections.
- **Nitrogen Dioxide (NO_2):** Although NO_2 is a primary pollutant when it is released directly from combustion sources, it can also be produced as a secondary pollutant when nitrogen oxides (NOx) are oxidized in the atmosphere. In addition to aggravating respiratory disorders like asthma, NO_2 can irritate the respiratory system and help create

ground-level ozone and particulate matter.

- **Sulfuric Acid (H_2SO_4) and Sulfate Particulate Matter (SO_4):** Sulfur dioxide (SO_2), a major pollutant released by sources like power plants and industrial facilities, oxidizes in the atmosphere to form these. Acid rain can damage infrastructure, soil, water bodies, ecosystems, and sulfuric acid and sulfate particulate matter. Cardiovascular and respiratory disorders can also be brought on by inhaling sulfate particulate matter.
- **Secondary Particulate Matter:** These particles can have a negative impact on climate, ecosystem health, visibility, and respiratory and cardiovascular health
- **Secondary Organic Aerosols (SOAs):** They are produced when atmospheric volatile organic compounds (VOCs) oxidize and then condense.

Similar to primary particulate matter, synthetic organic aerosols (SOAs) can have detrimental health effects, such as respiratory and cardiovascular problems. They also contribute to particulate matter pollution.

1.4 AIR QUALITY

Air quality refers to the condition or cleanliness of the air. It is determined by the presence and concentration of pollutants in the air. Common air pollutants include:

1. **Particulate Matter (PM):** tiny airborne particles that can enter the lungs through inhalation. PM is classified according to size, with PM2.5 being especially dangerous because of its capacity to enter the respiratory system deeply.
2. **Nitrogen Dioxide (NO_2):** a gas produced mostly in industrial and automotive settings when fossil fuels are burned. The respiratory system and respiratory disorders can be made worse by excessive NO_2 concentrations.
3. **Ozone (O_3):** Although ozone at high altitudes in the atmosphere shields humans from UV radiation from the sun, ozone at ground level may be dangerous. It develops as a result of chemical reactions between pollutants released by automobiles, power plants, industrial boilers, refineries, chemical plants, and other sources and sunlight
4. **Sulfur Dioxide (SO_2):** a gas produced during the burning of fossil fuels like coal and oil that contain SO_2 can aggravate acid rain, which can also cause breathing problems.

5. **Carbon Monoxide (CO):** a colorless and odorless gas produced when fossil fuels burn too slowly. Elevations of CO have the potential to be hazardous because they obstruct the body's oxygen transportation system.

1.4.1 EFFECTS OF AIR QUALITY

Air quality can have a major and wide-ranging impact on both the environment and human health. High concentrations of pollutants like particulate matter, ozone, nitrogen dioxide, sulfur dioxide, carbon monoxide, and volatile organic compounds are indicative of poor air quality, which can have a number of negative effects.

- **Asthma and Bronchitis**: Exposure to air pollutants can worsen asthma symptoms, increasing the frequency and severity of asthma attacks. It can also lead to bronchitis, characterized by inflammation of the bronchial tubes.
- **Chronic Obstructive Pulmonary Disease (COPD)**: Prolonged exposure to air pollution, particularly nitrogen dioxide (NO_2) and fine particulate matter (PM2.5), has been associated with the onset and progression of chronic obstructive pulmonary disease (COPD), which encompasses emphysema and chronic bronchitis.
- **Heart Disease**: Heart attacks, strokes, and high blood pressure are among the cardiovascular diseases that are linked to air pollution. Particularly associated with cardiovascular health problems are nitrogen dioxide (NO_2) and fine particulate matter (PM2.5).
- **Arrhythmias**: Air pollution can disrupt the normal rhythm of the heart, leading to arrhythmias (irregular heartbeats) and increasing the risk of cardiac events.
- **Cognitive Impairment**: Long-term exposure to air pollution, especially fine particulate matter (PM2.5), has been linked to a higher risk of developing neurodegenerative diseases like Parkinson's and Alzheimer's as well as cognitive decline.
- **Developmental Delays**: Exposure to air pollution during pregnancy has been associated with detrimental impacts on the development of the fetal brain, which may result in behavioral and cognitive issues in the offspring.
- **Reduced Fertility**: Certain air pollutants can impact reproductive health and lower fertility in both men and women.

- **Ecosystem Damage**: Air pollution can harm plants, animals, and ecosystems, leading to reduced biodiversity, damage to crops, forests, and water bodies, and disruption of ecological balance.
- **Climate Change**: Methane (CH_4) and carbon dioxide (CO_2) are two examples of air pollutants that are emitted in the greenhouse gases effect in the climate change and global warming.

1.4.2 APPLICATIONS OF AIR QUALITY

- **Healthcare**: Compressed air is extensively used in healthcare products and for patient care delivery.
- **Urban Planning**: Data on air quality is crucial for urban planning since it helps create cities where citizens are exposed to as little pollution as possible. To cut down on emissions caused by traffic, this entails laying out industrial zones, designing green areas, and improving transportation infrastructure.
- **Transportation**: Monitoring air quality near roadways helps manage traffic to reduce congestion and emissions. It also supports strategies for cleaner transportation, such as promoting electric vehicles and improving public transit.
- **Climate Change Mitigation**: Monitoring air quality involves keeping tabs on greenhouse gas emissions, such as nitrous oxide (N_2O), carbon dioxide (CO_2), and methane (CH_4). This data is essential for determining how different sources contribute to climate change and creating plans to cut greenhouse gas emissions.
- **Emergency Response**: During events like wildfires, industrial accidents, or other environmental disasters, real-time air quality monitoring is essential for assessing immediate health risks to the population. It aids emergency responders in making informed decisions regarding evacuations and public health interventions.
- **Health Research**: Epidemiological studies that look into the long-term health effects of air pollution use data on air quality. Scholars examine the correlation between air quality indices and diverse health outcomes, offering significant perspectives on the health hazards linked to particular pollutants.

- **Indoor Air Quality:** It's critical to keep an eye on indoor air quality in offices to protect workers' health and welfare. Reduced productivity and health problems like respiratory disorders can result from poor indoor air quality.
- **Technology Development**: Innovation in technologies targeted at lowering emissions and enhancing air quality is fueled by data from air quality monitoring. This includes creating smart city solutions, emission control technologies, and cleaner energy sources.

1.4.3 BENEFITS OF AIR QUALITY

- **Improved Respiratory Health:** Good air quality reduces the concentration of harmful pollutants, decreasing the incidence of respiratory problems such as lung problem like as asthma etc.
- **Cardiovascular Benefits:** Clean air is linked to a lower risk of heart-related issues, including heart attacks and other cardiovascular diseases.
- **Enhanced Cognitive Function:** High air quality has been associated with improved cognitive function and productivity, especially in indoor environments.
- **Decreased Allergies:** Cleaner air reduces the presence of allergens, leading to fewer allergic reactions and respiratory discomfort.
- **Longer Life Expectancy**: Better air quality is associated with longer life expectancy due to reduced exposure to harmful pollutants.
- **Environmental Preservation:** Good air quality supports biodiversity by preserving ecosystems and habitats for various plant and animal species.

1.4.4 IMPORTANCE OF AIR QUALITY PREDICTION

An Air Quality Prediction System is valuable for society as it helps prevent air pollution by enabling timely actions to control pollutants. Accurate air pollution forecasting provides essential information for urban planning to maintain environmental sustainability and reduce mortality risks associated with health problems. An air quality prediction system is beneficial for:

Health Alerts: Air quality forecasts provide early warning of impending episodes of air pollution, which enables public health authorities to issue alerts. This makes it easier for vulnerable groups to take preventative measures and stay out of the sun, such as people with respiratory disorders.

Infrastructure Development: In order for urban planners to make well-informed decisions regarding infrastructure development, air quality prediction is essential. This involves placing residential areas, schools, and hospitals to reduce pollution exposure.

Greenhouse Gas Monitoring: Monitoring air quality involves tracking greenhouse gas levels, which contribute to climate change. Predictions help assess the impact of emissions and develop mitigate on strategies.

Regulatory Compliance: Industries and regulatory bodies can use air quality predictions to assess and manage compliance with air quality standards and regulations, contributing to environmental protection and sustainable development.

1.4.5 Statistical data on Air quality

With levels consistently surpassing national ambient air quality standards (NAAQS), particulate matter pollution, particularly PM10 and PM2.5, became the most important air quality problem in India between 2015 and 2019. The highest concentrations were consistently found in Northern India, where PM2.5 levels were 100% higher and PM10 levels were 150% higher than NAAQS. High population density, industrial emissions, vehicle exhaust, and the burning of seasonal crop residues all contributed to the worst pollution in this area, especially in the Indo-Gangetic Plain (IGP). Particulate pollution was a national concern, as evidenced by the fact that southern India, despite being relatively less polluted, still exceeded the NAAQS by 50% for PM10 and 40% for PM2.5.

During this period, seasonal fluctuations were a significant aspect of particulate matter pollution. PM2.5 and PM10 levels were consistently highest during the winter months of December through February. The main cause of this was weather-related factors like temperature inversions and low wind speeds, which kept pollutants near the ground. During these five years, there were no notable long-term declining trends in PM10 or PM2.5 concentrations throughout India, despite considerable year-to-year variability. Toward the end of the period, northern and central India saw a slight drop in winter particulate levels, which may have been caused by early actions taken as part of the National Clean Air Program (NCAP), which was started in 2019. But this decrease was not steady or substantial enough to indicate that particulate pollution was being addressed.

Another significant source was vehicular pollution, which was made worse by the fast urbanization and rise in automobiles in places like Delhi, Mumbai, and Kolkata. The burning of seasonal crop residue in Punjab and Haryana was a major factor in the post-monsoon spike in PM2.5 levels, which had an impact on the quality of the air throughout northern India. PM10 concentrations were also increased by natural dust storms and dust resuspension from inadequate land management, particularly in arid and semi-arid areas.

India made significant strides in growing its network of air quality monitors during this time, which yielded more thorough and precise information on particulate matter pollution. Strong policy measures, such as more stringent industrial and vehicle emission regulations, sustainable farming methods to lessen stubble burning, and better urban planning to reduce dust resuspension, are desperately needed, according to these findings. The data from that time period also provided a crucial baseline for evaluating the success of air quality programs such as the National Clean Air program, which seeks to lower PM2.5 and PM10 levels by 20–30% by 2024.

1.5 AIR QUALITY INDEX (AQI)

In terms of human health, the Air Quality Index (AQI) is a standardized tool for reporting air quality (Feng and Yang, 2012). The AQI has become the industry standard for measuring air quality, providing details on the local atmospheric conditions. It measures the ratio of pollutant concentrations to the condition of atmospheric air in the area (Mamta and Basin, 2010). The AQI allows for ranking various sites based on their pollution levels, highlighting the most contaminated areas and the most common threats. The AQI typically considers several pollutants, including ground-level ozone, particulate matter (PM2.5 and

PM10), carbon monoxide, sulfur dioxide, and nitrogen dioxide. Table 1.1 shows the AQI quantifications based on significant pollutants.

Table 1.1 AQI Quantifications Based on Significant Pollutants

AQI Category	PM10 24 Hr μg/m^3	PM2.5 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
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 (1 01- 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+

They are six threshold level in air quality prediction model:

- Good (0–50): The air quality is considered acceptable and there is little to no risk to health.
- Moderate (51–100): While the air quality is acceptable, there may be a risk for some people, particularly those who are very sensitive to air pollution.
- Unhealthy for Sensitive Groups (101–150): The air quality may cause health issues for sensitive individuals, such as children, the elderly, and those with heart or respiratory conditions.
- Unhealthy (151-200): Everyone may experience negative health effects from the air quality, which is deemed unhealthy. More severe health effects might be experienced by sensitive people.
- Extremely Unhealthy (201-300): The population as a whole may experience serious health consequences due to the extremely unhealthy air quality. There may be health advisories regarding emergency conditions.
- Hazardous (301–500): The population as a whole is probably impacted by the hazardous air quality, health alerts regarding emergency situations.

This data not only aids resource allocation for air pollution control boards but also offers crucial insights into pollution caused by industrial emissions, serving as a feedback mechanism to encourage reduced pollution from such sources. By evaluating changes in air quality over time, the AQI helps predict air quality and keep pollution levels in check. It measures air pollution at highly quantifiable levels and is used to report air quality in specific locations, indicating how clean or polluted the air is and what health effects might be of concern to the general population.

1.6 RESEARCH GAP

Many researchers have concentrated on PM2.5 and PM10 to forecast India's air quality. In the literature review, the analysis was primarily restricted to a single city or region. Conventional air quality prediction models frequently perform less than optimally because they are unable to fully capture the complex patterns and dependencies found in the data. Furthermore, complex sequences can be efficiently extracted, processed, and non-linear features can be extracted using machine learning algorithms. However, they still have issues with over fitting and a tendency toward local optima**.**

1.7 PROBLEM STATEMENT

The health and climate are seriously threatened by air pollution. The sparse networks of stationary sensors used in traditional air quality monitoring techniques result in a limited amount of real-time data and limited spatial coverage. It is possible that the prediction models have not picked up enough information from the past data to produce predictions that are accurate. As a result, it is imperative to have a predictive model that can reduce the restriction and raise prediction accuracy at higher resolutions. In this work, deep learning algorithms are utilized to build air quality prediction models.

1.8 OBJECTIVES OF THIS RESEARCH

The research objectives undertaken in this work are highlighted below:

- To develop deep learning-based models for air quality prediction.
- To identify the temporal and spatial features using long-time delay-based locations for better air quality prediction model.
- To minimize the learning from long-term dependencies of air quality data using Transferred Stacked Bidirectional and Unidirectional Long Short-Term memory method
- To optimize the LSTM-based parameters by using Wasserstein distance.

1.9 CONTRIBUTION OF THE THESIS

In the initial phase of this research, an Improved Sparse Auto-encoder with Deep Learning (ISAE-DL) is implemented and utilizing a feed-forward neural network as a sparse auto-encoder. This approach integrates spatial and temporal features to collect data on particulate matter and meteorological conditions. Additionally, to gather pertinent data and create a prediction model using training data, neural network layers is merged into the fully functional feed forward neural network. The improved sparse auto-encoder method, is introduced to accurately predict air quality system.

In the second part of the research, a Voronoi-Based Clustering using Sparse Auto-Encoder with Deep Learning (VCSAE-DL) is implemented. It is developed to address the long-time delay in location-based air quality prediction. The Manhattan distance is used to group locations that are correlated both spatially and temporally. Clusters are formed iteratively with different centers until all the data points are covered. The clustered data and terrain information are then input into ANN, CNN, and LSTM models. The results from these models are combined and fed into the sparse auto-encoder for accurate air quality prediction.

In the third part of the research, a Transferred Stacked Bidirectional and Unidirectional Long Short-Term Memory (T-SBU-LSTM) model is implemented. For better air quality prediction, this technique reduces the learning problems brought on by long- term dependencies in LSTM. By employing transfer learning to move knowledge from lower temporal resolutions to higher temporal resolutions, the T-SBU-LSTM model increases prediction accuracy. The process of moving knowledge from one domain to another by identifying similarities in tasks, datasets, or models is known as transfer learning. This combined architecture enhances feature learning from large-scale spatial- temporal time series data by capturing both forward and backward dependencies. During this stage of the study, the air quality prediction is expanded from one site to multiple neighboring locations, with varying time delays ranging from short to long.

In the fourth part of the research, Wasserstein Distance-Deep Transfer Learning (WD-DTL) is proposed. This method is used to reduce the learning time process of Transfer

Learning. Using the Wasserstein distance, WD-DTL creates a model that learns invariant features between the source and target domains. First, plenty of data from the source domain is used to train a baseline LSTM model. The LSTM-based feature extractor's parameters are then optimized in order to minimize the estimated empirical Wasserstein distance. Through this adversarial learning process, transferable features from a source domain with known faulty labels can be utilized to diagnose a novel but relevant task in the absence of labelled samples. This technique minimizes computation time while handling huge datasets effectively.

1.10 ORGANIZATION OF THE THESIS

The thesis is organized as follows:

Chapter 1 discusses about the deep learning techniques, types of neural network, air pollution, air quality and contribution of this research work.

Chapter 2 briefly reviews the Machine learning Techniques, Deep learning Techniques, Transfer learning used for air quality prediction method. From the study, the key challenges are identified.

Chapter 3 describes the proposed Improved Sparse auto Encoder using Deep learning Algorithm.

Chapter 4 explains the Voronoi-Based Clustering Sparse Autoencoder using Deep learning algorithm.

Chapter 5 provides an explanation of Transferred Stacked Bidirectional and Unidirectional using Long Short Term Memory Algorithm.

Chapter 6 discusses the Wasserstein Distance using Deep Transfer Learning Algorithm.

Chapter 7 shows the performance comparison of Air quality Prediction with Deep learning Process.

Chapter 8 concludes with a summary of the findings and discusses the improvements to be made in the future.