INTRODUCTION

`

1 Chapter 1

1.1 Introduction

1.1.1 Air Pollution impacts

Air pollution has become a serious cause of concern for environment and health in general due to its severe impact on the quality of life of general population. It can be caused by a variety of sources and manifest itself in a variety of forms, such as gases, particulate matter, Poly Aromatic Hydrocarbons (PAHs), volatile organic compounds (VOCs) etc. Addressing air pollution is crucial for the well-being of both people and the planet, as cleaner air contributes to healthier lives, sustainable ecosystems, and a more stable climate [1, 2]. The deterioration of air quality can be ascribed to various factors such as population growth, industrial processes, fossil fuel combustion, substandard farming methods, and automobile emissions [3, 4]. In general, particulate matters (especially PM_{10} and $PM_{2.5}$), nitrogen dioxide (NO_x), carbon dioxide (CO_2), volatile organic compounds (VOCs), sulfur dioxide (SO_X), ammonia, ozone (O_3), benzene, and carbon monoxide (CO) are some of the most frequently released air pollutants from these activities. As per World Health Organization report 2016, India has 14 cities which belong to top 20 polluted cities of the world [5]. Among numerous air pollutants the high concentration of $PM_{2.5}$ (Particulate Matter $<=$ 2.5μm in diameter), has attracted the attention of various researchers and policymakers across the globe due to their health, environmental and industrial implications. Long-term exposure of $PM_{2.5}$ has been associated with an increased risk of respiratory infections, including pneumonia and bronchitis. $PM_{2.5}$ can enter the bloodstream after inhalation, and can become a cause of cardiovascular events, including heart attacks and strokes. Fine particulate matter can contain carcinogenic substances, and the deposition of these particles in the lungs over time may also contribute to the development of cancer. Evidence has been found to link $PM_{2.5}$ exposure to the onset of hypertension (high blood pressure). It can trigger inflammatory responses in the body, affecting not only the respiratory and cardiovascular systems but also other organs and tissues. Chronic inflammation is implicated in various diseases. Fine particulate matter can potentially reach the brain through various pathways. Workers in certain industries may be exposed to high levels of particulate matter, leading to occupational health risks. Individuals with other respiratory conditions such as asthma, chronic

`

obstructive pulmonary disease (COPD), bronchitis, etc. may experience more severe symptoms when exposed to $PM_{2.5}$. Higher values of $PM_{2.5}$ increases risk of giving birth to infants with low birth weight, which can have long-term health implications. Emerging studies indicate that $PM_{2.5}$ may be linked to neurodegenerative illnesses and may have negative effects on the central nervous system.

1.1.2 Air Pollution and Climate Change

The concentration of most of the pollutants in air are mostly governed by different meteorological factors such as temperature, wind direction, solar radiation, wind speed, rainfall, relative humidity, etc. [6]. Since the pre-industrial era (between 1850 and 1900), the Earth's surface has been warming steadily causing Global warming. The burning of fossil fuels and vehicle emissions, which increase the amount of heat-trapping greenhouse gases in the atmosphere, are the main causes of this warming. The effects of the rise in earth's temperature have detrimental effects in the form of threats to the ecosystem, spread of diseases, high mortality rates and loss of natural habitats etc. Another significant effect of air pollution is climate change, which is an alteration of long-term typical weather patterns. Owing to the shift in climatic patterns, some places become hotter, some wetter, and some drier than others. Extreme and unpredictable weather has become more common nowadays severely affecting all life on Earth. There is a strong interrelationship between climate change and air pollution on a local, regional, and global scale.

Particulate matter can influence climate by scattering and absorbing sunlight, affecting cloud formation, and influencing regional weather patterns. $PM_{2.5}$ can influence the Earth's energy balance through radiative forcing. The particles can either scatter sunlight back into space, leading to a cooling effect, or absorb and re-emit infrared radiation, contributing to a warming effect. The net impact depends on various factors, including the composition and properties of the particles. $PM_{2.5}$ can serve as cloud condensation nuclei (CCN), influencing cloud formation and properties. The presence of aerosols can affect cloud albedo (reflectivity) and its lifetime, impacting regional and global climate patterns. Changes in cloud properties due to aerosol interactions can alter the amount of sunlight reflected back into space, affecting the Earth's albedo. This, in turn, influences temperature patterns and climate dynamics. The distribution of $PM_{2.5}$ is often uneven, with higher concentrations in specific regions due to anthropogenic activities. This spatial variability can have localized climate effects, including changes in temperature, precipitation, and atmospheric circulation patterns. $PM_{2.5}$ can interact with greenhouse gases (GHGs) such as carbon dioxide (CO_2) and methane (CH_4). These interactions may influence the behavior and atmospheric lifetime of GHGs, potentially impacting their contribution to the greenhouse effect and global warming. Research on the role of particulate matter in climate change helps improve climate models and predictions, contributing to a better understanding of the Earth's atmosphere. Including accurate representations of aerosols, including $PM_{2.5}$, in global climate models is essential for improving the accuracy of climate projections. This helps researchers and policymakers understand the complex interactions between aerosols, clouds, and radiation.

The Intergovernmental Panel on Climate Change (IPCC) sixth assessment report 2023 [7], concludes that the unethical use of energy, consumption patterns, changes in land use, and lifestyles are all responsible for the continuous rise in greenhouse gas emissions worldwide between 2010 and 2019 at different levels - regional, national, and individual. This has resulted in unequal contributions to climate change. The impact of human-induced climate change is already evident through the influence on various climatic extremes worldwide. Consequently, this has had widespread negative effects on basic human needs like human health, food and water security, society, and economies, causing both losses to individuals and environment. Significantly, vulnerable communities are more affected, although they have little contribution to current climate change. Planning and executing adaptations has advanced in all areas and sectors, with varied degrees of success and proven advantages. Although there has been progress, gaps in adaptation still exist and will widen at the current rate of implementation. Certain ecosystems and geographical areas have reached their hard and soft limits for environmental adaptation. In certain areas and industries, maladaptation is taking place. The current global financial flows for adaptation are insufficient to support adaptation options and impede their implementation, particularly in developing nations. Since assessment report 5 (AR5), the scope of mitigation-related laws and policies has continuously increased. Keeping warming

`

below 2°C due to global GHG emissions in 2030 will become more challenging, and it is more likely that warming will surpass 1.5°C during the 21st century. The data are reported by nationally determined contributions (NDCs) on October 2021. All sectors and regions do not currently receive the levels of financing needed to meet climate goals, and the predicted emissions from policies that have been put into effect and those from NDCs differ. Many of the climate-related risks are higher than those assessed in AR5 for any given future warming level, and the projected long-term impacts would be higher than what is currently observed. Risks and the resulting losses and damages from climate change rises with global warming. Compound and cascade risks will become more complex and difficult to manage as a result of increased interaction between climatic and non-climatic risks. A significant, swift, and long-term decrease in greenhouse gas emissions worldwide can help to mitigate some inevitable and/or irreversible future changes. Rising global warming levels increase the probability of sudden and/or irreversible changes. Similar to this, increased global warming raises the probability of low-likelihood outcomes that could have extremely significant negative effects. The race to tackle climate change at all governance level is being accelerated by international climate agreements, growing national plans for climate action, and more public awareness. Some countries have achieved GHG emission reductions for more than ten years, as a result of mitigation policies, which have helped to lower the global energy and carbon intensity. Nowadays, there are numerous nearly emission-free options available for industry, buildings, transportation, and energy. Low-emission technologies are also becoming cheaper. Planning and executing adaptation strategies has produced a number of advantages; efficient adaptation strategies can lower the risks associated with climate change and promote sustainable growth. Though it still falls short of needs, global finance for adaptation and mitigation has been increasing since AR5.

The study of $PM_{2.5}$ in the context of climate change allows for a more comprehensive knowledge of the linkages between air quality and climate. Addressing both issues simultaneously can lead to more effective strategies for sustainable development and environmental management. Recognizing the role of $PM_{2.5}$ in climate change emphasizes the need for integrated strategies that address both air quality and climate goals. Policies aimed at reducing emissions of particulate matter can have co-benefits for both public health and climate.

1.1.3 Climate Change and PM2.5: An Inextricable Link

The relationship between climate change and $PM_{2.5}$ pollution is complex and multilayered, creating challenges for both climate science and public health. $PM_{2.5}$ is known for its severe health impacts, and its interactions with climate processes play a critical role in the Earth's atmospheric system. Understanding the link between these two phenomena is essential for developing policies that tackle environmental degradation while protecting human health.

$PM_{2.5}$ can be deeply penetrated into the lungs, making them dangerous to human health. These particles can penetrate into the bloodstream, leading to cardiovascular diseases, respiratory disorders, and even premature death. $PM_{2.5}$ comes from various natural and human-made sources, such as agricultural burning, biomass burning, combustion of fossil fuels, dust storms, volcanic eruptions, and secondary aerosols formed from precursor gases, such as sulfur dioxide (SO_2) and nitrogen oxides (NOx).

$PM_{2.5}$ can alter the Earth's radiative balance, directly impacting climate. The primary mechanism is via aerosols—particles suspended in the atmosphere, some of which are components of $PM_{2.5}$. Aerosols have two primary climate impacts depending on their type: cooling effects and warming effects. Aerosols like sulfate and organic carbon can reflect back sunlight into space, thus cools the Earth's surface. This phenomenon is sometimes called "global dimming" and thus partially compensates the effect caused by greenhouse gases (GHGs). Other aerosols, particularly black carbon (BC), absorb solar radiation, leading to localized atmospheric warming. Black carbon also contributes to warming when deposited on ice and snow surfaces, where it reduces reflectivity (albedo), increasing heat absorption and accelerating ice melt in polar and mountainous regions. This process has accelerated glacier melt in the Himalayas and Arctic, contributing to sea-level rise and changes in freshwater resources.

One of the most significant effects of $PM_{2.5}$ is its role in Arctic and sub-Arctic warming. Black carbon, a component of $PM_{2.5}$, plays a crucial role in accelerating ice melt in these regions. When black carbon falls on snow, it darkens the surface, reducing the albedo and causing the snow to absorb more heat. This effect is most pronounced in the Arctic, where temperatures are rising faster than anywhere else on the planet, contributing to the amplification of global warming known as “Arctic amplification”.

$PM_{2.5}$ also affects the climate indirectly by working as cloud condensation nuclei (CCN), facilitating cloud droplet formation. Clouds formed with smaller, more numerous droplets reflect more sunlight, a phenomenon known as the Twomey effect. This results in a net cooling effect, but these clouds are less efficient at precipitating (producing rainfall), which can lead to longer-lived clouds that alter precipitation patterns.

Increased cloud formation from aerosols may also lead to albedo changes over large regions, creating further cooling or warming impacts depending on cloud type, location, and duration. However, the full extent of these impacts is still an area of ongoing research.

Climate change and $PM_{2.5}$ interact in feedback loops that exacerbate each other’s effects. Climate change has been linked to longer and more intense wildfire seasons, especially in regions like the western United States, Australia, and southern Europe. These fires emit large amounts of $PM_{2.5}$, including black carbon, which contributes to both air pollution and further warming. Wildfires are expected to become more frequent as temperatures rise, leading to a vicious cycle where increased $PM_{2.5}$ emissions contribute to further climate warming.

As black carbon accelerates ice melt, it reveals darker surfaces of water or land, which absorb more heat. This feedback loop leads to more melting, enhancing global warming and regional changes in albedo. The disappearance of reflective ice leads to more solar absorption, exacerbating temperature rises.

Conversely, climate change can alter $PM_{2.5}$ levels by modifying atmospheric processes. Warmer temperatures, changes in precipitation patterns, and shifts in wind and weather systems affect how $PM_{2.5}$ is generated, transported, and removed from the atmosphere. Rising temperatures accelerate the chemical reactions that lead to the formation of secondary $PM_{2.5}$, such as sulfate, nitrate, and organic aerosols. The increase in temperature also promotes the formation of ozone, which in turn influences $PM_{2.5}$ formation through complex atmospheric chemistry. Precipitation plays a critical role in removing $PM_{2.5}$ from the atmosphere. As climate change alters

`

global precipitation patterns, some regions may experience reduced rainfall, allowing $PM_{2.5}$ to accumulate in the air. Dry, arid conditions also increase dust storms, which contribute to natural sources of $PM_{2.5}$. Climate change may increase the frequency of atmospheric stagnation events—periods of low wind speed and poor air mixing—which trap pollutants like $PM_{2.5}$ close to the ground, exacerbating air pollution levels in urban and industrial areas.

Both $PM_{2.5}$ and climate change pose significant threats to public health, and their combined effects are particularly harmful. Exposure to $PM_{2.5}$ is linked to various health issues. Some of the common examples of respiratory and cardiovascular diseases are chronic obstructive pulmonary disease (COPD), asthma, lung cancer, heart attacks and premature deaths due to the exacerbation of underlying health conditions.

Heatwaves become more frequent as a result of climate change, which also degrades air quality. Elevated temperatures have the potential to produce ground-level ozone, another harmful substance. When combined with $PM_{2.5}$, this combination can result in hazardous air, particularly for susceptible groups such as children, the elderly, and those with underlying medical conditions.

In regions affected by frequent wildfires, such as California or parts of Australia, the combined effects of smoke (rich in $PM_{2.5}$) and heat can lead to spikes in emergency room visits and hospitalizations. The World Health Organization (WHO) estimates that millions of premature deaths annually are linked to air pollution, and this number is expected to rise with climate change.

The effects of $PM_{2.5}$ and climate change are felt unevenly across the globe. Heavily industrialized regions like China and India experience high levels of $PM_{2.5}$, which can amplify local warming due to the black carbon component. These regions also face severe public health crises from air pollution.

In contrast, regions like the Arctic or high-altitude areas like the Himalayas are more vulnerable to the albedo effects of black carbon, where small increases in $PM_{2.5}$ can have disproportionate impacts on ice and snow melt. This is particularly concerning for global sea-level rise and water security for millions of people reliant on glacier-fed rivers.

Addressing the intertwined challenges of $PM_{2.5}$ pollution and climate change requires comprehensive policies that focus on reducing emissions of both greenhouse gases and aerosols. Some key strategies include utilizing renewable energy sources, such as hydropower, solar power, and wind, to lessen reliance on fossil fuels. This reduces both CO_2 and $PM_{2.5}$ emissions. Policies to limit emissions of black carbon from diesel engines, industrial processes, and biomass burning can help in reducing air pollution and climate impacts. Technologies like particulate filters for vehicles and improved cook stoves in developing countries are essential in reducing black carbon emissions. Global agreements like the Paris Agreement and initiatives under the Climate and Clean Air Coalition (CCAC) emphasize reducing short-lived climate pollutants (SLCPs), including black carbon, methane, and HFCs (hydrofluorocarbons). These pollutants contribute both to warming and air pollution, making them high-priority targets for climate action.

The interaction between climate change and $PM_{2.5}$ is a dynamic and multifaceted issue, requiring urgent and integrated responses. $PM_{2.5}$ plays a dual role in climate processes both contributing to warming and offsetting it through aerosol effects but its overall impact is detrimental to human health and the environment. As climate change alters the distribution and intensity of $PM_{2.5}$, it compounds the public health risks posed by air pollution. Coordinated global efforts that address both climate change and air quality are crucial for ensuring a sustainable and healthy future.

1.1.4 Co-Benefits of PM2.5 Reduction Strategies and Climate Mitigation

Fine particulate matter ($PM_{2.5}$), in particular, is a major source of air pollution that endangers public health and fuels global warming. Pollutants such as organic compounds, sulfate, and black carbon are included in $PM_{2.5}$, which is made up of particles smaller than 2.5 micro meters in diameter. Greenhouse gas (GHG) emissions, specifically those of carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O), are the primary drivers of climate change concurrently. As many pollutants that contribute to air pollution are also climate forcers, strategies to reduce $PM_{2.5}$ emissions can have a major positive impact on mitigating climate change. This study of the literature investigates the benefits that come from combining efforts to reduce $PM_{2.5}$ pollution with efforts to mitigate climate change.

`

Short-lived climate pollutants (SLCPs) include methane and black carbon, two air pollutants that cause $PM_{2.5}$. These pollutants have a significant, albeit transient, warming effect on the climate. For example, black carbon is a major component of $PM_{2.5}$ and, because of its capacity to absorb sunlight, is a powerful warming agent. Reducing black carbon emissions can significantly lower $PM_{2.5}$ levels and contribute to short-term climate mitigation. Examples of such strategies include burning biomass outdoors, cooking with biomass in homes, and using diesel engines. [8] demonstrate that reducing the harmful health effects of $PM_{2.5}$ exposure while also focusing on SLCPs like black carbon and methane could cut global warming by 0.5°C by 2050.

Emissions of GHG and $PM_{2.5}$ are mostly caused by the burning of fossil fuels in factories, power plants, and automobiles. There can be significant decreases in CO2 levels and $PM_{2.5}$ emissions by switching from coal and oil to greener energy sources like wind, solar, and hydroelectricity. [9] argue that a full switch to renewable energy reduces $PM_{2.5}$ emissions and, by gradually replacing carbon-intensive energy sources, also aids in mitigating climate change. Improved air quality can drastically lower the number of premature deaths linked to air pollution, making the co-benefits of such transitions especially noteworthy in urban areas.

Efficient industrial processes and transportation systems can contribute to a decrease in $PM_{2.5}$ and greenhouse gas emissions. Energy-efficient technologies, for example, reduce carbon output and particulate emissions by using less fossil fuel in industrial processes and low-emission vehicles. [10] shows that energy-efficient building and transportation policies lower CO_2 and $PM_{2.5}$, which benefits human health and the environment at the same time. Two notable examples are the promotion of electric vehicles (EVs) and investments in public transportation.

Agriculture contributes significantly to $PM_{2.5}$ and methane emissions because of its biomass burning and livestock production. Both air pollution and climate change can be mitigated by the adoption of sustainable agricultural practices, such as enhanced manure management, decreased use of synthetic fertilizers, and alternative cropping systems. Reducing the amount of crop residues burned outdoors and implementing conservation tillage are two practices that lower $PM_{2.5}$ emissions while simultaneously improving soil carbon sequestration, which serves as a climate mitigation tactic [11].

The emission of greenhouse gases and air pollution are both significantly increased in urban areas. Incorporating green spaces, encouraging public transportation, and decreasing dependence on automobiles can all contribute to urban planning that improves air quality and mitigates climate change. In addition to absorbing CO_2, trees and other vegetation also trap particulate matter, lowering $PM_{2.5}$ levels. [12] point out that urban green spaces have a dual positive impact on the environment and human health by improving air quality and aiding in climate adaptation.

Emission trading systems (ETS) and other carbon pricing policies, which discourage the use of fossil fuels, can unintentionally lower $PM_{2.5}$ levels. Research indicates that the implementation of a carbon tax, especially in areas with a high coal dependency, can result in a noteworthy decrease in air pollution levels concomitant with greenhouse gas reductions. Research by [13] discovered that the reduction of coal combustion and other carbon-intensive processes led to an improvement in air quality in areas where carbon pricing was instituted.

Lower emissions of GHGs and $PM_{2.5}$ are the result of climate policies that support renewable energy sources and lessen dependency on fossil fuels. One of the best examples of co-benefits is the switch from coal-fired power plants to renewable energy sources. [14] make the case that implementing solar, wind, and hydropower reduces air pollution, which benefits public health significantly and helps to mitigate climate change, particularly in areas with high population density.

Energy consumption and emissions are decreased by energy efficiency standards and the adoption of low-carbon technologies in residential and commercial buildings, such as electrification, better insulation, and the use of heat pumps. [15]. point out that retrofitting buildings for energy efficiency contributes to overall GHG reductions and helps lower $PM_{2.5}$ emissions related to heating. In colder climates, where traditional biomass and coal-burning stoves are major sources of indoor and outdoor air pollution, this is especially important.

Strategies for reducing $PM_{2.5}$ and efforts to mitigate climate change are closely related, with each providing reciprocal benefits. In addition to lowering the risks to the public's health, addressing $PM_{2.5}$ advances both immediate and long-term climate goals. Limiting the use of fossil fuels and encouraging the use of renewable energy sources are other climate mitigation strategies that lead to improved air quality. By

`

creating integrated strategies that optimize the benefits to health and the environment, policymakers can take advantage of these synergies and create a future that is more resilient and sustainable.

1.1.5 Modeling Studies

Various methodologies, including deterministic models, statistical models, and data driven models, have been documented in the literature for the purpose of forecasting air pollution. Deterministic models require specific knowledge about the air pollutant because they operate at the molecular level, taking into account its physical properties. Conversely, statistical models do not require specific knowledge about the pollutant and can deal directly with linear problems. Conventional studies on air pollution prediction rely on statistical techniques like the Autoregressive Integrated Moving Average (ARIMA), the Auto-Regression and Moving Average Model (ARMA), the Moving Average Model (MA), and the Autoregressive Model (AR).Numerous studies have used deterministic model [16-20], statistical model [21] and models based on artificial neural networks (ANN), for air pollution forecasting. Numerical models build upon previous knowledge and hypothesis theory. Due to its capacity to manage both linear as well as nonlinear data, artificial neural network (ANN)-based models have become increasingly popular in recent years. A number of ANN based architectures were used in air pollution forecasting [22–26]. Neuro-Fuzzy models [27] have also been used in literature for air pollution study. Multilayer perceptron (MLP) model was used by [28] for particulate matter forecasting. [29, 30] have used back-propagation neural network; and [31, 32] have used recurrent neural networks (RNN) for air quality forecasting. Researchers studying air pollution are gradually turning to deep learning techniques, an expanding field under artificial intelligence and machine learning paradigm. Long short-term memory (LSTM) network, one of deep learning architecture, has recently been primarily employed for air quality time series forecasting due to its ability to capture both long-term dependencies and short-term dependencies [33-38], however, other hybrid network topologies were also used [39-41]. The usefulness of encoder-decoder based networks for building predictive models using time-series data was further investigated using the LSTM network.[42-44].To forecast the particulate matter concentration, convolutional LSTM was used for extracting spatiotemporal information and CNN was used for parallel extraction of temporal characteristics [45, 46]. Similarly, the transfer learning BLSTM model was used to study the hourly, daily, and weekly air pollution predictions [47]. Since majority of the research work were centered on air pollution forecasting at one (or a small number of) monitoring stations, the application of such modeling techniques is somewhat restricted. With a population of ~1.4 billion and an area ranked seventh in the world, India is incredibly diverse, spanning latitudes of 8°4′ and 37°6′ and longitudes of 68°7′ and 97°25′. Therefore, models created for use in one city might not work well in another. The model architecture needs to be modified on an ongoing basis in order to be applied in different cities. As a result, it's crucial to have a uniform, less complex, and data-driven model that can be used throughout India with no need of making structural changes in the model. The aim of the current study was to develop a deep learning model of hybrid nature with a standardized architecture that could be applied across Indian cities. To determine which deep learning model performed the best, additional analyses of different encoder-decoder based models were conducted.

1.2 Literature Review

Air pollutant concentration predicting models can be classified in to following distinct type: Deterministic model, Statistical model and Machine learning model or Data driven model [48,49]. Machine learning model could be further subdivided into Artificial Neural Network model, Deep Learning model and recent Hybrid models.

1.2.1 Deterministic Models:

Based on physical conditions, deterministic models simulate diffusion and dispersion patterns, allowing researchers to examine air pollution properties at the molecular level [50,51].Community Multiscale Air Quality model (CMAQ) proposed by [52] is an example of deterministic model. Another famous model- Weather Research and Forecasting (WRF-Chem) introduced by [53] belongs to deterministic model [54]. Physical and Chemical models are included in the above two models for generating environmental parameters and transmission of pollutants. Simulation result of these models gives an understanding of different mechanism associated with the pollutant but could not make prediction up to expected level due to complex surface condition, incomplete pollutant data and expert priori knowledge of the theory [55]. Some of the other deterministic models are Nested Air Quality Prediction Modeling System (NAQPMS) [56], Chemical Transport Models (CTMs) [57], etc. The deterministic model uses default parameters and tested with a smaller number of observations and thus their result was not always accurate [58,59]. Expert priori knowledge, expensive computing, a smaller number of observations were some of the drawbacks of deterministic models.

1.2.2 Statistical models

Statistical models used for forecasting and classification problems produce output directly from the input supplied in data driven manner without going into the undergoing physical or chemical changes in the pollutant. A large number of inputs or observational data are used in statistical models. The statistical techniques most frequently employed in air pollution forecasting are Multiple Linear Regression (MLR) as used by [60], Autoregressive Integrated Moving Average (ARIMA) [61,62], Generalized Additive Models (GAMs) [63], Geographically Weighted Regression (GWR) [64]. Most statistical models consider a linear relationship between the input and output variable which is uncommon in real world problems. These models could not deal with nonlinear problems and this limits their capabilities in predicting accurate air pollutant forecasting [65].

1.2.3 Artificial Neural Network (ANN):

In the area of air quality prediction, shallow machine learning models start to thrive by addressing the shortcomings of statistical models. ANN were data driven, non-parametric, flexible and can handle nonlinearity. Random Forest (RF), Support Vector Machine (SVM) are examples of shallow machine learning models. [66] utilized an RF model to forecast Shanghai's PM2.5 and NO2 concentrations while taking transportation, weather, topography, land use, etc. into account. The model produced superior outcomes to the conventional regression model. [67] used SVM for air pollution. In the air pollutant monitoring system, shallow machine learning models may take in a variety of data types as input but overlook nearby factors [68]. It is also difficult to process the complex and massive spatiotemporal pollutant data, because a large portion of their performance is dependent on manually constructed features. Support vector regressions (SVR) model have been proposed by [69]. Other SVR applications in air pollution forecasting are also found in literature [70-72]. These models mostly predict current time-step pollution [73] instead of future pollution. Artificial Neural Network (ANN) is gaining popularity and perform better over statistics-based models [74]. Use of ANN and other hybrid models were found in literature for air pollutant forecasting [75-78].

ANN can be used as a feed forward [79,80] or feed backward [81,82] neural network. Recurrent neural networks (RNN) [83] were examples of feed backward neural networks. Different hybrid neural networks were also found in literature such as general regression neural network [84] and neuro-fuzzy neural network [85]. Other machine learning techniques such as Random Forest [86], Fuzzy Logic [87,88] were also found in literature used for air pollution prediction.

Recurrent Neural Network (RNN) which can learn temporal sequences in a time series data was found widely used in air pollution forecasting [89]. Elman neural network, a variant of RNN, was used [90] for the air pollutant concentration forecasting model. Another variant of RNN applying time delay [91] was found in literature applied for air pollutant concentration forecasting. One major drawback of traditional RNN was that their performance degrades when applied for long term sequences.

1.2.4 Deep Learning Models:

A subset of artificial intelligence and machine learning, deep learning is a rapidly developing field that has drawn many researchers to work in it. Development and literature review of Deep learning was well mentioned in [92-94]. In contrary to shallow machine learning models, these models have the ability to automatically identify important features and accept unprocessed data as input for prediction. Long Short Term Memory (LSTM) [95-98], Gated Recurrent Units (GRU) and Deep Recurrent Neural Network are some of the examples of implementation of Deep learning models.[99], proposed a deep learning model which can deal missing values

`

efficiently in spatiotemporal prediction problems. [100] proposed an effective convolutional neural network for air pollution prediction from raw input images or photos. [101] proposed a model for air quality estimation using autoencoders in deep learning architecture. [102] proposed a model called "long short-term memory neural network extended (LSTME)" to estimate air pollution concentration considering spatiotemporal correlations. The model can be used in long term dependency time series modeling for determination of best time lag. [103] introduced a spatial-temporal estimation model based on the properties of air quality and utilized LSTM architecture to capture long-term and short-term dependencies. [104] had proposed a deep learning approach for predicting air pollutant concentration using LSTM in South Korea and examined the efficiency of using time series data to create prediction machines using encoder-decoder networks. [105] made an effort to use a deep learning framework that combined RNN and LSTM to forecast the air quality classification (AQC) of three distinct American industrial cities. The findings show that, depending on the feature sets and data lengths, the prediction models were effective for daily AQC forecasting. In order to forecast particulate matter concentration, [106] created a hybrid model combining CNN and convolutional LSTM. CNN was utilized for parallel temporal feature extraction while Convolutional LSTM was employed for sequential spatiotemporal data. [107] used the BLSTM model with transfer learning to forecast air quality on an hourly, daily, and weekly basis. In [108] presented a deep learning model to forecast Beijing city's average daily $PM_{2.5}$ concentration value the following day. The model used convolutional neural networks and long short-term memory. [109] suggested a spatiotemporal deep learning model for air pollution prediction across cities. It has been applied to model air pollutant using the residual neural network (ResNet). The ResNet is the one of the most recent replacement of CNN that could learn long term spatial dependencies. [110] used the ResNet with LSTM to forecast $PM_{2.5}$ in Beijing; this model perform better than alternative models like general recurrent neural networks. The deep learning network model proposed by [111] consists of the ResNet and the ConvLSTM. Its goal is to thoroughly extract the temporal and spatial distribution features of pollutant concentration and meteorological data from multiple cities. [112] used encoder decoder based air pollution model with LSTM network.In the network used to monitor pollutants, could lead to the loss of irregular topological data. In order to overcome these shortcomings, the graph convolutional network (GCN) is suggested

`

as a solution. It analyzes graph-structured data in order to identify topological relationships [113].

Earlier the focus was on various traditional statistical models like ARIMA, VAR and dynamical models such as CMAQ, CAMx, etc. In recent years, neural networks have been extensively used in air pollutant forecasting, but there was a limited study on the implementation of advanced machine learning techniques in air pollutant forecasting. Despite having some of the most polluted cities in the world, researchers have not given as much attention to the analysis and forecasting of multistep ahead air quality scenarios in Indian cities. This research work aims to close this gap by developing a model for forecasting air pollution for different Indian cities. The present study intended to explore the rich features of deep learning techniques in forecasting ambient aerosol concentration and its response with respect to signal by noise ratio across the country with the following objectives.

1.2.5 Objectives

The objectives of the study can be outlined as:

- To characterize the stochastic properties of ambient $PM_{2.5}$ concentrations in different regions of India.
- To forecast $PM_{2.5}$ concentrations using different deep learning modelling approaches.
- To determine the best performing forecasting model framework for ambient $PM_{2.5}$ concentrations in India.