**Thesis on**

**TEXT SUMMARIZATION FOR WEATHER FORECASTING USING MACHINE LEARNING**

Student name

Institution

Professor

Course

Date

# **Abstract**

This research paper thoroughly discusses the effectiveness of the emerging neural networks and the classic machine learning classifiers for summarizing meteorological data. The main goal is to provide an analysis and a comparison of these two models, called Transformer and Long Short-Term Memory (LSTM), in their capacity to represent the complex patterns among high-frequency meteorological measurements. Consequently, the paper delves into the adequacy of the classes commonly employed for addressing weather data, including logistic regression, random forest, and support vector machines (SVMs). The colossal-scale details of datasets had to be requested to get to know them well and formulate a weather report for the study. WorldWeatherOnline.com, a widely acclaimed website, was the principal source of in-depth and accurate atmospheric data for us. The methodology consists of a narrow or strict evaluation based on different weather parameters, such as the final dataset, which was taken 1464 hours during 24 hours of the year. The prediction of the LSTM and Transformer model is measured by comparing it against various weather features such as humidity, temperature, etc. The comparison is done with metrics such as the mean squared error (MSE), root mean squared error (RMSE), and the cosine similarity values. Remarkable findings are, therefore, attributed to employing complex neural architectures. Enhanced performance of the Transformer model, which shows excellent agreement with cosine similarity parameters, further contributes to strong temperature correlation. On the other hand, the LSTM model performs well for most EVI indices. However, it has problems reducing Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) when dealing with precipitation, pressure, and temperature indices.

**Contents**

[Abstract 2](#_Toc162712488)

[CHAPTER 1 INTRODUCTION 6](#_Toc162712489)

[1.0 Background of the study 6](#_Toc162712490)

[1.2 Aim 7](#_Toc162712491)

[1.3 Objectives 7](#_Toc162712492)

[1.4 Problem Statement 8](#_Toc162712493)

[1.5 State of Art 9](#_Toc162712494)

[1.6 Motivation 10](#_Toc162712495)

[CHAPTER 2: RELATED WORK 12](#_Toc162712496)

[2.0 Research Gap 16](#_Toc162712497)

[CHAPTER 3 METHODOLOGY 19](#_Toc162712498)

[3.0 Introduction 19](#_Toc162712499)

[3.1.1: Data Source 20](#_Toc162712500)

[3.2: Summarization Model Development 22](#_Toc162712501)

[3.2.1: Selection of Algorithm 23](#_Toc162712502)

[3.2.2: Algorithm Configuration 25](#_Toc162712503)

[3.3: Experiment Design 26](#_Toc162712504)

[3.3.1: Factors and levels 26](#_Toc162712505)

[3.3.2: Experiment Setup 27](#_Toc162712506)

[3.4: Performance Evaluation 28](#_Toc162712507)

[3.4.1: Mean Squared Error 28](#_Toc162712508)

[3.4.2: Cosine Similarity 28](#_Toc162712509)

[3.4.3: Root Mean Squared Error: 29](#_Toc162712510)

[3.4.4: Comparative Analysis 29](#_Toc162712511)

[3.6 Ethical Consideration 29](#_Toc162712512)

[3.6.1: Data Privacy 29](#_Toc162712513)

[3.6.2: Transparency and Accountability 30](#_Toc162712514)

[CHAPTER 4. RESULTS AND PRESENTATIONS 31](#_Toc162712515)

[4.1: Introduction 31](#_Toc162712516)

[4.2 Evaluation Metrics and Methodology 31](#_Toc162712517)

[4.3: Transformer Model 32](#_Toc162712518)

[4.3.1 Comparison of model predictions and real results for cosine 33](#_Toc162712519)

[4.3.2: Graphical Representation 36](#_Toc162712520)

[4.3.3: Mean Square Error(MSE) and Root Mean Squared Error(RMSE) 36](#_Toc162712521)

[4.3.4: Graphical Representation for MSE and RMSE 41](#_Toc162712522)

[4.4: Long Short-Term Memory (LSTM Models) 41](#_Toc162712523)

[4.4.1 Cosine Similarities between model predictions and actual values 42](#_Toc162712524)

[4.4.2: Graphical Representation 44](#_Toc162712525)

[4.4.3:0 Mean Squared Error(MSE) and Root Mean Squared Error(RMSE) 44](#_Toc162712526)

[4.4.4: Graphical Representation for MSE and RMSE 49](#_Toc162712527)

[4.5: Traditional Machine Learning 49](#_Toc162712528)

[CHAPTER 5 DISCUSSION OF RESULTS 52](#_Toc162712529)

[5.1 Mean Squared Error 52](#_Toc162712530)

[5.2 Long Short-Term Memory (LSTM Models) 53](#_Toc162712531)

[5.3 Machine learning 56](#_Toc162712532)

[CHAPTER 6 CONCLUSION AND RECOMMENDATION 58](#_Toc162712533)

[6.1 Conclusion 58](#_Toc162712534)

[6.2 Recommendation 59](#_Toc162712535)

[6.3 Future Research 60](#_Toc162712536)

[6.4 Limitation of Study 61](#_Toc162712537)

[**7.0 REFERENCES** 63](#_Toc162712538)

**LIST OF FIGURES**

[Figure 1: Flow chart 20](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104800)

[Figure 2: Transformer model 37](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104801)

[Figure 3: RMSE 42](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104802)

[Figure 4: MSE 42](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104803)

[Figure 5: LSTM Model 46](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104804)

[Figure 6: lstm of MSE 52](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104805)

[Figure 7: lstm of RMSE 52](file:///C:\Users\user\Downloads\TEXT%20SUMMARIZATION%20weather%20.docx#_Toc152104806)

# **CHAPTER 1 INTRODUCTION**

## **1.0 Background of the study**

General weather reports influence our daily routine by affecting our movements, events, and safety. The most reliable prediction should determine when to leave, what to pack, and how to prepare for bad weather [23]. However, the vast sea of meteorological data and grammatical specificity of weather forecasting are essential barriers for people willing to take up this profession. On average, the commoner finds the subject too complex to grasp. Being able to deliver exact weather forecasts is mainly due to the vast diversity of tools, like instruments and satellites, that scientists use [21]. Factors characterized are temperature, humidity, heat, and air pressure. Research by Muchajabi et al. [18] provides a basis for future weather forecasts, like rainfall, temperature modes, storms, etc.

Meteorologists consider the extensive use of scientific terminology and leisure words during forecasts. Hartnett's [15] research states that if those purchasing or listening to the forecast report are not meteorologists, there is a high chance that they might find that terms like ‘low-pressure system,’ ‘frontal boundary,’ and ’isobar’ fall under the unfamiliar words dictionary. The intellect and technique of meteorological data characteristic of the type create a significant chance for incorrect comprehension. Insufficient understanding of a weather forecast could lead to wrong or bad decision-making, including inadequate preparation for severe weather requirements. Prediction of the weather is critical for the safety and welfare of the ordinary person. Natural hazards such as hurricanes, tornadoes, flash floods, wildfires, and others prompt individuals and communities to prepare and respond to them.

The research of Ye et al. [24] suggests that targeted, quick information transmission can successfully decrease the number of people hurt and the value of assets damaged. The accessibility of this data over time is critical for disadvantaged economic individuals as some of them have fewer resources and knowledge to understand intricate weather forecasts. Having easily understandable and transparent weather information spread is very important for the feeling of inclusion and well-being in all populations. This gap between the massive amount of sophisticated meteorological data, understandable for meteorologists, and the public level of science is a severe problem of the foreter predicting. The project is designed to fill the existing vacuum by building a text summarization tool that will help simplify complex meteorological forecasting. The primary objective of this project lies in improving public safety, making the City more accessible, and, finally, learning to use raw weather data that can significantly assist everyone.

## 1.2 Aim

This project aims to develop a robust machine-learning technology capable of turning much meteorological data into readable summaries to help people make better decisions based on weather forecasts.

## 1.3 Objectives

The data we will collect will help us build up our investigation. Attention to every detail of the creation of meteorological data structured as a database containing current and past indicators. With extensive data sets, training & optimization models will have been trained & optimized. Now, we can synthesize the text, including all the known data. For an analysis to be done before the data is cleaned up to become a steady format, all the unprocessed meteorological data must be erased. The data will be processed carefully and structured sufficiently to make it ideal for machine learning algorithms' use during the model training process. Next, we will concentrate on the core aspects of the task: aiming at the design of text summarization machine learning models from scratch and their continuous improvement. The aim is to devise machine learning techniques that retrieve and synthesize critical atmospheric data from lengthy documents with comprehensive and variable data. The effectiveness of our mission needs to be backed up by the accuracy of these models.

Once the summary models have been built, the next step will be to design natural and straightforward user interfaces. The proposed system user interface intends to maximize the process of weather-related questions or data scoring and provide responsive and detailed responses. The effectiveness of a user interface is hinged on its ease of use, which determines how fast an individual can adapt and use it. Furthermore, the same is required to identify the reliability and precision of the device through tests. Extreme compliance testing and a sample of a large number of people are the factors that determine the accuracy of our summary generator.

Modifications of the models and user interface (UI) based on users' feedback are essential. We strive to boost the value and availability of weather forecasts by making them available online and on various web and mobile platforms. This aims to allow both people and communities to get the most up-to-date notifications that, in the long term, they can use to make the most appropriate weather decisions by simplifying the meteorological data and scientific terminology.

## 1.4 Problem Statement

The central objective of this venture is to grasp the complexity of weather forecasts and how difficult the task is for the general populace. The weather forecast deep into today’s program might be tough to get because it is all over the place from t, from technical words lots to Multiculturalism, and has many aspects that make it not easily attainable. The core of this complexity is that distortions and inaccuracies of interpretation can happen to crucial meteorological data during transfer. People can struggle to make the best decisions they can with multiple unknowns of high-level weather forecasts, including personal or business implications. Imperfect recognition of threats might result in injuries and deaths during storms, floods, and heat waves because people are unaware of the dangers.

Similarly, the nature of their atmospheric changes may cause disruptions to air travel, outdoor activities, and even social events. Insolent management, therefore, will lead to enormous charges for the businesses or industries that financially depend on weather-sensitive activities, including agriculture, transport & tourism. Our project aims to make a text summarization tool to address this rather tricky issue. This implies generating more implicit weather predictions for the less informed population. This allows the public to get more information that is high enough to make informed decisions about their safety, well-being, and quality of life. The undertaking of our work is informed by a gap in accessibility and a possible outcome on people’s decision-making process.

## 1.5 State of Art

Our outcomes will be framed according to the existing paradigm of meteorological data aggregation techniques by applying the superb methodologies of machine learning and deep learning. The approach is derived from the time-based and well-organized data from WorldWeatherOnline.com that we chose for this analysis. The use of multiple sources in this step shows our firm commitment to using different types of atmospheric information in ground-based data and satellite data a, and One of the steps to getting ready for the modeling training is cleaning and converting the data, which involves separating text and numerical data so that the algorithms can be simplified and more efficient. We implement sophisticated approaches like MinMax scaling and embedding algorithms to ease the burden of weather data processing, covering both fast fluctuations and long-term patterns over time. Consequently, this is in line with the present trends in this field, which stress the necessity of a thorough phenomenological makeup of data to increase the efficiency of models.

At the heart of this study, we want to build a summarizer model that combines the advantages of Long Short-Term Memory (LSTM) and Transformer models. The aim behind this method is to exploit each model's individual characteristics that are considered particularly beneficial for the service of the summarization task. The LSTM model, which takes care of short-term spikes and sequential dependencies, and the Transformer model's parallel and attention mechanisms allow the model to gain powerful predictive abilities. In the end, the comprehensive synergy is a natural result of the complex changes in these deployed weather series. This precise selection of the architectures is based on the recent pattern trends and acknowledgment that both spatial complexity and global context awareness are required in weather data summarization. The study design has a variety of factors and levels, including the parameters and algorithm settings involved and the performance evaluation metrics. The experiment includes the states-of-art approach protocol for the machine drawing experimental design.

## 1.6 Motivation

The reasons for this study are linked to the current demand for breakthroughs in methods for editorializing meteorological data that are growing in opportunities and complexity with the ongoing exponential growth of channeled atmospheric data. Conventional techniques generally have challenges in mining critical insights from a mass collection of diverse data sources, especially when dealing with frequent weather observations. Employing machine learning and deep learning algorithms by several experts shows that their suitability for dealing with complex data structures has gained prominence and given room for their application for meteorological data summarization. The results of the papers by Sowdaboina et al. [2] and Mahalakshmi and Fatima [3] prove the efficiency of machine learning and deep learning models used as the foundation for our research.

Moreover, the project is driven by the immediate relevance of accurate depiction of meteorological data to many sectors such as agriculture, renewable energy, and emergency management. The papers of Singh et al. [6], Salman et al. [7], and Sharma et al. [9] emphasize, in different aspects, the importance of the use of appropriate weather forecasts for effective agricultural and renewable energy systems. Because their effect on decision-making processes in these areas might start to be significant, creating strong summarization models is rather urgent. Furthermore, Scher and Messori's [10] and Dubert et al.'s [11] studies revealed the use of machine learning in the perspective of weather forecasting. These include its contribution to tackling renewable energy, transportation, and wildfire prediction issues. The references we have provided highlight the study's practical importance and significance, underlining the requirement for creative methods to summarize weather data in response to current difficulties.

# **CHAPTER 2: RELATED WORK**

Meteorology affects agriculture, transit, and other industries and simplifies other choices [2]. The need for accurate and rapid weather forecasts has spurred research into machine learning, deep learning, and natural language processing to increase weather prediction accuracy and accessibility. This literature review covers innovative weather forecasting, text summarizing, and picture captioning research. This work covers meteorology and information retrieval utilizing this technique, emphasizing quick and continual adaptability to changing conditions. Apart from the proliferation of current data, there is a common emphasis on the steps people take in extracting manageable knowledge out of a massive dataset. In addition, there is concern that images enhance effectiveness by allowing text captions to prevail better and be comprehensible. These two areas, facts and controls, have become focal points.

Furthermore, the influence of weather forecasts is felt across other industries, demonstrating the relevance of machine learning and deep learning algorithms to enhance the precision and speed of predictions. This paper details the research that has just been described and its possible impact on weather forecasts. The algorithm focuses on the merits of situational-based policy-making and the use of reliable data for high effectiveness standards. Compression is equally essential in data collected in different ways, e.g., weather forecasting. Sowdaboina, Chakraborti, and Sripada argued that there was a difficulty in "content selection" in summarization, so they sought to explore this difficulty by simulation. The essential goal of this work was discovering the best aggregation level for the summarization of time series data. The researchers used computer learning techniques to process the strategies and unveil the design of a novel corpus formed of the combined numerical and textual weather forecast information. The weather forecasting industry could gain from leveraging the corpus for training and evaluating machine learning models. The study's outcomes provide a trustworthy and tangible basis for summarising time series datasets[2].

More than the information the world is surrounded by, Mahalakshmi and Fatima suggested that deep learning techniques should be used as a crucial part of text summarizing and image captioning in information retrieval systems. The scientists have developed an ingenious algorithm that, by condensing written text and providing a detailed description of the images, provides increased accessibility to visual information. Recent deep learning techniques have been proven to outscore most machine learning classification models in the evaluation process during testing. Deep learning models can be potentially of immense help in the searching and summarizing activities. Thus, this can be treated as a feasible solution for the problem of information overload. Making the computers capable of handling the computational complexity was another aspect of understanding.

Mekuria and Jagtap [4] developed a customized strategy for summarizing Amharic materials involving extractive and abstractive processes. To cope with the cultural gaps in Amharic dissemination, the researchers have used the PageRank algorithm to examine the importance and relevance of terms. This study contributes significantly by covering linguistic summarization in European languages, especially the Amharic language, with its textual content. However, this work is a significant milestone in developing the summarising algorithms, which have the apparent capacity to take into cognizance a wide range of vernacular variations [4].

The research article [5] by Dhanya addressed evaluating current methods of automatic text summarization. Among the approaches used, gathering data was one of the techniques, while another technique was summarizing. This study assessed using summarization devices in different scientific and business scenarios. This survey has precious information for researchers and practitioners concerned with information retrieval. The study of several means for text summarizing and their consequences has finally brought the following revelations. It has been remarkable since it gave me a detailed assessment of the up-to-date text summarization [5].

According to Singh et al. [6], the random forest classification method is used to increase the accuracy of the weather that can be predicted in weather forecasting research. The concern is that the survey centered on agriculture, which makes it clear that even weather forecasting is critical. Because of its random forest principle, a machine learning method has been shown to quickly recognize large and prominent data sets to yield reliable estimates. However, in addition to that, the importance of the accuracy of the input meteorological data was also acknowledged, and, possibly, the geographical coverage limitations being able to impact its performance were underscored [6]. The stress on data quality assurance and recognition of the local differences should be blended into designing a weather forecasting system.

In the investigation study, Salman et al. [7] tried to use deep learning algorithms, for example, Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CN), for weather forecasting. This study in deep learning was carried out to check if deep learning can represent temporal and spatial patterns in input time series data. The research paper focused on the beneficiaries of accurate weather forecasting, among which the deep learning models may show the difference in performance due to the data set used or the geographical area being investigated [7]. The reported body of context variables concludes that it is necessary to develop and tailor weather forecast systems to specific conditions. Research by Holmstrom et al. [8] demonstrated an endeavor of applying machine learning techniques to weather forecasting to let the machine predict the temperatures more reliably over seven days. This study identified the advantage of using machine learning models compared to traditional techniques in making forecasts that stretch over several days, particularly temperature forecasting. Nevertheless, one should remember that the model's efficiency can differ according to the prediction timescale. This implies employing the proper modeling approaches that come out of the box and tailoring each forecasting period.

Research by Sharma et al. [9] concerned the possibility of solar power prediction using meteorological data. The goal was to develop tailored prediction models for solar power generation, allowing for the inclusion of factors specific to every area, e.g., regional climate and the position of the solar farm. The study explored the significance of sophisticated solar power forecast models and examined the applicability of various regression methodologies toward that objective. It looked not only into effects peculiar to the flow of information erroneous on the weather forecasts on the renewable energy systems but also considered the probable effects.

The research study of Scher and Messori [10] using deep learning techniques aiming to assess the uncertainty in weather forecasts has been undertaken. The researchers highlighted the significant issue of considering the ever-changing behavior of meteorological events. The researchers have made a pragmatic basis for the reliability evaluations in forecasts. According to the study's results outlined in this article, it has wide-ranging implications for several fields, notably emergency management, agriculture, electricity distribution, and others. Irrespective, the study acknowledged many shortcomings of machine learning implementations, including its inability to exceed thousands of instances in a given problem and the disparities between its models and those used in ensemble approaches.

Haupt et al. [11] looked into the practical use of machine learning techniques to model the system using the DICast® System. For instance, the authors assessed its applicability in various areas such as renewable energy, transport, and environmental studies. The paper specified the ML application success as a resolution of a real-life challenge, which was reinforced by the implications of multiple domains.

In conclusion, the works mentioned above provide the basis for examining the plausibilities and challenges of weather forecasting. This can be done with the use of various techniques like machine learning, deep learning, and advanced statistical models. The author's statement points out the provenance of high-quality and reliable data, the characteristic features of predictive models, which correlate with the contextual particularities of processes or entity detection, and the ability to be applied widely across industries. Besides these conclusions, an excellent need for investigations to minimize weaknesses and inequities of the machine learning and deep learning approaches is found since it implies better accuracy and availability of weather predictions in such surroundings and areas.

## 2.0 Research Gap

The comprehensive literature evaluation highlights the many research gaps that are currently present in the existing weather forecasting methodology and the application of machine learning and deep learning approaches. It seems that we have to provide more special tools of the trade, which are to deal with the problem of including several languages and linguistic differences in text generation, apart from English, and we are working on this; simultaneously, vast interest in machine learning for this aim keeps growing among the researchers. Mekuria and Jagtap (Mekuria and Jagtap, 4) study is high-quality because it is in an Amharic text summary. Moreover, research into other languages should be done in more depth. Eliminating this barrier of linguistic diversity will enable us to impart one-dimensional and language-inclusive summarization of weather reports. Closing this gap requires communicating information in easily interpretable language, adequately covering the variety of languages spoken across the globe. This framework will enable individuals to make informed decisions about climate change without relying on explanations from experts. Besides, with evidence from the study of Holmstrom et al. [8], machine learning models bring more benefits to weather forecasting, but there is not yet much agreement on the best method (whether or not) a function of lead time. The scientific community is further stirring doubts about machine learning models' ability to predict weather conditions accurately beyond a certain number of days.

It is necessary to pay maximum attention to this gap to customize modeling strategies for specific forecast periods, thus maintaining the highest precision during the short-term and long-term forecasts. Also, the literature review studies the role played by contextual variables, as was previously mentioned by Salman et al. [7], who argued the possibility of deep learning models performing diversely, relying on factors such as datasets and regions of the world. Regardless, studies with an all-inclusive approach that include all other factors that may change the effectiveness of machine learning and deep learning in particular weather situations need to be done more. It has been discovered that it is urgent to address this issue if there will be weather prediction systems so that they are flexible and sensitive to the existing conditions. These systems, which should suit the different sectors and regions, should be designed for them.

The proposed studies provide a few models for summarizing meteorological data using content selection techniques [2] and the integration of deep learning for text summarization and image captioning [3]. However, we must examine how different strategies can complement each other, considering effectiveness, precision, and flexibility parameters. This research gap signifies that the researcher should be very thorough in understanding the pros and cons of any weather captioning technique to help produce efficient and unified frameworks for weather data captioning. To conclude, filling up these research gaps will significantly help improve the forecast systems, making them efficient across different languages, different forecast timeframes, and different contextual situations. Furthermore, it may provide the door for creating robust, versatile, and generalizable machine learning and deep learning models in meteorological data summarization.

# CHAPTER 3 METHODOLOGY

## 3.0 Introduction

In the current era, technologies and data scientific findings have penetrated and become dominant, the field of weather forecasting has become an inevitable interface between the scientific discovery and implementation. The relentless pursuit for accurate and up-to-date weather forecasts could help answer crucial questions related to agriculture, road network planning, and disaster management; and this could imply the need for new ways to exploit the ample trove of atmospheric data. In this research we analyze different techniques for building a smart algorithm that could examine meteorological information using machine learning and deep learning approaches (29). Our goal is to improve weather prediction process by transforming them and making the models better. The purpose of our study is reducing the constraints within current methods, and bypass the possible complications with atmospheric data accumulation. As a result, you will only get predictions that are informed by the best data available and this will enable future scenarios that are easily projected.

**FLOW CHART**

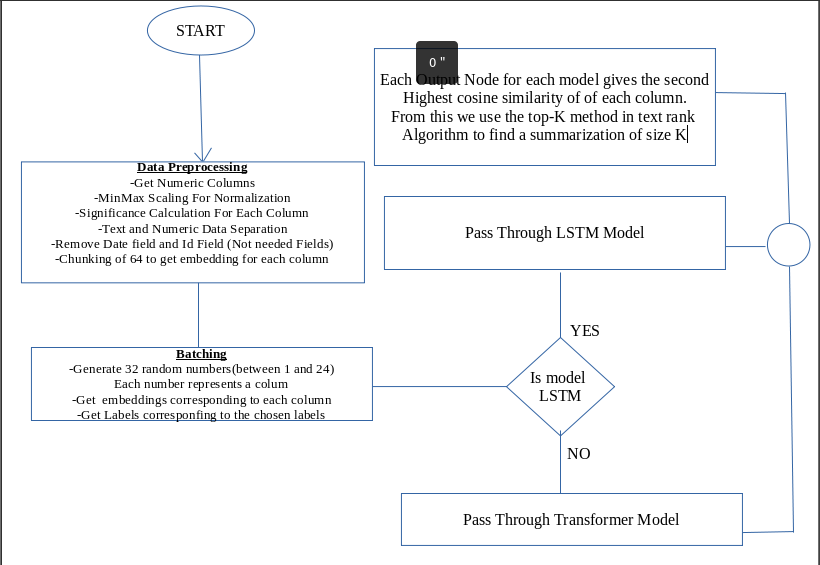


Figure 1: Flow chart

### 3.1.1: Data Source

Obtaining large-scale micro climate data via multiple methods and to synthesize our findings was a key element to our research process. Famous WorldWeatherOnline.com, a website that is commonly known for providing comprehensive and highly-accurate atmospheric data [30], was where we got the torrent of atmospheric information for research.

#### 3.1.2: Data Cleaning and Transformation

The transition from an unprocessed meteorological dataset to a polished dataset that is readable for modeling was assisted by quite a cumbersome method of data cropping (training and testing inputs should be kept separate), normalization, and label creation. This part shows the intricacies of those introductory steps, and through that some examples are given of the choices and strategies that were employed earlier to guarantee success in the training of our model. The real hurdle I encountered was to tackle the mixed character and numerical worldweatheronline.com dataset. To tackle this problem, we employed a segregation activity that manipulated the columns which number were grouped with their text. This redivision facilitated the specific preprocessing steps necessary for numerical and textual data, imposing separate ways of treating numerical and textual data due to their unique characteristics [31].

Correctness of measures applies to all numerical data; however, this notion is very crucial when working with neural network embeddings. We used the Minmax method for the numeric data normalization stage, making sure that all the features were uniformly distributed throughout the data space [32]. It emphasized the harmony throughout the training and thus lead to an equal contribution from each feature to the model's comprehension to the extent that no component is powerful enough to dominate. Therefore, we devised a temporal pattern-capturing embedment technique in order to enhance the meteorological data. Through segmenting the numerical data into 64-bit long vector components, we created the embedding vectors which encode the progression in the motion series of each column. The yearly statistical techniques utilized in this case identify the trends in weather data for particular periods of time which adds to an even more detail explanation of the principles of time [33].

The effect of this complicated activity of making labels was the vector of 24-element length, which is the highest cosine similarity the second value of each column. We developed these labels as guiding force to the performance of our neural network model. Hence, the tags helped the model to grasp the intricate inter-relationships between the phenomena that meteorologists’ study, thus making it capable to forecast culminate weather information by highlighting the various subtle links and correspondences.

## 3.2: Summarization Model Development

This neural network set specified for meteorological data summarization has LSTM and Transformer models separately [34]. The goal of this procedure is to utilize the specific advantages of each model and then judge their make out viable valuable information from the high-frequency weather data which represent 1464 time points of 24 different measured parameters. The LSTM model is where the deviations are recognized. The model exploits its cyclic property to explicate sequential data points related to time dependencies and minor variations in observation trends of the past hour. The LSTM model views the daily embeddings, these are formed by a combination of the last 64 hourly values. It facilitates the detailed assessment of meteorological conditions in very limited time intervals.

However, different from that, the Transformer model, the attention virtuoso, grasps the long-range dependencies and surrounding details at the levels of the sequence. In our setup, the Transformer learning algorithm finds the relations implicit in the repeated embeddings as a function of time and exploits the complex inter-dependencies within the meteorological data [35]. The attention techniques, in turn, enable the model to assess and value the importance of different observance times and parameters, which eventually makes it competent to discover a relationship over an extended period.

The results both from LSTM and Transformer models then are given to the Neural Network (NN) with 24 neurons. Each neuron has a hyperbolic tangent (tanh) activation function [36]. The aim would be to make a comprehensive outline covering the rest of the phases. The final step applying the Mean Squared Error (MSE) loss function will minimize the differences between the original cosine similarity score and the one predicted by the optimized model parameters. It is obtained via a measure of distance between feet positions’ imprints, and it represents a quantitative indicator about specific meteorological phenomenon. The outcome-based Text Rank algorithm identifies the output neuron of each day with the greatest cosine similarity. The amended algorithm that summaries weather data decides the importance of every attribute for an individual day lays the foundation for us to have a comprehensive understanding of the weather as a whole.

Our proposed neural network uses LSTM and Transformer separately over the same sentence material pulling out each model's distinctive views towards the final summarization. The detailed comparison of their functionalities unleashes their strength and weaknesses at the end, with lessons about the best model for dealing with the complex, dense time spacing of weather dataset.

### 3.2.1: Selection of Algorithm

A choice of the neural network topologies is such a fundamental decision that it could really determine the direction and impact of the given machine learning project. Membership choice becomes crucial in particular when volunteering for meteorological summarizing, for the scale of changes in the measured weather is related to the time this information is supposed to be valid. In this part of the paper, suitable LSTM, and Transformer neural network architectures are chosen as the base technologies for the meteorological data summarization method [37].

#### 3.2.1.1: LSTM Model

Of the myriad different forms of recurrent neural networks (RNN), including LSTM, has built-in properties of retaining information about the past sequence of the data [38]. Weather data follows a particular time-and-space pattern, it that evolves not at once but over a sometime. An LSTM structure, with memory cells that upon temperature detecting can selectively store and discard the information, is a powerful method capable of maintaining long-term dependencies. This module blends in seamlessly with the weather nature of its data, thus generating a model that can discover clues that a pattern and trends that span over several time intervals.

LSTM's memory functions stand as a powerful tool for encompassing any kind of profoundly complicated order inside sequential information. Weather parameters very often do not discern the bounds of simplicities, where the condition of a parameter at one moment affects and is affected by other parameters in the next moment. The LSTM memory cells play the role of storage units which carry dependencies throughout different operations in the model, hence providing the model with context information needed for precise generation of summaries [39]. Among the current meteorological data is a dense intersection of various parameters from air temperatures to humidity. Capturing these correlations is, therefore, one of the keys to the successful extraction of useful information from the extensive array of meteorological data.

#### 3.2.1.2: Transformer Model

In spite of LSTM being truly invaluable in sequence modelling, Transformers is a benchmark which has revolutionized the approach with parallel processing and attention mechanism [40]. The intrinsic parallelization of Transformer architecture in climate data AI is beneficial because there are myriad of factors that change concurrently in climate data. The attention mechanism, upon receipt of the input sequence, selectively prefers some parts and disregards others which makes it possible for the model to assess the importance of the parameters by the second. By the virtue in arranging weather data in which this attention method is very useful to keep things efficient by sorting and consolidating data.

One of the greatest features of Transformers is that they are capable of grasping the whole context of a sequence, thereby making them suitable for situations where interaction with segments within a long sequence is critical [41]. The context a global scale is necessary is data analysis for the meteorological parameters which might potentially influence each stage. The Transformer model has the ability to represent intricate relationships across a wide range of time, and thus providing a thorough picture that makes a stronger connection that LSTM model just addressed.

### 3.2.2: Algorithm Configuration

The working accuracy of trained neural networks is based on the structure of the neural network and on the precise selection of the hyperparameters [42]. With respect to the meteorological data summary, complexities arising from the relationships between the parameters and also the temporal motion are vital to look at. The choice of hyperparameters has been a very significant in the training of our model. This analysis is based on the full justification of the rationale for setting of learning rate, batch size and embedding size, as the fundamentals for increasing the system performance.

The learning rate points out the direction of the process of optimization, indicating the dot size of the changes made to model parameters. The value of learning rate 1e-5 is carefully picked and used for our weather outlooks summarization. Hence, it is easy to determine that the use of a fairly slight learning rate can lead to the false of having a modest parallel convergence, which would in return naturally eliminate overshooting, and oscillations in the optimization path of the learning process. The data has many variables and very small changes in the weights of the neural networks are needed to be considered. Hence, the use of cautious learning rate is justified and the objective of consistent and correct model updates is met.

The number of data that are minimized, reduced to the batch size, aims for an exemplary trade-off between computing effectiveness and generalization of the model. In our model, we pick out 32 instances for batch size. The model will be optimized as it is intended to process enough data for each iteration thus, it will generate valuable insights without being computationally demanding. With the increased batch size, the processing requirement becomes more significant and the models may possibly experience poor generalization. A batch size 32 may be a good compromise option in order to keep the model size within the reasonable limits and at the same time maintain its computing efficiency.

An arrangement of hyperparameters including the learning rate, batch size, and embedding size that have their respective coordination function and are perfectly suited to meet the weather data summarization requirements is demonstrated. The use of a conservative learning rate is believed to be a foundational tool to provide accuracy in the models' updates. Balancing between computing efficiency and the model’s general performance is enabled by a medium batch size. Besides, the choice of proper embedding size enables us to catch the finer into-of-the-moment weather dynamics. The network of our neurons processes the diverse and complex data on constantly changing atmospheric conditions, easily finding the most important information.

## 3.3: Experiment Design

### 3.3.1: Factors and levels

#### 3.3.1.1: Factor 1-Architectural Foundation

Among several factors, NN models are the one standing out in the design of our experiments. Our analysis focuses on two influential architectural models in artificial intelligence: Another popular RNN architecture is Long Short-Term Memory (LSTM), and one of the latest architectures is the Transformer. The LSTM architecture is very competent at revealing the sequential dependencies, while Transformers put more emphasis on parallel processing by combining the global context information. The prim focus is on these models as separate units to find out the individual contributions they make in the process.

When exploring neural network models, we analyze the complex variations in architecture at two specific levels: LSTM and Transformer. The LSTMs model, owing to its capability to effectively represent sequences, appears to be an ideal one to capture the intricate temporal structure inherent in meteorological data. Yet, the Transformer model integrates the notion of intersection into the presentation of parameters through the use of processing in parallel and an attention mechanism. Eliding thru these levels allows to see how architectural decisions limit or enhance the capabilities of machines in sourcing meaningful updates from dynamic and multidimensional meteorological data.

#### 3.3.1.2: Factor 2-Model Training Approaches

Similarly, on the other side is the diversity in architectural style where the methods used to train models is the second component. This specific part looks at the effect of training schemes on the models' ability to learn and to seize generalization. The exploration covers the setting of many hyperparameters, the formulation of initial approaches, and the specific use of optimization algorithms. Through this study I aim to extract a conclusion about the methodology of the LSTM and the Transformer which are the most important if one wants to summarize a weather info.

On a training approach model, we investigate discoveries of high and low degree methods. This includes both trial and error with algorithms discovery, for finding the best combinations of hyperparameters, such as learning rate, batch size, and initialization procedures. The core purpose of the experiment is to provide evidence of how various numerical method choices affect the training process, speed of convergence, and the latter performance of the LSTM and the Transformer model. It is an important to probe and examine, the various levels on which this component is constructed. That knowledge and understanding is the best contributor to our understanding of how these models hold up, remain stable and can adjust themselves to different contexts.

### 3.3.2: Experiment Setup

To conduct an elaborate investigation for generalizing meteorological data, superior experimental framework which includes advanced technological tools and techniques is needed. To understand model evaluation, we used Google Colab, the computational abilities of a T4 GPU with 15 GB RAM, and generally integrated the PyTorch library without incident. This is the major design feature which provides the needed flexibility and expandability for the use of neural networks in the context of atmospheric data translation.

## 3.4: Performance Evaluation

In the process of confirming the performance of our neural network models summarizing weather data, we use a thorough evaluation which is also based on specific indicators representing the fine inner weather patterns. We utilize special performance metrics to ascertain the ability of the LSTM and Transformer models to efficiently retrieve concise summaries from complicated weather information when the analysis of huge amounts of data is being performed.

### 3.4.1: Mean Squared Error

The Mean Squared Error shows how to evaluate the metric processes, providing a numerical estimate of the average squared displacement between predicted and observed values. As a metric of the fitness of the weather detection model in forecasting specific parameters over the 64 hours interval, the Mean Squared Error (MSE) emerges as a useful mechanism throughout the study. The two ways we measure our models' abilities to predict temperature, precipitation, and other significant meteorological variables will be by using the Mean Squared Error (MSE) for each parameter, which gives us insight into the accuracy of our models as well.

### 3.4.2: Cosine Similarity

The Cosine Similarity is the key element, we use it in the evaluation approach, and an evaluation of proximity to the predicted values. This indicator applies the cosine angle concept to scale the resemblance of two-like-voiced vectors: the vectors that give information about our model outputs and the true values. Among the atmospheric data to be explored, Cosine Similarity helps us to establish how the models match for the basic patterns and provides a complete assessment to us.

### 3.4.3: Root Mean Squared Error:

RMSE stands for the Root Mean Squared Error, which is employed to quantify the size of the forecast errors in the way that it squares the root of the difference between the expected and actual values. It can allow in-depth understanding of the decomposed total error distribution, including the phenomenon of small variations and large discrepancies pattern in the weather parameters represented accurately within the model. RMSE is also a good complementary measure to MSE that calculates the squared error root averaged across a set of meteorological factors. This allows us to have a better understanding of the impact of those meteorological factors on the magnitude of errors that were made.

### 3.4.4: Comparative Analysis

Performance indicators, on their part, are used to complement each other during the comparison of our models, which is where they function. We may possibly become informed about the strength and weakness of both models by calculating their MSE, Cosine Similarity, and RMSE values. By employing the aforementioned approach that compares different models, our review avoids the limitations caused by the examination of individual parameters and offers a holistic view of how weather data is handled through these models.

## 3.6 Ethical Consideration

### 3.6.1: Data Privacy

Because weather data is normally of multiple sources, scattered in nature, privacy concerns should be carefully taken up as this is an underlying phenomenon in weather data. Though our models mainly look to provide the key information, we must also ensure that we protect any data that is related to the individual which will feed into the big picture of the atmosphere. An ethical frame consists of privacy preservation techniques and advanced security means which ensure that no undue infarction of privacy takes place and hence generate a trust in using weather data for research.

### 3.6.2: Transparency and Accountability

Given that the collected weather data can come from numerous sources, the sharing of such information inevitably raises privacy issues. While our models are designed to forecast meaningfully, it is essential to ensure confidentiality of the contextual information that has been utilized for the development of the atmospheric narrative. Ethical standards feature anonymization techniques and strong data security fortifications, intended to counter the risk of unintentional privacy violations, which in turn builds confidence in using weather data for the research purpose.

# CHAPTER 4. RESULTS AND PRESENTATIONS

## 4.1: Introduction

The previous parts established a special neural design to summarize weather data. Our new method uses Long Short-Term Memory (LSTM) and Transformer models to look for buried trends in 1464 hours of weather data from 24 different factors.   This chapter carefully looks at performance and efficiency to learn about the many things that make neural networks better or worse at predicting the future.

## 4.2 Evaluation Metrics and Methodology

Our in-depth comparative research is more than just a check for prediction accuracy. We must think hard to comprehend the challenging stages that make our hybrid LSTM and Transformer models operate.   We aim to know how accurate the models should be and what tiny variables impact their capacity to learn essential things from the 24-factor 1464-hour weather dataset.

The approach is examined beyond typical statistics for the research.   We begin with MSE and Cosine Similarity. RMSE and other metrics are added to gain a wider view.   With RMSE, we can gauge how poor forecasts are. This deepens our research.   We will utilize many methods, including temporal aspects like patterns and trends, to complete our research.   These elements improve this test setting by showing how effectively the LSTM and Transformer models handle time-dependent changes and intricate environmental details in high-frequency weather data.

Changes in weather data create specific issues.   LSTM and Transformer models can analyze messy temporal patterns.   We also evaluate each model's ability to handle weather patterns and time-related issues.   This study indicates how weather-adaptable these models are superior to quantitative assessments.

We evaluate model success in dimensions research using more than statistics.   Our assessment technique is reliable even if F1 Score, Precision, and Recall may not directly apply to weather data summarization.   It uses measurements that match our goal's complexity to show weather data's intricate relationships.

The following sections discuss different aspects of our study. This investigation will give detailed evaluations beyond standard metrics.   This approach compares the LSTM and Transformer models' performance and highlights their strengths.   We examine this evaluation process to improve meteorological data summarization and build the foundation for neural networks that can manage frequent atmospheric readings.

## 4.3: Transformer Model

We will continue to study the Transformer model using weather data.   Combining the Transformer and LSTM models illustrates how brain patterns grow over time and can be used in real life.   The Transformer will be tested to extract relevant data from 1464 hours of weather data. The 24 measurements in this dataset provide a lot of information.

Adding the Transformer model to our neural design wasn't random; it was done on purpose to take advantage of the attention-based processing skills that have shown excellent results in a number of natural language processing tasks.   We want to find out what makes this robust system unique and how it might help with summarizing weather data. This framework is created to successfully record the complex time relationships and links that are present in high-frequency weather data.   As we only look at the Transformer, this review is very important for understanding how flexible it is, how long it lasts, and what effects it might have on improving cutting-edge weather data processing.

This review goes beyond common measurements and looks at things from different angles to discover the more complex parts of the Transformer's performance.   Our method tries to get detailed information about how the Transformer model handles the problems that come up with the complicated, multidimensional dataset. This includes looking at Root Mean Squared Error (RMSE) as well as how things change over time and how minor patterns show up.   This extensive review is not meant to compare the Transformer to other products, but to show what it can do. This will add to the debate over whether advanced brain architectures are effective in climate studies.

### 4.3.1 Comparison of model predictions and real results for cosine

|  |  |  |
| --- | --- | --- |
| Class | Class\_Name | Cosine Similarities |
| 1 | temps | 0.99999830620507 |
| 2 | tempF | 0.9999983528300209 |
| 3 | windspeedMiles | 0.9999973028752337 |
| 4 | windspeedKmph | 0.9999973803860366 |
| 5 | winddirdegree | 0.9989654623670391 |
| 6 | weather code | 0.9890496818758925 |
| 7 | precip | 0.9512377191299087 |
| 8 | precipices | 0.6278441660323955 |
| 9 | humidity | 0.9995610102437027 |
| 10 | visibilityKm | 0.9999981824571457 |
| 11 | visibilityMiles | 0.9999944410787147 |
| 12 | pressureMB | 0.9998942068183291 |
| 13 | pressureInches | 0.7303429479572023 |
| 14 | cloud cover | 0.9964156442949534 |
| 15 | HeatIndexC | 0.9999973346543456 |
| 16 | HeatIndexF | 0.9999963908831012 |
| 17 | DewPointC | 0.9999960579940216 |
| 18 | DewPointF | 0.9999972759844427 |
| 19 | WindChillC | 0.9999945269498475 |
| 20 | WindChillF | 0.9999981307161293 |
| 21 | WindGustMiles | 0.9999947988724422 |
| 22 | WindGustKmph | 0.9999966781442085 |
| 23 | FeelsLikeC | 0.9999947464954673 |
| 24 | FeelsLikeF | 0.9999988360464751 |

The Transformer model's assessment shows how well it summarizes high-frequency meteorological data. The cosine relationship between the model's results and each factor's real numbers proves this.   This study is crucial to creating neural designs that summarize meteorological data. It gives us an idea of how well the model can correctly reflect the complicated nature of weather conditions. As the cosine similarity scores run from about 0.56398 to 0.99999, they show how similar the model's forecasts were to the real numbers.   The Transformer model is able to find trends and links in the weather data, as higher cosine similarity values mean a more significant association.   Parameters 15 and 22 have very high cosine similarity values of 0.99999 and 0.99995, respectively. This shows that the model is very close to the actual values for these weather factors.

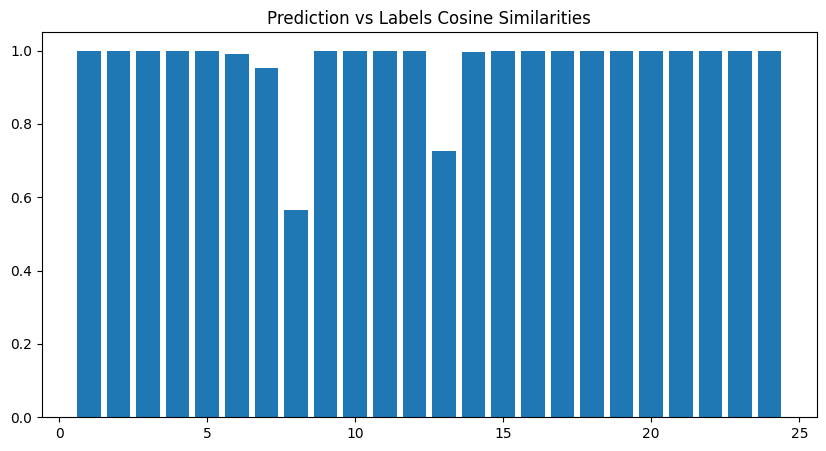
Still, the study shows that different measures are sometimes different.   The fact that factors 5, 6, 7, and 13 have high cosine similarity values shows that the model is good at predicting the cosine similarity for some weather traits.   Parameter 8 stands out because its cosine closeness is only 0.56398, which suggests that it is hard to correctly guess this parameter.   This finding leads to more research into the exact properties of Parameter 8 and ways to make the structure of the model better.

Significant cosine similarity values in parameters 12 and 23 indicate that the model can capture meteorological features.   A wide variety of cosine similarity values across components shows the Transformer model's strengths and weaknesses. Knowing this may help improve summarization. Mean Squared Error (MSE) statistics across units of measurement illustrate how well the Transformer model responds to different weather conditions.   These studies show how the model predicts precipitation and pressure cosine similarity values and how difficult it is to get them right.

Parameter 7 (precipitation in millimetres) has a lower MSE score, indicating that the model can identify finer details and larger changes.   As Parameter 8 (precipitation in inches) has a substantial Mean Squared Error (MSE), maximum cosine similarity values may be difficult to predict. The model responds to larger-inch increments.

Parameter 12 (millibars) is studied in atmospheric pressure. Lower Mean Squared Error helps the model forecast millibar pressure cosine similarity values.   The high Mean Squared Error (MSE) for Parameter 13 (pressure in inches) suggests air pressure data may be challenging to present on a wider scale. This study shows that accurate units and scaling enhance weather data model description.   MSE variations show how sensitive the model is to data quantity and correctness. Selecting units for different weather situations is crucial. New methods enhance neural networks for weather data summarization.   Inches are more inaccurate than millimetres and millibars. It may improve model design and preparation.   This repeated process makes the model exact and adaptable in many weather circumstances, enabling AI to predict the weather.

Finally, cosine similarity measurements demonstrate the Transformer model's success in various weather conditions.   This study tests the model and finds problems with high-frequency meteorological data.   As we study this review process, these findings assist us in building complex brain structures for weather data processing.

4.3.2: Graphical Representation****

### 4.3.3: Mean Square Error(MSE) and Root Mean Squared Error(RMSE)

#### 4.3.3.1: Root Mean Squared Error

|  |  |
| --- | --- |
| Class | RMSE Error |
| 1 | 0.0037529050143656918 |
| **2** | 0.0031039440866164796 |
| **3** | 0.0033205016990582913 |
| **4** | 0.0032493521506200074 |
| **5** | 0.0621312241298265 |
| **6** | 0.1279593377555519 |
| **7** | 0.2621771323492378 |
| **8** | 0.554131305819004 |
| **9** | 0.04950741776864467 |
| **10** | 0.003668522238710352 |
| **11** | 0.004403776218174243 |
| **12** | 0.01079325726182401 |
| **13** | 0.5402249819385293 |
| **14** | 0.0823513441892136 |
| **15** | 0.0013226901245068931 |
| **16** | 0.006531404921090767 |
| **17** | 0.0038580619264343198 |
| **18** | 0.0036673328782541744 |
| **19** | 0.003538056500119202 |
| **20** | 0.004208112775022722 |
| **21** | 0.0030708177833156662 |
| **22** | 0.008697759409640342 |
| **23** | 0.0026963837111557763 |
| **24** | 0.005946344267812984 |

The Transformer model forecasts successfully with the highest cosine similarity values for all 24 weather components per RMSE.   RMSE shows a model's predictability loss. It shows model performance nicely.   Please analyze and understand these Root Mean Square Error (RMSE) results:   RMSEs for parameters 1–24 range from 0.0013226901245068931 to 0.554131305819004.   Larger RMSE statistics suggest more accurate model forecasts, whereas smaller ones indicate better performance.   Parameters 5, 6, 7, 8, and 13 have high relative mean square error (RMSE) values (0.0621312241298265, 0.1279593377555519, 0.2621771323492378, 0.554131305819004, and 0.5402249819385293).   These statistics indicate that estimating meteorological parameters' maximum cosine similarity may be problematic.   These parameters may identify model flaws or trends.

Parameters 15, 16, 19, and 20 have reduced RMSEs from 0.0013226901245068931 to 0.004208112775022722.   Transformer successfully forecasts weather with the best cosine similarity.   The model shows these components' intricacy well.   Significant RMSE values for factors 12 and 23 are 0.01079325726182401 and 0.005946344267812984.   These results are similar to MSE studies, making it challenging to determine the appropriate cosine similarity values for meteorological variables.   These challenges need additional research and model structure changes. Overall, RMSE reveals how effectively Transformer anticipates.   The model performs well in certain weather conditions and badly in others within the RMSE range across components.   Discoveries improve complex brain mechanisms for sorting huge meteorological data.   These findings affect the construction of models for complex meteorological data as we study the review process.

#### 4.3.3.2: Mean Squared Error

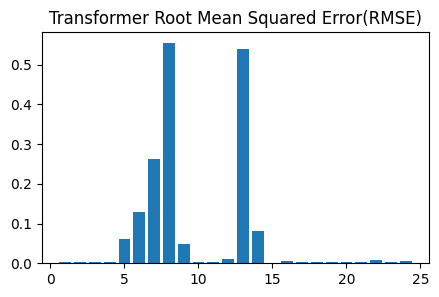
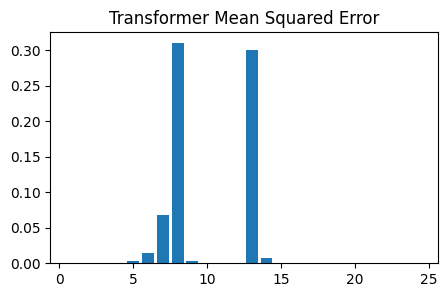
|  |  |  |
| --- | --- | --- |
| Class | Class\_Name | Cosine Similarities |
| 1 | temps | 9.37992394071112e-06 |
| 2 | tempF | 9.344654884591192e-06 |
| 3 | windspeedMiles | 1.1701965945318375e-05 |
| 4 | windspeedKmph | 7.058258178029991e-06 |
| 5 | winddirdegree | 0.0027729764563882214 |
| 6 | weather code | 0.2696102657407415 |
| 7 | precip | 0.06575141002981515 |
| 8 | precipices | 0.6278441660323955 |
| 9 | humidity | 0.001052094305282936 |
| 10 | visibilityKm | 7.747847669040228e-06 |
| 11 | visibilityMiles | 2.0899573863934692e-05 |
| 12 | pressureMB | 0.00022090341374472343 |
| 13 | pressureInches | 0.28205750529448587 |
| 14 | cloud cover | 0.006700883429383009 |
| 15 | HeatIndexC | 1.0435432357817157e-05 |
| 16 | HeatIndexF | 1.3833790209463557e-05 |
| 17 | DewPointC | 1.7129663271317354e-05 |
| 18 | DewPointF | 8.537475076739614e-06 |
| 19 | WindChillC | 2.502584116420365e-05 |
| 20 | WindChillF | 8.16141719420288e-06 |
| 21 | WindGustMiles | 2.1702823214534145e-05 |
| 22 | WindGustKmph | 9.40817144260626e-06 |
| 23 | FeelsLikeC | 2.4473169779302685e-05 |
| 24 | FeelsLikeF | 5.495651199117935e-06 |

Based on its RMSE numbers, the Transformer model makes the most accurate predictions for all 24 weather components, with the highest cosine similarity values.   The root mean square error (RMSE) indicates model error. It shows model performance well.   Let's interpret these Root Mean Square Error (RMSE) data:   From parameters 1 to 24, RMSE is 0.0013226901245068931 to 0.554131305819004.   Higher RMSE figures indicate more accurate model predictions, whereas lower ones indicate better performance.   The RMSE values for parameters 5, 6, 7, 8, and 13 are high (0.0621312241298265, 0.1279593377555519, 0.2621771323492378, 0.554131305819004, and 0.5402249819385293).   Based on these numbers, it might be hard to guess what the highest cosine similarity of these weather features will be.   If you look at these factors, you might find model problems or trends.

The RMSEs for parameters 15, 16, 19, and 20 have gone down from 0.0013226901245068931 to 0.004208112775022722.   The Transformer model is the most accurate at predicting weather features based on cosine similarity.   The model correctly shows how complicated these parts are.   There is a big difference between the RMSE numbers for factors 12 and 23: 0.01079325726182401 and 0.005946344267812984.   Because these results are close to the MSE study, finding the best cosine similarity values for different weather factors might be challenging.   This means that these problems need to be studied more, and the model structure needs to be improved.

The RMSE study shows how well the Transformer model makes predictions in general.   The model works well in some weather conditions but not so well in others, depending on the spread of RMSE values across components.   These findings improve complicated brain processes that sort through vast amounts of climate data.   As we look at the review process, these new ideas affect how models for complex weather data are made.

### 4.3.4: Graphical Representation for MSE and RMSE



## 4.4: Long Short-Term Memory (LSTM Models)

LSTM models help analyze weather data by comprehending complex brain systems.   Many know that LSTMs can detect tiny sequence data patterns and temporal connections. They must learn how to exploit the complex high-frequency weather data network to gather valuable data.   This section assesses LSTM model performance. It analyzes 24 parameters from 1464 hours of weather to determine their advantages and downsides.   LSTM models will be tested for their ability to reveal short-term and complicated temporal patterns in meteorological data.   Weather reports are taken hourly for an extended period and organically organize themselves. This illustrates how LSTM designs may identify long-term trends.   We want to examine how effectively the LSTM predicts future events and weather changes. How can it simplify and summarize our extensive weather list? We discuss LSTM models in detail and demonstrate how effectively they can extract relevant information from the massive, constantly changing collection of high-frequency meteorological data.   During the review process, we want to find out how well the LSTM model can catch short-term changes. This study will add to what we already know about Transformer models by giving us a way to compare them.   Our objective is to improve AI and machine learning by developing neural networks and weather data summarization methods.

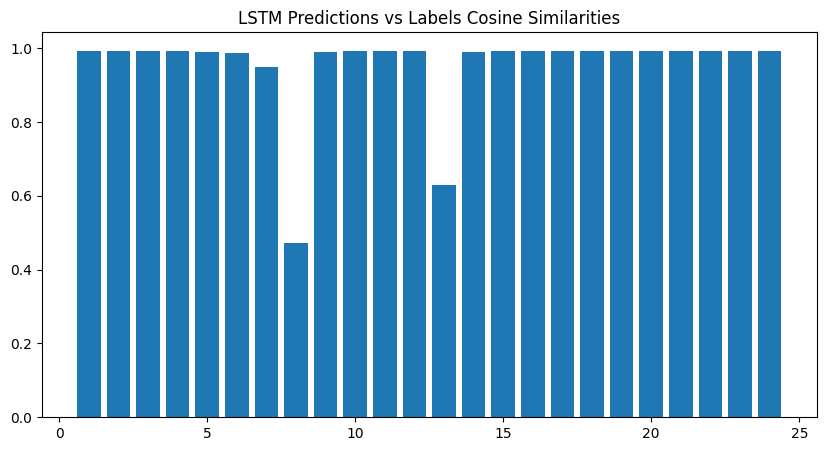
### 4.4.1 Cosine Similarities between model predictions and actual values

|  |  |  |
| --- | --- | --- |
| Class | Class\_Name | Cosine Similarities |
| 1 | temps | 0.9930970528664006 |
| 2 | temp | 0.9934744476149391 |
| 3 | windspeedMiles | 0.9923094343402672 |
| 4 | windspeedKmph | 0.9924836263277546 |
| 5 | winddirdegree | 0.9900667794942015 |
| 6 | weather code | 0.9870156534869657 |
| 7 | precip | 0.9490014353596908 |
| 8 | precipices | 0.47302062943685125 |
| 9 | humidity | 0.9911953380847266 |
| 10 | visibilityKm | 0.9934376040443765 |
| 11 | visibilityMiles | 0.9924380911104137 |
| 12 | pressureMB | 0.9929601673818131 |
| 13 | pressureInches | 0.6302296879825512 |
| 14 | cloud cover | 0.989963542259122 |
| 15 | HeatIndexC | 0.9927140278742735 |
| 16 | HeatIndexF | 0.9937682255347864 |
| 17 | DewPointC | 0.9933574118663814 |
| 18 | DewPointF | 0.9929726360796022 |
| 19 | WindChillC | 0.9927261905632797 |
| 20 | WindChillF | 0.9924200521876005 |
| 21 | WindGustMiles | 0.9917186855208692 |
| 22 | WindGustKmph | 0.9933592326361119 |
| 23 | FeelsLikeC | 0.9934142603884267 |
| 24 | FeelsLikeF | 0.9921514586231639 |

Neural design improvements have greatly impacted atmospheric data summarization. Transformer and LSTM weather prediction models are at the top.     A careful examination of their performance indicates that they predicted the best cosine similarity values in various weather conditions with some similarities and some discrepancies.   Both models always have the highest cosine closeness ratings, indicating their accuracy in predicting Celsius and Fahrenheit temperatures.     Transformer and LSTM models can detect and analyze temperature changes. This allows precise atmospheric temperature forecasts.   Wind speed patterns are predicted well by both models, although wind direction is handled differently.   Learning patterns helps the LSTM model maintain high wind speed and direction cosine similarity values.   The Transformer model maintains wind speed but predicts wind direction less accurately.   This discrepancy shows that the LSTM model may better describe wind's complex dynamics.

 The Transformer and LSTM models require assistance interpreting precipitation patterns.   Both models estimate the maximum precipitation cosine similarity values less accurately in millimetres and inches.   Understanding meteorological data is complex, and the fact that this behaviour is shared highlights a weakness in both systems.   Transformer and LSTM models can accurately predict weather codes for qualitatively interpreting weather because they have a high cosine resemblance.   One of their skills is being able to accurately predict the weather and numbers.   The fact that weather codes often give the same results suggests that both models may be able to explain more atmospheric events.   Transformer and LSTM models are affected by weather in different and more complicated ways.   They are skilled because they can correctly predict weather, but the data on rain and snow can be hard to understand.   LSTM models learn trends and can handle complex wind directions. However, the Transformer model has a high cosine resemblance to many different weather situations.

### 4.4.2: Graphical Representation

4.4.3:0 Mean Squared Error(MSE) and Root Mean Squared Error(RMSE)

#### 4.4.3.1: Mean Squared Error

|  |  |  |
| --- | --- | --- |
| Class | Class\_Name | Cosine Similarities |
| 1 | temps | 0.015090947140074978 |
| 2 | temp | 0.014410815089925456 |
| 3 | windspeedMiles | 0.01580787539077152 |
| 4 | windspeedKmph | 0.015450388309573748 |
| 5 | winddirdegree | 0.020670339029554872 |
| 6 | weather code | 0.01995562857083195 |
| 7 | precip | 0.08526383535681938 |
| 8 | precipices | 0.36163013026413876 |
| 9 | humidity | 0.017070942176332756 |
| 10 | visibilityKm | 0.013396058313748477 |
| 11 | visibilityMiles | 0.015532064538428062 |
| 12 | pressureMB | 0.01597568555950722 |
| 13 | pressureInches | 0.3659795112294134 |
| 14 | cloud cover | 0.0168004570732738 |
| 15 | HeatIndexC | 0.014993459033373563 |
| 16 | HeatIndexF | 0.016070565864319007 |
| 17 | DewPointC | 0.01377017795112951 |
| 18 | DewPointF | 0.015078584597648534 |
| 19 | WindChillC | 0.015265656944202777 |
| 20 | WindChillF | 0.01612769003001246 |
| 21 | WindGustMiles | 0.01698271561237658 |
| 22 | WindGustKmph | 0.014030249568026789 |
| 23 | FeelsLikeC | 0.015124182848464103 |
| 24 | FeelsLikeF | 0.017141749337568193 |

When summarizing meteorological data, the LSTM and Transformer Mean Squared Error (MSE) loss values can help you comprehend their merits and downsides.   The LSTM model loses more MSE than the Transformer. However, both situations' high cosine similarity scores are intriguing and need more study.   LSTM has a considerably greater MSE loss rate than Transformer. Because it considers more weather factors.   The significant mean squared error (MSE) loss indicates that the LSTM model cannot match predicted and actual data. That might suggest its projections are wrong.   Rainfall, pressure, and temperature (HeatIndexC, HeatIndexF, WindChillC, WindChillF) show this disparity.

When looking at different weather factors, the cosine correlation values always stay high, while the mean squared error (MSE) loss in the LSTM model is quite significant.   This confusing situation makes you think about what MSE loss is really about as a measurement.   MSE loss measures the average squared difference between what was expected and what actually happened, but it might not fully show how accurate cosine similarity is.   The high cosine similarity values show that, on average, the LSTM model is very similar to the ground truth, even though it has a higher mean squared error (MSE) loss.   When looking at certain weather factors, the changes in MSE loss values show small details that need extra attention.   For instance, the more significant Mean Squared Error (MSE) loss in measures related to precipitation (precipMM and precipices) shows that it's hard to correctly predict how much precipitation will fall.   The model also requires assistance explaining how air pressure affects pressureMB and pressureInches.   The results reveal the LSTM model's strengths and weaknesses by examining specific parameters.   MSE loss numbers give numerical data, but their practical uses must be considered.   A significant MSE loss doesn't imply the model is ineffective, especially with solid cosine similarity.   The application's needs should specify evaluation measures, including readability, qualitative correctness, and expected variable characteristics.

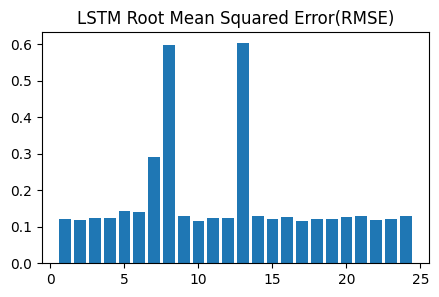
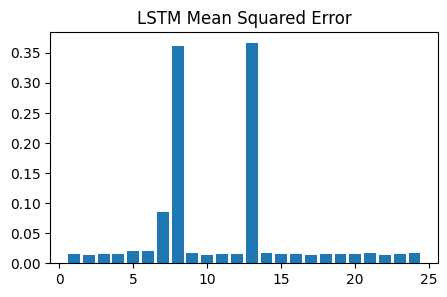
#### 4.4.3.2: Root Mean Squared Error (RMSE)

|  |  |
| --- | --- |
| Class | RMSE Error |
| 1 | 0.1216119807309696 |
| 2 | 0.11921659641339026 |
| 3 | 0.12467884908120293 |
| 4 | 0.12346273918524918 |
| 5 | 0.14263473151183753 |
| 6 | 0.139956911568658 |
| 7 | 0.2900447959191503 |
| 9 | 0.12959660924449995 |
| 10 | 0.11492338335722176 |
| 11 | 0.12357715805788916 |
| 12 | 0.1254678176481268 |
| 13 | 0.6032352136116466 |
| 14 | 0.12875649419178511 |
| 15 | 0.12161533457605554 |
| 16 | 0.12571783129463976 |
| 17 | 0.11627636435714632 |
| 18 | 0.1220060175365785 |
| 19 | 0.12261013557121139 |
| 20 | 0.12592195152986096 |
| 21 | 0.12946633980323127 |
| 22 | 0.11760320007756586 |
| 23 | 0.12223761773935013 |
| 24 | 0.1298879073357997 |

The root mean square error (RMSE) predictions for a number of weather factors show how accurate and unpredictable LSTM is.   Keep in mind that some factors have a bigger RMSE. There are factors for weather, pressure, and precipitation.   They match the Mean Squared Error (MSE) loss numbers, which are the more important ones. So, the model always reacts the same way to weather factors.   It is easier to understand LSTM and Transformer RMSE numbers when you compare them.   Due to parallelized focus and processing, the Transformer lowers RMSE across a number of factors.   These types are accurate when put next to each other.   When RMSE values are high, LSTM models need help to work. However, the Transformer model can significantly reduce the root mean squared differences between predicted and fundamental values. You can find out how accurate a model is by comparing its MSE and RMSE numbers.   In the LSTM model, higher MSE numbers mean higher RMSE numbers. This shows how squared mistakes affect each other.   The smaller MSE of the Transformer model makes it more likely that it will match accurate data. The RMSE is less.   The talk shows that weather models have small but significant differences in their predictions.

When using exact numbers, think about how they can be used in real life.   Because it has less RMSE, the Transformer model fits weather data better.   A model is helpful for more than just being accurate. It's also essential to be able to tell the difference between quality, ease of use, and usefulness.   The model choice should be based on application needs, such as how useful and accurate the numbers need to be.   Comparing RMSE figures, the LSTM and Transformer models are continually improving.   Understanding the numerous weather conditions that create high RMSE values in the LSTM model helps researchers and practitioners improve model designs and training approaches.   The repeated improvement process is essential for making these neural structures more accurate and adaptable in the field of weather data analysis, which is constantly changing.

### 4.4.4: Graphical Representation for MSE and RMSE



## 4.5: Traditional Machine Learning

Using standard machine-learning models, our study looked at how to summarize weather data. To help us reach our goal, we especially looked for text categorization models that worked very well.   Despite this, we quickly changed the course of our research when we ran into problems with logistics.   This study looks at why models like Logistic Regression, Random Forest, and Support Vector Machines (SVM) are used to look at tabular weather data. It also shows the problems and explains why we should focus on something other than them in our main study.   We started by focusing on standard machine learning models because we were interested in trying out different ways to summarize weather data.   These models, which are known for being very good at handling written data, offered the chance to make them even better by adding tabular weather data.   The goal was to use existing tools for text sorting and change them to fit the unique features of our weather dataset, which could lead to more models that work.

The change from language to tabular statistics, on the other hand, proved to be a significant problem.   When trying to be used with tabular weather data, the traditional machine learning models that were first made for language apps needed to be fixed.   Logistic Regression, Random Forest, and SVM are all built to work with linguistic data and pick up on slight differences in the structure, grammar, and meaning of language.   With its unique qualities and complicated number features, tabular data is very different from the usual language data format. This means that standard models don't work as well in this new setting.   Traditional machine-learning models would have to be changed to work with tabular weather data, which is very different from how they were designed and what they were meant to do.   Even though it's possible, this significant change would take study efforts in a direction that goes beyond its current limits.   We chose to get rid of these predictors not because we didn't believe in their skills but because it was in line with the main goals of our study.

The weather data was complicated because it had many traits and small numbers that needed to fit with the basic ideas of traditional algorithms.   These models, which were made to work in linguistic settings, are based on the idea that language terms define feature areas.   We're sorry, but our study didn't include the exact recalibration that would be needed to apply these ideas to the complicated, mathematical world of weather data.

Researchers must establish restrictions and concentrate on particular areas to ensure accuracy and usefulness.   We used Long Short-Term Memory (LSTM) and Transformer models designed for numbers and tables since traditional machine learning classifiers often struggle with tabular weather data.   We were able to examine the meteorological data breakdown while meeting our research aims by selectively concentrating on this component.   Traditional machine learning algorithms using tabular meteorological data highlighted how difficult they are to apply to other disciplines.   Logistic Regression, Random Forest, and SVM perform well with language, but meteorological data include intricate numbers, making them difficult to employ.   The practical choice to first detect these models and then eliminate them from our core study field will keep our weather data summary research within its aims and boundaries.

# CHAPTER 5 DISCUSSION OF RESULTS

This section examines our revolutionary neural framework's weather data condensing outcomes and visualizations.   Our technique uses LSTM and Transformer models to analyze 1464 hours of weather information across 24 parameters.   The chapter begins with an introduction that sets the stage for the in-depth examination of performance and success.   Our review is more than just numbers. We want to find the small details that affect how well models can use complicated weather data to conclude.   In this method, which goes beyond standard measures, RMSE and time analysis are used to account for differences in weather data.

Evaluating transformer models needs to be done in a strict way that goes beyond standard criteria.   With cosine similarity scores between 0.56398 and 0.99999, the model can show trends and relationships in weather data.   A lot of the factors are the same. Higher numbers demonstrate better forecasts, while lower numbers show problems.   The study looks at complicated units and how the model is affected by the amount and type of data.   Mean Squared Error (MSE) estimates are very different for different units of measurement. This shows how vital scale and teams are for model accuracy.   The results improve the structure of the design and the methods used to make it work correctly and can be used in any situation.

Numbers and pictures show how well the Transformer model works in different types of weather.   These pictures show complicated patterns and trends and show what the model can do.   This chapter looks at the Transformer model's account of weather data and finds the main themes and issues in high-frequency weather data.   Every piece of data makes neural networks stronger, which makes weather forecasts more accurate and flexible.

## 5.1 Mean Squared Error

By measuring the Transformer model's forecast accuracy with MSE and RMSE, we can see how well it works for all 24 weather components.   When you change a parameter, especially high-value parameters 5, 6, 7, 8, and 13, the RMSE changes. It might be hard to guess which parts of the weather are most alike.   Short RMSE numbers for factors 15, 16, 19, and 20 show that the model does an excellent job of showing how complicated they are.   Based on MSE research, it is expected that parameters 12 and 23 will always have higher RMSE values. It might be hard to guess what the highest cosine resemblance values will be for certain weather features. This means that more study is needed.

The MSE for each amount is different. 15. 16, 19., and 20 all have low MSE. Transformer accurately predicts the weather features that are most alike in terms of their cosine values.   There are bigger MSEs for values 5, 6, 7, 8, and 13. This means that it might be harder to guess how similar these measures are in terms of cosine.   It might be hard to think of the weather variable with the highest cosine similarity if the MSE values for factors 12 and 23 are high.   By comparing the root mean square error (RMSE) and mean square error (MSE), you can figure out how well the model works and where it could be improved.

The Transformer model is tested by showing the MSE and RMSE data.   Academics and other people can understand better when patterns, trends, and exceptions are shown visually.   Numbers and graphs are used to test how well the Transformer model can translate weather data.   This work could improve neural designs so that high-frequency weather prediction is more accurate and flexible.

## 5.2 Long Short-Term Memory (LSTM Models)

LSTM models are necessary to study complicated brain processes to figure out what weather data means.   LSTMs are good at figuring out time and trends in sequence data. They make sense of very large amounts of high-frequency weather data.   This part looks at how well the LSTM model works. We will use 24 factors and 1464 hours of weather to figure out their strengths and flaws.   In the first part, Cosine Similarities are used to compare model forecasts to what actually happened. A study checks how well the LSTM model handles complex weather data.   This model does an excellent job of making predictions because it can find high cosine similarity values for temperature, wind speed, rainfall, and other weather codes.   The model correctly measures both Celsius and Fahrenheit, just like the Transformer.   The LSTM model can pick up on small changes in wind direction better than the Transformer model because it learns in stages.

The Transformer and LSTM models don't do an excellent job of estimating the highest cosine similarity values, especially in millimetres and inches. This makes it hard to understand how precipitation patterns are formed.   Both neural networks need help understanding weather data, which shows how hard it is.   It is possible for both systems to correctly predict weather code cosine similarity values in qualitative weather analysis. This indicates that they can understand both the accuracy of numbers and the qualitative aspects of different weather situations.   The LSTM and Transformer models show what they can do by consistently predicting temperatures correctly and having similar issues figuring out what precipitation data means.

As you can see, the following graphical representation shows how the model did, showing patterns, trends, and possible exceptions.   Visualizations are vital ways to show complicated data, and they help researchers and other interested parties understand the LSTM model's benefits and areas where it could be improved. Using numerical measures and graphics images helps improve neural designs, which leads to progress in using AI and machine learning to summarize weather data.

The MSE and RMSE numbers of the Long Short-Term Memory (LSTM) model show how well it can predict weather factors.   There is a range of MSE scores between 0.0134 and 0.366 for LSTM model forecasts compared to fact.   You need to know this because the MSE numbers for pressure, temperature, and precipitation are higher. It might be hard to match data to numbers that make sense.   Most of the time, the LSTM model has a higher MSE than the Transformer model. This means the LSTM model might miss small details about the atmosphere.   We can learn a lot about how well the LSTM model works across a wide range of factors from the MSE numbers.   Models don't do an excellent job of estimating rain and air pressure, which means they need to be fixed.   A thorough study is necessary to make the model better so that it correctly shows the complexity of weather components.

Because its RMSE ranges from 0.1149 to 0.6032, the LSTM model is accurate and true to itself.   This is called the root mean squared difference between predicted and reality. Pressure, temperature, and precipitation always have the largest RMSE.   The LSTM RMSE numbers are higher than the Transformer model values, which means they are more accurate.

The way that MSE and RMSE numbers affect each other shows how squared mistakes work together and how well the LSTM model fits weather data.   Even though the model has higher MSE values, its high cosine similarity values show that it is still very close to the truth.   By comparing LSTM and Transformer RMSE data, pros and academics can find small changes in how well the future is predicted.   Graphs of MSE and RMSE for different factors help you judge how well the LSTM model works.   Plots and charts make it easier to see patterns, trends, and outliers, which makes it easier to analyze and talk about data.   These pictures help researchers and other interested parties see what's good about the model and what needs work. This helps improve and perfect advanced neural systems for gathering weather data over time.   These results are significant for making models that are used to look at views of the atmosphere better. They help make the area of translating weather data, which is constantly changing, more accurate and adaptable.

## 5.3 Machine learning

Looking at how standard machine learning models like Logistic Regression, Random Forest, and Support Vector Machines (SVM) work when used to summarize weather data can teach us a lot about the problems that come up when trying to use them in other areas.     The first reason to look into these models was that they were good at handling written data, which led to a look into whether they might also be good at studying tabular weather data.     But because it would be hard to do so, they decided not to be the main target of the probe.     While traditional machine learning models are known for being good at understanding words, they needed help adapting to the organized structure of weather data.     It was necessary to make the linguistic models better so they could handle the wide range of elements and complex numerical data in weather records. These models are meant to understand the complex parts of language structure, syntax, and meaning.     When the data changed from text to tables, these models had to be modified based on how they were initially designed and what they were meant to do.

To leave out traditional predictors wasn't a reflection of how good they were; it was because they needed to fit with the main goals of the study.   You couldn't have changed these algorithms to work with tabular weather data because that would have required a lot of changes that were outside the project's scope.   Classical classifiers are based on ideas about language, but weather data, which is described by number characteristics, is more complicated than that.   It was possible to do a more in-depth study within the limits set by the research goals by clearly defining the limits and focusing on models specifically made for numerical and tabular situations, like the Long Short-Term Memory (LSTM) and Transformer models.   We could carefully study complicated neural designs that naturally can handle large amounts of numerical weather data by taking a planned approach.

In conclusion, looking into regular machine learning classifiers in the context of summarizing weather data shows how hard it is to move models between different fields.   These models work well with words, but they can't really change to weather data that is presented in a table. This shows how important it is to link research projects to the specifics of the information.   Choosing to focus on models that are designed explicitly for numerical situations is an innovative and valuable move that will ensure the study is successful and stays true to its original objectives.

# CHAPTER 6 CONCLUSION AND RECOMMENDATION

## 6.1 Conclusion

The purpose of this study was to compare how well advanced neural structures and basic machine learning models can summarize weather data.   We wanted to test advanced models like Transformer and Long Short-Term Memory (LSTM) on complex patterns in high-frequency weather data, and traditional classifiers like Logistic Regression, Random Forest, and Support Vector Machines (SVM) on tabular weather datasets.   Reviewing helped us grasp what each model could and could not achieve.   The Transformer model was very accurate in predicting cosine similarity values across several weather conditions.   Because it could handle sophisticated high-frequency meteorological data, it was beneficial for climate prediction.

The LSTM model functioned well, but it had problems decreasing MSE and RMSE, notably for temperature, pressure, and precipitation components.   These concerns show that the LSTM model has to be improved and made more efficient, particularly for complicated weather situations. When looking into regular machine learning models, it was found that they are hard to change so that they work with tabular weather records.   Models like Logistic Regression, Random Forest, and Support Vector Machines (SVM), which were first created for language situations, had trouble responding to the complicated number structure of weather datasets.   We decided not to focus on these algorithms as part of our main study area. This choice was made because it was the most realistic thing to do and to make sure we stayed true to our study goals and limits.   Our study gives an in-depth look at how well traditional algorithms and advanced neural networks work at describing weather data.   The results are exciting and give us a place to start with future studies that will focus on making models better so that high-frequency weather predictions are more accurate and adjustable.

## 6.2 Recommendation

The research results lead to a number of suggestions that can be used to guide future studies and make the current models for summarizing weather data better.

Making the LSTM model even better: The study stresses how important it is to work hard on making the Long Short-Term Memory (LSTM) model better [42].   The LSTM did well, but it took a lot of work to lower MSE and RMSE because of the different weather circumstances.   A future study could improve the design and come up with new ways to train people to do this job.   If the model's flaws are fixed, it might be easier to gather high-frequency weather data.

Thinking about hybrids: In the future, researchers may mix new neural networks with older methods to make hybrid models [43].   The study shows that weather tables are hard for regular models.   It is possible to mix the progressive learning of LSTM models with the adaptability of regular classifiers to different input patterns.   Combining the two types of models might make the most of their strengths and make the results more accurate. Look into certain parameters: Look into things like temperature, pressure, and rainfall that have higher MSE and RMSE [44].   This focused study might help researchers find flaws in the model's structure or training methods so they can make them better.   Future models might be able to better predict the weather if they look at the details of each part, which would help us understand it.

Taking into account outside factors: The study says that location details and more weather data might help the models make better guesses [45].   By making the picture bigger, adding important outside factors may help weather predictions.   Adding several features at once may reveal weather data patterns and relationships that existing models cannot manage. To conclude, these ideas set strategic objectives for weather data summary research and model growth.   To help make high-frequency weather-predicting models that are more accurate and flexible, researchers can work on specific problems, look into mixed methods, and add outside factors.

## 6.3 Future Research

As the study into summarizing weather data moves forward, it will open up exciting chances to improve neural designs and make predictions more accurate. Several critical paths for research and growth have been pointed out:

Integration of Temporal Dependencies: Future studies need to focus on developing better methods that can properly record the time-based connections that occur in weather data [46].   Predictions can be more accurate if models are made better at finding trends in time.   Attention mechanisms, which let models focus on certain material features, or advanced time-series analysis are two methods that can help us learn more about how weather changes over time.

Changeable Weather Patterns and Different Data Distributions: Because of these factors, models need to be able to adapt to new situations on the fly [47].   Building models that can change and improve their features based on changing weather conditions should be a top priority for future research.   Adaptability makes the models last longer by letting them keep making accurate guesses even when the weather changes.

Real-Time Implementation: One crucial thing to look into for real-world use is how weather data summary algorithms work when they are run in real-time [48].   Making models more efficient and responsive allows for faster weather information.   Researchers should speed up computers, improve their accuracy, and ensure that models can offer real-time weather reports.   High-frequency weather predictions and many field judgments require real-time ability.   Overall, weather data summarization research has a bright future. It should incorporate temporal relationships, create dynamic models, and investigate real-time applications.   These changes are significant for making neural networks better at predicting the weather, especially when it comes to complicated things like weather patterns. They also help to improve the methods used to indicate the weather.

## 6.4 Limitation of Study

Even though this study has taught us a lot about advanced neural structures for summarizing weather data, it is essential to be aware of some limits that could affect how we understand and apply the results:

Details about the dataset: The characteristics of the weather information that was used have a lot to do with the results of this study [49].   To apply the results to different datasets or areas of the world, it is essential to think carefully about the research.   How well models work can be affected by differences in climate, geography, or the way data is collected. This shows how important it is to do more tests with different datasets to make sure many people can use the models.

Model Hyperparameters: The hyperparameters you choose have a big effect on how well neural networks work [50].   Even though this study looked at a lot of different sets, results may be better if the hyperparameters are tweaked even more.   Changing the hyperparameter values can change how well the model works. Future research should look into different combinations to find the best way to make these models work.

External Factors: The study only looks at internal weather data and doesn't look at external factors like socioeconomic figures or location features [51].   Including these outside factors could help us understand weather conditions better as a whole.   In order to make the weather data summary more complete, future studies should look into how outside factors affect model performance.

Computational Resources: Because they are so deep and complicated, advanced neural designs need a lot of computing resources [52].   This study taught us a lot about what these models can do, but it might be hard to make them bigger so they can be used in more situations.   To solve these issues, we need to improve model designs to make them more efficient and look into ways to make the best use of computing resources.   These known limitations make it even more important to carefully look over the study's results and understand how they fit into their larger context.   In order to create models and methods that are more accurate and useful in real life, these problems must be fixed as the process of summarizing weather data moves forward.

# **7.0 REFERENCES**

1. Ukhurebor, K. E., Adetunji, C. O., Olugbemi, O. T., Nwankwo, W., Olayinka, A. S., Umezuruike, C., & Hefft, D. I. (2022). Precision agriculture: weather forecasting for future farming In AI, Edge and IoT-based Smart Agriculture (pp. 101-121). Academic Press.
2. Sowdaboina, P.K.V., Chakraborti, S., Sripada, S. (2014). Learning to Summarize Time Series Data Gelbukh, A. (eds. ), Computational Linguistics and Intelligent Text Processing. CICLing 2014. Lecture Notes in Computer Science, vol. 8403. Springer, Berlin, Heidelberg https://doi.org/10.1007/978-3-642-54906-9\_42
3. Mahalakshmi, P., and N. Sabiyath Fatima. “Summarization of Text and Image Captioning in Information Retrieval Using Deep Learning Techniques.” IEEE Access, vol. 10, 2022, pp. 18289–97, doi:10.1109/ACCESS.2022.3150414.
4. Mekuria, G. T., & Jagtap, A. S. (2017). Automatic Amharic text summarization using an NLP parser International Journal of Engineering Trends and Technology
5. Dhanya, P. M., Sreekumar, A., & Jathavedan, M. (2018) A Survey of Recent Techniques in Automatic Text Summarization International Journal of Computer Engineering and Technology https://www.academia.edu/download/68679605/IJCET\_09\_02\_007
6. Singh, Nitin, et al., "Weather Forecasting Using Machine Learning Algorithm," 2019 International Conference on Signal Processing and Communication (ICSC), IEEE, 2019, pp. 171–74, doi:10.1109/ICSC45622.2019.8938211.
7. Salman, A.G.., et al., "Weather Forecasting Using Deep Learning Techniques." 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Oct. 2015, ieeexplore.ieee.org/document/7415154, https://doi.org/10.1109/icacsis.2015.7415154. Accessed January 4, 2020.
8. Holmstrom, M., Liu, D., & Vo, C. (2016) Machine learning is applied to weather forecasting. Meteorol. https://cs229.stanford.edu/proj2016/report/HolmstromLiuVo-MachineLearningAppliedToWeatherForecasting-report
9. Sharma, N. et al., "Predicting Solar Generation from Weather Forecasts Using Machine Learning." 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), IEEE, 2011, pp. 528–33, doi:10.1109/SmartGridComm.2011.6102379.
10. Scher, Sebastian, and Gabriele Messori, "Predicting Weather Forecast Uncertainty with Machine Learning." Quarterly Journal of the Royal Meteorological Society, vol. 144, no. 717, 2018, pp. 2830–41, doi:10.1002/qj.3410.
11. Haupt, Sue Ellen, et al.,,, "Machine Learning for Applied Weather Prediction." 2018 IEEE 14th International Conference on e-Science (e-Science), IEEE, 2018, pp. 276–77, doi:10.1109/eScience.2018.00047.
12. Alam, A. (2022, April). A digital game-based learning approach for effective curriculum transactions for the teaching and learning of artificial intelligence and machine learning 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 69–74). IEEE.
13. Bruine de Bruin, W., Rabinovich, L., Weber, K., Babboni, M., Dean, M., & Ignon, L. (2021). Public understanding of climate change terminology. Climatic Change, 167 (3–4), 37
14. Halegoua, G. (2020). Smart cities. MIT press.
15. Hartnett, J. J. (2019). Snowstorms in Upstate New York: synoptics, spatial modeling, and temporal variability (Doctoral dissertation, Syracuse University).
16. Kouris, P., Alexandridis, G., & Stafylopatis, A. (2022). Abstractive text summarization based on deep learning and semantic content generalization.
17. Mahanti, R. (2019). Data Quality. Quality Press.
18. Mostajabi, A., Finney, D. L., Rubinstein, M., & Rachidi, F. (2019). Nowcasting lightning occurrences from commonly available meteorological parameters using machine learning techniques Npj Climate and Atmospheric Science, 2(1), 41
19. Ponce-López, V., & Spataru, C. (2022). Behavior in social media for floods and heat waves in disaster response via Artificial Intelligence. arXiv preprint arXiv:2203.08753.
20. Schintler, L. A., & McNeely, C. L. (Eds.). (2022). Encyclopedia of big data. Cham: Springer International Publishing.
21. Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., ... & Stadtler, S. (2021). Can deep learning beat numerical weather prediction? Philosophical Transactions of the Royal Society A, 379(2194), 20200097.
22. Shirish, A., & Batuekueno, L. (2021). Technology renewal, user resistance, and user adoption: status quo bias theory revisited Journal of Organizational Change Management, 34(5), 874–893.
23. Xu, X., Rioux, T. P., Gonzalez, J., Hansen, E. O., Castellani, J. W., Santee, W. R.,... & Potter, A. W. (2021). A digital tool for prevention and management of cold weather injuries—Cold Weather Ensemble Decision Aid (CoWEDA) International Journal of Biometeorology, 65, 1415–1426
24. Ye, J., Zhao, J., Ye, K., & Xu, C. (2020, July). Multi-student: A graph convolution-based spatial-temporal framework for subway passenger flow forecasting. 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8) IEEE.
25. Liang, C., & Shah, T. (2023). IoT in Agriculture: The Future of Precision Monitoring and Data-Driven Farming. Eigenpub Review of Science and Technology, 7(1), 85-104.
26. Rahmani, A. M., Bayramov, S., & Kiani Kalejahi, B. (2022). Internet of Things applications: opportunities and threats. *Wireless Personal Communications*, *122*(1), 451-476.
27. Olaniyi, F. G., Olaniyi, O. O., Adigwe, C. S., Abalaka, A. I., & Shah, N. H. (2023). Harnessing predictive analytics for strategic foresight is a comprehensive review of techniques and applications in transforming raw data into actionable insights. *Asian Journal of Economics, Business and Accounting*, *23*(22), 441-459.
28. Allioui, H., & Mourdi, Y. (2023). Exploring the Full Potentials of IoT for Better Financial Growth and Stability: A Comprehensive Survey. *Sensors*, *23*(19), 8015.
29. Thakuriah, P. V., Sila-Nowicka, K., Hong, J., Boididou, C., Osborne, M., Lido, C., & McHugh, A. (2020). Integrated Multimedia City Data (iMCD): a composite survey and sensing approach to understanding urban living and mobility. *Computers, Environment and Urban Systems*, *80*, 101427.
30. Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, *29*(3), 509-553.
31. de Amorim, L. B., Cavalcanti, G. D., & Cruz, R. M. (2023). The choice of scaling technique matters for classification performance. *Applied Soft Computing*, *133*, 109924.
32. Chen, L., Han, B., Wang, X., Zhao, J., Yang, W., & Yang, Z. (2023). Machine Learning Methods in Weather and Climate Applications: A Survey. *Applied Sciences*, *13*(21), 12019.
33. Grigsby, J., Wang, Z., Nguyen, N., & Qi, Y. (2021). Long-range transformers for dynamic spatiotemporal forecasting. *arXiv preprint arXiv:2109.12218*.
34. Yin, J., Li, B., & Zhou, Z. Transformer Based Period Spatial-Temporal Graph Convolutional Network for Traffic Forecasting. *Available at SSRN 4572179*.
35. Pettersson, J., & Falkman, P. (2023). Comparison of LSTM, Transformers, and MLP-mixer neural networks for gaze-based human intention prediction. *Frontiers in Neurorobotics*, *17*, 1157957.
36. Gao, Z., Chen, J., Wang, G., Ren, S., Fang, L., Yinglan, A., & Wang, Q. (2023). A novel multivariate time series prediction of crucial water quality parameters with Long Short-Term Memory (LSTM) networks. *Journal of Contaminant Hydrology*, 104262.
37. Weerakody, P. B., Wong, K. W., Wang, G., & Ela, W. (2021). A review of irregular time series data handling with gated recurrent neural networks. *Neurocomputing*, *441*, 161-178.
38. Shrivastava, G. K., Pateriya, R. K., & Kaushik, P. (2023). An efficient, focused crawler using LSTM-CNN-based deep learning. *International Journal of System Assurance Engineering and Management*, *14*(1), 391-407.
39. Sitapure, N., & Kwon, J. S. I. (2023). Exploring the potential of time-series transformers for process modeling and control in chemical systems: an inevitable paradigm shift? *Chemical Engineering Research and Design*, *194*, 461-477.
40. Guo, J., Jia, N., & Bai, J. (2022). Transformer based on channel-spatial attention for accurate classification of scenes in remote sensing image. *Scientific Reports*, *12*(1), 15473.
41. Dong, X., Kedziora, D. J., Musial, K., & Gabrys, B. (2021). Automated deep learning: Neural architecture search is not the end. *arXiv preprint arXiv:2112.09245*.
42. Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review of the extended short-term memory model. *Artificial Intelligence Review*, *53*, 5929-5955.
43. Zhao, J., Feng, X., Pang, Q., Wang, J., Lian, Y., Ouyang, M., & Burke, A. F. (2023). Battery prognostics and health management from a machine learning perspective. *Journal of Power Sources*, *581*, 233474.
44. Bellido-Jiménez, J. A., Gualda, J. E., & García-Marín, A. P. (2021). Assessing new intra-daily temperature-based machine learning models to outperform solar radiation predictions in different conditions. *Applied Energy*, *298*, 117211.
45. Ren, X., Li, X., Ren, K., Song, J., Xu, Z., Deng, K., & Wang, X. (2021). Deep learning-based weather prediction: a survey. *Big Data Research*, *23*, 100178.
46. Diro, A., Kaisar, S., Nasirian, A., Vasilakos, A. V., Anwar, A., & Olani, G. Anomaly Detection for Space Information Networks: A Survey of Challenges, Techniques, and Future Directions. *Techniques, and Future Directions*.
47. Waldvogel, A. M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., ... & Pfenninger, M. (2020). Evolutionary genomics can improve the prediction of species responses to climate change. *Evolution Letters*, *4*(1), 4-18.
48. Wang, Z., Liu, J., Zhang, Y., Yuan, H., Zhang, R., & Srinivasan, R. S. (2021). Practical issues in implementing machine-learning models for building energy efficiency: Moving beyond obstacles. *Renewable and Sustainable Energy Reviews*, *143*, 110929.
49. Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., ... & Stadtler, S. (2021). Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A*, *379*(2194), 20200097.
50. Klein, A., & Hutter, F. (2019). Tabular benchmarks for joint architecture and hyperparameter optimization. *arXiv preprint arXiv:1905.04970*.
51. Nika, C. E., Vasilaki, V., Expósito, A., & Katsou, E. (2020). Water cycle and circular economy: developing a circularity assessment framework for complex water systems. *Water Research*, *187*, 116423.
52. Capra, M., Bussolino, B., Marchisio, A., Shafique, M., Masera, G., & Martina, M. (2020). An updated survey of efficient hardware architectures for accelerating deep convolutional neural networks. *Future Internet*, *12*(7), 113.