|  |
| --- |
| **Team no –12** |

Submitted by-

|  |  |
| --- | --- |
| 1. Harshal Sanjiv Patil | (50604782) |
| 1. Mrudula Chandrakant Deshmukh | (50605669) |
| 1. Siddhi Sunil Nalawade | (50613176) |

**INTRODUCTION: -**

This project aims to compare two well-known methods for predicting short-term temperature changes: ARIMA, a traditional statistical approach, and LSTM, a type of machine learning model. Predicting temperatures is important for weather forecasting

And will offer valuable insights for industries like agriculture, air travel planning, and transportation etc.

This project uses a large collection of daily temperature records from many U.S. cities, covering several decades. This detailed data provides a strong basis for testing how well each method works at predicting short-term temperature changes and understanding their strengths and weaknesses.

By comparing these two approaches, we aim to evaluate their accuracy and reliability under different conditions, explore potential hybrid solutions, and analyse the practical applications of each method in real-world scenarios. Our approach includes processing the data, models’ implementations, comparing their performance, and analysing the results. The results of this comparison will give useful information for industries that need temperature forecasts, like farming, air travel, city planning, and transportation. This can help these industries make better decisions and run more smoothly.

1. **What are we trying to do?**

We are trying to create and test two different models, ARIMA and LSTM, to see how well they can predict short-term temperature changes. By comparing the results, we want to find out which method gives more accurate forecasts for short periods, making weather predictions more reliable.

1. **How is it done today? What are the limits of current practice?**

Currently, weather forecasting uses a combination of traditional statistical models and advanced machine learning techniques. Traditional models like ARIMA are commonly used because they can handle linear trends and seasonality in weather data. However, ARIMA struggles with complex, non-linear weather patterns, which limits its effectiveness for short-term and local predictions (Murat et al., 2018; Salman et al., 2018).

Advanced models such as LSTM (Long Short-Term Memory) are becoming more popular because they can capture complex patterns over time, making them more effective for chaotic and non-linear weather predictions (Siami-Namini et al., 2018; De Saa and Ranathunga, 2021). Despite this, LSTM models require significant computational resources and large datasets, which limits their widespread use, particularly in localized forecasting scenarios (Li & Qian, 2018; Salman et al., 2018).

To address the limitations of both methods, hybrid models combining ARIMA and LSTM have been developed. These models take advantage of ARIMA's ability to model linear trends and LSTM’s ability to capture non-linear dependencies, resulting in improved accuracy. Several studies have demonstrated the effectiveness of hybrid ARIMA-LSTM models for weather forecasting, showing their potential to outperform either model when used alone (Salman et al., 2018; Li & Qian, 2018; Abdallah et al., 2020).

1. **What’s in our approach? Why will it be successful?**

Our approach involves directly comparing ARIMA and LSTM models for short-term temperature forecasting using a comprehensive dataset of daily temperature records from various U.S. cities. This allows us to assess the strengths of both traditional statistical methods and modern deep learning techniques. ARIMA is well-suited for capturing linear trends and seasonal patterns in weather data, while LSTM excels in modeling non-linear, complex temporal dependencies (Murat et al., 2018; De Saa and Ranathunga, 2021).

We believe this approach will be successful because of the extensive historical data available, which will enable thorough training and testing of both models. Furthermore, by comparing both methods on the same dataset, we can determine which model performs best under different conditions (Siami-Namini et al., 2018). Additionally, we plan to investigate a hybrid ARIMA-LSTM model, inspired by the work of Salman et al. (2018), which showed that merging these models could significantly improve prediction accuracy compared to using either one alone.

The hybrid approach, which leverages ARIMA’s strength in capturing linear trends and LSTM’s capacity to handle non-linear patterns, has been shown to outperform standalone models in various studies (Salman et al., 2018; Abdallah et al., 2020). Combining these methods allows us to exploit their respective strengths, providing a more accurate and robust forecasting solution, particularly for short-term temperature predictions.

1. **What are the effects, who cares?**

This research on short-term temperature forecasting has a broad impact across several sectors. While meteorologists and climate scientists are direct beneficiaries, other industries rely heavily on accurate temperature forecasts.

In the energy sector, power companies use these forecasts to manage supply and demand, ensuring that they can handle energy usage spikes during extreme weather conditions (Siami-Namini et al., 2018; Salman et al., 2018). Accurate temperature forecasts help farmers make informed decisions about planting, harvesting, and irrigation, which can increase crop yields and reduce losses due to unpredictable weather changes (De Saa & Ranathunga, 2021; Murat et al., 2018).

For the transportation industry, particularly aviation, reliable weather predictions are essential for flight safety and scheduling, allowing airlines to avoid hazardous conditions and reduce delays (Salman et al., 2018). Local governments also depend on precise weather forecasts for emergency preparedness and disaster response, enabling more effective resource allocation in extreme weather events (Murat et al., 2018).

Moreover, sectors such as construction and retail benefit from accurate forecasts, allowing them to plan activities and manage inventories based on expected weather conditions, reducing operational disruptions and financial losses (Abdallah et al., 2020; Ding, 2020). Finally, improving temperature forecasting can lead to better public safety, resource management, and cost savings across a wide range of industries (Li & Qian, 2018; Nghiem et al., 2021).

.

1. **If we’re successful, what difference and impact will it make? How do we measure it (e.g., via user studies, experiments, ground truth data)?**

Success in this project would lead to more accurate short-term temperature forecasts, which would greatly help decision-making in agriculture, energy, and emergency management. By reducing errors in predictions, we can assist these industries in planning better and using their resources more effectively. For example, Siami-Namini et al. (2018) found that LSTM models can cut prediction errors by 84-87% compared to ARIMA.

To measure our success, we will use common metrics like Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) to see how well ARIMA, LSTM, and hybrid models perform compared to standard forecasts. These metrics are widely used in the field, as shown in Salman et al. (2018), where RMSE helped track improvements in accuracy for hybrid ARIMA-LSTM models. Better forecasts can have a big impact across various sectors. De Saa and Ranathunga (2021) highlight that accurate weather predictions are essential for agriculture, transportation, and other industries.

1. **What are the risks and payoffs?**

The main risks of this project stem from the computational demands of LSTM models and the potential for overfitting. (Salman et al. (2018)) highlight the complexity of running LSTM models, especially when dealing with large datasets, which can require significant computational resources. (Siami-Namini et al. (2018)) further note that LSTM models may overfit when trained on limited or low-quality datasets, leading to inaccurate predictions. Additionally, there is the risk that our models may not generalize well across different geographic regions or time periods, limiting the broader applicability of the findings (De Saa & Ranathunga, 2021).

ARIMA, while simpler to implement and computationally less demanding, struggles with non-linear and complex patterns in weather data (De Saa & Ranathunga, 2021; Murat et al., 2018). This could lead to limitations in ARIMA’s effectiveness for short-term temperature forecasting, especially in rapidly changing or chaotic weather conditions.

However, the potential payoffs are substantial. By improving the accuracy of short-term temperature forecasts, this project could benefit industries such as agriculture, where better forecasts help with planting and harvesting (De Saa & Ranathunga, 2021), and the energy sector, which can better manage supply and demand during extreme weather conditions (Siami-Namini et al., 2018; Salman et al., 2018). (Salman et al. (2018)) demonstrated that combining ARIMA and LSTM models can yield more accurate results than using either method alone, which suggests a strong payoff if the hybrid approach is successful.

In addition, clarifying the strengths and weaknesses of ARIMA and LSTM will contribute to the development of more robust forecasting models. (Siami-Namini et al. (2018)) found that LSTM significantly reduces forecasting errors compared to ARIMA, indicating that this project could lead to advances in weather prediction techniques and practical applications, especially if we successfully integrate the strengths of both models (Abdallah et al., 2020).

1. **How much will it cost?**

The main costs for this project will come from the computing power needed for processing data and training models, especially for complex ones like LSTM. Salman et al. (2018) point out that LSTM can require a lot of processing power. To keep costs low, we plan to use free weather datasets and open-source tools, as suggested by Siami-Namini et al. (2018) in their study comparing ARIMA and LSTM models. We will also utilize existing R libraries and frameworks to help implement the models, which will further reduce development expenses.

1. **How long will it take?**

We expect the project to take about 2 months, starting with learning the R programming language and its statistical packages, followed by getting familiar with ARIMA and LSTM models for time series forecasting. The next steps will involve collecting and preparing the data, then developing and testing the models. Afterward, we’ll analyze the results and refine the models as needed. The final phase will be documenting the findings and preparing the project report

1. **What are the midterm and final” exams” to check for success? How will progress be measured?**

To measure progress and success, we will conduct midterm and final evaluations. Midterm checks will involve ensuring that we have completed data collection and preprocessing, implemented initial versions of the ARIMA and LSTM models, and produced some preliminary results using a test dataset. For the final evaluation, we will focus on comparing the performance of the ARIMA, LSTM, and any hybrid models using key metrics like RMSE and MAE. We will also analyze each model's strengths and weaknesses in different forecasting scenarios, conclude which method or combination works best for short-term temperature prediction, and document the entire process, results, and findings with visualizations to demonstrate forecast accuracy

**References:**

* Comparative Studies and Hybrid Approaches siddhi

1. (Afan Galih Salman, 2018)
2. (Ranathunga, 2020)Comparison between ARIMA and Deep Learning Models for Temperature Forecasting [Eranga De Saa](https://arxiv.org/search/cs?searchtype=author&query=De+Saa,+E), [Lochandaka Ranathunga](https://arxiv.org/search/cs?searchtype=author&query=Ranathunga,+L)
3. [A Comparison of ARIMA and LSTM in Forecasting Time Series (Sima Siami-Namini et al.)](https://ieeexplore.ieee.org/abstract/document/8614252)

* LSTM Models for Temperature Forecasting

1. [Time Series Forecasting using LSTM and ARIMA](https://thesai.org/Downloads/Volume14No1/Paper_33-Time_Series_Forecasting_using_LSTM_and_ARIMA.pdf)
2. [Enhancing Stock Price Prediction Method Based on CNN-LSTM hybrid model](https://drpress.org/ojs/index.php/HBEM/article/view/14760)
3. [Applying Bayesian inference in a hybrid CNN-LSTM model for time-series prediction](https://ieeexplore.ieee.org/abstract/document/9924783)

* Comparative Studies and Hybrid Approaches