DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

BANGLADESH ARMY UNIVERSITY OF SCIENCE & TECHNOLOGY (BAUST)
SAIDPUR CANTONMENT, NILPHAMARI

(Thesis Proposal)

Application for the approval of B.Sc. Engineering Thesis
(Computer Science & Engineering)

Session: Winter 2023 **Date:** 9 MAY, 2023

1. **Name of the Students(with ID) :**
   - Mst. Habiba Hena Sumi (200101070)
   - Most. Jannat-Ul-Fredoush (200101068)
   - Talath Un Nabi (200101076)

2. **Present Address :**Bangladesh Army University of Science & Technology, Saidpur Cantonment, Nilphamari.

3. **Name of the Supervisor :**Md. Mamun Hossain

   Designation : Assistant Professor
   Department of Computer Science & Engineering
   Bangladesh Army University of Science & Technology

4. **Name of the Department:** Computer Science & Engineering

   Program :B.Sc.Engineering

5. **Date of First Enrolment in the Program :**29.12.2019

Tentative Title : EXPLORING ADVANCED MODELS FOR TIME SERIES-BASED WEATHER FORECASTING IN BANGLADESH: A COMPARATIVE ANALYSIS OF ARIMA, SARIMA, FBPROPHET, LSTM, AND BILSTM MODELS**.**

6. Introduction:

The weather conditions in various areas have changed due to global warming. Where there used to be rain, there is now frequently dryness, and regions that formerly saw flooding are turning desert. The rivers are ceasing to flow, and the rain has caused areas that were once covered in water to suddenly become green again. Coastal water levels have risen due to sea level rise, putting several coastal cities in danger of being swamped in the near future. In Saudi Arabia, the weather remains incredibly hot; however, now we are witnessing the onset of rainfall and the beginning of tree growth. Moreover, in the Anog region, there has been a complete transformation in the weather, different from the past 100 years[1]. On the other hand, in Bangladesh, although there is more rainfall in the same location, over the last 3 years, we have observed a completely different pattern. It is presenting a completely opposite picture compared to five years ago. Future weather predictions, driven by advancements in temperature forecasting, empower us to anticipate and adapt to changing climatic conditions, aiding diverse sectors from agriculture to infrastructure planning. Integrating temperature and rain predictions provides a comprehensive outlook, facilitating informed decision-making to mitigate the impact of weather events on various aspects of our daily lives(agriculture to industry, from traveling to daily commuting.

7. Background and Present State of the Problem:

In recent years, various time series forecasting models have been extensively explored for weather prediction, showcasing a diverse range of methodologies and their application to meteorological data. The AutoRegressive Integrated Moving Average (ARIMA) model, a classical statistical approach, has been widely used in weather forecasting, as evidenced by studies such as Badal et al. (2022) [2], which compared ARIMA with Multilayer Perceptron (MLP) and Bidirectional Long Short-Term Memory (Bi-LSTM) models. Another classical model, the Seasonal AutoRegressive Integrated Moving Average (SARIMA), was employed by Ray et al. (2021) [3] to forecast monthly rainfall and temperature in South Asian countries. Comparisons of ARIMA with other models, such as Adaptive Neuro-Fuzzy Inference System (ANFIS), were conducted by Rahman et al. (2013) [4], highlighting ARIMA's superiority in certain performance metrics. More advanced models, including the Facebook Prophet model, have gained traction, with Sulasikin et al. (2021) [5] favoring Prophet over SARIMA and Long Short-Term Memory (LSTM) for monthly rainfall prediction in Central Jakarta. LSTM and Stacked LSTM models have been employed for temperature prediction, as demonstrated by studies like Karevan et al. (2019) [6]. Bidirectional Long Short-Term Memory (BiLSTM) models have been utilized for various applications, including daily reference evapotranspiration prediction (Roy et al., 2022) [7] and COVID-19 case forecasting (Said Erradi Aly Mohamed, 2021) [8]. Hybrid approaches, such as the combination of Stacked ResNet-LSTM and CORAL for air quality prediction (Cheng et al., 2022) [9], showcase the trend toward integrating multiple models for enhanced accuracy. Additionally, studies like Mollick et al. (2023) [10] have introduced stacking ensemble models, leveraging various base models for air temperature prediction. These advanced models, however,

come with challenges, including the need for substantial data and potential difficulties in generalizing findings. The hybrid approach of Neural Prophet and LSTM proposed by Prasad et al. (2023) [11] for rainfall prediction represents an innovative fusion of classical and deep learning techniques. Despite the promising results demonstrated by these models, it is crucial to acknowledge limitations, such as the sensitivity to parameter tuning, reliance on historical data, and challenges in handling complex meteorological phenomena. In summary, the field of weather forecasting has evolved significantly, moving from traditional statistical models like ARIMA and SARIMA to advanced deep learning models such as LSTM, BiLSTM, and hybrid approaches, each offering unique strengths and limitations in capturing the intricate patterns of meteorological data. The choice of model depends on the specific characteristics of the dataset and the forecasting task at hand.**Objective with Specific Aims and Possible Outcome**

Interacting to following Aims:

i. To collect up to date historical weather data of Bangladesh.
ii. To apply Time Series and Deep Learning models(ARIMA, SARIMA, Prophet, LSTM, BiLSTM) to capture comprehensive patterns in weather conditions.
iii. To predict future trend & weather of Bangladesh to uncover latent hierarchical patterns in the weather dataset for a deeper understanding.
iv. To enhance/optimize performance through advanced modeling and observational techniques in terms of MSE, RMSE, MAE, MAPE.
v. To inform government about our research findings.

Possible Outcome:

The purpose is to create a comprehensive weather forecasting model that uses deep learning and time series approaches to improve prediction accuracy, allowing for better decision-making and resource management in the face of more unexpected and extreme weather events.

8. Outline of Methodology Design

We obtained a comprehensive dataset by using the Open-Meteo historical weather API for data collection. After data collection, we handled missing values by locating and then removing them from the data.

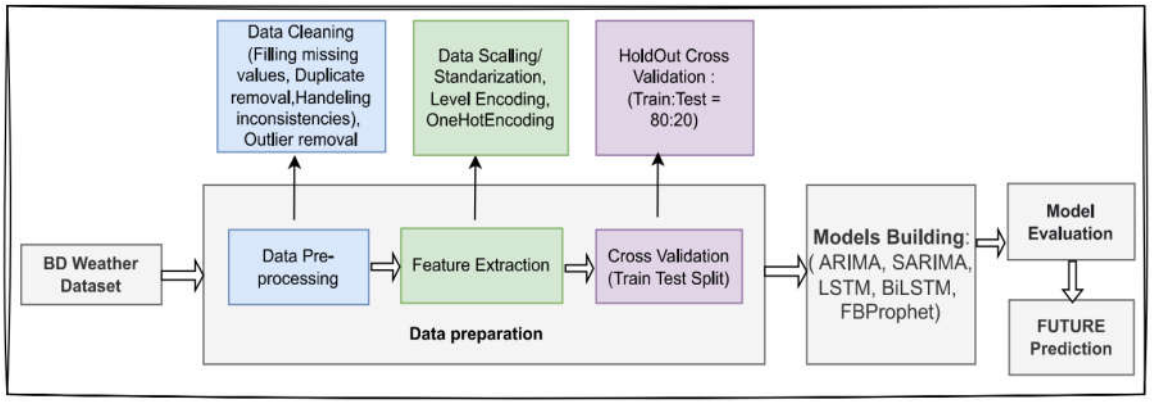

Figure: Flow Chart of Proposed Weather Forecasting System

1) Data Collection: We collected the dataset from the Open-Meteo website.
2) Data Preparation:
    - Handle Missing Values
    - Reshape Data
    - Data Scaling
3) Train-Test Split: The total dataset consists of 5735 data points. We divided the data into 80% for training (4586 data points) and 20% for testing (1149 data points).
4) Models for Training: We used ARIMA, SARIMA, Prophet, LSTM, and Bi-LSTM models for training.
5) Model Evaluation: Calculated Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) to evaluate the performance of our models.
6) Forecast: Predicted Future Weather for the Next 365 Days.

9. Resources Required to Accomplish the Task

Google Colab, Kaggle, Turnitin, Grammarly – Subscription

Nvidia RTX 4090 GPU, Intel Core i5 10^{th} Gen Processor

10. References

1. F. M. Alkolibi, "Possible effects of global warming on agriculture and water resources in saudi arabia: impacts and responses," Climatic change, vol. 54, no. 1-2, pp. 225–245, 2002.

2. M. K. I. Badal and S. Saha, "Performance analysis of deep neural network models for weather forecasting in bangladesh," in Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering: TCCE 2021. Springer, 2022, pp. 81–92.
3. S. Ray, S. S. Das, P. Mishra, and A. M. G. Al Khatib, "Time series sarima modelling and forecasting of monthly rainfall and temperature in the south asian countries," Earth Systems and Environment, vol. 5, pp. 531–546, 2021.
4. M. Rahman, A. S. Islam, S. Y. M. Nadvi, and R. M. Rahman, "Comparative study of anfis and arima model for weather forecasting in dhaka," in 2013 international conference on informatics, electronics and vision (ICIEV). IEEE, 2013, pp. 1–6.

5. . Sulasikin, Y. Nugraha, J. I. Kanggrawan, and A. L. Suherman, “Monthly rainfall prediction using the facebook prophet model for flood mitigation in central jakarta,” in 2021 International Conference on ICT for Smart Society (ICISS). IEEE, 2021, pp. 1–5.
6. C. B. A. Satrio, W. Darmawan, B. U. Nadia, and N. Hanafiah, “Time series analysis and forecasting of coronavirus disease in indonesia using arima model and prophet,” Procedia Computer Science, vol. 179, pp. 524–532, 2021.
7. D. K. Roy, T. K. Sarkar, S. S. A. Kamar, T. Goswami, M. A. Muktadir, H. M. AlGhobari, A. Alataway, A. Z. Dewidar, A. A. El-Shafei, and M. A. Mattar, “Daily prediction and multi-step forward forecasting of reference evapotranspiration using lstm and bi-lstm models,” Agronomy, vol. 12, no. 3, p. 594, 2022.
8. A. B. Said, A. Erradi, H. A. Aly, and A. Mohamed, “Predicting covid-19 cases using bidirectional lstm on multivariate time series - environmental science and pollution research,” May 2021. [Online]. Available: https: //link.springer.com/article/10.1007/s11356-021-14286-7.
9. X. Cheng, W. Zhang, A. Wenzel, and J. Chen, “Stacked resnet-lstm and coral model for multi-site air quality prediction,” Neural Computing and Applications, vol. 34, no. 16, pp. 13 849–13 866, 2022.
10. T. Mollick, G. Hashmi, and S. R. Sabuj, “A perceptible stacking ensemble model for air temperature prediction in a tropical climate zone,” Discover Environment, vol. 1, no. 1, p. 15, 2023.
11. G. L. V. Prasad, B. R. Teja, T. Haribabu, G. N. Pavani, D. Karunamma, and K. Vivek, “A hybrid time series rainfall prediction model using neural prophet and ls tm,” in 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS). IEEE, 2023, pp. 1582–1587.

11. Cost Estimation

a) Cost of Materials:
GPU cost: $1599 (Rtx 4090 GPU)
CPU cost: $589 (i9-13900K)
Google Colab Pro+: $49.99(Yearly)
Turnition: $9(yearly)
Grammarly: $15(yearly)

b) **Cost of Repot Printing and Binding:** $50

c) Others: $29.99

12. Committeefor Advance Studies and Research (CASR)

Meeting No: Resolution No: Date:

13. Number of Under-Graduate Students Working with the Supervisor at Present:12

<table>
<tr><th>Signature of the Students</th><th>Department of CSE</th></tr>
<tr><td>Jannat</td><td rowspan="2">

Signature of the Supervisor</td></tr>
<tr><td>Sumi</td></tr>
<tr><td>Anik</td><td rowspan="2">

Signature of the Head of the Department</td></tr>
<tr><td></td></tr>
</table>