Journal Home Page: https://sjes.univsul.edu.iq/

Research Article:

Climate Forecasting in Sulaymaniyah City Using Deep Learning Techniques

Rebin Abdulkareem Hamaamin [1,a,*]
Omar Mohammed Amin Ali [2,b]
Shahab Wahhab Kareemc[3,c]

[a]Department of Computer Science, College of Sciences, Charmo University, Chamchamal, Sulaymaniyah, Iraq.
[b]Department of Information Technology, Chamchamal Technical Institute, Sulaimani Polytechnic University, Sulaymaniyah, Iraq.
[c] Department of Technical Information Systems Engineering, Technical Engineering College, Erbil Polytechnic University, Erbil, Iraq.

Article Information

Article History:
Received: April 27th , 2025
Accepted :July 28, 2025
Available online: August 1st, 2025

Keywords: Weather, Temperature, Prediction models Weather, Bi-LSTM model, GRU model, GBT model, Deep learning

About the Authors:

Corresponding author:
Rebin Abdulkareem Hamaamin
E-mail: rebin.abdulkarim@chu.edu.iq

Researcher Involved:
Omar Mohammed Amin Ali
Asst. Prof. Dr.Shahab Wahhab Kareem

DOI https://doi.org/10.17656/sjes.10196

Abstract

A neural network model was used to categories and predict the weather in Sulaymaniyah, which is located in the Kurdistan Region of Iraq. The use of neural network models (NNM) for climate data analysis has advanced; the accuracy and application of neural network models for climate data analysis have significantly advanced in recent years. Using NNM affected the prediction; the dataset-based forecasting approach uses a single-column time series to depict the future. A real-time weather station dataset from the Sulaimani Meteorological and Seismological Directorate is used in the implementation model. The data collection includes implementation data from the weather station for an iterative neural network model that forecasts future climate and displays historical propagation. From 1993 to 2023, the dataset includes daily data on average temperature, humidity percentage, and precipitation for each of the twelve months from January to December. It was gathered every day for thirty years and includes information about Suleimani City's maximum and minimum temperatures, average temperatures, humidity levels, and rainfall over that time. The models were used to predict relative average temperatures. As the results indicate that Bi-LSTM and GRU outperformed GBT in both training and testing.

1. Introduction

Weather forecasting predicts the future weather conditions for a certain region by utilizing both science and technology. A model of the atmosphere is developed utilizing sophisticated technological computations and quantitative data on the present meteorological conditions in order to forecast its behavior. Supercomputers are frequently used for weather forecasting since the state of the atmosphere is dependent on a variety of factors, such as temperature, wind, rain, and humidity. Since the weather has a significant impact on many aspects of both human and animal existence, precise predictions are necessary for a variety of jobs, including those involving autonomous vehicles that depend on anticipating the state of the outside world. Accurate weather forecasts are essential for outdoor activities like farming since they may be used to predict how the weather will alter biodiversity and impact the environment. Significant temperature variations can have an impact on agriculture and wildlife, while dry weather and high winds can cause soil erosion. Because agricultural yields depend on the weather, which in turn impacts the economy, human health, and energy consumption, extreme weather will have an influence on food security.[1][2]. Weather forecasting serves to preserve the standard of living while also promoting public health and alleviating the effects of the recession. It is crucial to consider how weather variations affect people's safety and well-being, because agricultural planning depends on it, it is helpful in the agriculture industry. Farmers may make scientific decisions about their goods by using weather forecasts to choose whether to plant or postpone. Weather forecasting aids autonomous cars in making judgments that reduce traffic jams and accidents, which are solely reliant on the identification and forecasting of external environmental factors like air visibility, rainfall, and so on. Forecasting and analog weather prediction gained popularity in the late 19th century, but the barometer and other crude techniques were among the several employed in ancient times to predict the weather. Until the end of 1950, synoptic weather forecasting was the oldest method. Another traditional technique for forecasting the weather was numerical weather prediction (NWP), which made use of a mathematical model of the atmosphere. The statistical technique, which focused on forecasting future weather based on historical data, was the last phase of NWP. Given the importance of weather forecasting in our daily lives, many regions of the world now accept it as a valid technique. A lot of researchers are focusing on using machine learning to predict the weather.

2. Literature Review

Researchers have recently developed a variety of methods to increase the accuracy of weather forecasting. The researcher has endeavored to present studies on deep learning in this area. A range of meteorological applications have used artificial neural networks as well as artificial intelligence [3][4]. fuzzy logic and big data analytics methods for weather forecasting [5][6]. The blue color shows the location of the study area, the location of Sulaymaniyah province in Iraq, show Figure 1. The random forest method, one of the well-established machine learning techniques, is the foundation of the study conducted by [7], to estimate relative humidity (RH). Air quality forecasting of relative humidity uses the random forest algorithm as an effective technique. The RH prediction extracted through Advanced System for Process Engineering (ASPEN) - Hygrotech Systems (APSEN HYSYS) serves as the basis to run this technique and resolve complicated estimate problems while the products are connected in MATLAB. An evaluation on RF model prediction performance compared the random forest model to support vector machine (SVM) regression model results. The obtained results showed that the RF model demonstrated superior performance than the SVM model by 74.4%. A study [8]. demonstrates the evaluation of relative humidity prediction abilities through long short-term memory (LSTM) which represents a machine learning method. The LSTM model processed relative humidity predictions through analysis of dataset information collected at a weather station. A forecast of relative humidity derived from synoptic station measurements required training with time series data from 2008 to 2009. LSTM models apply prediction analysis to RH07, RH13, RH18 and RH average time intervals for performance evaluation. Machine learning represents an advanced forecasting method for RH prediction since models based on time-series data using LSTM deliver promising results. The authors at [9]. employed machine learning models to efficiently validate weather forecasting accuracy in northwest Bangladesh. The researchers conducted predictions based on thirty years of temperature and wind and rainfall and humidity observations collected from 1986 through 2017. The research utilizes ELM techniques jointly with ANN frameworks to solve weather prediction through evaluation with MAE and RMSE metrics. The ELM algorithms provide superior performance to ANN algorithms for forecasting rainfall alongside temperature and wind speed and humidity according to RSME, MAE, MASE, PP and CC evaluation standards. Research by [10]. employed CNN together with BI-LSTM networks for the prediction of multivariable nonlinear time series. For multi-step prediction the performance comparison included CONV-BI-LSTM together with LSTM models. The data contains numerous meteorological variables extending from temperature to humidity measurements. The measurements originated from Beijing weather station. The Conv-Bi-LSTM network achieves higher accuracy than the single LSTM model within the prediction of temperature as well as humidity measurements. The deep learning strategy remains crucial for generating accurate weather predictions [11]. This study creates a weather prediction model through the combination of Bi-LSTM (Bidirectional Long Short-Term Memory) and 1D CNN (1-Dimensional Convolution) algorithms. Three meteorological elements including 2-m temperature, 2-m relative humidity and 10-m wind speed are forecasted through the Bi-LSTM and 1D-CNN model system. The study utilized the "OBS" and "RMAPS" databases as their information source. The experimental results simplify the evaluation between the new hybrid model consisting of 1D-CNN and Bi-LSTM integration versus the FNN model built from 1D-CNN with LSTM and BI-LSTM components. Weather prediction benefits most from the implementation of the hybrid model. The implementation of one-hot encoding led to better accuracy results when predicting weather conditions in unforeseen neighborhood areas. In [12] will employ five classifications "Cold," "Cool," "Normal," "Warm," and "Hot" as well as other meteorological variables, including temperature, humidity, rainfall, and wind speed from 2000 to 2019, to forecast the temperature for the following three days. Over the last 20 years, information has been collected from the Bureau of Meteorology, Climatology, and Geophysics (BMKG) in Bandung. Weather parameters are predicted using RNN and LSTM models; pre-processed data is transformed as input data via segmentation, interpolation, feature extraction, and normalization. The accuracy was 90.92 percent using Adam's optimization model with 100 epochs in the training data and 80.36 percent in the test data. Therefore, the optimization model, data volume, and data sharing may influence the final results. This work uses vibrational mode decomposition (VMD) and extreme learning machine (ELM) to develop a novel combination model for electric load prediction many steps ahead of time [13]. The Australian energy market data comes from two sources which include New South Wales (NSW) and

Queensland (QLD). Experimental findings revealed three essential points during the research. The recommended forecasting algorithm delivered better accuracy than one-step and multi-step forward electric load forecasting methods. Results indicated that VMD made a higher number of mistakes once compared to ELM. The DE method served to enhance both thresholds and initial weights for the ELM model in order to boost its forecasting effectiveness. The research builds a data-driven forecast model for weather applications through the implementation of LSTM according to [14]. The method known as T-LSTM provides Transudative LSTM as a time series estimation tool which leverages local data points. Researchers obtained meteorological data and characteristics on minimum and maximum temperatures from the Meteorological Underground website. LSTM achieves better performance than transudative LSTM only when different sequence lengths and transfer functions are employed. Numerous studies prove that inductive as well as transudative LSTM models use forecast methods equivalent to the most sophisticated weather prediction algorithms despite minimal available data. The performance of the inductive LSTM model under evaluation increases from April and May into November and December. The addition of novel weight data to temporal convolutional networks (TCN) and long short-term memory (LSTM) represents [15] approach for weather enhancement and prediction. The researchers have developed Arbitrage of Forecasting Expert (AFE) as a new approach to determine its suitability for short-term weather forecasting versus standard techniques including Autoregressive Integrated Moving Average (ARIMA), Vector Auto Regression (VAR), Vector Error Correction Model (VECM), standard regression (SR), support vector regression (SVR) and random forest (RF). The proposed neural network model serves for short-term weather forecasting which gets compared to WRF model predictions. The WRF model operates on GRIB data. We collected a total of twelve meteorological variables from January to May of 2018. As a result, the MISO method outperforms MIMO in terms of MSE values. Compared to the LSTM, the Bi-LSTM generated lengthier forecasts with more accuracy. Therefore, the Bi-LSTM could potentially enhance accuracy. The deep learning model often produces better predictions for up to 12 hours when compared to the WRF model. The goal of [16] is to introduce a new framework for automatically extracting weather and visual conditions information from street-level images. Depending on the image analysis, these conditions can include weather states like foggy, clear, snowy, and rainy; time periods such as dusk, dawn, night, and day; and lighting conditions including glare. Using CNN (Convolutional Neural Network), frequently referred to as a "weather net," the researcher collected data from many Kaggle sources. Consequently, utilizing computer vision and deep learning, the proposed weather net has shown outstanding performance in identifying the different kinds of photo pairings. The purpose of this work [17] is to accurately estimate the rainfall status using a unique hybrid machine learning approach: Intense Neural Network Mining (INNM) is the name of the proposed methodology. The INNM approach uses two distinct machine learning logics the Rapid Miner and the Back Propagation Neural Network to describe how rainfall situations are predicted. The researcher utilized a new set of data from the Chennai Regional Meteorological Center. The next portions of this study provide strong validation of this finding in a graphical process. Consequently, the prospective proposal of INNM yields a forecasting accuracy of almost 96.5% with a base error rate of 0.04%. A neural network is used to forecast the weather [18]. A fully connected neural network (FCNN) is recommended for classifying weather data. A significant quantity of data was gathered by the Indian Meteorological Department (IMD) experts. Temperature, dew point, humidity, wind type, wind speed, wind gust, and pressure are the sample values. Consequently, in some metrics like UA, PA, and KC, the performance of the FCNN model is much better than that of the fine Gaussian SVMs (FGs). A few of the publications considered in the literature review are summarized in Table 1.

Table 1 presents a comparison of a few earlier literature papers. The table displays the articles' publication date, purpose, datasets, applicable methodologies, and outcomes.

3. Materials and Methods

Deep learning algorithms are included in this research. Figure 2 illustrates how we use actual meteorological and seismological data from Suleimani City to forecast temperature. After acquiring the dataset, the pre-processing activity involves applying Min-Max normalization. Three deep learning models Bi-LSTM, GRU, and GBT are used to build a neural network for weather forecasting. The networks used to forecast temperature in this study provide a clear explanation of the model's implementation processes. Following the findings, the assessment procedure is carried out utilizing metrics such as MAE, RSME, and R^2, which are thoroughly described in the Materials and Methods section.

3.1. Dataset

Deep learning models require a large amount of data to produce the best model or a high-fidelity system. While quantity is vital, data quality is equally important, even with the most advanced machine learning algorithms. Whether a machine learning or deep learning model works well depends on the quantity, quality, and relevance of the dataset. The first and most crucial step in weather forecasting is thought to be data collection. The study's primary data collection sources for putting certain deep learning models into practice are the Suleimani Meteorology and Seismology. Meteorological observatory stations collect different types of climate data at different times. One station may have many sensors that measure humidity, precipitation (rain), average temperature, maximum temperature, and minimum temperature. We have attempted to obtain the precise data set here. The data collection is first gathered on a Daily's basis for 30 years (1993–2023). Table 2 displays the data for the first month of 1993, while Table 3 displays the data for the final month of 2023. The data set includes several factors, such as the maximum temperature, minimum temperature, average temperature, humidity, and precipitation.

3.2. Preprocessing Modules

A step in the data mining and analysis process called data pre-processing converts unprocessed data into a format that computers and machine learning algorithms can understand. This study views data processing as the initial phase of the approach, following data gathering. Suleimani City's meteorology and seismology provide the dataset. The missing value procedure, which entails averaging the missing values, is the first stage since the parameter has some missing values. Data cleaning is the most crucial preprocessing phase since it prepares the data for training. This procedure makes use of the min-max normalization formula. One of the most important data processing techniques is data normalization. It could disclose a new range, which is crucial for forecasting and prediction. We use the min-max normalization technique on all of the dataset's parameters to build the models.

3.3. Bi-LSTM Model

Bi-LSTM networks utilize advanced recurrent neural network architecture that improves learning capabilities through its dual-directional processing mechanism which works on the input sequence from both ends. The parallel processing method of the model retrieves information from both previous and upcoming time steps to build better contextual understanding. The Bi-LSTM configuration integrates double LSTM units where each layer conducts information processing from opposite sequence directions. The feature representation strength increases when both layers provide their outputs for final integration. The implemented Bi-LSTM model begins with a first layer containing 128 units for both types of direction then proceed to a second layer with 64 units each way. The network uses dropout mechanisms to avoid overfitting and fully connected blocks to mold the output results for predictions. The system contains a total of 534,765 trainable parameters. The refined Bi-LSTM model achieves learning stability and enables access to more information thus making it appropriate for handling complex sequential data. The proposed Bi-LSTM model presents its architectural structure through the Figure 3 illustration [23].

3.4. GRU Model

Time series forecasting reaches enhanced efficiency through combination of bidirectional recurrent layers and dropout regularization within the Enhanced Gated Recurrent Unit (GRU) model. The model contains two stacked bidirectional GRU layers that have 128 units in the first layer to return sequences for forward and backward dependency tracking. The subsequent GRU layer contains 64 units which identifies advanced sequential characteristics before the information enters the fully connected layer. The model contains dropout layers set to 0.2 that are placed between GRU layers to reduce overfitting while improving generalization capabilities. By utilizing a dense network with 32 ReLU nodes additional feature details become noticeable to the system. The prediction of the target value is the job of the last dense layer. The model is optimized using the Adam optimizer with a learning rate of 0.001, and it employs the Mean Squared Error (MSE) loss function to minimize prediction errors. This architecture ensures efficient learning of temporal dependencies while maintaining robustness against overfitting [24]. Show Figure 4.

3.5. GBT Model

Gradient Boosting Trees (GBT) is a type of ensemble learning that creates predictive models by teaching weak learners (decision trees) one after the other to fix mistakes made by earlier learners. The model gradually improves accuracy by training each new tree on the residual errors of the prior trees. The model uses a loss function to get the best performance and learning rate control to keep it from becoming too good at what it does. GBT is highly effective for structured data, handling both regression and classification tasks. It provides feature importance insights and can be fine-tuned using hyperparameters like tree depth, number of estimators, and learning rate [25]. Show Figure 5.

3.6. Evaluation Model

Three different measures were utilized for the purpose of evaluation: the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), and the R^2 score. RMSE is a measurement that determines the average size of the mistake, with higher errors being given more weight. A smaller root mean square error (RMSE) suggests higher performance, see equation (1). MAE is a method that determines the average absolute error while considering all mistakes in the same manner. When the MAE is less, it implies that the forecast may be more accurate, see equation (2). A value that is closer to one indicates that the model is doing better. The R^2 score is a measurement of how well the predictions match the actual data see equation (3).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \quad (1)$$

- y_i: Actual Value
- x_i: Predicted Value
- n: Total number of observations

$$MAE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|y_i - x_i|} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - x_i)^2}{\sum_{i=1}^{n}(y - \bar{x})^2} \quad (3)$$

- $\bar{x}$: Mean of actual values.
- The closer R^2 is to **1**, the better the model.

4. Result

This study assesses various deep learning models utilizing a real dataset to determine the most effective and accurate predictions of the weather parameter, average temperature. The study aims to implement the models Bi-LSTM, GRU, and GBT. To create and evaluate models, the researcher used deep learning methods and the Python programming language. Seismology and meteorology in Suleimani might provide the dataset. We gather the dataset every day.
The Results section presents daily data collected, incorporating elements such as maximum temperature, lowest temperature, humidity, and precipitation, which are affected by fluctuations in the average temperature value. The initial procedure involved splitting the data into an 80% training set and a 20% testing set. The data scale was transformed using min-max normalization.

4.1. Bi-LSTM Model

Using meteorological data that has been gathered every day for fourteen years (1993–2023), the Bi-LSTM model is used to forecast average temperature. ReLU is the activation function for the Bi-LSTM model. There are 75 epochs and 32 batches. The same metrics, RMSE, MAE, and R^2, are utilized for all models employed in this research to display the average temperature forecast and outcome, as shown in Figure 6. Figure 7 displays the Bi-LSTM result model.

4.2. GRU Model

The GRU model is employed to estimate average temperature using a meteorological dataset collected daily over a span of fourteen years (1993-2023). The model consists of two stacked bidirectional GRU layers, where the first layer contains 128 units and returns sequences to capture temporal dependencies in both forward and backward directions. The total number of epochs is 75. All models included in this study utilize the same metrics: RMSE, MAE, and R^2, to demonstrate the results and predictions of average temperature. As seen in Figure 8. The GRU outcome model is depicted in Figure 9.

4.3. GBT Model

Using a meteorological dataset that has been collected every day for fourteen years (1993–2023), the GBT model is used to estimate the average temperature. When training the XGBoost Regressor model, GBT uses settings such as learning rate=0.01, n estimators=1000, and max depth=6. There are 75 epochs in all. The same metrics RMSE, MAE, and R^2 are used by all of the models in this study to illustrate the average temperature forecasts and outcomes. as seen in Figures 10. As a result, Figure 11 shows the GBT outcome model.
The performance of the GBT, GRU, and Bi-LSTM models over 75 epochs is displayed in Figures 12, 13, and 14, respectively. A pattern of prediction errors over time is depicted in Figure 12, which displays the Root Mean Squared Error (RMSE) for each of the three models. The Mean Absolute Error (MAE) for the same models is displayed in Figure 13, which draws attention to the differences in absolute error that occur throughout the training process. Last but not least, Figure 14 illustrates the R^2 (R-squared) values for the GBT, GRU, and Bi-LSTM models. These scores demonstrate the prediction accuracy of the models as well as how well they match the data.

5. Discussion

The Bi-LSTM, GBT and GRU models received assessment across training and testing datasets through three performance indicators which consisted of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) and R^2 Score. The Bi-LSTM model produced outstanding results through its low training set MAE of 1.0044 and RMSE of 1.3073 and high R^2 score of 0.9833. The testing data demonstrated a minor decrease in performance with an MAE value of 1.9237 alongside RMSE value of 2.5334 resulting in an R^2 score of 0.9399. The GBT model produced training error values that were higher than the other models since it achieved MAE scores of 3.5744 and RMSE scores of 4.2577 which yielded a R^2 score of 0.8230. The evaluation on the testing set revealed a slight overfitting issue because

the MAE reached 3.6265 while RMSE reached 4.3145 but R^2 maintained a higher value of 0.8257. During training the GRU model achieved similar performance to the Bi-LSTM model through its I.9701 MAE and I.3067 RMSE and I.9833 R^2 score. In the testing scenario the GRU model demonstrated an average error rate of 1.8293 measured by MAE while its RMSE value reached 2.4232 and its R^2 score amounted to 0.9450 indicating it maintained a stable yet slightly weaker performance than training. The Bi-LSTM together with GRU offered superior performance compared to GBT mainly because of their higher R^2 prediction accuracy score. The Bi-LSTM showed the best overall performance with the highest R^2 score on the training dataset and a small drop on the testing set, which can be attributed to some level of overfitting. The GBT model's higher error metrics suggest that it was less effective at generalizing to unseen data compared to the deep learning models. These findings are summarized in Table 4 and visually represented in Figure 15.
Table 5 presents the findings of a performance comparison between the proposed system and certain state of the art techniques.

6. Conclusion

This research employed a data science methodology to investigate the potential use of deep learning techniques for forecasting average temperatures. Predicting the weather is an essential application for 365 days of the year. This research predicts the relative average temperature, a critical meteorological element, using deep learning algorithms. The objective of this study is to use three deep-learning models to anticipate the average temperature for each day. Using real data and meteorological conditions, we trained a number of deep learning algorithms, such as Bi-LSTM, GBT, and GRU, to forecast the average temperature. We compared the outcomes of training and testing to see how well the predictions performed. The average temperature was well predicted by models based on data that was collected every day for many years. The models had low RMSE, MAE, and R^2 errors. In general, the data set used for prediction is split. It was gathered every day for thirty years and includes information about Suleimani City maximum and minimum temperatures, average temperatures, humidity levels, and rainfall over that time. One of the other goals of this investigation is to test and compare the models used to predict relative average temperatures. After applying these models to the daily data, we found that GRU outperforms the other models, achieving an RMSE of 1.3067. On the other hand, the results indicate that Bi-LSTM and GRU outperformed GBT in both training and testing. When they were trained, Bi-LSTM and GRU had lower RMSE (1.3073 and 1.3067) and MAE (1.0044 and 0.9770), but GBT made more mistakes (RMSE = 4.2577, MAE = 3.5744). Bi-LSTM and GRU did better in tests (RMSE = 2.5334, 2.4232), and their higher R^2 values (0.9399, 0.9450) showed that they were stable and accurate.

Acknowledgments

The authors would like to thank, Geography Department, Faculty of Education, Koya University, for his assistance.

References

[1] Cifuentes, J. et al. (2020) ‘Air temperature forecasting using machine learning techniques: A review’, Energies, 13(6), pp. 1–28. doi: 10.3390/en13164215.

[2] Pratyush Reddy, K. S. et al. (2020) ‘IoT based Smart Agriculture using Machine Learning’, Proceedings of the 2nd International Conference on Inventive Research in Computing Applications, ICIRCA 2020, (July), pp. 130–134. doi: 10.1109/ICIRCA48905.2020.9183373.

[3] Dewitte, S. et al. (2021) ‘Artificial intelligence revolutionises weather forecast, climate monitoring and decadal prediction’, Remote Sensing, 13(16), pp. 1–12. doi: 10.3390/rs13163209.

[4] Donadio, L., Fang, J. and Porté-Agel, F., 2021. Numerical weather prediction and artificial neural network coupling for wind energy forecast. Energies, 14(2), p.338. http://dx.doi.org/10.3390/en14020338

[5] Harahap, A.M., Suwilo, S. and Sembiring, R.W., 2021, February. The mamdani fuzzy logic engineering analysis for determining weather forecast. In Journal of Physics: Conference Series (Vol. 1783, No. 1, p. 012039). IOP Publishing.

[6] Fathi, M. et al. (2022) ‘Big Data Analytics in Weather Forecasting: A Systematic Review’, Archives of Computational Methods in Engineering, 29(2), pp. 1247–1275. doi: 10.1007/s11831-021-09616-4.

[7] Qadeer, K. et al. (2021) ‘Developing machine learning models for relative humidity prediction in air-based energy systems and environmental management applications’, Journal of Environmental Management, 292(April). doi: 10.1016/j.jenvman.2021.112736.

[8] Hutapea, M. I. et al. (2020) 'Prediction of relative humidity based on long short-term memory network', AIP Conference Proceedings, 2221(March). doi: 10.1063/5.0003171.

[9] Rizvee, M. A. et al. (2020) 'Weather Forecasting for the North-Western region of Bangladesh: A Machine Learning Approach', 2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020, pp. 1–6. doi: 10.1109/ICCCNT49239.2020.9225389.

[10] Wang, R. (2018) Proceedings of 2017 9th International Conference On Modelling, Identification and Control, ICMIC 2017, Proceedings of 2017 9th International Conference On Modelling, Identification and Control, ICMIC 2017.

[11] Fu, Q. et al. (2019) 'Multi-stations' weather prediction based on hybrid model using 1D CNN and Bi-LSTM', Chinese Control Conference, CCC, 2019-July, pp. 3771–3775. doi: 10.23919/ChiCC.2019.8866496.

[12] Rahayu, I.S., Djamal, E.C., Ilyas, R. and Bon, A.T., 2020, August. Daily temperature prediction using recurrent neural networks and long-short term memory. In Proceedings of the International Conference on Industrial Engineering and Operations Management (pp. 2700-2709).

[13] Lin, Y. et al. (2017) 'An ensemble model based on machine learning methods and data preprocessing for short-term electric load forecasting', Energies, 10(8). doi: 10.3390/en10081186.

[14] Karevan, Z. and Suykens, J. A. K. (2020) 'Transductive LSTM for time-series prediction: An application to weather forecasting', Neural Networks, 125, pp. 1–9. doi: 10.1016/j.neunet.2019.12.030.

[15] Hewage, P. et al. (2021) 'Deep learning-based effective fine-grained weather forecasting model', Pattern Analysis and Applications, 24(1), pp. 343–366. doi: 10.1007/s10044-020-00898-1.

[16] Priyanka, N. Sreeja Luis, Manogna Palli and Niharika Uppu . (2021) 'Weather Prediction Using Deep Learning Techniques', pp. 11069–79, Vol 12, Issue 05, MAY /2021, ISSN NO:0377-9254

[17] Sakthivel, S. and Thailambal, G., (2021). Effective procedure to predict rainfall conditions using hybrid machine learning strategies. Turkish Journal of Computer and Mathematics Education, 12(6), pp.209-216. https://turcomat.org/index.php/turkbilmat/article/view/1291.

[18] Hemalatha, G., Rao, K.S. and Kumar, D.A., (2021), November. Weather prediction using advanced machine learning techniques. In Journal of Physics: Conference Series (Vol. 2089, No. 1, p. 012059). IOPPublishing.

[19] Singh, N., Chaturvedi, S. and Akhter, S. (2019) 'Weather Forecasting Using Machine Learning Algorithm', 2019 International Conference on Signal Processing and Communication, ICSC 2019, pp. 171–174. doi: 10.1109/ICSC45622.2019.8938211.

[20] Heydari, M. et al. (2020) 'Application of holt-winters time series models for predicting climatic parameters (Case study: Robat Garah-Bil station, Iran)', Polish Journal of Environmental Studies, 29(1), pp. 617–627. doi: 10.15244/pjoes/100496.

[21] Sharma, N. et al. (2011) 'Predicting solar generation from weather forecasts using machine learning', 2011 IEEE International Conference on Smart Grid Communications, SmartGridComm 2011, pp. 528–533. doi: 10.1109/SmartGridComm.2011.6102379.

[22] Zhan, C. et al. (2019) 'Daily rainfall data construction and application to weather prediction', Proceedings - IEEE International Symposium on Circuits and Systems, 2019-May(May). doi: 10.1109/ISCAS.2019.8702124.

[23] Sun, H., Cui, Q., Wen, J., Kou, L. and Ke, W., (2024). Short-term wind power prediction method based on CEEMDAN-GWO-Bi-LSTM. Energy Reports, 11, pp.1487-1502. https://doi.org/10.1016/j.egyr.2024.01.021

[24] Daraghmi, E.Y., Qadan, S., Daraghmi, Y., Yussuf, R., Cheikhrouhou, O. and Baz, M., (2024).

From Text to Insight: An Integrated CNN-BiLSTM-GRU Model for Arabic Cyberbullying Detection. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3431939.

[25] Rizkallah, L.W. Enhancing the performance of gradient boosting trees on regression problems. J Big Data 12, 35 (2025). https://doi.org/10.1186/s40537-025-01071-3

[26] Jin, S. et al. (2021) 'Application of deep learning methods in biological networks', Briefings in Bioinformatics, 22(2), pp. 1902–1917. doi: 10.1093/bib/bbaa043.

[27] Chhetri, M. et al. (2020) 'Deep BLSTM-GRU model for monthly rainfall prediction: A case study of Simtokha, Bhutan', Remote Sensing, 12(19), pp. 1–13. doi: 10.3390/rs12193174.

[28] Khan, M. I. and Maity, R. (2020) 'Hybrid Deep Learning Approach for Multi-Step-Ahead Daily Rainfall Prediction Using GCM Simulations', IEEE Access, 8(MI), pp. 52774–52784. doi: 10.1109/ACCESS.2020.2980977

[29] Rahman, F. I. (2020) 'Short Term Traffic Flow Prediction Using Machine Learning - Knn, Svm and Ann With Weather Information', International Journal for Traffic and Transport Engineering, 10(3), pp. 371–389. doi: 10.7708/ijtte.2020.10(3).08.

التنبؤ بالمناخ في مدينة السليمانية باستخدام تقنيات التعلم العميق

المستخلص

تم استخدام نموذج الشبكة العصبية لتصنيف الطقس والتنبؤ به في السليمانية، التي تقع في إقليم كردستان العراق. وقد تقدم استخدام نماذج الشبكة العصبية (NNM) لتحليل بيانات المناخ؛ وتطورت دقة وتطبيق نماذج الشبكة العصبية لتحليل بيانات المناخ بشكل كبير في السنوات الأخيرة. أثر استخدام NNM على التنبؤ؛ حيث يستخدم نهج التنبؤ القائم على مجموعة البيانات سلسلة زمنية أحادية العمود لتصوير المستقبل. يتم استخدام مجموعة بيانات محطة الطقس في الوقت الفعلي من مديرية الأرصاد الجوية والزلازل في السليمانية في نموذج التنفيذ. يتضمن جمع البيانات بيانات التنفيذ من محطة الطقس لنموذج الشبكة العصبية التكراري الذي يتنبأ بالمناخ المستقبلي ويعرض الانتشار التاريخي. من عام 1993 إلى عام 2023، تتضمن مجموعة البيانات بيانات يومية عن متوسط درجة الحرارة ونسبة الرطوبة وهطول الأمطار لكل شهر من الاثني عشر شهرًا من يناير إلى ديسمبر. تم جمعها كل يوم لمدة ثلاثين عامًا وتتضمن معلومات عن درجات الحرارة العظمى والصغرى في مدينة السليمانية ومتوسط درجات الحرارة ومستويات الرطوبة وهطول الأمطار خلال تلك الفترة. استُخدمت النماذج للتنبؤ بمتوسط درجات الحرارة النسبية. وتشير النتائج إلى أن نموذجي Bi–LSTM و GRU تفوقا على نموذج GBT في كلٍّ من التدريب والاختبار.

الكلمات المفتاحية:

الطقس، درجة الحرارة، نماذج التنبؤ الطقس، نموذج Bi–LSTM، نموذج GRU، نموذج GBT، التعلم العميق.

Tables and Graphs and pictures

Figure 1. The blue color shows the location of the study area, the location of Sulaymaniyah province in Iraq.

Table 1: A comparison of the reviews of papers

Reference	Dataset(s)	Modules	Results
[8]	The weather in Indonesia as reported by the Synoptic (9601) Station.	LSTM	RH testing = 0.40, RH training = 0.46
[9]	BMD dataset (Bangladesh meteorological department).	ANN module and ELM.	Compared to ANN, ELM offers a 70% higher performance rate and 95% accuracy.
[11]	Observation RMAPS	1D-CNN, Bi-LSTM, LSTM, FNN, 1D- CNN+LSTM, 1D- CNN+Bi-LSTM.	1D-CNN+Bi-LSTM = 0.42, FNN = 0. 34, LSTM=0.40, 1D-CNN=0.40, 1D-CNN + LSTM = 0.40, And Bi-LSTM = 0.41.
[12]	BMKG (Bandung Meteorology Climatology and Geophysics)	RNN module and LSTM.	training of Adam's optimization accuracy= 90.92%, testing of Adam's optimization accuracy= 80.36%, SGD's training data accuracy = 87.24%, SGD's testing data accuracy was 76.48%.
[17]	C-RMC (Centre Regional Meteorological Chennai)	RM, INNM and BPNN.	Predicting accuracy = 96.5% and the lowest error ratio = 0.04%
[18]	Indian metrological Department (IMD)	FCNN and FGS	FGS (%) for classes 1: UA=79.63, PA=68.24. UA=89.14, PA=92.46, KC=0.867 FCNN (%) for classes 1: UA=85.45, PA=71.48, UA=92.74, PA=96.46, KC=0.945
[19]	The system forecasts the likelihood of rain for the current day using real-time data from sensors that measure temperature, pressure, and humidity.	Classification with random forests	A score is utilized to determine the correctness of the model, and 87.90% accuracy is obtained.
[20]	Information was gathered from the Robat Gharah-Bil Station in Golestan, Iran, over a thirty-year span (from 1981 to 2010).	The model of Holt-Winters Models of Additive Multiplication	Rainfall: MAPE = 0.35, RMSE = 0.49; Sunlight: MAPE = 0.79, RMSE = 0.01.
[21]	historical sun intensity and NWS projections over a ten-month period	SVM-RBF and PPF modules	SVM-RBF accuracy =51%. PPF accuracy = 27%.
[22]	More than thirty meteorological variables and five basic climatic groups made up the restored climate.	CNN, RNN , LSTM, XGBoost and GBDT.	Suggested strategy’s is close = 80% accuracy.

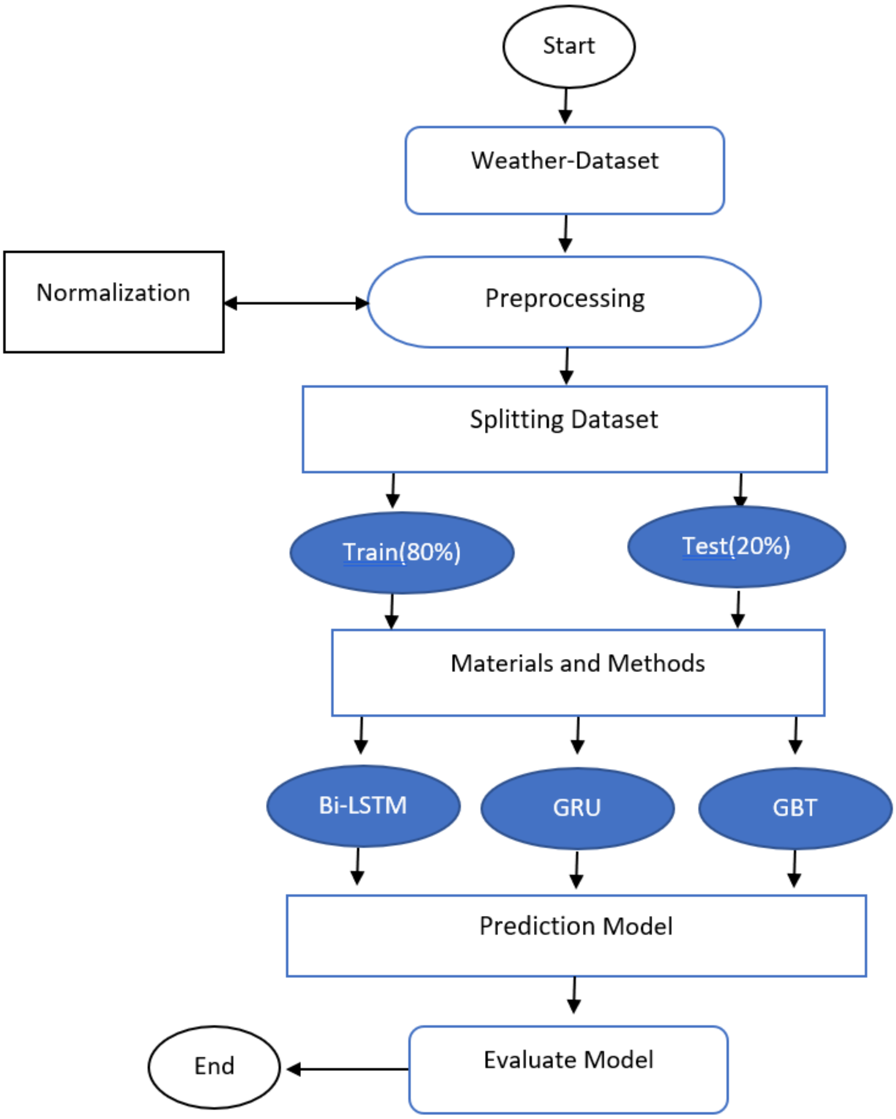

Figure 2: The General Diagram of Materials and Methods

Table 2. Dataset for three parameters of weather from month January in 1993

Date	Max Temperature	Min Temperature	Average Temperature	Humidity %	Precipitation
1/1/1993	8.1	-1.6	3.25	55	0
1/2/1993	8.3	-0.7	3.8	65	0
1/3/1993	11.4	2.4	6.9	71	0
1/4/1993	11	4	7.5	65	0
1/5/1993	9.8	3	6.4	68	3
1/6/1993	9	4	6.5	86	3.9
1/7/1993	10	4	7	83	3.9
1/8/1993	9.8	5.3	7.55	80	11
1/9/1993	4.6	0.5	2.55	89	0.5
1/10/1993	5.7	1.9	3.8	80	13
1/11/1993	6	2.4	4.2	86	0
1/12/1993	6.5	3.4	4.95	57	0
1/13/1993	7.3	3	5.15	59	0
1/14/1993	5.2	-0.8	2.2	47	0
1/15/1993	5.8	-1	2.4	68	0.5
1/16/1993	3.9	-1.4	1.25	60	0
1/17/1993	6.1	-2.6	1.75	46	0
1/18/1993	5.5	-4	0.75	57	0
1/19/1993	7	-4.6	1.2	33	0
1/20/1993	9.8	-2.2	3.8	46	0
1/21/1993	9	-2	3.5	47	0
1/22/1993	8.7	-1	3.85	60	0
1/23/1993	10	-1	4.5	58	1.2
1/24/1993	9.8	-9	0.4	54	0
1/25/1993	9	-1.2	3.9	64	1
1/26/1993	12.1	3.1	7.6	53	0
1/27/1993	13.8	2.2	8	55	10.7
1/28/1993	15	7	11	38	0
1/29/1993	10.3	7	8.65	64	0
1/30/1993	13	1.5	7.25	68	0
1/31/1993	12.2	2.2	7.2	57	0

Table 3. Dataset for three parameters of weather from month December in 2023.

Date	Max Temperature	Min Temperature	Average Temperature	Humidity %	Precipitation
12/1/2023	19.2	6.2	12.7	63	0
12/2/2023	19.5	5.1	12.3	61	0
12/3/2023	19.5	5.1	12.3	63	0
12/4/2023	20.2	5.6	12.9	65.5	0
12/5/2023	20.3	5.5	12.9	62	0
12/6/2023	20.2	8	14.1	60	0
12/7/2023	18.6	10.4	14.5	63	0
12/8/2023	19	8.5	13.75	63.5	0
12/9/2023	18.6	10.3	14.45	61.5	0
12/10/2023	18.9	10.2	14.55	70	0
12/11/2023	18.6	8	13.3	72	0
12/12/2023	18.2	11	14.6	63	3.2
12/13/2023	11	10	10.5	91	16.3
12/14/2023	15.3	7.7	11.5	85.5	0.4
12/15/2023	15	5.5	10.25	80.5	0
12/16/2023	17.1	4.3	10.7	75.5	0
12/17/2023	16.5	5.1	10.8	68	0
12/18/2023	15.2	5.9	10.55	64	0
12/19/2023	14	6	10	54.5	0
12/20/2023	15	5.6	10.3	63	0
12/21/2023	15	5	10	70.5	3.4
12/22/2023	9.5	6.5	8	90.5	8.2
12/23/2023	11.2	8.2	9.7	83	15.7
12/24/2023	11.2	9.5	10.35	85.5	11.9
12/25/2023	15.2	7.5	11.35	77	0
12/26/2023	13.9	6	9.95	74.5	0
12/27/2023	16.4	4.9	10.65	71	0
12/28/2023	16	7.8	11.9	69	0
12/29/2023	17	5.4	11.2	68	0
12/30/2023	17.6	6.5	12.05	66.5	0
12/31/2023	15.7	4.1	9.9	79.5	0

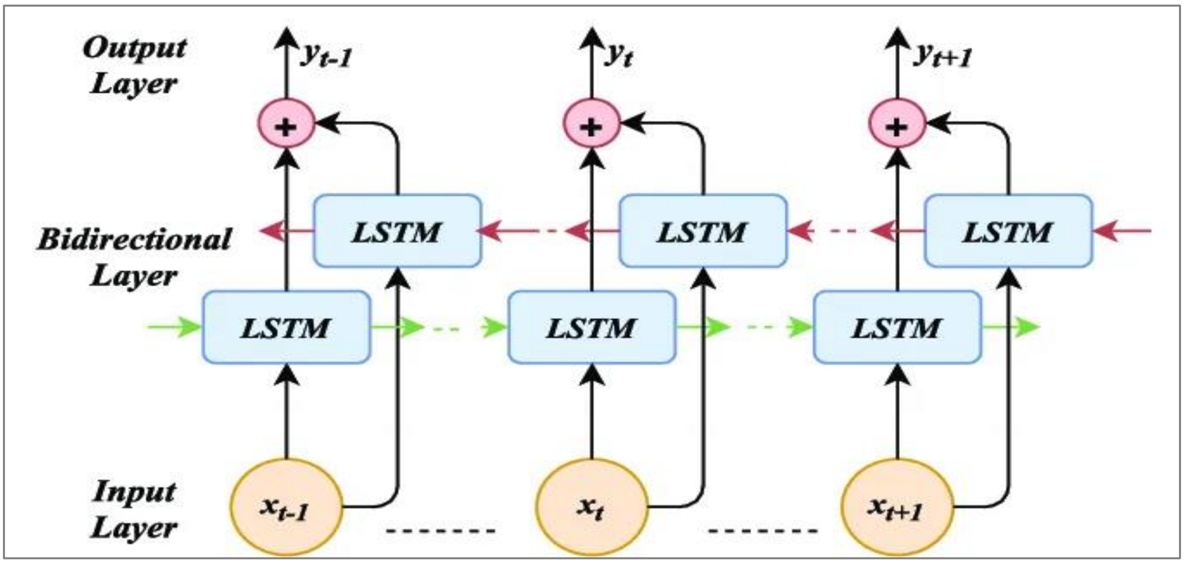

Figure 3: The Diagram of Bi-LSTM module.

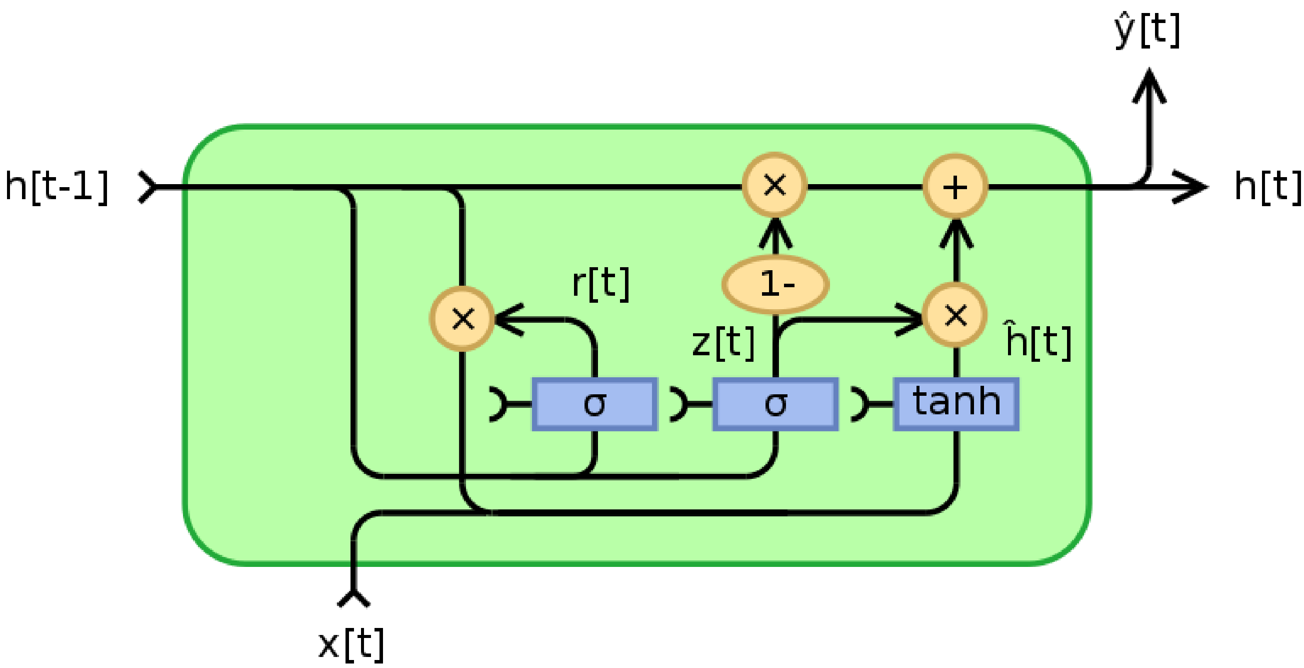

Figure 4: The Diagram of GRU module.

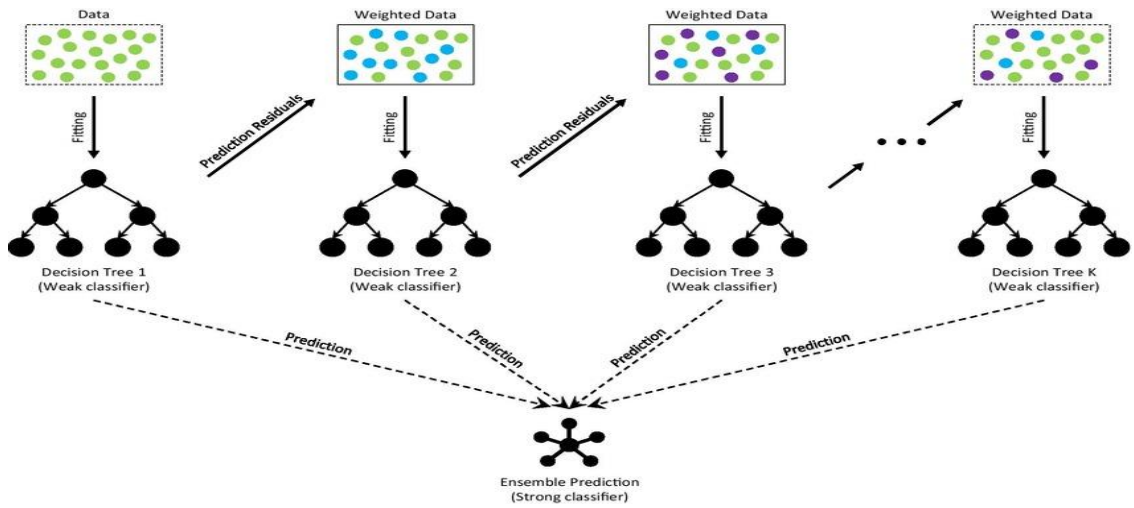

Figure 5: The Diagram of GBT module.

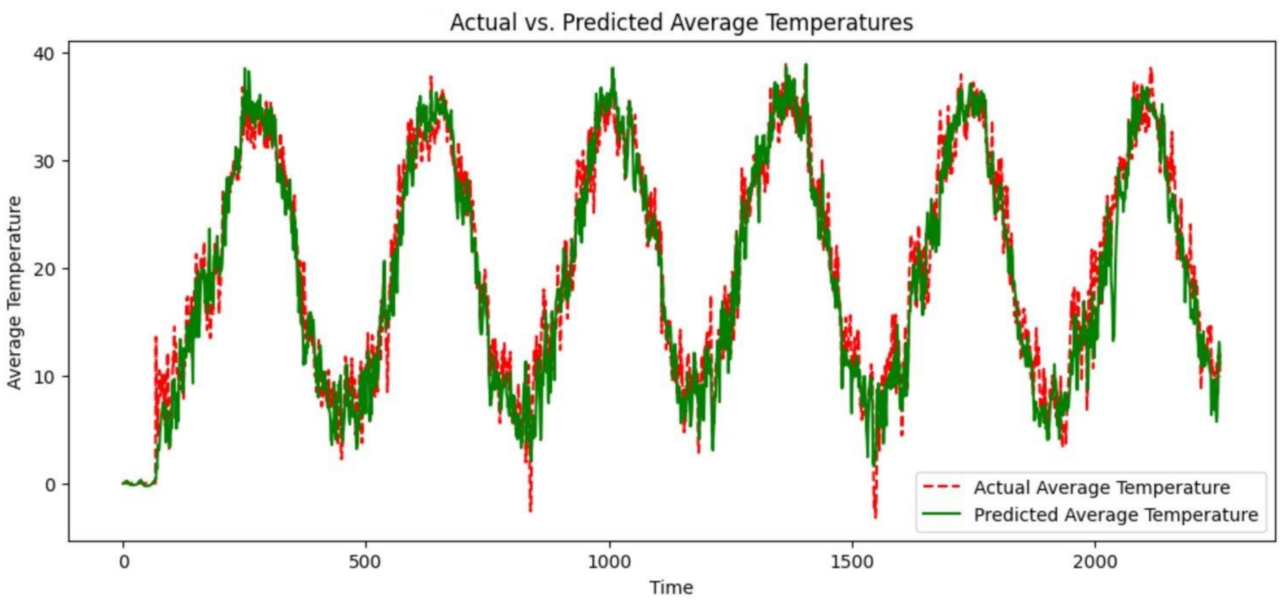

Figure 6. Prediction of The Bi-LSTM model

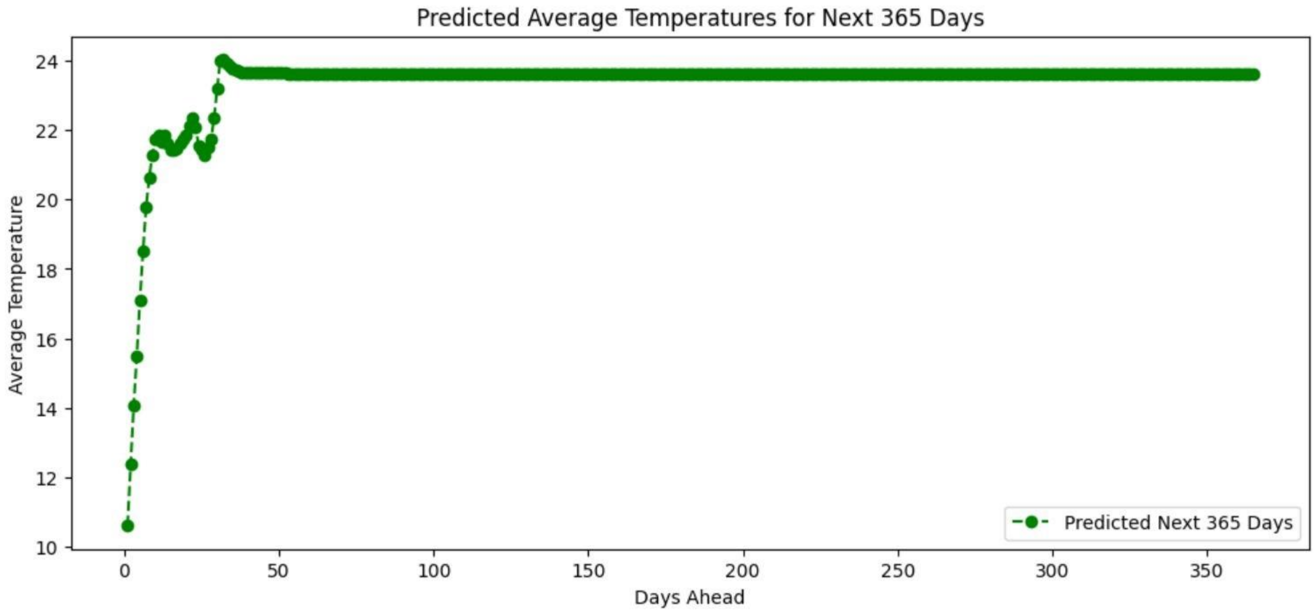

Figure 7. Bi-LSTM result model

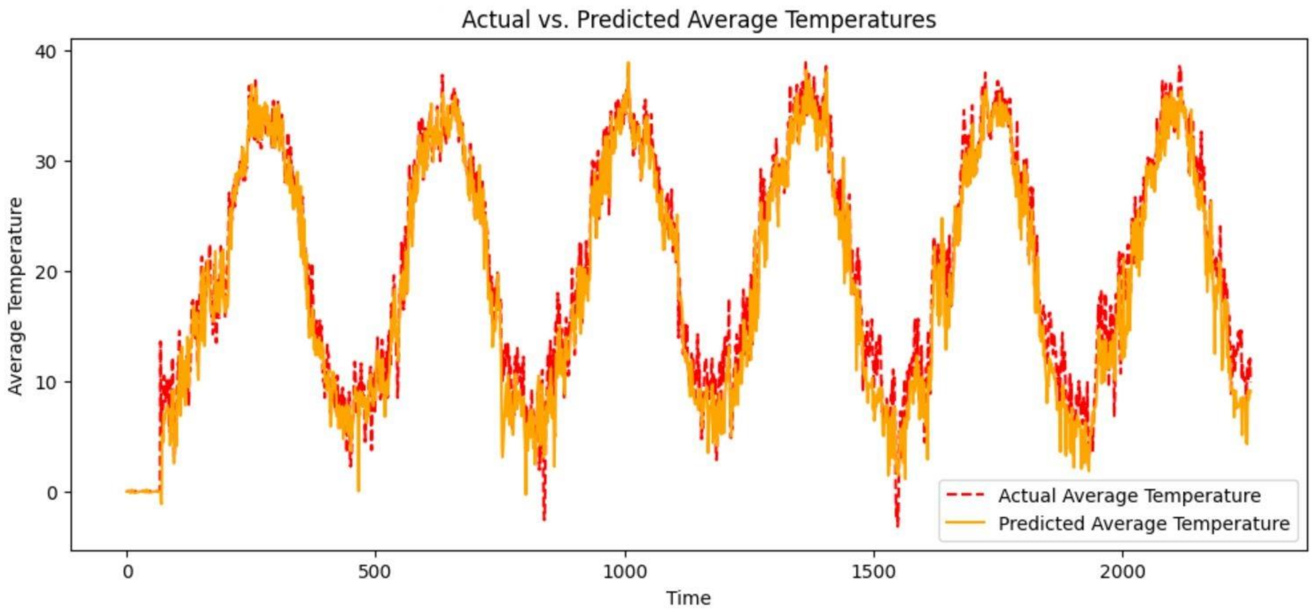

Figure 8. Prediction of The GRU model

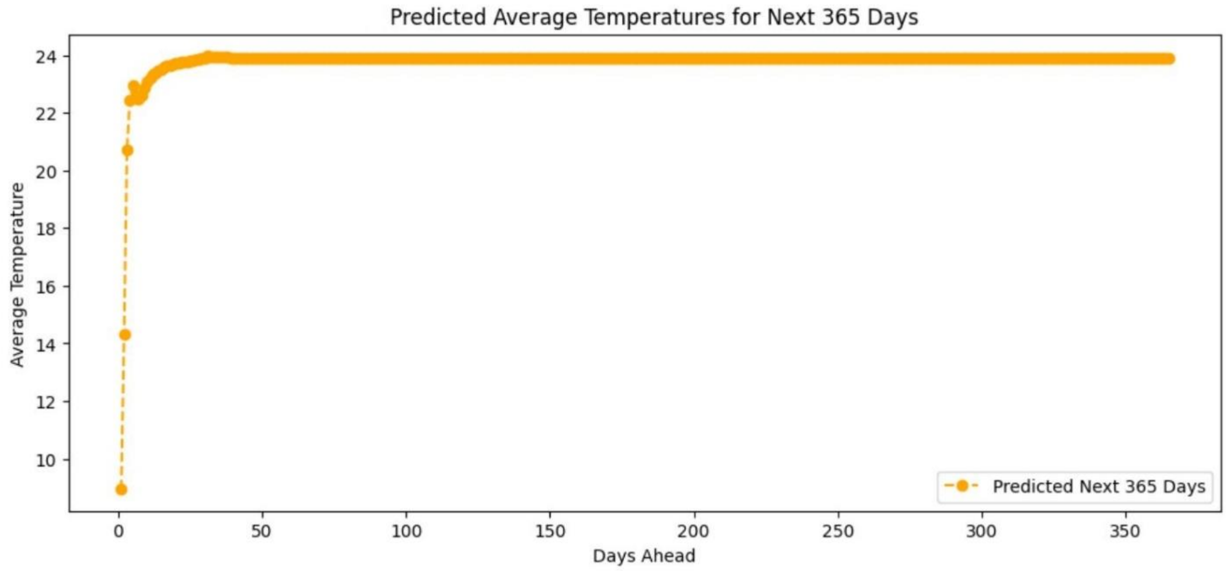

Figure 9. GRU result model.

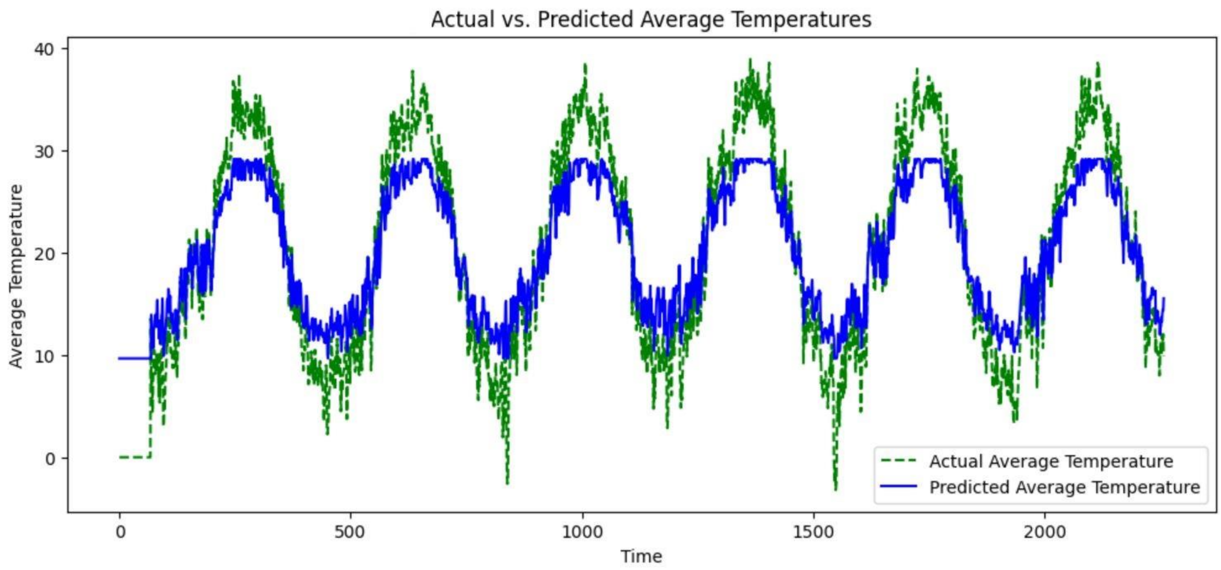

Figure 10. Prediction of the GBT model

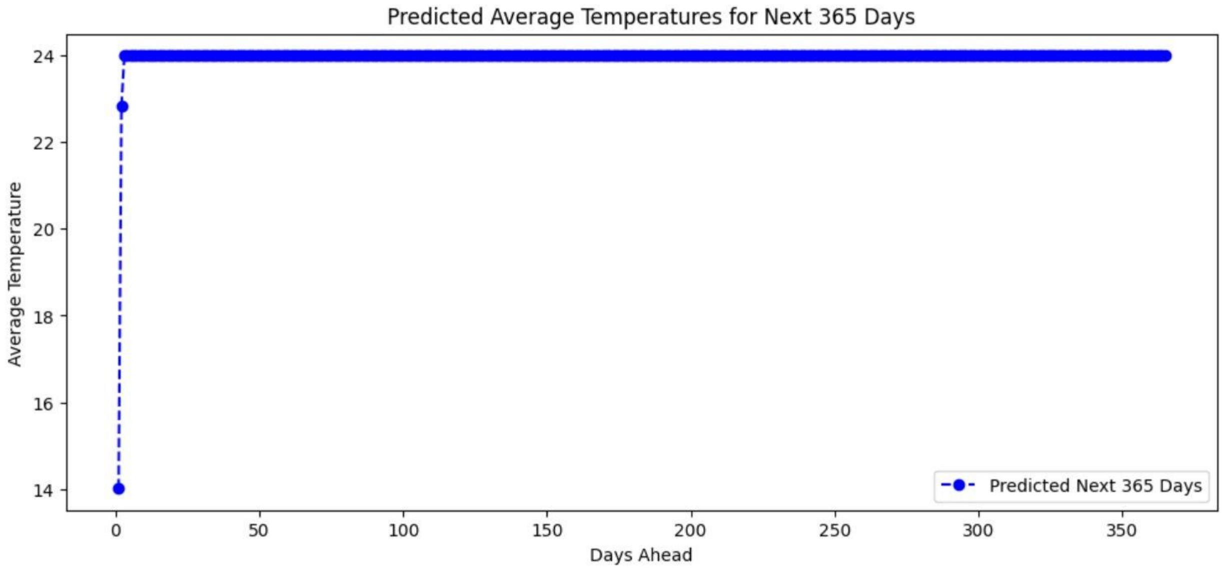

Figure 11. GBT result model

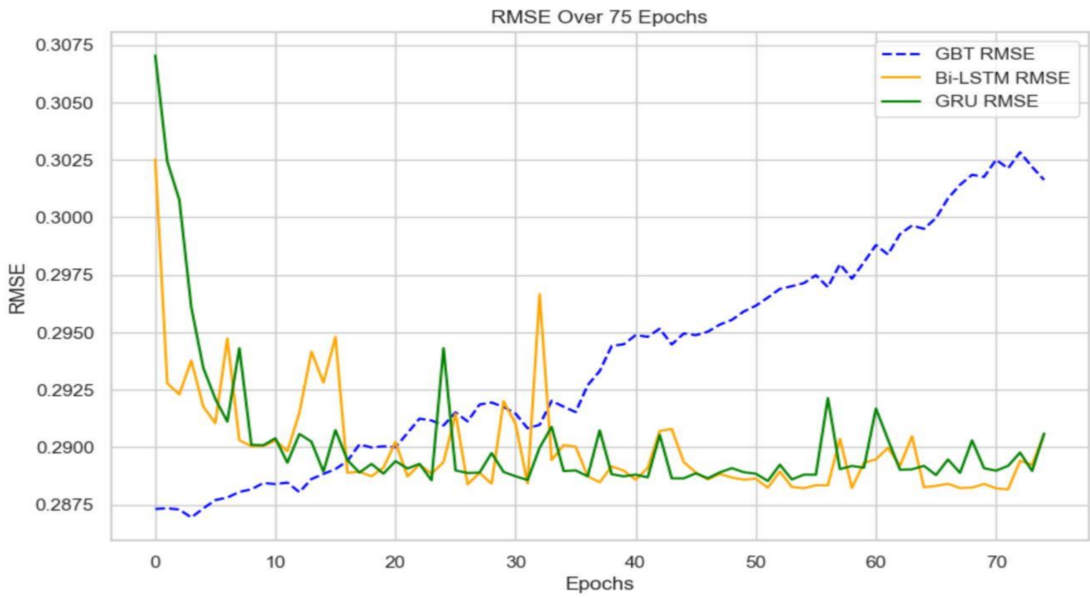

Figure 12. The RMSE for all models

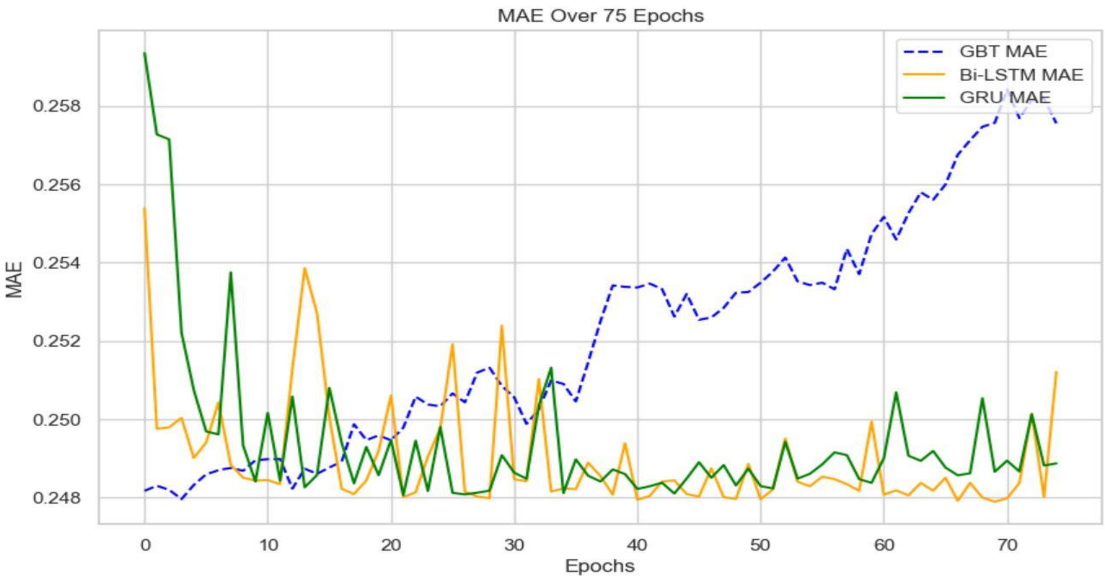

Figure 13. The MAE for all models

Figure 14. The R^2 for all models

Table 4: The performance models

Model	Train			Test		
	RMSE	MAE	R^2	RMSE	MAE	R^2
Bi-LSTM	1.3073	1.0044	0.9833	2.5334	1.9237	0.9399
GBT	4.2577	3.5744	0.8230	**4.3145**	3.6265	0.8257
GRU	**1.3067**	**0.97701**	**0.9833**	2.4232	**1.8293**	**0.9450**

Figure 15. Visually represented modules

Table 5: Comparison with relevant research and the current state of the art.

Reference	Dataset(s)	Modules	Results
[26]	Meteorological stationin Beijing, parameters temperature and humidity.	CONV-BI-LSTM and LSTM	LSTM's RMSE (RH) = 15.32, CONV-BI-LSTM's RMSE (T) = 2.47, and RMSE(RH) = 14.11 3.21 is the LSTM's RMSE (T).
[27]	NCHM (National center Hydrology and Meteorology)	CNN, LSTM, GRU and BLSTM	MSE score for the BLSTM-GRU model is 0.0075. The 1-D CNN, LSTM, GRU, and BLSTM MSE scores are 0.013.
[28]	The only purpose of the Scientific Python Development Environment (spyder) notebook is to handle and prepare data.	ONV1D-MLP and SVR modules	R = 0.25 to 0.55, RMSE = 6.6 to 24.19, and NSE = 0.28 to 0.11.
[29]	https://www.accuweather.com/	ANN, KNN and SVM.	R-squared value = 0.948. mean average percentage error =14.348. compared to SVM and ANN, KNN yields a more accurate output.
Proposed	Forecast the weather in Sulaymaniyah, a city in the Kurdistan Region/Iraq.	BI-LSTM, GRU and GBT	RMSE MAE R^2 Bi-LSTM 1.3073 1.0044 0.9833 GBT 4.2577 3.5744 0.8230 GRU 1.3067 0.97701 0.9833