Prediction of Average Land Temperature using Bi-directional LSTM

Aryan Debray (2105020)
Navonil Ghosh (2105041)
Roudrak Saha (2105058)
Amiya Kumari (2105100)
Diyanshi Sharma(2105456)

Abstract

Predicting average land temperature is crucial for climate change research and adaptation planning. Linear regression, a simple but powerful machine learning algorithm, can be used for this purpose. This paper investigates the use of linear regression to predict average land temperature using average land temperature uncertainty, maximum land temperature, maximum land temperature uncertainty, minimum land temperature, minimum land temperature uncertainty, average ocean and land temperature and average ocean and land temperature uncertainty.

A Bi-directional LSTM model was trained and evaluated on a dataset by Berkeley Earth which is affiliated with Lawrence Berkeley National Laboratory. The Berkeley Earth Surface Temperature Study combines 1.6 billion temperature reports from 16 pre-existing archives.. The model achieved an R-squared value of 0.9997 on the test set, indicating accurate average land temperature predictions. The importance of each predictor variable in the model was also analyzed.

This research illustrates the proficiency of neural networks in anticipating average land temperature. The model developed here has the capacity to provide projections of future land temperature across diverse climate change scenarios. These projections offer valuable insights for shaping strategies and decisions related to climate change adaptation with a remarkable level of accuracy.

1. Introduction

Big data analytics is revolutionizing climate science and environmental research. This project demonstrates the power of big data analytics in harnessing climate data to predict average land temperature using Recurrent Neural Networks (RNN).

Weather forecasting plays a crucial role in various aspects of our lives, including safety and preparedness, agriculture and farming, transportation, energy management, natural resource management, healthcare, emergency response, construction and infrastructure planning, tourism and recreation, scientific research, economic decision-making, and environmental protection.

RNNSs are a category of artificial neural networks tailored for handling sequential data. Their specialized design enables them to grasp connections between various elements of a sequence, even when these elements are widely spaced. Bidirectional Long Short-Term Memory (BiLSTM) is a variant of RNNs capable of analyzing sequential data in both forward and backward directions. This bidirectional processing capability empowers BiLSTM models to better capture extensive dependencies in data compared to unidirectional LSTM models, which are restricted to processing data in one direction.

2. Literature Review

This literature review examines existing studies in meteorology, highlighting the role of forecasting in decision-making. The goal is to assess the current state of knowledge, identify gaps, study results achieved from different models and set the stage for the research's distinctive contribution.

In 2019[1], the research conducted by Singh N. introduced a weather forecasting system aimed at addressing the specific needs of remote areas. The primary focus of the study was to develop an innovative solution utilizing data analytics and machine learning algorithms, including random forest classification, for accurate weather predictions. The research paper outlines the creation of a cost-effective and portable system dedicated to weather forecasting.

In a paper authored by Jayasingh in 2022,[2], various machine learning models are introduced, employing historical data for training and subsequently predicting weather with superior accuracy compared to conventional methods. The assessment of these models based on accuracy demonstrates their superiority, positioning them as advanced techniques for more efficient and timely weather predictions.

Holmstrom's paper from 2016[3] focused on predicting maximum and minimum temperatures over a seven-day period using past two days' weather data. The study utilized a linear regression model and a modified functional regression model capable of capturing weather trends. Despite both models being outperformed by professional weather forecasting services, the performance gap narrowed for later-day forecasts. There is a suggestion that, over longer time scales, the models in the study could potentially outperform professional forecasts. Significantly, the linear regression model outperformed the functional regression model, indicating that the two-day data window might be insufficient for the latter to capture significant weather trends. The study implies that extending the forecast window to four or five days could enable the functional regression model to surpass the linear regression model.

In this paper(2015)[4], the Narvekar. M. review various machine learning techniques that have been used for weather forecasting. They conclude that artificial neural networks with backpropagation perform the best, with minimal error. Neural networks are able to learn patterns in the data and make predictions based on those patterns. The authors suggest that using 28 input parameters, such as temperature, rainfall, humidity, cloud condition, and weather of the day, can improve the accuracy of the forecasts.

T. Anjali's approach to temperature prediction, as proposed in the research conducted in 2019[5], involved the utilization of three machine learning models: Multiple Linear Regression (MLR), Artificial Neural Network (ANN), and Support Vector Machine (SVM). The study encompassed a comparative analysis using weather data collected from Central Kerala over the period 2007 to 2015. The evaluation of experimental results employed metrics such as Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Correlation Coefficients (CC). Notably, the assessment of these metrics and CC collectively affirmed that MLR emerged as a more precise model for temperature prediction when contrasted with ANN and SVM.

Paras et al. present a straightforward model for weather prediction through regression analysis [6]. The study involves recording time-series weather data at a specific station, predicting parameters such as maximum temperature, minimum temperature, and relative humidity based on correlation values in the weather data series across various periods. Additionally, relative humidity is forecasted using a time series of maximum and minimum temperature and rainfall. Rainfall category is estimated by considering features of maximum and minimum temperature, along with relative humidity. In a related study, Teresa Jacobson et al. conduct an illustrative investigation using linear regression on seasonal weather patterns to enhance the learning process [7].

In 2019[8], Bin Wang and collaborators devised an end-to-end deep learning framework for weather forecasting. Their method, which integrated historical data and Numerical Weather Prediction (NWP) knowledge, introduced a unique negative log-likelihood error (NLE) loss function for simultaneous forecasting and uncertainty quantification. Evaluating the approach on a Beijing weather dataset, they demonstrated a 47.76% increase in accuracy, surpassing traditional metrics and NWP and setting a new benchmark.

In 2018[9], Scher Sebastian and team explored machine learning for predicting weather forecast uncertainty using large-scale atmospheric conditions. Their approach, utilizing artificial convolutional neural networks trained on past forecasts, efficiently assigned confidence values to medium-range forecasts. While less accurate than ensemble models, this method outperformed alternatives without numerical forecasts, with the main challenge being the availability of sufficient past forecast data for training.

In 2022[10], de Castro and the team developed machine learning models for forecasting severe convective weather near São Paulo airports, focusing on atmospheric discharge (AD) data. Using classical thermodynamic indices and AD information from 2001 to 2017, the models achieved high accuracy in predicting convective events and their severity, with 96.7% accuracy in occurrence and 86.7% for severe events over a 30-day hindcast. The study discussed detection errors and emphasized the model's effectiveness.

3. Proposed Methodology

This proposed methodology outlines the steps for time series prediction using a Bidirectional LSTM model, including data loading, preprocessing, feature scaling, model building, training, and evaluation.

3.1. Loading and Understanding Data:

- Import the necessary libraries, including pandas for data manipulation.
- Load the dataset and display the first few rows.

3.2. Data Preprocessing:

- Check for missing values and handle them appropriately.
- Drop rows with missing values.

3.3. Feature Scaling:

- Utilize StandardScaler to standardize the features.

3.4. Train-Test Split:

- Split the data into training and testing sets using train_test_split.

3.5. Reshape Data for LSTM:

- Reshape the training data to be compatible with LSTM input shape.

3.6. Build and Compile Bidirectional LSTM Model:

- Create a sequential model using Keras with Bidirectional LSTM layers and a Dense layer for output.

- Compile the model with Mean Squared Error (MSE) loss and Adam optimizer.

3.7. Train the Model:

- Train the Bidirectional LSTM model using the training data.

3.8. Prepare and Evaluate Test Data:

- Reshape the test data.

- Make predictions using the trained model and evaluate performance metrics.

3.9. Evaluate Model Performance:

- Utilize metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared.

4. Working Model

Initially, the data was preprocessed by dropping the null values. This is done to remove any incomplete or missing data from the dataset. Next, the data was scaled and reshaped from 2D to 3D. This is done to make the data compatible with the Nested Bidirectional LSTM model, which requires 3D input data.

The functioning of the model commences with the input data, a sequence of temperature readings, undergoing processing through the initial bidirectional LSTM layer. This layer is responsible for understanding long-range dependencies within the data and extracting relevant features. The extracted features are subsequently forwarded to the second bidirectional LSTM layer, which further refines and comprehends intricate patterns in the data. The output from this layer is then directed to a dropout layer, which selectively deactivates some neurons to prevent overfitting. The resulting output from the dropout layer is fed into the dense layer, responsible for predicting both high and low temperatures for the subsequent day.

During training, the model aims to minimize the Mean Squared Error (MSE) loss function. This loss function quantifies the disparity between the predicted prices and the actual prices. The Adam optimizer is employed to iteratively adjust the weights of the model, facilitating the minimization of the MSE loss function and enhancing the model's predictive accuracy.

5. Results.

The model developed predicts the average land average temperature using Bi-directional LSTM giving promising accuracy metric results-

Mean Absolute Error (MAE) : 0.043614865938822425

Mean Square Error (MSE) : 0.0030596765871330383

Root Mean Square Error (RMSE) : 0.055314343412292605

R^2: 0.9998332254224851

The accuracy of a model that predicts the average land temperature using Linear Regression, Support Vector Machines (SVMs) and Decision Trees in comparison to the model using Bi-directional LSTM has also been shown-

Model	Model Used	MAE	MSE	RMSE	R^2
Our Model	Bi-directional LSTM	0.043	0.0030	0.055	0.9998
Model 1	Linear Regression	0.15	0.07	0.26	0.996
Model 2	Support Vector Machine	0.57	0.25	0.53	0.91
Model 3	Decision Tree	1.34	1.58	1.92	0.87

Table 1

The Bi-directional LSTM model outperforms others in predicting land temperatures as shown in Fig.1, excelling in all measured aspects. BiLSTMs are better suited for weather prediction than traditional LSTM models because they can learn long-range dependencies in data from both the past and the future, and they can process data in both the forward and backward directions. This allows BiLSTMs to learn more informative representations of the data and to make more accurate predictions. This makes it the top choice for temperature forecasts.

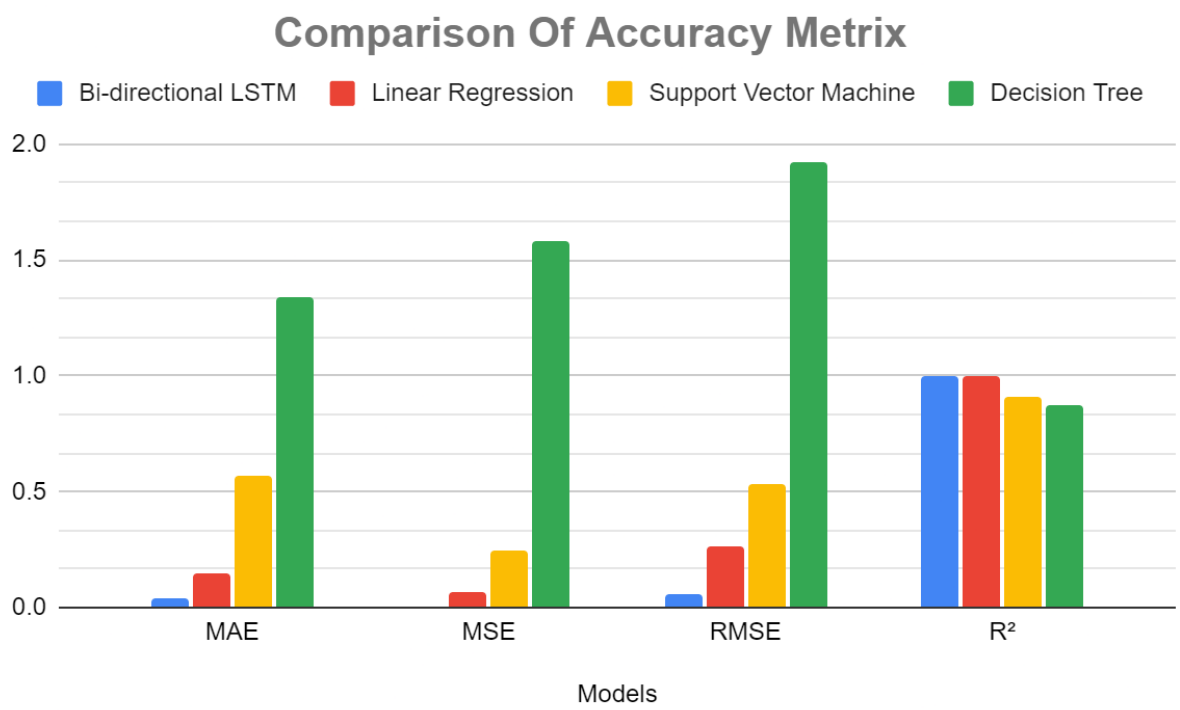

Figure 1. Comparison of accuracy metrics between models implemented using different algorithms.

6. Future Scope and Scalability

The utilization of bidirectional LSTM for temperature prediction offers several advantages, such as its potential extension to predict various weather variables, diverse time and spatial scales, and the incorporation of additional data sources and advanced machine learning techniques. To enhance the paper's future prospects, an exploration of bidirectional LSTM's efficacy in predicting temperature across different climatic zones could provide a more comprehensive understanding of the model's performance in varied environments, potentially leading to more accurate predictions. Comparing the performance of bidirectional LSTM with alternative machine learning models for temperature prediction would offer valuable insights into the strengths and weaknesses of different approaches.

Improving the scalability of the model is crucial, and one approach is the development of parallel training algorithms, enabling quicker training on large datasets. This enhancement facilitates real-time temperature prediction, expanding the model's applicability. Additionally, quantizing the model can reduce its size and memory requirements, rendering it more deployable on resource-constrained devices like mobile devices.

The model presently can be used to predict the very next day's temperature. But given the extremely high accuracy of the model, the same model can be used to predict the temperature after a week or few weeks as well. This can be done by predicting the next day's temperature and using it to predict the temperature on the day after. This can be done till the temperature is found on a desirable date.

In summary, a Bidirectional LSTM model holds significant potential for future advancements and scalability. The model's versatility allows for extensions and improvements across various dimensions, making it applicable to diverse domains such as weather forecasting, climate change research, and energy management.

7. Conclusion

In conclusion, this paper presents a robust and accurate approach to predicting average land temperature using a Bidirectional LSTM model, showcasing its superiority over traditional machine learning algorithms such as Linear Regression, Support Vector Machines, and Decision Trees. The Bidirectional LSTM model, trained and evaluated on a comprehensive dataset from Berkeley Earth, achieved exceptional accuracy with an R-squared value of 0.9998 on the test set. The comparison of accuracy metrics across different models solidifies the Bidirectional LSTM's efficacy in temperature prediction.

The results indicate that the model exhibits outstanding performance, surpassing other machine learning models in metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared. The paper addresses the scalability of the model, proposing potential expansions to predict diverse weather variables, explore different climatic zones, and integrate additional data sources. To enhance scalability and deployment on resource-constrained devices, the paper recommends the development of parallel training algorithms and model quantization.

In summary, the Bidirectional LSTM model showcases substantial potential for advancing weather prediction and scalability, positioning it as a valuable asset for diverse applications in climate science and environmental research.

8. References

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

[2] Jayasingh, S.K., Mantri, J.K., Pradhan, S. (2022). Smart Weather Prediction Using Machine Learning.

[3] Machine Learning Applied to Weather Forecasting Authors: Mark Holmstrom, Dylan Liu, Christopher Vo Institution: Stanford University Date: December 2016

[4] M. Narvekar and P. Fargose, "Daily Weather Forecasting using Artificial Neural Network," in Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1713-1717, 2014.

[5] T. Anjali, K. Chandini, K. Anoop and V. L. Lajish, "Temperature Prediction using Machine Learning Approaches," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur, India, 2019, pp. 1264-1268, doi: 10.1109/ICICICT46008.2019.8993316.

[6] Paras, Sanjay Mathur. "A simple weather forecasting model using mathematical regression." Indian Research Journal of Extension

[7] Jacobson, Teresa, Josh James, and Neil C. Schwertman. "An example of using linear regression of seasonal weather patterns to enhance undergraduate learning." Journal of Statistics Education 17, no. 2 (2009).

[8] Jin X, Wang J, Guo S, Wei T, Zhao Y, Lin Y and Wan H. (2024). Spatial–temporal uncertainty-aware graph networks for promoting accuracy and reliability of traffic forecasting. Expert Systems with Applications. 10.1016/j.eswa.2023.122143. 238. (122143). Online publication date: 1-Mar-2024.

[9] cher, S, Messori G. Predicting weather forecast uncertainty with machine learning. Q J R Meteorol Soc. 2018; 144: 2830–2841.

[10] de Castro, J.N., França, G.B., de Almeida, V.A. et al. Severe Convective Weather Forecast Using Machine Learning Models. Pure Appl. Geophys. 179, 2945–2955 (2022).