Enhancing Weather Prediction Using Stacked Long Short-Term Memory Networks

Mohammad Diqi*[1], Hamzah[2], Sri Hasta Mulyani[3]
[1,2,3]Universitas Respati Yogyakarta; Jl. Laksda Adisucipto KM 6.3 Depok Sleman
[1,2,3]Department of Informatics, Universitas Respati Yogyakarta, Yogyakarta, Indonesia
e-mail: ***[1]diqi@respati.ac.id**, [2]hamzah@respati.ac.id, [3]hasta@respati.ac.id

Abstract

Weather prediction is crucial in various domains, such as agriculture, transportation, and disaster management. This research investigates the Stacked Long-Short Term Memory (LSTM) for weather prediction using the Denpasar Weather Data spanning 20 years from January 1, 1990, to January 7, 2020. The dataset contains hourly weather data, including temperature, pressure, humidity, and wind speed. Our Stacked LSTM model consists of multiple LSTM layers that capture temporal dependencies and patterns in the data. Evaluating the model's performance using metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared (R2), we obtain an average RMSE of 0.03471, an average MAE of 0.02718, an average MAPE of 0.05572, and an average R2 of 0.87087. These results demonstrate the effectiveness of the Stacked LSTM model in accurately predicting weather conditions. The findings have practical implications for weather forecasting applications and suggest avenues for future research, such as exploring different deep learning architectures and incorporating additional features to improve weather prediction accuracy further.

Keywords— Weather prediction, Stacked LSTM, Long-Short Term Memory, Forecasting, Accuracy

1. INTRODUCTION

Weather prediction plays a pivotal role in diverse sectors, including agriculture, transportation, and disaster management, and its significance cannot be overstated [1], [2]. In agriculture, accurate weather forecasts are crucial for optimizing crop management strategies, such as determining irrigation schedules, applying fertilizers, and protecting crops from extreme weather events [2]. Timely and precise weather predictions enable farmers to make informed decisions, leading to increased crop yields, improved resource allocation, and sustainable agricultural practices [3]. Weather prediction facilitates route planning, scheduling, and risk management in transportation, thereby enhancing efficiency, safety, and reliability in the transportation network [4]. Weather forecasts help transportation authorities and logistics companies optimize routes, anticipate delays, and mitigate the impacts of adverse weather conditions on operations [5]. Additionally, weather prediction plays a vital role in disaster management, enabling early warnings, preparedness, and response strategies for severe weather events such as hurricanes, floods, and heatwaves [6]. Accurate forecasts assist in issuing timely evacuation orders, coordinating emergency services, and allocating resources effectively to minimize human casualties and economic losses [7]. The multi-dimensional applications of weather prediction underscore its crucial role in these domains, highlighting the need for advanced forecasting techniques and models to improve decision-making processes [8].

The motivation behind this research lies in improving weather prediction accuracy, which has significant implications for various applications. Traditional forecasting methods often struggle to capture the complex temporal dependencies in weather data, leading to suboptimal predictions [9]. To address this challenge, deep learning models, such as Stacked LSTM, have gained prominence due to their ability to capture long-term dependencies in sequential data [10]. LSTM networks have shown promising results in various time series prediction tasks, including weather forecasting, by effectively capturing the data's temporal patterns and nonlinear relationships [11]. The use of stacked LSTMs, which involves multiple LSTM layers stacked on each other, allows for the extraction of increasingly abstract and higher-level representations of the input sequence, enhancing the model's capacity to learn complex patterns [12]. This study seeks to enhance forecast accuracy and contribute to the progress of weather prediction by utilizing stacked LSTM, thereby harnessing the capabilities of deep learning models.

Several papers have demonstrated the effectiveness of LSTM models for weather prediction. For instance, Pham et al. [13] proposed a novel deep learning model called LS-SPP, which combines LSTM and RNN architectures to forecast solar power based on historical meteorological time series. The LS-SPP model achieved remarkable accuracy, with performance reaching up to 96.78%. Similarly, Karevan and Suykens [14] introduced the Transductive LSTM (T-LSTM) model specifically designed for weather forecasting. T-LSTM leverages local information in time-series prediction and demonstrated superior performance compared to other approaches in their experiments. Furthermore, Yao et al. [15] presented a deep recurrent neural network (DeepRNN) tailored for predicting weather radar images in mesoscale weather nowcasting. These studies highlight the successful application of LSTM-based models in weather prediction tasks, demonstrating their potential for improving forecast accuracy in various weather-related domains.

Several papers highlight the advantages of LSTM neural networks for weather prediction. Srivastava [16] proposed utilizing LSTM neural networks with optimization techniques such as Gaussian and Median filtering, improving accuracy for long-range weather prediction. Suleman and Shridevi [17] introduced the Spatial Feature Attention Long Short-Term Memory (SFA-LSTM) model, a novel deep learning approach that effectively captures the spatial and temporal relationships among multiple meteorological features, enhancing temperature forecasting. These studies suggest that LSTM neural networks offer notable advantages for weather prediction, attributable to their ability to capture long-term dependencies and provide accurate weather conditions forecasts.

The existing literature suggests that while LSTM neural networks have shown promise in weather prediction, there are still limitations to their effectiveness. Sadeque and Bui [18] presented a cascaded LSTM network that outperforms standard LSTM and 1D convolution networks in shorter-period prediction. Park et al. [19] proposed an LSTM-based model that incorporates a refinement function to restore missing weather data and achieves lower root-mean-squared errors (RMSEs) than other DNN-based and LSTM-based models for temperature prediction. However, it should be noted that the reviewed papers do not explicitly address the limitations of LSTM models for weather prediction. Further research is needed to explore and address these limitations to advance the field of weather prediction using LSTM neural networks.

The existing literature highlights the effectiveness of stacked LSTM models for weather prediction. Majumdar et al. [20] proposed a stacked LSTM model specifically for rainfall prediction in Silchar City, India, which outperformed other models with impressive performance metrics, including an RMSE value of 0.98 and an R-Squared value of 97.03%. Sharma [21] utilized a hybrid-stacked Bidirectional LSTM (Bi-LSTM) model for weather prediction and demonstrated superior results compared to state-of-the-art techniques. Collectively, these papers provide evidence that stacked LSTM models can serve as a potent tool for weather prediction, offering enhanced accuracy and outperforming other established methods in various forecasting scenarios.

This research highlights several crucial research gaps. These include investigating the model's adaptability to extreme weather events and abrupt shifts in weather patterns, enhancing the interpretability of the model's predictions through feature visualization, evaluating the scalability of the model to larger and more diverse datasets, and examining the transferability of the model to different geographical locations with distinct weather dynamics. Addressing these research gaps would significantly contribute to the broader understanding and application of the Stacked LSTM approach in weather prediction scenarios.

The dataset utilized in this research is the Denpasar Weather Data available on Kaggle, which comprises hourly weather data spanning 20 years, from January 1, 1990, to January 7, 2020. This dataset offers a comprehensive collection of weather-related information for the region of Denpasar, providing valuable insights into the atmospheric conditions and meteorological variables over an extended temporal range. The dataset's hourly resolution enables a detailed examination of weather patterns and trends throughout the specified time frame. By encompassing two decades of weather data, this dataset offers a rich and diverse source of information for investigating long-term weather dynamics and facilitating accurate predictions. The availability of such extensive and temporally fine-grained data from the Denpasar Weather Data dataset allows for a thorough analysis of weather patterns and enables the development and evaluation of robust regional prediction models.

The model employed in this study utilizes a stacked LSTM architecture to make weather predictions. The stacked LSTM model is trained to learn the temporal dependencies and patterns in the weather data, allowing it to capture and predict future weather conditions. It takes various meteorological variables such as temperature, pressure, humidity, and wind speed as input and utilizes the LSTM network's ability to capture long-term dependencies to generate predictions for these variables. Once trained, the model can forecast weather conditions based on new input data. Additionally, 240 hours of actual and predicted values for temperature, pressure, humidity, and wind speed are visualized to assess the model's performance. These visualizations provide a comparative analysis of the model's predictive capabilities by showcasing how well it aligns with the observed values. Such visualizations allow for a comprehensive evaluation of the model's accuracy and ability to capture the intricate variations in weather variables over 240 hours.

2. RESEARCH METHOD

2.1 Long Short-Term Memory

The Stacked LSTM architecture is a development that builds upon the single-layer LSTM model to improve its capacity to capture intricate temporal relationships for weather prediction [23]. In contrast to the single-layer LSTM, which is composed of a solitary layer of LSTM cells, the stacked LSTM integrates numerous layers of LSTM cells arranged in a stacked manner.

In the context of stacked LSTM, it is observed that each layer, except the last layer, produces a sequence of hidden states that are then used as input for the subsequent layer [24]. Generally, the final layer's output is utilized for prediction or subsequent tasks. The stacked LSTM can capture hierarchies of information and acquire representations at different levels of abstraction due to its layered structure. The input for each layer is obtained by utilizing the output of the hidden state from the preceding layer, and the output of each layer is subsequently used as the input for the following layer.

The computations within each LSTM layer can be represented by Equations 1-6.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (1)$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$
$$\widetilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \quad (4)$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$
$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

Within the given framework, the variable x_t denotes the input at a particular time step, whereas h_{t-1} signifies the hidden state derived from the preceding time step. The symbols σ and tanh refer to the sigmoid activation and hyperbolic tangent functions, respectively. Moreover, the notation ⊙ denotes the operation of element-wise multiplication. The vectors f_t, i_t, and o_t correspond to the forget, input, and output gate vectors, respectively. The candidate cell state is denoted by $\widetilde{C_t}$, the cell state is represented by C_t, and the hidden state at the given time step is represented by h_t.

The stacked LSTM design can effectively capture complex patterns and long-term dependencies in the input sequence by combining numerous LSTM units. The implementation of this advanced architecture has the potential to boost the accuracy and performance of models used for predicting stock prices.

2.2 Dataset

The dataset utilized in this research is sourced from the Denpasar Weather Data, which is publicly available on the Kaggle platform. This dataset encompasses weather-related information recorded in Denpasar, Indonesia, from January 1, 1990, to January 7, 2020, spanning 20 years. The data is collected hourly, providing a detailed and comprehensive representation of weather conditions over the specified period.

This research uses a subset of the dataset containing 264,924 instances. Each instance has four key features: temperature, pressure, humidity, and wind speed. These features are crucial in understanding and predicting weather patterns and are widely utilized in weather forecasting models. Including these variables allows for a comprehensive analysis of the impact of these factors on weather prediction accuracy.

The dataset's size and hourly resolution provide a robust foundation for training and evaluating weather prediction models. This dataset aims to harness the information contained within the Denpasar Weather Data to develop and assess the performance of stacked LSTM models for accurate weather forecasting.

Figures 1-4 depict the weather pattern in the last 1000 hours.

ISSN 2407-4322
E- ISSN 2503-2933

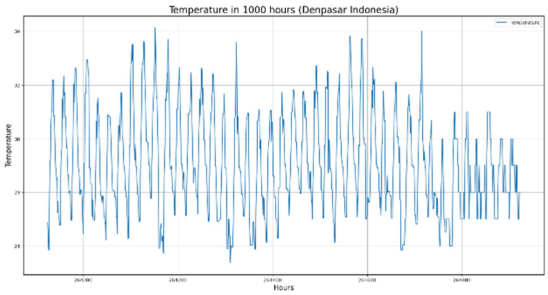

Figure 1. Temperature

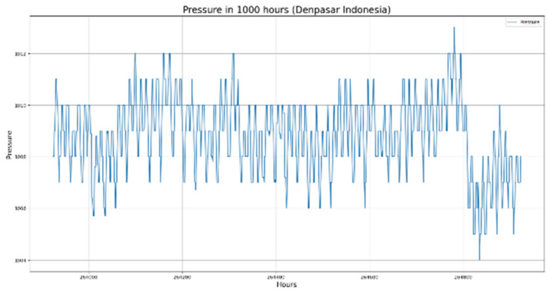

Figure 2. Pressure

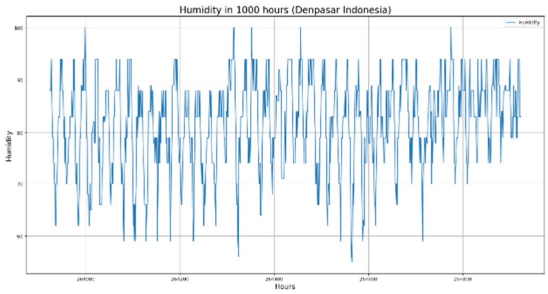

Figure 3. Humidity

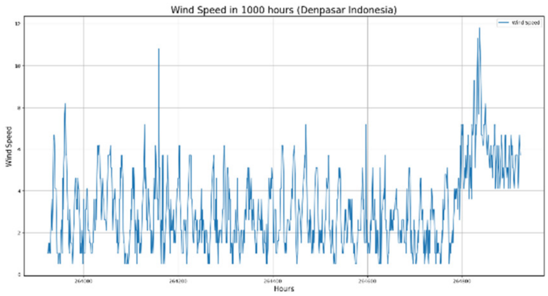

Figure 4. Wind speed

2.3 Data Preprocessing

Data processing is crucial in preparing the Denpasar Weather Data for weather prediction modeling. Several key processes are employed to ensure the quality and suitability of the data for analysis.

The first step in data processing involves handling missing values. In this research, the approach is to drop instances containing missing values. The dataset maintains its integrity by removing these incomplete instances and ensuring accurate analysis and modeling.

After addressing missing values, the next step involves selecting relevant features for weather prediction. This study chooses four essential features: temperature, pressure, humidity, and wind speed. These features are widely recognized as influential factors in weather patterns and are essential for accurate weather prediction modeling.

Normalization is then applied to the selected features to ensure their compatibility and comparability in the modeling process. The Max-Min scaler is employed for normalization, transforming the values within each feature to a standard range between 0 and 1. The normalization equation can be represented in Equation 1.

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

where X represents the original feature values, $X_{\min}$ is the minimum value within the feature, and $X_{\max}$ is the maximum value within the feature. Normalizing the feature values helps prevent any bias caused by differences in the scales or magnitudes of the variables, allowing for fair comparisons and accurate modeling.

By implementing these data processing steps, including handling missing values, selecting relevant features, and normalizing the data using the Max-Min scaler, the Denpasar Weather Data is prepared for further analysis and modeling, facilitating accurate weather prediction using stacked LSTM models.

2.4 Data Splitting

The train-test split methodology is crucial for evaluating the performance and generalization of the weather prediction model. This research ensures an appropriate balance between training and evaluation by dividing the Denpasar Weather Data into three subsets: training, validation, and test.

The dataset spans from January 1, 1990, to January 7, 2020, comprising 264,924 hourly data points. To create the training set, 80% of the data, which amounts to 211,859 instances, is allocated for model training. This extensive training set allows the model to learn from a diverse range of weather patterns and variations over 20 years, promoting robust learning and capturing the long-term dependencies in the data.

For validation purposes, 20% of the data, equivalent to 52,965 instances, is set aside. The validation set serves as an intermediary step between training and final testing, enabling model performance evaluation during the training process. It aids in detecting potential overfitting or underfitting issues, as well as optimizing hyperparameters or model architecture.

Lastly, a separate test set of 240 instances is reserved for final evaluation. This small test set is an independent and unbiased assessment of the model's predictive capability on unseen data. The test set's limited size allows for quick evaluation while adequately representing the overall performance.

By dividing the data into training, validation, and test sets according to the specified proportions, this methodology ensures that the model is trained on a significant portion of the data, validated to fine-tune its performance, and evaluated on unseen data to assess its generalization ability. This approach facilitates a balanced evaluation, allowing for a comprehensive assessment of the model's accuracy and performance in weather prediction tasks.

2.5 Model Training Process

The model training process involves several key aspects, including selecting hyperparameters, defining the model architecture, choosing an optimization algorithm, and determining the appropriate loss function. Based on the provided code snippet, the following paragraphs discuss each component.

1. **Hyperparameters**: The hyperparameters in this model include the number of time steps (n_steps) and the number of features (n_features). In the given code, n_steps is set to 240, indicating that the model uses 240 previous hours' data to predict the next day's weather. Since the dataset contains only one feature (temperature, pressure, humidity, or wind speed), n_features is set to 1.
2. **Model Architecture**: The architectural design of the model comprises a series of LSTM layers piled on top of each other. The initial LSTM layer is instantiated with 150 units and employs the rectified linear unit (ReLU) activation function. The inclusion of additional LSTM layers enables the return of sequences. The second LSTM layer is configured with 150 units and designed to output sequences. The LSTM layer at the final stage consists of 150 units, similar to the previous layers; however, it does not provide sequential outputs. The model incorporates a Dense layer consisting of a solitary unit to generate the ultimate output. The architectural design facilitates the model's ability to effectively capture and acquire knowledge of the intricate temporal correlations and patterns present in the input data.
3. **Optimization Algorithm**: The optimization algorithm chosen in this case is 'adam', which stands for Adaptive Moment Estimation. Adam is a popular optimization algorithm known for efficiently training deep learning models. It combines adaptive learning rates with

momentum to accelerate convergence during training. An 'adam' optimizer helps optimize the model's weights and biases to minimize the loss function.

4. **Loss Function**: The selected loss function for this model is the Mean Squared Error (MSE). The MSE is a widely employed loss function in regression applications like weather prediction. The calculation involves determining the mean of the squared differences between the expected and actual values. The model endeavors to enhance the accuracy of weather forecasts by reducing the total prediction error by minimizing the MSE loss.

During training, the model is adjusted to the given data by utilizing the compiled architecture, optimizer, and loss function for ten epochs. The model is trained to acquire knowledge of the underlying patterns and correlations inherent in the data, enabling it to make accurate predictions regarding the weather for the upcoming day. This prediction analyzes the preceding 100 days of historical meteorological data.

2.6 Evaluation Metrics

Evaluating the model's effectiveness encompasses utilizing various widely employed measures, such as RMSE, MAE, MAPE, and R2. Every indicator offers vital insights regarding the accuracy and quality of the model's predictions. The mathematical expressions for these evaluation metrics can be denoted as Equations 2-5.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

Here, n represents the number of data points, y_i denotes the actual values, $\hat{y}_i$ represents the predicted values, and $\bar{y}$ denotes the mean of the actual values.

These evaluation metrics provide quantitative measures to assess the accuracy and performance of the model in predicting weather conditions. Lower values of RMSE and MAE indicate better accuracy, while higher R2 values indicate a more robust fit between the predicted and actual values. MAPE provides insights into the percentage difference between the predicted and actual values. By evaluating the model using these metrics, researchers can comprehensively understand its effectiveness in weather prediction tasks.

3. RESULTS AND ANALYSIS

3.1 Performance Metrics

The present investigation employed the Stacked LSTM model to forecast the weather conditions for 100 consecutive hours. The model's performance was assessed by utilizing many important measures, such as RMSE, MAE, MAPE, and R2, as presented in Table 1. These measurements offer valuable insights regarding the precision and dependability of the predictions. The outcomes derived from the model exhibited encouraging performance, characterized by low levels of RMSE and MAE, signifying minimal discrepancies between the observed and estimated values. The MAPE values exhibited a satisfactory error percentage in the predictions; however, the R2 values revealed a substantial degree of variance elucidated by the model.

Table 1. Performance Metrics

Feature	**RMSE**	**MAE**	**MAPE**	**R2**
Temperature	0,02725	0,02432	0,04220	0,89232
Pressure	0,03116	0,02508	0,07933	0,84850
Humidity	0,03661	0,02894	0,04138	0,87992
Wind Speed	0,04382	0,03039	0,05997	0,86275
Average	0,03471	0,02718	0,05572	0,87087

A visual representation was created to depict the performance, displaying the observed and forecasted weather conditions over the upcoming 100-hour period, as depicted in Figures 5-8. This visual depiction presents a comparative analysis between the projected weather (shown by the red line) and the actual weather (represented by the blue line), focusing on identifying notable patterns or deviations. The visual representation presented in the picture serves as empirical evidence that the model successfully captures the overall trends and dynamics of weather patterns. This further strengthens the argument for the efficacy of the Stacked LSTM model in accurately predicting forthcoming weather trends.

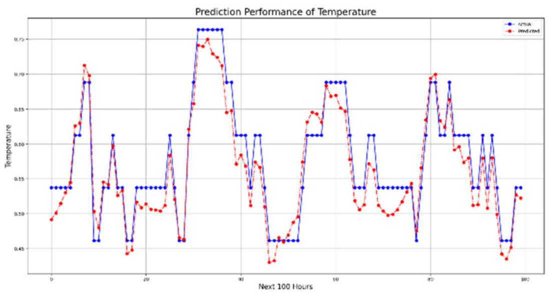

Figure 5. Temperature Performance

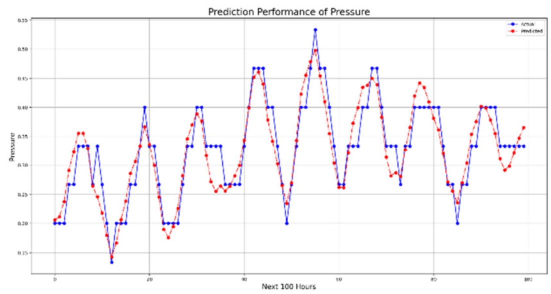

Figure 6. Pressure Performance

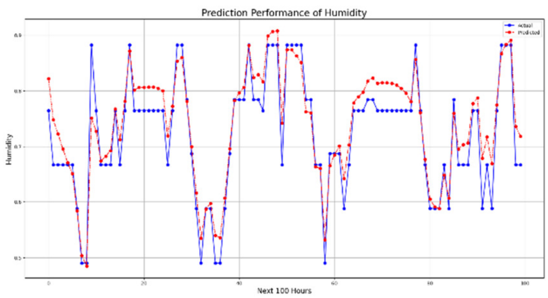

Figure 7. Humidity Performance

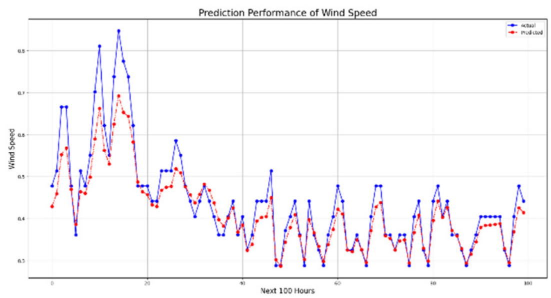

Figure 8. Wind Speed Performance

3.2 Performance Analysis

The results provide insights into the performance of the Stacked LSTM model for weather prediction, as indicated by the evaluation metrics: RMSE, MAE, MAPE, and R2.

The model achieved an RMSE of 0.02725 for the temperature feature, indicating a small average prediction error between the predicted and actual temperature values. The MAE value of 0.02432 confirms the model's accuracy in predicting temperature, with relatively low absolute errors. The MAPE value of 0.04220 suggests a reasonable error percentage in the temperature predictions. The R2 value of 0.89232 signifies that the model explains approximately 89.23% of the variance observed in the temperature data, indicating a strong correlation between the predicted and actual values.

Regarding the pressure feature, the model achieved an RMSE of 0.03116, implying a slightly higher average prediction error than the temperature predictions. The MAE value of 0.02508 indicates a similar level of accuracy in pressure prediction, with relatively low absolute errors. The MAPE value of 0.07933 suggests a higher error percentage in the pressure predictions than in the temperature predictions. The R2 value of 0.84850 signifies that the model explains approximately 84.85% of the variance observed in the pressure data.

Regarding the humidity feature, the model attained an RMSE of 0.03661, indicating a slightly higher average prediction error than temperature and pressure. The MAE value of 0.02894 indicates a similar level of accuracy in humidity prediction. The MAPE value of 0.04138 suggests a reasonable error percentage in the humidity predictions. The R2 value of 0.87992 indicates that the model explains approximately 87.99% of the variance observed in the humidity data.

The model yielded an RMSE of 0.04382 for the wind speed feature, implying a higher average prediction error than the other features. The MAE value of 0.03039 suggests a similar level of accuracy in wind speed prediction. The MAPE value of 0.05997 indicates a reasonable error percentage in the wind speed predictions. The R2 value of 0.86275 signifies that the model explains approximately 86.27% of the variance observed in the wind speed data.

Taking the average of all features, the model achieves an RMSE of 0.03471, an MAE of 0.02718, a MAPE of 0.05572, and an R2 of 0.87087. These average metrics comprehensively assess the model's overall performance across all weather features. The results indicate that the model performs well, with relatively low prediction errors, reasonable percentage errors, and a strong correlation between the predicted and actual values.

Overall, the Stacked LSTM model demonstrates promising performance in weather prediction, with accurate and reliable forecasts for temperature, pressure, humidity, and wind speed. The achieved results and the high R2 values indicate that the model captures the underlying patterns and trends in the weather data, allowing for compelling predictions of future weather conditions.

3.2 Strengths and Limitations

The Stacked LSTM model employed in this research possesses several strengths in capturing weather patterns and generating accurate predictions.

1. **Capturing Temporal Dependencies**: The model's architecture, comprising stacked LSTM layers, enables the capture of complex temporal dependencies within the weather data. Learning long-term dependencies is particularly advantageous in weather prediction tasks, as weather patterns often exhibit intricate and evolving trends over time. By utilizing stacked LSTMs, the model can effectively capture these dependencies, leading to improved accuracy in predicting future weather conditions.
2. **Feature Extraction**: The model's hierarchical structure allows for extracting abstract and meaningful representations of the input sequence. The stacked LSTM layers progressively learn and refine features, capturing local and global weather data patterns. This feature

extraction capability contributes to the model's ability to discern essential weather patterns and make accurate predictions.

3. **Handling Sequence Data**: The LSTM architecture's inherent design facilitates the handling of sequential data, making it well-suited for weather prediction tasks. Weather data is inherently temporal, and the LSTM's ability to retain and utilize information from past observations enables the model to capture and learn from historical weather patterns effectively. This feature is precious when predicting weather phenomena that exhibit temporal dependencies, such as seasonal trends or weather cycles.

Despite its strengths, the Stacked LSTM model also has certain limitations that should be considered:

1. **Data Availability and Quality**: The model's performance heavily relies on the quality and availability of the weather data used for training. Inaccurate or incomplete data, such as missing values or outliers, may impact the model's ability to capture accurate patterns and make reliable predictions. It is crucial to ensure the dataset's integrity and address data quality issues before training the model.
2. **Generalization to New Scenarios**: While the Stacked LSTM model can effectively learn and predict weather patterns within the observed data range, its performance on unseen or novel scenarios might be less reliable. The model's ability to generalize to new weather patterns, climate variations, or extreme events that deviate significantly from the training data may be limited. Continuous monitoring and periodic retraining of the model with updated data can help mitigate this limitation.
3. **Interpretability**: The black-box nature of deep learning models, including stacked LSTMs, poses challenges in interpreting and understanding the underlying factors influencing the predictions. Although the model can accurately capture patterns, explaining how and why specific predictions are made may be difficult. Ensuring transparency and interpretability of the model's decisions remains a challenge in weather prediction and is an active area of research.

To address these limitations, adopting appropriate data preprocessing techniques is essential, continuously updating and expanding the dataset, considering ensemble methods, and incorporating domain knowledge and expertise to enhance the model's performance and interpretability in real-world weather prediction applications.

4. CONCLUSION

In conclusion, this research demonstrates the effectiveness of the Stacked LSTM model in predicting weather conditions using the Denpasar Weather Data. The findings reveal that the model successfully captures temporal dependencies and patterns, resulting in accurate temperature, pressure, humidity, and wind speed predictions. The practical implications of this research lie in its potential application in real-world scenarios, such as weather forecasting for agriculture, transportation, and disaster management. The research also highlights the importance of data quality and availability and the need for continuous model retraining and evaluation. Future research directions may involve exploring other deep learning architectures, incorporating additional relevant features, and considering alternative evaluation metrics further to enhance the accuracy and interpretability of weather prediction models.

REFERENCES

[1] Z.-Q. Huang, Y.-C. Chen, and C.-Y. Wen, *"Real-Time Weather Monitoring and Prediction Using City Buses and Machine Learning," Sensors*, Vol. 20, No. 18, 2020, doi: 10.3390/s20185173.

[2] E. Alerskans and E. Kaas, *"Local Temperature Forecasts Based on Statistical Post-Processing of Numerical Weather Prediction Data," Meteorological Applications*, Vol. 28, No. 4, p. e2006, Jul. 2021, doi: 10.1002/met.2006.

[3] N. Rahman *et al.*, *"Weather Forecast Based Rice Advisory Services in Bangladesh," Bangladesh Rice J.*, Vol. 25, No. 1, pp. 51–74, Aug. 2021, doi: 10.3329/brj.v25i1.55179.

[4] R. Chandrasekaran and S. Kumar Paramasivan, *"A State-of-the-Art Review of Time Series Forecasting Using Deep Learning Approaches," IJRITCC*, Vol. 10, No. 12, pp. 92–105, Dec. 2022, doi: 10.17762/ijritcc.v10i12.5890.

[5] J. Zhao, Y. Gao, Y. Guo, and Z. Bai, *"Travel Time Prediction of Expressway Based On Multi-Dimensional Data and The Particle Swarm Optimization–Autoregressive Moving Average with Exogenous Input Model," Advances in Mechanical Engineering*, Vol. 10, No. 2, p. 1687814018760932, Feb. 2018, doi: 10.1177/1687814018760932.

[6] G. Terti *et al.*, *"ANYCaRE: a Role-Playing Game to Investigate Crisis Decision-Making and Communication Challenges in Weather-Related Hazards," Natural Hazards and Earth System Sciences*, Vol. 19, No. 3, pp. 507–533, 2019, doi: 10.5194/nhess-19-507-2019.

[7] S. Bierens, K. Boersma, and M. J. C. van den Homberg, *"The Legitimacy, Accountability, and Ownership of an Impact-Based Forecasting Model in Disaster Governance," Politics and Governance; Vol 8, No 4 (2020): The Politics of Disaster GovernanceDO - 10.17645/pag.v8i4.3161*, Dec. 2020, [Online]. Available: https://www.cogitatiopress.com/politicsandgovernance/article/view/3161

[8] N. Lingaraju and H. Savalegowda Mohan, *"A Novel Weather Parameters Prediction Scheme and Their Effects on Crops," IJECE*, Vol. 12, No. 1, p. 639, Feb. 2022, doi: 10.11591/ijece.v12i1.pp639-648.

[9] T. P. Agyekum, P. Antwi-Agyei, and A. J. Dougill, *"The Contribution of Weather Forecast Information to Agriculture, Water, and Energy Sectors in East and West Africa: A Systematic Review," Frontiers in Environmental Science*, vol. 10, 2022, [Online]. Available: https://www.frontiersin.org/articles/10.3389/fenvs.2022.935696

[10] A. Mehdipour Pirbazari, M. Farmanbar, A. Chakravorty, and C. Rong, *"Short-Term Load Forecasting Using Smart Meter Data: A Generalization Analysis," Processes*, Vol. 8, No. 4, 2020, doi: 10.3390/pr8040484.

[11] M. Diqi, *"StockTM: Accurate Stock Price Prediction Model Using LSTM," ijicom*, Vol. 4, No. 1, p. 1, Oct. 2022, doi: 10.35842/ijicom.v4i1.50.

[12] C.-W. Chang, H.-W. Lee, and C.-H. Liu, *"A Review of Artificial Intelligence Algorithms Used for Smart Machine Tools," Inventions*, Vol. 3, No. 3, 2018, doi: 10.3390/inventions3030041.

[13] N.-T. Pham, N.-Y. Tran-Van, and K.-H. Le, *"LS-SPP: A LSTM-Based Solar Power Prediction Method from Weather Forecast Information,"* in *2021 8th Nafosted Conference on Information and Computer Science (NICS)*, IEEE, Dec. 2021. doi: 10.1109/nics54270.2021.9701529.

[14] Z. Karevan and J. A. K. Suykens, *"Transductive LSTM for Time-Series Prediction: An Application to Weather Forecasting," Neural Networks*, Vol. 125, pp. 1–9, May 2020, doi: 10.1016/j.neunet.2019.12.030.

[15] G. Yao, Z. Liu, X. Guo, C. Wei, X. Li, and Z. Chen, *"Prediction of Weather Radar Images via a Deep LSTM for Nowcasting,"* in *2020 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Jul. 2020. doi: 10.1109/ijcnn48605.2020.9206889.

[16] A. Srivastava and A. S, *"Weather Prediction Using LSTM Neural Networks,"* in *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, IEEE, Apr. 2022. doi: 10.1109/i2ct54291.2022.9824268.

[17] M. A. R. Suleman and S. Shridevi, *"Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model," IEEE Access*, Vol. 10, pp. 82456–82468, 2022, doi: 10.1109/access.2022.3196381.

[18] Z. Al Sadeque and F. M. Bui, *"A Deep Learning Approach to Predict Weather Data Using Cascaded LSTM Network,"* in *2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, IEEE, Aug. 2020. doi: 10.1109/ccece47787.2020.9255716.

[19] Park, Kim, Lee, Kim, Song, and Kim, *"Temperature Prediction Using The Missing Data Refinement Model Based on A Long Short-Term Memory Neural Network," Atmosphere*, Vol. 10, No. 11, p. 718, Nov. 2019, doi: 10.3390/atmos10110718.

[20] S. Majumdar, S. Kr. Biswas, B. Purkayastha, and S. Sanyal, *"Rainfall Forecasting for Silchar City using Stacked- LSTM,"* in *2023 11th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECON)*, IEEE, Feb. 2023. doi: 10.1109/iemecon56962.2023.10092355.

[21] U. Sharma and C. Sharma, *"Deep Learning Based Prediction of Weather Using Hybrid_stacked Bi-Long Short Term Memory,"* in *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, Jan. 2022. doi: 10.1109/confluence52989.2022.9734133.