International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 6, December 2024, pp. 7057~7066
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i6.pp7057-7066

Enhancing El Niño-Southern oscillation prediction using an attention-based sequence-to-sequence architecture

Karli Eka Setiawan, Renaldy Fredyan, Islam Nur Alam
Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia

Article Info

Article history:

Received Mar 26, 2024
Revised Jul 24, 2024
Accepted Aug 14, 2024

Keywords:

Climate change monitoring
Deep learning
Enso prediction
Luong attention
Nino index forecasting
Seasonal climate prediction
Seq2seq

ABSTRACT

The ability to accurately predict the EI Nino-Southern oscillation (ENSO) is essential for seasonal climate forecasting. Monitoring the Pacific Ocean's surface temperature has many benefits for human life, including a better understanding of climate and weather, the ability to predict summer and winter, the ability to manage natural resources, serving as a reference for maritime transportation and navigation needs, serving as a reference for climate change monitoring needs, and even serving as a renewable energy source by utilizing high sea surface temperatures. This study introduces a deep learning (DL) model with AttentionSeq2Luong model as our proposed model to the ENSO research community. The present study showcases the capability of our proposed model to effectively forecast the forthcoming monthly average Nino index compared to the baseline seq2seq architecture model. For the dataset, this study utilized monthly observations of Nino 12, Nino 3, Nino 34, and Nino 4 between January 1870 and August 2022. The brief result of our experiment was that applying Luong Attention in the seq2seq model reduced the RMSE error by around 0.03494, 0.04635, 0.03853, and 0.03892 for forecasting Nino 12, Nino 3, Nino 34, and Nino 4, respectively.

Corresponding Author:

Karli Eka Setiawan
Computer Science Department, School of Computer Science, Bina Nusantara University
Jakarta, Indonesia
Email: karli.setiawan@binus.ac.id

1. INTRODUCTION

In the system of global climate, the ocean serves as a thermal sink. Through the interchange of mass and energy, it is crucial in managing and maintaining the global environment. Oceans have recently warmed significantly as a result of the earth's energy imbalance (EEI), which results in the storage of over 93% of the planet's heat increase [1]. It is vital to comprehend the present and forthcoming oceanic conditions in order to maintain a worldwide environment that is as conducive and pleasant for human existence as feasible. To comprehend how the ocean contributes to climate change, it is important to understand two essential dynamic variables: subsurface temperature and salinity. Through heat absorption and storage throughout the most recent period of global warming, the oceans have warmed dramatically. The level of water warmth has dramatically grown in previous decades [2]. Some researchers shown that the worldwide upper ocean warmed considerably since 1993 [3]. Recently, the level of warmth absorption in the ocean floor beneath 300 m has risen [4]. The oceanic system boosts absorbance of heat, that stimulates the oceans' temperature rise. The surface air temperature is the primary climate change metric that is widely utilized. In consequence, the ocean thermal content in 2019 attained an all-time high [5]. In contrast, ocean warming studies must also consider the level of salinity of the water. Prior research established the salinity process that describes how warming indications travel starting at the top ocean up to the intermediate ocean, demonstrating the critical importance of saline circulation to the evolution of oceanic temperatures [6].

Furthermore, the salinity level in the oceans influences the worldwide natural cycle [7]. Ocean salinity and temperature are also linked to thermohaline expansions, which substantially contribute to rising sea levels [8]. It is crucial to identify and predict subsurface oceanic salinity and temperature features in order to gain a deeper knowledge about ocean's processes and variations [9]. The scarcity of in situ data severely restricts studies of causes and mechanism in the ocean interior [10], which introduces uncertainty into analyses and predictions of ocean warming [11]. With the exception of high latitudes, a constant observation of the whole world's ocean from 0 to 2,000 m has been made during the Argo era (since 2004). Before the Argo era, data were few. Recently, the utilization of neural networks and deep learning for Earth observation has increased. For instance, some researchers used climatic parameters such as air current and temperature (among other factors) to calculate sea surface temperature (SST) in order to reproduce the yearly and seasonal variation of SST prior to the advent of satellites [12]. Some academics predicted SST at multiple sites in the Baltic Sea spatial clustering, which enabled them to gain an understanding of the daily changes in SST historically and in the future, as well as an understanding of marine ecosystems and public management [13]. Krasnopolsky *et al.* [14] built artificial neural networks to determine the color of the ocean without these data (*i.e.*, simulating a sensor malfunction). The AI model developed by them estimates the Chlorophyll A concentration using as inputs satellite sea surface height, warmth level, salinity levels, and in situ Argo salinity and vertical warmth characteristics, as well as some further information (longitude, latitude, and time frame) [14]. During inference (the reconstruction phase), the model does not utilize observed Chlorophyll a concentration at a specific spot, nor does it use data from adjacent grid points to determine Chlorophyll a saturation level. Only during the training phase does the network encounter chlorophyll a measurement. Renosh *et al.* [15] created a dataset of entangled particles. using model and in situ data, smart maps, and satellite data.

Deep learning (DL), a technique that automatically finds the pattern in the large amount of data to be shaped in black box formulation, has the potential to address these challenges in meteorology and geophysics, particularly extreme learning machine (ELM) cluster [16], DL with recurrent neural network (RNN) [17], or long short-term memory (LSTM) to predict rainfall by using spatiotemporal data [18]. This study reported a successful experiment utilizing a Luong attention-based deep learning seq2seq model we called AttentionSeq2Luong as proposed to forecast the anomaly in the warm phase of the ocean surface in the middle Pacific. The proposed deep learning model was compared with a seq2seq model without an attention mechanism. The seq2seq is mostly implemented in many-to-many predictions, but this research applied it to many-to-one predictions, which made this research special. Another reason for employing seq2seq architecture is that the Pearson correlation coefficient (PCC) values in the dataset are diverse and are not dominated by strong PCC values [19]. Meanwhile, when the dataset is dominated by a high PCC value, the seq2seq architecture does not perform well.

2. RELATED WORKS

Until now, long-term sequences of multisource sea surface measurements from satellite remote sensing have been produced, but all of them are all restricted to the ocean surface [10]. Past study indicated that the deeper ocean remote sensing (DORS) approach, when combined with a data set of type float, can have an opportunity to determine the ocean subsurface implicitly from satellite-based images. Data integration and mathematical modelling [20], a dynamic conceptual approach [21], and an intuitive statistical technique are all examples of DORS methodologies [22]. Due to their intricacy and ambiguity, the precision of mathematical and dynamic modeling techniques for massive subsurface modelling and prediction cannot be promised. Utilizing deep learning approach namely 3D-EddyNet, Feng *et al.* [23] experimentally approximated mesoscale three-dimensional sea heat formations. A number of experts reported a consequential observational pattern mapping using satellite altimetry to determine the four-dimensional shape of the Southern Ocean [24]. In a period of tremendous oceanographic information, however, computational models powered by data, especially those based on advanced deep learning approaches, outperform excellently and can reach superior precision in the DORS field and implementation.

Intuitive mathematical and statistical models, such as the linear regression model [25], the empirical orthogonal function-based approach [26], the model of weighted-based geographically regression [27]. Even though regular methods of machine learning have contributed greatly to DORS research, they are unable to compensate for the temporal and spatial characteristics of oceanic data. Deep learning approach have an opportunity to circumvent constraints and improve outcomes. Long short-term memory (LSTM) accumulates time-series data and enables learning process, whereas convolutional neural networks analyze the geographical information details of input and enable spatial learning. Deep learning has been used successfully in many other industries, but in geosciences, it is nevertheless in its earliest phases and shows great promise in the DORS discipline as ocean observation data continues to grow [28]. Although certain investigations have employed deep learning techniques for oceanic mapping, the technologies are still not widely implemented on oceans [29]. Available data in the field of oceanography are nonlinear in space and time [30], and the deep learning approach with dual-directional LSTM, namely bidirectional LSTM (Bi-LSTM), is able to understand these complex and influenced by time characteristics and improve precise forecasts. The idea may be applicable for reconstructing time-series subsurface properties and deriving other crucial oceanic variables, such as stream, turbulence, and others.

3. METHOD

3.1. Data

El Nino and La Nina are contrary stages of the El Nino-Southern oscillation that may result in worldwide catastrophic weather events such as drought and inundation. El Nino is a balmy phase characterized by a higher-than-average SST in the middle and eastern equatorial Pacific Ocean. As shown in Figure 1, there are four Nino areas for monitoring SST in the tropical Pacific Ocean: Nino 1+2, Nino 3, Nino 4, and Nino 3.4 zones. There are four anomaly event categories in ENSO research, such as weak (0.5–0.9 °C), moderate (1.0–1.4 °C), strong (1.5–1.9 °C), and exceptionally strong (>2.0 °C), while the associated negative numbers are utilized for categorizing the intensity of the La Nina phenomenon.

The National Oceanic and Atmospheric Administration site (*https://origin.cpc.ncep.noaa.gov*) provides access to historical ONI data, where this research obtained our dataset. This research paper utilized historical ONI data from January 1870 to August 2022 spanning fifteen decades. The dataset contains 1,831 monthly data points. The dates of each data point correspond to the quarter's middle month. The ONI value is the average of the SST anomaly values for the previous three months. February's ONI number is the mean of January through March's ONI numbers. The datasets used in this research are depicted in Figure 2(a) by explaining Nino 12, Nino 3, Nino 34, and Nino 4 data from January 1870 until August 2022 with y axis in degrees Celcius and x axis explaining monthly timesteps for 152 years. This research used Pearson correlation coefficient values as in Figure 2(b) to analyze the correlation between all variables explained as in (1), where the PCC value is denoted as r_{xy}, the compared variables are denoted as x and y, and the mean values of each parameter illustrated as $\bar{x}$ and $\bar{y}$.

$$r_{xy} = \frac{\sum(x_i - \bar{x}) \sum(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}}) \quad (1)$$

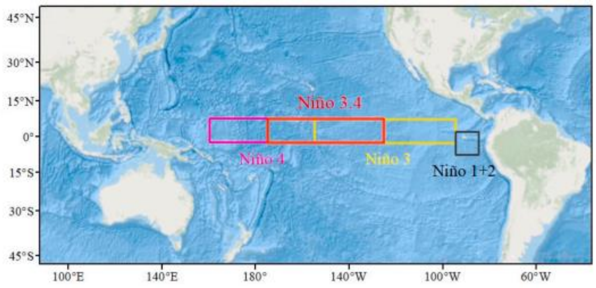


Figure 1. Nino regions

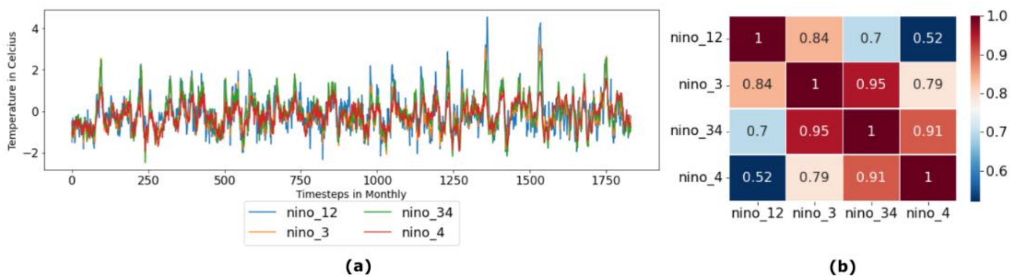


Figure 2. Dataset illustration with (a) line charts of monthly sea surface temperature anomalies and (b) Pearson correlation coefficients (PCC) between all Niño indices

3.2. Models

Encoder-decoder or sequence-to-sequence (seq2seq) is one of the architectures of deep learning that emerged from machine translation development [31]. Seq2seq architecture consist of two layers namely encoder and decoder. This research decided to implement LSTM into both encoder and decoder layers. Because LSTM is ideally suited for time series prediction, as well as any other task requiring temporal memory [32]. As the history of the beginning of seq2seq, seq2seq architecture was accurately and effectively applied in translating English to French with very long texts [33]. The Figure 3 illustrate the calculation and process of LSTM, which consist of three gates, such as forget, input, and output gate [32]. This study examined two seq2seq models, simple LSTM and AttentionSeq2Luong, to predict one timestep in the future (t+1) based on two years of data with 24 previous timesteps Figure 4.

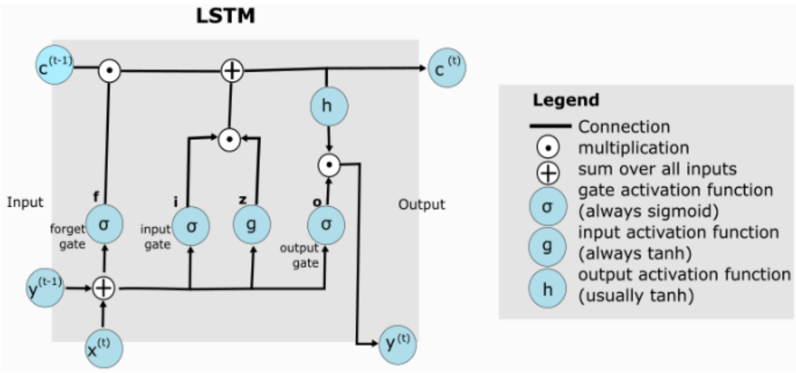


Figure 3. LSTM architecture illustration

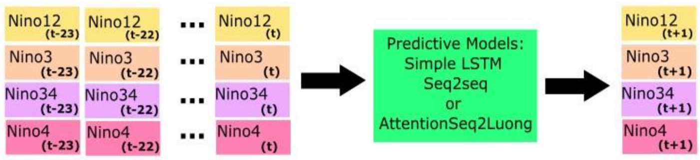


Figure 4. Prediction scenario

3.3. Simple Seq2seq LSTM model

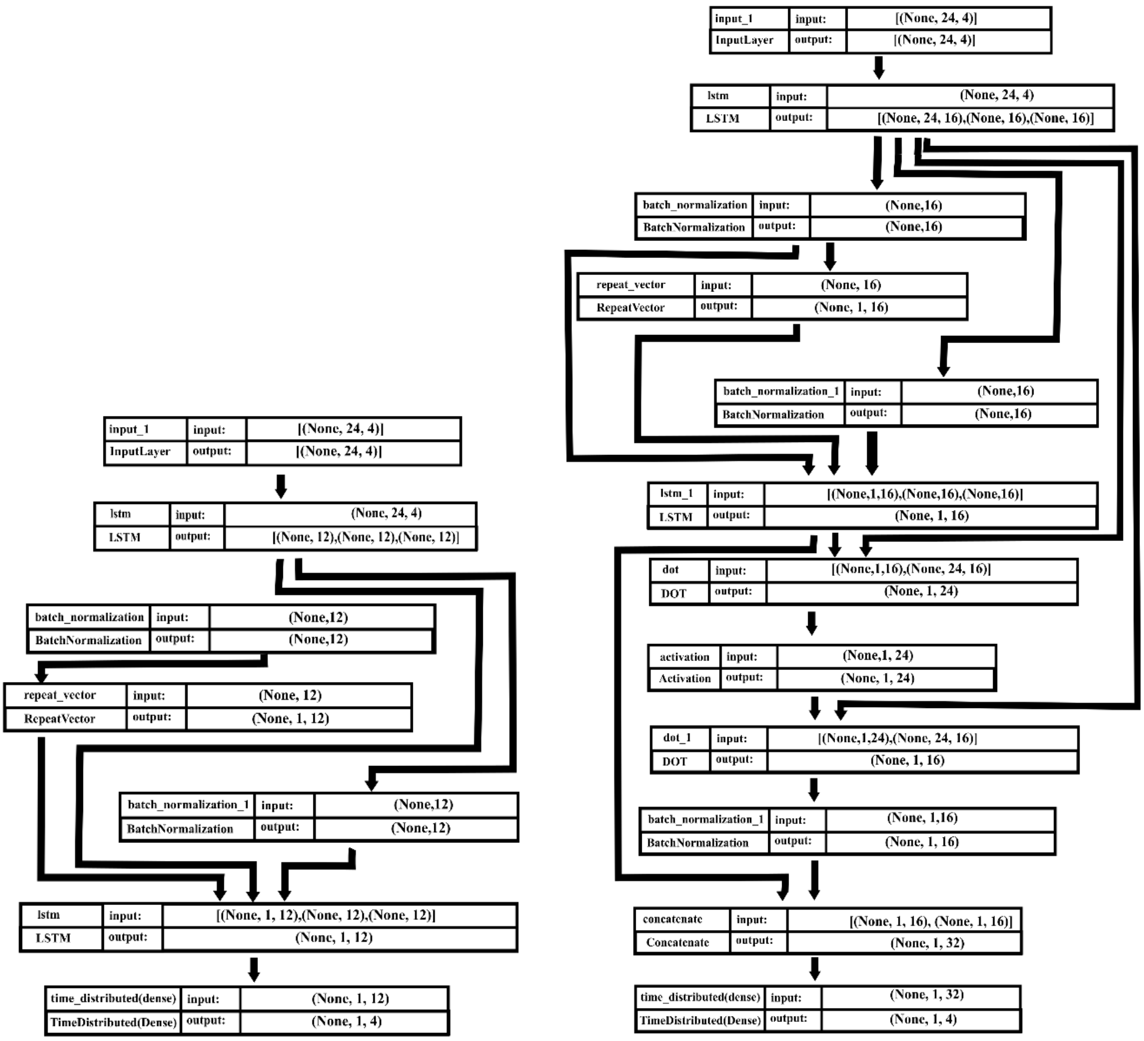
The architecture of the comparison model, simple Seq2seq LSTM, is illustrated in Figure 5(a). This research adopted the seq2seq architecture in our previous research on indoor climate prediction [19]. As in our previous, the research batch normalization was still applied because it can boost the training process to achieve convergence quickly [34]. The repeat vector layer seems useless in this experiment because this research only predicted one timestep in the future, however, it is still permissible to be implemented.

3.4. Proposed models with AttentionSeq2Luong

Figure 5(b) depicts the architecture of our proposed model, the AttentionSeq2Luong. This architecture has been used in our previous research, especially in predicting indoor climate data [33]. Due to the similarity in dataset characteristics seen in the PCC score as shown in Figure 2, this research tried to reuse our previous architecture models for monitoring the anomaly of the warm phase of the oceanic surface in the middle of the Pacific. The architecture differed little from simple LSTM seq2seq, but the Luong attention mechanism was included. In Figures 5(a) and 5(b), the first dot layer represented Luong attention using dot scoring by implementing a Keras dot layer for combining hidden states from the encoder and decoder layers. The second dot layer represented the context vector for merging the dot score results after activation with the encoder's hidden states. The final stage is the concatenate layer, which combines the context vector after batch normalization with the decoder's hidden states.

3.5. Hyperparameter settings

The differences between simple seq2seq and proposed models is the execution of the Luong attentiveness technique, which made this research implementing the same hyperparameters setting to both models. In our experiment, several number neurons in LSTM layers were tested to find the best setting for this case with neuron={8, 16, 32, 48, 64}. The results showed that 16 neurons were the best for this case. The learning rate, number of epochs, and batch size were respectively set to 0.01; 100; and 256. Both models utilized Adam optimizer. Meanwhile for prediction scenario, this research predicted one timestep in the future based on two years data or 24 previous timesteps.



(a) (b)

Figure 5. Proposed deep learning architecture using (a) a simple Seq2seq LSTM architecture and (b) an attentionSeq2Luong architecture

4. RESULTS AND DISCUSSION

4.1. Preprocessing dataset

This research did not implement any normalization and standardization. The datasets were split into two parts with the proportion 80% of original datasets to be a train set and 20% of original dataset to be a test set as illustrated in Figure 6. This approach allowed for a clear evaluation of the performance of the models without preprocessing alterations. The results reflect the ability of the model to handle raw data, providing insights into its robustness and generalization capabilities.

4.2. Model training

To measure the training process in this research, the loss plot in MAE, as shown in Figure 7(a) for simple seq2seq and Figure 7(b) for attentionSeq2Luong, was used to monitor the model in learning process. The dataset used for the training process were the train sets colored blue in Figure 6, where 80% of the train sets were used to train the model and the other 20% were used to validate the training process. A quick glance at Figure 7 shows that the slit between training and validation loss using simple seq2seq LSTM was wider than the slit between training and validation loss using AttentionSeq2Luong. All models were trained using a relatively small dataset, allowing this research to be sufficiently executed on Google Collaboratory with the TensorFlow version 2.8.2 and Keras version 2.8.1 libraries.

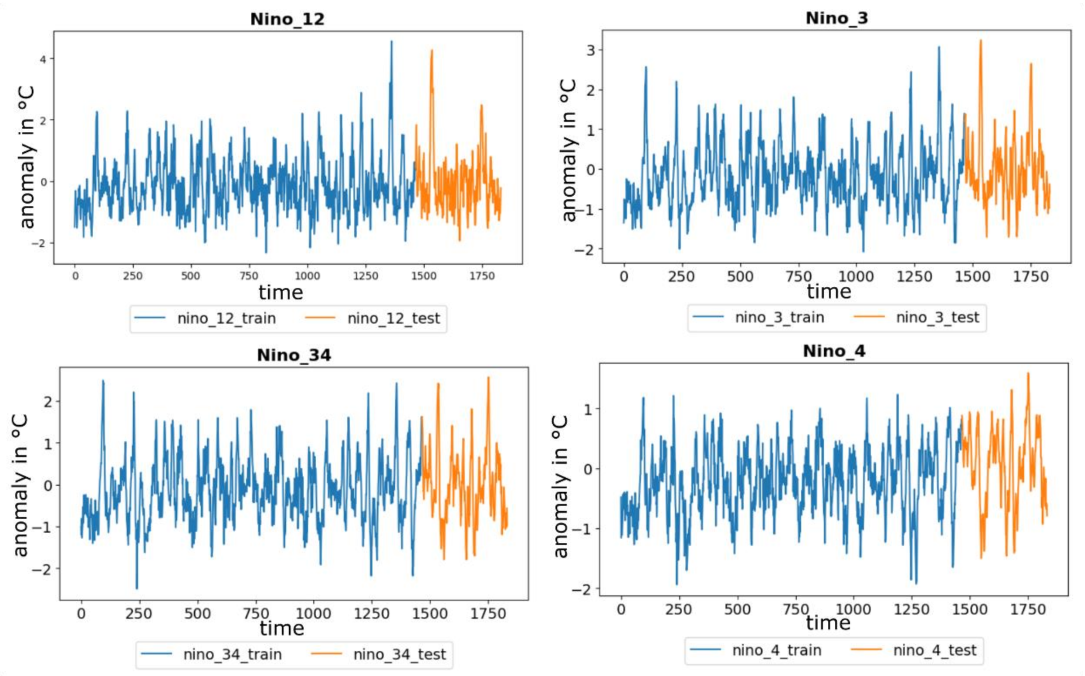
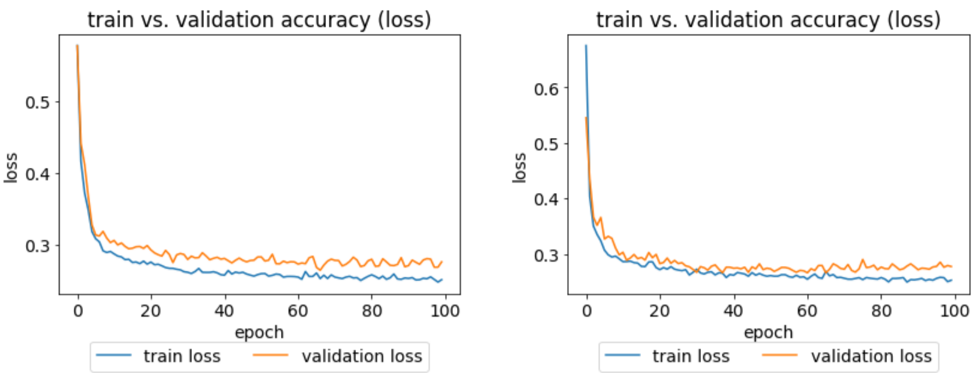


Figure 6. Split dataset



(a) (b)

Figure 7. Train and validation loss plot in MAE with (a) Seq2seq LSTM and (b) attentionSeq2Luong

4.3. Model testing and research result

This study tested the models with testing data to compare their performance. By using testing data, Table 1 compare the prediction results of Nino 12, Nino 3, Nino 34, and Nino 4 to the ground truth data using MAE, RMSE, and R^2. Meanwhile, the comparison of both models' prediction results with ground truth data is depicted in Figure 8. From Table 1, overall, the implementation of AttentionSeq2Luong model gave significant improvements in almost all predictions. Applying Luong Attention in the seq2seq model reduced the RMSE error by around 0.03494, 0.04635, 0.03853, and 0.03892 for forecasting Nino 12, Nino 3, Nino 34, and Nino 4, respectively. Applying Luong attention also boosted the coefficient of determination values for forecasting Nino 12, Nino 3, Nino 34, and Nino 4 by 0.27497, 0.14723, 0.08050, and 0.16375, respectively. In MAE, our proposed model improved the prediction of Nino 3, Nino 34, and Nino 4 by decreasing the error around 0.00597, 0.02124, and 0.03229, respectively, but increased the prediction of Nino 12 by around 0.01263.

Table 1. Prediction results

	Metrics	Simple Seq2seq	Luong Attention-based Seq2seq
Nino 12	MAE	**0.409369**	0.421831
	RMSE	0.596675	**0.561740**
	R^2	0.207550	**0.482527**
Nino 3	MAE	0.308916	**0.302944**
	$RMSE$	0.421389	**0.375040**
	R^2	0.617100	**0.764337**
Nino 34	MAE	0.268810	**0.247574**
	$RMSE$	0.347810	**0.309283**
	R^2	0.758974	**0.839480**
Nino 4	MAE	0.246585	**0.214295**
	$RMSE$	0.300187	**0.214295**
	R^2	0.612627	**0.776377**

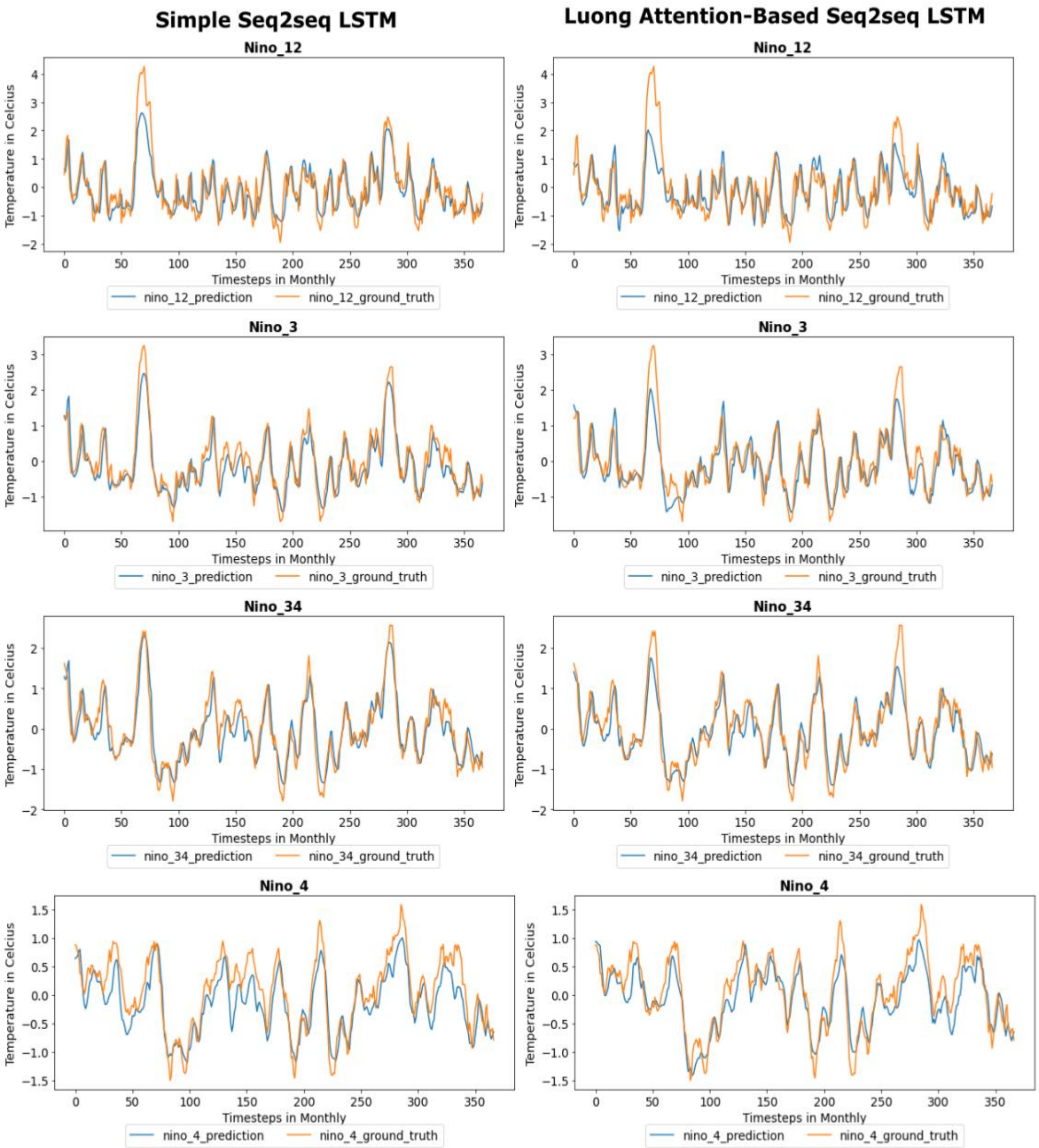


Figure 8. Testing results for predicting Nino 12, 3, 34, and 4

From explained data comparison of prediction result using data testing between base model using simple seq2seq and our proposed model using AttentionSeq2Luong, it can be concluded that our proposed model can outperform the base model, which can bring positive impact to ENSO research and time-series prediction research. Compared to our previous research in ENSO using other deep learning with temporal transformer-based models, this research using deep learning with seq2seq-based models resulted in a much better positive improvement result with a lower error in MAE and RMSE [35]. This research found that deep learning using a seq2seq-based approach is promising to be explored more in time-series research, as proven by the successful results of this research.

5. CONCLUSION

By conducting experiments using the Nino anomaly dataset, specifically Nino 12, Nino 3, Nino 34, and Nino 4, and implementing deep learning models for the time-series data, this research made a significant contribution to the monitoring of warm phase anomalies at the sea surface in the middle pacific. The results showed that our proposed models with the AttentionSeq2Luong model outperformed the LSTM seq2seq models without the Luong attention implementation when predicting the next timestep in the future based on 24 previous timesteps (predicting the next month's data based on two years of previous data). Based on the data from our research, we conducted a successful experiment using a Luong attention-based deep learning seq2seq model, which we named AttentionSeq2Luong. This model was proposed to forecast the anomaly in the warm phase of the ocean surface in the middle Pacific. This model may have positive impacts on our community, including improved predictive accuracy, enhanced understanding of ENSO dynamics, practical application and policy implications, cross-disciplinary insight, community preparedness and awareness, and data utilization in handling big data.

Another attention-based mechanism in deep learning research is Bahdanau attention-based, which may give significant improvement in Pacific Ocean surface temperature monitoring research for future research to be explored further. There is a belief that time series forecasting is a difficult task, so incorporating new variables such as atmospheric data such as wind pattern and atmospheric pressure, oceanic variables such as sea-level height level, ocean current, and subsurface ocean temperature, and precipitation pattern data may enhance the prediction outcomes. Collaboration with oceanographers is required for future research aimed at gaining knowledge about our planet and preventing damage that will affect future humanity.

REFERENCES

[1] B. Meyssignac *et al.*, "Measuring global ocean heat content to estimate the earth energy imbalance," *Frontiers in Marine Science*, vol. 6, Aug. 2019, doi: 10.3389/fmars.2019.00432.

[2] L. Cheng *et al.*, "2018 continues record global ocean warming," *Advances in Atmospheric Sciences*, vol. 36, no. 3, pp. 249–252, Mar. 2019, doi: 10.1007/s00376-019-8276-x.

[3] G. C. Johnson and J. M. Lyman, "Warming trends increasingly dominate global ocean," *Nature Climate Change*, vol. 10, no. 8, pp. 757–761, Aug. 2020, doi: 10.1038/s41558-020-0822-0.

[4] L. C. Allison *et al.*, "Towards quantifying uncertainty in ocean heat content changes using synthetic profiles," *Environmental Research Letters*, vol. 14, no. 8, Aug. 2019, doi: 10.1088/1748-9326/ab2b0b.

[5] L. Cheng *et al.*, "Record-setting ocean warmth continued in 2019," *Advances in Atmospheric Sciences*, vol. 37, no. 2, pp. 137–142, Feb. 2020, doi: 10.1007/s00376-020-9283-7.

[6] G. Knorr *et al.*, "A salty deep ocean as a prerequisite for glacial termination," *Nature Geoscience*, vol. 14, no. 12, pp. 930–936, Dec. 2021, doi: 10.1038/s41561-021-00857-3.

[7] S. Bao, R. Zhang, H. Wang, H. Yan, Y. Yu, and J. Chen, "Salinity profile estimation in the Pacific Ocean from satellite surface salinity observations," *Journal of Atmospheric and Oceanic Technology*, vol. 36, no. 1, pp. 53–68, Jan. 2019, doi: 10.1175/JTECH-D-17-0226.1.

[8] Cazenave *et al.*, "Global sea-level budget 1993-present," *Earth System Science Data*, vol. 10, no. 3, pp. 1551-1590, doi: 10.3929/ethz-b-000287786.

[9] W. Lu, H. Su, X. Yang, and X.-H. Yan, "Subsurface temperature estimation from remote sensing data using a clustering-neural network method," *Remote Sensing of Environment*, vol. 229, pp. 213–222, Aug. 2019, doi: 10.1016/j.rse.2019.04.009.

[10] H. Su, T. Zhang, M. Lin, W. Lu, and X.-H. Yan, "Predicting subsurface thermohaline structure from remote sensing data based on long short-term memory neural networks," *Remote Sensing of Environment*, vol. 260, Jul. 2021, doi: 10.1016/j.rse.2021.112465.

[11] G. Wang, L. Cheng, J. Abraham, and C. Li, "Consensuses and discrepancies of basin-scale ocean heat content changes in different ocean analyses," *Climate Dynamics*, vol. 50, no. 7–8, pp. 2471–2487, Apr. 2018, doi: 10.1007/s00382-017-3751-5.

[12] P. P. Sarkar, P. Janardhan, and P. Roy, "Prediction of sea surface temperatures using deep learning neural networks," *SN Applied Sciences*, vol. 2, no. 8, Aug. 2020, doi: 10.1007/s42452-020-03239-3.

[13] C. Dutheil, H. E. M. Meier, M. Gröger, and F. Börgel, "Understanding past and future sea surface temperature trends in the Baltic Sea," *Climate Dynamics*, vol. 58, no. 11–12, pp. 3021–3039, Jun. 2022, doi: 10.1007/s00382-021-06084-1.

[14] V. Krasnopolsky, S. Nadiga, A. Mehra, E. Bayler, and D. Behringer, "Neural networks technique for filling gaps in satellite measurements: application to ocean color observations," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–9, 2016, doi: 10.1155/2016/6156513.

[15] P. Renosh *et al.*, "Construction of multi-year time-series profiles of suspended particulate inorganic matter concentrations using machine learning approach," *Remote Sensing*, vol. 9, no. 12, Dec. 2017, doi: 10.3390/rs9121320.

[16] R. Fredyan, M. R. N. Majiid, and G. P. Kusuma, "Spatiotemporal analysis for rainfall prediction using extreme learning machine cluster," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 13, no. 6, pp. 2240–2248, Dec. 2023, doi: 10.18517/ijaseit.13.6.18214.

[17] R. E. Caraka, R. C. Chen, H. Yasin, S. Suhartono, Y. Lee, and B. Pardamean, "Hybrid vector autoregression feedforward neural network with genetic algorithm model for forecasting space-time pollution data," *Indonesian Journal of Science and Technology*, vol. 6, no. 1, pp. 243–266, Jan. 2021, doi: 10.17509/ijost.v6i1.32732.

[18] R. Fredyan and G. P. Kusuma, "Spatiotemporal convolutional LSTM with attention mechanism for monthly rainfall prediction," *Communications in Mathematical Biology and Neuroscience*, 2022, doi: 10.28919/cmbn/7761.

[19] K. E. Setiawan, G. N. Elwirehardja, and B. Pardamean, "Sequence to sequence deep learning architecture for forecasting temperature and humidity inside closed space," in *2022 10th International Conference on Cyber and IT Service Management (CITSM)*, Sep. 2022, pp. 1–7, doi: 10.1109/CITSM56380.2022.9936008.

[20] S. Temitope Yekeen and A.-L. Balogun, "Advances in remote sensing technology, machine learning and deep learning for marine oil spill detection, prediction and vulnerability assessment," *Remote Sensing*, vol. 12, no. 20, Oct. 2020, doi: 10.3390/rs12203416.

[21] H. Yan, H. Wang, R. Zhang, J. Chen, S. Bao, and G. Wang, "A dynamical-statistical approach to retrieve the ocean interior structure from surface data: SQG-mEOF-R," *Journal of Geophysical Research: Oceans*, vol. 125, no. 2, Feb. 2020, doi: 10.1029/2019JC015840.

[22] Y. Jeong, J. Hwang, J. Park, C. J. Jang, and Y.-H. Jo, "Reconstructed 3-D ocean temperature derived from remotely sensed sea surface measurements for mixed layer depth analysis," *Remote Sensing*, vol. 11, no. 24, Dec. 2019, doi: 10.3390/rs11243018.

[23] P. Feng, Z. Fu, L. Hu, S. Wu, Y. Wang, and F. Zhang, "3D-EddyNet: a novel approach for identifying three-dimensional morphological features of mesoscale eddies in the ocean," *Journal of Marine Science and Engineering*, vol. 11, no. 9, Sep. 2023, doi: 10.3390/jmse11091779.

[24] A. R. Gray, "The four-dimensional carbon cycle of the Southern Ocean," *Annual Review of Marine Science*, vol. 16, no. 1, pp. 163–190, Jan. 2024, doi: 10.1146/annurev-marine-041923-104057.

[25] B. Ai *et al.*, "Convolutional neural network to retrieve water depth in marine shallow water area from remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2888–2898, 2020, doi: 10.1109/JSTARS.2020.2993731.

[26] J. Huang, Y. Luo, Y. Li, J. Shi, X. Zheng, and J. Wang, "Analysis of sound speed profile in the South China Sea based on empirical orthogonal function algorithm," in *2021 OES China Ocean Acoustics (COA)*, Jul. 2021, pp. 166–171, doi: 10.1109/COA50123.2021.9520009.

[27] H. Su, L. Huang, W. Li, X. Yang, and X. Yan, "Retrieving ocean subsurface temperature using a satellite-based geographically weighted regression model," *Journal of Geophysical Research: Oceans*, vol. 123, no. 8, pp. 5180–5193, Aug. 2018, doi: 10.1029/2018JC014246.

[28] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, Feb. 2019, doi: 10.1038/s41586-019-0912-1.

[29] A. Barth, A. Alvera-Azcárate, C. Troupin, and J.-M. Beckers, "DINCAE 2.0: multivariate convolutional neural network with error estimates to reconstruct sea surface temperature satellite and altimetry observations," *Geoscientific Model Development*, vol. 15, no. 5, pp. 2183–2196, Mar. 2022, doi: 10.5194/gmd-15-2183-2022.

[30] C. Chapman and A. A. Charantonis, "Reconstruction of subsurface velocities from satellite observations using iterative self-organizing maps," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 617–620, May 2017, doi: 10.1109/LGRS.2017.2665603.

[31] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1724–1734, doi: 10.3115/v1/D14-1179.

[32] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[33] K. Eka Setiawan, G. N. Elwirehardja, and B. Pardamean, "Indoor climate prediction using attention-based sequence-to-sequence neural network," *Civil Engineering Journal*, vol. 9, no. 5, pp. 1105–1120, May 2023, doi: 10.28991/CEJ-2023-09-05-06.

[34] J. Bjorck, C. Gomes, B. Selman, and K. Q. Weinberger, "Understanding batch normalization," *Advances in Neural Information Processing Systems*, vol. 2018-Decem, pp. 7694–7705, 2018.

[35] R. Fredyan and K. E. Setiawan, "An investigation into improving El Niño-Southern Oscillation prediction based on temporal transformer architecture," *Communications in Mathematical Biology and Neuroscience*, 2024, doi: 10.28919/cmbn/8371.

BIOGRAPHIES OF AUTHORS

Karli Eka Setiawan received an S.Si. degree in mathematics from Padjadjaran University in 2015 and an M.Kom. degree in computer science from Bina Nusantara University in 2023. He achieved outstanding grades of excellence, summa cum laude, and best graduate in his master's study. His research interests cover artificial intelligence, computer vision, data science, data security, blockchain, and mathematical modeling. He is an associate lecturer specialist S2 in the School of Computer Science at Bina Nusantara University. He also has experience as a credit analyst at a conventional bank, which is part of a state-owned enterprise. He can be contacted at email: karli.setiawan@binus.ac.id.

Renaldy Fredyan received a master's degree in computer science from Bina Nusantara University in Indonesia in 2023. He is currently employed by Indonesia's Bina Nusantara University as a lecturer in the Department of Computer Science. His areas of interest in research include time series, machine learning, pattern recognition, and computer vision. He can be contacted at email: renaldy.fredyan@binus.ac.id.

Islam Nur Alam is a lecturer at Bina Nusantara University (BINUS). He has 2 years of experience as a data science researcher with a proven track record of building successful algorithms and predictive models for image classification using convolutional neural networks. He is highly adept in clustering and classification, content-based filtering, data analysis, and visualization. Currently, he is also training individuals in data science skills, particularly focusing on natural language processing and machine translation tasks. He can be contacted at email: islam.alam@binus.ac.id.