

Time series prediction based on machine learning: A case study, temperature forecasting in Vietnam

Ngo Thi Thanh Hoa, Chu Thi Quyen*, Nguyen Thi Cam Ngoan

Hanoi University of Industry - No 298 Cau Dien street, Bac Tu Liem district, Hanoi

*Corresponding author: chuthiquyen_cntt@hau.edu.vn.

Received 4 Nov 2022; Revised 24 Dec 2022; Accepted 2 Feb 2023; Published 28 Feb 2023.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.85.2023.152-162>

ABSTRACT

In recent years, there has been a surge in interest in the subject of machine learning for prediction. In this study, a temperature dataset of Vietnam's stations is examined in order to anticipate temperature. Several forecasting models are used to accomplish this goal. First, a traditional time-series forecasting approach such as Seasonal Autoregressive Integrated Moving Average is used (SARIMA). Then, more complex approaches such as XGBoost, Encoder-Decoder, and Prophet are used. The models' performances are compared using several accuracy assessment methods (e.g., Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE)). The findings demonstrate the superiority of the deep learning approach over the other methods.

Keywords: Machine learning; Temperature forecast; Deep learning; Time series; SARIMA; XGBoost.

1. INTRODUCTION

The temperature of the air is a typical meteorological variable that indicates how hot or cold it is. It influences not only plant and animal development and reproduction, but also practically all other climatic factors such as evaporation rate, relative humidity, wind speed, wind direction, and precipitation patterns [1].

In Vietnam, temperature data such as average temperature, maximum temperature, and lowest temperature are utilized to synthesis, quantify, and evaluate cold air. This motivates forecasting and research work on temperature [2].

The goal of this research is to use historical data to forecast future temperature values. As it has been mentioned before, seasonal times-series forecasting is critical in strategic decision-making and future activity planning. We would also like to call into question the capacity of machine learning models to do seasonal forecasting, which is frequently used in a variety of applications. Furthermore, we would want to examine the performance of several traditional and advanced time-series forecasting algorithms in order to determine which method is the best.

In recent years, machine learning methods have been known as a powerful tool and have been applied in different fields especially in forecasting. The following issues will be addressed by this research: (a) How well will machine learning approaches perform for seasonal item forecasting? (b) Which of these time-series forecasting approaches, classic and new, will produce the most accurate results?

To the best of our knowledge, this is the first study to assess the performance of four forecasting models to predict the daily temperature, including traditional approaches and sophisticated techniques such as XGBoost and Prophet methods. In this investigation, the capability of encoder-decoder neural network which is mostly used for machine translate is also investigated for forecasting temperature.

The rest of this document is outlined as follows. In the next section, we make a statement to clarify the problem. The third section is a literature review is provided which discusses different approaches for forecasting. And the fourth section is devoted to the unified presentation of four methods used for time series modeling. The fourth section describes the data, the pre-processing

steps, and the metrics to evaluate models. In section 5, the performances of different approaches are reported, and the results are compared. Finally, section 6 is dedicated to the conclusions.

2. PROBLEM DESCRIPTION

In this paper, we have taken daily temperature series of several cities of Vietnam. Each series contains rainfall recorded from 2014 to 2019. Our aim is to forecast the temperature on the next 7 days in those cities.

Construction of input/output pairs: Let $\{x_1, x_2, \dots, x_N\}$ stands for a rainfall time series. It can be re-constructed into a series of delay vectors as

$$X_t = x_t, x_{t+\tau}, x_{t+2\tau}, \dots, x_{t+(k-1)\tau}$$

where $X_t \in \mathbb{R}^k$, τ is the delay time as a multiple of the sampling period, and k is the embedded dimension. Suppose that X_t is a vector, then the available historical data can be summarized as a set of pairs $\{X_t, x_{t+T+(k-1)\tau} : t = 1, \dots, n\}$.

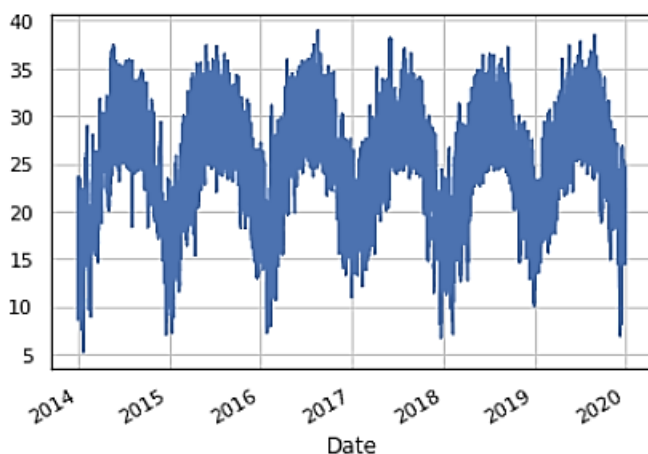


Figure 1. Typical time series for temperatures.

Based on Fig. 1 which displays seasonality in its pattern is temperature. It is not surprising that the temperature of this category reached their peak in the summer season.

3. RELATED WORKS

The precision of the prediction is an essential aspect of temperature forecasting. As a result, significant efforts have been undertaken to improve the accuracy of this procedure. In this field's publications, accuracy is acquired by comparing the actual temperature to the expected results and utilizing error measuring methods such as RMSE and MAE. Some relevant works are reviewed in this section.

3.1. Classical time-series forecasting methods

There is no one optimum method for solving time-series forecasting challenges [3]. To deal with time-series forecasting, each situation may require a distinct strategy. Moving Average (MA) is one of the most basic forecasting approaches for time series with no discernible seasonal trend [4]. A more advanced variant of MA known as Autoregressive Integrated Moving Average (ARIMA) has been employed in various articles. For example, Ahmad et al. [5] predicted by combining the wavelet technique with ARIMA model and an artificial neural network (ANN) applied to monthly maximum and minimum temperature data. The authors [6] effectively used multivariate ARIMA to predict the average global temperature. Yang et al. [7] recently applied a mixed ARIMA and neural network for network traffic forecasting. In [8], they employed the ARIMA model to anticipate bitcoin prices based on the effect of social media in a recent study.

Another traditional forecasting approach is seasonal ARIMA (SARIMA). This approach has been effectively applied in a variety of situations, including forecasting tourism demand [9] and estimating vehicular traffic flow [10]. However, Hamzaçebi [11] said that SARIMA's linear nature and inability to recognize nonlinear and highly volatile patterns might restrict prediction.

3.2. Machine learning methods

Many ML-based techniques to forecasting have been investigated in the literature. Artificial neural networks (ANN) and support vector machines (SVM) are the most extensively used methodologies in air temperature time series analysis. Most ANN models created to forecast temperature values, in particular, are represented by multilayer perceptron neural networks (MLPNN) and radial basis function neural networks (RBFNN) [12-17], with the Levenberg-Marquardt and gradient descent optimization methods being the most commonly utilized. When it comes to SVM models, the majority of the research in the field has used radial function base kernels [18-21]. In terms of performance, SVM has shown superior performance metrics than traditional ANNs [22] from 1 to 20 steps-ahead on a global scale.

Tree-based machine-learning algorithms, such as random forest [23] also is applied to historical load data in Kuala Lumpur (2000 to 2012) and its performance is compared with some alternative models.

3.3. Deep learning

Machine learning and deep learning techniques are useful tools for accurate forecasting. On a regional scale, however, modern Deep Learning (DL) techniques have been developed, with high accuracy values. The Recurrent Neural Networks (RNN) are essentially networks made up of loops that allow them to remember events from the past. As a result, they can be extremely beneficial in time-series prediction.

Specifically Convolutional and Long Short Term Memory (LSTM) RNN have been employed to anticipate hourly air temperature with very tiny errors for one step ahead [24]. Roesch and Günther [25] developed a similar technique to review yearly, monthly, and daily trends related with air temperature time series.

3.4. Prophet method

Because the Prophet technique is one of the most recently launched forecasting methods by the Facebook firm [26], its capacity has not been thoroughly explored. Prophet forecasting outcomes were comparable to traditional forecasting methods in a research on monthly temperature predictability done by [27]. However, in another study on hospital discharge volume prediction [28], the results were more accurate than the SARIMA model results. Furthermore, Subashini, Sandhiya, Saranya, and Harsha [29] suggested a Prophet time series model to anticipate website traffic, which yielded promising results. ARIMA and PROPHET were used by Satrio, Darmawan, Nadia, and Hanafiah [30] to anticipate COVID illness in Indonesia.

4. FORECASTING METHODS

The aim of this study is to perform a time-series forecasting on a seasonal item with the help of different models such as SARIMA, XGBoost, Prophet, and LSTM. In this section, we proposed recommended approaches for tackling the challenges associated with datasets to predict daily temperature. And also in this investigation, the model with the most accurate result is chosen as the best model.

4.1. SARIMA

Seasonal ARIMA or SARIMA model [31] is used to forecast seasonal timeseries, which is indicated by ARIMA (p, d, q) (P, D, Q) m, where p, d, and q are non-seasonal parameters and P, D, and Q are seasonal factors. They display the order of the autoregressive component, the degree

of the initial differencing, and the order of the moving average component. Furthermore, the seasonal parameter m can be defined as the cycle length. Eqs can express the entire model (1-2).

$$\Phi(B^m)\Phi(B)\nabla_m^D\nabla^d X_t = \Theta(B^m)\theta(B)Z_t \tag{1}$$

with

$$\nabla_m X_t = X_t - X_{t-m}, \nabla X_t = X_t - X_{t-1} \tag{2}$$

4.2. XGBoost

The XGBoost algorithm, a scalable machine learning method for tree boosting, will be explained in this section.

Boosting is a machine learning strategy for regression and classification issues. It creates a weak learner at each step and adds it to the overall model. Gradient Boosting Machines (GBM) are created when the weak learner for each step is dependent on the gradient direction of the loss function.

Assume we have data $\mathcal{D} = \{(x_i, y_i): i = \overline{1, \dots, n}, x_i \in \mathcal{R}^k, y_i \in \mathcal{R}\}$, so we have n observation with k features each and with a corresponding variable y . Let us define y_i as the outcome of an ensemble represented by the generalised model.

$$y_i = \phi(x_i) = \sum_{m=1}^M f_m(x_i) \tag{3}$$

In this scenario, f_m is a regression tree, and $f_m(x_i)$ is the score assigned by the m -th tree to the i -th observation in the data. In order to select functions f_m , we wish to minimize the following regularised objective function.

$$\mathcal{L}(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_m \Omega(f_m) \tag{4}$$

The loss function is denoted by l . To avoid overcomplicating the model, the penalty term is added as follows:

$$\Omega(f_m) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \tag{5}$$

where γ, λ are parameters determining the penalty for the number of leaves T and the magnitude of leaf weights w , respectively.

This penalty term is the first characteristic that distinguishes XGBoost from other tree boosting algorithms. Its goal is to prevent over-fitting and simplify the models generated by this technique.

An iterative approach is employed to minimize the objective function. We wish to add f_j in the j -th iteration to minimize the following objective function:

$$\mathcal{L}^j = \sum_{i=1}^n l(y_i, \hat{y}_i^{(j-1)} + f_j(x_i)) + \Omega(f_j) \tag{6}$$

We can simplify this function and derive a formula for loss reduction after the tree split from a given node using the Taylor expansion:

$$\mathcal{L}_{split} = \frac{1}{2} \left[\frac{(\sum_{i \in \mathfrak{X}_L} g_i)^2}{\sum_{i \in \mathfrak{X}_L} h_i + \lambda} + \frac{(\sum_{i \in \mathfrak{X}_R} g_i)^2}{\sum_{i \in \mathfrak{X}_R} h_i + \lambda} - \frac{(\sum_{i \in \mathfrak{X}} g_i)^2}{\sum_{i \in \mathfrak{X}} h_i + \lambda} \right] \tag{7}$$

where \mathfrak{X} represents a subset of the observations available in the current node and \mathfrak{X}_L and \mathfrak{X}_R represent subsets of the observations available in the left and right nodes following the split. The g_i and h_i functions are defined as follows:

$$g_i = \partial_{\hat{y}_i^{(j-1)}} l(y_i, \hat{y}_i^{(j-1)}), h_i = \partial_{\hat{y}_i^{(j-1)}}^2 l(y_i, \hat{y}_i^{(j-1)}) \tag{8}$$

This formula determines the optimal split at every given node. It is solely determined by the loss function (by the first and second-order gradients) and the regularisation value. It is clear that this technique can optimize any loss function provided by the user, as long as the first and second-order gradients are provided.

In addition to offering the regularized loss function, XGBoost has two extra features to prevent over-fitting (when compared to ordinary GBM). To begin, the weights of each new tree can be scaled down by a constant. It lessens the influence of a single tree on the final score, leaving more opportunity for future trees to enhance the model. Column sampling is the second feature. It operates similarly to random forests in that each tree is constructed using simply a column-wise sample from the training dataset.

4.3. LSTM encoder-decoder

The main objective of this part is to showcase how deep stacked bidirectional LSTMs can be applied to time series data as a sequence to sequence (Seq-2-Seq) based encoder-decoder model. As the name implies, sequence to sequence models accept a series of features as input and output a target sequence as a continuation of the input target sequence (it can forecast 'n' time steps into the future).

It consists of two parts: the encoder, which generates a context vector (encoding) of the input sequence, and the decoder, which decodes and predicts the targets.

In this paper we use the bidirectional LSTMs (BLSTM) which have two recurrent components: one that moves ahead and one that moves backward. The forward component computes the hidden and cell states similarly to a normal unidirectional LSTM, but the backward component computes them in reverse-chronological order, starting from time step T_x to 1. The idea behind employing a backward component is that we are allowing the network to perceive future data and adjust its weights accordingly. This may assist the network in capturing dependencies that would not have been recorded by the normal (forward) LSTM. Because of its ability to capture dependencies in the input sequence rather effectively, BLSTM is also a go to beginning method for most natural language processing (NLP) jobs.

The hidden and cell states of the forward component differ from those of the backward component in BLSTMs. To get an encoding, the hidden and cell states of the forward component must be concatenated with those of the backward component.

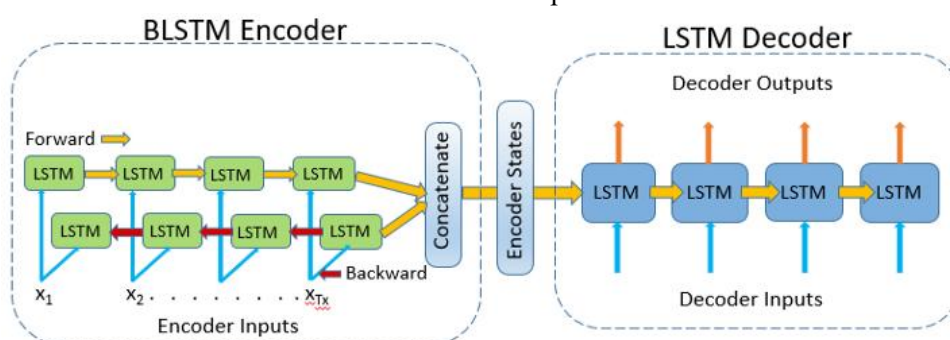


Figure 2. The architectures of LSTM encode-decoder.

4.4. Facebook Prophet

Facebook Prophet is an open-source time-series model generation algorithm that combines some classic principles with some novel twists. It excels at modeling time series with numerous seasonality and avoids some of the limitations mentioned above in other methods. It is composed of three functions plus an error term: growth $g(t)$, seasonality $s(t)$, holidays $h(t)$, and error $e(t)$:

$$y(t) = g(t) + h(t) + s(t) + \epsilon_t \quad (9)$$

where $y(t)$ is the additive regressive model; $g(t)$ is the trend factor; $h(t)$ is the holiday component; $s(t)$ is the seasonality component and \mathcal{E}_t is the error term. The trend factor $g(t)$ can be modeled in two ways:

- *Logistic growth model*: This model illustrates the growth in phases; in the first stage, the development is essentially exponential; after that, the saturation stage is reached; and from there, the growth is linear. The model may be written in (10):

$$g(t) = \frac{L}{1 + e^{-k(t-t_0)}} \tag{10}$$

In this mathematical model, L represents the model curve's maximum value; is the growth rate; and is the value at the sigmoid point.

- *Logistic Growth model*: This is Prophet's default setting. It employs a collection of piecewise linear equations with varying slopes between change points. When linear growth is utilized, the growth term will resemble the traditional $g(t) = mt + b$ from middle school, with the exception that the *slope*(m) and *offset*(b) are variable and will change value at each changepoint.

5. EXPERIMENTAL SETUP

In this section we discuss implementation of the SARIMA, XGBoost, LSTM and Prophet methods for comparison.

5.1. Dataset

We use temperature data is retrieved every 3 hours each day from the stations in Vietnam presented in Table 1. The data span from 2014 to 2019 (i.e. 17528 values). Since we need a large data set without missing values, we do not use more recent data. Data from January 2014 to December 2017 are used as the training set, while data from January 2018 to December 2019 are used as the testing set.

Table 1. The temperature stations.

No.	Dataset name
1.	BAC QUANG
2.	QUANG HA
3.	TAM DAO
4.	CAT TIEN
5.	KY ANH
6.	A LUOI
7.	NAM DONG

5.2. Statistical evaluation of model performance

The predicted and observed temperature concentration values were compared using three measures of predict accuracy computed from the test datasets: the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), and the Coefficient Correlation (also called r).

The RMSE is defined as:

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

The MAE is as

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

The r is defined as

$$r = \frac{\sum_{i=1}^m (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^m (y_i - \bar{y})^2 \sum_{i=1}^m (\hat{y}_i - \bar{\hat{y}})^2}}$$

where $y_i, i = 1 \div m$ are observed values for the temperature, \bar{y} is their mean and $\hat{y}_i, i = 1 \div m$ are predict values, $\bar{\hat{y}}$ is their mean.

The strength of the linear relationship between the variables is indicated by the sign of r . When r is close to 1, the two variables have a strong linear connection. If r is close to zero, the two variables have no linear relationship. When r approaches -1 , the two variables exhibit a weak (negative) linear connection.

6. RESULTS AND ANALYSIS

The experimental results of our experiments and various research studies are presented and discussed in this section. Firstly, we evaluate the performance of four algorithms on various datasets and compare the performance of models. Secondly, a discussion is conducted to analyze the results obtained from evaluation metrics.

6.1. Performance comparison of four algorithms

In this subsection, the performance of these 4 algorithms is examined using 3 different evaluation metrics. It can be seen from the table 2, 3, 4 that the RMSE, MAE and r values of LSTM are all better than the other three models. If we use RMSE as the evaluation criterion, the LSTM model has the best performance, followed by XGBoost, SARIMA and Prophet. Similarly, the MAE value of LSTM is smaller than that of XGBoost, SARIMA and Prophet. It is well-known that the MAE is less sensitive to outliers than the RMSE. That may explain why the trends of MAE and RMSE are analogous in four models. Figure 3 to figure 9 show the graphs of the prediction of LSTM model and the true values for 7 datasets. By comparing the experimental results of the four models, the LSTM method is more efficient for temperatures prediction so it is a good alternative to existing models for temperatures forecasting.

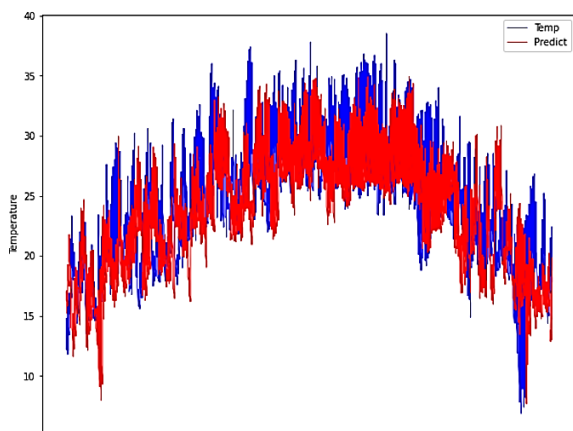


Figure 3. Temperature forecasting for BAC QUANG station.

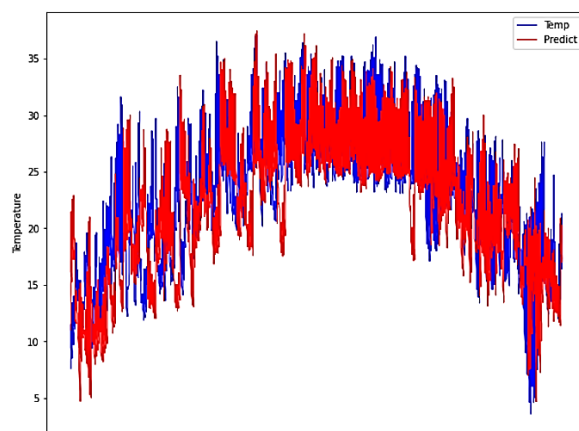


Figure 4. Temperature forecasting for QUANG HA station.

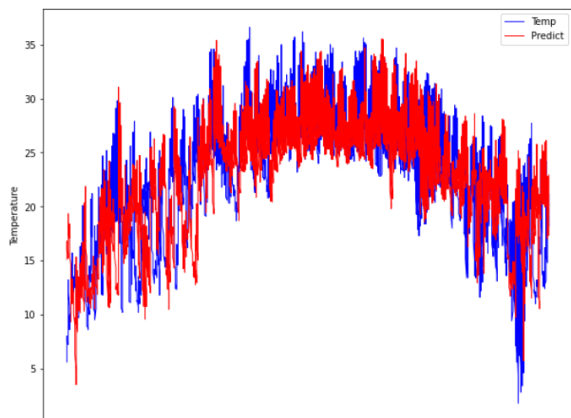


Figure 5. Temperature forecasting for TAM DAO station.

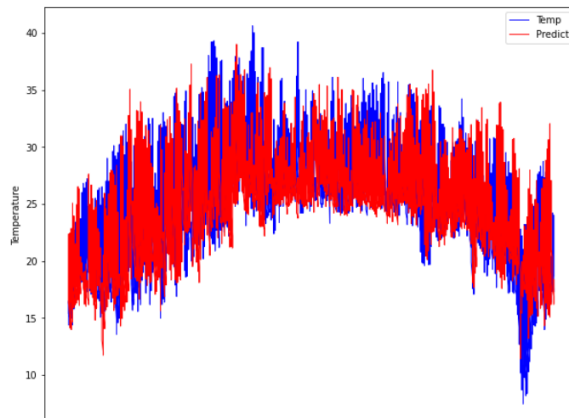


Figure 6. Temperature forecasting for CAT TIEN station.

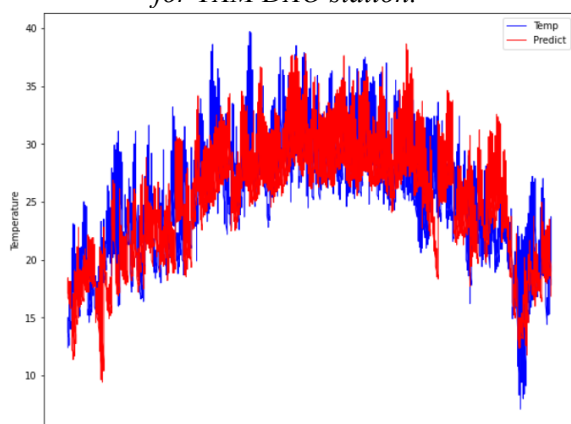


Figure 7. Temperature forecasting for KY ANH station.

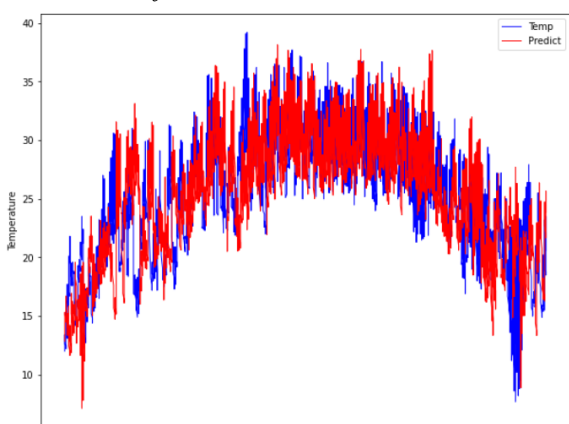


Figure 8. Temperature forecasting for A LUOI station.

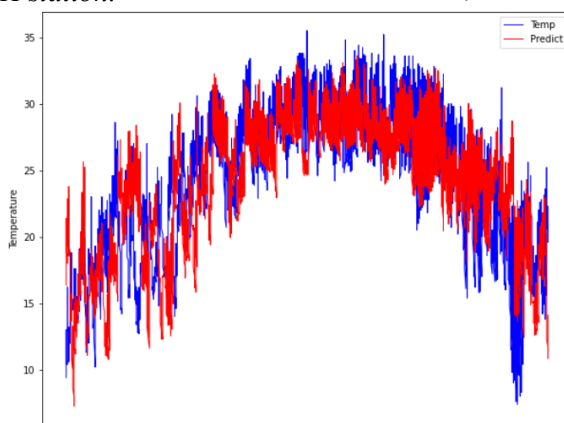


Figure 9. Temperature forecasting for NAM DONG station.

6.2. Contributions to literature

The research contributions of this study are significant. First, this is the study that has compared the performances of more different forecasting models, such as classical and advanced techniques applied on a real-world dataset. Second, we have demonstrated the ability of deep learning models such as LSTM on seasonal time-series forecasting. Third, we have contributed to finding how competitive are neural networks time-series forecasting compared with traditional

univariate methods. After performing a rigorous evaluation, the superior performance of LSTM was confirmed on 7 temperature datasets. Finally, this study has proven that XGBoost model are quite good candidates for predicting the future patterns of seasonal items.

Table 2. The comparison RMSE values of four models.

No.	SARIMAX	XGBoost	LSTM	Prophet
1	7.9576	4.8785	3.2511	17.1160
2	8.4006	4.7651	4.1366	20.2354
3	8.4481	5.2967	3.1249	18.6728
4	6.5694	4.9958	2.8958	16.0360
5	8.9642	5.2283	4.0031	21.0852
6	8.2748	4.9890	3.7674	24.5881
7	9.2046	4.8036	3.4504	23.8692

Table 3. The MAE value of four models.

No.	SARIMAX	XGBoost	LSTM	Prophet
1	6.9197	3.8442	2.5207	13.486
2	7.2924	3.6324	3.0303	17.326
3	7.4073	4.3008	2.4130	16.0529
4	5.6399	3.8031	2.1204	13.7714
5	7.8837	4.2629	2.9848	18.0842
6	7.2499	4.0467	2.7862	21.3880
7	8.1137	4.1042	2.5220	20.7011

Table 4. The r value of 4 models.

No.	SARIMAX	XGBoost	LSTM	Prophet
1	37.2%	72.2%	80.1%	-12.8%
2	37.1%	71.6%	76.5%	-7.6%
3	37.1%	72.4%	81.5%	-13.9%
4	51.5%	65.1%	81.1%	-15.2%
5	32.3%	73.0%	76.6%	-9.2%
6	31.8%	72.4%	73.0%	-12.7%
7	27.4%	79.1%	80.1%	-15.9%

6.3. Implications for practice

The derived results of this research can be applied in practice to forecast future temperature of any day with a similar seasonal pattern. Moreover, by performing the necessary data pre-processing and tuning the hyperparameters, the provided models can be applied to many different datasets. This has significant implications for advanced strategic decision-making and planning.

7. CONCLUSIONS AND FUTURE WORKS

Time-series forecasting objective is more complex to learn as it requires modeling approaches to associate the correlation factor with time and simultaneously with past observations. A good forecasting model should be able to observe both trends as well as seasonality in data. In this work, we provided a comprehensive review of 4-time series forecasting techniques for temperature data. These techniques are widely used in other fields, such as stock market prediction, energy consumption forecasting, sales forecasting, etc. We evaluated the above-listed set of representative forecasting techniques on seven different time series data-sets and logged our observations as a summary.

In summary, two models (in four models) including LSTM and XGBoost models, are capable of forecasting temperature in the dataset within 7 stations. It should be noted that temperature

range between 5 and 35. As a result, a model with an RMSE of approximately 3.5 can be considered satisfactory.

The outcomes of this article are dependent on the dataset used. The same models may be applied to other seasonal time series. However, the parameters must still be tuned in order to get the best forecasting model. Future study should investigate further prospective research directions for this work. More complicated LSTM and deep learning models can be tested to enhance the outcomes. In addition, multivariate time-series forecasting can be explored. Furthermore, constructing hybrid models that combine classical and current forecasting approaches might be beneficial and a promising direction for future study.

Acknowledgement: *The research presented in this paper was supplied data by Vietnam National Centre for Hydro - Meteorological Forecasting Center Hanoi, Vietnam.*

REFERENCES

- [1]. IFRC, "World Disaster Report 2020," International Federation of Red Cross and Red Crescent Societies, (2020).
- [2]. N. L. Duc et. al., "Research and quantitative rainfall forecasting from HRM and GSM model products," Vietnam Journal of Hydrometeorology, vol. 592, pp. 17-27, (2010).
- [3]. G. P. Z. a. D. M. Kline, "Quarterly Time-Series Forecasting With Neural Networks," IEEE Transactions on Neural Networks, vol. 18, pp. 1800-1814, (2007).
- [4]. S. M. P. Chopra, "Supply Chain Management". Strategy, Planning & Operation, German: Gabler, (2007).
- [5]. K. H. M. J. B. A. Ahmad Hasan Nury, "Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh," Journal of King Saud University - Science, vol. 29, no. 1, pp. 47-61, (2017).
- [6]. C. N. B. a. B. E. Reddy, "Predictive data mining on Average Global Temperature using variants of ARIMA models," in IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012), (2012).
- [7]. X. L. W. Q. Y. Z. W. Z. C. T. H. Yang, "A network traffic forecasting method based on SA optimized ARIMA-BP neural network," Computer Networks, vol. 193, (2021).
- [8]. S. R. H. P. S. S. P. C. Tandon, "How can we predict the impact of the social media messages on the value of cryptocurrency? Insights from big data analytics," International Journal of Information Management Data Insights, vol. 1, no. 2, (2021).
- [9]. R. L. Carey Goh, "Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention," Tourism Management, vol. 22, no. 5, pp. 499-510, (2002).
- [10]. L. A. H. Billy M. Williams, "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results," Journal of Transportation Engineering, vol. 129, no. 6, pp. 664-672, (2003).
- [11]. C. Hamzaçebi, "Improving artificial neural networks' performance in seasonal time series forecasting," Information Sciences, vol. 178, no. 23, pp. 4550-4559, (2008).
- [12]. T. Miyano and F. Girosi, "Forecasting Global Temperature Variations by Neural Networks," Cambridge Artificial Intelligence Laboratory: Cambridge, MA., USA, (1994).
- [13]. H. Hippert, C. Pedreira and R. Souza, "Combining neural networks and ARIMA models for hourly temperature forecast," in The IEEE-INNS-ENNS International Joint Conference on Neural Networks, Como, Italy, (2000).
- [14]. I. Tasadduq, S. Rehman and Bubshait, "Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia," Renew. Energy, vol. 11, pp. 71-77, (2001).
- [15]. Maqsood, M. Khan and A. Abraham, "An ensemble of neural networks for weather forecasting," Neural Computing Application, vol. 13, pp. 112-122, (2004).
- [16]. B. Smith, R. McClendon and G. Hoogenboom, "Improving air temperature prediction with artificial neural networks," International Journal Computing Intelligent, vol. 3, pp. 179-186, (2006).
- [17]. K. Abhishek, M. Singh, S. Ghosh and A. Anand, "Weather Forecasting Model using Artificial Neural Network," Procedia Technology, pp. 311-318, (2012).
- [18]. H. Mori and D. Kanaoka, "Application of support vector regression to temperature forecasting for short-term load forecasting," in Proceedings of the 2007 International Joint Conference on Neural

- Networks, Orlando, FL, USA, (2017).
- [19]. E. Ortiz-García, S. Salcedo-Sanz, C. Casanova-Mateo, A. Paniagua-Tineo and J. Portilla-Figueras, "Accurate local very short-term temperature prediction based on synoptic situation Support Vector Regression banks," Atmos. Res., (2012).
- [20]. R. Chevalier, G. Hoogenboom, R. McClendon and J. Paz, "Support vector regression with reduced training sets for air temperature prediction: A comparison with artificial neural networks," Neural Computing Application, vol. 20, pp. 151-159, (2010).
- [21]. A. Paniagua-Tineo, S. Salcedo-Sanz, C. Casanova-Mateo, E. Ortiz-García, M. Cony and E. Hernández-Martín, "Prediction of daily maximum temperature using a support vector regression algorithm," Renew. Energy, vol. 36, pp. 3054-3060, (2011).
- [22]. I. Abubakar, H. Chiroma, A. Zeki and M. Uddin, "Utilising key climate element variability for the prediction of future climate change using a support vector machine model," Int. J. Glob. Warm., vol. 9, pp. 129-151, (2016).
- [23]. Z. Z. H. Wai Yan Nyein Naing, "Forecasting of monthly temperature variations using random forests," Asian Research Publishing Network Journal of Engineering and Applied Sciences, vol. 10, no. 21, pp. 10109-10112, (2015).
- [24]. P. Hewage, M. Trovati, E. Pereira and Behera, "A. Deep learning-based effective fine-grained weather forecasting," Pattern Analysis, pp. 1-24, (2020).
- [25]. I. Roesch and T. Günther, "Visualization of Neural Network Predictions for Weather Forecasting," Comput. Graph. Forum, pp. 209-220, (2018).
- [26]. S. J. a. L. B. Taylor, "Forecasting at Scale," The American Statistician, vol. 72, no. 1, pp. 37-45, (2018).
- [27]. K.-P. L. C.-S. L. P.-T. C. Ping-Feng Pai, "Time series forecasting by a seasonal support vector regression model," Expert Systems with Applications, vol. 37, no. 6, pp. 4261-4265, (2010).
- [28]. T. H. M. Jr, A. M. Pellegrini and R. H. Perlis, "Assessment of Time-Series Machine Learning Methods for Forecasting Hospital Discharge Volume," JAMA Network Open, (2018).
- [29]. A. S. K. S. S. a. U. H. Sandhiya, "Forecasting Website Traffic Using Prophet Time Series Model," International Research Journal of Multidisciplinary Technovation, vol. 1, no. 2, pp. 1-8, (2019).
- [30]. W. D. B. U. N. N. H. Christophorus Beneditto Aditya Satrio, "Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET," Procedia Computer Science, vol. 179, pp. 524-532, (2021).
- [31]. G. A. Rob J Hyndman, "Forecasting: principles and practice", OTexts, (2018).
- [32]. Y. B. J. Q. J. W. P. Y. S. W. Mingju Gong, "Gradient boosting machine for predicting return temperature of district heating system: A case study for residential buildings in Tianjin," Journal of Building Engineering, vol. 27, (2020).

TÓM TẮT

Dự báo chuỗi thời gian dựa trên học máy: Trường hợp nghiên cứu về dự báo nhiệt độ ở một số trạm Việt Nam

Trong những năm gần đây, sử dụng các mô hình học máy để dự báo ngày càng được quan tâm. Trong nghiên cứu này, chúng tôi sẽ sử dụng dữ liệu nhiệt độ ở các trạm của Việt Nam để dự báo nhiệt độ. Nghiên cứu có sử dụng một số mô hình dự báo. Đầu tiên, phương pháp dự báo theo chuỗi thời gian truyền thống như ARIMA theo mùa (SARIMA). Sau đó, các phương pháp tiếp cận phức tạp hơn như XGBoost, Encoder-Decoder và Prophet được sử dụng. Hiệu suất của các mô hình được so sánh với nhau sử dụng một số độ đo đánh giá độ chính xác (ví dụ: sai số toàn phương trung bình (RMSE) và sai số trung bình tuyệt đối (MAE)). Các kết quả đã khẳng định tính ưu việt của phương pháp học sâu so với các phương pháp khác.

Từ khóa: Học máy; Dự báo nhiệt độ; Học sâu; Chuỗi thời gian; SARIMA; XGBoost.