International Journal of Statistical Sciences *ISSN 1683-5603*
Vol. 26(1), March, 2026, pp 119-134
DOI: https://doi.org/10.3329/ijss.v26i1.88851

Predicting and Identifying the Finest Model for Forecasting Maximum Temperature in Rajshahi District using Machine Learning Approaches

Most. Shamole Akter [1,‡], Md. Maniruzzaman[2], Md. Ashiqur Rahman Shah[1], Md. Sifat Hossain[1] and Md. Monimul Huq[1,*]

[1]Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh
(Email: shamolestat@gmail.com; sifat.stat@gmail.com; ashique.shah.bd@gmail.com)

[2]Statistics Discipline, Khulna University, Khulna-9208, Bangladesh
(Email: monir.stat91@gmail.com)
‡ These authors contributed equally to this work.
* Correspondence should be addressed to Md. Monimul Huq
(Email: monimul@ru.ac.bd)

[Received January 10, 2026; Accepted February 28, 2026]

Abstract

Temperature plays a significant role in driving climate change. Climate change tries to distinguish changes within the cruel values of temperature arrangement over time. Compared to other locales the most extreme temperature of Rajshahi City is high in the summer season and low in the winter season. The main objective of this paper was to identify the best model for forecasting maximum temperature in Rajshahi. Firstly, we determined classical ARIMA model then applied seven machine learning based algorithms like as support vector machine(SVM), Bayesian Regularized Neural networks(BRNN), Neural Network(Nnet), Elastic Net(ENet), Least Absolute Shrinkage and Selection Operation(Lasso), Relevance Vector Machines With Linear Kernel(RVML) for forecasting temperature data. Machine learning (ML) based models were upgraded utilizing arbitrary look with tune length 1000 and bootstrap-based 10 time's cross-validation. In this paper, we considered the month-to-month greatest temperature forecast in Bangladesh's Rajshahi division. We utilized a set of classical and machine learning based approaches strategies to anticipate the most extreme temperature of the Rajshahi locale. Although other machine learning like ARIMA performed great too this result demonstrated that the ENet showed more proficient and way better than other models. The values of MAE, RMSE, MAPE and R^2 are 0.97, 1.29, 2.74 and 89%. Hence the comes about such as execution of the forecasting model may be progressed by utilizing ENet compared to routine ARIMA & other models, particularly for most extreme temperature expectations. This investigation can be useful for organizations mindful of checking climate and defining arrangements.

Keywords: Monthly Maximum Temperature, ARIMA, SVM, avNNet, ENet, BRNN, nnet, Lasso, RVMS.

AMS Subject Classification: Primary 62M10, 62M20; Secondary 62J07, 62J12, 62G08, 68T07, 68T05.

1. Introduction

The variation in global or regional climates over time is referred to as climate change. Climate change reflects a change in atmospheric variability or atmospheric conditions over decades to millions of years. Recently used the term climate change refers to current environment changes, which rise in average temperature referred to as global warming. Global climate change has already been a great impact on the environment. Glaciers are melting, oceans and animals ranges are shifting, summers are getting hotter, and winters are getting colder (Clark et al., 2003). Climate change is a problem for Bangladesh's economy and environment. Bangladesh's economy relies heavily on the weather, particularly in the agricultural and industrial sectors (Ayers et al., 2014). But due to, climate change it has adversely affected the economy of Bangladesh. Rajshahi city is one of the ten divisions of Bangladesh. Farakka Barrage in Rajshahi also plays an important role in climate change due to reduced Padma River flows. As a result, the weather in Rajshahi is getting worse day by day. Forecasting of climate time series represents one of the greatest crucial and difficult operational tasks for meteorological services around the world. Accurate forecasting of climate time series can be critical in combating the negative effects of climate change in Bangladesh. Because time series forecasting is critical, it is critical that the appropriate model be fitted to the underlying climate time series. The classical method depends on Autoregressive Integrated Moving Average (ARIMA) model which invented by Box and a Jenkins (1976) is commonly utilized since the resulting model is straightforward to comprehend and interpret. So over the years, ARIMA has been routinely employed to predict climatic variables (Nanda et al., 2013; Sultana and Hasan, 2015; Doulah, 2018). Kaushik research group forecasted monthly rainfall and temperature data using the seasonal ARIMA model (Kaushik and Singh, 2008). The findings demonstrated that the seasonal ARIMA model delivers accurate monthly rainfall and temperature forecasts. Another researcher examined maximum mean temperature data collected from six regions in Bangladesh using the ARIMA model. Rajshahi, Dhaka, Khulna, Chittagong Barisal, and Sylhet had the highest monthly temperatures (Syeda, 2013).

Machine learning techniques for time series modeling have gotten much attention in recent years. Machine learning techniques can identify complex non-linear relationships. Machine learning approaches has been widely applied in a multitude of time series forecasting applications, particularly for climatic variables such as rainfall, temperature, humidity and wind forecasting. The outcomes showed that they significantly outperform better than classical approaches**.** The ANN model was to determine the maximum and minimum summer temperatures over India (De and Debnath, 2009). They find the statistics for these three months from 1901 to 2003. Use a Multilayer Perceptron Neural Network (MLPNN) to forecast the average monthly temperature increase India during the monsoon months. As a result, models for maximum and lowest temperature were developed by each monsoon month from 1901 to 2003. Consequently, in all situations, the forecast error was less than 5%. Another researcher presented a method for analyzing and forecasting temperature and humidity values for the future using a clustering technique (Badhiye, Chatur and Wakode, 2012). The K-nearest neighbor (K-NN) procedure was used in the study to forecast the values of climate temperature and relative humidity parameters. The authors claimed that the temperature and humidity forecasts were nearly 100% accurate. The authors used artificial neural network for time series forecast (Khashei and Bijari, 2011). They showed that artificial neural networks (ANNs) is an excellent, general-purpose solution for high-accuracy pattern classification, recognition, grouping, and in particular, time series prediction. The researcher proposed a model for support vector approach to forecast average weather temperature of Bangladesh (Shafin, 2019). Also, another authors proposed a wind speed prediction model based on machine learning methods (Mohandes et al., 2004). The authors developed an algorithm of neural network which is called support vector machine (SVM) to forecast wind speed and compared its performance to that of their multi-layer perceptron (MLP). They discovered that for all systems, SVM outperforms MLP. In forecasting, time series data of water consumption was compared the forecasting performance of the ANN and seasonal ARIMA models (BuHamra, Smaoui and Gabr, 2003). The outcomes demonstrated that the model of ANN performed better than the seasonal ARIMA model. In Iran, The authors compared SVM and ARIMA in Daily River flow forecasting (Uzundumlu, Oksuz and Kurtoglu, 2018). The performance of the models suggested that the flow could be easily forecasted using SVM from available data. Another model was evaluated, the ARIMA, ANN, and SVM forecasting approaches for time series data (Kim, 2003). The SVM model was a more effective method than that of the ANN & ARIMA model. The authors conducted a comparison of ANN, MLR, and ARIMA wind speed forecasting techniques (Arzu, M. Rafiuddin Ahmed and Khan, 2020). In comparison to the regression and ARIMA models, the ANN model was identified to be more workable and accurate in forecasting wind speed.

Although machine learning methods have attracted increasing attention and a substantial amount of research has been carried out, no significant research works have been found in Rajshahi in modeling and more accurately forecasting climatic time series variables using these methods. In many research used machine learning based models but for the first time, an attempt has been made in this study to Enet machine learning model performance better for modeling and forecasting climatic variables in Rajshahi along with the commonly used classical models ARIMA. The comparison of the models and using the most efficient model is very important to have more accuracy in forecasting climatic variables.

2. Materials and Methods

2.1.1 Data Description

We used secondary data to predict and model climate variables in Rajshahi. For this research, we used monthly maximum temperature data collected from the Bangladesh Meteorological Department (BMD). The variable is available from May 1982 to September 2024 in the selected area.

2.1.2 Graphical Representation

Graphical representations are a very well-informed and effective technique for determining whether the data in time series are stationary. If the data series shows a strong upward trend or a gradually fluctuating trend, it means the data series will be non-stationary. On the other hand, a series can be non-stationary without showing a persistent upward or downward trend. If the series is non-stationary, we need to take the difference to check stationarity.

2.1.3 Correlogram Test

A Correlogram generally called an Auto Correlation Function (ACF) graph is a graphical representation of a time series in data that changes over time. This Correlogram depicts the auto-correlation among data sets at various points in time. It can be used to test the unpredictability of data collection by producing auto-correlations given sample values collected at various time delays. It can be defined as

$$\eta_k = \frac{\sum_{t=1}^{n}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}; \text{for } k = 1,2,3 \ldots \ldots \ldots \ldots \ldots \quad (1)$$

2.1.4 Unit Root Test

The test of unit root is now the most popular technique for determining whether something is stationary or non-stationary. This unit root test used an autoregressive model to determine if a time-series variable was non-stationary. Augmented Dickey-Fuller (ADF) test, Phillip-Perron (PP) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are several tests used to assess the stationary of the variables (Dickey and Fuller, 1981; Kwiatkowski et al., 1992). Augmented Dickey-Fuller and Phillip-Perron justified the unit root test but in the absence of a unit root, this model was considered a random walk that is non - stationary at level. In the lack of such a null hypothesis, ADF and PP are parametric and non-parametric tests, but both are stationary. The test of Kwiatkowski-Phillips-Schmidt-Shin is generally used to determine stationarity, although the null hypothesis differs from the previous two. Feedback is always based on the findings of a significant number of tests.

2.1.5 Auto Regressive Integrated Moving Average (ARIMA) Model

The (Geurts, Box, and Jenkins, 1976) approach is a widely utilized Autoregressive Integrated Moving Average (ARIMA) model in time series analysis that can identify the appropriate models to generate forecasts. Analyzing time series data, the majority of that series deals with which are non-stationary and indicates that the random process's fundamental features fluctuate with time. In this situation, we need to create a model describing a non-stationary sequence that is transformed into a stationary series just an absolute difference one or many times. If $\Delta^d y_t$ is a stationary series, we say y_t is homogeneous non-stationary of order d and y_t is indeed an autoregressive integrated moving average procedure in orders p, d, and q, also known as ARIMA (p, d, q). The method of ARIMA (p, d, and q) procedure is defined in its general form as follows:

$$\nabla^d y_t = \varphi_1 \nabla^d y_{t-1} + \varphi_2 \nabla^d y_{t-2} + \cdots + \varphi_p \nabla^d y_{t-p} + \epsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (2)$$

Where p is an autoregressive process, d is integrated, and q is moving average parts of the model.

2.1.6 Box- Jenkins Methodology

The (Geurts, Box and Jenkins, 1976) approach is such a collection of processes of discovering and forecasting time series models with an autoregressive integrated moving average (ARIMA) class. In 1976, the Autoregressive Integrated Moving Average (ARIMA) model was developed by the well-known statisticians George Box and Gwilym Jenkins to determine the most suitable time series models for forecasting. Box-Jenkins demonstrates how to include seasonal components in SARIMA models. In summary, the following are the major steps in developing a Box-Jenkins forecasting model**.**

2.2.1 Model Identification

At first, we must ascertain if the data series is steady as well as whether substantial seasonality must be represented. The rank of both p and q from ACF and PACF represent the essential components for choosing ARIMA models. The ACF determines the level of linear dependency between time series measurements spread by just a lag k revealing the series stable level. The ACF and PACF aid components are needed to detect a number of the following criteria: time delays when large correlations arise, seasonality of such a series, as well as trend in series averages or variances.

2.2.2 Estimation

After determining the suitable p and q quantities, the ARIMA model parameters must be estimated. An ARIMA model's parameters can be calculated using the ML approach. Other approaches, such as the conditional LS and the unconditional LS, are available. If the outliers are believed to be independently normal, the ML technique is equal to the LS method inside this event of such an AR model.

2.2.3 Diagnostic checking

Our next target is to determine whether the model appropriately matches the data because some other ARIMA model might perform best. The validation process is crucial in two ways: confirming that the model's error terms are random and determining whether the estimated coefficients are statistically significant. The purpose of this method of modeling is really to make the smallest model feasible. Whereas if the model accurately describes the data, we may accept it; alternatively, we must go back to the identification step and choose the most suitable model.

2.2.4 Forecasting

One of the most important evaluations of a model's effectiveness has always been its capacity to forecast. Without samples, projections provide a more accurate picture of how efficiently the model is working in general. We should retain some data after the sample without sample projections. The model with the minimum predicting accuracy could be needed for the data.

2.3 Some Machine Learning Based Algorithms

2.3.1 Support Vector Machine (SVM) Approach

Support vector machines (SVMs) are based on the core idea that in order to achieve classification performance, a categorization hyper plane must be found that maximizes the empty space through both sides of a higher dimensional space. Support vector regression (SVR) is the term used when SVMs are employed to solve regression issues (Vapnik, Golowich and Smola, 1997; Basak, Pavlova and Shapiro, 2007). The formula for the linear function f(x) is

$$y(x) = w^t \varphi(x) + h \tag{3}$$

Where the high-dimensional feature spaces are demonstrated by $\varphi(x)$ that from the input space, x is nonlinearly mapped. Those analyses will be performed using only a kernel function that must gratify Mercer's situation (Flake and Lawrence, 2002). The SVM kernel function and parameter tuning are critical to improve predictive performance. Now at moment, it has a wide range of kernel functions accessible (Schölkopf, 2002). Therefore, selecting the best kernel trick for a given task has continued to be difficult. Three commonly used kernel functions (linear, polynomial, and radial basis function) were employed to build the SVM algorithm in this work (here, d = 1 for polynomial, sigma = 1.5 for radial basis function, C =2(2:9), epsilon=seq (0, 1, 0.5)). (Vapnik, Golowich and Smola, 1997; Schölkopf, 2002) provide excellent introductions to SVM. In this study the main tuning parameter, known as the radial basis fraction was C = 0.03135127, provide excellent result for predictive performance.

2.3.2 Averaged Neural Network (avNNet) Approach

The Model Averaged Neural Network (avNNet) methodology used to forecast indoor temperature including the number of size (7) and weight decay (3) evaluate the performance by (Alawadi et al., 2022). Authors (Appelhans et al., 2015) used fourteen machine learning algorithms were applied to predict spatial temperature patterns and then assessed against the widely-used kriging method. Utilizing a 10-fold cross-validation approach, regression tree models typically outperformed both linear and non-linear regression models. Among these, the stochastic gradient boosting model demonstrated the highest individual performance, followed by Cubist, random forest, and model-averaged neural networks; with the exception of the neural networks, all are regression tree-based algorithms. Although these machine learning models outperformed kriging in quantitative evaluations, the visual clarity of the resulting air temperature maps remains inconclusive.
Four machine learning algorithms like as random forests (RF), neural networks (Nnet), averaged neural networks (avNNet) and support vector machines (SVM) was used (Meyer et al., 2016) for rainfall area detection and rainfall rate assignment using MSG SEVIRI data over Germany. But they didn't provide suitable model for rainfall.

Averaged Neural Network (avNNet) involves training multiple neural networks on bagged subsets of the data and averaging their predictions to reduce variance and improve stability. In our paper, we used key hyperparameters including the number of hidden units (size=12), weight decay (decay=0.0169014), and bagging (bag) are tuned using cross-validation to optimize model performance.

2.3.3 Bayesian Regularized Neural Network (BRNN) Approach

In this study, a Bayesian Regularized Neural Network (BRNN) was employed for regression using the brnn package in R. Authors (Ye et al., 2021) applied BRNN for predicting daily rainfall from sea level pressure data and evaluate the best model for predicting the performance. BRNN and traditional statistic-based data-driven fault detection and identification methods was used by (Sun et al., 2020) in some manufacturing industries. But the first time, The BRNN-based approach was used for predicting the maximum temperature in Rajshahi. It provides uncertainty estimates that enable concurrent fault detection in chemical processes, direct identification of faults, and analysis of fault propagation. This method set to method = 'brnn', adds regularization to reduce over fitting and improve generalization. The primary tuning parameter was the number of neurons (neurons=1).

2.3.4 Neural Network (Nnet) Approach

The Neural Network model (Nnet method) trains a single-layer neural network for classification or regression by adjusting weights to minimize prediction error. To detect rainfall areas and estimate rainfall rates over Germany, four machine learning algorithms random forests (RF), neural networks (Nnet), averaged neural networks (avNNet), and support vector machines (SVM) were applied using MSG SEVIRI data by (Appelhans et al., 2015). However, these algorithms did not produce a model that could accurately predict rainfall. The paper suggested the model's key hyperparameters number of hidden units (size=6) and weight decay (decay=0.02624695) are tuned through cross-validation to balance complexity and prevent over fitting.

2.3.5 Elastic Net (ENet) Approach

The Elastic Net model (enet method) combines L1 (lasso) and L2 (ridge) regularization for regression, balancing variable selection and regularization. At present many researcher usage Enet algorithms in several sectors. A novel elastic net with regression coefficients method (Enet-BETA) has been introduced to identify key variables in the spectrum data created by the authors (Liu and Li, 2017). This Enet-BETA approach not only isolates essential variables to facilitate clear interpretation of quality but also enhances the stability and practicality of the developed model. The proposed key tuning parameters include fraction= 0.7573842 (which controls the proportion between lasso and ridge) and lambda= 0.08222577 (which determines the strength of regularization).

2.3.6 Least Absolute Shrinkage and Selection Operator (Lasso) Approach

The Lasso regression approach employs L1 regularization to encourage sparse solutions by reducing the less significant coefficients to zero. The authors (Wanishsakpong and Notodiputro, 2024) used Ridge Regression and Lasso techniques to compare which model was best to predict daily maximum temperature in Thailand. Ridge regression outperformed Lasso regression when assessing Mean Squared Error (MSE) and the coefficient of determination (R^2). The Ridge regression model indicated that daily maximum temperatures were significantly influenced by factors such as the average, maximum, and minimum humidity levels; the previous day's minimum humidity; the average wind speed; the previous day's average and maximum wind speeds; as well as the average, maximum, and minimum solar radiation, including these values from the previous day. In this study the main tuning parameter, known as the fraction was 0.6525013, which adjusts the level of regularization applied, achieving a balance between model simplicity and predictive performance.

2.3.7 Relevance Vector Machine (RVMLinear) Approach

The Relevance Vector Machine with a Linear Kernel (rvmLinear method) is a regression model that applies Bayesian inference to achieve sparse solutions, similar to support vector machines but with probabilistic predictions. It requires no tuning parameters, making it straightforward to implement. This model effectively identifies key data points (relevance vectors), enhancing interpretability while maintaining high predictive accuracy on linear regression tasks. For predicting air temperature a data drive model was developed by (Patel and Joshi, 2023). This study explored the air temperature prediction capabilities of various data-driven modeling frameworks, including static models such as artificial neural networks (ANN), Gaussian process regression (GPR), support vector regression (SVR), relevance vector machines (RVM), linear regression, and regression trees, as well as transient models like long-short term memory (LSTM) networks and nonlinear autoregressive neural networks with external input (NARX). In the static analysis, GPR yielded the highest accuracy, with an average error of 0.56 °C, followed closely by ANN and SVR, which had average errors of 0.60 °C and 0.68 °C, respectively. In the transient analysis using experimental data, NARX outperformed LSTM for standard operations, with average errors of 0.83 °C and 1.07 °C, respectively. The results suggest that data-driven models can still achieve reliable predictions even when input variables deviate slightly from the training range.

2.4 Model Selection Criteria

Modeling criteria are extremely important for accurate forecasting. There are numerous algorithms available for selecting the best model. Choosing the best algorithm is a difficult question. The main recommendation is to initially establish that the measure of forecasting error is most suited to the particular circumstances at stake. We choose the best model also with the lowest Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) values (Islam and Rahman, 2018). They can be recognized by

$$\mathrm{MAE} = \frac{1}{\mathrm{N}}\sum_{1}^{\mathrm{N}}|\mathrm{y}_{\mathrm{real}} - \mathrm{y}_{\mathrm{forecast}}| \tag{4}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{1}^{\mathrm{N}}(\mathrm{y}_{\mathrm{real}} - \mathrm{y}_{\mathrm{forecast}})^2}{\mathrm{N}}} \tag{5}$$

$$\mathrm{MAPE} = \frac{1}{\mathrm{N}}\sum_{1}^{\mathrm{N}}\left|\frac{(\mathrm{y}_{\mathrm{real}} - \mathrm{y}_{\mathrm{forecast}})}{\mathrm{y}_{\mathrm{real}}}\right| * 100 \tag{6}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{1}^{\mathrm{N}}(\mathrm{y}_{\mathrm{real}} - \mathrm{y}_{\mathrm{forecast}})^2}{\sum_{1}^{\mathrm{N}}(\mathrm{y}_{\mathrm{real}} - \bar{\mathrm{y}})^2} \tag{7}$$

3. Results and discussion

3.1 Forecasting ability of ARIMA Model

Using Box-Jenkins' (1976) famous modeling philosophy, we chose the best ARIMA model for monthly maximum temperature data. For forecasting purposes, the monthly maximum temperature series is split into two sections. The training series of 70% is the initial part and the second section, 30% is the test series. First, we created a line graph using the monthly maximum temperature data. We have seen from the time series graphic that the monthly maximum temperature exhibits seasonality with a seasonality of 1 year in the data set. The seasonal pattern is moving through a constant value (Figure 1).

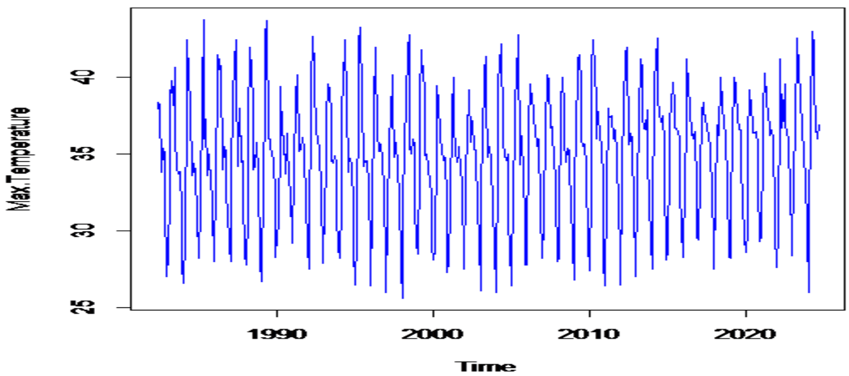

Figure 1: Plot of Monthly Maximum Temperature in Rajshahi.

To test the stationarity of the data, we used the Correlogram test. We used the Correlogram with lag length 24 as follows (Figure 2).We found that the ACF had substantial spikes at numerous delays because the dataset incorporates seasonal variations. We have discovered that the PACF had large spikes at certain delays. To eliminate seasonal variance, the underlying data set must have seasonal differencing.

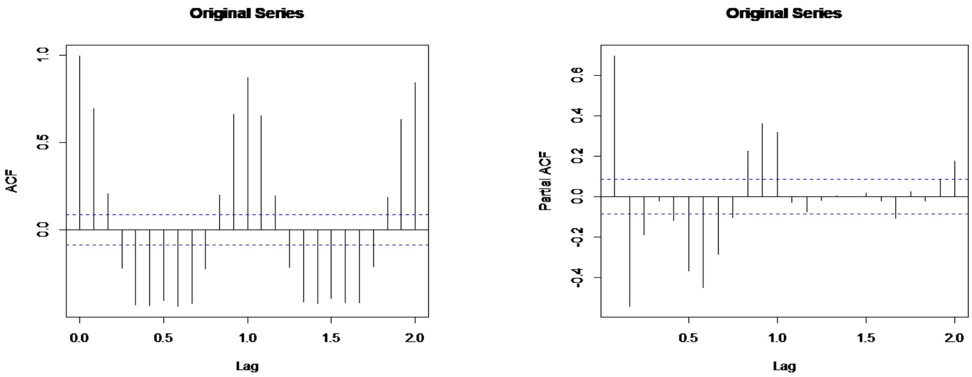

Figure 2: Plot of ACF and PACF of Monthly Maximum Temperature in Rajshahi.

Following the seasonal difference, we present the graph is almost stationary. Also, we construct seasonal autocorrelation (SAC) and seasonal partial autocorrelation (SPAC) for various lags.

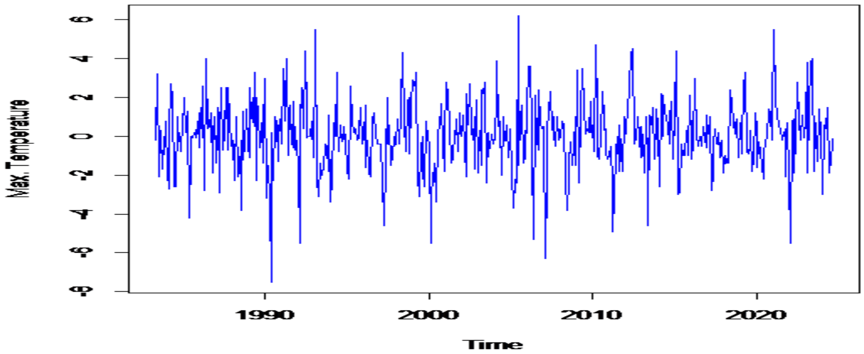

Figure 3: Plot of Seasonal Differencing Monthly Maximum Temperature in Rajshahi.

Therefore, the ACF and PACF of seasonally difference series support the series to be stationary as the AC and PAC values at the lags are within the 95% confidence limits.

Furthermore, we have used basic unit root tests to determine whether the de-seasonalized series was stationary. The Augmented Dicky-Fuller (ADF) test, created by (Dickey and Fuller, 1981), the Phillips-Perron (PP) test, introduced by Phillips and Perron in 1988, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are the most widely used unit root tests, developed by (Kwiatkowski et al., 1992). The Augmented Dicky-Fuller (ADF) test value is -7.22, the p-value being 0.01; the Phillips-Perron (PP) test value is -18.45 with a p-value of 0.01 and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test value is 0.015 with a p-value of 0.01, appropriately mentioned that the observation series remains at the point the 5% level of statistical significance and suggests that seasonal lagged series don't have a unit root.

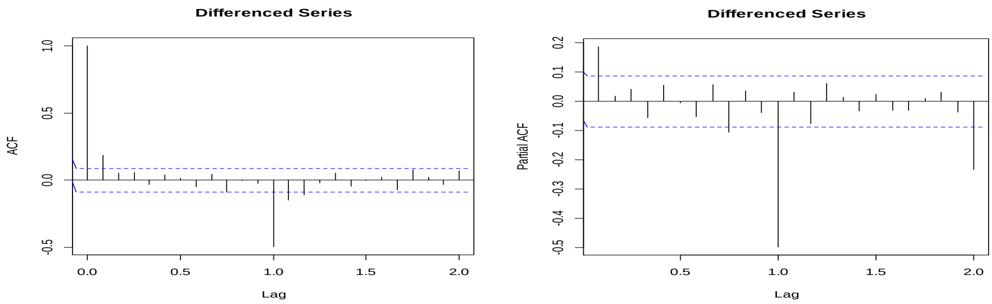

Figure 4: Seasonal Difference of Monthly Maximum Temperature data for ACF and PACF

From the Correlogram of SACF and SPACF, we detected a tentative seasonal ARIMA model by using the Box-Jenkins modeling philosophy. Five ARIMA models at different values of p,d, and q stand for non-seasonal, and P, D, and Q stand for seasonal. These models are shown in Table 1. When the model is analyzed, the test period is forecasted. After forecasting the data by using the forecasting values and actual values of the test series, we get the forecasting accuracy values MAE, RMSE, MAPE, and R^2. Out of these seven models, one is selected comparing the other models, and which values are minimum of the criteria MAE, RMSE, MAPE, and R^2. That model is the best.

Table 1: Tentative seasonal ARIMA models for Monthly Maximum Temperature.

Model	MAE	RMSE	MAPE	R^2
$(2,1,2)\ (2,1,0)_{12}$	1.058	1.367	2.997	88%
$(0,1,1)\ (0,1,1)_{12}$	1.117	1.381	3.145	87%
$(2,1,3)\ (1,1,1)_{12}$	1.138	1.398	3.202	87%
$(1,1,1)\ (1,1,1)_{12}$	1.167	1.425	3.278	87%
$(1,1,0)\ (1,1,0)_{12}$	1.174	1.431	3.297	87%
$(1,1,2)\ (2,1,1)_{12}$	1.110	1.456	3.121	86%

We observed that the value of the criterion for a model with bold numerals that model is the best model rather than other models concerning those criteria. Hence, the seasonal ARIMA $(2,1,2)(2,1,0)_{12}$ model like an accurate model for the average monthly maximum temperature in Rajshahi. The estimated seasonal ARIMA $(2,1,2)\ (2,1,0)_{12}$ model has been shown in Equation-8.

$$\nabla^d y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \theta_3 \varepsilon_{t-3} + \varphi_{12} y_{t-12} + \theta_{12} \varepsilon_{t-12} \qquad (8)$$

3.2 Diagnostic Checking for Estimated ARIMA Model

Diagnostic checking for an ARIMA (Autoregressive Integrated Moving Average) model involves visual and statistical analysis of the residuals (errors) to ensure the model's adequacy in capturing the underlying data structure. The diagnostic plots help identify issues like autocorrelation, heteroscedasticity, or non-normality in the residuals. Figure-5 showed a standardized residuals plot, ACF of residuals plot, Ljung-Box plot and Normal Q-Q plot of standardized residuals were used to check the diagnostic of the ARIMA $(2,1,2)\ (2,1,0)_{12}$ model.

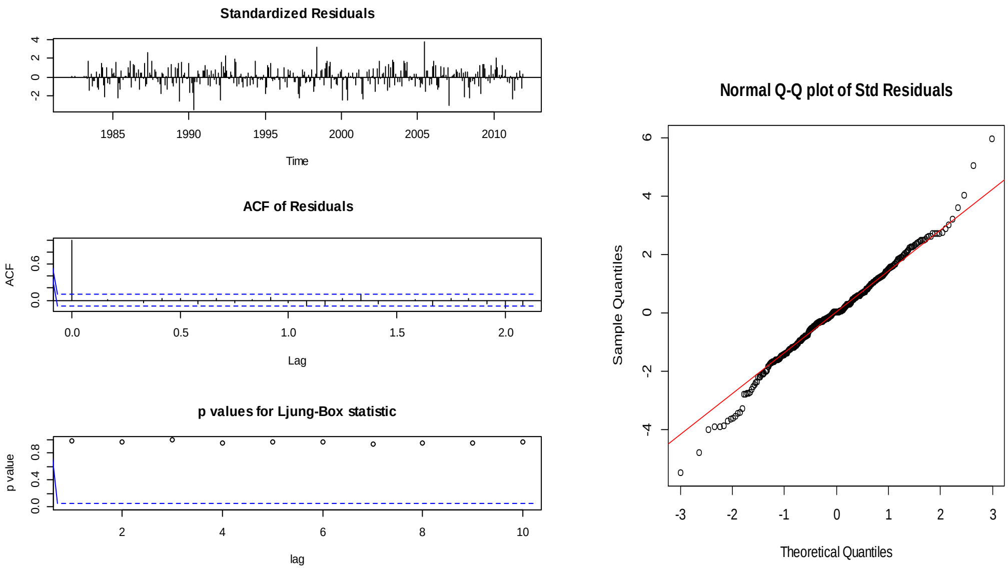

Figure 5: Diagnostic checking plot of estimated ARIMA model

3.3 Some Proposed ML-based Models

The primary distinction among the proposed ML-based models lies in their varying approaches to feature selection, learning algorithms, and adaptability to different data patterns are given in below table-2:

Table 2: Some MLbased model for Monthly Maximum Temperature data

Model	**MAE**	**RMSE**	**MAPE**	R^2
SVM Linear	1.03	1.39	2.90	87%
avNNet	1.03	1.36	2092	88%
BRNN	1.05	1.37	2.96	88%
nnet	1.01	1.34	2.85	88%
Lasso	0.98	1.31	2.78	88%
ENet	0.97	1.29	2.74	89%
RVMLinear	0.98	1.30	2.79	89%

The values of MAE, RMSE, MAPE and R^2 of ENet model are 0.97, 1.29, 2.74 and 89% respectively. So we decided to use the precise ENet model well through random search with tuneless 1000 and 10 times bootstrap cross-validation. Providing a lower value for predicting future reliability requirements signifies that the models believe they are improving.

3.4 Comparison of the forecasting ability among ARIMA & ENet

Now that we have the data, we can compare how well the forecasting models ARIMA & ENet performed. Table 3 displays the actual and forecasted values of the test series using the classical model ARIMA and machine learning base model ENet. Four forecasting accuracy metrics: MAE, RMSE, MAPE, and R^2 were used to assess the predicting capabilities of these models.

Table 3: Comparative performance among ARIMA & ENet models for Monthly Maximum Temperature data.

Model	MAE	RMSE	MAPE	R^2
ARIMA	1.058	1.367	2.997	88%
ENet	0.97	1.29	2.74	89%

Table 3 shows that the accuracy rate of the ENet model is lower than the ARIMA. So, we concluded that the ENet model outperforms the ARIMA models in terms of maximum temperature forecasting in Rajshahi. Finally, we believe ENet is the best model for predicting maximum temperatures in Rajshahi. Figure 5 shows that the ENet model forecasting values are very close to the actual series.

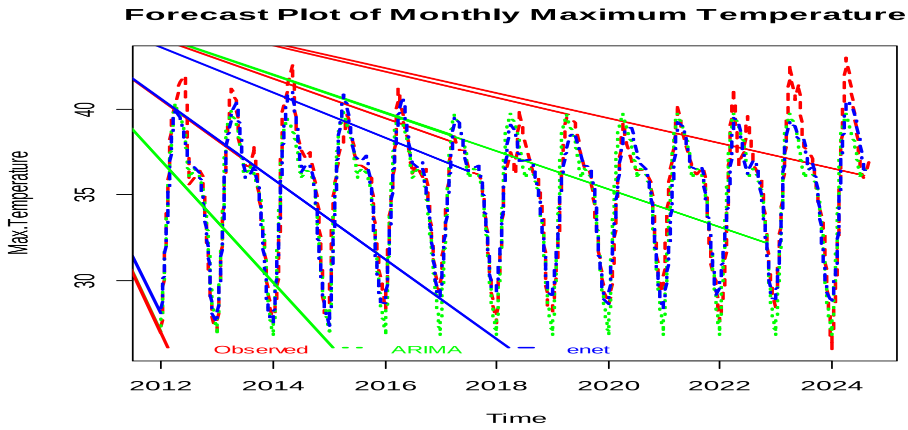

Figure 6: Actual series along with the forecasted series using ARIMA, ENet model for Monthly Maximum Temperature data.

3.5 Key difference between the proposed ML-based system and existing system for temperature prediction:

Several studies have explored the effectiveness of various classifiers for predicting different outcomes. (Arzu et al., 2020) reviewed several classifiers including MLR, ARIMA, and ANN, to predict wind speed using the University of the South Pacific dataset. They found that MLR achieved an MAE of 1.17, an RMSE of 0.85, a MAPE of 15.36, and an R^2 of 30%. As a complementary study, (Rangaraj et al., 2020) also used ARIMA, ANN, and LSTM to predict another outcome on the data of the NREL web portal. Here the LSTM achieved an MAE of 0.92, an RMSE of 1.24, and a MAPE of 13.07. (Arzu et al., 2018) used the University of the South Pacific dataset to evaluate a variety of classifiers such as MLR, ARIMA, and ANN. The MLR achieved an MAE of 1.17, an RMSE of 0.85, a MAPE of 15.36, and an R^2 of 30%. Another prediction models ANN and SARIMA was developed by (Cadenas and Rivera, 2006). They showed that the better performances MAE of 2.37, an RMSE of 8.28, and a MAPE of 28.50 were obtained by SARIMA. The prediction is done using comisio'n federal de electricidad data. (Ramesh et al., 2015) used ANN and GP based models to predict a specific outcome using data provided by the national center for environmental prediction. They show that ANN model achieved an MAE of 0.59, an RMSE of 0.81, a MAPE of 2.44, and an R^2 of 87% respectively. (Maldonado-Correa et al., 2021) used MLP, LSTM, CNN, and hybrid models to predict different outcomes with villonaco wind farm data. The hybrid model achieved an MAE of 0.14, an RMSE of 0.09, and a MAPE of 144.26. (Patil et al., 2022) used MLR to predict a different outcome using the meteorological data, achieving an MAE of 0.11, an RMSE of 0.18, and an R^2 of 69%. (Faniband et al., 2020) also used ANN based model to predict a specific outcome in Qaisumah that achieved RMSE and MAPE of 0.09 and 6.65, respectively. Finally, our study for 2024 using the weather data from the Bangladesh meteorological department aims to use a combination of ARIMA, SVM, avNNet, BRNN, Nnet, Lasso, ENet and RVMS based approaches to predict monthly maximum temperature of Rajshahi region in Bangladesh. Our current results showed that Enet based model achieved the highest value of R^2 (89%) compared to existing models.

Table 4: The Key difference between the proposed study and previous studies.

Author	Year	DS	Classifiers	MAE	RMSE	MAPE	R^2
Arzu et al.	2020	USP	**MLR**, ARIMA, ANN	1.17	0.85	15.36	30%
Rangaraj et al.	2020	NREL web porta	ARIMA, ANN, **LSTM**	0.92	1.24	13.07	--
Arzu et al.	2018	USP	MLR, ARIMA, **ANN**	1.48	1.19	29.73	55%
Cadenas and Rivera	2006	CFE	ANN, **SARIMA**	2.37	8.28	28.50	
Ramesh et al.	2015	NCEP	**ANN**, GP	0.59	0.81	2.44	87%
Maldonado-Correa et al.	2021	VWF	MLP, LSTM, CNN, **Hybrid**	0.14	0.09	144.26	--
Patil et al.	2022	MD	MLR	0.11	0.18	Inf.	69%
Faniband et al.	2020	Qaisumah	ANN	--	0.09	6.65	--
Proposed	2024	BMD	ARIMA, SVM, avNNet,BRNN, nnet, Lasso, **ENet**, RVMS, Lars	0.97	1.29	2.74	89%

The bold value indicates the proposed method results.

USP University of the south pacific, HTD historical temperature data, CFE comisio´n federal de electricidad, NCEP national centre for environmental prediction, VWF villonaco wind farm, MD meteorological data, BMD Bangladesh meteorological department, MLR multiple linear regression, LSTM long short-term memory, ETS error trend seasonality, KNN k-nearest neighborhood, RF random forest, GP gaussian process, MLP multi-layer perceptron, CNN convolutional neural network.

4. Conclusion and policy implications

4.1. Conclusion

This study compared maximum temperature forecasting performances among the classical models, ARIMA models and some recently developed ML-based models. The findings showed that Enet based obtained outstanding performances for forecasting monthly maximum temperature in other models. For all climatic variables, it was observed that the ENet model exhibits greater forecasting accuracy than the ARIMA model after using the error detection algorithms.

4.2 Recommendation

From the findings of the study and conclusion that this Enet based model forecasts monthly maximum temperatures more accurately than the ARIMA models, benefits, policymakers, and decision makers should use it to forecast. The weather forecasting authority should use machine learning methods rather than traditional methods for better and more accurate forecasting of climatic variables, thereby assisting in the prediction of any climatic disaster or hazard with reasonable accuracy, which has enormous implications in disaster preparedness or management.

Acknowledgement

The authors would like to extend their gratitude to the Bangladesh Meteorological Department (BMD) for supplying the data utilized in this study. The authors appreciate the reviewer's careful reading and thoughtful remarks that helped refine this work.

References

[1] Ahmadi, A. et al. (2014). Hybrid model for weather forecasting using ensemble of neural networks and mutual information, International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3774–3777. doi: 10.1109/IGARSS.2014.6947305.

[2] Alawadi, S. et al. (2022). A comparison of machine learning algorithms for forecasting indoor temperature in smart buildings, Energy Systems, 13(3), pp. 689–705. Available at: https://doi.org/10.1007/s12667-020-00376-x.

[3] Appelhans, T. et al. (2015). Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania, Spatial Statistics, 14, pp. 91–113. Available at: https://doi.org/10.1016/j.spasta.2015.05.008.

[4] Arzu, A., Ahmed, M. Rafiuddin and Khan, M.G.M. (2020). Wind Speed Forecasting using Regression, Time Series and Neural Network Models: A Case Study of Suva, Proceedings of the 22nd Australasian Fluid Mechanics Conference AFMC2020 [Preprint], (December). Available at: https://doi.org/10.14264/ccee311.

[5] Arzu, A., Ahmed, Mohammed R. and Khan, M.G.M. (2020). Wind Speed Forecasting using Regression, Time Series and Neural Network Models: A Case Study of Suva, Proceedings of the 22nd Australasian Fluid Mechanics Conference AFMC2020 [Preprint]. Available at: https://doi.org/10.14264/CCEE311.

[6] Ayers, J. M. et al. (2014). Mainstreaming climate change adaptation into development: A case study of Bangladesh, Wiley Interdisciplinary Reviews: Climate Change, 5(1), pp. 37–51. Available at: https://doi.org/10.1002/wcc.226.

[7] Badhiye, S. S., Chatur, P. N. and Wakode, B. V (2012). Temperature and Humidity Data Analysis for Future Value Prediction using Clustering Technique : An Approach, International Journal of Emerging Technology and Advanced Engineering, 2(1), pp. 88–91.

[8] Basak, S., Pavlova, A. and Shapiro, A. (2007). Optimal asset allocation and risk shifting in money management, Review of Financial Studies, 20(5), pp. 1583–1621. Available at: https://doi.org/10.1093/rfs/hhm026.

[9] BuHamra, S., Smaoui, N. and Gabr, M. (2003). The Box-Jenkins analysis and neural networks: Prediction and time series modelling, Applied Mathematical Modelling, 27(10), pp. 805–815. Available at: https://doi.org/10.1016/S0307-904X(03)00079-9.

[10] Clark, D. et al. (2003). Climate Change Climate change : 923(August), pp. 920–923. Available at: https://education.nationalgeographic.org/resource/climate-change.

[11] De, S. S. and Debnath, A. (2009). Artificial Neural Network Based Prediction of Maximum and Minimum Temperature in the Summer Monsoon Months over India, Applied Physics Research, 1(2), pp. 37–44. Available at: https://doi.org/10.5539/apr.v1n2p37.

[12] Dickey, D. A. and Fuller, W. A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, Econometrica, 49(4), p. 1057. Available at: https://doi.org/10.2307/1912517.

[13] Doulah, M.S.-U.- (2018). Forecasting Temperatures in Bangladesh: An Application of SARIMA Models, Environmental Science [Preprint], (November). Available at: https://www.semanticscholar.org/paper/Forecasting-Temperatures-in-Bangladesh%3A-An-of-Doulah/1e5c582fb704dbf0c173420641b5cb29abd3b2ee.

[14] Flake, G. W. and Lawrence, S. (2002). Efficient SVM regression training with SMO, achine Learning, 46(1–3), pp. 271–290. Available at: https://doi.org/10.1023/A:1012474916001.

[15] Hayati, M. and Mohebi, Z. (2007). Application of Artificial Neural Networks for Temperature Forecasting, International Journal of Electrical and Computer Engineering, 1(4), pp. 662–666.

[16] Hewage, P. et al. (2020). Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station, Soft Computing, 24(21), pp. 16453–16482. Available at: https://doi.org/10.1007/S00500-020-04954-0/FIGURES/17.

[17] Hossain, M. et al. (2015). Forecasting the weather of Nevada: A deep learning approach, Proceedings of the International Joint Conference on Neural Networks, 2015-Septe. Available at: https://doi.org/10.1109/IJCNN.2015.7280812.

[18] Islam, S. and Rahman, J. (2018). Comparative Study of ARIMA and Artificial Neural Network Methods for Temperature Forecasting in Rajshahi, pp. 77–84.

[19] Kaushik, I. and Singh, S. (2008). Seasonal ARIMA Model for Forecasting of Monthly Rainfall and Temperature, Journal of Environmental Research And Development, 3(2).

[20] Khashei, M. and Bijari, M. (2011). A new hybrid methodology for nonlinear time series forecasting, Modelling and Simulation in Engineering, 2011. Available at: https://doi.org/10.1155/2011/379121.

[21] Kim, K. J. (2003). Financial time series forecasting using support vector machines, Neurocomputing, 55(1–2), pp. 307–319. Available at: https://doi.org/10.1016/S0925-2312(03)00372-2.

[22] Kwiatkowski, D. et al. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root?, Journal of Econometrics, 54(1–3), pp. 159–178. Available at: https://doi.org/10.1016/0304-4076(92)90104-Y.

[23] Liu, W. and Li, Q. (2017). An efficient elastic net with regression coefficients method for variable selection of spectrum data, PLoS ONE, 12(2), pp. 1–13. Available at: https://doi.org/10.1371/journal.pone.0171122.

[24] Márquez-Vera, M. A. et al. (2021). Inverse fuzzy fault model for fault detection and isolation with least angle regression for variable selection, Computers & Industrial Engineering, 159, p. 107499. Available at: https://doi.org/https://doi.org/10.1016/j.cie.2021.107499.

[25] Meyer, H. et al. (2016). Comparison of four machine learning algorithms for their applicability in satellite-based optical rainfall retrievals, Atmospheric Research, 169, pp. 424–433. Available at: https://doi.org/10.1016/j.atmosres.2015.09.021.

[26] Mohandes, M. A. et al. (2004). Support vector machines for wind speed prediction, enewable Energy, 29(6), pp. 939–947. Available at: https://doi.org/10.1016/j.renene.2003.11.009.

[27] Nanda, S. K. et al. (2013). Prediction of Rainfall in India using Artificial Neural Network (ANN) Models, International Journal of Intelligent Systems and Applications, 5(12), pp. 1–22. Available at: https://doi.org/10.5815/ijisa.2013.12.01.

[28] Paras, S., Kumar, A. and Chandra, M. (2007). A feature based neural network model for weather forecasting, World Academy of Science, Engineering …, 4(3), pp. 209–216.

[29] Patel, D. and Joshi, Y. (no date). Data driven modeling advancements for air temperature predictions in data centers, Numerical Heat Transfer, Part B: Fundamentals, pp. 1–18. Available at: https://doi.org/10.1080/10407790.2023.2280745.

[30] Schölkopf, B. (2002). Slides- Learning with kernels, Journal of the Electrochemical Society, 129(November), p. 2865. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.167.5140&rep=rep1&type=pdf.

[31] Shafin, A. A. (2019). Machine Learning Approach to Forecast Average Weather Temperature of Bangladesh, Global Journal of Computer Science and Technology: D Neural & Artificial Intelligence, 19(3), p. pp.38-41.

[32] Sharaff, A. and Roy, S. R. (2018). Comparative Analysis of Temperature Prediction Using Regression Methods and Back Propagation Neural Network, Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018, pp. 739–742. Available at: https://doi.org/10.1109/ICOEI.2018.8553803.

[33] Shi, X. et al. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, Advances in Neural Information Processing Systems, 28.

[34] Sultana, N. and Hasan, M. M. (2015). Forecasting Temperature in the Coastal Area of Bay of Bengal-An Application of Box-Jenkins Seasonal ARIMA Model, Journal of Civil and Environmental Research, 7(8), pp. 149–160.

[35] Sun, W. et al. (2020). Fault detection and identification using Bayesian recurrent neural networks, Computers and Chemical Engineering, 141. Available at: https://doi.org/10.1016/j.compchemeng.2020.106991.

[36] Syeda, J. (2013). Trend and Variability Analysis, and Forecasting of Maximum Temperature in Bangladesh, Journal of Environmental Science and Natural Resources, 5(2), pp. 119–128. Available at: https://doi.org/10.3329/jesnr.v5i2.14803.

[37] Uzundumlu, A. S., Oksuz, M. E. and Kurtoglu, S. (2018). Future of fig production in Turkey, Journal of Tekirdag Agricultural Faculty, 15(2), pp. 138–146.

[38] Vapnik, V., Golowich, S. E. and Smola, A. (1997). Support vector method for function approximation, regression estimation, and signal processing, Advances in Neural Information Processing Systems, pp. 281–287.

[39] Ye, L. et al. (2021). Bayesian regularized neural network model development for predicting daily rainfall from sea level pressure data: Investigation on solving complex hydrology problem, Complexity, 2021. Available at: https://doi.org/10.1155/2021/6631564.

[40] Yonekura, K., Hattori, H. and Suzuki, T. (2019). Short-term local weather forecast using dense weather station by deep neural network, Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018, pp. 1683–1690. Available at: https://doi.org/10.1109/BIGDATA.2018.8622195.