Original Research Article

Comparative analysis of Machine Learning Models in Prediction of Maximum Temperature in the Raichur City

.

ABSTRACT

|  |
| --- |
| Nowadays, many Machine Learning (ML) models gaining prominence due to their accuracy in predicting the time series data. Because of, availability of, different machine learning models, it is important to know the difference between them. A python-based and open-source application, known as Orange software (Version 3.38.1), used to train the ML models. The study was conducted in the Raichur city, Karnataka, India. The data of 45 years were collected which includes a dependent variable i.e., maximum temperature and independent variables i.e., relative humidity, wind speed, surface pressure and precipitation. Six machine learning models namely, Linear Regression (LR), k-Nearest Neighbor (kNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), along with two ensemble models such as Random Forest (RF) and Adaptive Boosting (AB) were selected to analyze the data. Each model was trained by trial and error then the trained model was validated/predicted with the observed values. The seven performance indicators were recorded for each model and compared. Out of six ML models, the ANN model perform better in all seven performance indicators. Specifically, it was found that the R2 value of the ANN model in both training and testing stages was and respectively. The ANN model occupies rank , after comparing six ML models, with accuracy in validating/predicting the maximum temperature. Hence, the ANN model can be suggested for validation/prediction of maximum temperature around the regions of the Raichur city. |

*Keywords: Machin Learning Models, Ensemble Models, Temperature Prediction,   
 Performance Indicators.*

1. INTRODUCTION

Temperature is a key parameter in the global climate change scenario. A small change temperature results in large impact on the anthropogenic activities. The presence of nonlinearities in temperature time series, furthermore the geographic distribution of correlation exponents exhibits well defined clustering [1]. The largest individual daily temperature trends occur during winter and early spring, when substantial warming is observed [2]. Temperatures are likely to increase in entire calendar year, but the changes in winter season are expected to be prominent [3].

Nowadays, machine learning techniques gathering more importance because of their level of accuracy in the prediction. The goal of machine learning is to program computers to use example data or past experience to solve a problem [4-5]. The most comprehensive treatment available of neural networks from an engineering perspective [6]. One day ahead temperature forecast shows that MLP network has a good performance and reasonable prediction accuracy was achieved [7]. ANN models developed with reduced average prediction error by increasing the number of distinct observations used in training, adding additional input terms that describe the date of an observation, increasing the duration of prior weather data included in each observation, and reexamining the number of hidden nodes used in the network [8]. The applicability of ANN approach by developing effective and reliable nonlinear predictive models for weather analysis also compare and evaluate the performance of the developed models using different transfer functions, hidden layers and neurons to forecast maximum temperature [9].

SVMs, a generation learning system based on recent advances in statistical learning theory and are now established as one of the standard tools for machine learning and data mining [10]. Time series data of daily maximum temperature was analyzed to predict the maximum temperature of the next day based on the daily maximum temperatures for a span of previous n days referred to as order of the input [11]. SVMr technique applied was fully described, including some bounds on the machine hyper-parameters in order to speed up the SVMr training process [12].

The effect of urban heat island using temperature as the independent variable was studied using multiple linear regression model and the accuracy of the predicted values was depicted by comparing the predicted and measured values [13]. Multiple linear regression model was used to predict the average temperature of a day based on past meteorological data and features [14].

Prediction of temperature was done for number of days and it was found that K-Nearest Neighbor produced satisfying results with more accuracy [15]. A hybrid approach, K-Nearest Temperature Trends (KNTT) which identifies a cluster of K years, showing nearest temperature trends to that of the year of the missing value date [16]. A k-nearest neighbor (KNN) model validated to simulate and eventually produce a one-day forecast of mean water temperature [17].

Recently, there has been a notable tendency towards employing ensemble learning methodologies for simulation and prediction purposes. The advancement of ensemble methods, including the resampling ensemble methods (e.g., bagging, boosting, and dagging) in different application fields [18]. A bound for the mean squared generalization error is derived that shows that the decrease in error from the individual trees in the forest depends on the correlation between residuals and the mean squared error of the individual trees [19]. A weighted feature selection to select the most useful features, which can both distinguish the majority of all samples and the previous misclassified samples [20]. An efficient temperature forecasting model based on hybrid principal component analysis empowered machine learning techniques used to predict the test datasets [21].

Literature revealed that several machine learning models are available to predict and validate. A large scale comparison study for the major machine learning models for time series forecasting. The models considered are multilayer perceptron, Bayesian neural networks, radial basis functions, generalized regression neural networks, K-nearest neighbor regression, CART regression trees, support vector regression and Gaussian processes [22]. The SVM was found to perform significantly better than the MLPNN (multilayer perceptron neural network) in terms of mean square error and root mean square error, although computational times for the two models are statistically equal [23]. Extreme learning machine model was found to perform better than random forest, backpropagation neural networks and generalized regression neural networks [24]. The use of a decision tree, an artificial neural network, bagging and gradient boosting to forecast the compressive strength of concrete at high temperatures [25]. The performances of various machine learning methods such as support vector machine, adaptive neuro-fuzzy inference system and decision tree are compared to infill missing air temperature [26]. The ANN machine learning method was the best method for predicting temperature values when using dewpoint, relative humidity, wind speed, sea pressure and 1-hour precipitation values as dependent variables [27].

Therefore, in this paper six different ML methods were trained using the daily temperature data of the Raichur city. Various performance indicators for the six ML predictions were used to identify the best ML method.

**2. METHODOLOGY**

**2.1 Study Area**

The research region is in the northern section of India's Karnataka state, as seen in Figure 1. It is positioned between the Krishna and Tungabhadra rivers and is located at 16°20' N latitude and 77°37' E longitude (Deccan Plateau) (see [Table 1](#_bookmark0)). The location is in the monsoon climatic region, with yearly temperatures ranging from 18 to 45 degrees Celsius. It is rich in deposits from the Krishna and Tungabhadra rivers. Droughts have a significant impact on agricultural activities.



**Fig. 1. Inset map outlining the location of the Raichur city in Karnataka, India.**

From the NASA’s POWER CERES MERRA-2 climate data repository, discrete data for the Raichur city was mined for climate data from April 1st, 1981 to March 31st, 2025. The climate parameters such as maximum temperature, relative humidity, wind speed, surface pressure and precipitation are taken into consideration (see Table 2). The daily temperature defined as the dependent variable, and the other parameters were defined as independent variables.

**Table 1. Raichur city weather Metadata**

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Latitude (Degree)** | **Longitude**  **(Degree)** | **Elevation from sea level**  **(meter)** |
| Raichur | 16.2122 | 77.345 | 412.37 |

**Table 2. Features and parameters used in this study**

|  |  |  |
| --- | --- | --- |
| **Data** | **Role** | **Temporal scale and Unit** |
| Maximum Temperature | Target | Daily, °C |
| Relative humidity | Feature | Daily, % |
| Wind Speed | Feature | Daily, MPH |
| Surface pressure | Feature | Daily, milibar |
| Precipitation | Feature | Daily, mm |

**2.2 Machine Learning Methods**

**2.2.1 Single Model ML Methods**

Single model ML methods follow the procedure of only one ML method, whereas ensemble model ML methods follow multiple ML methods. Four different single model ML methods, namely Linear Regression (LR), k-Nearest Neighbor (kNN), Support Vector Machine (SVM) and Artificial Neural Network (ANN), were used and compared in this paper.

**Linear Regression (LR)**

The Linear Regression (LR) is one of the simplest and most common machine learning algorithms. It works based on the best fit criterion. It provides a linear relationship between a predictor and one or more response variables. Multiple linear regression algorithm can be used to train the model by applying the feature selection method to select the most decisive features out of all the features available [14]. LR uses different types of regularization such as no regularization, ridge regression regularization, lasso regression regularization and elastic net regression regularization. Lasso regression minimizes a penalized version of the least squares loss function with L1-norm penalty and ridge regularization with L2-norm penalty. Whereas, the elastic net is a regularized regression method that linearly combines the L1 and L2 in the ratio which gives good results.

**k-Nearest Neighbor (kNN)**

The k-Nearest Neighbor (kNN) algorithm searches for k closest training value in feature space and uses their average as prediction. kNN is used to develop a system that uses numeric historical data to forecast the climate of a specific region [15]. The imputation method relies solely on the available daily maximum temperature data set [16]. The criteria for this model are set the number of nearest neighbors, the various distance parameter and different weights. Most common types of metrics can be used with kNN are Euclidean, Manhattan, Maximal and Mahalanobis whereas, two different weights can be used in this model are uniform and distance.

**Support Vector Machine (SVM)**

The Support Vector Machine (SVM) algorithm separates the attribute space with a hyperplane by maximizing the margin between different classes. The general framework of support vector learning includes, regularized linear learning models, theoretical bounds, convex duality and sparseness of the dual-kernel representation [10]. It is a kernel based technique used for forecasting, classification and regression applications [23]. The optimum value can be obtained by choosing different types of kernels such as Linear, Polynomial, Radial bass function and Sigmoid. Meanwhile, the values of the gamma constant g, degree of kernel d and a smoothness constant c are crucial. This model also uses optimization parameters such numerical tolerance and iteration limit.

**Artificial Neural Network (ANN)**

The Artificial Neural Network (ANN) algorithm uses a multi-layer perceptron (MLP) with backpropagation. A multilayer neural network can approximate any smooth, measurable function between input and output vectors by selecting a suitable set of connecting weight and transfer functions [7]. The ANN model with reduced average prediction error was developed by increasing the number of distinct observations used in training, adding additional input terms and reexamining the number of hidden nodes used in the network [8]. The logistic sigmoid function, the hyperbolic tan function, the rectified linear unit function and identity can be used as activation function for the hidden layer. This model includes solver for weight optimization such as quasi-Newton method, stochastic gradient descent, stochastic gradient-based optimizer and also regularization term parameter.

**2.2.2 Ensemble-Model ML methods**

Ensemble machine learning method is a compound model that combines several base ML models to make better prediction. Ensemble models can achieve better predictive performance by reducing the noise or error between observed and predicted data. Ensemble models consists of various approaches based on different methodologies including, stacking methods, averaging methods, stacking, bagging, boosting and dagging approaches [18]. In this paper, two ensemble models namely, Random Forest (bagging) and Adaptive Boosting (boosting) were taken into consideration for temperature time series analysis.

**Random Forest**

The Random Forest (RF) algorithm is a bagging ensemble model that uses an ensemble of decision trees. Each tree is developed from a bootstrap sample from the training data. A subset of attributes is drawn arbitrarily during the development of individual tree. Out of these attributes the best attribute is selected for the split. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them [19]. The number of trees at each node, number of attributes considered at each split, replicable training, balancing class distribution are the basic properties of this model. This model also uses growth control parameters such as limit depth of individual trees and smallest subset that can be split.

**Adaptive Boosting (AB)**

The adaptive Boosting (AB) algorithm is a boosting ensemble model that combines weak learners to boost their performance. It has no random elements and grows an ensemble of trees by successive reweighting of the training set [20]. The base estimator of this model is a tree and one can set the parameters such as number of estimators, learning rate and fixed seed for random generator. This model uses mainly three different types of regression loss function namely linear, square and exponential.

**2.3 Performance Measure**

The evaluation metrics such as (1) coefficient of determination (R2), (2) The index of agreement (), (3) Nash-Sutcliffe efficiency (NSE), (4) Mean absolute error (MAE), (5) Root mean square error (RMSE), (6) Percent bias (PBIAS), and (7) the RMSE- observations standard deviation ratio (RSR) were used to measure the ability of prediction of different ML algorithms [28-29].

The degree of collinearity between observed and simulated values can be measured from both Pearson's correlation coefficient () and coefficient of determination (R2). The correlation coefficient (), which ranges from to , measures the linear relationship between the parameters. Here, means no linear relationship exists, whereas or mean positive and negative correlation respectively. R2 interprets the proportion of the variance in the simulated values that is predictable from the observed values. Typically, R2 varies from to where the higher values suggest less variance in the observed dataset [29].

The agreement index () is the standardized agreement measurement between the predictions and observations given by equation (1) and varies between to . Here, means a perfect agreement between the observed and predicted values, and means no agreement [28-29].

(1)

Nash-Sutcliffe efficiency (NSE) is a normalized statistic determining the relative magnitude between the noise and information. NSE variers between to for the acceptable performance and can be described by equation (2) [28-29].

(2)

Mean absolute error (MAE), and root mean square error (RMSE) can be described using the equations (3-4) below. Both MAE and RMSE indicate perfect fit when close to . However, RMSE is a better tool than MAE since it increases significantly when there are significant differences between the observed and simulated values [28-29].

(3)

(4)

Percent bias (PBIAS) calculates the average tendency of the simulated data to be larger or smaller than their observed counterparts as in equation (5). Positive and negative values indicate underestimation and overestimation bias, respectively [28-29].

(5)

RMSE-observations standard deviation ratio (RSR) is the ratio of the RMSE and the standard deviation of the measured data as shown in equation (6). Better model performance means lower RSR or, in other words, lower RMSE [28-29].

(6)

For equations (1 through 6), is the total number of observations, is the th observation of the temperature time series data, is the th simulated value of the temperature time series, is the mean of the temperature time series. [Table 3](#_bookmark2)summarizes the criteria for the before mentioned statistical methods used in finding a better predictive model.

**2.4 Machine Learning Tool for Implementation**

Data sorting and filtering using Python’s Jupyter Notebook 7.0.8 were done in this study. First, statistical trends were conducted using Python’s Jupyter Notebook 7.0.8. Then ML algorithms were carried out using an open-source data mining software called Orange 3.38.1, a python-based programming language, developed by the University of Ljubljana [30,31].

**3. RESULTS AND DISCUSSION**

**3.1 Data Correlation**

The general performance ratings of the statistical indices must be checked in the data before applying machine learning algorithms. The collinearity among the features and between the features and the target must also be checked.

**Table 3. General performance ratings of the statistical indices**

**(Adopted from [28,29])**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Performance Rating** | **RANGE** | | | | | |
| **RSR** | **NSE** | **PBIAS (%)** | **RMSE** | **MAE** | **INDEX OF AGREEMENT** |
| Very Good | ~ | ~ | ≤ | ~ |  |  |
| Good | ~ | ~ | ~ | ~ | Not Defined | Not Defined |
| Satisfactory | ~ | ~ | ~ | ~ | Not Defined | Not Defined |
| Unsatisfactory | > | < | ≥ | ~ |  |  |

Collinearity among the features and target data were evaluated. Table 4 shows that the collinearity among the features was relatively low as required for learning methods. The surface pressure and wind speed were higher than the rest but remained within an acceptable range.

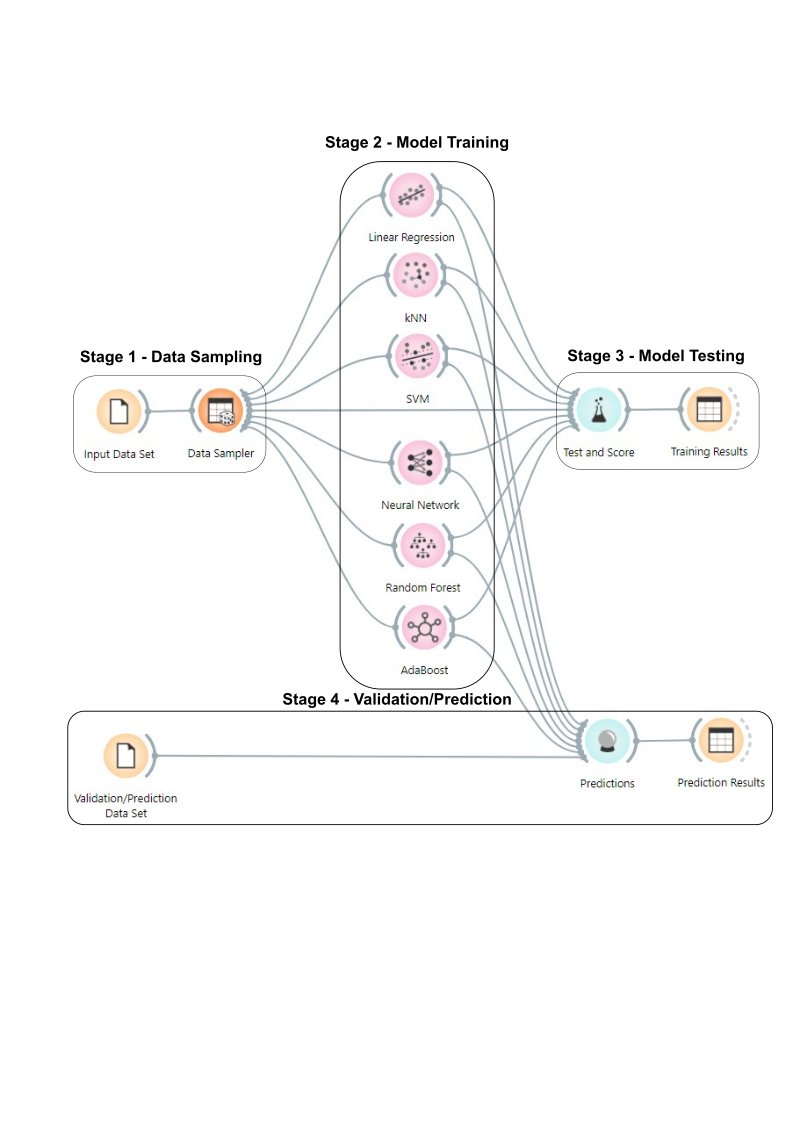
**Table 4. Features collinearity and correlation matrix**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Features** | **Relative Humidity** | **Wind**  **Speed** | **Surface Pressure** | **Precipitation** |
| Relative Humidity |  |  |  |  |
| Wind Speed |  |  |  |  |
| Surface Pressure |  |  |  |  |
| Precipitation |  |  |  |  |

**3.2 Training and Prediction Results**

The prediction of maximum temperature in the study area can be done by using different ML methods for the provided feature data sets. The combined data set contained records of daily maximum temperature, relative humidity, wind speed, surface pressure, and depth of precipitation from to . The training set included records for the date range from to . Training data were then randomly split into two parts, for training, for testing. Data from to were preserved for the validation/prediction step used in deciding on the best method. In this study, we used the daily climate data, including relative humidity, wind speed, surface pressure, and precipitation, to predict the maximum temperature for the day of interest.

A python-based and open-source application, known as Orange software (Version 3.38.1), used to train the ML models. The schematic of the ML model training and prediction steps showed in Figure 2. There are three stages involved in training and test data sets. The data sampler split the input data into training and test data sets in the first stage. In the second stage, the training data set was introduced to the ML models and the training process started. Each of the models was tested using the test data set in the third stage. The training stage was done many times using trial and error to obtain the best results. When the best results for the training stage were achieved, the validation/prediction data set was introduced to the trained models, and the results were compared to the observed data set.



**Fig. 2. Schematic of the workflow and the Orange Software workspace**

**Fig. 3. Temperature time series including the Training and Validation data sets**

The accuracy of the training step was measured by seven unique indices, namely the coefficient of determination (R2), the index of agreement (d), the Nash-Sutcliffe efficiency (NSE), the mean absolute error (MAE), the root mean square error (RMSE), the percent bias (PBIAS), and the RMSE-observations standard deviation ratio (RSR). The optimized results for every ML method were explained more in detail in the following section.

The Linear Regression (LR) algorithm was optimized by minimizing the mean squared error (MSE) with regularization parameter (alpha) value of . Elastic net regression was chosen as regularization type with mixing of L1 (Ridge regression) and L2 (Lasso regression) at and respectively. The correlation between the observed values and the tested values was depicted in Figure 4 (a) with the value of R2 as .

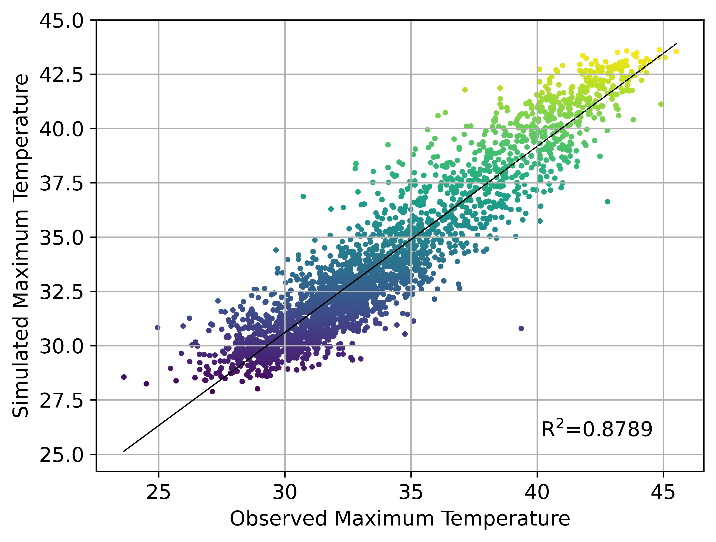
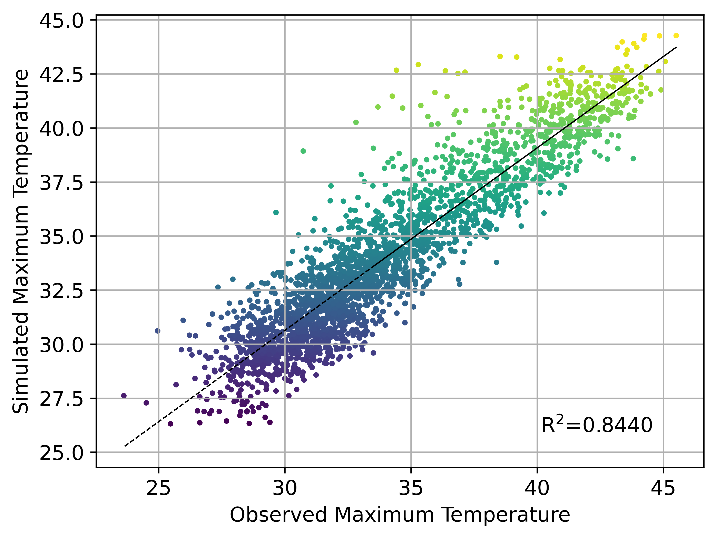
The k-Nearest Neighbor (kNN) algorithm optimized through repeated attempts. The best results found by setting number of neighbors (k) to seventeen. Mahalanobis was selected as distance metric and distances as the weighting parameter. Figure 4 (b) depicts the correlation between the observed values and test results with R2 value of .

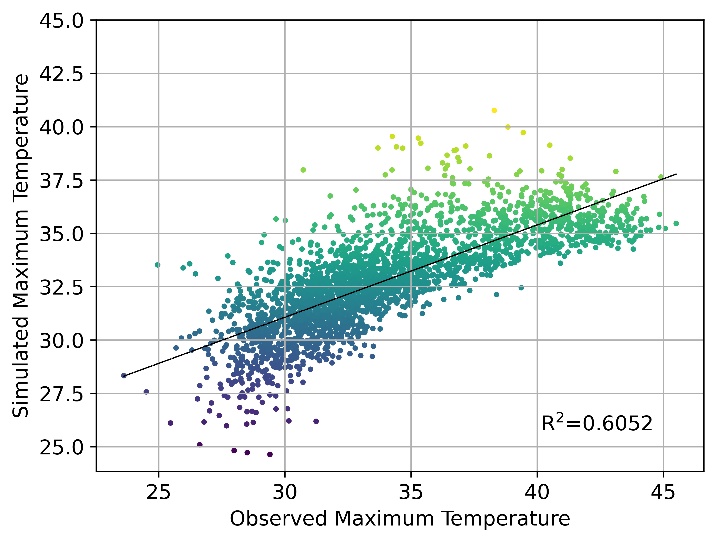
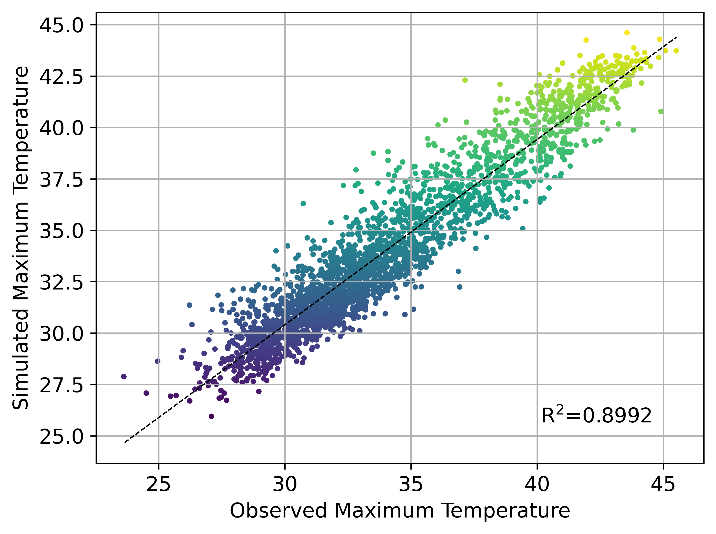
The Support Vector Machine (SVM) algorithm was set at cost and ε values of and respectively. Sigmoid kernel function is selected with gamma (g) value of and coefficient (c) value of . Optimization parameters such as numerical tolerance and iteration limit were set at and respectively. Figure 4 (c) shows the correlation between the observed values and tested values with R2 value of .

Figure 4 (d) represents the correlation between the observed values and the tested values with R2 value of . The Artificial Neural Network (ANN) algorithm trained with neurons in the hidden layers of the neural network. Logistic function Limited-memory-Broyden-Fletcher-Goldfarb-Shanno (L-BFGS-B) were chosen as activation function and solver of the algorithm respectively. Setting the Regularization parameter (alpha) at and the maximum number of iterations at .

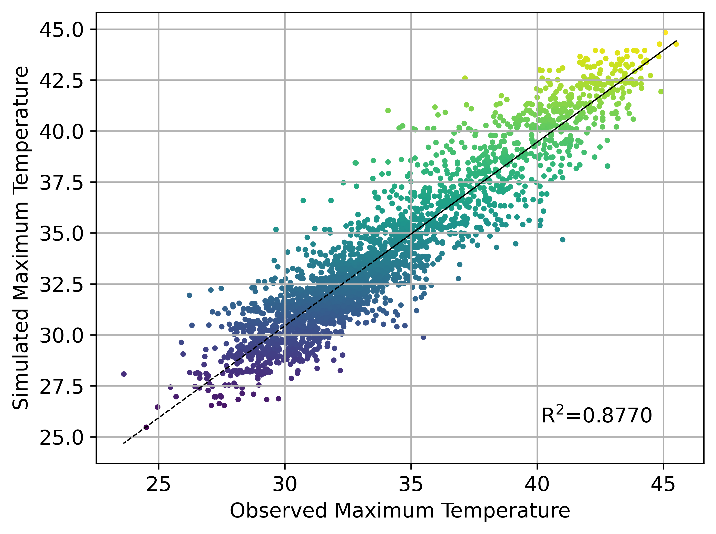
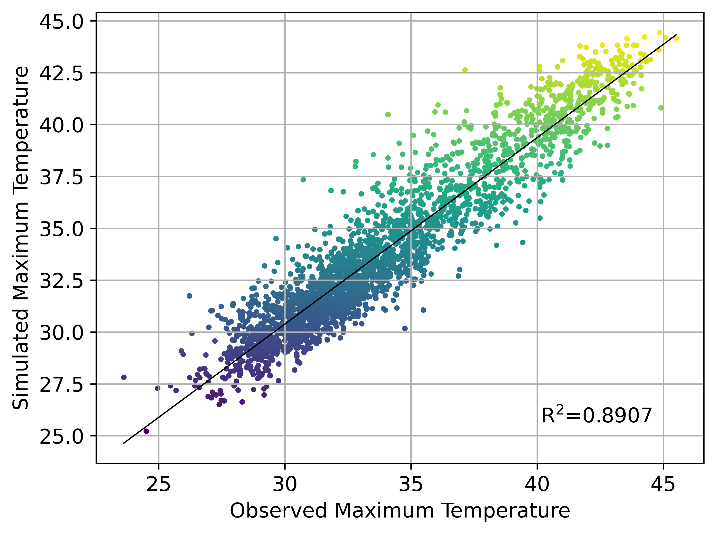
The Random Forest (RF) algorithm was the first ensemble model used in this paper. Adjust the basic properties of the model such as number of trees and number of attributes considered at each split to and respectively. The subsets containing smaller than trees should not split and the depth of individual trees must be set to . Figure 4 (e) shows the correlation between the observed values and the tested values with R2 value of .

The second ensemble model used in this paper was the Adaptive Boosting (AB) algorithm. Base estimator used in this model was decision tree and the number of estimators are chosen as . The learning rate of each weak learner is set at . The linear loss function was selected while fixing the reproducibility parameter at . Figure 4 (f) depicts the correlation between the observed values and the tested values with R2 value of .



**(a) (b)** 

**(c) (d)**



**(e) (f)**

**Fig. 4. Correlation between observed and test outputs for (a) Linear regression   
 (b) kNN, (c) SVM, (d) ANN, (e) Random Forest, and (f) AdaBoost**

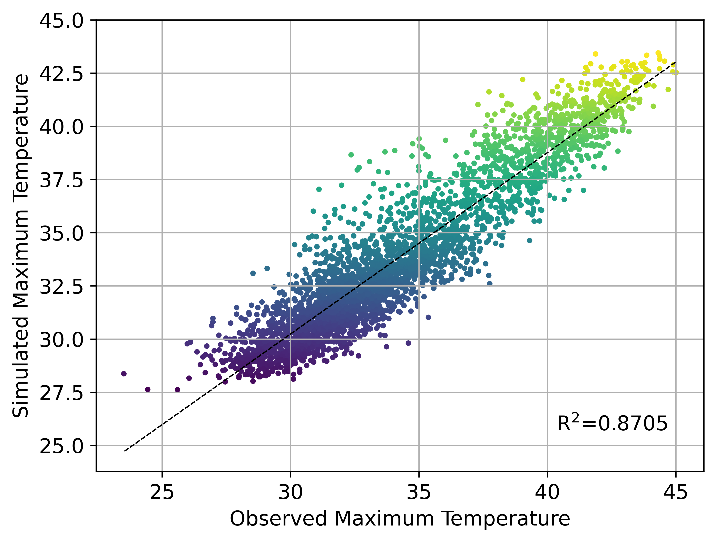
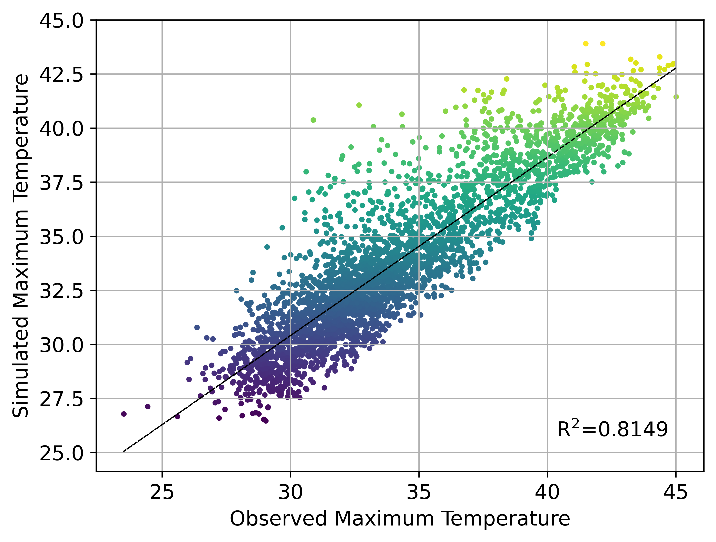
The value of R2 is not sufficient to judge the accuracy of the model and hence the values of the remaining six performance indicators also important in identifying the best model. Table 5 shows the values of seven performance indicators of each model along with R2. All six machine learning models were well trained with good results. But, the performance of the SVM model was not up to the satisfactory mark with less R2 value of . It is observed from the table that the best performance showed by ANN model with highest value of NSE = , index of agreement = and lowest value of RSR = , MAE = , PBIAS = , and RMSE = . Out of six trained models, the ANN model emerged as the best trained model followed by RF, kNN, AB, and LR whereas the worst trained model was SVM with lowest value of NSE = , index of agreement = and highest value of RSR = .

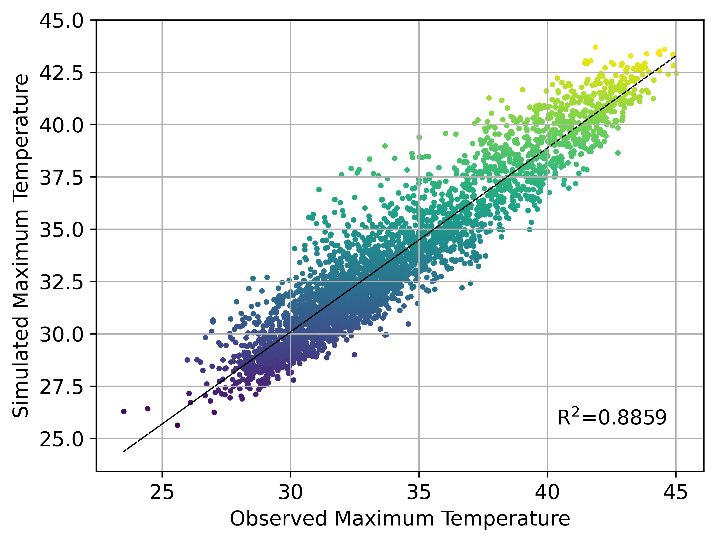
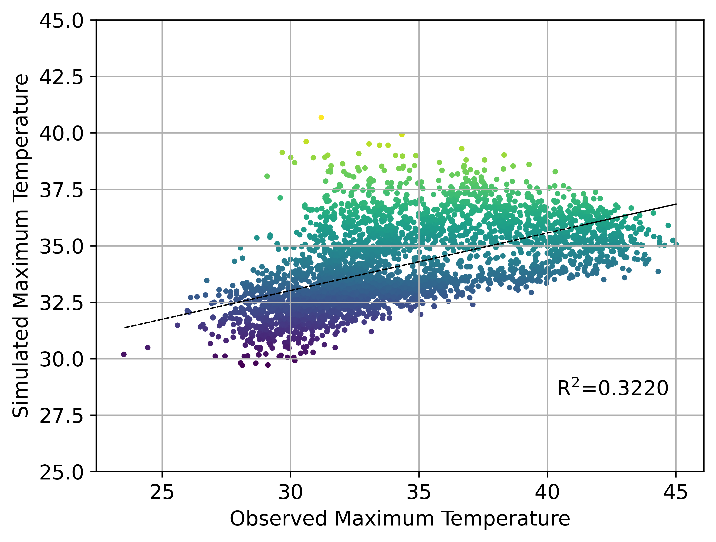
**Table 5. Performance ratings for the ML models for the test dataset.**

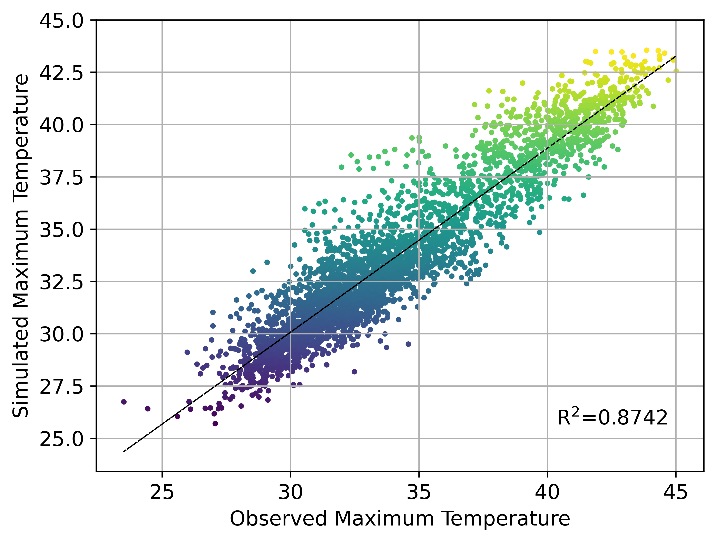
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **METHOD** | **RSR** | **NSE** | **MAE** | **INDEX OF AGREEMENT** | **PBIAS (%)** | **RMSE** | **R2** |
| Linear Regression |  |  |  |  |  |  |  |
| kNN |  |  |  |  |  |  |  |
| SVM |  |  |  |  |  |  |  |
| ANN |  |  |  |  |  |  |  |
| Random Forest |  |  |  |  |  |  |  |
| Adaptive Boostiing |  |  |  |  |  |  |  |

After training the six machine learning models, the validation/prediction of these models will be carried out at stage 4 as shown in Figure 2. For this, the daily data from to were used to validate including maximum daily temperature as dependent variable, whereas relative humidity, wind speed, surface pressure, and precipitation as independent variables. These data were introduced to the trained models and the resulting maximum temperature predictions were compared to the observed values.

Figure 5 shows the correlation between the observed values and predicted values for all six models. The best model fit was observed for ANN with R2 = , followed by RF, kNN, AB, and LR. Whereas, the least accurate prediction was seen by SVM with R2 = 0.3220.



**(a) (b)**

**(c) (d)**

**(e) (f)**

**Fig. 5. Correlation between observed and validation outputs for (a) Linear regression   
 (b) kNN, (c) SVM, (d) ANN, (e) Random Forest, and (f) AdaBoost**

The values of performance indicators of each ML model for validation/prediction data set tabulated in Table 6. Along with R2, the ANN model also predicts well with respect to other six performance indicators such as highest NSE = 0.8859, index of agreement = 0.9690 and lowest RSR = 0.3378, MAE = 0.9858, PBIAS = 0.0291 and RMSE = 1.2752. The worst predictions shown in the case of the SVM model with lowest NSE = 0.3220, index of agreement = 0.6835 and highest RSR = 0.8234.

**Table 6. Performance ratings for the ML models for validation/prediction data set.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **METHOD** | **RSR** | **NSE** | **MAE** | **INDEX OF AGREEMENT** | **PBIAS (%)** | **RMSE** | **R2** |
| Linear Regression | 0.4302 | 0.8149 | 1.1982 | 0.9470 | 0.0353 | 1.5877 | 0.8149 |
| kNN | 0.3599 | 0.8705 | 1.0318 | 0.9644 | 0.0305 | 1.3284 | 0.8705 |
| SVM | 0.8234 | 0.3220 | 1.1963 | 0.6835 | 0.0351 | 1.4945 | 0.3220 |
| ANN | 0.3378 | 0.8859 | 0.9858 | 0.9690 | 0.0291 | 1.2752 | 0.8859 |
| Random Forest | 0.3547 | 0.8742 | 1.0350 | 0.9655 | 0.0306 | 1.3484 | 0.8742 |
| Adaptive Boosting | 0.3734 | 0.8606 | 1.1011 | 0.9614 | 0.0325 | 1.4433 | 0.8606 |

The comparison of six machine learning models based on the seven performance indicators was shown in Table 7. Ranking was allotted depending on the overall accuracy of the model. It was observed from the table that the ANN model outperformed the other models including the ensemble models. The ANN model exhibits the very good agreement in all seven performance indicators in comparison with other models. Hence, the ANN model achieves rank 1. Out of seven indicators, the RF model performed well in four indicators namely, RSR, NSE, index of agreement and R2. Therefore, rank 2 was given to the RF model. Similarly, the kNN model gets the rank 3, the AB model gets the rank 4. In deciding the rank 5 and rank 6, it very close between the LR model and the SVM model. Even though the SVM model not shown good performance four indicators, but manage to perform better than the LR model in three indicators namely, MAE, PBIAS and RMSE. Hence, the rank was given to the LR model and the SVM model gets the rank 6.

**Table 7. ML methods ranking based on their performance**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Rank** | **RSR** | **NSE** | **MAE** | **INDEX OF AGREEMENT** | **PBIAS (%)** | **RMSE** | **R2** | **Overall** |
| 1 | ANN | ANN | ANN | ANN | ANN | ANN | ANN | ANN |
| 2 | RF | RF | kNN | RF | kNN | kNN | RF | RF |
| 3 | kNN | kNN | RF | kNN | RF | RF | kNN | kNN |
| 4 | AB | AB | AB | AB | AB | AB | AB | AB |
| 5 | LR | LR | SVM | LR | SVM | SVM | LR | LR |
| 6 | SVM | SVM | LR | SVM | LR | LR | SVM | SVM |

Figure 6 depicts the comparison of the observed values and ANN predicted values. It was noticed that the accuracy of the ANN model was seen in good agreement with the observed values. The average squared difference between the observed maximum temperature and ANN predicted maximum temperature was found to be reduced.

**Fig. 6. Comparing ANN results with the observed value**

4. Conclusion

In the era of machine learning models, the main goal of this study was to compare the efficiency of the six machine learning models in predicting maximum temperature of the Raichur city. For this purpose, 45 years of data i.e., from 4/1/1981 to 3/31/2025, of Maximum temperature used as dependent variable whereas Relative humidity, Wind speed, Surface pressure and Precipitation as independent variables was taken into consideration. The data was analyzed using the Orange software, a python-based open-source application, by splitting into training data set and validation/prediction data set. The results observed from the study are as follows :

1. Among the six machine learning models, the ANN model trained well than the remaining models with best results in all seven performance indicators followed by the RF model, the kNN model, the AB model and the LR model at the training and testing stage. Whereas the SVM model remains as the worst performer.
2. At the validation/prediction stage, the accuracy of the ANN model was high by obtaining the good output in RSR, NSE, MAE, index of agreement, PBIAS, RMSE and R2. But, the SVM model unable to produce the better output leading to low accuracy in prediction.
3. The ANN model outperformed not only the ensemble ML models i.e., the RF model and the AB model, but also other ML models. Hence, out of six ML models, the ANN model secured the first rank with 89% accuracy.

Overall, the ANN model was the best model for validation/prediction of maximum temperature with independent variables such as relative humidity, wind speed, surface temperature and precipitation. It was suggested from the study that the ANN model can be considered as a best model for validation/prediction of maximum temperature around the regions of Raichur city.

**COMPETING INTERESTS DISCLAIMER:**

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

**Disclaimer (Artificial intelligence)**

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

References

1. Bartos I, Janosi IM. Nonlinear correlations of daily temperature records over land. Nonlin. Processes Geophys. 2006;13:571-576.
2. Bonsal B, Zhang X, Vincent L, Hogg W. Characteristics of Daily and Extreme Temperatures over Canada. J. Clim. 2001;14:1959–1976.
3. Revadekar JV, Kothawale DR, Patwardhan SK, Pant GB, Rupa Kumar K. About the observed and future changes in temperature extremes over India. Natural Hazards. 2012; 60:1133-1155.
4. Alpaydin E. Introduction to Machine Learning. MIT Press: Cambridge MA USA; 2009.
5. Topp SN, Barclay J, Diaz J, Sun AY, Jia X, Lu D et al. Stream temperature prediction in a shifting environment: Explaining the influence of deep learning architecture. Water Resources Research. 2023;59(4):e2022WR033880.
6. Haykin S. Neural Networks. Prentice Hall: New York NY USA; 1994.
7. Hayati M, Mohebi Z. Application of artificial neural networks for temperature forecasting. World Acad. Sci. Eng. Technol. 2007;28:275–279.
8. Smith BA, McClendon RW, Hoogenboom G. Improving air temperature prediction with artificial neural networks. Int. J. Comput. Intell. 2006;3:179–186.
9. Abhishek K, Singh M, Ghosh S, Anand A. Weather Forecasting Model using Artificial Neural Network. Procedia Technol. 2012;4:311–318.
10. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and other Kernel-Based Learning Methods. Cambridge University Press: Cambridge MA USA; 2000.
11. Radhika Y, Shashi M. Atmospheric Temperature Prediction using Support Vector Machines. Int. J. Comput. Theory Eng. 2009;1(1):55–58.
12. Paniagua-Tineo A, Salcedo-Sanz S, Casanova-Mateo C, Ortiz-García E, Cony M, Hernández-Martín E. Prediction of daily maximum temperature using a support vector regression algorithm. Renew. Energy. 2011;36:3054–3060.
13. Sindhu PM, Ramith B, Pooja S, Prajwal S, Sai N. Prediction of temperature using linear regression. ICEECCOT. 2017:1-6.
14. Ishu G, Harsh M, Deepak R, Ashutosh KS. MLRM : A Multiple linear regression based model for average temperature prediction of a day. arXiv preprint arXiv:2203.05835. 2022.
15. Badhiye SS, Nilesh US, Chatur PN. KNN technique for analysis and prediction of temperature and humidity data. Int. Journal of Comp. Appl. 2013;61(14):7-13.
16. Kiani K, Saleem K. K-nearest temperature trends : A method for weather temperature data imputation. Proc. 2017 Int. Conf. on Information System and Data Mining. 2017:23-27.
17. St-Hilaire A, Ouarda TB, Bargaoui Z, Daigle A, Bilodeau L. Daily river water temperature forecast model with a k-nearest neighbor approach. Hydrological Processes. 2012;26(9):1302-1310.
18. Mohammad ZK, Okke B, Marzieh F, Reinhard H. Ensemble machine learning paradigms in hydrology : A review. Journal of Hydrology. 2021;598:126266.
19. Breiman L. Random forests. Machine Learning. 2001;45:5–32.
20. Wang Y, Feng L. An adaptive boosting algorithm based on weighted feature selection and category classification confidence. Applied Intelligence. 2021:1-22.
21. Sen S, Saha S, Chaki S, Saha P, Dutta P. Analysis of PCA based adaboost machine learning model for predict mid-term weather forecasting. Computational Intelligence and Machine Learning. 2021;2(2):41-52.
22. Nesreen KA, Amir FA, Neamat E, Hisham E. An empirical comparison of machine learning models for time series forecasting. Econometric reviews. 2010;29(5-6):594-621.
23. Abubakar A, Chiroma H, Zeki A, Uddin M. Utilising key climate element variability for the prediction of future climate change using a support vector machine model. Int. J. Glob. Warm. 2016;9:129–151.
24. Yu F, Ningbo C, Weiping H, Lili G, Daozhi G. Estimation of soil temperature from meteorological data using different machine learning models. Geoderma. 2019;338:67-77.
25. Ayaz A, Krzysztof AO, Mariusz M, Furqan F, Imran M, Afnan N. Comparative study of supervised machine learning algorithms for predicting the compressive strength of concrete at high temperature. Materials. 2021;14(15):4222.
26. Katipoglu OM. Prediction of missing temperature data using different machine learning methods. Arabian Journal of Geosciences. 2022;15(1):21.
27. Babak A, Khairul H, Joel P, Saman E. Evaluation of machine learning methods application in temperature prediction. CRPASE : Transc. Civil and Envir. Engg. 2022;8(1):1-12.
28. Shcherbakov MV, Brebels A, Shcherbakova NL, Tyukov AP, Janovsky TA, Kamaev VA. A survey of forecast error measures. World Appl. Sci. J. 2013;24:171–176.
29. Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science. 2021;7:623.
30. Demšar J, Curk T, Erjavec A, Gorup Č, Hočevar T, Milutinovič M et al. Orange: data mining toolbox in Python. J. Mach. Learn. Res. 2013;14:2349–2353.
31. Naik A, Samant L. Correlation review of classification algorithm using data mining tool: WEKA, Rapidminer, Tanagra, Orange and Knime. Procedia Comput. Sci. 2016;85:662–668.