IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12631

Analysis of Machine Learning Techniques for Weather Forecasting

D. Sandhya Rani[1] and Swetha P[2]

Assistant Professor, Department of Electronics and Communication Engineering, Sri Indu College of Engineering and Technology, Ibrahimpatnam, Ranga Reddy District, Telangna -501510[1]

Assistant Professor, Department of Electronics and Communication Engineering, Sri Indu College of Engineering and Technology, Ibrahimpatnam, Ranga Reddy District, Telangana -501510[2]

Abstract: In many applications, weather forecasting (WF) research is an essential endeavor. These applications always require accurate WF. The purpose of this study is to validate different machine learning (ML) classifiers for wet weather prediction. Several machine learning techniques, including support vector machines (SVM), decision trees (DT), and artificial neural networks (ANN), are validated and tested using the Koggle weather dataset. Extended surveys of ML and NN-based WF techniques are also included in this study. The different ML-based classifiers are validated by comparing their prediction accuracy. The five data classes are classified using four features. Overall, it is discovered that the verified accuracy of ANN and Random Forest is 84.35% each.

Keyword: Weather Forecasting, Rain, Classifier, Machine Learning. SVM, ANN, DT and Random Forest etc.

I. INTRODUCTION

Weather forecasting (WF) is being revolutionized by machine learning (ML), which is providing exciting new developments in addition to the conventional physics-based models. This is an explanation of the recent technological developments. There are many advantages of the using ML for WF prediction. These methods using training may improve precision for WF particularly with regard to short- term forecasts. Use of ML allow to combine information from various sources to provide deeper insights. The noisy of partial weather data can be easily managed. There are many features used as the components of the WF systems the most common components are shown in the Figure 1.

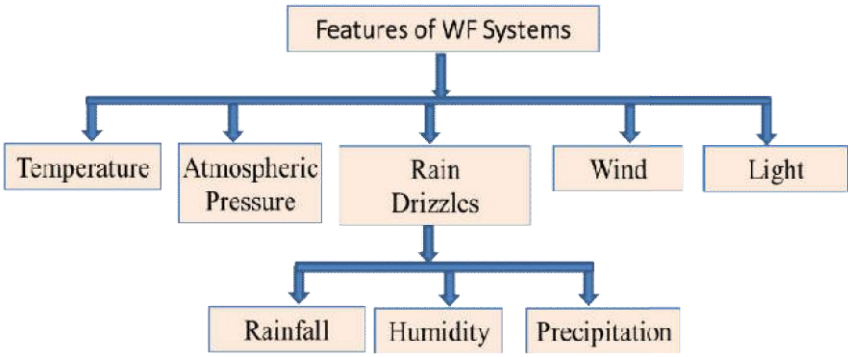

Figure 1 Components of the WF systems to be taken as features

Applications of WF System:
Extensive applications serve as a strong incentive for WF research. Since there is a constant need for accurate weather forecasts in almost every part of life, their significance cannot be overstated.

Disaster weather alerts & Management
Severe weather alerts including advisories, that the national weather services issue when severe or dangerous weather is anticipated, are a significant component of modern weather forecasting.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12631

Air Traffic
Given that a higher percentage of plane crashes worldwide have weather-related causes, reliable weather forecasting is crucial since the aviation sector is particularly weather sensitive.

Navigation
The direction and speed of the wind, the height and frequency of the waves, the tides, and the precipitation can all have a substantial impact on the commercial and recreational use of rivers.

Agriculture
WF are used by farmers to determine what tasks to perform on any given day. For instance, it is only practical to dry day during dry weather. Crops like corn, wheat, and cotton can be destroyed by extended dry spells. Drought can destroy crops, but the dried remnants can be made into silage, which can be used in place of cattle feed. Crops suffer greatly from freezes and frosts in the spring and autumn.

Utility companies
The WF systems are crucial for electricity and gas businesses as they help them predict demand which is highly influenced by the weather. To calculate the degree of heating (day) as well as cooling (cooling day) that will occur, they utilise a quantity known as the degree day. Based on an average day's temperature of 650 F (180C), these amounts have been calculated. One degree Fahrenheit corresponds to one heating degree day, and one degree Fahrenheit corresponds to one cooling degree day.

Private Sector
Private businesses are increasingly paying for customised weather forecasts in order to boost revenue or minimize losses. Supermarket companies, for instance, might adjust the merchandise on their shelves in anticipation of shifting consumer buying patterns under various weather conditions. Investing in commodities, such as futures for oranges, maize, soybeans, and oil, can be done using weather forecasts.

Military applications
Military weather forecasters inform the war combatant community about the weather, just like their commercial sector counterparts do. Military weather forecasters give pilots pre-flight weather briefs and offer military sites real time resource protection services.

II. LITERATURE REVIEW

Wide range of the ML techniques has been used for the accurate production of the Weather in the country and abroad researches. This section of the paper has presented the extensive survey of the relevant WF methodologies.

1. Classifier based ML based WF systems
There are many ML methods adopted for improving the accuracy of WF systems based on applications and feature set requirements. This section has reviewed SVM and time series based prediction systems, summary is given in Table 1.

Table 1 Summary and discussion of the classifier based ML WF methodologies.

Authors	Methodology	Approach
Bogdan Bochenek et al [1]	Deep Learning, Random Forest, (RF) Artificial Neural Networks, Support Vector Machine,	Presented survey of weather forecasting RF is for Beginners and ANN need knowledge.
A H M Jakaria et al.[2]	High Performance Computing (HPC) environment which consumes a large amount of energy.	it is beneficial to leverage the weather station data from multiple city data and observed RMSE for 10 cites.
Bengtsson, L et. al [3]	Forecast of weather overview	Various applications of WF systems have been described
Rina Mahakud et.al. [4]	method has a lower skill than ensemble weather forecast models in predicting forecast uncertainty.	The performance is evaluated using the **Fi score** for forecast model learning algorithm fir prediction error FI=0.62

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12631

Bin Wang et al [5]	Numerical weather prediction NWP, is compared to negative log error (NLE).	approach was evaluated on a public dataset from weather stations in Beijing, China and improves accuracy by 47.76%.
B. Arca and P. Duceet al [6]	Both ARIMA and ANN models were developed using four years of hourly meteorological data from three meteorological stations of SAR.	The use of meteorological variables from numerical weather forecast gave better results than those obtained from ARIMA and ANN (98%) models.
Sorkun, murat Cihan et al/ [7]	variation, namely a long short-term memory (LSTM) unit is used	In further experiments, the effect of combining most effective parameters was investigated and, as a result.
Proposed	Multi class classification and using SVM, RF,DT and ANN	The cross features are considered to apply classifiers Decision tree, RF, ANN xGBoost for prediction.

2. Review of Neural Network Based WF

Himani Tyagi et al The study on predicting the weather utilising previous data sets is presented in this publication. Due to the complicated and complex nature of atmospheric pattern, conventional methods are not as efficient or effective. The use of artificial neural networks is a popular approach to overcoming these issues. In order to predict temperatures for every day of the year, the suggested artificial neural network (ANN) applies various neurons, layers that are concealed, and function transfers to assess the accuracy of the constructed networks. The mean-square error (MSE) is a criterion used for choosing the suitable hypothesis. Mohammed Al-Shawwa et al The use of an ANN model for predicting the environmental temperature was created and evaluated in this study. Numerous variables that could impact moisture or temperature were found. The artificial neural network (ANN) algorithm takes into account a variety of elements as input parameters, including the environment around it, location or closeness of water substrates, the impact of plants, and the degree of elevation above or beneath the level of the ocean. A.E. Ruano et al This paper discusses the building of prediction neural network simulations for the temperature of indoor air that will be utilised to air conditioning systems controlling. A full description of how to create off the grid neural network with radial basis functions models using multi-objective genetic algorithms is provided. Such models that use data function better when compared to a physically based multi-node system. The simulations are produced using weather and ecological information that were gathered by a far-flung data collecting system from an intermediate school in the southern region of Portuguese. Sotiris Pap Antoniou et al This paper's objective is to describe the creation and assessment of classification methods utilising neural networks for forecasting of outside temperatures employing data collected from 4 European cities: Mollet, Spain; Granada, Spain; Chania, Greece; and Ancona in Italy. To determine which neural networks topology—feed ahead, waterfall, and elman—is best for each city, several configurations are being examined. By contrasting the projected and actual ambient temperatures, the correctness of the forecast is confirmed. Ashrafi, Kh et al This structure takes into account both natural and man-made factors that impact the amount of sunlight that reaches the atmosphere in order to forecast the increase in local temperatures caused by global warming. The average temperatures, condensation point humidity percentage, solar radiation, wind speed, cloud cover, precipitation, station-level pressure (QFE), & greenhouse emissions are the inputs of the artificial neural network (ANN) model. One period, six months, twelve months, and twenty-four months prior to the collected information constitute the input information used to predict the month mean temperature.

Brian A. Smith et al Reducing crop loss from destructive freezes necessitates precise weather forecasting techniques. A better model for predicting Georgia's temperatures was created by tweaking the settings of an already-existing synthetic neural network (ANN) model and adding seasonally data. For every model, multiple neural networks were instantiated and trained in order to contrast different models.

Imran Maqsood et al This research offers a comparison of many neural network simulations for Vancouver, British Columbia, Canada for forecasting the weather. We utilised a year's worth of data, including the wind speed and monthly maximum and lowest temperatures, to create the simulations. We educated an Elman has Retractable Artificial Neural Network (ERNN) and a Multiple layers of fabric Perceptron is (MLP) utilising the Levenberg-Mar and one-step-secant methods.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12631

G. Mihalakakou et al In this investigation, the airborne temperature over time is analysed and modelled using a neural network-based method. The forecasts made by the model can be highly helpful in the fields of weather, environmental sciences, and energy-related tasks, such as controlling passively and standard air conditioners to maintain the temperature within structures. Utilising multiple-layer back propagation systems, projected hourly readings of the surrounding temperature over a number of years were estimated by utilising the information extracted from its historical value.

Hrachya Astsatryan et al, in order to enhance the daily ambient temperature forecasting for as long as twenty-four hours in the Armenia region (Armenia), the study proposes to adopt a weather forecasting strategy that utilises artificial intelligence techniques & approaches. One of the driest areas of Armenia, it Hermon Valley experiences soaring temperatures and soaring winds 120–160 days a year due to its extreme heat as well as low humidity levels. The method makes use of enormous databases suitable for satellites study at various frequency and sensitivities as well as earth observational information obtained from many meteorology stations. Cong Li et al This research proposes a new method for utilising weather prediction data for predicting the daily ST. The daily average ST forecast and the ST amplitude, or the disparity in the hourly ST with the typical daily ST, are the two components that the approach superimposes to get the hourly ST prediction. We chose nine environmental indicators and merged two chronological variables as input variables for forecasting the mean for the day ST based on the relationship analysis results. Eight weather variables and one chronological variable were chosen as input for the ST loudness prediction challenge. Trang Thi Kieu Tran et al This paper offers a thorough analysis of methods based on artificially generated neural networks (ANNs), including short-term long-term memory (LSTM), recurrent neural networks (RNN), and others, that have been applied to air temperatures forecasting.

The works produced between 2005 and 2020 are highlighted. The review demonstrates that models using neural networks are potentially useful instruments for weather prediction. There is currently no agreement on the best strategy currently in use, despite the fact that ANN-based systems have been used extensively to estimate air temperature because of them quick processing speed and capacity for managing complicated issues. Emre Yakut et al The goal of the current work is to create models utilising ANN, ANFIS, and SVMs approaches for the average monthly air temperature readings in Turkish. The simulations were developed using daily air temperatures (a minimum, highest, and mean) obtained from eight TSMS sites covering the years 1963–2015. The data included longitude, latitude and longitude height, and date. Using R2, MSE, MAPE, and RRMSE, the results of the ANN, ANFIS, and SVMs systems were compared. In young Park et al, in this research, we use actual observed weather information to present a new neural network based temperatures forecast model. To do this, an enormous volume of high-quality modelling data is required; nevertheless, the data must be flawless.

However, as it is impossible to quantify data that has been overlooked, gathering meteorological data has limitations. Because of this, the data that have been gathered may be erroneous or contain long gaps. Consequently, in order to recover lost weather information, the suggested temperatures forecasting approach is applied to improve the information that is missing. Furthermore, the suggested model makes use of a long-term short-term memory (LSTM) network, which is a type of recurrent network neural network that is well-suited for modelling historical data because temperatures is periodic. Seongyoep Jeong et al numerous interrelated elements, such as location, geography, and time, influence weather. In order to forecast temperatures in South Korea, information from numerous areas must be used. In order to do this, we study the application of time-series weather information from an autonomous weather observatory and picture data from a local information integration and forecasting systems (RDAPS) in deep neural-network-based temperatures forecasting systems. The characteristics derived from the time-lapse observable along with the RDAPS picture data are represented by a bidirectional long- and short-term memory (BLSTM) framework and a convolutional neuronal network (CNN) theory, in order to fit these separate kinds of material into the same model.

III. CHALLENGES OF WF SYSTEMS

There are numerous approaches to weather forecasting. The effectiveness of forecasting is dependent on the various properties that are taken into account. Research is needed to determine how machine learning models, such as SVM, KNN, and RF, can be modelled for forecasting and how to improve the accuracy of weather forecasting.

- It's still difficult to understand how machine learning models arrive at their judgements.
- Good, varied data is essential for training models that work.
- The potential for combining ML with physics-based methods is enormous.
- Another research area is the application of the Classification problem for weather data and the
- comparison of the efficient methods.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12631

- Testing the efficiency of rain classification is crucial. Since the data is typically imbalanced, pre-processing is necessary to improve performance.
- The selection of the suitable ML method for forecasting is an another open challenge thus it is required to validate various classifiers first.

IV. PROPOSED CASE STUDY AND VALIDATION OF CLASSIFIERS

With the help of the Koggle weather dataset, a variety of machine learning (ML) classifiers, including Support Vector Machines (SVM), Decision Trees (DT), and Artificial Neural Networks (ANN), are validated and tested in this research in order to predict rainy weather. Additionally, an extended survey of ML and NN based WF methodologies is provided in the paper. The proposed flow chart of the WF is given in the Figure 2. It is clear that first we load the raw dataset and extract the available features train and validate the ML based modes to test the classification accuracy. In this paper the various SVM based classifiers and the ANN are validated on the raw Koggle weather forecasting database.

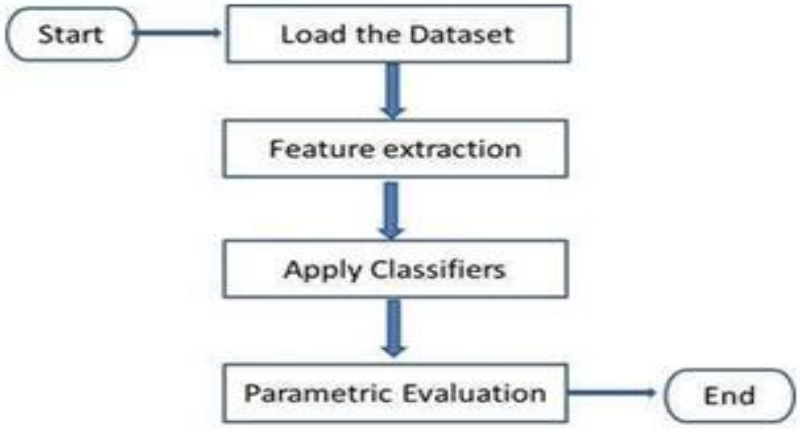

Figure 2. Flow chart of classification process for WF

V. RESULTS OF VALIDATION FOR WEATHER DATASET

The WF using ML is validated in this work. We have used 4 algorithms to do the classification of the Weather forecast dataset containing total records 1461 records with 4 features Precipitation, Temp_max, temp_min and wind and weather column (target column) with five target values – drizzle, rain, sun, fog and snow. Following are the results. The parametric performance of accuracy, precision and recall validation is shown in Figure 4 and the respective Fi score are shown in the Figure 5.

Table 2 Validation of Classifiers for WF

Algorithm	Efficiency	F1- Score	Precision	Recall
Decision Tree [16]	76.74	1.84	62.24	61.41
ANN [16]	81.54	2.21	74.36	73.69
SVM	82.99	0.7776	73.98	82.99
ANN	84.35	0.7997	84.09	84.35
Random Forest	84.35	0.7954	77.40	84.35

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 6, June 2025

DOI: 10.17148/IARJSET.2025.12631

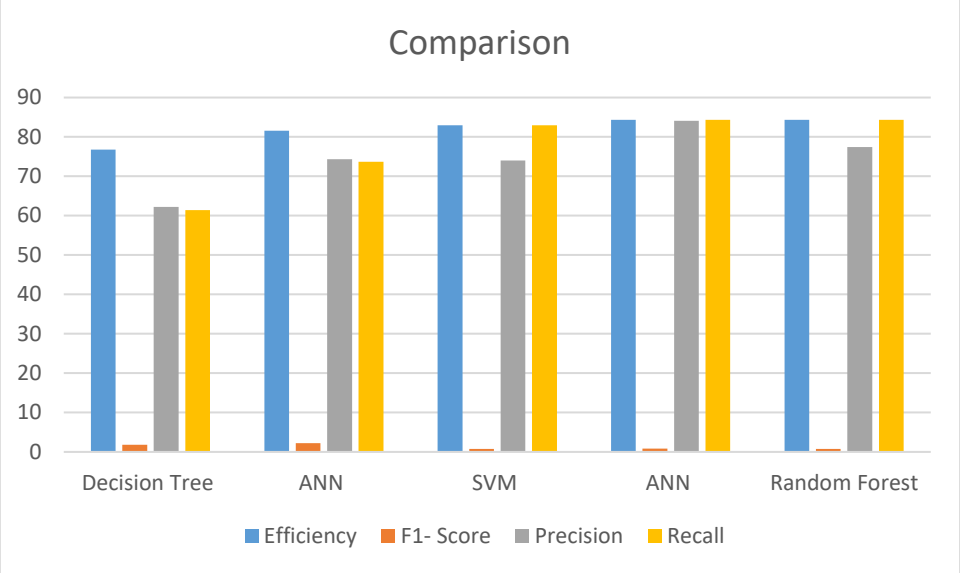

Figure 3 Comparison of performance measures

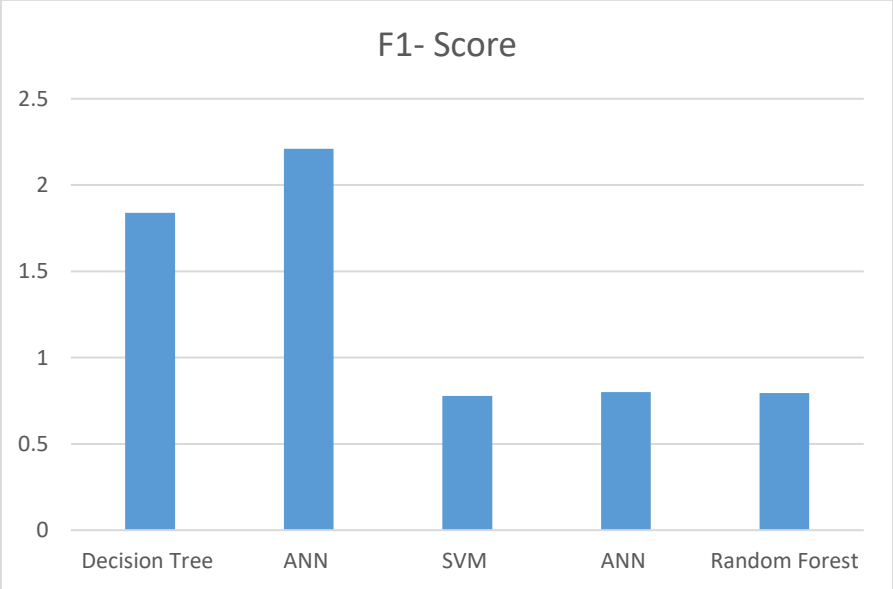

Figure 4 Validated F1 scores for classifiers

VI. CONCLUSIONS AND FUTURE SCOPES

The purpose of this study is to evaluate different machine learning (ML) classifiers for the prediction of precipitation. The Koggle weather dataset is used to test and validate a variety of machine learning techniques, including decision trees (DT), artificial neural networks (ANN), and support vector machines (SVM). Extended surveys of ML and ANN based WF approaches were also presented in the paper. Prediction accuracy is contrasted in order to verify the different machine learning-based classifiers. The five classes of data are classified using four features. It can be concluded that the significant lower accuracy is observed and maximum of 84.35 % for ANN is achieved.

There is significant chance of improvement The lower performance is due to data imbalance problem and lack of features correlation. In future the data balancing methods and extended feature set can be used for improving the accuracy. The testing and training is Implemented using the python coding.

DOI: 10.17148/IARJSET.2025.12631

REFERENCES

[1]. Bochenek B, Ustrnul Z. Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives. Atmosphere. 2022; 13(2):180. https://doi.org/10.3390/atmos13020180

[2]. A H M Jakaria, Md Mosharaf Hossain, Mohammad Ashiqur Rahman "WSmart Weather Forecasting Using Machine Learning:A Case Study in Tennessee" arXiv:2008.10789 https://arxiv.org/abs/2008.10789

[3]. Bengtsson, L. The weather forecast. PAGEOPH 119, 515–537 (1980). https://doi.org/10.1007/BF00878157

[4]. Mahakud, Rina & Pattanayak, Binod & Pati,Bibudhendu. (2022). A Hybrid Multi-class Classification Model for the Detection of Leaf Disease using XGBoost and SVM. International Journal of Engineering Trends and Technology. 70. 298-306. 10.14445/22315381/IJETT-V70I10P229.

[5]. Wang, Bin & Lu, Jie & Yan, Zheng & Luo, Huaishao & Li, Tianrui & Zheng, Yu & Zhang, Guangquan. (2019). Deep Uncertainty Quantification: A Machine Learning Approach for Weather Forecasting. 2087-2095. 10.1145/3292500.3330704.

[6]. Arca, Bachisio & Duce, Pierpaolo & Spano, Donatella & Snyder, Richard & Fiori, Michele. (2004). Use ofNumerical Weather Forecast and Time Series Models for Predicting Reference Evapotranspiration. IVInternational Symposium on Irrigation of Horticultural Crops 664. 664. 10.17660/ActaHortic.2004.664.2.

[7]. SORKUN, MURAT CIHAN; INCEL, ÖZLEM DURMAZ; AND PAOLI, CHRISTOPHE (2020) "Timeseries forecasting on multivariate solar radiation data using deep learning (LSTM)," Turkish Journal ofElectrical Engineering and Computer Sciences: Vol. 28: No. 1, Article 15. https://doi.org/10.3906/elk-1907-218

[8]. Himani Tyagi "Weather - Temperature Pattern Prediction and Anomaly Identification using Artificial NeuralNetwork" International Journal of Computer Applications (0975 – 8887) Volume 140 – No.3, April2016 DOI: 10.5120/ijca2016909252

[9]. Mohammed Al-Shawwa "Predicting Temperature and Humidity in the Surrounding Environment UsingArtificial Neural Network" International Journal of Academic Pedagogical Research (IJAPR) ISSN: 2000-004X Vol. 2 Issue 9, September – 2018, Pages: 1-6

[10]. A.E. Ruano "Prediction of building's temperature using neural networks models" Received 21 March 2005;received in revised form 29 August 2005; accepted 2 September 2005 doi: 10.1016/j.enbuild.2005.09.007

[11]. Sotiris Papantoniou "Prediction of outdoor air temperature using neural networks: Application in 4European cities" http://dx.doi.org/10.1016/j.enbuild.2015.06.054

[12]. Ashrafi, Kh "Prediction of Climate Change Induced Temperature Rise in Regional Scale Using NeuralNetwork" Int. J. Environ. Res., 6(3):677-688, Summer 2012 ISSN: 1735-6865

[13]. Brian A. Smith "An Enhanced Artificial Neural Network for Air Temperature Prediction" World Academy of Science, Engineering and Technology 7 2005.

