International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *20 August 2023* ***Revised:*** *09 October 2023* ***Accepted:*** *24 October 2023*

Machine Learning Techniques for Prediction of Climate Change in Nagpur using various Optimized Models

[1]Bharati Masram , [2]ArchanaTiwari , [3]Nita Nimbarte, [4]Rachana Sawadh, [5]Sanjay Balamwar

[1]Assistant Professor ,Department of Electronics & Telecommunication Engineering ,
Yeshwantarao Chavan College of Enginering
Nagpur, India
e-mail: bharatimasram@gmail.com

[2]Assistant Professor ,Department of Electronics Engineering ,
Shri Ramdeobaba College of Enginering
Nagpur, India
e-mail: tiwariar@rknec.edu

[3]Assistant Professor ,Department of Electronics & Telecommunication Engineering ,
Yeshwantarao Chavan College of Enginering
Nagpur, India
e-mail: nitangp@gmail.com

[4]Research Scholar,Department of Electronics & Telecommunication Engineering ,
line 2: Yeshwantarao Chavan College of Enginering
Nagpur, India
e-mail: sawadhrachana@gmail.com

[5] Maharashtra Remote Sensing Application Centre (MRSAC)
R & D Department
Nagpur,India
e-mail: sanjaybalamwar@gmail.com

Abstract— The Global temperatures are being affected by climate change. Nagpur's climate is a mix of wet and dry conditions. In the summer, the city experiences the driest weather. Because the city is approaching the extreme situation of climate change, it is vital to check the city's climate change parameters. In this study, we look at machine learning methods for predicting the climate, specifically the temperature factor for Nagpur. Three algorithms are discussed ,compare and contrast the Logistic Regression model, the Random Forest Classifier model, and the Auto Regressive approach model to find the optimum model for our prediction needs.

Keywords-Prediction, Machine learning, logistic regression model, Random Forest Classifier, Model using Auto regressive method, Dickey Fuller test.

I. INTRODUCTION

Changes in climatic conditions are the most visible manifestation of Nagpur's global warming. Nagpur, which is located in the heart of India, has been subjected to a variety of harsh weather conditions. Urbanization's effects and climate change's consequences are dangerously converging. The city contributes a significant amount to the effects of climate change. Cities utilize approximately 80% to 90% of energy and emit 50% to 60% of carbon dioxide, which is a considerable quantity. Specifically, industrialization, fossil fuels, hydropower, carbon dioxide and other harmful gases, and so on. Nagpur has a tropical climate with dry weather for the majority of the year. Estimating the likely effects of climate change necessitates the prediction of local temperatures. And compare them to see which model is best for our prediction Analyzing data through data analysis is the most popular and efficient computing method. When compared to physically base down scaling procedures, these compute at a faster rate and are easier to derive local climatic variables. Climate change is inevitable, and forecasted changes in rainfall patterns, humidity levels, and temperature ranges are expected to have an impact on many water supplies and biological systems, including agriculture. Climate-related phenomena such as unpredictable rainfall and temperature swings have a negative impact on global food supply and the economy. Rural inhabitants are particularly sensitive to adverse weather conditions due to their heavy reliance on agricultural products for their subsistence. Due to the rise in temperature, the city's water storage is also at maximum. Temperature, humidity, and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *20 August 2023* ***Revised:*** *09 October 2023* ***Accepted:*** *24 October 2023*

wind pressure are the most vital components in this environment [1][2]. We examine all factors in relation to yearly patterns in order to comprehend climate change. After that, we use our data to make predictions. We do this by employing machine learning to create the model and forecast weather [3][4].

II. SURVEY OF ENVIRONMENTAL CONDITION IN NAGPUR

A. Literature Survey

Climate-related phenomena such as unpredictable precipitation and temperature swings have a negative impact on global food supply and the entire economy. Due to their strong reliance on agricultural produce for survival, farming communities are particularly vulnerable to and at danger from unfavorable climate conditions. Additionally, due to the rise in temperature, the water storage in town is at capacity [5]. Temperature, moisture, and wind pressure are the key issues in this environment. We tend to study all data according to yearly trends and detect temperature variation throughout this time. Following that, we have a proclivity to apply our understanding to forecasting. For this, we usually use machine learning to create a model that mistreats the climate and predicts it. As the temperature rises, so does the demand for water. Because it has rained every few days, Associ*ate in Nursingnum once a year* each year has been an exception. Otherwise, the mercury consistently hovers around 45 degrees astronomer toward the end of the Gregorian calendar month and into May, and in May, typically during a strong heat wave, it reaches around 48 or 49 degrees astronomer. [6].

B. Water supply to Nagpur

- Water is drawn from irrigation reservoirs by the Nagpur Municipal Corporation (NMC). Nagpur gets water from the Navegaon Khairi dam on the Pench River and the Kanhan River.. Nagpur requires 640 million liters per day (MLD) of water for the majority of the year. During the Gregorian calendar month of June, demand climbs to around 720 MLD.
- As a result, this is frequently all about the water of Nagpur, which has an impact due to changes in the environment. Nagpur's earth science is similar to that of the surrounding parts of the district because it has a similar structure for rains and therefore for water. [4][5] The Nagpur Municipal Corporation (NMC) and the Nagpur Metropolitan Area (NMA), which also includes the NMC, the Kamthi Camp Board, the two municipal cities of Kamthi and Kalmeshwar, and 262 neighboring villages, are separate organizations that develop and carry out the town's development and welfare policies. Historically, the district's precipitation has been evenly distributed, with little variance in yearly precipitation from year to year. The average annual precipitation is 1161.54 mm, with approximately ninety per cent falling during the south-west monsoon season, peaking in July. We tend to go through the entire reading of environmental condition alteration during this work.

III. EASE OF USE OF ML

For each survey, we can infer that it is required to analyze the climate of Nagpur and, more importantly, it is necessary to forecast upcoming events using cutting-edge technology. As a result, we tend to analyze it using the Python programming language. Python is a high-level, general-purpose, and extremely popular programming language. Python programming language (latest Python 3) is used in web development, machine learning applications, and all other cutting-edge technology in the software package industry. Following that, we usually train the model for prediction purposes. We often use victimization machine learning approaches to compare two to a few algorithms for accuracy. Machine learning is a branch of computer science known as artificial intelligence. The goal of machine learning is to understand the structured knowledge of information and to match that data into models that people can understand and use. [7] Despite the fact that machine learning is a branch of computing, it is distinct from traditional process techniques. Algorithms are sets of explicitly coded directions used by computers to calculate or solve problems in old computing. Machine learning algorithms, on the other hand, train computers on knowledge inputs and utilize applied mathematics analysis to produce values that fall inside a specific range. Machine learning assists computers in developing models using sample knowledge in order to adjust decision-making processes based on knowledge inputs as a result of this [8]. Machine learning has being used by almost every technology user these days.

IV. ANALYTICAL STUDY AND RESULTS

After analyzing the data, train and test the model in python using jupyter notebook on several strategies to predict climate change[9][10], then choose the best model as shown in Figure .1 to deploy in a production application.

A. Modelling

We must construct a batch to train and test data in order for the prediction to be accurate. We must divide the data into two domains. To develop a model, you'll need to use all of the measurements that are included in the datasets.

Generalized formula for the prediction

$$Y = X_1 + X_2 + X_3 ..... + X_N . \quad (1)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *20 August 2023* ***Revised:*** *09 October 2023* ***Accepted:*** *24 October 2023*

Where,

Y = dependent variable (predicted outcome variable)
X = independent variable (predictor variables)

Split and train python has in build library named as sklearn.model_selection. We have to split the data into two parts as one is test and other is trained.
The test size is always less than that of the train size .In our model we are assuming or declared the test size is 25% and other hand the train size 75%.random state- to maintain the certainty of the program.

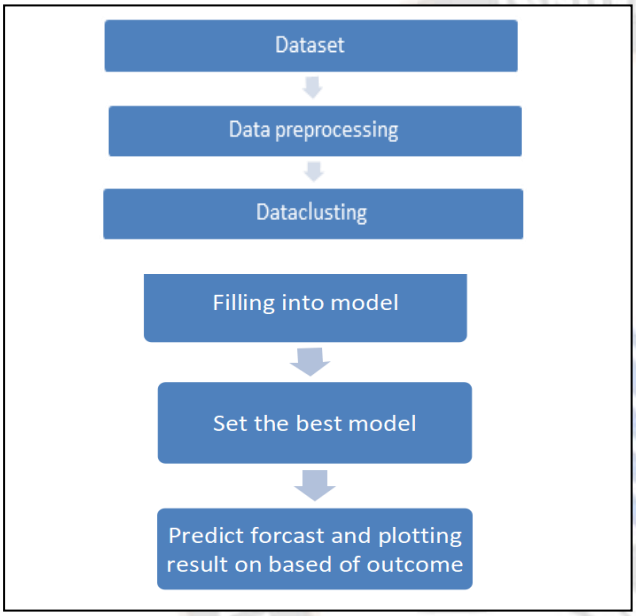

Figure.1.Process of Modelling

B. Algorithm Application on Data Sets

Logistic Regression model: As seen in Figure 2, logistic regression sometimes referred to as logistic model or logistic regression is a mathematical model used in statistics to estimate (guess) the likelihood of an event occurring given some prior data. With binary data, the event either occurs (1) or it does not occur (0), logistic regression can be used.

Formula for logistic regression is given as

$$Z = C + w_1Y_1 + w_2Y_2 + \ldots\ldots + w_nX_n \quad (2)$$

Where Y1, Y2 …., Yn being features of the observation, and w1, w2… wn are the 'weights' of each feature. For the most prominent feature is in the final decision in that is the higher its weight. [10][11].

Results using Logistic Regression:

In this model ,any real-valued number can be transformed into a value between 0 and 1 using the logistic, but never precisely at either end.

$$LogisticFunction = \frac{1}{1+e^{-value}} \quad (3)$$

Where value is the actual numerical value that you want to transform and e is the base of the natural logarithms (Euler's number or the EXP ( ) function in your spread sheet). Below Figure 2.shows a plot of the numbers between -5 and 5 transformed into the range 0 and 1 using the logistic function.

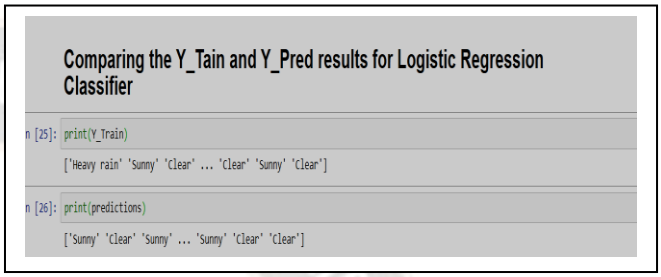

Figure.2 .Model of Train and test data using Logistic regression

The function follows supervised learning method and it dedicated towards classification task, In logistic regression model, a linear regression equation to produce discrete binary outputs .In our data as a primary data is not classified frequency .so the output of train and test data is not showing the result with the accuracy.

In this model the data sets of the train and test box contradict each other's data set and then prediction using this method is not accurate for this experiment.

Model using Random Forest Classifier: A decision tree creates classification or regression models within a tree structure. It gradually divides the information set into smaller Associate in nursing subgroups while simultaneously creating a related call tree. A tree with both calls and leaf nodes is the final product. A choice node can have two or more branches. A leaf node represents a category or call. In an extremely tree, the root node is the highest call node and corresponds to the most straightforward predictor [11][12][13].Each categorical and numerical information will be handled using call trees. Random Forest creates many call trees that are linked together for a more accurate forecast. The Random Forest model is based on the idea that a group of unrelated models (individual call trees) performs considerably better than they do independently.

Result using Random Forest Classifier:

The decision tree-based Random Forest Classifier in Figure 3 is typically trained using the "bagging" technique. The fundamental tenet of the bagging method is that the output is improved by combining different learning models. Random forest builds many decision trees and combines them to

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *20 August 2023* ***Revised:*** *09 October 2023* ***Accepted:*** *24 October 2023*

produce a more accurate and trustworthy prediction. [14][15][16].

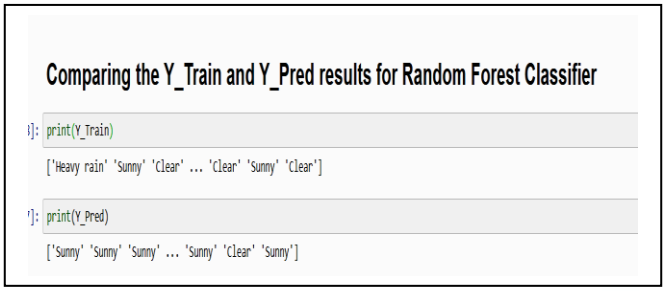

Figure.3. Model of Train and test data using Random Forest classifier

A random forest is a Meta estimator that is created by fitting several decision tree classifiers to different dataset subsamples and using averaging to increase predictive accuracy and reduce over fitting. With respect to the logistic regression model [17][18] [19].

If bootstrap = True (the default),
The sub-sample size is controlled by the max samples parameter; otherwise, each tree is constructed using the entire dataset.

But on heavy side of the train data it could not perform the accuracy with respect to the test data So this classifier is not used for the forecasting of the data.[20][21].

Model using Auto Regressive Method

A time series model known as auto-regression uses observations from previous time steps as the input to a regression equation to forecast the value at the following time step. It is a straightforward idea that can generate accurate forecasts for a number of time series problems.. Proposed system goes over how to use Python to create an auto-regressive model for time series forecasting. A model that uses auto-regression makes the assumption that previous time step observations are useful for forecasting the value at the following time step. The term "correlation" is used to describe the connection between variables. When two variables change simultaneously in the same direction (e.g., go up or down), there is a positive correlation. When the variables move in opposite directions as their values alter (for instance, when one goes up and the other goes down), this is known as a negative correlation. We used a graphical representation as shown in Figure 4 to check the comparison of test and prediction results.

Auto-regressive model predicts up-to 200 values. That means it works accurately for our prediction. We consider on x-axis count of prediction Vs. temperature. An autoregressive model predicts output behavior based on act summarized calculation of the train and test data outcome. It corrects the accuracy when there is some correlation between values in a time series and the values that precede and succeed them .

The model in this experiment creates a linear regression of the data in the current series against one or more past values in the same series by only using past data to model the behavior. At some time t, the outcome variable (Y) value is directly correlated with the predictor variable (X), much like "regular" linear regression. The difference between simple linear regression and AR models is that Y depends on X and Y's prior values.

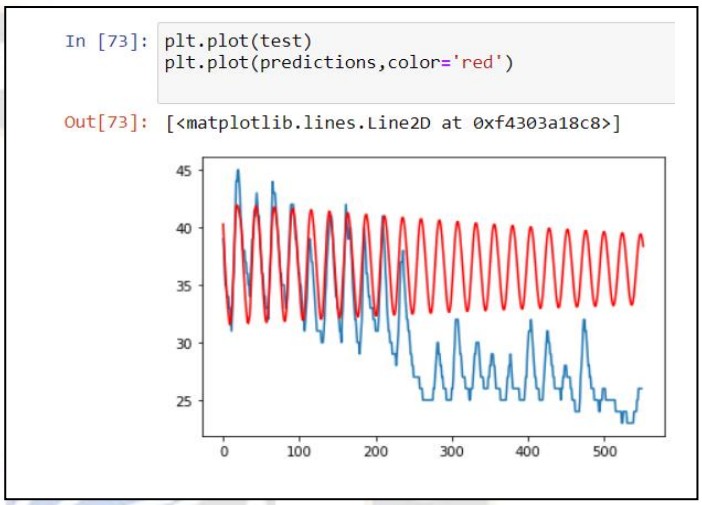

Figure.4. Graphical representation of output data using Auto regressive method

The outcome shows the 98% of accuracy with respect to the accrual data that means it perfectly fit for our data and it draws up to 300 number of prediction

V. TIME SERIES ANALYSIS AND PREDICTION

Before Resampling

The basic analysis on Time Vs. Average temperature is as shown in Figure 5.In these graph TSA shows the rise and fall between the averages of the temperature. According to government study 2003 year records highest temperature and it can be proved from this graph.

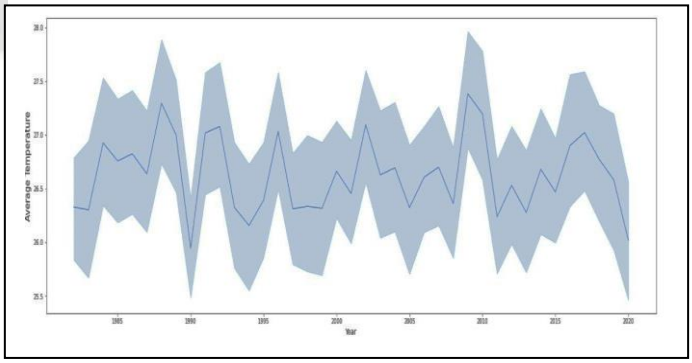

Figure 5. Time series analysis of the average of the temperature (before resampling)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 20 August 2023 **Revised:** 09 October 2023 **Accepted:** 24 October 2023*

After Resampling

The Basic analysis now for after resampling can be shown as in Figure 6. The city survey calculates , the sudden rise into gap of 2-3 years .This situation occurs maximum in time interval of 2 years .In 2009 the city records highest average temperature according to the graph. City calculate sudden fall in year of 2005 and 2020. This fall is mainly due to less emission of the CO_2 and other parameter to reduce in change of climatic condition. Before resampling our TSA analysis it shows the boundary around average of the temperature, get distracted to understanding data.

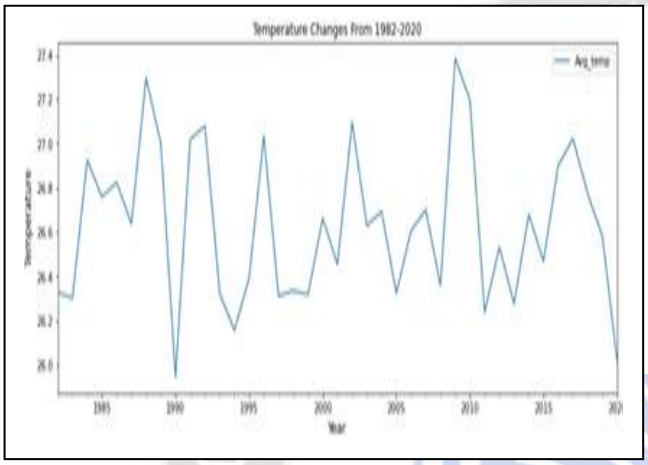

Figure 6. Time series analysis of the average of the temperature (After resampling)

Dickey Fuller Test Results:

The augmented Dickey Fuller (ADF) test predicted results as shown in Figure .7 for unit roots was conducted for all the time series used for the study. First graph shows the time series analysis after resampling trend how patters work fluctuate with respect to average temperature seasonal it checks the frequency of the graph residual length of the time series analysis according to our data we see that this test show us the categorical frequency of the time interval as you can see the seasonal data repeats the pattern after every 2-3 year. That means it has continuous frequency. This frequency helps us to forecasting purpose.

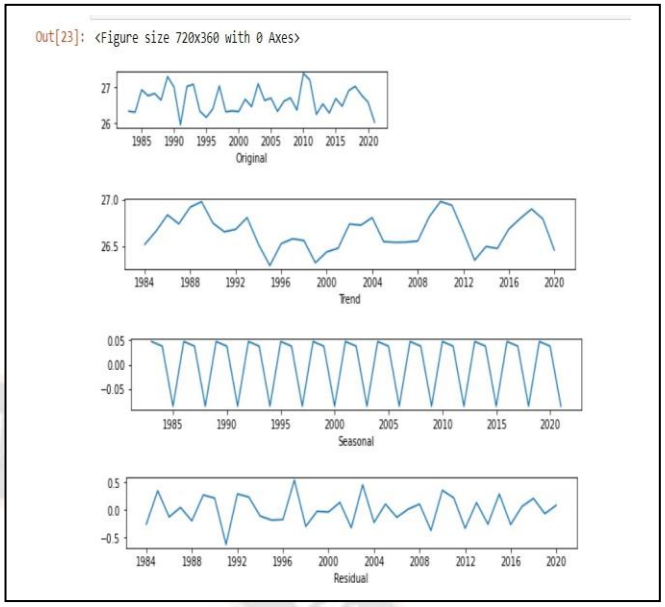

Figure 7. Dickey Fuller Test - Predicted Results

5 Conclusion

Prediction of climate change using computational way is one of the important topic and mainly in cities and advance more than other system. We believe that our work design will perform important role in climatic change prediction. This system has many advance technique like analysis of past data using time series analysis and prediction using machine learning like Logistic Regression model, random forest model, auto-regressive model.

Based on this collected data, we showed how to use the Machine Learning algorithm to forecast future mean weather temperatures. Machine learning is one of the most advance technique and useful for prediction. The train models for various parameters are tested accordingly and fit the model .The model next compare with the algorithms and fit the best model for prediction purpose. This results are gives us accuracy for readings.

Three models are compared for accuracy, and the best model is chosen. We showed how to choose statistically significant predictors based on reliable statistical techniques by using the random forest, logistic regression, and auto-regressive model.

Using this data, we then fitted a prediction model based on a training subset. Using the auto-regressive model we can predict climate changes using time series. As finally we select auto-regressive model for prediction as in this we choose it because it gives better result as the other two models.
Based on inputs from a testing subset and this fitted model, we predict the expected values and assess the prediction's accuracy, which shows a respectable level of accuracy.

This model gives us accuracy according to the future 200 count. This is best result as we get according to the model this could help us for storage of water supply, vegetation, fuel, road construction, and other supply which depend on climate change

ACKNOWLEDGMENT

The first author expresses his gratitude to Maharashtra Remote Sensing Application Centre (MRSAC), Nagpur for providing us the assistant for the research work .She is extremely thankful to all people who directly or indirectly guided us for the idea behind the study of Climate change in Nagpur. Second Author also would like to thanks to Mr. Sanjay Balamwar, Associate Scientist MRSAC, and Nagpur for his constant support for analyzing and confirmation to the result provided above.

REFERENCES

[1] V. A. Vuyyuru, G. Apparao and S. Anuradha, "Rainfall Prediction Using Machine Learning Based Ensemble Model," *2021 5th International Conference on Information Systems and Computer Networks (ISCON)*, Mathura, India, 2021, pp. 1-4, doi: 10.1109/ISCON52037.2021.9702409.

[2] D. Fister a , J. Pérez-Aracil a,∗ , C. Peláez-Rodríguez a , J. Del Ser b,c , S. Salcedo-Sanz a "Accurate long-term air temperature prediction with Machine Learning models and data reduction techniques"Volume136,2023,110118,ISSN,15684946,https: //doi.org/10.1016/j.asoc.2023.110118.

[3] Wanie M. Ridwan, Michelle Sapitang, Awatif Aziz, Khairul Faizal Kushiar, Ali Najah Ahmed, Ahmed El-Shafie,Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia,Ain Shams Engineering Journal,Volume 12, Issue 2,2021,Pages 1651-1663,ISSN20904479,https://doi.org/10.1016/j.asej.2020.09.011.

[4] M. Karamouz , M. Fallahi , S. Nazif and M. Rahimi Farahani " Long Lead Rainfall Prediction Using Statistical Downscaling and Articial Neural Network Modeling",Transaction A: Civil Engineering Vol. 16, No. 2, pp. 165-172 (2009) .

[5] Navneet Kumara , Bernhard Tischbeinb , Jürgen Kuschec , Patrick Lauxd, , Mirza K. Bege , Janos J. Bogardi," Impact of climate change on water resources of upper Kharun catchment in Chhattisgarh, India", Journal of Hydrology: Regional Studies,Voume 13 (2017), Pages 189- 207.

[6] Cities and Climate Change :Responding to an Urgent Agenda, World bank Report

[7] Guide to Climate Change adaptation in cities world bank report

[8] Subhajit Bandopadhyay," Does Elevation Impact local level climate change?An anaysis based on fifteen years of daily diurnal data and time series forecasts",pacific science review A:Natural science and Engineering volumes 18,issue 3,(2016),Pages 241-253.

[9] F.Giogi,L.O.Mearns,"Probability of regional ,Climate Change based on the Reliabilty Ensemble averaging (Rea)Method "Geophysical research Letters ,Vol.30,No.12,1629,(2003).

[10] Micheal Qu Joe Wan and Xianjun Hio,"Analyis of Diurnal Air Temperature range Change in the Continental united states",article in weather and climate extremes.(2014).

[11] Climate Change And Its Impact on Nagpur's Water Supply International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181,Vol. 4 Issue 10, October-2015.

[12] HaoruiChen,i Wang, ZhanyiGao, Yaqiong Hu "Artificial Neural Network Approach for Quantifying Climate Change and Human Activities Impacts on Shallow Groundwater Level" A case study of Wuqiao in North China Plain

[13] Bharati Masram, Rachana Sawadh,Sanjay Balamwar," Analysis and Prediction of climate Change with Artificial Intelligence Due to Multiple parameters in Nagpur", Zeichen Journal ,ISSN No: 0932-4747 , Volume 6, Issue 12, 2020, pp 19-23.

[14] Rachana Sawadh, B .Y. Masram, Sanjay Balamwar, "Analysis of Past Behaviour Of Climate Change on Basis Of Multiple Parameters In Nagpur " Zeichen Journal ISSN No: 0932-4747Page No:672-680 .DOI:15.10089.ZJ.2020.V6I3.285311.2272.

[15] Judith A. Ouellette, Wendy Wood,"Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior", 1998, Vol. 124, No. 1, 54-74 article American Psychological Association, Inc. 1998, Vol. 124, No. 1, 54-74 0033-2909198/$.

[16] D.Deva Hema, Anirban Pal, Vineet Loyer and Rajeev Gaurav, "Global Warming Prediction in India using Machine Learning", International Journal of Engineering and Advanced Technology (IJEAT), vol. 9, no. 1, October 2019, ISSN 2249–8958

[17] M. Sharifzadeh, A. Sikinioti-Lock and N. Shah, "Machine-learning methods for integrated renewable power generation: A comparative study of artificial neural networks support vector regression and Gaussian Process Regression", Renewable and Sustainable Energy Reviews, vol. 108, pp. 513-538, 2019.

[18] S. Parkpoom and G. P. Harrison, "Analyzing the Impact of Climate Change on Future Electricity Demand in Thailand", IEEE Transactions on Power Systems, vol. 23, no. 3, pp. 1441-1448, 2008.

[19] Climatic change and the Mediterranean: L. Jeftic, J.D. Milliman and G. Sestini (Editors), 1992. Edward Arnold Publ., Sevenoaks, Kent. Hardcover 673 pp. Price: £89.50 ISBN 0-340-55329-4

[20] Nita Nimbarte, Aniket Nagpure, Badal Sanodiya, Harshal Sevatkar, Sanjay Balamwar, "Knowledge Based Classifier and Pattern Recognition Technique for Satellite Image Analysis", IEEE GCAT 2022 Conference 07th -09th Oct 2022, Bangalore.

[21] N. Y. Yun Ng, H. Gopalan, V. S. G. Raghavan and C. C. Ooi, "Day-Ahead Forecasting for the Tropics with Numerical Weather Prediction and Machine Learning," *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Singapore,

Singapore, 2022, pp. 125-130, doi: 10.1109/ICARCV57592.2022.10004290.

AUTHORS PROFILE

Bharati Masram has received her Ph.D degree in the area of VLSI signal processing under the research centre of YCCE from RTM Nagpur University, India in 2020. She received her master degree and B.E degree in Electronics Engineering from Yeshwantrao Chavan College of Engineering (Autonomous) in 2010 and 2004 respectively.Her research interest is mainly in VLSI Signal Processing, Communication; 3D Image processing etc With this she she is also now working on machine learning for the analysis and prediction of weather of Nagpur city due to unpredictive nature of climatic change in current .She has published her total 20 research based papers in national, international conferences and also in International journal. She had on her name 3 published & 2 patent granted in National & Interbnational level .

Archana Tiwari was born in 1983 in Nagpur, Maharashtra, India. She graduated from Nagpur University in Maharashtra with a B.E. in Electronics and Communication in 2004 and an M.Tech in VLSI Design in 2010,respectively. She is attending Nagpur University to complete her Ph.D. in antenna design.From 2005 to 2006, she was a lecturer in the Shri Ramdeobaba College of Engineering'sElectronics Design Technology department is in Nagpur. She has been an assistant professor at Shri Ramdeobaba College of Engineering in Nagpur's Electronics Engineering Department since 2009. She is the author of one book chapter, filed one patent and has more than 10 research papers to her credit. Her research interests include Microwave Engineering, AntennaDesign, Electromagnetism, Waves and Propagation.

Dr. Nita M. Nimbarte (Rehpade) is an Assistant Professor in the Department of Electronics and Telecommunication, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, INDIA. She specializes in the Digital Image Processing domain. She has presented her work in 13 Journals and 12 International Conferences. She is a life member of IETE and ISTE professional society.

Rachana Sawadh has received her master degree in the area of Machine learning from YCCE under RTM Nagpur University, India in 2021. .She had published her total 05 research based papers in national, international conferences and also in International journal.

Er.Sanjay Balamwar Working as a Senior Scientist, Maharashtra Remote Sensing Applications Centre, Dept. Of Planning Govt. of Maharashtra, VNIT Campus, S.A. Road Nagpur -10 from last 21 Years. He had his Ph.D. in Investigations of Dielectric properties and related thermodynamics of some polar liquids at Microwave Frequency submitted to RTM, Nagpur University, Nagpur. Received various awards at state and central level including the AGI India Awards 2022 for prestigious project of MahaBHUMI which will going to transformvarious government schemes and infrastructure development programs in the State of Maharashtra.