Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777548

PREDICTION OF CLIMATE CHANGE AND TEMPERATURE DETECTION USING DEEP LEARNING

Dr. B.Jaishanthi [1] , Dr. B.Swaminathan [2], S.S.Bhaviya [3], Dr.J.Premalatha[4], Dr.Beulah Jayakumari[5]

[1]Associate Professor, [2,5]Professor, [3,4]Assistant Professor
[1,2,3]Dept of Computer Science & Engg.,
[4]Dept. of Electronics and Communication,
[5]Dept. Of Information Technology
[1,3,5]Tagore Engineering College,
[2] Saveetha school of Engineering,
[4] Chennai Institute Technology

Abstract - Weather changes in last few years have been increasing enormously and it is expected to increase more in the future. Therefore, it is a wearisome process to investigate larger forms of climate, temperature changes in the data and perform predictive analysis of the same using traditional methods. This paper aims to project temperature and seasonality changes using predictive analysis with the help of various machine learning techniques such as Time series forecasting using ARIMA & SARIMAX, Linear Regression and Auto-Correlation technique. The system proposed serves as a tool which takes in the climatic changes from huge amount of data as input and predicts the future temperature with max, min and average temperature in an efficient manner. We will be predicting the temperature change from 1992-2024 with the detailed forecast and changes from 2020- 2024 and predicting the accuracy in the changes. Predictive analytic model internment relationships among various features in a data set to assess risk with a particular set of conditions to assign a weight or score. These patterns of weight/score found in historical data can be used for predicting the future climatic and temperature changes.

Index Terms – Linear Regression, ARIMA, SARIMAX,Auto- correlation.

I. INTRODUCTION

Weather change is a crucial challenge in the recent years. It influences many factors of the environmental ecosystems such as, bio-diversity, soil erosion and changes in sea-water level .Weather forecasting alleviates the economic crisis and stimulates better public health to maintain the quality of life. Safety and well-being of human is highly implacable by the climate changes. It also assists in the agricultural field as it is an essential part of planning the farming operations. Farmers have to make optimal decisions for cultivation of crops using prediction of weather, whether to undertake or withhold the sowing operation. The consequences of unseasonal changes in weather and their potential negative effects on host plants and pests are very well known. Unexpectedly high temperatures may lead to lower plant productivity and more pests on farm. Industries like energy consumption and food security can also benefit from weather forecasting. As the key problem of weather forecasting, air temperature prediction has manifold benefits for the environment, agriculture and industry. The impact of temperature on morbidity and mortality can be assessed at both the seasonal and daily level. Extreme temperature changes Due to harsh environment, arises the lack of access to safe water and food, it can also cause Heat-aggravated and respiratory illness. Forecast of the temperature, soil surface temperature, energy consumption, and solar radiation is related to ambient temperature forecasting. Temperature forecasting is useful in understanding the probability of storm, wildfires, drought and flood occurrence in an area. . This leads to an incomplete understanding of the atmospheric processes, so it restricts weather prediction up to a 1 month time period, because beyond that weather forecasts are significantly unreliable. But Machine learning is relatively robust to most atmospheric disturbances as compared to traditional methods. Another advantage of machine learning is that it is not dependent on the physical laws of atmospheric processes.

Data mining with machine learning are the best way toward extracting vital data from the extensive information collection. The procedure of concentrate important data portrayed as information revelation that can be connected on any extensive informational index.

Temperature prediction is an notoriously sophisticated and resource consuming task. Changes in Temperature are caused by many factors. Parameterization of those features is a difficult task to achieve due to their dynamic nature. Recent development in the field of artificial intelligence, machine learning can help provide computationally less expensive solutions. We can estimated the forecasts using several machine learning techniques without a need of wide mathematical calculations by evaluating historical temperature data. The purpose of this paper is to study current research on temperature forecasting and compare machine learning techniques in the field. Various inputs to these machine learning models were also tested to determine the usefulness of each, as measured by their contribution to lowering the difference between predicted values and the ultimate observed ground truth. Machine Learning algorithms have been widely used for complicated data. Pattern analysis and recognition of Temperature data can be simplified with use of machine learning algorithms. Air temperature data is classified as part of time series statistics up. Hence, use of Linear regression, Auto correlation algorithms to estimate the future value of temperature seems a plausible solution.

The purpose of this thesis was to analyse current research on weather forecasting and compare machine learning techniques in the field. Various inputs to these machine learning models were also trained and tested to determine the usefulness of each, as measured by their contribution to lowering the difference between predicted values and the ultimate observed ground truth. It was found that including weather forecast data in the prediction models resulted in a 0.56 % reduction in mean absolute error (RMSE) for one-hour predictions when compared to using historical observations alone, and a 40.2% reduction in RMSE for 24-hour predictions. Results from several machine learning techniques were compared, with Random Forests achieving the lowest error rate. In addition, inclusion of weather forecasts from nearby areas resulted in a 4.6% lower mean absolute error (MAE) in one-hour predictions and a 20.9% reduction in 24-hour predictions when averaged across the more than 1city is studied.

II. RELATED WORK

In the study conducted, various machine and deep learning techniques are available. An early paper in this field by B. Ustaoglu , tests three different kind of ANN based methods like radial basis function (RBF), feed-forward back propagation (FFBP) and generalized regression neural network (GRNN) were studied. It compares the answers with traditionally used Time series forecasting using ARIMA AND SARIMAX model, multiple linear regression (MLR) and Auto-correlation model and obtained notable improvements over MLR outputs. A paper by Z. Karevan [10] defines a black box idea: use of Machine learning methods such as Elastic Net and k-NN for the process of feature selection then trains model using Support Vector Machine Regression with Least Square loss function to forecast maximum and minimum temperature. .Later, R. Isabelle used Recurrent Convolutional Neural Networks for weather forecasting and visualization [11] in which they propose use of convolution filters +LSTMs. The results were substantially better in comparison with popular methods. Yet another technique was proposed by P. Hewage [4], where they used multiple features like pressure, temperature, precipitation, wind, humidity and moisture to predict future value of the same feature. This was executed by implementing several machine learning and deep learning algorithms like TCN [13], LSTM with multi-input multi-output and multi-input single-output methods.

The MSE of the model is remarkably low, but model does not satisfy the criteria recommended by A. Kumar [3]—the performance of the model can be acceptable if: NMSE ≤ 0.5. Satellite images also used as input to the model; when combined with estimates for clear-sky solar radiation, the forecasts can be quite effective was proposed by Miller et al., 2013 and Linares-Rodriguez et al., 2013.

Sub-Dimension	Possible Values
A: Test Planning	Testing cost estimation
B: Test Case Management	Test case prioritization Test case refienement Test case evaluation
C: Debugging	Fault localization Bug prioritization Fault prediction

Table 1: Testing general activity dimension

III. PROPOSED WORK

The main objective of this study is to analyse temperature and seasonality change in region (temperature and seasonality changes like winter, spring, summer) for approx. 10-20 years on the basis data recorded at different place across tamilnadu and predicts the temperature changes from 2020-2024. The temperature change can be predicted more accurately while comparing with the existing work

The objectives of the work are:

1) To forecast the seasonality and temperature changes for next 10 years.

2) To create a graphical interface based on the prediction, and to make it easier for understanding.

There are various Machine Learning Algorithms to predict data like , linear regression, Support Vector Machine, elastic net, lasso, light gm regressor, Auto Regressive Model, Time series analysis using arima and sarimax. We have tried many of the algorithm to get the highest accuracy. All the methods have different working procedure. The working of some of those predicting techniques are explained as follows:

1) Linear Regression

Linear Regression is a technique that gives a relationship between an independent variable or explanatory variable and dependent variable or scalar variable. In this method the relationships are model using linear predictor function. The data is trained by this method in the model. Linear predictor function is used to create an object of that function which is used
for prediction. After object creation, the data is forecasted for future.

2) Support Vector Regression

Support Vector Machine has a component called Support Vector Regression. The same standards that SVM adheres to are likewise observed by SVR. When compared to other methods, the prediction procedure for support vector regression is complex. Calculating the algorithm is more challenging. Linear Regression provides the highest level of accuracy out of all the techniques. In comparison to other techniques, linear regression has a less complexity.

3) Time Series Forecasting using ARIMA and SARIMAX:

The acronym ARIMA, which stands for "Auto Regressive Integrated Moving Average," refers to a class of models that uses a time series' own previous values—specifically, its own lags and lagged prediction errors—to "explain" the time series in order to predict future values.
ARIMA models can be used to model any "non-seasonal" time series that has patterns and isn't just random noise. An ARIMA model is characterized by 3 terms: p, d, q where,

- p is the order of the AR term
- q is the order of the MA term
- d is the number of differencing required to make the time series stationary

If a time series, has seasonal patterns, then you need to add seasonal terms and it becomes SARIMA, short for 'Seasonal ARIMA'.

4) Elastic Net:

Elastic Net first emerged as a result of critique on lasso, whose variable selection can be too dependent on unstable data. As a solution to it combine the penalties of ridge regression and lasso to get the best of both methodss. Elastic Net aims at minimizing the following loss function: where α (alpha) is the mixing parameter between ridge ($\alpha = 0$) and lasso ($\alpha = 1$). Some modules which are required to develop the Global Warming Prediction System are briefly explained below:

Data Collection module:

Module consists of collecting the raw data from different data set. Then the data set is changed as according to the requirement. This raw data cannot be used directly for prediction purpose. It needsto to be cleaned and pre-processed.

Data Pre-processing module:

The data is cleaned in this module. After cleaning the data, it is grouped as per requirement. Grouping of cleansed data is known as data clustering.

Then it needs to be checked for missing value in the data set or not. It there is any missing value then change it with any default value. After that if any data need to change its format, it is done. Total process before the prediction is known is data pre- processing. After that the data is used for the prediction and forecasting step.

a) Data Prediction and forecasting:

In this step, the pre-processed data is taken for the prediction. This prediction can be done in any process which are mentioned above. But the Linear Regression algorithm scores more prediction accuracy than the other algorithm. So, in this project the linear regression method is used for the prediction. For that, the pre-processed data is splitted for the train and test purpose. Then a predictive object is created to predict the test value which is trained by the trained value. Then the object is used to forecast data for next few years.

b) Visualization:

In this step, the predicted and forecasted data is used to provide a graphical interface separately. At first the predicted data is plotted in a graph separately with the help of matplot library. Then the forecasted data of temperature is plotted in graph with proper scale. Then the greenhouse gases forecasted data are plotted in a single graph with proper scale.

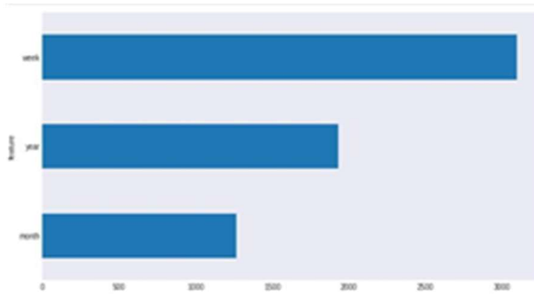


Fig 1: Visualization of temperature change by month, year, week

5) Auto Regressive Model:

Autocorrelation and Partial Autocorrelation

- Identification of an AR model is often best done with the PACF.
- For an AR model, the theoretical PACF “shuts off” past the order of the model. The phrase “shuts off” means that in theory the partial autocorrelations are equal to 0 beyond that point. Put another way, the number of non-zero partial autocorrelations gives the order of the AR model. By the “order of the model” we mean the most extreme lag of x that is

used as a predictor. If both variables change in the same direction (e.g. go up together or down together), this is called a positive correlation. If the variables move in opposite directions as values change (e.g. one goes up and one goes down), then this is called negative correlation.

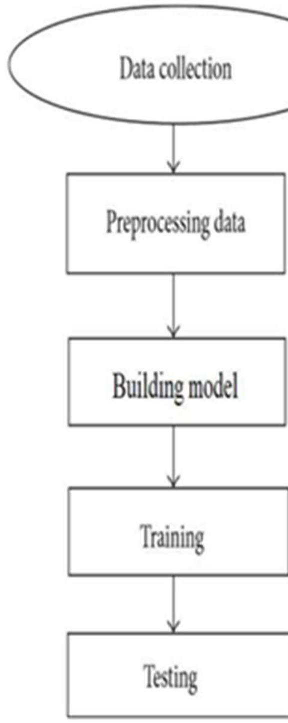
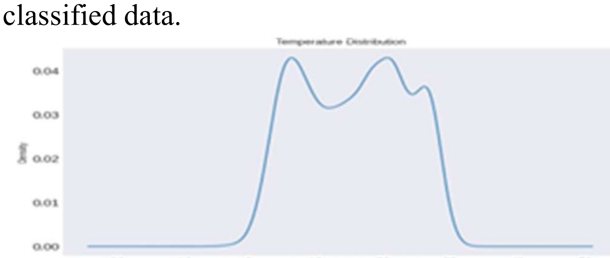


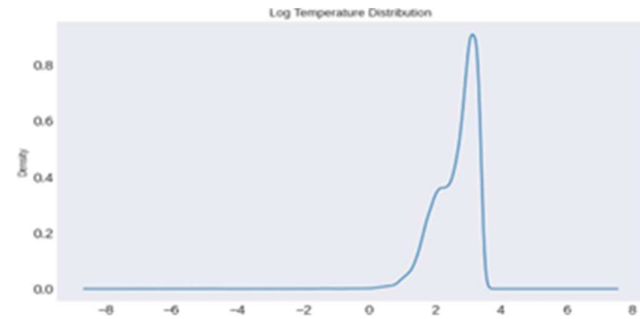
Fig2: Data flow model

IV. EXPERIMENTAL RESULTS

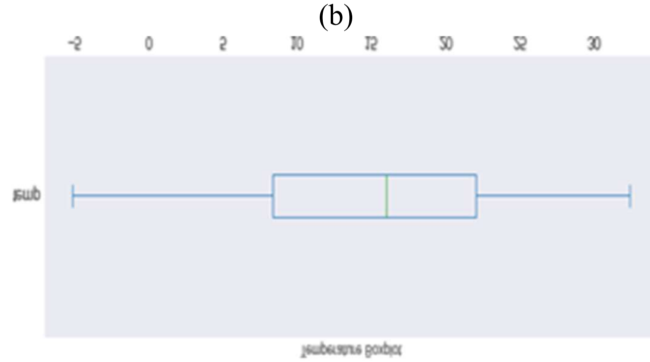
In this paper, the temperature changes from 1992-2024 can be predicted using predictive analysis using various machine learning algorithms and we will be predicting the future temperature based on the existing data which we took(a) we train, test various machine learning techniques(b) and we will be predicting the future temperature and seasonal changes(c), loss functions and one optimizer have been used to find the loss and accuracy of the proposed model(e). The result in test 1 has better accuracy compare with other tests so at this case we conclude the result here that best accuracy is obtained(f). The results are better if and only if we train ,and test a wide variety of data so that we can give an approximation of temperature change based on the classified data.



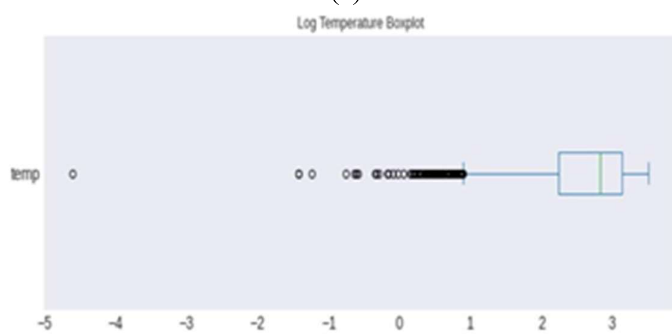
(a)


Log Temperature Distribution
Density

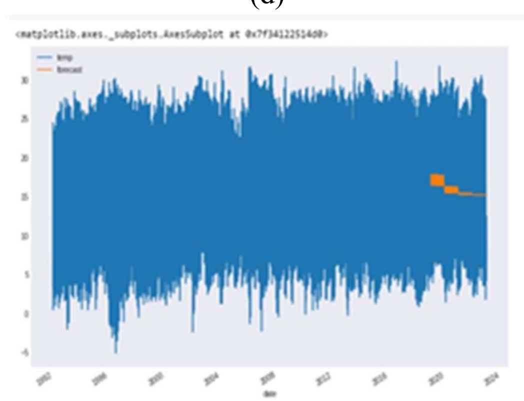
(b)



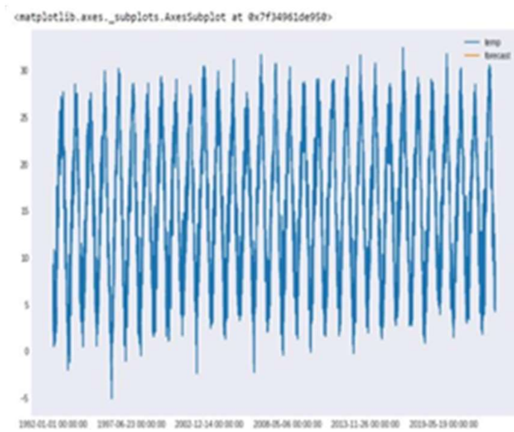
(c)


Log Temperature Boxplot
temp

(d)


<matplotlib.axes._subplots.AxesSubplot at 0x7f34122514d0>
temp
forecast

(e)



(f)

V CONCLUSIONS AND FUTURE WORK

In this paper, a model for temperature forecasting is proposed. The objective of this study was to forecast the daily mean temperature using three machine learning models, EDA AND DATA PRE- PROCESSING and smoothing technique. Model showed acceptable and significantly high performance in terms of root mean squared error, mean squared error and mean absolute error on predicting average next day temperature of chennai. Model used the previous 'n' day's temperature where the optimal trade-off resolution method selected the value of n. Smoothing dataset played an important role in performance of the model. In our work we will be predicting the weather from 1992-2024 in whic+h for some particular date and year we can tell how much temperature is there and produce the approximate results based on our evaluation criteria. Results are intrinsically a high and accurate as demonstrated as it is steady for exceptions and forecasting, so one approach to enhance the straight relapse show is by accumulation of more information using linear regression and Auto-Correlation. Showing that the decision of model was efficient and effective that its expectations can be enhanced by promote accumulation of information under the proposed scheme. For future scope the same can be incorporated over apache spark for concurrent prediction of weather whereas the same can be compare with the results obtained from sensors.

v.REFERENCES

- Kumar, J. Luo and G. Bennett, "Statistical Evaluation of Lower Flammability.
- Ackerman, K. V., and E. T. Sundquist. 2008. Comparison of two U.S. power-plant carbon dioxide emissions data sets. Environmental Science & Technology 42(15):5688-5693.
- Adams, P. N., and D. L. Inman. 2009. Climate Change and Potential Hotspots of Coastal Erosion Along the Southern California Coast—Final Report. CEC-500- 2009-022-F, Sacramento, California Energy Commission.
- Adger, W. N., I. Lorenzoni, and K. O'Brien. 2009a. Adaptation now. In Adapting to Climate Change: Thresholds, Values, Governance. W. N. Adger, I. Lorenzoni, and K. L. O'Brien, eds. Cambridge: Cambridge University Press.
- Adger, W. N., J. Paavola, S. Huq, and M. J. Mace, eds. 2006. Fairness in Adaptation to Climate Change. Cambridge, MA: MIT Press.

- Adger, W. N., S. Agrawala, M. M. Q. Mirza,
- Conde, K. L. O'Brien, J. Pulhin, R. Pulwarty, B. Smit, and K. Takahashi. 2007. Assessment of adaptation practices, options, constraints and capacity. In Climate Change 2007: Impacts, Adaptation and Vulnerability: Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.
- Adger, W. N., S. Dessai, M. Goulden, M. Hulme, I. Lorenzoni, D. R. Nelson, L. O. Naess, J. Wolf, and A. Wreford. 2009b. Are there social limits to adaptation to climate change? Climatic Change 93(3- 4):335-354.
- Adler, M. D., and E. A. Posner. 2006. New Foundations of Cost-Benefit Analysis. Cambridge, MA: Harvard University Press. Agrawala, S. 2004. Adaptation, development assistance and planning: Challenges and opportunities. IDS Bulletin—Institute of Development Studies 35:50-54.
- Agrawal, A. 2008. The role of local institutions in adaptation to climate change. In Social Dimensions of Climate Change Workshop. Washington, DC: Social Development Department, The World Bank.
- Agrawal, A., and N. Perrin. 2008. Climate Adaptation, Local Institutions, and Rural Livelihoods. Ann Arbor, MI: International Forestry Resources and Institutions Program, University of Michigan.
- AGU (American Geophysical Union). 2009. Geoengineering the Climate System. A Position Statement of the American Geophysical Union (adopted by the AGU Council on December 13, 2009). Washington, DC: AGU.
- Airame, S., J. E. Dugan, K. D. Lafferty, H. Leslie, D. A. Mcardle, and R. R. Warner. 2003. Applying ecological criteria to marine reserve design: A case study from the California Channel Islands. Ecological Applications 13 (1):S170-S184.
- Akbari, H., M. Pomerantz, and H. Taha. 2001. Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. Solar Energy 70(3):295-310.
- Akbari, H., S. Menon, and A. Rosenfeld. 2009. Global cooling: Increasing world-wide urban albedos to offset CO2. Climatic Change 94(3-4):275-286.
- Albrecht, A., D. Schindler, K. Grebhan, U. Kohnle, and H. Mayer. 2009. Storminess over the North- Atlantic European region under climate change—a review. Allgemeine Forst And Jagdzeitung 180(5-6):109- 118.
- Aldy, J. E., and R. N. Stavins. 2007. Architectures for Agreement. Cambridge, MA: Cambridge University Press.
- Alheit, J., and E. Hagen. 1997. Long-term climate forcing of European herring and sardine populations.1997. Long-term climate forcing of European herring and sardine populations. Fisheries Oceanography 6:130-139.
- Allan, J. D., M. A. Palmer, and N. L. Poff. 2005. Climate change and freshwater ecosystems. Pp. 272-290 in Climate Change and Biodiversity. T. E. Lovejoy and L. Hannah, eds. New Haven, CT: Yale University Press.
- Alley, R. B., J. Marotzke, W. D. Nordhaus, J.
- T. Overpeck, D. M. Peteet, R. A. Pielke, Jr., R.Alley, W.
- M. 1984. The Palmer Drought Severity Index— Limitations and assumptions. Journal of Climate and Applied Meteorology 23:1100-1109.

- D. Marutho, S. Hendra Handaka, E. Wijaya and Muljono, "The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News," 2018 Distance (LFD) using Four Hazardous Release Models", Process Safety Progress, 12(1), pp.
- Fung. 2005. Drier summers cancel out the CO2 uptake enhancement induced by warmer springs. Proceedings of the National Academy of Sciences of the United States of America 102:10823-10827.
- Grimm, N., Groffman, P., Staudinger, M., and Tallis, H. (2015). Climate change impacts on ecosystems and ecosystem services in the United States: process and prospects for sustained assessment. Climatic Change, 135(1), 97-109.
- Hammer G.L., J.W. Hansen, J.G. Phillips, J.W. Mjelde, H. Hill, H.A. Love, and A.Potgieter. 2001. Advances in Application of Climate Prediction in Agriculture. Agricultural Systems. 70(2- 3)(November):515-553.
- In AMS (American Meteorological Society). 2009. Geoengineering the Climate System. A Policy Statement of the American Meteorological Society (adopted by the AMS Council on July 20, 2009). Washington, DC: AMS. and minimum temperature time series by three artificial neural network methods. Met. Apps, 15:431-445.
- International Seminar on Application for Technology of Information and Communication, 2018, pp. 533-538, doi: 10.1109/ISEMANTIC.2018.8549751.
- J Cifuentes, G Marulanda, A Bello, J Reneses. Air Temperature Forecasting Using
- K.H., Shin, et al. “Dynamical Prediction of Two Meteorological Factors Using the Deep Neural Network and the Long Short Term Memory (1).”
- Keras Metrics—Keras documentation [WWW Document]. [Cross Reference]
- Lazo, J. K., Hosterman, H. R., Sprague- Hilderbrand, J. M., and Adkins, J. E. (2020). Impact- Based Decision Support Services and the Socioeconomic Impacts of Winter Storms, Bulletin of the American Meteorological Society, 101(5), E626-E639.
- M. L. Parry, O. F. Canziani, J. P. Palutikof, C.
- E. Hanson, and P. J. Van Der Linden, eds. Cambridge:M.
- L. Parry, O. F. Canziani, J. P. Palutikof, C. E. Hanson, and P. J. Van Der Linden, eds. Cambridge:Cambridge: Cambridge University Press.
- M. Norman. (2011). Parameterizations: Representing key processes in climate models
- Machine learning - Intuitive understanding of 1D, 2D, and 3D convolutions in convolutional neural networks - Stack Overflow. [Cross Reference] Machine Learning Techniques: A Review. Energies. 2020; 13(16):4215.
- Banu, N. A. S. R. E. E. N., and A. Mrunalini. "Empowering young farmers in the context of climate change." Int J Agric Sci Res (IJASR) 7.2 (2017): 449-456.
- Dympep, A. L. E. T. H. E. A., and R. J. Singh. "A test to measure knowledge of farmers on mitigation and adaptation practices of climate change in hill agricultural system." International Journal of Agricultural Science and Research 7.1 (2017): 21-28.
- Khairandish, Mohammad Omid, R. Gurta, and Meenakshi Sharma. "A hybrid model of faster R-CNN and SVM for tumor detection and classification of MRI brain images." Int. J. Mech. Prod. Eng. Res. Dev 10.3 (2020): 6863-6876.
- Kumar, B. Satish, and Y. Kalyan Chakravarthy. "Prediction Of Optimal Torques From Gait Analysis Applying The Machine Learning Concepts." International Journal Of

Mechanical And Production Engineering Research And Development 9.4 (2019): 685-698.

- Nawalagatti, Amitvikram, and R. Kolhe Prakash. "A comprehensive review on artificial intelligence based machine learning techniques for designing interactive characters." International Journal of Mathematics and Computer Applications Research (IJMCAR) 8.3 (2018): 1-10.
- Danthala, S. W. E. T. H. A., et al. "Robotic manipulator control by using machine learning algorithms: A review." International Journal of Mechanical and Production Engineering Research and Development 8.5 (2018): 305-310.