# Rainfall Prediction For Crop Production Using Machine Learning Algorithms

**Swaraj Bhandare[1], Sadhana Ghuge[2], Vrushali Saka[3], Prashant Paulbudhe[4], Dipali Pawar[5]**
[1, 2, 3, 4] Dept of Information Technology
5Asst.Professor, Dept of Information Technology
[1, 2, 3, 4, 5] Sinhgad Instituteof Technology and Science Pune, Maharashtra, India

***Abstract-*** *Rainfall prediction model are previously based on single regression algorithms have been developed in India until now. In this research work does a study of 6 different regression algorithms to get the more accurate one. The present systems are not that accurate for complex data present. India is an agriculture country and agriculture is considered as main economy driven source. To make this biggest economic source successful it is needed to make a proper estimation of rainfall. The research is focused on getting the effects of different meteorological parameters in rainfall estimation along with an exploration of approaches which were used for prediction of rainfall, machine learning, and their limitation. This research work can be used by all sources and channels to predict and broadcast the predicted rainfall.*

***Keywords***- Machine learning, Rainfall prediction, Linear regression, Random forest.

## I. INTRODUCTION

In today's Era, rainfall is the leading responsible factor for almost happening things across the world. In our country India agriculture is the main economic driven source and 80% of population from village side do agriculture as their main income source. In pandemic situation also only sector with positive GDP rate.

Apart from this in many parts over the world the rainfall prediction is very much necessary because many areas can plan the rain harvesting based on the shortage of water they have to face in future and this can be happened by predicting the rainfall. This project based on predicting the rainfall using 6 different regression algorithms. The project performs the comparative study of linear regression, ridge regression, lasso regression, support vector machine, random forest and accordingly portrays the efficient approach for rainfall prediction.

First of all data collection is performed the data from dataset here is collected from kaggle which contains Subdivision wise monthly and yearly rainfall in India of 1901-2015. Then secondly data preprocessing is performed by using the pandas and numpy libraries from python. Data Visualization is then performed which is the main factor here is does the graphical representation which represent the data in form of various graphs like bar graph and line graph.

So the main aim behind this research is to resolve the issues faced by existing systems which can give us the proper and accurate estimation to assist the nation to grow when it comes to agriculture and rainfall related factors.

## II. LITERATURE SURVEY

In the first paper (1) author Aymen E Khedr shows the food shortage in rural areas of Egypt country. It proposes a system which will give the estimation of amount of crops needed to satisfy the Egyptian citizens. This system is build for prediction using Artificial Neural Network(ANN) along with the help of MLP function in WEKA. It helps to predict and estimate the amount of the crops needed up to the next year.

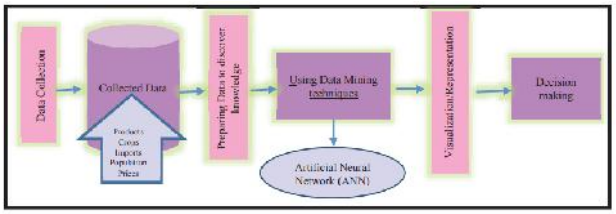


**Fig1:Proposed Framework**

In this paper(2) author Afan Galih Salman researches deep learning techniques for weather forecasting. It do a differentiate study of the prediction performance of Recurrence Neural Network, Conditional restricted Boltzmann machine and Convolution Network models. The outcome will be very useful for weather forecasting for various applications like flight navigation. In this section, briefly describe the four existing studies on predicting the large-scale atmospheric state

in the medium-range with a focus on the data, methods and evaluation.

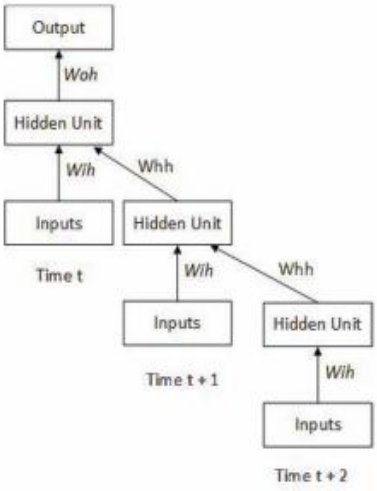


**Fig 2: Unfolded Recurrent Neural Network Model**

In this research(3) author Niketa Gandhi makes a comparative study of classification algorithms BayesNet and NaiveBayes to predict the rice crop yield for a specific kharif season for Maharashtra a leading agriculture state in India. This is mainly done by selecting all the districts in Maharashtra state by considering the different parameters like min, max, average temperature . The outcome is drawn at the end is the performance of BayesNet was far better compared to NaiveBayes.. After deep analysis of the techniques some significant improvement are proposed which can be used for increasing the accuracy.Critical analysis of the existing system is the main integral part of any research work because it mainly leads us to the errors and challenges of the research work which was already faced by the authors. By knowing these challenges we can simply know where to hit the target and it will be easy to formulate a new model based on the limitations of the existing one.

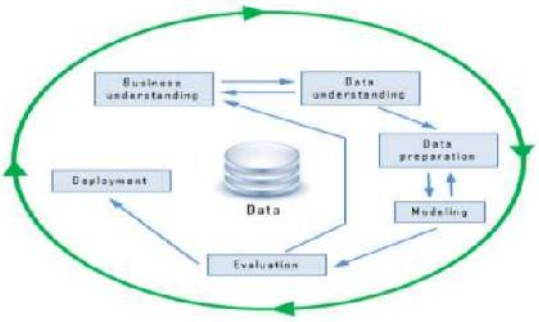


**Fig 3: Outline of the Model used**

In this paper(4) author Rupanjali D Baruah states multiple linear regression techniques for forecasting the future yield prediction in Tea cultivation with climatic change in the time period of 1999 to 2006. The parameters here selected for the research here are temperature, rainfall, relative humidity, evaporation for different four regions Which are South Bank, North Bank, Upper Assam and Cachar of Assam. The outcome comes at the last is that the Tea production estimation which is forecasted for four different regions were validated for the future yield production (for Next upcoming years) and the developed system can be used to predict the Tea production for each region with precision.

After successful MLR analysis, four tea production estimation equations were derived for the four regions and the equations thus generated represented the relation between the meteorological parameters and yield specific to the particular region. The equations derived are given below as equation (1), (2), (3) and (4) for South bank, Upper Assam, North Bank and Cachar, respectively.

PSOUTHBANK = 383773–16656.9*S – 5.5612*R–17276.8*Tdiff (1)

PUPPERASSAM=235133.3–5790.38*S–7.22811*R-4772.87*Tdiff (2)

PNORTHBANK=245962.8–23078.4*S–19.5684*R +1698.916*Tdiff (3)

PCACHAR=7126.21–6493.54*S–2.90149*R+ 1475.789*Tdiff (4)

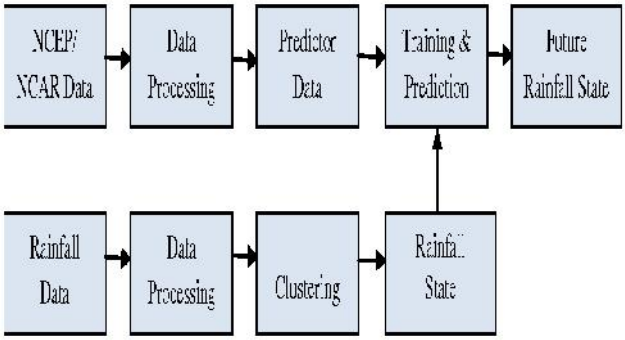


**Fig 4: Proposed System Architecture**

## III. CONCLUSION

Regression algorithm was used as modeling techniques for prediction of the rainfall. Five different regression algorithms were used here which are linear regression, Lasso, Ridge, SVM, Random forest algorithm. The

experimental outcome shows that the accuracy and performance of the Random forest algorithm was far better as compared to other regression algorithms. Accurate forecast using the correct dataset with minimum null values and correct parameters will result in accurate rainfall prediction forecast in the future which will lead to resolve the issues faced by existing systems which can give us the proper and accurate estimation to assist the nation to grow when it comes to agriculture and rainfall related factors.

## REFERENCES

[1] Aymen E Khedr, Mona Kadry, Ghada Walid(2015), "Proposed framework for Implementing Data Mining Techniques to enhance Decisions in Agriculture sector Applied case on Food Security Information Center Ministry of Agriculture,Egypt",international conference on Communication, management and Information technology(ICCM).

[2] Afan Galih Salman, Bayu Kanigoro, Yaya Heryadi, "Wheather forecasting using deep learning Techniques", 2015 International Conference on Advance Computer science and Information Systems(ICACSIS)

[3] Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong, "Predicting Rice crop yield using Bayesian Networks", 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India.

[4] Rupanjali D. Baruah, R.M. Bhagat, Sudipta Roy, L.N. Sethi "Use of Data Mining Technique for Prediction of tea yield in the face of Climate Change of Assam, India ", 2016 International Conference on Information Technology.

[5] Niketa Gandhi, Leisa J. Armstrong "Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India using data mining Techniques", 2016 IEEE International Conference on Advances in Computer Applications (ICACA)

[6] Kuljit kaur, Kalnwalpreet singh Attwal "Effect of Temperature and Rainfall on Paddy Yield using data mining ", 2017 7th international conference on Cloud Computing, Data Science and EngineeringConfluene.