# Effective Farming Using Machine Learning Algorithm in Climate Change

S.Priyanga[1], S. Soniya[1], R.Kowsalya[1] K.Prasannakumar[2]
[1]Student, Computer Science and Engineering, Adhiparasakthi Engineering College Melmaruvathur, Kancheepuram, Chennai,India
[2]Assistant Professor, Department of Computer Science and Engineering, Adhiparasakthi Engineering College Melmaruvathur, Kancheepuram, Chennai, India

**Abstract - In accuracy agriculture, real time and exactly produced information is organize and unstructured datasets. As accuracy agriculture produces more information in the unstructured shape and momentum look into slant is to discover learned data from them. The necessity of past cultivating is to enhance the excellence of farming items & administrations by decreasing speculation rate. In proposed methodology, to store the unstructured data in server, then it to be processed using streaming algorithms and finally get the result in visualization using machine learning. It seems to show the prediction of cultivation, save the farmers time, gave more profit in short duration instead of losing their growth on natural disasters. The necessity of cultivating is the effective utilization on cultivation using machine learning algorithm. Machine learning algorithm is a predictive modelling technique which gives relationship between dependent and independent variables which was so helpful in predicting unstructured data on agribusiness). In this methodology, data sets have to be collected from trained data sets. The visualization gives overview of farmers able to cultivate or not in this season.**

**Keywords-visualization, machine learning, cultivation, farming, natural disasters, prediction.**

## 1. INTRODUCTION

The term Big Data refers to all the data which were processed across the globe at an unprecedented rate. The data were implemented in either structured or unstructured format. There is a need to convert Big Data into Business Intelligence that enterprises can readily deploy. Better data leads to better decision making and an improved way to strategize for organizations regardless of their size, geography, market share, customer segmentation such other categorizations. Machine Learning is considered to be an idea to learn from experience without being explicitly programmed. Instead of implementing into code, we can able to feed data to the generic algorithm, and it builds logic based on the data given. The data being very massive, the time taken to compute is increased, and this is where Machine Learning comes into action, to help people with large data in minimum time. Data drives the modern organizations of the world and hence making sense of this data and unravelling the various patterns and revealing unseen connections within the vast collection of data becomes critical and a hugely rewarding endeavour indeed. If big data and cloud computing are gaining importance for their contributions, machine learning as technology helps analyse huge collection of data, easing the task of data scientists in an automated process and gaining equal importance and recognition. A multi-dimensional big data system has been established and is still promptly developing, which has enabled us to observe and monitor changes on a global scale for the climate parameters.

## 2. MOTIVATION OF REASEARCH

The idea discussed on big data is to produce the enormous profit on agribusinesses and also climate malfunctions on crops is determined. Here the machine learning concepts also indulged to provide exactness farming and enormous information gatherings. The yield administration and disaster administration is also clarified through frequently used algorithms like regression algorithm. Thus the researches exaggerate profitability of on farmer welfare.

### 2.1 PROPOSED SYSTEM

In the proposed methodology is used to store the unstructured data (climate gauging dataset) in server, then it to be processed using streaming algorithms and finally get the visualization using regression algorithm in machine learning. The visualization gives the result and prediction of farmers able to cultivate or not in this season. It seems to show the prediction of cultivation, save the farmers time, gave more profit in short duration instead of losing their growth on natural disasters. It saves the time complexity to give an accurate prediction of crop data's and soil data's and climate data sets.

### 2.2 CONCERN IN ACCURACY AGRICULTURE

Accuracy agriculture through consist of different coats for example, application coats, store and preparing coats and infrastructure coats. In the application coats, information obtaining instruments, electronic arrangements, programming and advancement stages are available. The capacity and association of enormous information require an alternate framework and stage. The present distributed computing arrangements give such colossal measure of capacity and administration. The conveyed and parallel frameworks make a part in the signs.

### 3.1 PROCEDURE AND IMPLEMENTATION

The software construction, defect diagnosis and prevention strategies are synchronized to reduce development risk, time and cost. The main objective of testing at this phase is to eliminate errors, increase quality as well as efficiency of the software,

before it proceeds to the Quality Assurance phase. Development Testing can be classified into different types, depending on the Company's expectations: - data flow analysis metrics analysis, code coverage analysis, Static Code Analysis, Traceability, Peer Code Review, etc. After Development but before Testing, developers use exhaustive Code Review Checklist. This checklist is used to check the Functionality of the codes along with exception handling and correct propagation of errors; their Readability and Maintainability including naming conventions, standard formatting, intuitive naming of variables, flexibility of the code and verifications of whether the code is configurable or not; and Performance or Efficiency that will address all language specific pitfalls, and will include areas like Concurrency management, Memory leaks, Buffer overflows, Array indexing, pointer arithmetic, proper termination of threads, etc.

### 3.2 SYSTEM ARCHITECTURE AND EXPLANATION

Big data have also been playing a vital role in predicting when it is incorporated with climate science, for instance weather forecasting, natural disasters monitoring and early warning, energy consumption forecasting, traffic forecasting, etc. By applying corresponding data mining techniques (the detailed introduction of data mining techniques can be found in it allows knowledge discovery of the potential relationships and causal inferences, which further contribute to the modelling and predicting. However, optimization requires comprehensive theoretical understanding as well as adequate big data analytics skills to structure the optimal model infrastructure so as to maximise performance, efficiency and utility, or, in some cases, for achieving sustainable development. The soil datasets, climate datasets and crop data is collected and processed through the data processing which constitutes filtering redundant, inconsistent and missing values. This processed information is passed to Pre-processed data contains weather crop and soil information. In this information some mentionable data were collected like test data, validation data and training data. Then the trained data and predicted data were processed through Machine Learning Algorithm. If the result is accurate, then the farmer can cultivate and if it's not they have to change the cultivation period accordingly. This figure briefly explains on the agriculture production and enhancement through the machine learning concepts. Accordingly to climate is also predicted easily by datasets like test, trained and predict data on agriculture.

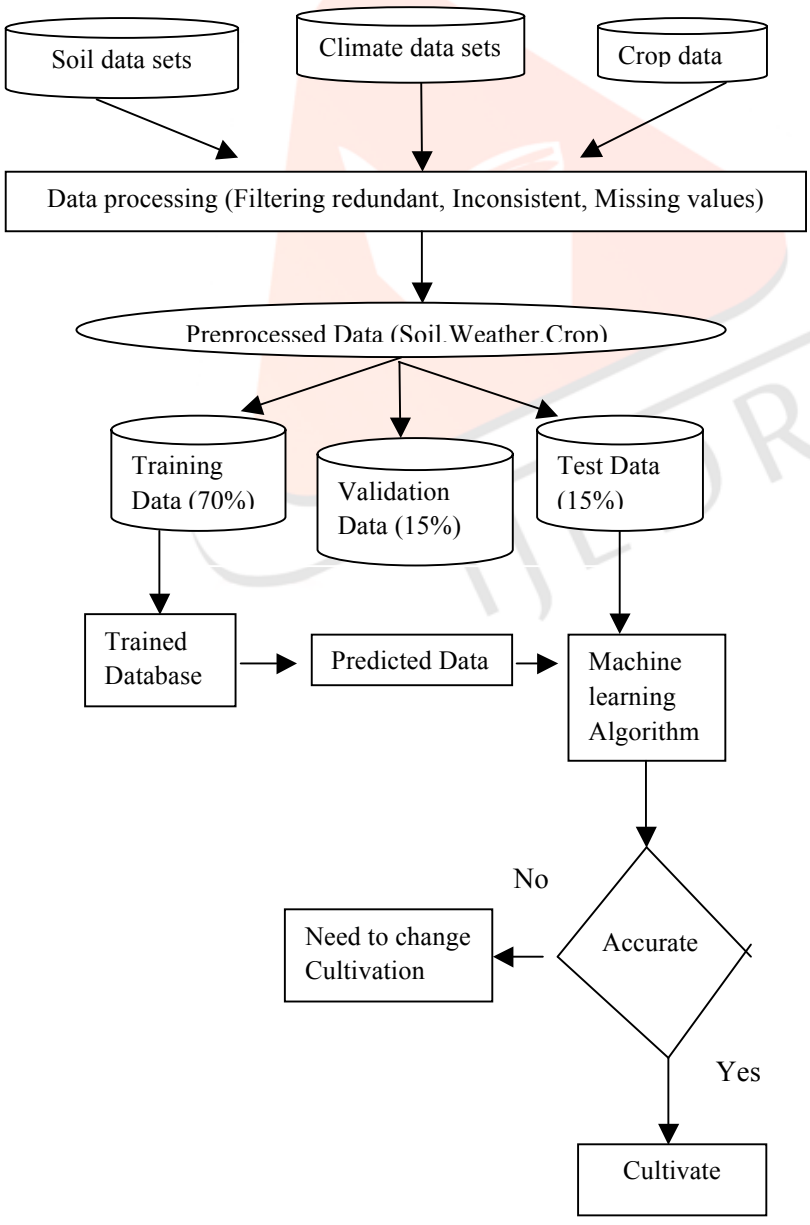

**Figure 3.1: System Architecture**

The above figure illuminates the structure on processing how the climate is identified with maximum profit.

## 4. METHODLOGY

Identity management (ID management) is a broad administrative area that deals with identifying individuals in a system (such as a country, a network, or an enterprise) and controlling their access to resources within that system by associating user rights and restrictions with the established identity. Identity management (IdM) is the task of controlling information about users on computers. Such information includes information that authenticates the identity of a user, and information that describes information and actions they are authorized to access and/or perform. It also includes the management of descriptive information about the user and how and by whom that information can be accessed and modified.

Managed entities typically include users, hardware and network resources and even applications.In traditional collaborative filtering systems, the amount of work increases with the number of participants in the system. New recommender system technologies are needed that can quickly produce high quality recommendations, even for very large-scale problems. To address these issues, we have explored item-based collaborative filtering techniques. Item-based techniques first analyze the user-item matrix to identify relationships between different items, and then use these relationships to indirectly compute recommendations for users. Finally, we experimentally evaluate our results and compare them to the basic k-nearest neighbor approach. Our experiments suggest that item-based algorithms provide dramatically better performance than user-based algorithms, while at the same time providing better quality than the best available user-based algorithms. In pattern recognition, the ***k*-nearest neighbour's algorithm (*k*-NN)** is a non-parametric method used for classification and regression. In both cases, the input consists of the *k* closest training examples in the feature space. The output depends on whether *k*-NN is used for classification or regression

In k-NN classification, the output is a class membership. The *k*-NN algorithm is among the simplest of all machine learning algorithms. Both for classification and regression, it can be useful to assign weight to the contributions of the neighbours, so that the nearer neighbours contribute more to the average than the more distant ones. This can be thought of as the training set for the algorithm, though no explicit training step is required. A shortcoming of the k-NN algorithm is that it is sensitive to the local structure of the data. The algorithm is not to be confused with *k*-means another popular machine learning technique. Item-based collaborative filtering is a model-based algorithm for making recommendations. In the algorithm, the similarities between different items in the dataset are calculated by using one of a number of similarity measures, and then these similarity values are used to predict ratings for user-item pairs not present in the dataset. By using this algorithm we have to predict the ratings of the item with the help another user's recommendation and suggestion about the item and place which is suitable or not for the particular person.

The above algorithm is used to calculate the highest rating about the item set and able to choose which location to take a decision by the user. The user details and item details are maintained by admin also he has the responsibilities in collaborative filtering storage. Random forest algorithm is a supervised classification algorithm. To model a greater number of decision trees to create the forest you are not going to use the same apache of constructing the decision with information gain index approach. The working natures of the random forest algorithm were described to be as below.

### 4.1 Random Forest pseudo code:

1. Randomly select "k" features from total "m" features, where k << m
2. Among the "k" features, calculate the node "d" using the best split point.
3. Split the node into daughter nodes using the best split.
4. Repeat 1 to 3 steps until "l" number of nodes has been reached.
5. Build forest by repeating steps 1 to 4 for "n" number times to create "n" number of trees.

## 5. CONCLUSION AND DISCUSSION

Climate change has contributed to changing patterns of extreme weather across the globe, from longer and hotter heat waves to heavier rains. From a broad perspective, all weather events are now connected to climate change. While natural variability continues to play a key role in extreme weather, climate change has shifted the odds and changed the natural limits, making certain types of extreme weather more frequent and more intense. The understanding of how climate change affects extreme weather are still developing, evidence suggests that extreme weather may be affected even more than anticipated. Extreme weather is on the rise, and the indications are that it will continue to increase, in both predictable and unpredictable ways.

## REFFERENCES

[1] Ulrich, Benjamin Justus, Dennis Loehr, "A Privacy Preserving System for Cloud Computing", International Conference on Computer and Information Technology, 2011.

[2] Yong Wang, Ping Zhang , "Enhance Big Data Security in Cloud Using Access Control" , International Conference on Advances in Big Data Analytics, 2017.

[3] Kire Jakimoski, "Security Techniques for Data Protection in Cloud Computing", International Journal of Grid and Distributed Computing Vol. 9, No. 1 (2016), pp.49-56.

[4] Iynkaran Natgunanathan, Yong Xiang, Guang HUA, Song Guo, (2016), IEEE Access January 2016.

[5] Banks, David, John S. Erickson, and Michael Rhodes, "Toward cloud-based collaboration services", In Usenix Workshop HotCloud. 2009.

[6] Elham Abd Al Latif Al Badawi1 & Ahmed Kayed, "Survey of enhancing the data security of cloud computing environment by using the data Segregation Technique", IJRRAS 23 (2) - May 2015.

[7]R. Banyal, P. Jain, and V. Jain, "Multi-factor authentication framework for loud computing", Fifth International Conference on Computational Intelligence, Modeling and Simmulation, 2013, Pp 105-110.
[8] H. Kim, and S. Timm, "X.509 Authentication and Authorization in femi cloud", IEEE/ACM 7th International Conference on Utility and Cloud Computing, 2014, Pp 732-737.
[9] Sweeney L, "k-anonymity: a model for protecting privacy", International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 10, No.5, pp.557–570.