CLIMATE CHANGE DETECTION IN BIG DATA ANALYTICS

Suthi R  
PG Scholar,  
Department of Computer Science and Engineering,  
Sri Ramakrishna Engineering College, Coimbatore

Kingsy Grace R  
Associate Professor,  
Department of Computer Science and Engineering,  
Sri Ramakrishna Engineering College, Coimbatore

Abstract- Big data plays a vital role for Prediction of climate change detection big data plays a vital role. The climate challenges the impact of day to day life of temperature, moisture, precipitation. Despite the urgency, data science has little impact on further understanding of climate data. The data science describes opportunities to mine large climate datasets, for mining climate data and traditional big data approaches. The proposed project mainly focuses on data, methods, and applications for big data to fulfill climate sciences and applications. The Artificial Neural Network (ANN) and Classification and Regression Tree (CART) algorithms are used to detect the climate changes using Big Data Analytics. The proposed ANN algorithm provides better results in terms of climate.

Keywords: Traditional big data approaches, Artificial Neural Network, Classification and Regression Tree

1. INTRODUCTION
The predictive analytics in big data used to detect climate change and is used to extract the values of climate change detection. For predicting climate change analytics ANN and CART algorithms are used for predicting the climate change analysis. The big climate data used for predicting the various seasonal changes occurred in each state denotes rainfall, precipitation Temperature, latitude, longitude. For, prediction of each climatic condition, climate datasets are collected from Madurai stations. Climate change is a global problem that is having widespread effects on weather on every continent and in every nation. It is altering environments and impacting on biodiversity, human health and sustainable economic and social development. The climate refers to a general description of long term features of the weather in Madurai station location, such as the average temperature or rainfall for each month of the year, the weather as being a daily expression of the fluctuations of climate around the long-term average pattern. World’s climate increased, scientists began using the term climate change to distinguish these trends from normal climate variation. The structure of manuscript is explained below: Section 1 introduces big data analytics in climate change detection using ANN and CART algorithms. Section 2 reviews the recent work in big climate data analytics. Section 3 reviews the proposed framework for seasonal climate changes. Section 4 discusses the proposed method for seasonal climate change detection. Algorithms are combined for predicting various seasonal changes in climatic predictions. Section 5 reviews Results and Discussions for predicting climatic conditions. Section 6 predicts the Performance Evaluation for various climatic conditions, and finally, section 7 concludes the work.

2. RELATED WORK
Research work in Big Data Analytics in climate change detection is to develop a data science for understanding the major field of global climatic change and has an impact of human life for extracting accurate insight in large climate data [3]. Climate data is observed from various advanced sensor technologies and is used to represent the seasonal changes [4]. NASA developed the Climate-Analytics-as-a-Service (CAaaS) in cloud computing to enable the best performance in following domains such as data proximal analytics, scalable data administration (big data), software machine virtualization, high-performance computing, adaptive analytics and synchronized API development [5]. The impact of big data and cloud computing has been increased noticeably [9]. Researchers from Brazil has developed Platform as a Service (PaaS) based cloud computing platform called EUBrazilCC project to store and analyze big climate data [10]. Researchers from bioinformatics and climate modeling have identified the correlation between the disease and climate with the help of big data analytics. Merging of climate and health data
becomes a significant role in big data analytics. For climate change detection various methods are used.

The climate change detection techniques used to predict climatic conditions using day wise climatic data and season wise climatic data. For day wise climatic condition it used to predict climate using previous day climate. For season wise climatic condition it used to predict season wise based on seasonal inter-annual or long-term time scales. The climatic conditions are highly sensitive to initial conditions and Climate projections are distinct from climate change predictions. Climate projections depends upon emission, concentration, radiation forcing scenarios, which are based on assumptions concerning, for future socio-economic and technological developments.

B. Artificial Neural Network in season wise prediction:

The Artificial Neural Network in season wise prediction denotes the seasonal inter-annual or long-term time scales. The ANN process are specified based on the mean square error, relative or absolute error, determination coefficient as well as the central statistical moments such as mean, standard deviation, minimum and maximum values. The number of input and output variables directly gives the number of neurons in the input and output layers. The number of hidden layers and hidden nodes seemed insufficient, in comparison with the number of input and output nodes, to reserve the higher order relationship. The climate parameters like maximum temperature, minimum temperature and relative humidity use the features extracted over different periods as well as from the climate parameter time series. The ANN model for step-by-step prediction to climate over India during summer-monsoon datasets was available at the website (http://climate.noaa.gov).

C. CART [Classification and Regression Tree]:

CART is an increasingly popular form of statistical analysis available through the statistical package. CART operates by splitting the data till ending points, or terminal nodes, are achieved using preset criteria by analyzing all explanatory variables. The binary division of a single explanatory variables, reduces the deviance in the response variable

A. Artificial Neural Network in day wise prediction:

Artificial Neural Networks (ANNs) are developed in recent years with development of computer technology and artificial intelligence. It is mostly used by researchers in light of increasing concerns over ecological issues such as global warming and is highly accurate when used for short-term predictions, despite global climate change research, effects the interactions on study of long-term experimental research. As shown in Figure 3.1, x is the neuron input, w is the adjustable input weight, and θ is offset signal used for modeling the neurons excitation threshold. u(·) is input value in ANN and f(·) is output of ANN are the basic function and activation function.

Figure 2.1 Climate change detection using ANN

In Fig 2.1. The various datasets of climatic conditions are collected from the Madurai station, of attributes such as temperature, latitude, longitude, elevation, precipitation of daily data.

Figure 3.1 General Neural model

Figure 3.2 Structure of CART

Classification and Regression Tree (CART) Analysis, is a simple and powerful analytic tool that helps to determine the variables in a climate dataset, and it helps researchers to develop a model. This technical method describes a decision tree of (CART) by practical applications and implementation strategies in R. Based on analysis of (CART) it become popular and valuable in climatic fields. CART represents a snapshot of data variables and can be used as a first step in constructing an informative model or a final visualization of important associations. In a large climatic change detection project, statisticians can use CART to present preliminary data to predict climate
changes or other seasonal changes for statistical results with practice knowledge and intuition. The statistical processes based on classification and regression tree analysis is very similar, but it is important to clearly distinguish the two. For a response variable which has classes, often 0-1 binary. The dataset must be organized into groups by the response variable classification.

4. CLIMATE PREDICTION IN R

A. Artificial Neural Network in climate change prediction:

New dataset was scaled and saved into R Tool named as climate dataset. A neural network is normalization of data. For, predicting the values of the neural network can be accurately compared to actual data. To train a neural network model using neural net package, the neural net library is loaded into R, to predict the changes in climatic conditions using ANN algorithm. By using neural network model it denotes train set and test set of ANN model. Training is the data set used to adjust the weights of the nodes to build the model initially. Usually the proportion is around 70%, but varies with the nature of the dataset and the model. To train the dataset we use, This data is crucial as it should reflect all of the heterogeneity of the problem

\[
\text{train} \leftarrow \text{data [ index, ]}
\]

The test data is used to check the final performance of a developed model. The data set ideally should be independent of the training and validation set. The model is now allowed to predict with inputs from a new set of data. Therefore, the results generated will reflect the model's success/failure. This would be real world implementation.

\[
\text{test} \leftarrow \text{data [ -index, ]}
\]

Figure 2.1 ANN model in R tool

By using neural network package in R tool it provides a graphical representation model based on weight of each connection nodes. The black lines show the connections between each layer and the weights on each connection while the blue lines show the bias term added in each step. The bias can be thought as the intercept of a linear model.

\[
\text{tree}(\text{factor}(Y) \sim \sum(x), \text{data=Weather data})
\]

while in \text{rpart} classification is specified in the method option:

\[
\text{rpart}((Y)\sim\sum, \text{method="class", data=Weather data})
\]

Figure 4.1 CART model in R tool

CART reveals the most important factors associated with maximum temperature rates is the percent of privately collected data sets. For each node , the right branch of the node is conditional on the node being true, and the left is conditional on the node being false. The numbers at the bottom of the terminal branches indicate the mean of the maximum temperature rate in each data subset. The summary of the model shows a deviance of 0.0244, or a root mean square of 0.12 (12%), which indicates that our model is actually pretty good given its simplicity, the nodes identifies a good deal of the variation in temperature rates.

5. RESULTS AND DISCUSSIONS

This paper uses ANN method to find the differences of each sample value from the target value. ANN control charts are also called as time-weighted control chart used to monitor the small shifts in the mean of a process. The traditional ANN control chart is used to monitor the changes in rainfall, precipitation, maximum temperature, minimum temperature, humidity, wind speed and sunshine.
Experimental results prove that Artificial Neural Network based climate change detection algorithm performed well when compared with Classification and Regression Tree (CART) method. The CART algorithm identifies the changes in the seasonal climate with the 9am to 3pm based on speed of wind and direction of wind.

6. PERFORMANCE EVALUATION

ANN control charts are called as time-weighted control chart used to monitor the small shifts in the mean of a process. The neural network of differences is between the values and the average (target value), i.e., the mean is subtracted from each values. The ANN control chart is used to identify the neural net of the deviations of each sample value from the target value for, finding the maximum and minimum temperature of CART and ANN.
To find the maximum and minimum temperature based on the datasets collected, used to predict the day wise climatic conditions.

Based on the precipitation occurs, it used to provide the hourly calculation of predicting changes occurred in climatic conditions using ANN and CART algorithms.

7. CONCLUSION AND FUTURE WORK

ANN is an expensive step for identifying climate changes in the field of both effective and accurate calculation. The proposed system uses the algorithm of ANN and CART algorithms for finding the changes occurred due to climatic changes. A Predictive Analysis used to determine the Scalable data storage platform and efficient change detection algorithms are required to monitor the climate changes. The proposed system performance is verified using climate data set. To demonstrate the effectiveness, scales of data set is reduced by using data cleaning technique in order to decrease the time and cost, to improve the accuracy of current changes in climatic conditions. The future work is to improve the result for predicting climatic conditions. To predict climatic conditions such as Temperature, Humidity, Wind Speed and Rain fall are to be used for analysing the seasonal diseases due to climatic change.

REFERENCES


[10] Sandro fiore, Marco mancini, Donatello elia, Paola nassisi, Francisco vilar brasileiro, Ignacio blanquer, “Big Data analytics for climate change and biodiversity in the eubrazilcc federated cloud infrastructure” ,ACM pp. 978-1-4503-3358-0/15/05,2015