International Journal on Recent and Innovation Trends in Computing and Communication Data Mining Techniques for Weather Prediction: A Review

Article *in* International Journal on Recent and Innovation Trends in Computing and Communication · August 2014

CITATIONS
9

READS
665

2 authors, including:

Divya Chauhan
Government College Rampur Bushahr HImachal Pradesh Shimla
8 PUBLICATIONS **15** CITATIONS

SEE PROFILE

International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 2 Issue: 8 2184 – 2189

Data Mining Techniques for Weather Prediction: A Review

Divya Chauhan
Department of Computer Science
Himachal Pradesh University
Shimla 5, India
dvcherish90@gmail.com

Jawahar Thakur
Department of Computer Science
Himachal Pradesh University
Shimla 5, India
jawahar.hpu@gmail.com

***Abstract*—** Data mining is the computer assisted process of digging through and analysing enormous sets of data and then extracting the meaningful data. Data mining tools predicts behaviours and future trends, allowing businesses to make proactive decisions. It can answer questions that traditionally were very time consuming to resolve. Therefore they can be used to predict meteorological data that is weather prediction. Weather prediction is a vital application in meteorology and has been one of the most scientifically and technologically challenging problems across the world in the last century. Predicting the weather is essential to help preparing for the best and the worst of the climate. Accurate Weather Prediction has been one of the most challenging problems around the world. Many weather predictions like rainfall prediction, thunderstorm prediction, predicting cloud conditions are major challenges for atmospheric research. This paper presents the review of Data Mining Techniques for Weather Prediction and studies the benefit of using it. The paper provides a survey of available literatures of some algorithms employed by different researchers to utilize various data mining techniques, for Weather Prediction. The work that has been done by various researchers in this field has been reviewed and compared in a tabular form. For weather prediction, decision tree and k-mean clustering proves to be good with higher prediction accuracy than other techniques of data mining.

Keywords- *Data Mining, Decision Trees, Artificial Neural Network, Regression, Clustering.*

*****

I. INTRODUCTION

Data mining [13] is a process which finds useful patterns from large amount of data. Data mining can also be defined as the process of extracting implicit, previously unknown and useful information and knowledge from large quantities of noisy, ambiguous, random, incomplete data for practical application. It is a powerful new technology with great potential to help companies focus on the most important information in their databases. It uses machine learning, statistical and visualization technique to discover and predict knowledge in a form which is understandable to the user. Prediction is the most important technique of data mining which employs a set of pre-classified examples to develop a model that can classify the data and discover relationship between independent and dependent data.
Weather prediction is the application of science and technology to predict the state of the atmosphere for a given location. It is becoming increasingly vital for scientists, agriculturists, farmers, global food security, disaster management and related organizations to understand the natural phenomena to plan and be prepared for the future [17,37,19,35].The art of weather prediction began with early civilizations using reoccurring astronomical and meteorological events to help them monitor seasonal changes in the weather. Throughout the centuries, attempts have been made to produce forecasts based on weather changes and personal observations. Many meteorological instruments were being refined during the previous centuries. Other related developments that are, theoretical, and technological developments, also contributed to our knowledge of the atmospheric weather conditions. Weather prediction is an important goal of atmospheric research. Hence changes weather condition is risky for human society [3,5,15].It affects the human society in all the possible ways.
Weather prediction is usually done using the data gathered by remote sensing satellites. Various weather parameters like temperature, rainfall, and cloud conditions are projected using image taken by meteorological satellites to access future trends. The satellite based systems are expensive and requires complete support systems. The variables defining weather conditions varies continuously with time, prediction model can be developed either statistically or by using some other means like decision tree, artificial neural networks, regression, clustering techniques of data mining. Weather prediction is a form of data mining which is concerned with finding hidden patterns inside largely available meteorological data [31].
Rest of the paper is organized as follows. Section II narrates the background study of data mining and weather prediction. Section III discusses the literature review of various data mining techniques used for predicting weather. Section IV gives the comparison of work done by researchers. Finally, the paper is concluded in section V.

II. BACKGROUND STUDY

A. *Data Mining*

Data mining is the science and technology of exploring data in order to discover unexplored patterns. Traditionally, data acquisition was considered as one of the most important stages of data analysis [36]. The data had to be collected manually so the quantity was also small. So the decisions were based on limited information. But now, gathering data has become easier and storing it has become inexpensive. Unfortunately, as the amount of information increases, it becomes harder to

understand it. Data mining is a matter of considerable importance and necessity for the accessibility and abundance of this information in the database. Data Mining can be defined as the process of extracting useful information and knowledge from large amount of unstructured and structured data, which is also an effective means of discovering knowledge [14]. It has got many applications [20]. Data mining appeared as a means of coping with the exponential growth of data and information. Data mining sift through large databases in search of interesting pattern and relationships among instances. In practice, data mining provides many tools by which large amount of data can be analyzed automatically. There are steps to the process of data mining which are run iteratively: preprocessing, analysis, data exchanging.

There are various data mining techniques [7,24,32,33] such as: Classification, Prediction, Clustering, Association, Outlier Detection and Regression.

The prediction discovers relationship between independent variables and relationship between dependent and independent variables. There are various algorithms of classification and prediction [8,18,26]. Some of them are Decision Tree, Artificial Neural Networks, Support Vector Machines (SVM), Bayesian Classification and Regression. There are several criteria for evaluating the prediction performance of algorithm [3].

B. Weather Prediction

The various methods used in prediction of weather are [30]:

1) *Synoptic weather prediction*: It is the traditional approach in weather prediction. Synoptic refers to the observation of different weather elements within the specific time of observation. In order to keep track of the changing weather, a meteorological center prepares a series of synoptic charts every day, which forms the very basic of weather forecasts. It involves huge collection and analysis of observational data obtained from thousands of weather stations.
2) *Numerical weather prediction*: It uses the power of computer to predict the weather. Complex computer programs are run on supercomputers and provide predictions on many atmospheric parameters. One flaw is that the equations used are not precise. If the initial stage of the weather is not completely known, the prediction will not be entirely accurate.
3) *Statistical weather prediction*: They are used along with the numerical methods. It uses the past records of weather data on the assumption that future will be a repetition of past weather. The main purpose is to find out those aspects of weather that are good indicators of the future events. Only the overall weather can be predicted in this way.

III. LITERATURE REVIEW

There are many studies that support the applicability of data mining techniques for weather prediction.

E. G. Petre [10] presented a small application of CART decision tree algorithm for weather prediction. The data collected is registered over Hong Kong. The data is recorded between 2002 and 2005. The data used for creating the dataset includes parameters year, month, average pressure, relative humidity, clouds quantity, precipitation and average temperature. WEKA, open source data mining software, is used for the implementation of CART decision tree algorithm. The decision tree, results and statistical information about the data are used to generate the decision model for prediction of weather. The way the data is stored about past events is highlighted. The data transformation is required according to the decision tree algorithm in order to be used by WEKA efficiently for weather prediction.

M. A. Kalyankar and S. J. Alaspurkar [23] used data mining techniques to acquire weather data and find the hidden patterns inside the large dataset so as to transfer the retrieved information into usable knowledge for classification and prediction of weather condition. Data mining process is applied to extract knowledge from Gaza city weather dataset. This knowledge can be used to obtain useful predictions and support the decision making process. Dynamic data mining methods are required to build, that can learn dynamically to match the nature of rapidly changeable weather nature and sudden events.

F. Oliya and A. B. Adeyemo [17] investigated the use of data mining techniques in predicting maximum temperature, rainfall, evaporation and wind speed. C4.5 decision tree algorithm and artificial neural networks are used for prediction. The meteorological data is collected between 2000 and 2009 from the city Ibadan, Nigeria. A data model for the meteorological data is developed and is used to train the classifier algorithms. The performance of each algorithm is compared with the standard performance metrics and the algorithm with the best result is used to generate classification rules for the mean weather variables. A predictive neural network model is also developed for weather prediction and the results are compared with the actual weather data for the predicted period. The results shows that given enough training data, data mining technique can be efficiently used for weather prediction and climate change studies.

Abhishek Saxena *et al.* [2] presented the review of weather prediction using artificial neural networks and studied the benefit of using it. It yields good results and can be considered as an alternative to traditional meteorological approach. The study expressed the capability of artificial neural network in predicting various weather phenomena such as temperature, thunderstorms, rainfall, wind speed and concluded that major architecture like BP, MLP are suitable to predict weather phenomenon. But due to the nonlinear nature of the weather dataset, prediction accuracy obtained by these techniques is still below the satisfactory level.

M. Kannan *et al.* [32] described empirical method technique using data mining to make a short term prediction of rainfalls over specific regions. The three months rainfall data of a particular region for five years is analyzed. Accurate and timely weather prediction is a major challenge for research community. Classification technique is used to classify the reason for rainfall in the ground level. Clustering technique is used to group the element that is particular area occupied by rainfall regions and the rainfall is predicted in a particular region. Multiple linear regression model is adopted for prediction but the results give the rainfall data having some approximate value not a predictor value.

Gaurav J. Sawale and Sunil R. Gupta [12] proposed an artificial neural network method for the prediction of weather for future in a given location. Back Propagation Neural Network is used for initial modeling. Then Hopfield Networks are fed with the result outputted by BPN model. The attributes include

temperature, humidity and wind speed. Three years data of weather is collected comprising of 15000 instances. The prediction error is very less and learning process is quick. This can be considered as an alternative to the traditional meteorological approaches. Both algorithms are combined effectively. It is able to determine non- linear relationship that exists between the historical data attributes and predicts the weather in future.

P.Hemalatha [27] implemented data mining methods for guiding the path of the ships during sailing. Global Positioning System is used for identifying the area in which the ship is currently navigating. The attributes of weather data includes climate, humidity, temperature, stormy. The weather report of the area traced is compared with the existing database. The analyzed dataset is provided to the decision tree algorithm, C4.5 and ID3. The decision obtained regarding the weather condition is instructed to the ship and the path is chosen accordingly. A close cooperation between the statistical and computational communities provides synergy in data analysis. Few continuous attributes need to be altered as ID3 cannot directly deal with the continuous ranges.

Subana Shanmuganathan and Philip Sallis [34] examined the use of data mining methods to search for the patterns in the ad-hoc weather conditions, such as time of the day, month of the year, wind direction, speed, and severity using a data set from a single location. The historical weather data, between 2008 and 2012 is used from telemetry devices installed in a vineyard in the north of New Zealand. It is shown that using data mining techniques and the local weather condition recorded at irregular intervals can produce new knowledge relating to wind gust patterns for vineyard management decision making. From the data repository, instances relating to the Kumeu River vineyard are extracted for a period of four years (2008–2012). The data collected is cleaned to remove all readings that are outside of Kumeu record readings. The final 86,418 instances and their distribution over the 12 months are presented. The decision tree algorithms used are C5, Quest, CRT and CHAID. SOM is used for the clustering purpose. Multilayered supervised ANN is used for predicting the wind gust. Data mining techniques and statistical methods are run using SPSS. It provides a good tool for analyzing adhoc dataset.

A.R.W.M.M.S.C.B. Amarakoon [1] proposed a system that uses the historical weather data and applies the data-mining algorithm "K-Nearest Neighbor (KNN)" for classification of these historical data into a specific time span. The k nearest time spans is then further taken to predict the weather of Sri Lanka. The day to day weather data is collected for complete one year. It generates accurate results within a reasonable time for months in advance. It is concluded that KNN is beneficial to dynamic data, the data that changes or updates rapidly and provides better performance as compared to the other techniques. Integrating feature selection techniques can even give more accurate results.

Kavita Pabreja [16] demonstrated the derivation of sub-grid scale weather systems from NWP model output products using data mining techniques which is not possible through normal MOS technique. Data mining technique, clustering, when applied on divergence and relative humidity can provide an early indication of formation of cloudburst. K mean clustering is used for two days data of real life case of cloudburst. An effort is made towards providing timely and actionable information of these events using data mining techniques in supplement with NWP models. One shortcoming is found that it cannot be used for long term predictions.

S. S Badhiye *et al.* [4] used clustering technique with K-Nearest Neighbor method to find the hidden pattern inside the large dataset related to weather so as to transfer the retrieved information into usable knowledge for classification and prediction of climate condition. Temperature and humidity is acquired for a particular time interval. High prediction accuracy is acquired for temperature and humidity. The software can be embedded with the data logger system for the analysis and prediction of parameters in remote areas.

Pinky Saikia Dutta and Hitesh Tahbilder [28] predicted monthly Rainfall of Assam by using data mining technique. Traditional statistical technique -Multiple Linear Regression is used. The data include Six years period between 2007 and 2012 which is collected locally from Regional Meteorological Center, Guwahati, Assam, India. The data is divided into four month for each season. Parameters selected for the model are minimum temperature, maximum temperature, mean sea level pressure, wind speed and rainfall. The performance of this model is measured in adjusted R-squared implemented in C#. Some parameters like wind direction is not included due to constraints on data collection which could give more accurate result. Acceptable accuracy is given by prediction model based on multiple linear regression.

Neha Khandelwal and Ruchi Davey [25] predicted the rainfall of a year by using different 4 climatic factors temperature, humidity, pressure and sea level and thereby using the dataset for calculating drought possibilities in Rajasthan. Certain factors are extracted using data mining techniques. Then correlation analysis is applied on the dataset and correlation is found in the factors. The factors with positive correlations are selected and used for regression analysis. MLR is used for regression analysis for predicting rainfall. Then statistical analysis is applied on that data for finding drought possibility. For drought possibility standard deviation, variance of coefficient, drought indices and drought perception are used. Only one parameter rainfall is considered for analyzing drought condition whereas other climate factors may influence the condition to a wide range. Therefore it is not so accurate.

Z. Jan *et al.* [32] developed new accurate and sophisticated systems for Seasonal to inter annual climate prediction using data mining technique, K-Nearest Neighbor (KNN). It uses numeric historical data to predict the climate of a specific region, city or country months in advance. Dataset consist of 10 years of historic data with has 17 attributes, i.e. mean temperature, Max Temp, Min Temp, Wind Speed, Max Wind Speed, Min Wind Speed, Dew Point, Sea Level, Snow Depth, Fog, gust, SST, SLP, etc., with 40000 records for 10 cities. The dataset uses data cleansing to deal with noisy and missing values. It is stored in MS ACCESS format. It can predict a huge set of attribute at the same time with high level of accuracy. The predicted result of KNN is easier to understand. It cannot incorporate to reflect the global changes (ENSO events) but can work correctly with the areas not prone to these global effects.

Soo-Yeon Ji et al. [33] predicted the hourly rainfall in any geographical regions time efficiently. The chance of rain is first determined. Then only if there is any chance of rainfall, the hourly rainfall prediction is performed. Although quite a lot

methodology have been introduced to predict hourly prediction, most of them have performance limitations because of the existence of wide range of variation in data and limited amount of data. CART and C4.5 are used to provide outcomes, which may provide hidden and important patterns with transparent reasons. About 18 variables were used from weather station. For validation purpose, 10 fold cross validation method is performed. CART gives slightly better performance than C4.5. Considering the chances, only a small number of instances are left for prediction which makes it hard to predict.

S. Kannan and S. Ghosh [29] contributed towards developing methodology for predicting state of rainfall at local or regional scale for a river basin from large scale climatological data. A model based on K- mean clustering technique coupled with decision tree algorithm, CART, is used for the generation of rainfall states from large scale atmospheric variables in a river basin. Daily rainfall state is derived from the historical daily multi-site rainfall data by using K-mean clustering. Various cluster validity measures are applied to observed rainfall data to get the optimum number of clusters. CART is used to train the data of daily rainfall state of the river basin for 33 years. The methodology is tested for the Mahanadi River in India. The change expected in the river basin due to global warming is given by the comparisons of the number of days falling under different rainfall states for the observed period and the future predicted. CART algorithm proved to be good in predicting the daily rainfall state in a river basin using statistical downscaling.

IV. COMPARISON OF DATA MINING TECHNIQUES

According to the previous work done by researchers presented in the literature review, a comparison can be done. Various data mining techniques are used to predict different parameters of weather like humidity, temperature, wind gust. Various attributed used for the comparison are applications, authors, data mining techniques, algorithms, attributes, time period, dataset size, accuracy percentage, advantages and disadvantages. They yield different results with their cons and pros. The main consequence of this fact is formulated by the 'no-free lunch theorem', which states that there is no universally best data mining algorithm. This triggers the need to select the appropriate learning algorithm for a given problem. For weather prediction, decision tree and k-mean clustering proves to be good with higher prediction accuracy than other techniques of data mining. Regression technique could not find accurate value of prediction. However, approximate value could be retrieved. It is also observed that with the increase in dataset size, the accuracy first increases but then decreases after a certain extent. One of the reasons may be due to over fitting of training dataset. The work done by different researchers and their comparison is jotted down in Table I.

TABLE I
COMPARISON OF A DATA MINING TECHNIQUES FOR WEATHER PREDICTION

Authors	Applications	Techniques	Algorithms	Attributes	Time Period	Dataset Size	Accuracy	Advantages	Disadvantages
P. Hemalatha [27]	Weather prediction for ship navigation	Decision tree	C4.5, ID3	Climate, Humidity, Stormy, Temperature	4-5 location	20- 30 instances	-	Verifiable performance	Do not handle continuous range data directly.
E. G. Petre [10]	Weather prediction	Decision tree	CART	Pressure, clouds quantity, humidity, precipitation, temperature	4 years	48 instances	83%	Good prediction accuracy	Data transformation is required. Extra computation required.
S Yeon *et al.*[33]	Hourly rainfall prediction	Decision tree	C4.5, CART	Temperature, wind direction, speed, gust, humidity, pressure	3 years	26280 instances	99%, 93%	High prediction accuracy	Small data is left for prediction.
S Kannan , S Ghosh [29]	Daily rainfall prediction in river basin	Decision tree, Clustering	CART, k-Mean clustering	Temperature, MSLP, pressure, wind, rainfall	50 years	432000 instances	-	Grouping of multisite rainfall data in clusters	Small data is left for prediction. No verification is done.
F Oliya, AB Adeyemo [17]	Weather Prediction and Climate Change Studies	Decision tree, ANN	C4.5, CART, TLFN	temperature, rainfall, evaporation, wind speed	10 years	36000 instances	82%	Best network is selected for prediction	Accuracy varies highly with size of training dataset
P Sallis, S Shanmuganathan [34]	Wind gust prediction	Decision tree, ANN	C5.0, CRT, QUEST, CHAID, SOM	Dew point, humidity, temperature, wind direction, wind speed	4 years	86418 instances	99%, 85%	Good for analyzing ad hoc dataset	Data recorded at irregular intervals. Do not handle continuous data.
GJ Sawale [12]	Weather prediction general	ANN	BPN, Hopfield networks	Temperature, humidity, wind speed	3 years	15000 instances	-	Combining both gives better prediction accuracy	Attribute normalization is required
Amarakoon [1]	Climate prediction in Sri Lanka	ANN	KNN	Temperature, humidity, precipitation, wind speed	1 year	365 instances	-	Beneficial for dynamic data.	Need to integrate feature selection techniques.

S Badhiye *et al* [4].	humidity and temperature prediction	Lazy learning, clustering	KNN, K-mean clustering	Temperature, humidity	-	-	100% approx	Suitable for multi-modal classes.	Cannot predict data in remote areas
Z Jan *et al.* [38]	Inter annual climate prediction	Lazy learning	KNN	Wind speed, dew point, seal level, snow depth, rain	10 years	40000 instances	96%	Long term accurate results with large set of attributes.	Cannot incorporate to reflect global changes.
M. A. Kalyankar, S. J. Alaspurkar [23]	Meteorological data analysis	Clustering	K- mean clustering	Temperature, humidity, rain, wind speed	4years	8660 instances	-	Good prediction accuracy	Dynamic data mining methods required.
K Pabreja [16]	Cloud burst predicion	Clustering	K- mean clustering	Temperature, humidity	2 days		100% clustering	Supplement with NWP models.	Not good for long term predictions.
PS Dutta, H Tahbilder [28]	Rainfall prediction	Regression	MLR	Min and max temperature, wind direction, humidity, rainfall	6 years	72 instances	63%	Acceptable accuracy.	Attribute elimination required for better accuracy
M Kannan *et al.*[32]	Short Term Rainfall prediction	Regression	MLR	Min and max temperature, wind direction, humidity, rainfall	3 months for 5 years	450 instances	52%	Can work even with small dataset	Instead of accurate, an approximated value is retrieved.
N Khandelwal , R Davey [25]	Drought prediction	Regression	MLR	Rainfall, sea level, humidity, temperature	1 year	365 instances	-	Coorelation and statistical analysis is also applied.	Verification is not done.

IV. CONCLUSION

This paper presents a survey that using Data mining techniques for weather prediction yields good results and can be considered as an alternative to traditional metrological approaches. The study describes the capabilities of various algorithms in predicting several weather phenomena such as temperature, thunderstorms, rainfall and concluded that major techniques like decision trees, lazy learning, artificial neural networks, clustering and regression algorithms are suitable to predict weather phenomena. A comparison is made in this paper, which shows that decision trees and k-mean clustering are best suited data mining technique for this application. With the increase in size of training set, the accuracy is first increased but then decreased after a certain limit.

ACKNOWLEDGMENT

I acknowledge my sincere and profound gratitude to my guide, Er. Jawahar Thakur, for his valuable guidance, dedicated concentration and support throughout this work. I also acknowledge my sincere gratitude to authorities of Himachal Pradesh University, Summerhill and other teaching staff of Computer Science for their help and support. I am also thankful to my friends for their cooperation.

REFERENCES

[1] A.R.W.M.M.S.C.B. Amarakoon, "Effectiveness of Using Data Mining for Predicting Climate Change in Sri Lanka", 2010.
[2] Abhishek Saxena, Neeta Verma, Dr K. C. Tripathi, "A Review Study of Weather Forecasting Using Artificial Neural Network Approach", International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 11, November – 2013.
[3] Auroop R Ganguly, and Karsten Teinhaeuser, "Data Mining for Climate Change and Impacts," IEEE International Conference on Data Mining, 2008.
[4] Badhiye S. S., Dr. Chatur P. N., Wakode B. V., "Temperature and Humidity Data Analysis for Future Value Prediction using Clustering Technique: An Approach", International Journal of Emerging Technology and Advanced Engineering, 2250-2459, Volume 2, Issue 1, January 2012.
[5] Badhiye S. S., Wakode B. V., Chatur P. N. "Analysis of Temperature and Humidity Data for Future value prediction", IJCSIT Vol. 3 (1), 2012.
[6] Boris Mirkin, "Clustering: A Data Recovery Approach", Second Edition, Chapman and Hall/CRC, October, 2012.
[7] Cohen, J., Cohen P., West, S.G., & Aiken, L.S. Applied multiple regression/correlation analysis for the behavioral sciences, 2nd ed., Hillsdale, NJ: Lawrence Erlbaum Associates, 2003.
[8] D. Singh, A. Ganju, A. Singh, "Weather prediction using nearest neighbor model" Current Science, vol. 88, no. 8, 25: 1283-1289, April 2005.
[9] DL Gupta, A. K. Malviya, S. Singh, "Performance Analysis of Classification Tree Learning Algorithm", International Journal of Computer Application, Vol 55, 2012.
[10] Elia Georgiana Petre "A Decision Tree for Weather Prediction", Buletinul, Vol. LXI No. 1, 77-82, 2009.
[11] Folorunsho Olaiya, Adesesan Barnabas Adeyemo, "Application of Data Mining Techniques in Weather Prediction and Climate Change Studies", I.J. Information Engineering and Electronic Business, 51-59 ,July- 2012.
[12] Gaurav J. Sawale, Dr. Sunil R. Gupta, "Use of Artificial Neural Network in Data Mining For Weather Forecasting", International Journal Of Computer Science And Applications Vol. 6, No.2, Apr 2013.
[13] Han J., Kamber M. Data Mining: Concepts & Techniques, Morgan & Kaufmann, 2000.
[14] International Conference on 31, 163-167, doi: 10.1109/ WISM, 2013.
[15] J. Han, M. Kamber. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2000.
[16] Kavita Pabreja, "Clustering technique to interpret Numerical Weather Prediction output products for forecast of Cloudburst", International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 3 (1) , 2996 - 2999, 2012.
[17] Kaya, E.; Barutçu, B.; Menteş, S. "A method based on the van der Hoven spectrum for performance evaluation in prediction of wind speed". Turk. J. Earth Science, 22, 1–9, 2013.
[18] L. Ertoz, M. Steinbach, V. Kumar, "Finding Clusters of Different Sizes, Shapes, and Densities in Noisy, High Dimensional Data", In Proc. of the 3rd SIAM International Conference on Data Mining, San Francisco, CA,USA,2003.
[19] Linnenluecke, M.K.; Griffiths, A.; Winn, M., "Extreme weather events and the critical importance of anticipatory adaptation and organizational resilience in responding to impacts", Business Strategy Environment, 21, 17–32, 2012.

[20] Lior Rokach, Oded Maomom,"Data Mining with Decision Tree: Theory and Application", World scientific publishing Co. Pte Ltd., 2008.
[21] Lior Rokach, Oded Maomom,"Data Mining with Decision Tree: Theory and Application", World scientific publishing Co. Pte Ltd., 2008.
[22] M.Kannan, S.Prabhakaran, P.Ramachandran, "Rainfall Forecasting Using Data Mining Technique", International Journal of Engineering and Technology Vol.2 (6), 397-401, 2010.
[23] Meghali A. Kalyankar, S. J. Alaspurkar, " Data Mining Technique to Analyse the Metrological Data", International Journal of Advanced Research in Computer Science and Software Engineering 3(2), 114-118, February – 2013.
[24] Nathalie Japkowicz, Mohak Shah, "Evaluating Learning Algorithms: A Classification Perspective" First Edition, Cambridge University Press, 2011.
[25] Neha Khandelwal, Ruchi Davey, " Climatic Assessment Of Rajasthan's Region For Drought With Concern Of Data Mining Techniques", International Journal Of Engineering Research and Applications (IJERA), Vol. 2, Issue 5, 1695-1697,September-October 2012.
[26] P. Hall, B.U. Park, R. J. Samworth, "Choice of neighbor order in nearest-neighbor classification", Annals of statistics 36(5), 2135-2152, 2008.
[27] P.Hemalatha, "Implementation of Data Mining Techniques for Weather Report Guidance for Ships Using Global Positioning System", International Journal Of Computational Engineering Research Vol. 3 Issue. 3 , march 2013.
[28] Pinky Saikia Dutta, Hitesh Tahbilder, "Prediction Of Rainfall Using Data mining Technique Over Assam", Indian Journal of Computer Science and Engineering (IJCSE), Vol. 5 No.2 Apr-May 2014.
[29] S. Kannan , Subimal Ghosh, "Prediction of daily rainfall state in a river basin using statistical downscaling from GCM output", Springer-Verlag, July- 2010.
[30] S. Kotsiantis and, "Using Data Mining Techniques for Estimating Minimum, Maximum and Average Daily Temperature Values", World Academy of Science, Engineering and Technology, 450-454, 2007.
[31] Sarah N. Kohail, Alaa M. El-Halees, "Implementation of Data Mining Techniques for Meteorological Data Analysis", IJICT Journal Volume 1 No. 3, 2011.
[32] Simon S. Haykin, "Neural Networks: A Comprehensive Foundation", Second Edition, Prentice Hall International, 1999.
[33] Soo-Yeon Ji, Sharad Sharma, Byunggu Yu, Dong Hyun Jeong, "Designing a Rule-Based Hourly Rainfall Prediction Model", IEEE IRI 2012, August – 2012.
[34] Subana Shanmuganathan and Philip Sallis, "Data Mining Methods to Generate Severe Wind Gust Models", 5, 60-80, Atmosphere 2014.
[35] Taylor, R.G.; Scanlon, B.; Döll, P.; Rodell, M.; R. Beek, V.; Wada, Y.; Longuevergne, L.; Leblanc, M.; Famiglietti, J.S.; Edmunds, M.; et al., "Ground water and climate change", National Climate Change, 3, 322–329, 2012.
[36] University of Alberta, Osmar R. Zaïane, "Chapter I: Introduction to Data Mining", CMPUT690 Principles of Knowledge Discovery in Databases, 1990.
[37] Willenbockel, D, "Extreme Weather Events and Crop Price Spikes in a Changing Climate, Illustrative Global Simulation Scenarios"; Oxfarm Reserach Reports, 2012.
[38] Zahoor Jan, M. Abrar, Shariq Bashir, and Anwar M. Mirza, "Seasonal to Inter-annual Climate Prediction Using Data Mining KNN Technique", Springer-Verlag Berlin Heidelberg, CCIS 20, 40