International Journal of Computations, Information and Manufacturing (IJCIM) 3(2) -2023

Contents available at the publisher website: G A F T I M . C O M

International Journal of Computations, Information and Manufacturing (IJCIM)

Journal homepage: https://journals.gaftim.com/index.php/ijcim/index

Rainfall Prediction using Big Data Analytics: A Systematic Literature Review

Mahwish Anwer Cheema [1], Madeeha Saqib [2], Sardar Zafar Iqbal [3]

[1] *National College of Business Administration & Economics, Pakistan*

[2,3] *Department of Computer Information Systems, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia*

A R T I C L E I N F O

Keywords:
Rainfall prediction; big data analytics, SLR

Received: Nov, 02, 2023
Accepted: Nov, 23, 2023
Published: Dec, 22, 2023

A B S T R A C T

With major ramifications for agriculture, water resource management, and disaster planning, rainfall prediction is an essential component of weather forecasting. The use of big data analytics techniques has become more commonplace in recent years as a means of improving rainfall prediction models' accuracy and dependability. The goal of this systematic literature review is to present a thorough summary of the state of the art in the field of big data analytics-based rainfall prediction research. The first section of this paper provides a thorough examination of the basic ideas and procedures used in rainfall prediction models. It emphasises how crucial it is to incorporate data from a variety of sources into predictive models, such as social media, meteorological, and remote sensing data. This complete overview of the state-of-the-art in big data analytics-based rainfall prediction is provided by this systematic literature review. It highlights the need for multidisciplinary collaboration between meteorologists, data scientists, and domain specialists to further develop the subject of rainfall prediction and its applications. It also identifies gaps in the current research and recommends prospective directions for future studies.

1. INTRODUCTION

Big data contains a large volume of data that includes both structured and unstructured data. And huge volume of data is not an easy task to store, analyze process, share, visualize and manage with the traditional database and software techniques (Shabariram, et al., 2016). Big data analysis in learning, data exists in large volume. It is applied where we want to check hidden patterns e.g. in health care, consumer demand and adds etc. In big data, it takes tasks through clustering the data and the main purpose of big data analytics is to help in decision making (Shatnawi, M. et al., 2020) By entering in the era of big data analysis which is already applied in many fields, e-business cane make efficient and better decisions to meet their needs (Zhang, et al., 2021). Big data analysis gathered the multiple information e.g. from health organizations. It brings new information and with the help of big data analysis unstructured data convert into structured data. Also, quality of data matters more than the quantity of data. Due to increase in information the data is getting huge. The development in big data brought a lot of benefits for technology. Cloud computing and big data analysis are two sides of coins. Because big data cannot be controlled with one device so cloud computing plays a vital role in distributing properties. When there is need to handle data with a proper algorithm then machine learning also plays a great role in big data analysis. Big data also provide huge information for internet of things (IOT). Unstructured data can become structured or

semi-structured by using the specific machine learning algorithm. Use of machine learning can solve the problems related to big data (Molinari and G. Nollo, 2020). Data evaluated based on cluster tendency, we need to detect the cluster tendency. Some are the existing techniques: Visual assessment tendency (VAT), improved VAT (iVAT) and Spectral- based VAT (spec VAT). These techniques used for small datasets but recently a method is developed for assessment of cluster tendency for big data is big VAT (R. Hou, et al., 2020).

Time series plays a prominent role in rainfall prediction. The data for time series can be gathered on different time periods such as hourly, daily, weekly, monthly, quarterly and yearly [7]. Time series is divided into single variate and multivariate methods[8]. Data need to be accurately efficiently prepared in order to take out functional and valuable information. Thus, the evaluation of new tools for connecting with big data has become a sensitive issue. An important and prominent factor of the nature of big data is that data collection related to specific study is usually fetched over time at different points, resulting in big data time series.

In literature many researchers already have talked about that prediction and forecasting are the most challenging tasks in term of convolution and technologically. Forecasting for rainfall is based on the time series and this idea can get from statistical theory. Four major issues: determination of effective lengths of intervals, repeated fuzzy sets, and trend associated with fuzzy sets, and defuzzification operation can be deal in fuzzy time series (FTS) (Rajendra Prasad, et al., 2021). The quantity of rainfall is a main factor in the field of agriculture. This natural event highly effect the weather.it is also revealed in many studies that the quantity of rainfall also depends on other weather parameters. In Asia the south-eastern really depends on the annual rainfall. Many researchers already have given a lot of models for predicting rainfall. Artificial neural network (ANN) is a most dominant model for predicting rainfall. Due to heavy quantity of rainfall flood warning is the major problem. This quantitative forecasting can be willingly used by the flood warning system to expand the lead time for warning. Some uncertainties become very complex for prediction models and flood is one of most prominent problem. To overcome this complexity machine learning (ML) provides batter models for weather prediction (S. Aftab, M. Ahmad, et al., 2018). In many natural disasters landslides are one of the most destructive have affected several mountainous regions across the globe. The big subjection to landslide possibility has made the Indian Himalayas collect the heighten attentiveness by the landslides community . A study revealed that two prominent tools for time series analysis and forecasting are Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) (Y. Sakurai, Y. Matsubara, et al., 2017) The authors have reviewed several rain forecasting models by NN (Neural Networks) like FFNN, RNN, and TDNN.

Further this SLR contains three sections. In second section namely related work, describes literature on rainfall prediction and forecasting using different techniques of big data. In next section namely research protocol, detailed discussion on literature of rainfall prediction using big data analytics. In last section, there is results and discussion of whole study.

2. RELATED WORK

Big data has thoroughly changed the ways through which human community embrace natural disaster management scheme to minimize human agonize and economic losses. There is less accuracy in the modeling for prediction and forecasting of rainfall due to two main reasons. First, there are no same conditions for every rainfall due to change in some factor, so in result there is not accurate prediction for rainfall. Second, quantity of factors are not sure and also variables are not constant which effects the rainfall (P. Singh, 2018). The development of the intelligent computing methods, many researchers already have proposed rainfall prediction methods, Artificial Neural Network being one of the most well-known method. Feed-forward back propagation algorithm, layer recurrent algorithm and feed-forward distributed time delay algorithm are three methods have used to create a model of artificial neural network for rainfall prediction in a coastal region of India (S. Chatterjee. et al.,m 2018) For Summer monsoon rainfall prediction of India on monthly and seasonal time series by using data of past five years, used error back propagation algorithm. The artificial neural network is used for rainfall

prediction of Hyderabad, region of India and then compare with ARIMA technique . (Mosavi, P. Ozturk, and K. W. Chau, 2018). Prediction of rainfall in Thailand, the Back Propagation Neural Network is used for reveling an acceptable accuracy. ARIMA model used to predict the rainfall and also it gave future trend of rainfall using existing data in past times (S. Bahri, et al., 2019). Many researches used of Box-Jenkins ARIMA model to forecast the temperature data, rainfall, humidity, rice productions etc (A. Dikshit, et al., 2020). In this study, researcher proposed nearest neighbors modeling and Map Reduce Method techniques for prediction of crop decisions using on big data analytics which is different from conventional methods. Researcher composed datasets from China Meteorological Administration (1995-2014) and gave results after analyzed the data. It focused on agricultural mining from time aspect and it gave better and efficient results compared to conventional methods (M. Yu, C. Yang, et al., 2018) Map Reduce algorithm shown in fig.1. In this study, researcher used ANN for the purpose of forecast monthly and seasonal rainfall in Queensland, Australia. As input this network takes meteorological or climatic circulation components, such as temperature and El Nino index. They also discussed about the input collection and optimization of ANN (N. Sultana and M. M. Hasan, 2015). Research in this study initiate a model based on DBN for forecasting of rainfall. In two cites of China namely Zunyi and GuiYan, efficiency of the model is manifest.

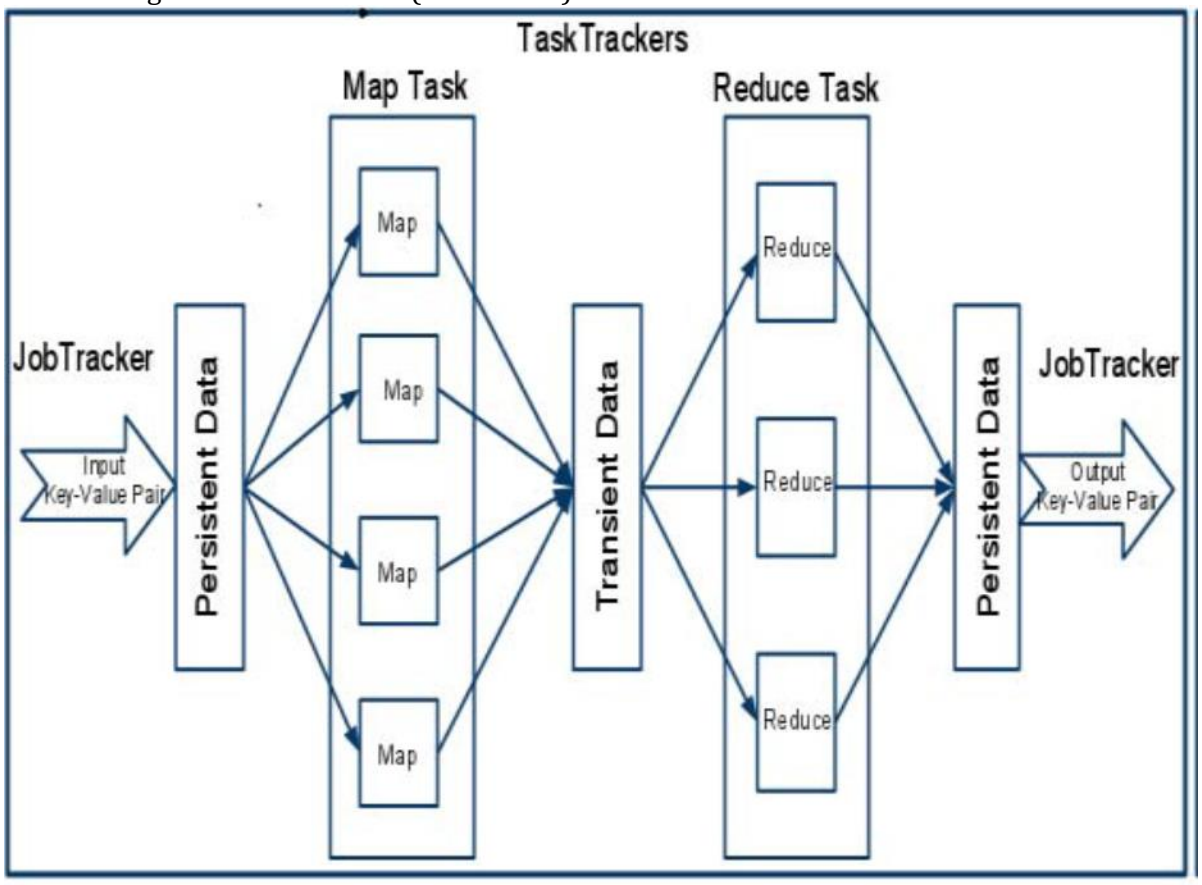

Fig. 1. Map Reduce Algorithm

3. RSEARCH PROTOCOLE

Systematic literature review (*SLR*) with a high standards is one which obtain its purpose by providing the compress details and facts of essential research topic for a specific time span. A detailed research methodology with step by step directions is needed to manage a productive *SLR*. In (W. Fan, C. Chong, et al., 2016),

the evaluation of *SLR* includes three key points with sub tasks shown in fig. 2.

From the beginning of SLR the research questions already have recognized. The

purpose of research is specified and extended through the research questions shown in table1.

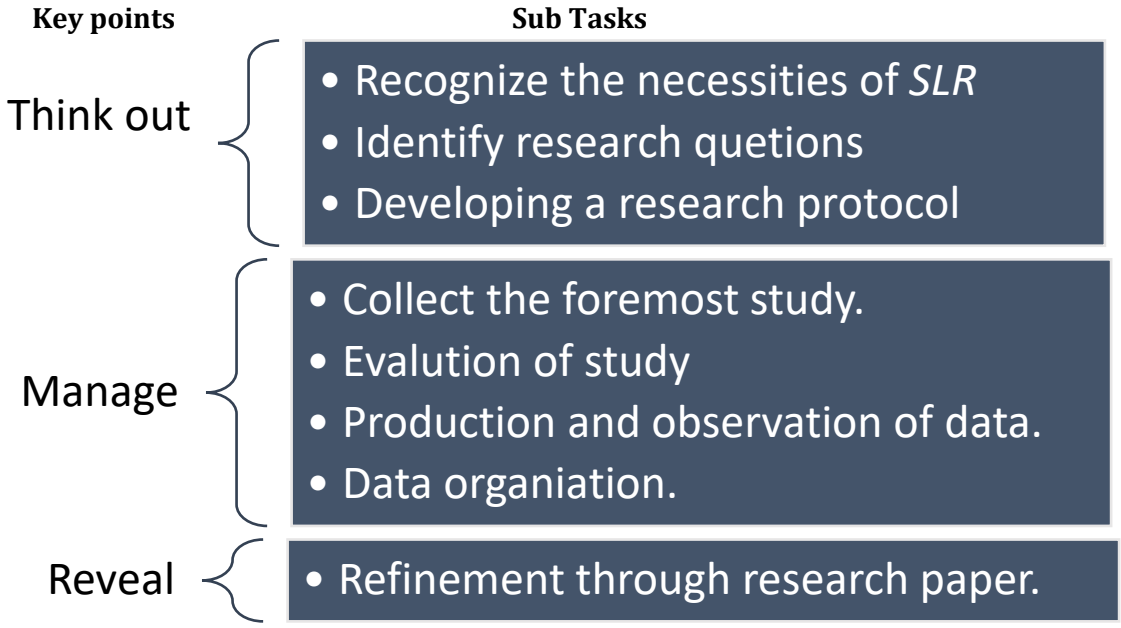

Fig. 2. Key points and sub tasks of *SLR*

Table 1. Identification of research questions

RQ1.	Which big data analytics techniques are used / proposed for rainfall prediction?
RQ2.	How the performance of prediction and forecasting techniques is assess?
RQ3.	Which type of data is used for prediction and forecasting?
RQ4.	For which location the rainfall prediction and forecasting is performed?
RQ5.	Which elements affect the prediction and forecasting results?

In detailed literature analysis, Full text of 15 selected articles were analyzed and then 7 most relevant research papers are shortlisted for critical review.

3.1. MapReduce and Optimized Deep Network for Rainfall Prediction in Agriculture

In J. Abbot and J. Marohasy, (2014), talked about that there are existing models which are not effective for rainfall prediction in some factors: humidity, rainfall and rainfall recorded in past years. To handle this situation they proposed in this research two main models. (i) Deep learning network which built in the MapReducd that allows productive handling of big data analytics and also this framework works on multiple sources disregarding of weather the data is structure or unstructured. The MapReduce framework refine visualization and come up with a productive prediction environment with a highly climbable environment. The MapReduce programming include two tasks, in first it create a map and then in second it reduced to engaged in mapping the input data. The map and reduce tasks of the MapReduce framework utilizes the *S-SGD* algorithm-based convLSTM. (ii) The development of the SSGD-based convLSTM model for predicting the rainfall and this model handled the time series in best way. convLSTM predicts the future using the inputs and the past states corresponding to its neighbors, which is facilitated using the convolutional operator $*$ and Hadamard product. In this research for experimental work they performed analysis on MATLAB and took six datasets which contains the weather data of the rainfall predicted in entire India and the state,

Tamil Nadu, from the years 1901 to 2015. Three datasets are taken from Indian database and next three datasets are taken from Tamil Nadu database contains rainfall prediction based on month-wise, quarterly-wise and yearly-wise respectively. To assess the performance of model they used two validity measures: percent root mean square difference *(PRD)* and mean square error *(MSE).* They conclude that proposed model performed better than existing models.

3.2. Rainfall Prediction Using Data Visualisation Techniques

Due to the difficulty in analyzing rainfall and the capacity to get information from the previous datasets, data visualisation performs an essential part in the analysis of the large amount of data and it could provide a great aid to big data analytics. (P. Zhang, et al., 2018), by visualising data can make sense of large data sets. Visualising also motivate to extend patterns trends and differentiating various data. By extend these patterns has great use in Search Engine Optimisation (*SEO*). Visualising data shows all the complexities of confound subjects which is make use of in the field big data analytics. In this paper authors have taken datasets the India Meteorological Department (*IMD*) Govt. of India under Govt. Open Data License – India based on the monthly and annually rainfall detail of 36 meteorological sub-divisions and every state and union territory from 1901 to 2015 of India.

The methodology of this paper involves several steps. (i) Analyzing the problem: this step can be done by studding the datasets and then type of visualisation can be selected for specific dataset. (ii) Cleaning the data set: by data cleaning means it can find out the faulty, insufficient, immaterial and mistaken parts of the data and remove or replace them with median values. (iii) Data processing: After cleaning the dataset data processing can be done and used to attain the quality standards of a dataset. This step includes the substitution of NaN values with median values and then grouping of data with respect to their respective states. (iv) and, algorithm and methods: two graphs have used, bar graph is used to visualise the both levels that highs and lows of rainfall in each state and maximum and minimum rainfall of all the states is compared and line graph is used for visualising the annual rainfall of each state from 1901-2015. They discussed about minimum rainfall state Territories and minimum rainfall state Territories. Author showed by using line graph in this paper that West Rajasthan has the least rainfall in India over 20000mm in a century followed by Saurashtra & Kutch and Haryana Delhi & Chandigarh and average annual rainfall in West Rajasthan is around 300mm in minimum rainfall state Territories. In minimum rainfall state Territories Coastal Karnataka has the highest rainfall in India. The average annual rainfall is around 3400mm. The results acquired in this paper can be used for the prediction of rainfall with the help of regression which can be of great benefit in the field of agriculture.

3.3. Development of advanced artificial intelligence models for daily rainfall prediction

In O. Manoj and J. P. Ananth (2020), author has given different models for daily rainfall prediction in Hoa Binh province, Vietnam. In this study, five parameters: maximum temperature, minimum temperature, wind speed, relative humidity and solar radiation have used as input for rainfall prediction. And five methods have used in this research: (i) Artificial Neural Networks (*ANN*). When many machine learnings work together and process multiplex input through a framework form it becomes neural network, so *ANN* based on organizational and obligations of biological neural network. After learning process *ANN* process net data and can give results as output without having input. In this research, the scaled conjugate gradient algorithm was applied for training the neural network as it accelerates the convergence by guiding the search in the conjugate directions. (ii) Adaptive network based fuzzy inference system (*ANFIS*). It took input first and went through five layers of *ANFIS* and give overall output. *ANFIS* is a combination of adaptive networks and a fuzzy inferences system. (iii) Particle Swarm Optimization (*PSO*). It was propose for solving multiplex optimization problems. (iv) Support Vector Machines (SVM). It was developed using statistical learning theory. The main purpose of SVM algorithm is to find a hyperplane in an *N* dimensional space that distinctly separates data points, maximizing the distance between two datasets and (v) Validation criteria. In this research paper, they have used six validity measures to check the performance of proposed model which are Correlation Coefficient (R) and Mean Absolute

Error (*MAE*), Score Skill (*SS*), Probability of Detection (*POD*), Critical Success Index (*CSI*) and False Alarm Ratio (*FAR*). In this study, rainfall data was collected from meteorological gauge located at Latitude 20.763 m and Longitude 105.312 m in the Cao Phong district, Hoa Binh province, Vietnam (Fig. 4). Weather data was obtained from the Global Weather Data for SWAT. They collected total 3653 samples during the period of January 01, 2004 to December 31, 2013. The proposed model of this paper carried out four steps in modeling for prediction of rainfall: generation of datasets, construction of the models, validation of the models, and robustness analysis.

3.4. Short-term Rainfall Forecasting Using Multi-layer Perceptron

In Y. K. Joshi, (2020), they have proposed Dynamic Regional Combined short-term rainfall Forecasting approach (*DRCF*) using Multi-layer Perceptron (*MLP*). In the section of related work of this paper they have talk about atmospheric model and machine learning approaches. Atmospheric models basically based on numerical method. In the situation of uncertainty, the results numerical methods deviate from real situation, so we can get results on numerical method by forecasting. In machine learning approach first is based on time series and it can get from statistical theory. In rainfall prediction time series plays a vital role. They have talk about other approach metrological. It used in rainfall forecasting mainly include Support Vector Machine (*SVM*), Support Vector Regression (*SVR*), Grey Forecasting (*GF*) and Artificial Neural Network (*ANN*), and they give better results for rainfall prediction then statistical theory. In processing data, where are too many variables it become complex. Principal component analysis (*PCA*) simplifies the complexity of variables and gives the more information with less variables. In this paper, *PCA* is used to reduce the dimension of thirteen physical factors and then it serves to the multi-layer perception (*MLP*) as input. They have talk about two ways of *MLP* in but have adopted only second method due to it is fast and cheap. The method is find the structure of *MLP* namely greedy algorithm. The results in *MLP* can get from the input layer to output layer. The experimental data include 2015-2017 years' Micaps surface mapping data, altitude (500hPa) mapping data and numerical forecasting results released by the China Meteorological Administration.

3.5. Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman's rho tests and ARIMA model

In this study, B. Thai, L. Minh, (2019), firstly talk about that there are many countries in which there is heavy rainfall in summer monsoon and low rainfall in remaining time period. In south Asia Bangladesh is one of these countries. There are many climate factors which effect the agriculture and rainfall is most prominent. Bangladesh Meteorological Department (*BMD*) has 34 weather stations covering the country from which weather information is collected including rainfall data. Due to new establishment of few stations they do not have log-term data on rainfall. In this research paper they have selected 14 stations out of 34 stations and every station contains the data more than 60 years. They have use non-parametric test due to the fear of outliers namely Mann–Kendall and the Spearman's rho tests. They have used a model in this study is advanced the AutoRegressive Integrated Moving Average (*ARIMA*). Basically, this model contains three parts; autoregressive (*AR*), moving average (*MA*) part and both the parts are integrated in the model by the differencing order. After achieving the static of data, four steps includes in *ARIMA* model application: model identification, model parameter estimation, diagnostic of the model adequacy and forecast. They have checked the skewness of data in this study which is positively skewed and varies between 1.044356 and 1.149507. Also they showed the kurtosis results of data which are leptokurtic. They checked results of data by using descriptive statistics.

3.6. Survey on Weather Prediction using Big Data Analytics

In M. A. Rahman, (2016), they discussed that uncertainty cause huge losses such as flood, earth quick, storm, cyclones etc. In present days weather forecasting is a challenging task. Many rainfall prediction models are available in the field of statistics and in literature but these models could not provide more accurate results. Prediction for future damages is most important task which can be predicted by bothering big data of past years. Especially for agricultural field rainfall prediction

is very important in many countries who real depend on it such as India. In this study, researcher stated that many authors used models for rainfall prediction such as Radial Basis Function Network, *BPA* (Back Propagation Algorithm), *SVM* (Support Vector Machine) and *SOM* (Self Organization Map). This study mainly discuss about survey report on early prediction of weather conditions. Due to instability and large volume of weather data set do not give accurate result with *NWP* technique, so here for prediction the straight storm formation prediction system is required. This paper talked about four different type of weather forecasting namely Very short-scale forecast: hour's basis (1-5), Short-scale forecast: 6 hours - few days (week basis), Medium-scale forecast: month's basis (1-10) and Long-scale forecast: Yearly basis. They used map reduced algorithm and linear regression method, and conclude that they perform better and also provide better results.

3.7. Applying big data beyond small problems in climate research

The adequacy of big data and the connected to trained shifts are of specific significance for climate research for three reasons. Firstly, already huge volumes of contemporary climate data are anticipated to expand further in both volume and difficulty over the coming years and decades. Secondly, proceed towards commonly connected

Table 2. Answers of research questions

with big data have entered climate research. Third, climate models are embedded in scientific theory and it is one of the main cause for confidence in their forecast. In climate variables two important and major variables are temperature and rainfall, because they play a prominent role in finding the impact of climate change in a specific region. Rainfall and temperature drifts have been widely investigated for the semi-arid and humid regions of the world. Daily data of rainfall of 35-year (1979–2013), maximum air temperature, and minimum air temperature for 31 grid points, located in the study area, were utilized to generate annual and seasonal (pre-monsoon, monsoon, post monsoon and winter) time series. There are many ways to calculate the volume of data such as by transection (P. Zhang, et al., 2016).

4. RESULT AND DISCUSSION

In this study, seven research paper have been shortlisted and answers of research questions given in table 2.

Which big data analytics techniques are used / proposed for rainfall prediction?	In all above mentioned studies, various climatic attributes/variables from existing weather data were used as predictors for the objective of prediction/forecasting. In all selected studies, researches used combined and modified techniques for rainfall prediction and to prove better efficiency of models.
How the performance of prediction and forecasting techniques is assess?	The main purpose of every researcher is to propose various models/methods/designs with accurate results of rainfall prediction/forecasting. The performance of models/methods/designs given by researchers in all selected research papers were assessed by comparing with existing models and for the performance check of proposed models, statistics measurements used.
Which type of data is used for prediction and forecasting?	All selected paper used existing weather data for rainfall prediction and for the exercise objective used administered big data analytics techniques. Un-ruled techniques of big data analytics were also used in compound of administered techniques. Many climatic attributes were used as predictors including rainfall polarity, rainfall measure, minimum temperature, maximum temperature, wind speed, and humidity etc. Researchers stated, use of more features is not authentic for better efficiency in prediction rather irrelevant attributes could

	affect the performance.
For which location the rainfall prediction and forecasting is performed?	In above detailed literature reviews mostly locality used of region in India. For rainfall prediction locations used situated in Bangladesh and Australia.
Which elements affect the prediction and forecasting results?	After the evaluated review of selected research papers, it is concluded that following given elements could affect the rainfall prediction results: existing weather data: which is selected for exercising the algorithms of big data, climatic attributes: which are used as predictors, location: for which the rainfall prediction has to be performed, surrounding environment, pre-processing techniques and most importantly the used model/technique/method.

Which big data analytics techniques are used / proposed for rainfall prediction? In all above mentioned studies, various climatic attributes/variables from existing weather data were used as predictors for the objective of prediction/forecasting. In all selected studies, researches used combined and modified techniques for rainfall prediction and to prove better efficiency of models.

How the performance of prediction and forecasting techniques is assess? The main purpose of research is to propose various models/methods/designs with most substantial results of rainfall prediction/forecasting. The performance of models/methods/designs given by researchers in all selected research papers were assessed by comparing with existing models and for the performance check of proposed models, statistics measurements used.

Which type of data is used for prediction and forecasting? All selected paper used existing weather data for rainfall prediction and for the exercise objective used administered big data analytics techniques. Un-ruled techniques of big data analytics were also used in compound of administered techniques. Many climatic attributes were used as predictors including rainfall polarity, rainfall measure, minimum temperature, maximum temperature, wind speed, and humidity etc. Researchers stated, use of more features is not authentic for better efficiency in prediction rather irrelevant attributes could affect the performance.

For which location the rainfall prediction and forecasting is performed? In above detailed literature reviews mostly locality used of region in India. For rainfall prediction locations used situated in Bangladesh and Australia.

Which elements affect the prediction and forecasting results? After the evaluated review of selected research papers, it is concluded that following given elements could affect the rainfall prediction results: existing weather data: which is selected for exercising the algorithms of big data, climatic attributes: which are used as predictors, location: for which the rainfall prediction has to be performed, surrounding environment, pre-processing techniques and most importantly the used model/technique/method.

REFERENCES

C. P. Shabariram, K. E. Kannammal, and T. Manojpraphakar, "Rainfall analysis and rainstorm prediction using MapReduce Framework," 2016 Int. Conf. Comput. Commun. Informatics, ICCCI 2016, pp. 7–10, 2016, doi: 10.1109/ICCCI.2016.7479954.

M. Shatnawi, M. B. Yassein, and I. Jalabneh, "The Use of Big Data Analytics in Medical Images: A Survey," 2020 11th Int. Conf. Inf. Commun. Syst. ICICS 2020, pp. 294–298, 2020, doi: 10.1109/ICICS49469.2020.239514.

X. Li, Y. Zhang, Y. Li, K. Yu, and Y. Du, "Study of E-business applications based on big data analysis in modern hospital health management," Inf. Syst. E-bus. Manag., no. 0123456789, 2021, doi: 10.1007/s10257-021-00520-9.

A. Molinari and G. Nollo, "The quality concerns in health care Big Data," 20th IEEE Mediterr. Electrotech. Conf. MELECON 2020 - Proc., pp. 302–305, 2020, doi: 10.1109/MELECON48756.2020.9140534.

R. Hou, Y. Q. Kong, B. Cai, and H. Liu, "Unstructured big data sanalysis algorithm and simulation of Internet of Things based on machine learning," Neural Comput. Appl., vol. 32, no. 10, pp. 5399–5407, 2020, doi: 10.1007/s00521-

019-04682-z.
K. Rajendra Prasad, M. Mohammed, L. V. Narasimha Prasad, and D. K. Anguraj, "An efficient sampling-based visualization technique for big data clustering with crisp partitions," Distrib. Parallel Databases, no. 0123456789, 2021, doi: 10.1007/s10619-021-07324-3.
S. Aftab, M. Ahmad, N. Hameed, M. S. Bashir, I. Ali, and Z. Nawaz, "Rainfall prediction using data mining techniques: A systematic literature review," Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 5, pp. 143–150, 2018, doi: 10.14569/IJACSA.2018.090518.
Y. Sakurai, Y. Matsubara, and C. Faloutsos, "Mining and Forecasting of Big Time-series Data: Part 1A - Similarity search , pattern discovery and summarization," Proc. 2015 ACM SIGMOD Int. Conf. Manag. Data - SIGMOD '15, pp. 919–922, 2015, [Online]. Available: http://dl.acm.org/citation.cfm?doid=2723372.2731081.
P. Singh, "Rainfall and financial forecasting using fuzzy time series and neural networks based model," Int. J. Mach. Learn. Cybern., vol. 9, no. 3, pp. 491–506, 2018, doi: 10.1007/s13042-016-0548-5.
S. Chatterjee, B. Datta, and N. Dey, "HYBRID NEURAL NETWORK BASED RAINFALL PREDICTION SUPPORTED BY," 2018, doi: 10.14311/NNW.2018.28.027.
A. Mosavi, P. Ozturk, and K. W. Chau, "Flood prediction using machine learning models: Literature review," Water (Switzerland), vol. 10, no. 11, pp. 1–40, 2018, doi: 10.3390/w10111536.
S. Bahri, N. Zoghlami, M. Abed, and J. M. R. S. Tavares, "BIG DATA for Healthcare: A Survey," IEEE Access, vol. 7, pp. 7397–7408, 2019, doi: 10.1109/ACCESS.2018.2889180.
A. Dikshit, R. Sarkar, B. Pradhan, S. Segoni, and A. M. Alamri, "Rainfall induced landslide studies in indian himalayan region: A critical review," Appl. Sci., vol. 10, no. 7, pp. 1–24, 2020, doi: 10.3390/app10072466.
M. Yu, C. Yang, and Y. Li, "Big data in natural disaster management: A review," Geosci., vol. 8, no. 5, 2018, doi: 10.3390/geosciences8050165.
S. K. Mohapatra, A. Upadhyay, and C. Gola, "Rainfall prediction based on 100 years of meteorological data," 2017 Int. Conf. Comput. Commun. Technol. Smart Nation, IC3TSN 2017, vol. 2017-Octob, pp. 162–166, 2018, doi: 10.1109/IC3TSN.2017.8284469.
N. Sultana and M. M. Hasan, "Forecasting Temperature in the Coastal Area of Bay of Bengal-An Application of Box-Jenkins Seasonal ARIMA Model," J. Civ. Environ. Res., vol. 7, no. 8, pp. 149–160, 2015.
W. Fan, C. Chong, G. Xiaoling, Y. Hua, and W. Juyun, "Prediction of Crop Yield Using Big Data," Proc. - 2015 8th Int. Symp. Comput. Intell. Des. Isc. 2015, vol. 1, pp. 255–260, 2016, doi: 10.1109/ISCID.2015.191.
J. Abbot and J. Marohasy, "Input selection and optimisation for monthly rainfall forecasting in queensland, australia, using artificial neural networks," Atmos. Res., vol. 138, pp. 166–178, 2014, doi: 10.1016/j.atmosres.2013.11.002.
P. Zhang, Y. Jia, L. Zhang, J. Gao, and H. Leung, "A deep belief network based precipitation forecast approach using multiple environmental factors," Intell. Data Anal., vol. 22, no. 4, pp. 843–866, 2018, doi: 10.3233/IDA-173427.
O. Manoj and J. P. Ananth, "MapReduce and Optimized Deep Network for Rainfall Prediction in Agriculture," Comput. J., vol. 63, no. 6, pp. 900–912, 2020, doi: 10.1093/comjnl/bxz164.
Y. K. Joshi, U. Chawla, and S. Shukla, "Rainfall prediction using data visualisation techniques," Proc. Conflu. 2020 - 10th Int. Conf. Cloud Comput. Data Sci. Eng., pp. 327–331, 2020, doi: 10.1109/Confluence47617.2020.9057928.
B. Thai, L. Minh, T. Le, K. T. Bui, and V. Minh, "Development of advanced arti fi cial intelligence models for daily rainfall prediction," Atmos. Res., vol. 237, no. August 2019, p. 104845, 2020, doi: 10.1016/j.atmosres.2020.104845.
P. Zhang, Y. Jia, J. Gao, W. Song, and H. Leung, "Short-Term Rainfall Forecasting Using Multi-Layer Perceptron," IEEE Trans. Big Data, vol. 6, no. 1, pp. 93–106, 2018, doi: 10.1109/tbdata.2018.2871151.
M. A. Rahman, L. Yunsheng, and N. Sultana, "Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman's rho tests and ARIMA model," Meteorol. Atmos. Phys., vol. 129, no. 4, pp. 409–424, 2017, doi: 10.1007/s00703-016-0479-4.

.