International Journal of Innovations in Engineering and Technology (IJIET)
http://dx.doi.org/10.21172/ijiet.151.07

Rain Fall Prediction Using Bigdata Analytics

Sumanta Kumar Singh[1], Ashis Kumar Mishra[2]
[1,2]*Department of Computer Science and Engineering,*
College of Engineering and Technology, Bhubaneswar, Orissa, India

Abstract- Now a day's big data is a new driver of the world economic and societal changes. As I know that Big data defined by extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions and it also known as 'predictive analytics'. Though there are so many definitions available on internet but this is only concern about data analytics using some methods and how its implement to extract some use full information. The world's data collection is reaching a tipping point for major technological changes that can bring new ways in decision making managing our health, cities, finance and education. While the data complexities are increasing including data's volume, variety, velocity and veracity, the real impact hinges on our ability to uncover the 'value' in the data through Big data Analytics technologies. Here using some new Analytics are potential breakthroughs include updated algorithms, methodologies, systems and applications in Big data Analytics that discover useful and hidden knowledge from the Big Data efficiently and effectively.
Keywords: Predictive Analysis, Variety, Velocity, Veracity, Methodologies, algorithms:

I. INTRODUCTION

Big data Analytics is a area of computer science and Artificial intelligence concerned with interactions between computers and human language along with historical record. It focuses on developing efficient algorithms to maintain rain fall prediction accurate and to make their collecting information accessible to computer correct [1].
The goal is to design and build new prediction result that will analyze, understand, and generate that rain fall possibility. If we go inside more details there is five 5'vs of big data value, velocity, veracity, volume, variety which are generated when source and destination both device communicating each other [2] .

II. PROPOSED MODEL

2.1 Propose model of Big Data Analytics –

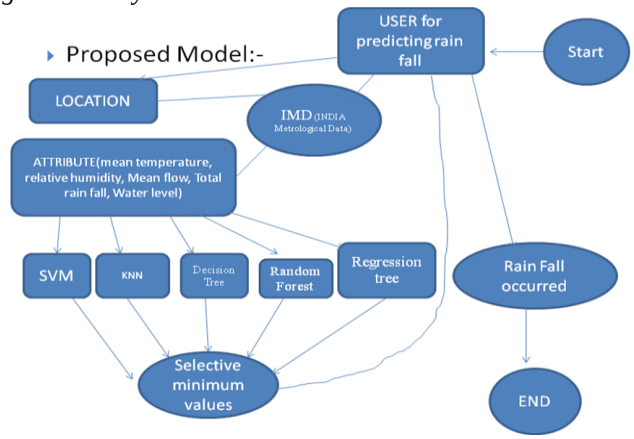

Fig 1 Propose model of Big Data Analytics

This model is describe about a single user can predict rainfall by taking some historical data of IMD and that will be the final result which is less value or minimum value among all these models. This picture depicts all the details about inside operation of rain fall prediction.

III. EXPERIMENT AND RESULT

The test set for this evaluation experiment randomly selected from internet. Weka 3.8 for calculating the optimum value and making the rain fall prediction of that particular area.
The proposed scheme is tested using all method like support vector machine, k-nearest neighbor, decision tree, Random forest, Regression tree. From the simulation of the experiment results, we can draw to the conclusion that this method is robust to many kinds of the methods.

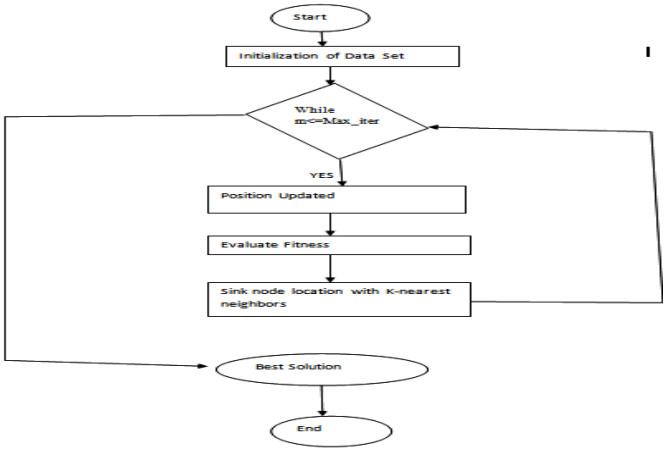

Fig 2 Flow chart Diagram for KNN

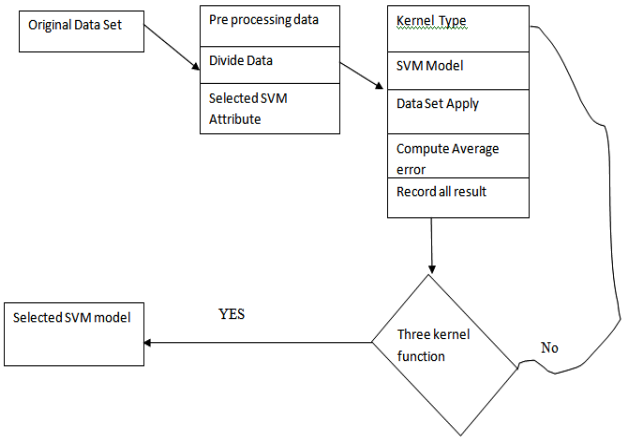

Fig 3 Flow Chart Diagram for Support Vector Machine

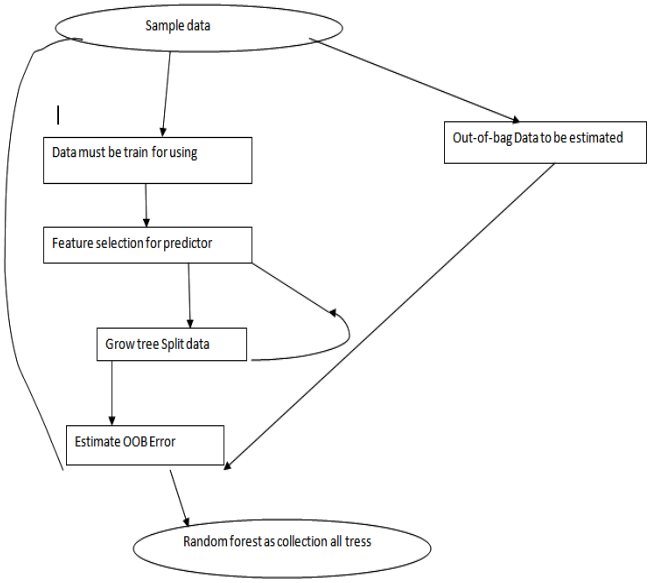

Fig 4 Flow Chart Diagram For Random Forest

Figure 4. (a) K-nearest neighbour (b) Support vector Machine (c) Random forest

Table -1 Experiment Result

	TP Rate	FP Rate	Precision	Recall
K-nearest neighbor	0.680	0.580	0.670	0.766
	0.325	0.210	0.480	0.380
Avg. KNN value	0.502	0.395	0.575	0.573
Support vector Machine	0.785	0.689	0.753	0.790
	0.444	0.250	0.585	0.450
Avg. SVM value	1.007	0.4695	0.669	0.62
Random forest	0.778	0.600	0.700	0.778
	0.400	0.222	0.500	0.400
Avg. RF value	0.643	0.465	0.629	0.643

IV.CONCLUSION

The performance of Table 1 is our proposed method of various data sets attribute, where our KNN is the better performance than other.

V. REFERENCE

[1] Foto N. Afrati, Anish Das Sarma, David Menestrina, Aditya Parameswaran, Jeffrey D. Ullman, “Fuzzy Joins Using MapReduce” 1084-4627/12 $26.00 © 2012 IEEE DOI 10.1109/ICDE.2012.66.
[2] A. Rajaraman and J. D. Ullman, Mining of massive Datasets, 20th ed. Cambridge: Cambridge University Press (Virtual Publishing), 2012
[3] C. L. Wu and K. W. Chau, "Prediction of rainfall time series using modular soft computing methods,"Engineering Applications of Artificial Intelligence, vol. 26, no. 3, pp. 997–1007, May. 2012.

[4] Panel, C. Change, and P. Ivonne, Climate change 2013: The physical science basis: Working group I contribution to the fifth assessment report of the intergovernmental panel on climate change, Intergovernmental Panel on Climate Change, Ed. Cambridge: Cambridge University Press, 2014.
[5] Abhishek, A. Kumar, R. Ranjan, and S. Kumar, "A rainfall prediction model using artificial neural network," in Control andSystem Graduate Research Colloquium (ICSGRC), 2012, IEEE, 2012. [Online]. Available: 10.1109/ICSGRC.2012.6287140. Accessed: Nov.6, 2016.
[6] R. Venkata Ramana, B. Krishna, S. R. Kumar, and N. G. Pandey, "Monthly rainfall prediction using Wavelet neural network analysis," Water Resources Management, vol. 27, no. 10, pp. 3697–3711, Jun. 2013.
[7] B. Wang et al., "Rethinking Indian monsoon rainfall prediction in the context of recent global warming," Nature Communications, vol. 6, p.7154, May 2015.
[8] Neelam Mishra, Hemant Kumar Soni, Sanjiv Sharma “Development and Analysis of Artificial Neural Network Models for Rainfall Prediction by Using Time-Series Data” I.J. Intelligent Systems and Applications, 2018, 1, 16-23 Published Online January 2018 in MECS , Published: 08 January 2018.
[9] D. Gupta and U. Ghose, "A comparative study of classification algorithms for forecasting rainfall," 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Sep. 2015.
[10] Joseph and R. T K, "Rainfall prediction using data mining techniques," International Journal of Computer Applications, vol. 83,no. 8, pp. 11–15, Dec. 2013.
[11] Nancy Kansal, Vineet Kansal PhD, An Efficient Data Mining Approach to Improve Students’ Employability Prediction , International Journal of Computer Applications (0975 – 8887) Volume 178 – No. 47, September 2019.
[12] P. Kumswat, Ki. Attakitmongcol and A. Striaew, "A New Approach for Optimization in Image Watermarking by Using Genetic Algorithms", IEEE Transactions on Signal Processing, Vol. 53, No. 12, pp. 4707-4719, December, 2005.
[13] H. Daren, L. Jifuen,H. Jiwu, and L. Hongmei, "A DWT-Based Image Watermarking Algorithm", in Proceedings of the IEEE International Conference on Multimedia and Expo, pp. 429-432, 2001.
[14] C. Hsu and J. Wu, "Multi-resolution Watermarking for Digital Images", IEEE Transactions on Circuits and Systems- II, Vol. 45, No. 8, pp. 1097-1101, August 1998.
[15] R. Mehul, "Discrete Wavelet Transform Based Multiple Watermarking Scheme", in Proceedings of the 2003 IEEE TENCON, pp. 935-938, 2003.
[16] J. Joseph and R. T K, "Rainfall prediction using data mining techniques," International Journal of Computer Applications, vol. 83, no. 8, pp. 11–15, Dec. 2013.