Big Data in Weather Forecasting

Vaishnavi P N
Student, Department of ECE
Coorg Insitute of Technology,
Ponnampet, Karnataka – 571216

Rizwan M A
Student, Department of ECE
Coorg Insitute of Technology,
Ponnampet, Karnataka – 571216

Dechakka M P
#Assistant Professor,
Department of ECE
Coorg Insitute of Technology,
Ponnampet, Karnataka – 571216

***Abstract*—The interpretation of a massive amount of data plays a vital role, likewise with the climate and weather. The data must be accumulated, assessed and filed. A lot of human activities are weather-dependent like agriculture, power production and pollution. A sustainable solution to overcome the drawbacks of the existing weather monitoring system that takes global warming parameters into account is needed. In this paper, we design an elegant solution using map reducing algorithm and Hadoop database system that lessens the burden of computing weather forecasts with a bulk amount of data and elevates the life expectancy in disaster-prone areas.**

Keywords—BigData;Global warming; Pre-Processing; Weather Monitoring.

I. INTRODUCTION

Big storms might yield despair, but the cataclysmic weather can be downright devastating. Economic losses from hurricanes, floods, typhoon and other natural calamities top 200 billion dollars annually with the impact felt everywhere [1]. Natural disasters affect more than a third of global G.D.P [2]. The scariest element about climate change for industries is not the heightened severity of storms, but an increasingly unpredictable world where disasters catch everyone by surprise. So the last couple of decades, politicians continued to decide if climate change was real or not when industries and Tech worlds turned to a new cloud for weather prediction. This revolution is the BigData revolution with its massive storage farms and supercomputers, which has formed a rapidly growing world of hyper real-time forecasting.

Approximately 40 Exabyte's of data is generated every month by a single user [3]. Imagine the same information being multiplied by five billion users a month. We are walking into an era where human beings can worry about how to gather more data and process it, instead of how to reduce a large amount of stored data and archives. Users are no longer intimidated by volumes and seek to gain further insight into businesses, government and society. Big Data is a new approach to experiencing the world and decision making. Big Data turns unstructured data into actionable data to gain insights over the data traditional systems cannot store, process, and analyze.

Now, the question is, what role does BigData play in weather forecasting? BigData has turned meteorologists into magicians. The nature of the atmosphere makes forecasts inaccurate and highly competent to a vast extent. Forecasting weather requires analysis of a large set of data at an instantaneous rate. The size and complexity of weather patterns make it very difficult to predict the weather a few days in advance [7]. Researchers have to refine their instruments and techniques to analyze and make more accurate weather forecasts. But in the presence of BigData, researchers use data describing the physics of the atmosphere to create a mathematical model of how the weather is likely to unfold.

Fig 1.1 6 V's of Big Data

Predicting the climate with preceding data and existing weather parameters for predictive interpretation in a data warehouse requires high computing and agile query languages using spatial data and high cache memory. BigData algorithms help program logic using supercomputers and create a presumptive plan to restructure the nature of the operation. BigData ensures accuracy in a circumstance where the approximated original data is slightly different while the predictions are based on the transformed data. It establishes models with the help of never-ending packets of data from various points across the surface of the earth

II. APPROACHES OF WEATHER FORECASTING

The climate is one of the most powerful and important forces on the planet. A network of devices scan accumulate data from all over the world and depict the climatic changes. Another vital ingredient is space, the satellites helps us observe details and store it before it transforms into the atmosphere. In this section, we provide a depiction of the diverse approaches used so far.

PERSISTENCE and TRENDS:Persistence forecasting is a simple and easy means for weather forecasting. It processes the present weather conditions to predict the upcoming changes in the weather. It is a suitable approach in regions close to the equator like Ecuador where the climate remains ceaseless most of the year. The mechanism is useful in both the brief term and long-term forecast scenarios. The method is called persistence and trends because the meteorologists predict the current weather acknowledging past trends. The same can be performed by every individual irrespective of their skills.

Let us consider an example when the weather is rainy with mild thunderstorms. In such an atmosphere, we can equate the weather for the upcoming week to be the same. Yet, the persistence method of forecasting may not be the most accurate. Let us consider another example when the weather is humid with dust storms. In such an atmosphere, it may be difficult to predict the weather patterns for the upcoming week. The method works effectively in areas with consistent weather conditions. If the weather transforms steadily, the method habitually breaks down and becomes inaccurate. However, it plays a valuable role in predicting long-range weather or climate forecasts. Thus, persistence is eminent for seasonal weather.

SYNOPTIC FORECASTING:The synoptic methodology is one of the oldest means to predict the weather. Though a traditional method, it involved various types of analysis using several empirical rules. The rules used the physical, climatological and kinematic elemental nature to conduct analysis. In the 1860s, men who tried to predict the changes in weather faced suspicion and ridicule. Back then, there were just individual forecasters instead of meteorologists. The absence of scientific tools and mechanisms, the forecasters used weather balloons to estimate the rate, temperature, direction and winds in the atmosphere. With the development of synoptic rules, convenience improved while the forecasters

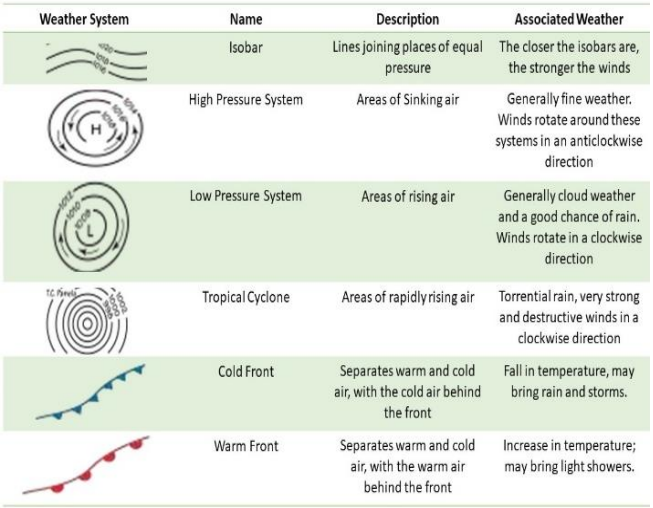

Fig 2.1 Types of Weather patterns in Synoptic map

predicted shorter foretells. With the synoptic approach, the forecast stated the occasional states of weather with discrete atmospheric information and varying heights of the atmosphere. In a synoptic approach, a forecaster attempts to predict changes in the state of atmosphere from the opening state using theoretical knowledge and expertise. However, the foretells demanded precise information on the existing conditions of the atmosphere and discrete heights that were similar to past occurrences.

The **synoptic weather patterns** in the above figure reveal weather patterns for all states of weather. The white lines are termed isobars. The weather patterns are drawn out by examination over different locations, all transcribed at the same moment in time. The main advantage of synoptic is the capability to interpret complex atmospheric equations drawn over synoptic charts [5]. Yet, the prediction is subjective and can vary among meteorologists.

STATISTICAL FORECASTING: The statistical approach uses the previous event to predict the changes in weather shortly. The statistical approach is an analogue method that scrutinizes the accumulated data to perceive events that mimic the current state of the weather. The mechanism is intricate and considered in rare scenarios because it is bias to statistical probability and favorable if affirmative to the past. The results are never precise and if found to be accurate, varies between 40 to 50% [8]. The smallest discrepancies in data can alter the outcome resulting in a crisis for regions prone to natural calamities and compile may not be a leading means for weather prediction. Another aspect is different businesses can process varying results which can be downright devastating and cause the downfall of organizations. These reasons make statistical forecasting inferior to other methods and unfavorable for the weather forecast.

NUMERICAL WEATHER PREDICTION: Numerical Weather Prediction (NWP) is a model-based approach. The medium of prediction complete relies on the study of the atmosphere and oceans [8]. The computers, massive supercomputers analyze the accumulated and ongoing data produced to form forecasting models based on the varying conditions of the atmosphere. The obtained software models are prior reviewed by meteorologists before forecast. The inaccuracy in equations must be observed to avoid errors in measurement and distribution of software models for forecasting.

However, the NWP consumes a vast amount of time to process the forecast results. It may not be considered as a prime disadvantage because previously obtained results can be used until the latest results are available. A smaller NWP processing can be performed on a small-scale of data. Thus, resulting in a significant demerit because as human beings can instantaneously predict the weather by observing the climatic changes, the NWP has to execute to process new results which are unfortunately time-consuming.

III. IMPLEMENTATION

The weather forecasting using Big Data is implemented using the MapReduce algorithm. The algorithm performs two vital tasks, called Map and Reduce. The Mapper class implements the map while the Reducer class implements reduce [4]. The MapReduce makes it easier and constructively reliable to scaledata processing over multiple data nodes. It analyzes the given data and predicts the forecast for the day. Hence, in this design, it can predict the weather, thunderstorms, snowfall, a beautiful day for a stroll and other significant details that most individuals assess at the start of a day.

The design comprises the following phases:
Phase 1 - Accumulation
Phase 2 - Storage
Phase 3 - Map Reduce
Phase 4 - Illustration

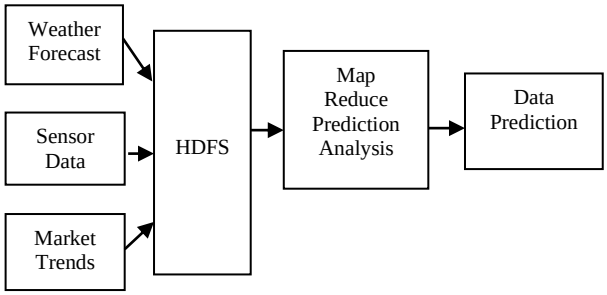

Fig 3.1 System Composition

Accumulation:In the first phase, the design accumulates data from various sources like weather forecasts, sensor data, market trends and social media. In this design, the data acquired is from various online gatherings. The database management system stores the assembled data. The data remains in the database system until transferred to a storehouse. On transfer, the data subconsciously deletes.

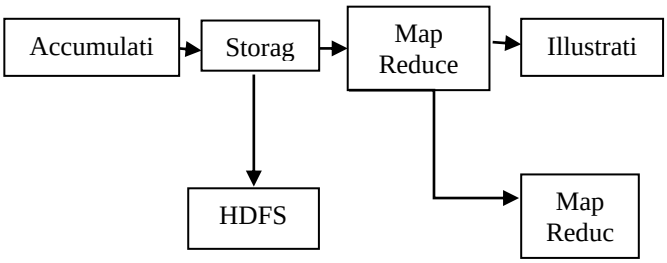

Fig 3.2 Flow Diagram

Storage:The accumulated data is transferred to the primary data storage system. This system is called the Hadoop Distributed File System (HDFS). The HDFS performs the processing using the MapReduce programming algorithm. The HDFS operates as a Master-Slave system [6]. The system framework consists of a single name node synchronized with a backup node. The backup node performs the process of compacting and editing files stored in the operating system into new files. The data nodes, however, can be multiple. The data nodes break down and save the transmitted data in data blocks. The system can multitask while completing Read/Write operations. The HDFS is one of the main elements of Hadoop followed by Map Reduce.HDFS is scalable and performs parallel processing to avoid failures.

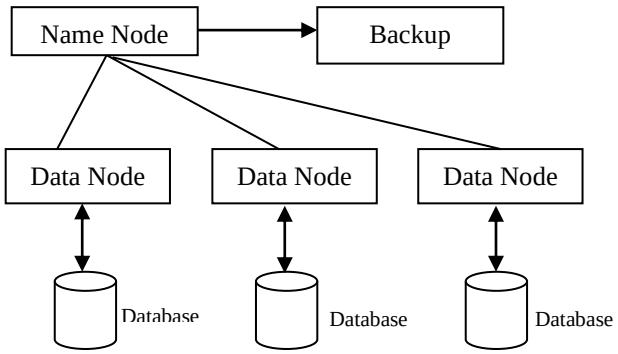

Fig 3.3 Architecture of HDFS

Map Reduce: This step deduces the stored data using the map-reduce framework. As illustrated in figure 3.4, the first step of the algorithm collects different weather predictions. The next step comprises splitting of input into multiple blocks of data. Once split, the data undergoes sorting and matching data of one set against the data of another set. This

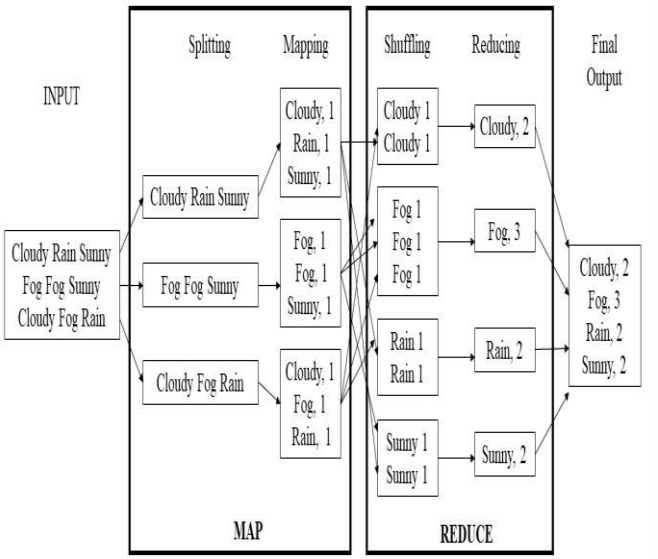

Fig 3.4 Working of MapReduce algorithm

method is called mapping. Then, the data is rearranged in the process of shuffling. Once shuffled, the data is reduced to individual blocks. The final forecast is the data with the highest occurrences. From figure 3.4, FOG is the most common event and can be deduced as the final forecast. The entire process is an efficient way of sorting and rearranging a cluster of data into individual sets. The MapReduce uses various mathematical algorithms to perform the named operations**.**

Illustration: The final phase represents the filtered data. The filtering helps focus on the most accurate and relevant weather predictions. The data can be illustrated pictorially using line graphs, bar graphs, histograms or pie. As concluded in the above Figure 3.4, the varying distribution of weather is illustrated in Figure 3.5.

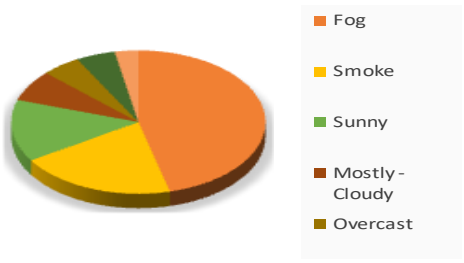

Fig 3.5 Pictorial Illustration of attained data

IV. CONCLUSIONS

Attaining a precise prediction is favorable. The weather is getting messier and storms are getting bigger, and they are showing up at unlikely places. While the clouds get unrulier, it is imperative to keep the weather observed. The map reducing algorithm renders forecasts at an unprecedented resolution. With algorithmic updates and improved initialization by incorporating high-resolution datasets, map reducing brings weather stories to life. The design helps citizens living in areas prone to environmental disasters and assist farmers in adapting to the climate.

REFERENCES

[1] Wiston, Modise & Mphale, Kgakgamatso. (2018). Weather Forecasting: From the Early Weather Wizards to Modern-day Weather Predictions. Journal of Climatology & Weather Forecasting. 06. 10.4172/2332-2594.1000229.

[2] Snook, Nathan & Kong, Fanyou & Brewster, Keith & Xue, Ming & Thomas, Kevin & Supinie, Timothy & Perfater, Sarah & Albright, Benjamin. (2019). Evaluation of Convection-Permitting Precipitation Forecast Products Using WRF, NMMB, and FV3 for the 2016–17 NOAA Hydrometeorology Testbed Flash Flood and Intense Rainfall Experiments. Weather and Forecasting. 34. 10.1175/WAF-D-18-0155.1

[3] X. Wu, X. Zhu, G. Wu and W. Ding, "Data mining with big data," in IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 1, pp. 97-107, Jan. 2014, doi: 10.1109/TKDE.2013.109.

[4] Adam, Khalid & Majid, Mazlina & Fakherldin, Mohammed & Mohamad Zain, Jasni. (2017). A Big Data Prediction Framework for Weather Forecast Using MapReduce Algorithm. Advanced Science Letters. 23. 11138-11143. 10.1166/asl.2017.10237.

[5] Big Data Quality Framework: Pre-Processing Data in Weather Monitoring Application Ashish Juneja, Nripendra Narayan Das Deaprtment of Computer Science & EngineeringFaculty of Engineering & Technology

[6] H. Jain and R. Jain, "Big data in weather forecasting: Applications and challenges," 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, 2017, pp. 138-142, doi: 10.1109/ICBDACI.2017.8070824.

[7] T H, Udayashankara & BM, Sadashiva & M., Madhukar. (2016). Impact of Climate Change on Rainfall Pattern and Reservoir Level. Journal of Water Resource Engineering and Management. 3. 10-14.

[8] S. Kothapalli and S. G. Totad, "A real-time weather forecasting and analysis," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, 2017, pp. 1567-1570, doi: 10.1109/ICPCSI.2017.8391974.

[9] S. Madan, P. Kumar, S. Rawat and T. Choudhury, "Analysis of Weather Prediction using Machine Learning & Big Data," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), Paris, 2018, pp. 259-264, doi: 10.1109/ICACCE.2018.8441679.