

# Weather Predictions using Big Data

Abey Thomas <sup>1</sup>, Naganath Bhat <sup>2</sup>

P.G. Student, Department of MCA, Acharya Institute of Technology, Soldevanahalli, Bengaluru, India<sup>1</sup>

P.G. Student, Department of MCA, Acharya Institute of Technology, Soldevanahalli, Bengaluru, India<sup>2</sup>

**ABSTRACT:** The usage of Big Data to predict the future, in consideration of weather is explained. So that, we can use the analysed data to prevent flight delays. Big Data is actually extremely large data sets that may be analysed computationally to predict the patters and trends with the associations. Here, we will be collecting the large data set from the historical weather reports and using Big Data analytics the future weather is predicted faster. The outcome is avoiding a big SOCIAL ISSUE owing to the flight delays and flight cancellations disrupting the everyday life of many around the globe. The system forecasts the weather with the accuracy of 92%.

**KEYWORDS:** Big Data, Weather Forecast, Airline.

## I. INTRODUCTION

The weather is an uncontrollable factor that can go too ease and also too worse so fast. The rapid weather changes can completely disrupt the humans in their livelihood. On considering the air transportations, this is a major factor that can destroy the schedule in no time. In fact, nearly 70% of flight delays and cancellations are caused as a result of severe weather.

In real world, the major weather events like snowstorms and severe weathers are skyrocketing the number of flight delays and cancellations today.

Our computational system uses Big Data as an input to eradicate this problem. Big Data for weather forecast is actually extremely large data sets holding the current and previous weather reports that are used for analysing computationally and to predict the patters and trends with the associations of each data. This system fetches data from various international weather report databases and also the historical data on flight delays and cancellations that are publically available in Internet. The collected data is then categorized on hourly basis and maintained for analysis.

At present, satellite prediction is used for weather forecast, which can give approximate data for a limited range of period. Using Big Data, this limitation can be removed and exact forecast is made available for long range of time. The recorded data is used to design a statistical model, which predicts the severity of weather by making use of time-delayed feed forward neural network. We will be using sliding window algorithm for the future weather prediction. The suitable windows from the past year and also the data from one week prior to the date for which the weather has to be predicted is used for the analysis and prediction.

The reason for using sliding window algorithm is that, the previous year data alone will not provide the accurate prediction. So we will be using the one week prior data with the ongoing week's data for the actual prediction. The sliding window algorithm is very efficient in calculating the smallest variation in data from the previous year database.

This system can help airline companies and even militaries to take precautions on knowing the future. Thus they can anticipate the flight delays and preventing them.

13<sup>th</sup>-14<sup>th</sup> March 2018

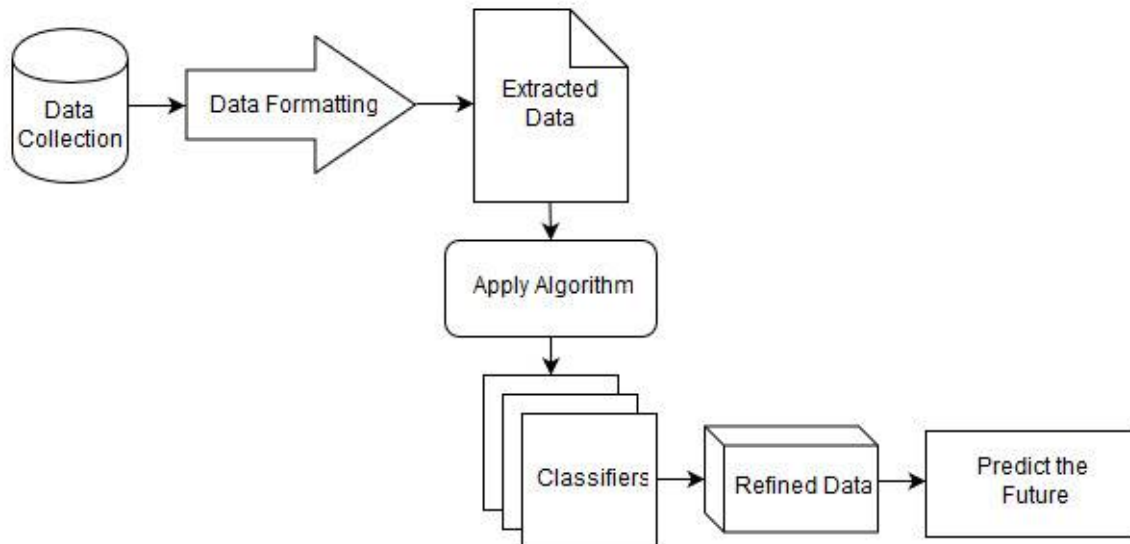
Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India

## II. PROBLEM STATEMENT

Accurate prediction of weather conditions is a difficult task due to the dynamic nature of atmosphere. The variables and their range always vary from place to place. The weather report data from the past 10 years will be of huge size consuming more than petabyte of data. Hence the cost of applying the big data will be unpredictable. Possibility is there for misuse and security loss. Accuracy cannot be always sustained as the climatic changes are not man made. Also unforeseen external factors may alter the weather condition.

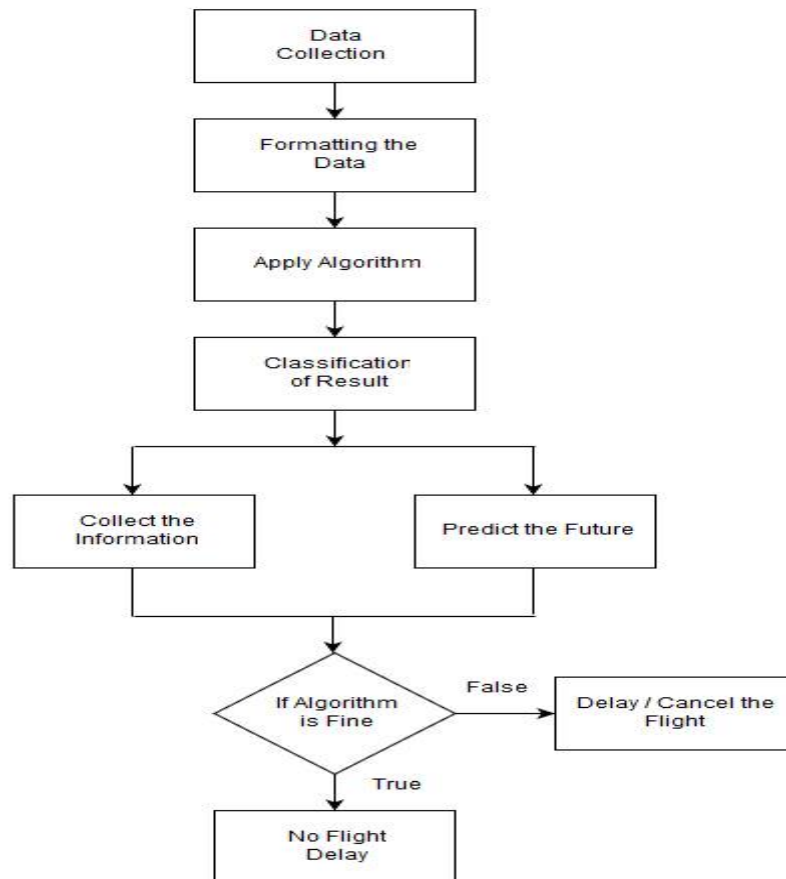
## III. DATA FLOW



13<sup>th</sup>-14<sup>th</sup> March 2018

Organized by

Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India



**Data Flow Diagram**

#### IV. IMPLEMENTATION DETAILS

The Climate Corporation generates daily weather report in excess of 50 terabytes of data. To store and manage this huge data, we need Big Data. Through Big Data, the weather reports of even the previous decades can be provided to companies and also for future prediction. Only after analysing a big set of data, nearly accurate predictions can be made. Thus we make use of Big Data.

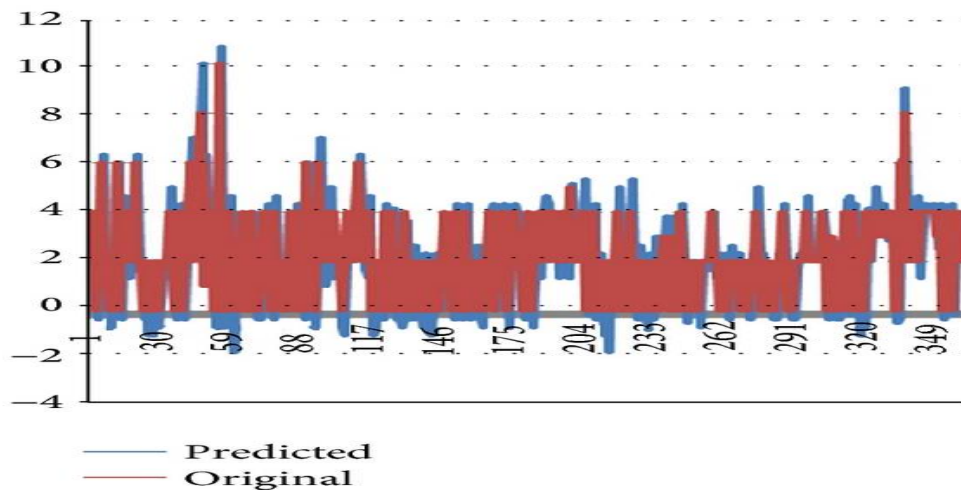
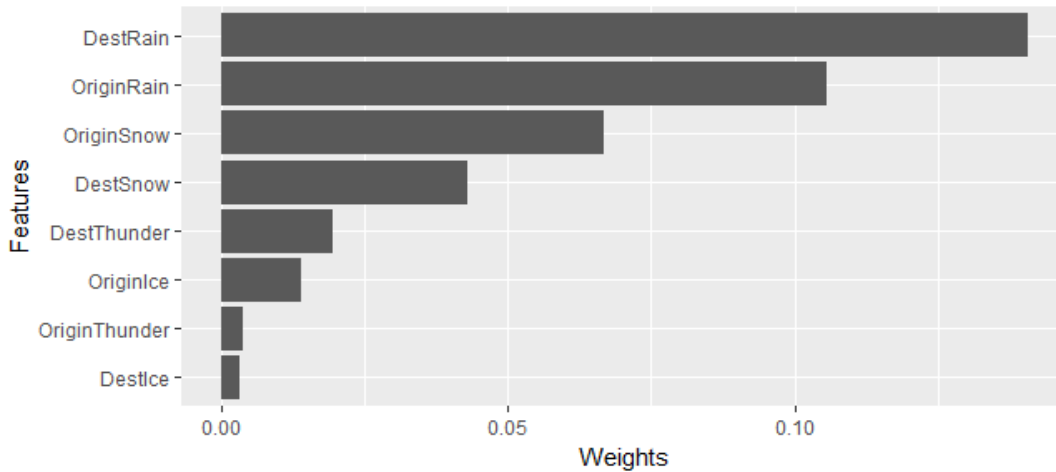
The weather report collected on daily basis is filtered and data mining is carried out using the Oracle Data Mining Tool. SQL provides a better data mining facility by digging data out of database tables, views, and schemas.

The Sliding Window algorithm is used for analysis and prediction. Through this algorithm, to predict the weather on a particular day, we will be fetching the weather report on the same period of past few years and also the variations in climate seen for that particular year. Multiple sliding windows are generated from the past year’s data. Those sliding windows that matches the current year’s variations in climate are fetched. These sliding windows are used for predicting the weather.

The outcome of the Algorithm after processing the data is classified on the label of required, may be required and essential. The vital information from the Algorithm’s output will be store in the database for future purpose and prediction.

The prediction is analysed and we come to a conclusion whether the flight can take off on that particular day or not.

## V. RESULTS



## REFERENCES

- [1] <https://www.hindawi.com/journals/isrn/2013/156540/>
- [2] <https://www.slideshare.net/pushkarzagade95/big-data-to-avoid-weather-related-flight-delays-ppt>
- [3] <http://phys.org/news/2014-12-big-weather-related-flight.html>
- [4] <https://github.com/mctrinkle/OnTimeFlights>
- [5] K. Izumi, K. Goto and M. Matsui, "Analysis of Financial Markets' Fluctuation by Textual Information," The 23rd Annual Conference of the Japanese Society for Artificial Intelligence, 2009.
- [6] S. Murayama, K. Okuhara and H. Ishii, "Innovation in Manufacturing Premise by New Finding Obtained from Accident Relapse Prevention Report," Proceedings of The 13th Asia Pacific Management Conference, Melbourne, Australia, vol. 13, pp. 1124-1129, 2007.

**International Journal of Innovative Research in Science, Engineering and Technology**

*An ISO 3297: 2007 Certified Organization*

*Volume 7, Special Issue 6, May 2018*

**2<sup>nd</sup> National Conference on "RECENT TRENDS IN COMPUTER SCIENCE & INFORMATION TECHNOLOGY (RTCSIT'18)"**

**13<sup>th</sup>-14<sup>th</sup> March 2018**

**Organized by**

**Departments of Computer Science & Information Science Engineering, Sri Krishna Institute of Technology, Bangalore, India**

- [7] Kyung-Joong. Kim, Myung-Chul. Jung and Sung-Bae. Cho, "KeyGraph-based chance discovery for mobil contents management system," International Journal of Knowledge and Intelligent Engineering System, Vol.11, pp. 313-320, 2007.
- [8] Álvarez-Martínez, J., Gómez-Villar, A., and Lasanta, T. 2016. The use of goats grazing to restore pastures invaded by shrubs and avoid desertification: A preliminary case study in the Spanish Cantabria Mountains. Land Degradation and Development 27(1), 3-13. doi:10.1002/ldr.2230.
- [9] Brevik, E.C., Cerdà, A., Mataix-Solera, J., Pereg, L., Quinton, J.N., Six, J., and Van Oost, K. 2015. The interdisciplinary nature of SOIL, SOIL, 1, 117-129, doi:10.5194/soil-1-117-2015.
- [10] de Araújo, A.S.F., Eisenhauer, N., Nunes, L.A.P.L., Leite, L.F.C. and Cesarz, S. 2015. Soil Surface-Active Fauna in Degraded and Restored Lands of Northeast Brazil. Land Degradation and Development 26(1), 1-8. doi:10.1002/ldr.2247. Ibáñez, J.J., Pérez-Gó.