IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

TO ANALYZE AND INFER A COMPLETE SOLUTION FOR WEATHER MANAGEMENT USING BIG DATA ANALYSIS

P Venu Gopal , V Krishna Teja , D Jaswanth
IV YEAR,CSE DEPARTMENT, SRM CHENNAI

SUPERVISOR: Mrs.Antony Vigil

Abstract- Weather plays a significant part in each and every phase of human life. It has a straight influence on all the sectors of this society. Weather predictions have a great prominence in the agricultural science, vacation industry and administration agencies. Prior knowledge of meteorological conditions will be of great support for us to make ourselves prepared for any kind of adverse climatic changes. Several factors like pressure, humidity, wind speed, temperature etc, have a crucial role while analysing the climatic conditions. The data of weather obtained from several sensors for monitoring different weather parameters always gets produced at a large scale, this data gets growing. The improvement of our capacity to predict the climate and weather has been the principal objective of the scientists and practitioners throughout the global community. In support of critical decision-making, the increase in the demand for weather forecasting has grown up very quickly from the past few decades

Keywords- Classification, Naïve Bayes, Prediction, Weather Forecasting

I. INTRODUCTION

Over the last few years, big data has become one of the buzzwords in the IT world. It was initially developed by companies that needed to manage rapidly growing data, such as data generated from network data, research or business simulation, or other data. The structure of some of these companies based on indexing and big data use is important. Additionally, many cities are becoming smart. Therefore, many sensor devices used in smart cities can be used to measure air quality. This allowed weather services to collect and analyze large amounts of data such as temperature and humidity. These different gauges measure temperature, humidity, and other factors to predict rainfall and other things: refugee camps, food shortages, medical facilities, etc. However, since the information cannot be trusted, many people may obtain false information and take risks. Many variables, such as temperature and humidity, can be analyzed without scalability issues. The speed of data processing needs to increase rapidly to keep everything running efficiently and effectively.

1.1. BPN and Hopfield Network

In this study, advanced modeling was performed with back propagation neural (BPN) networks. The results obtained from the BPN model are stored in the Hopfield network. In BPN, the information and imaging process has 3 neurons, the hidden process has 5 neurons, and the Hopfield network performs image processing with the help of information [5]. The system must be heated or the air velocity or humidity must be sufficient to create equilibrium as the ultimate goal. This process will continue and with each iteration the bias and estimated weight should be adjusted until they converge.

1.2. The Models: RNN, CRBM and CN models

The aim of this project is to explore the potential of deep learning in climate science. In the last decade, research on deep integration [14] and electronic models [15] has been conducted and has become the basis of design-based deep learning for deep structures. Three prediction models will be examined in this review, specifically: (i) Relational Neural Networks (RNN), (ii) Restricted boltzmann Machines (CRBM) and (iii) Convolution Networks (CN) [8]. Each of these samples will be prepared and tested using the specified weather conditions. For each model, parameter learning algorithms such as CRBM and gradient descent for CN are used to obtain the testing error under preorder, similar to neural networks, and compare it with the model's prediction time.

1.1 ANN and Decision Tree

Artificial neural network (ANN) and decision tree (DT) are used to analyze meteorological data to create classification rules for the use of meteorological data weather. The neuron model has three main features; these include (i) the arrangement of synapses and connections, each determined by its own weight/strength (ii) the aggregator used to: info signals, weighted by particular neurons neural connections (iii) Activation function used to limit the amplitude of neuronal output [6] MLP network was prepared by introducing the backward learning method. Predictions are made using decision trees.

II. METHODOLOGY

In this article the system predicts future weather based on current weather data. Apply mining techniques(e.g. Chi_square test and Naïve base statistics) to the dataset to extract important information from the dataset. System MethodologyScreen

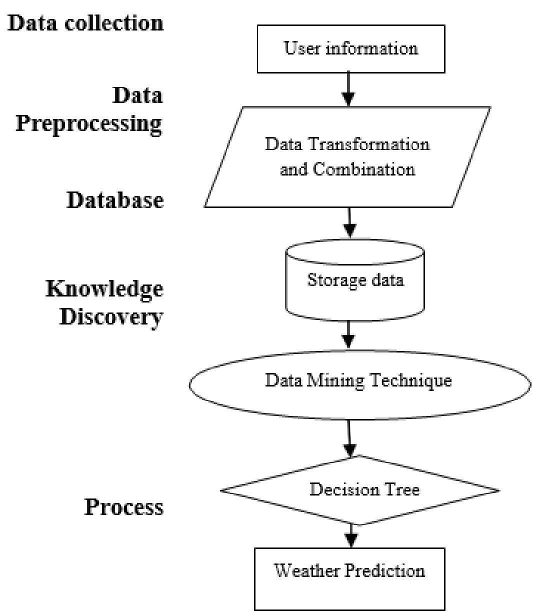


Fig-1

2.1. Data Collection and Preprocessing

The first stage of data search is data collection and preliminary preparation. The important stage is preliminary data because only valid data can create correct data. The information used in this project is collected by users. Although data has many characteristics, initial data only considers important data and ignores others. The data is then converted into a format suitable for searching the data . Four features are used to describe the weather forecast. They are indicated by the following words:

2.2. Data Mining Technique

The converted data is stored in the data stored by the user. Therefore previous storage is not used. Once real-time data is collected, data mining technology is used to predict weather data. The expression library can be consulted to manipulate expressions and expressions.

2.3. Decision Tree

Decision trees created from training data help make predictions. Creating the decision tree is done by selecting the best available features to split set of samples in most effective manner. The decision tree for this proposed system is figured below in fig 2:

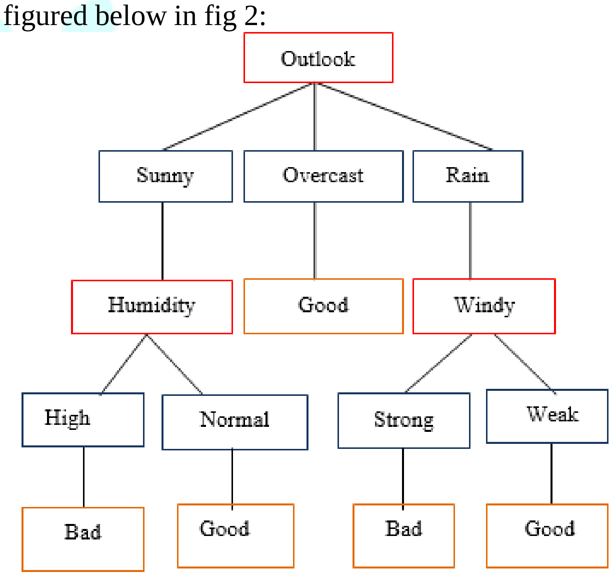


Fig-2

2.4. Numerical Weather Prediction (NWP)

Use the power of computerd for prediction. Complex computers, also known as forecast models, run on super computers and provide predictions of many climate variables such as temperature, pressure, wind and precipitation. Experts study how factors in computer forecasts interact to create weather days. The flaw in the NWP approach is the equation the model uses to simulate uncertain weather conditions. If the initial state is unknown, the computers prediction of how the initial state will change will be completely wrong. In this technology, the behavior of air is represented by equations based on physical laws that control air movement, air pressure, and other information. This method has been found to be suitable for moderate prediction. Combining weather data with modern weather equipment and transmitting the data through telemetry communication and appropriate agricultural support will mean that marine plants can plan to farm in other ways according to the weather.

III.PROPOSED MODELS

3.1. Sliding window algorithm

This project focuses on predicting one-day weather. For this purpose, in addition to the weather forecasts of the last seven days, two-week weather forecasts of the last few years are also taken into account. Suppose we need to predict the weather on March 30, 2022, then we will consider the weather from March 23, 2022 to March 29, 2022, and the weather from March 23 to March 29 last year. Then calculate the daily changes over the year. This change is also based on two weeks of data from the previous year, four main weather conditions will be considered in this study Maximum temperature, minimum temperature, humidity Maximum temperature, minimum temperature, humidity, and preceptation. Therefore, the size of the change for the current year will be represented by a 7x4 matrix. Likewise, the size of the previous year's matrix will be 14 x 4. First step right here, place the 14 x x 4 matrix on the window sash. Therefore, 8 sliding windows can be made, each measuring 7 x 4.

3.2. Analysis module

Objective predictive values were analyzed as independent variables. It is used only to find relationships between variables and make predictions. Regression models differ depending on the type of relationship between the dependent variables and the independent variables considered and the number of independent variables used. Create an AQI or PM2.5 measurement model that predicts air/temperature based on various parameters such as humidity, ppm and air quality.

Dataset 1

This data includes features such as air temperature, humidity, and AQI (PM2.5) as target variables. This file is very short, only 150 entries.

Dataset 2

This data contains approximately 24 parameters, including Ppm, humidity, and air temperature as target variables.

It was used in the design of BERT. Transformers are now used regularly and our version is almost identical to the original.

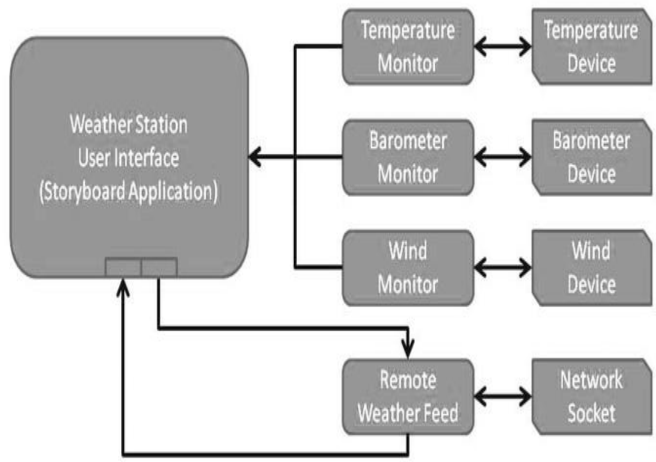


Fig-3

3.3. Naïve bayes Algorithm

Naïve bayes algorithm is a classification system based on bayes theorem. Naïve Bayes is easy to create and is useful for large data sets. We can see the future probability using the negative Bayes equation. The equation is as follows :

$$P(c|x)=P(x|c)*p(c)/p(x)$$

Where (cx) is the future probability of category (c, target), P(c) is the prior probability of this category, P(xc) is the probability of the forecaster by category, and P(x) is the forecaster. Before this happens.

Conditions for weather forecasting in our project are as follows:

Category:

C1: Weather Forecast = 'Good',

C2: Cloudy Weather Wind = 'Not Good'.

To find the largest category Ci, calculate P (X Ci) * P (Ci):

P(air = good x) P(air = good) [P(O =s air = good)) . P(T = c air = good). P(H = h air = good) · P(W =t air = good)]

P(air = bad x) P(air = bad). [P(O = s air = bad) . P(T = c air= bad). P(H = s air=bad) . P(W = t air= bad]

IF P (Weather = Good X) <P (Weather = Bad X) , so classify

IV. RESULTS AND DISCUSSION

In this project, a better solution for better weather management has been analyzed and inferred, at its best accuracy rate 71.9%. Efficient weather analyzing and forecasting has been done using processes like Sliding algorithm and regression analysis by obtaining datasets from meteorological departments.

After using the following code for importing the data module for weather forcasting the data is ready to be compiled in all the models.

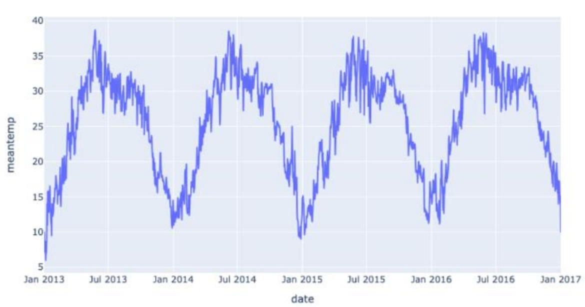
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px

data = pd.read_csv("DailyDelhiClimateTrain.csv")
print(data.head())
```

The dataset results displayed as:

	date	meantemp	humidity	wind_speed	meanpressure
0	2013-01-01	10.000000	84.500000	0.000000	1015.666667
1	2013-01-02	7.400000	92.000000	2.980000	1017.800000
2	2013-01-03	7.166667	87.000000	4.633333	1018.666667
3	2013-01-04	8.666667	71.333333	1.233333	1017.166667
4	2013-01-05	6.000000	86.833333	3.700000	1016.500000

The observations for the accuracy can be inferred from graph of each model's accuracy:



4 The mean temperature of delhi climate
Fig-4

PROPHET MODEL:

```
from prophet import Prophet
from prophet.plot import plot_plotly, plot_components_plotly
model = Prophet()
model.fit(forecast_data)
forecasts = model.make_future_dataframe(periods=365)
predictions = model.predict(forecasts)
plot_plotly(model, predictions)
```

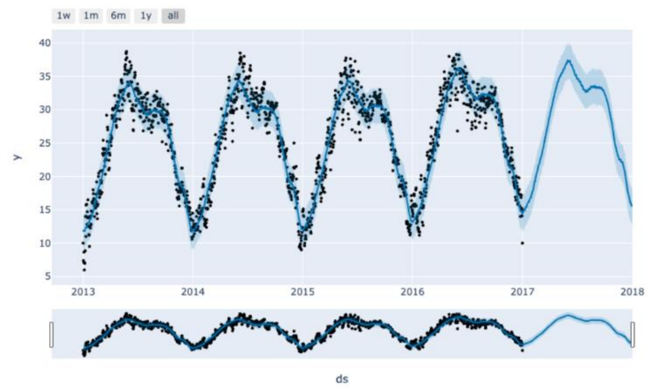


Fig-5

CONCLUSION

Based on the traditional weather management systems, an efficient and advanced weather management system which analyzes and infers a better solution has been proposed in this paper. However, when larger datasets are used, the predictions have slight variations in the accuracy. The next research goal is mainly to find an even more efficient weather prediction and forecasting system for future purposes, and strive to achieve a better performance with the larger datasets by using a smaller structure as used.

REFERENCES

[1] Mary Nsabagwaa, Maximus Byamukamab, Emmanuel Kondelaa, "Towards a robust and affordable Automatic Weather Station ", journal homepage: www.elsevier.com/locate/deveng.

[2] Ravi Kishore Kodali and Snehashish Mandal "IoT Based Weather Station" 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT) 978-1-5090- 5240-0/16/$31.00, IEEE, (2016)

[3] Ravi Kishore Kodali and Archana Sahu "An IoT based Weather Information Prototype Using WeMos" 2016 2nd International Conference on Contemporary Computing and Informatics (ic3i), 978-1-5090-5256- 1/16/$31.00, IEEE, (2016)

[4] Zi-Qi Huang, Ying-Chih Chen and Chih-Yu Wen, "Real-Time Weather Monitoring and Prediction Using City Buses and Machine Learning", Vols. 3 to 21 Published 10 September (2020)

[5] Kavya Ladi, A V S N Manoj, G V N Deepak, "IOT Based Weather Reporting System

to Find Dynamic Climatic Parameters", International Conference on Energy,
Communication, Data Analytics and Soft Computing (ICECDS-2017)
[6] P. Susmitha, G. Sowmyabala "Design and Implementation of Weather Monitoring and
Controlling System", International Journal of Computer Applications (0975 – 8887)
Volume 97– No.3, (July 2014)
[7] Tanmay Parashar1, Shobhit Gahlot2, Akash Godbole3, Y.B. Thakare4 "Weather
Monitoring System Using Wi-Fi", (IJSR) ISSN (Online): 2319-7064 Index Copernicus
Value (2015): 78.96, 2015
[8] Nutant Sabharwal, Rajesh Kumar, Abhishek Thakur, Jitender Sharma "A LOW-COST
ZIGBEE BASED AUTOMATIC WIRELESS WEATHER STATION WITH GUI AND
WEB HOSTING FACILITY" e-ISSN: 1694-2310 | p-ISSN: 1694-2426, Vol. 1, Spl. Issue
2 (May 2014)
[9] M. Prasanna, M. Iyapparaja, M. Vinothkumar, B Ramamurthy, S.S. Manivannan," An
Intelligent Weather Monitoring System using Internet of Things", International Journal of
Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue4,
November (2019)
[10] Mircea Popa and Catalin Iapa "Embedded Weather Station with Remote Wireless Control", 19th Telecommunications forum TELFOR 2011 Serbia, Belgrade, November 22-24, 2011, 978-1-4577-1500-6/11/$26.00, IEEE, 2011
[11] T.R.V. Anandharajan G. Abhishek Hariharan, K.K. Vignajeth, R. Jijendiran4kushmita "Weather Monitoring Using Artificial Intelligence" 2016 International Conference on Computational Intelligence and Networks. 2375-5822/16 $31.00, IEEE DOI 10.1109/CINE.2016.26, (2016)
[12] Yashaswi Rahul, Rimsha Afreen, Divya Kamini "Smart weather monitoring and real-time alert system using IoT" International Research Journal of Engineering and Technology (d) Volume: 05 Issue: 10, Oct(2018)
[13] Ravi Sharma, Shiva Prakash and Pankaj Kumar, "Methodology, Applications and Challenges of WSN IoT", IEEE International Conference of Electrical and Electronics Engineering (ICE3—2020), jointly organized by Madan Mohan Malaviya University of Technology Gorakhpur and North Dakota State University, Fargo, USA, Pid197, Feb. 14-15, (2020)
[14] Garima Verma, Shiva Prakash, "A Comparative Study Based on Different Energy Saving Mechanisms Based on Green Internet of Things (GIoT)", IEEE 8 the International Conference on Reliability, Infocom Technology and Optimization (ICRITO-2020), IEEE Conference Record Number 48877, Amity University, Noida, India, Pid631, June 4-05, (2020)
[15] Ferdin Joe John Joseph "IoT Based Weather Monitoring System for Effective Analytics "International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-4, April (2019)
[16] Tanmay Agarwal M.S.Antony Vigil,Rishabh Jain,Abhinav Chandra "A Novel Data-driven optimal methodology for detecting ship from Sar images based on Artificial Intelligence"
[17] Antony Vigil, W Abelwin Pereira, Divya Jaikanth Naicker, J Anto Levin "SMART DRIP IRRIGATION AND REAL TIME MONITORING SYSTEM USING IOT AND DATA ENCRYPTION ALGORITHM"
[18] Kamgba FA, Edet CO, Njok AO. Effects of some meteorological parameters on wind energy potential in Calabar, Nigeria. Asian journal of physical and chemical. sciences. 2017 Nov 16:1-7 [19] Hussain M. Mahmud I. pyMannKendall: a python package for non-parametric Mann Kendall family of trend tests. Journal of Open Source Software. 2019;4(39):1556. DOI: 10.21105/joss.01556 [20] Sen PK. Estimates of the regression coefficient based on Kendall's tau. Journal of the American statistical association. 1968 Dec 1;63(324): 1379-1389