E-ISSN 2581 – 7957
P-ISSN 2277 – 3916
CVR Journal of Science and Technology, Volume 23, December 2022
DOI: 10.32377/cvrjst2312

A Machine Learning Perspective for Data Analytics in Solar Powered Weather Station using IoT

S. Praveen Chakkravarthy[1], P. Bharath Chandra[2], D. Anwar Basha[3] and G. Sai Kiran[4]
[1]Assoc. Professor, CVR College of Engineering/ECE Department, Hyderabad, India
Email: dr.praveen@cvr.ac.in
[2]UG Schloar, CVR College of Engineering/ECE Department, Hyderabad, India
Email: bharathchandra2002@gmail.com
[3]UG Schloar, CVR College of Engineering/ECE Department, Hyderabad, India
Email: bashadanwar@gmail.com
[4]UG Schloar, CVR College of Engineering/ECE Department, Hyderabad, India
Email: gsaikiran502@gmail.com

***Abstract*: In the context of actions for small cities, CVR Solar Powered Weather Station has been developed to monitor changes in Weather. Users can find out the weather changes in the area and can plan their day-to-day activities. In several important productive areas, such as farming, the climate plays a crucial role. These days, there is a lot of climate change, which is why it is bothersome that old weather forecasts are growing closer and less accurate. Therefore, miles are crucial to embellishing and modifying the weather forecast model. To cater the needs of current socio-economic changes, a solar powered IoT Based Weather Monitoring System plays a vital role in predicting changes in weather.**

A Machine Learning Model to Predict Weather Conditions in Mangalpalli Area is presented in the article.

***Index Terms:* LoRaWAN Gateway, Machine Learning Model, Photosynthetically Available Radiation, Total Solar Radiation**

I. Introduction

LoRaWAN weather station enables one to measure atmospheric conditions to provide information for weather forecasts and to study the weather and climate. The Main Process Unit (MPU) Consists of various sensors that include

- Rain Gauge Sensor
- Temperature/Humidity/Pressure sensor,
- Wind Speed/direction sensor,
- Illumination sensor,
- CO2 sensor,
- Rain/Snow sensor,
- PM2.5/10 sensor,
- PAR (Photosynthetically Available Radiation) sensor,
- Total Solar Radiation sensor

Main process device WSC1-L is an outdoor LoRaWAN RS485 end node. It has a built-in lithium-ion backup battery and is fueled by external 12-volt solar power. The LoRaWAN wireless protocol is used by the WSC1-L to read values from a variety of sensors and transfer the sensor data to an IoT server. The WSC1-L can function with a typical LoRaWAN Gateway and is completely compatible with the LoRaWAN Class C protocol. A woody area utilized for weather forecasting is a part of a system of information and statistics analysis techniques. Weather must be considered, including temperature, rain, humidity, pressure and other protection, as one of the greatest natural barriers in all aspects of our existence, remarkable.

The purpose of the proposed work is to format accurate weather forecasts. Long-term climate change on Earth will occur, and its effects on current and upcoming generations are both unknown. Our ability to forecast end-of-life climates is a fantastic opportunity to provide information so that stadium insurers can make educated wishes for the future of the globe.

The proposed model aims in a way to govern the state of staff inconsistencies and inequalities and performs its function of accurately predicting the weather. There are three basic types of load forecasts: the first is the short-term prediction, which involves estimating demand from a few hours to a few days. Second, there are long-term forecasts, which aim to predict demand from a few years to a few months, and medium forecasts, which aim to predict demand from a few weeks to months. Various potential load forecasting techniques, including the Similar Day Approach, Regression, Time Series Analysis, Artificial Neural Networks, Expert Systems (rule based), Fuzzy Logic, and Support Vector Machine (SVM) are evaluated and presented. In order to exploit these predictions online for efficient energy management, this Work covers the fundamental concept of load forecasting utilizing ML algorithms in an IoT setting.

Internet of Things is a novel paradigm combining telecommunications [1] and any kind of device using sensors. The Internet of Things (IoT) is seen as an innovation and financial wave in the global data sector. The Internet of Things (IoT) is a sophisticated system that connects everything to the Internet in order to exchange data and transmit through devices that can detect it in accordance with established norms. It succeeds in achieving the goal of keenly identifying, tracking, following, overseeing, and observing things. It is an expansion and augmentation of an Internet-based system that increases communication between people or between people and things or between things and things. According to the IoT vision, many objects around us will be connected to systems in some way.

Due to the extraordinary potential of the idea that almost every device may be viewed and controlled remotely through a link to the Internet, the Internet of Things (IoT) concept has recently attracted a great deal of interest from the scientific and industrial sectors. Such widespread

Received on 31-10-2002, Revised on 04-11-2022, Accepted on 17-11-2022

connectivity would make a variety of services possible in many different contexts. Cities might profit from smart lighting control, more effective trash management, and ongoing infrastructure monitoring, for instance.[2]

Connected sensors can be used in industrial settings to continuously monitor the production process, allowing for the quick detection or even prediction of failures, while in the agricultural industry, the extensive collection of environmental data, such as temperature and soil moisture, can increase the quantity and quality of soil production while lowering costs. Various application scenarios could also include health monitoring, home security, and home automation.[3]

The term "smart environment" refers to an environment that is self-protecting and self-monitoring when it is equipped with sensor devices, microcontrollers, and various software applications. When an occurrence takes place in such an environment, the dashboard shows warnings. Intelligent environmental monitoring can help to monitor and manage the consequences of environmental changes on people, plants, and animals.

To gather the information needed to forecast the behavior of a certain area of interest, sensor devices are positioned at various points. The primary goal of the proposed Work is to create and put into place an effective monitoring system that will allow the required parameters to be supervised remotely over internet while the data collected from the sensors is stored in the cloud and the approximated trend is projected on the web browser.

In this work, a wireless embedded computing system is proposed as a method for monitoring the temperature, humidity, and CO levels, or any parameter value crossing its threshold value ranges, for example, CO levels in the air in a particular area exceeding the normal levels, etc., in the environment. Furthermore, the answer provides advanced remote monitoring for data collected.

The last day's weather scenario is entirely dependent on weather prediction difficulties to determine how much the weather might vary in the future correspondingly. The concept of renewable solar systems is presented in this proposed work, along with several factors that have a significant impact on how well weather conditions are appreciated. These factors include the impact of solar radiation because of environmental reactions and reflections, which alter temperature conditions and consequently humidity conditions. Another important component that greatly affects climate conditions, including air velocity, air density, air direction, and air coolness, is wind speed. Thus, the suggested circumstances and variables have a significant impact on how accurately humans can predict the weather each day.

II. Literature Survey

A. Conventional Weather Stations

Surface observations are ones that are made close to the surface. Instrument shelters around 2.0 metres above the ground are typically where temperature and humidity instruments are kept. The shelters are designed to provide enough ventilation while also shielding the instruments from direct sunlight, precipitation, and moisture. In the past, these shelters' designs have differed from nation to nation, and within nations, shelters have changed over time. In certain cases, systematic biases were introduced by the shelters or the tools. A thermometer that routinely reads temperatures that are excessively high or low is an illustration of a systematic mistake. Multiple tools and techniques are used to measure precipitation. The most popular technique is to gauge how deep the water is inside a container.

B. Modern Observing Systems

Meteorological Services will need to update their organisational structure and service model to provide the best service to all users who require meteorological support, continuously provide the users with more reliable data, and put to use the products and innovations created by modern technology in the field of meteorology for both domestic and international users.

On the other hand, AWOSs can offer useful data and goods for a population's general safety and wellbeing as well as the numerous related economic advantages that can be obtained from these systems. In a number of crucial areas, including environmental monitoring for general forecasting and severe weather conditions, transport safety for road, rail, sea, and air vehicles, and educational and research purposes for the present and future understanding of global climatic conditions, the use of a modern automated surface observation system can meet these requirements.

In [4], the author suggested a reliable and cost-effective automatic weather station. The author of this paper explains how the weather prediction system is evolving into a key issue in every weather extreme event that has a negative impact on both lives and property. In order to improve weather, forecast abilities and increase resilience to the effects of unfavourable weather report conditions, the accuracy of weather data is one of the key problems.

Any device running server software is also considered to be a server. Servers are responsible for managing network resources. The Internet-based services and information are connected by LAN and made freely available to consumers via smart phones, internet browsers, or other browser-based devices in order to boost the system's intelligence, adaptability, and efficiency. Through this effort, people can learn about climatic changes. It functions well when it is accurate and efficient. So, the purpose of this work is for the author to develop a weather monitoring system using IoT. This work is straightforward to construct because it uses both hardware and software. The concept's creator uses a different sensor to collect climatic information, which is subsequently stored in the cloud.

III. System Model for Prediction Using Machine Learning

A. Random Forest Algorithm

Popular machine learning algorithm Random Forest is a part of the supervised learning methodology. It can be applied to ML issues involving both classification and regression. It is built on the idea of ensemble learning, which is a method of integrating various classifiers to address difficult issues and enhance model performance.

Multiple models that are combined to solve classification and regression issues are referred to as ensemble algorithms. By examining a training set of data, classification seeks to pinpoint the discrete set of a new observation.

Competitive and cooperative ensemble forecasting are two subcategories of ensemble forecasting. According to a survey done by Soares et al., successful ensemble ML algorithms can be used to solve regression problems. [5] Ren et al. presented state-of-the-art ensemble algorithms in 2015, with a focus on forecasting wind and solar power [6].

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead, then depending on a single decision tree, the random forest uses forecasts from each tree and predicts the result based on the votes of the majority of predictions. More trees in the forest result in increased accuracy and mitigate the overfitting issue.

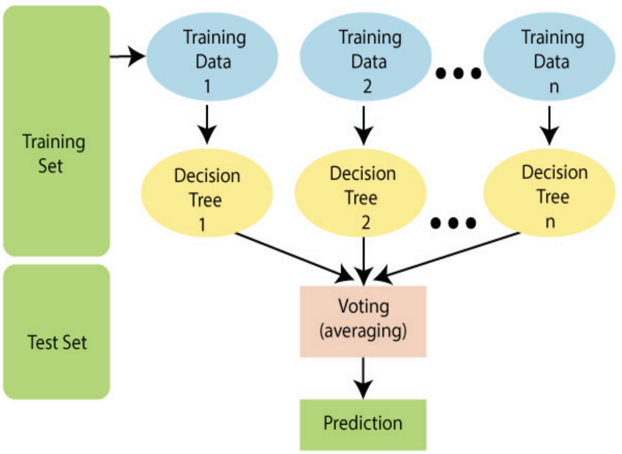


Figure 1. Various Techniques to train dataset

Figure 1 describes the various training techniques of the Random Forest classifier algorithm.

B. How does the Random Forest algorithm work?

First, N decision trees are combined to generate the random forest, and then predictions are made for each tree that was produced in the first phase. The Working process can be explained in the below steps and diagram:

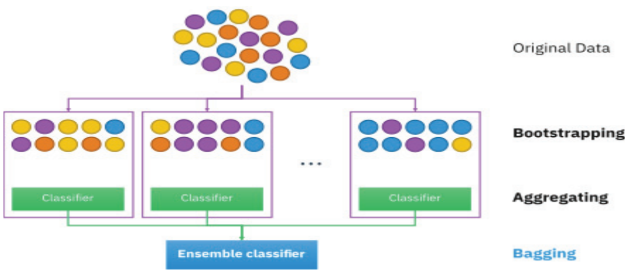


Figure 2. Various methodologies

Figure 2 explains various methodologies involved in training the raw data set to get an accurate prediction.

Step-1: Select random K data points from the training set.
Step-2: Build the decision trees associated with the selected data points (Subsets).
Step-3: Choose the number N for decision trees that you want to build.
Step-4: Repeat Step 1 & 2.
Step-5: For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

The ensemble method employed by random forest is bagging, sometimes referred to as Bootstrap Aggregation. A random sample is chosen from the data set using bagging. As a result, each model is created using the samples (Bootstrap Samples) that the Original Data gave, with a replacement process known as row sampling. Bootstrap refers to this stage of row sampling with replacement. Each model is currently trained independently, producing results. After merging the outputs of all the models, the final decision is made based on a majority vote. Aggregation is the process of aggregating all the results and producing a result based on a majority vote.

This model was created using a Random Forest classifier, and it has an accuracy of 85.0%. It was trained using 10,000 datasets, while we also tried K-Nearest Neighbours training. However, the accuracy of the Random Forest classifier is higher than that of other models.

IV. Related Work

When compared to traditional weather monitoring systems, the smart weather monitoring system is quite compact and simple to install. The Smart Weather Monitoring System's use of far less expensive sensors makes this work very cost-effective. Data from the sensors may also be sent to a web page that is accessible from any location in the world. Because there are fewer parts, the smart weather monitoring system's maintenance costs are also quite low. The Smart Weather Monitoring System's sensors collect and analyse data that is used to predict the weather. These sensors can easily detect any sudden change in the forecast, because of their high speed. Weather alerts are provided in advance.

The recommended method considers both the practical application of various sensing modalities and their best integration. Additionally, machine learning strategies and deep learning architectures are applied to the various inputs, and the outcome from both strategies is combined to create the final decision, which predicts rainfall and provides an appropriate judgement on how much irrigation should be done. In contrast to current practises, the proposed method has offered a technology-based solution that would be advantageous to the agricultural and scientific communities due to its portability and use of edge analytics, where input is processed, and output is given at the device level only without the use of cloud platforms, the internet, or Wi-Fi.

Wi-Fi, Bluetooth, and Zigbee are a few examples of common wireless network-based protocols that can be used as IoT infrastructure. However, all these protocols have drawbacks, including low range coverage and high-power consumption, and are therefore unsuitable for use in distributed applications with constrained resources. LoRa, which offers long-distance wireless communication, low power consumption, and a relatively low price, but has a limited data transmission capacity, can be the primary answer to this issue [7].

An outdoor Lora WAN gateway is the DLOS8. Through Wi-Fi, Ethernet, 3G or 4G cellular (the latter two supported by an extra module), you can connect a LoRa wireless network to an IP network.

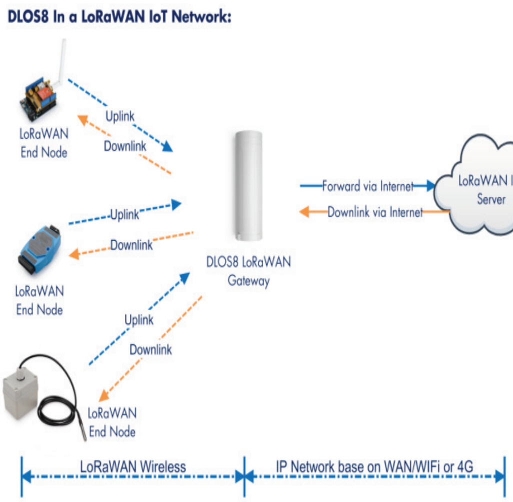


Figure 3. Architecture of LoRaWAN

Figure 3 shows a typical LoRa network. It is designed to allow low-powered devices to communicate with Internet-connected wide area networks [WAN] applications over long-range wireless connections [8]

Many weather monitoring systems are designed by deploying various sensors. Proposed model is Solar Powered IoT Weather Station Using Rain Gauge which senses the values from the LoRaWAN Weather Sensors and sends them to Network Server through LoRaWAN Gateway. LoRaWAN provides unlicensed spectrum so that it's cost free. End Device WSC-1 collects data from various sensors through RS485 Converter board and the data is sent to the gateway from WSC-1 Node. In the Things Mate dashboard, one can see the latest values and the variations of parameters in a plotted graphs and one can know whether it is preferred to stay in the surroundings by Alerts shown in Dashboard indicating Co2, PM2.5, PM10 Levels.

Figure 4. Working Model

Figure 4 delineates the working model of the weather station along with the solar panel and all the sensors as discussed in table 1.

V. Results and Discussions

More consistent and trustworthy meteorological data must be obtained and delivered as soon as possible to those who are affected in order to meet the growing needs of the developing world. There is a great need for meteorological data assistance in many industries today, including aviation, transportation, agriculture, construction, tourism, health, justice, security, national defence, sports, written press, and visual press.

Since the atmosphere is alive, it must be regularly monitored by noting any notable alterations and phenomena. Only AWOS—Automated Weather Observing Systems can be used to do this. Continuous observations cannot be made with non-automated technologies.

Our network and application servers are both hosted on the things mate server. The data on the network server is unstable. Between the gateway and the application server, the network server serves as a medium. The data is erased from the network server as it is saved in the application server.

Cloud management helps one to access the data, and one can monitor the changes happening around them and take preventive measures.

TABLE I.
List of Sensors Deployed

S.NO	Atmospheric Parameters used	Range
1	Wind speed range	(0-30m/s)
2	Wind direction	(0-360°)
3	C02	(0-5000ppm)
4	Pm2.5/10	(0-1000µg/m3)
5	Temperature sensor	(-30 ~ 70°C)
6	Humidity	(0 -100%RH)
7	Illuminance	(0-2/20/200Klux)
8	Rain Guage	(0-4mm/minute)
9	Pressure Sensor	(10-1100hPa) (hPa=Hecta pascals)
10	Total Solar Radiation Wavelength	(300nm to 3000nm)
11	Photosynthetically available radiation	(0-2500µmol/m2•s)
12	Rain/snow detect	Yes/No

E-ISSN 2581 – 7957
P-ISSN 2277 – 3916

DOI: 10.32377/cvrjst2312

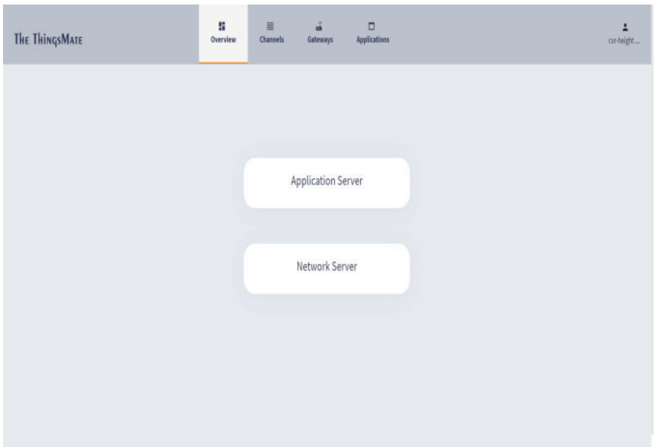


Figure 5. Things mate webpage

Figure 5 shows the application and network server where one can monitor the total number of sensors connected to LoRaWAN gateway. Our network and application servers are both hosted on the things mate server.

Figure 6. LoRaWAN Gateway

Figure 6 depicts a working model of the outdoor gateway which is the backbone of this research work. It enables the communication between the end device and end user. The DLOS8 is an open source outdoor LoRaWAN Gateway. It lets you bridge a LoRa wireless network to an IP network via Wi-Fi, Ethernet, 3G or 4G cellular (3G/4G is supported by optional modules).

TABLE II.
TEMPERATURE READINGS

Days	Sensor Data(^{0}C)	Reference Data(^{0}C)
Day 1	23	25
Day 2	25	26
Day 3	26	27
Day 4	28	29

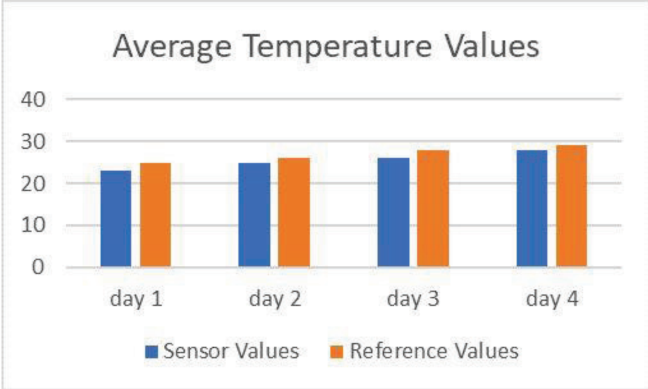


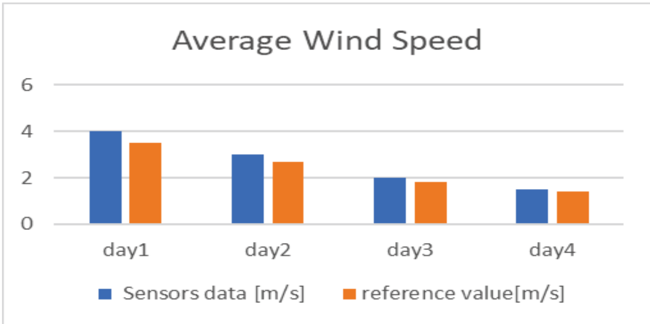
Figure 7. Trends of Temperature Data

From Figure 7 it can be clearly observed that the data received from the Temperature Sensor is relatively equal to that of the Reference Data.

TABLE III.
WIND SPEED READINGS

Days	Sensor Data (m/s)	Reference Data(m/s)
Day 1	4	3.5
Day 2	3	2.7
Day 3	2	1.8
Day 4	1.5	1.4

Figure 8. Trends of Wind Speed



From Figure 8, it can be seen clearly that the wind speed is relatively the same when compared with the reference values.

TABLE IV.
HUMIDITY READINGS

Days	Sensor Data (%)	Reference Data (%)
Day 1	95	93
Day 2	96	97
Day 3	98	97
Day 4	99	98

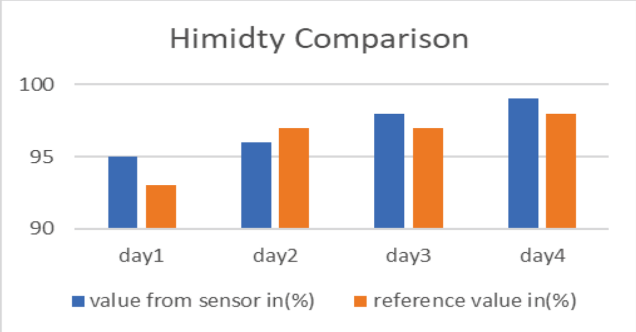


Figure 9. Trends of Humidity

From Figure 9, it can clearly see that humidity levels are relatively the same.

VI. Conclusions

Weather forecasting is a difficult task but a crucial area of study. because it pertains to our everyday lives. One can forecast the weather using machine learning and deep learning algorithms based on a variety of input features. To achieve better results, IoT approaches can be effectively combined with machine learning and deep learning. The location and timing of the weather station have a significant impact on prediction accuracy. The outcomes demonstrated that deep learning methods and mixed machine learning techniques can both improve accuracy. This concept offers a practical solution for ongoing environmental monitoring through the deployment of weather stations, the development of a smart environment, and the protection of public health from pollution.

References

[1] L. Atzori, A. Iera, and G. Morabito, “The Internet of Things: A survey,” computer. Networks, vol. 54, no. 15, pp. 2787–2805, 2010.
[2] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of things for smart cities,” IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22–32, 2014
[3] D. Bandyopadhyay and J. Sen, “Internet of things: Applications and challenges in technology and standardization,” Wireless Personal Communications, vol. 58, no. 1, pp. 49–69, 2011.
[4] Mary Nsabagwaa, Maximus Byamukamab, Emmanuel Kondelaa, “Towards a robust and affordable Automatic Weather Station “, journal homepage: www.elsevier.com/locate/deveng.
[5] J. Mendes-Moreira, C. Soares, A. M. Jorge, and J. F. de Sousa, “Ensemble approaches for regression: A survey,” ACM Comput. Surv., vol.45, no. 1, pp. 1–10, Nov. 2012.
[6] Y. Ren, P. N. Suganthan, and N. Srikanth, “Ensemble methods for wind and solar power forecasting: A state-of the-art review,” Renew Sustain. Energy Rev., vol. 50, pp. 82–91, Oct. 2015.
[7] R. S. Sinha, Y. Wei, and S. H. Hwang, “A survey on LPWA technology: LoRa and NB-IoT,”ICT Express, vol. 3, issue 1, pp 14-21, March 2017.
[8] A. Augustin, J. Yi, T. Clausen, and W. M. Townsley, “A study of Lora: Long range & low power networks for the internet of things,” Sensors (Switzerland), vol. 16, issue 9, pp 1466,2016.

The project is funded by Department of Science and Technology (DST), under NEWGEN IEDC scheme for the Academic year 2021-2022.