Kiran S. Todekar, et. al. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 10, Issue 10, (Series-IV) October 2020, pp. 20-24

RESEARCH ARTICLE OPEN ACCESS

IoT Based Distributed Environmental Monitoring System and Weather Prediction using Machine Learning Approach

Kiran S. Todekar, Prof.Shashikant S. Hippargi
Electronics and Telecommunication Engineering N B Navalesinhgad college of Engineering Solapur, India
Electronics and Telecommunication Engineering N B Navalesinhgad college of Engineering Solapur, India

ABSTRACT
Today environment monitoring becomes important for humans to ensure a safe and wealthy life. Monitoring requirements are extremely different depending on the environment, leading to specially appointed usage that needs adaptability. The proposed system describes an implementation of Wireless Sensor Network that can be adjusted to various applications. And it also inserts the adaptability required to be conveyed and updated without necessity of arranging complex infrastructures. The solution is based on small autonomous wireless sensor nodes, small wireless receivers connected to the Internet, and a cloud architecture which provides data storage and delivery to remote clients. The solution permits supervisors on-site not only to monitor the current situation by using their smart-phones but also to monitor remote sites through the Internet. All measurements are stored at different levels to guarantee a safe back-trace and to access data stored in case of network failure or unavailability. The proposed system is useful in predicting environmental condition like rain, also, to measure the temperature, humidity, and presence of pressure.
Keywords-*IoT, Machine learning, Node MCU, Temperature Sensor, humidity Sensor, firebase Cloud .*

Date of Submission: 10-10-2020 Date of Acceptance: 26-10-2020

I. INTRODUCTION

The environmental care has become one of the biggest concerns for almost every country in the last few years. Even though the industrialization level has been increasing without any control in the last decades, the current situation is clearly changing towards more environmentally friendly solutions. Water and air quality are essential to maintain the equilibrium between human development and a healthy environment. It is also important to notice that by means of looking for a more efficient production in factories both pollution and consumption of natural resources can be decreased. Processes, such as boiling, drying, binding, and so forth, are being carried out by almost every kind of the current factories. Those processes are responsible of a great amount of gas emissions and polluted water discharges. Although the majority of the factories have their own sewage plants, it is crucial to measure the quality of the waste water that is being poured into the public sewer.

In reality, clean air is a basic requirement for daily life. Air pollution affects human health and considered as a major serious problem globally, especially in countries where gas and oil industries are ubiquitous. According to the United States Environmental Protection Agency (USEPA), the air quality is characterized by measuring certain gases that affect the human health, which are: carbon monoxide (CO), ground-level ozone (O3), and hydrogen sulfide (H_2S).

The main intention of environmental monitoring is not only to gather data from a number of locations, but also to provide the information required by scientists, planners, and policy-makers, to enable those making decisions on managing and improving the environment, in addition to presenting helpful information to end-users. There are huge efforts are carried out to improve the air quality in both environments: indoors and outdoors. Habitat and environmental monitoring represent an important class of sensor network applications. Recent advances in low-power wireless network technology have created the technical conditions to build multi-functional tiny sensor devices, which can be used to sense and observe physical phenomena. Wireless Sensor Networks (WSNs) are currently an active research area due to their wide range applications including military, medical, environmental monitoring, safety, and civilian. Many environmental monitoring examples of WSNs are already presented in the literature and developed for different purposes.

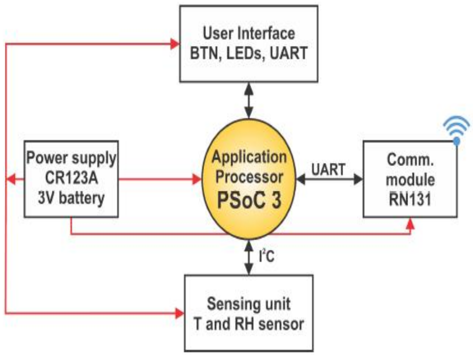

Figure 1: Block Diagram

The two sensors that communicate using Wi-Fi technology are based on the same hardware, the difference between the two consisting in the protocol that was used, namely UDP or HTTP. The generic architecture of the devices based on Wi-Fi technology is presented in Fig. 1. It consists of the application processor, a CY8C3246PVI-147 programmable system on chip microcontroller (PSoC 3) produced by Cypress Semiconductor, a wireless local area network (LAN) module, RN-131C/G , and the temperature and relative humidity sensor (DHT22), all powered by using a battery . The choice of using a separate application processor removes the possibility of interfering with the communication stack on the WLAN module. Therefore, the processor in the developed devices is in charge with performing all the actions for the proper operation of the device, namely, power management, acquisition of data from the sensing unit, and communication. For transmitting the data to a base station, a serial link between the PSoC 3 device and the communication module, and an API (application programming interface), called WiFly, and are used. The motivation for selecting the RN-131C/G wireless module consists in its low-power operation, providing 4 μA during sleep and short 210 mA pulses during transmission. The development of other cheaper wireless modules based on the IEEE 802.11 set of standards, such as the ESP8266 from Express if, will multiply the range of possible solutions. However, with a current larger than 20 μA in deep sleep mode, newer approaches for achieving power efficiency in the designs using it will have to be investigated.

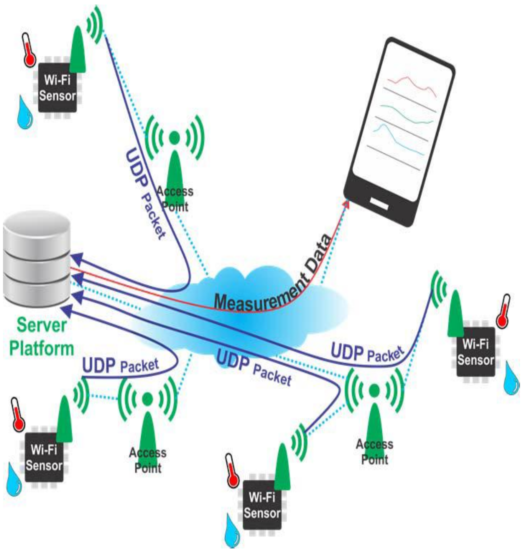

Figure 2: Environmental Monitoring System Based on UDP communication

UDP-based cyber-physical system for monitoring the temperature and relative humidity in the environment or ambient in this advantage of the existing IEEE 802.11 infrastructure for sending measurement information to a cloud-based server platform and provides the possibility of visualizing the data from every device with an Internet connection. Using UDP allows the low-power operation of the Wi-Fi sensors, because of its connectionless nature. Furthermore, this protocol provides lower packet sizes, increased speeds and low latency, compared to TCP/IP. All these come at a price, namely, a loss in transmission reliability, because there is no acknowledge message received for the packets being sent. The server platform can reside in the cloud, or can be an UDP listener running on a computer, that can interpret the received data, store them in a database, and provide the possibility of visualizing and processing them according to the user's needs, through a Web server.

II. METHODOLOGY

For determining temperature and humidity of atmosphere we are using temperature and humidity sensor which will help in predicting environmental conditions using machine learning Algorithm. The system is placed at locations and data from location is collected by the server as shown in figure. The server stores and displays the current values of all parameters.

A look up table is generated which contains the values of temperature and humidity and is used for predicting the current environmental conditions like if humidity is more and temperature is less then chances of rain is more etc.

The future state of the atmosphere is computed by solving numerical equations of thermodynamics and fluid dynamics. But this traditional system of differential equations that govern the physical model is sometimes unstable under disturbances and uncertainties while measuring the initial conditions of the atmosphere.

But Machine learning is relatively robust to most atmospheric disturbances as compared to traditional methods. Another advantage of machine learning is that it is not dependent on the physical laws of atmospheric processes.

Linear Regression: it uses all the features present in the dataset and gives a linear graph combining high and low temperatures. Linear regression is not used for weather classification of each day because this algorithm cannot be used with classification data.

Functional Regression: The second algorithm to be used is a type of functional regression. It looks for historical weather patterns which are similar to the present day weather patterns, and then it predicts the future weather condition.

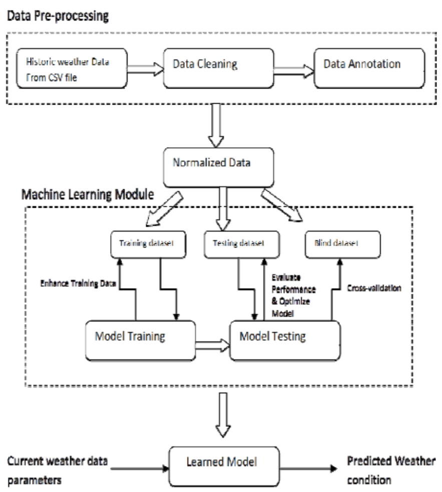

Figure 3: Architectural Model for Weather Prediction

III. DATASET ANALYSIS WITH MODEL

The data enclosed in our dataset is classified into the following categories:-

i) Temperature
ii) Atmospheric pressure
iii) Humidity

Temperature is a measure of the degree of hotness or coldness of the surroundings. It, like all weather conditions, varies from instance to instance. Similarly, atmospheric pressure and humidity, that plays a vital role in predicting whether an area will receive precipitation or not, is also included in the dataset. Details about fog and dew point are included in the dataset as well, as they only contribute to improving the accuracy of the predictions made by the prediction models.

Given below is a tabular representation of the data collected in the dataset:

Date Time	Temperature	Humidity	Pressure
04-08-2020 13:30	27	60	950
04-08-2020 13:31	27	60	950.5
04-08-2020 13:32	26.9	60.5	951
04-08-2020 13:33	26.9	60.3	951.2
04-08-2020 13:34	26.9	60.2	951.1

Table No.1 Data Set

3.1Test Metrics

3.1.1 Scikit-Learn library in Python

Scikit-learn are a free machine learning library for Python. It highlights different algorithms like support vector machine, random forests, and k-neighbors, and it likewise underpins Python numerical and scientific libraries like NumPy and SciPy.

The library has functions like accuracy score(), RandomForest(),Regressor () and many other very useful regression functions that enable us to make accurate predictions.

Given below is a snapshot of the use of the library in the code

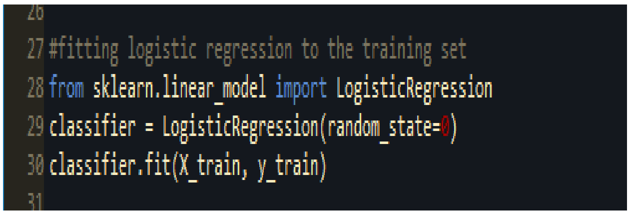

Figure 4: Implementation of Scikit-Learn Library for Logistic Regression

3.2 Pandas

Pandas are an open source, BSD-authorized library giving superior, simple to-utilize information structures and information investigation apparatuses for the Python programming language.

Pandas have been heavily utilized in the development of this project. Given below are a few snapshots from the code

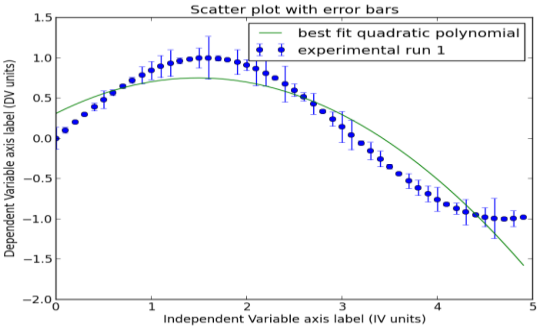

Figure 5: Model output

3.2.1 Confusion Matrix

A confusion matrix is a procedure which abridges the execution of an order calculation. It is a synopsis of forecast results on a characterization issue. Characterization precision can be deceiving there are an unequal number of perceptions in each class or if there are in excess of two classes in the dataset. Ascertaining a disarray lattice gives a superior thought of what the characterization demonstrate is getting right and the sorts of blunders it is making.

	Predicted Value	**Predicted Value**
Real Value	True Positive(TP) Reality: Rain ML model predicted: Rain	False Positive(FP) Reality: No Rain ML model predicted: Rain
Real Value	False Negative(FN) Reality: Rain ML model predicted: No Rain	True Negative(TN) Reality: Benign ML model predicted: No Rain

IV. TEST SETUP

The test process is already in-built in our system. The testing process taking place just after the model is trained. After the completion of the training process, we analyze each data entry in the test set. In order to analyze each entry, we use descriptors to extract features. Now we compare these feature values with the feature values which were initially retained using the train set. The comparison is done according to the Machine Learning model used and finally the output for each entry is received. Since each data entry is already labeled, we can compute accuracy by comparing the predicted value with the received value

V. CONCLUSION

All the machine learning models: linear regression, various linear regression, polynomial linear regression, logistic regression, random forest regression and Artificial neural systems were beaten by expert climate determining apparatuses, in spite of the fact that the error in their execution reduced significantly for later days, demonstrating that over longer timeframes, our models may beat genius professional ones.

Linear regression demonstrated to be a low predisposition, high fluctuation model though polynomial regression demonstrated to be a high predisposition, low difference model. Linear regression is naturally a high difference model as it is unsteady to outliers, so one approach to improve the linear regression model is by gathering of more information. Practical regression, however, was high predisposition, demonstrating that the decision of model was poor, and that its predictions can't be improved by further accumulation of information. This predisposition could be expected to the structure decision to estimate climate dependent on the climate of the previous two days, which might be too short to even think about capturing slants in climate that practical regression requires. On the off chance that the figure were rather founded on the climate of the past four or five days, the predisposition of the practical regression model could probably be decreased. In any case, this would require significantly more calculation time alongside retraining of the weight vector w, so this will be conceded to future work.

REFERENCES

[1]. The system proposed by George Mois, Member IEEE et.al[1] ,"Analysis of Three IoT Based Wireless Sensors for Environmental Monitoring "
[2]. The Mr.S.MShirsath, IJAREEIE et.al [2], "IoT based smart environmental monitoring using wireless sensor network" .
[3]. The Mr.T.Dhineesh et.al [3]. "Analysis of IoT based wireless sensors for environmental monitoring in agriculture."
[4]. The MrsReshmaS .Mundhe et.al [4]. "Environmental monitoring for IoT application based on wireless sensor network. "
[5]. The Dr. R sasikumar et.al [5], "Environmental monitoring system using IoT"
[6]. The research article on "weather forecasting using sliding window algorithm" by Kvyantum Inc. et.al
[7]. The Mr.Jitchashivang and et.al [7], "weather prediction for Indian location using machine learning"
[8]. Mrs Divyachauhan and Jawahar Thakur et.al [8],"Data mining technique for weather prediction. "
[9]. Christy Kunjumonet.al[8],"Survey on Weather Forecasting using Data Minin
[10]. Mohammad Wahiduzzaman, Eric C. J. Oliver, Simon J Wotherspoon, Neil J. Holbrook, "A

climatological model of North Indian Ocean tropical cyclone genesis, tracks and landfall".

[11]. Jinglin Du, YayunLiu ,Yanan Yu and Weilan Yan, "A Prediction of Precipitation Data Based on Support Vector Machine and Particle Swarm Optimization (PSO-SVM) Algorithms"

[12]. Prashant Kumar, Atul K. Varma, " Atmospheric and Oceanic Sciences Group, EPSA, Space Applications Centre (ISRO), Ahmedabad, IndiaAssimilation of INSAT-3D hydro-estimator method retrieved rainfall for short-range weather prediction"

[13]. Prashant Kumar, C. M. Kishtawal, P. K. Pal, "Impact of ECMWF, NCEP, and NCMRWF global model analysis on the WRF model forecast over Indian Region"

[14]. H. Vathsala, Shashidhar G. KoolagudiPrediction, "Model for peninsular Indian summer monsoon rainfall using data mining and statistical approaches"

[15]. Mark Holmstrom, Dylan Liu, Christopher Vo, "Machine Learning applied to weather forecasting", Stanford University, 2016.

[16]. GyaneshShrivastava, SanjeevKarmakar, Manoj Kumar Kowar, " Application of Artificial Neural Networks in Weather Forecasting: A Comprehensive Literature Review", International Journal of Computer Application, 2012.

[17]. MeeraNarvekar, PriyancaFargose, "Daily weather forcasting using Artificial Neural Network", International Journal of Computer Application, 2015.

Kiran S. Todekar, et. al. "IoT Based Distributed Environmental Monitoring System and Weather Prediction using Machine Learning Approach." *International Journal of Engineering Research and Applications (IJERA)*, vol.10 (10), 2020, pp 20-24.