2023 IEEE 8th International Conference for Convergence in Technology (I2CT)
Pune, India. Apr 7-9, 2023

Smart Electricity, Water and Air Quality Monitoring System

Sakshi Kamal
Computer Science and Engineering
RV College of Engineering
Bangalore, Karnataka
Email: sakshikamal.cs18@rvce.edu.in

Sai Keerthana Arun
Computer Science and Engineering
RV College of Engineering
Bangalore, Karnataka
Email: saikeerthana.cs18@rvce.edu.in

Dr Mamatha T
Computer Science and Engineering
RV College of Engineering
Bangalore, Karnataka
Email: mamathat@rvce.edu.in

Dr Anala M R
Computer Science and Engineering
RV College of Engineering
Bangalore, Karnataka
Email: analamr@rvce.edu.in

2023 IEEE 8th International Conference for Convergence in Technology (I2CT) | 979-8-3503-3401-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/I2CT57861.2023.10126191

Abstract—**Innovative technology is maneuvering the growing global issues related to the environment. This encompasses climatic changes, depletion of natural resources, the impact of industrialization on the environment, pollution caused due to increasing traffic etc. leading to the exploitation of natural resources and pollution of the environment. This paper presents a smart resource management system for predicting and forecasting the quality of air and water and usage of energy resources. Efficient techniques and machine learning algorithms have been compared to predict the quality of these resources. Monitoring and forecasting the usage and quality of these resources will help in effective utilization, precaution against diseases and reduction of cost in the long run.**

Index Terms—**Sustainable, Machine Learning, IOT, Energy resources, Tracking**

I. Introduction

Cost of Urbanization is the non-uniform consumption of basic resources like water, energy and deterioration of quality of water and air. In the broader sense the quality of air we breathe, water we consume has drastically decreased due to high pollution in these areas. The water bodies like rivers and lakes are contaminated due to discharge of wastewater and sewage into these water bodies. This results in excess of nitrogen, phosphorous and other toxic chemicals polluting the water bodies. This influences whole ecosystems and all the inhabitants of the planet. Exposure to such environment for longer period of time can cause several health-related issues. Electrical energy is an important non-renewable energy which is very essential in every social aspect of life. Conserving energy plays a vital role in contribution to economy of a country. Manual readings are error prone and tedious. In the conventional system there is no communication processes between the power providing company and the consumer premises for sending the meter readings and for generating real time billing. Periodic monitoring of energy, water consumption on daily, weekly or fortnight basis would give a better control of resource utilization. With everything around us getting smarter like smart homes, smart grids which holistically aims at simplifying the life of people and driving towards optimal usage of resources around us. A smart solution to monitor the resource consumption by members of the residence would give a better hold on usage of water or energy. Setting up threshold, sending alert messages and warnings on the usage of resources and providing user friendly mobile app are the advantages of the proposed system has compared to rigid conventional system which lacks flexibility, prone to errors and not control over the resource consumption.

Countries all over the world are coming up with solutions to prevent the overuse of such resources. For example, the European Union created an action plan with the aim of increasing energy efficiency and also on changing the energy behavior of the consumer. [1] Many organizations often keep track of their usage through meters and have access to easily accessible data. However, most of them fail to take further action and analyze trends data and assess organizational suitability. In the organization, the analysis and evaluation of energy data can help in compliance with environmental legislation, reduce carbon emissions, increase savings and promote a positive Corporate Social Responsibility (CSR) commitment. [2]

Devices can be attached with a sensor to monitor the energy and water consumption which is sent to the cloud. Analytics is done on the data collected at the cloud and registered user can monitor their resource usage. By applying machine learning algorithms to the data collected, area wise prediction of required resources usage can be carried out. Smart distribution of the resources according the the forecasts can be made. Different load distribution can be done at different period of time throughout the day based on the needs and usage of resources.

The objective of the work is to monitor, track and predict the usage of resources i.e., energy and water by residential and commercial buildings. For tracking electricity usage, a system of constant feedback is created, wherein if the usage crosses a certain threshold the user will get a notification on their phone regarding the same. Water and Air quality are tracked using live sensor data which will send collected sensor values in PPM (Parts per Million) and NTU for air and water quality respectively, to a web application via nodemcu. Random Regression model is used to predict the water and air quality. This information is available to the registered user via a mobile application.

II. Literature Review and Related Work

Utility meters can make use of automated meter reading technology for collecting the consumed data for billing purpose. The movement of mechanical dials are translated into digital signal hence does not require physical intervention for meter reading. The utility company receives this data via telephone line or cable. There are various implementation of Automated Meter reading depending on the communication module used. The communication to the utility company can be either wired or wireless, can use handheld devices, telephone wires, radio frequency waves etc. With AMR manual meter reads are becoming obsolete. Using AMR the data can be communicated in only one direction i.e. to the energy supplier. The customer does not have the facility to monitor his energy consumption. Wireless automation of AMR is implemented by leveraging the advantages of GSM module. The ARM7 LPC2148 microcontroller is used as the core for collecting the data and performing the control operations such as breaking the circuit. The information is then sent to the customers mobile phone using GSM modem. The implementation here is two way, the customer can send commands via his mobile device which retrieves the information from ARM7 LPC2148 microcontroller via GSM module. [3]

Meter readings in rural areas are done manually. Automating the meter readings using wireless zigbee as the communication technology for rural areas reduces the erroneous meter readings and wastage of man power. The Zigbee CC2430 is used which is System-on-Chip (SoC) solution tailored for zigbee applications. [4]

A complete system using arduino board is implemented for automating and monitoring meter reading which eliminates the human interference. A mobile app is developed where the users can set target consumption and they will receive notification on reaching the target. The system is implemented using arduino board for processing, hall sensors are connected to arduino which calibrates the electric meter readings. The readings are stored in database using Ethernet shield. Finally the user can check his energy usage and receive notification via android app, web interface and server. [5]

This paper addresses the challenge of energy consumption estimation by introducing a review of key ways to balance energy use in the field of computer technology, developed through machine learning programs. Its main focus is on defining high-level approaches to power consumption, especially in data mining and convolution neural networks. The compilation of the tested papers provides the necessary guidelines for exposing the use of force to the machine learning audience interested in incorporating energy as metrics in the design of machine learning programs. To illustrate the benefits of integration, two application cases have been introduced, showing, from data mining purposes and net- work networks, how to use different measurement methods. The benefits of continuing research in energy measurement can help machine learning researchers gain valuable insights into the design of machine learning systems. [6]

To increase the usage of renewable energy and to minimize the impact on the environment due to the usage of non-renewable energy IoT can be used. Use of IoT in power systems is reviewed in this paper by considering the IoT technology along with cloud computing and data analysis platform. The paper also reviews the various challenges faced while including IoT in energy sector. Focusing on the fullness of the IoT sensor ecosystem includes common sensors used to monitor energy, such as humidity, temperature, light, etc., with a study on the actuators and the communication technologies such as Bluetooth, Zigbee, etc. [5]

An IoT-based electric meter monitoring system that uses the android system to reduce manual efforts to measure power units to address users' concerns about excessive use of electricity. Arduino Uno and optical sensor are used to download electric beats. To reduce human error and cost in power consumption, a low-cost wireless network is used in digital power meters and mobile applications that can automatically translate unit meter. [7]

A smart grid with IoT includes a high-speed response and storage system and power consumption. This type of communication using IoT makes Smart Grid operation more reliable and more efficient without high demand times (because stored energy is used for demand). It also creates a better environment for collecting energy from distributed generation and enhances the use of performance in a smart grid environment using IoT. IoT for Smart Grid is considered in this paper which aims at better energy efficiency by saving energy by reducing wastage. [8]

The main purpose of power forecasting is the use of soft-computing techniques that play an important role in addressing the challenges of renewable energy use through various data mining techniques, including processing historical load data and features of the loading time series. This paper analyzes the potential for energy consumption from renewable energy sources and non-renewable energy sources. A robust machine-based learning method, including multilayer perceptron (MLP), vector regression support (SVR), and CatBoost, is suggested in this power forecast paper. Complete comparisons are made, taking into account the results obtained using other predictive methods. [9]

The project introduces a system that provides information on residential energy use by monitoring and forecasting energy. The system was sent to a living room with solar PV and electricity consumption was monitored for 28 days using an online cloud-server database. In addition, a variety of retrospective techniques and machine learning algorithms, such as direct and polynomial alignment, vector regression support (SVR) and Random Forest, were trained and used to identify a model that provides excellent accuracy in predicting total electricity use at the end of the month. Random Forest Regressor gave the slightest error of 0.58% and was also relatively easier to implement for a larger data set. [10]

In the existing systems surveyed, there are apps catering to one particular energy source only and not necessarily to a combination of two or more resources. Therefore through this paper we aim to create a platform wherein one can have access to monitoring and controlling their usage as and when required.

III. Proposed Work

The proposed work is composed of two systems- an IoT system and a Machine Learning system. The IoT system consists of sensors to monitor air and water quality and electricity usage, that are connected to the main circuit board of the Arduino Uno. These values are communicated to the Web App via APIs, where the internet connection is facilitated by the Node MCU. The Machine Learning system is employed with the task of forecasting or predicting the air and water quality parameters, for which various machine learning models have been tried and tested. The tracked sensor values as well as the predicted values are shown to the user via a dashboard. The scope is limited to electricity usage and Air and Water quality but can be expanded to other resources tracking as well.

A. Electricity Usage

Smart meters are very crucial when it comes to the sustainable consumption of electricity. However, the main drawback with respect to smart meters is that they aren't very affordable. With affordability in mind, this paper focuses on a solution that mimics the functionality of a smart meter at a much affordable price. The sensor used is ACS712 Electricity sensor, based on the hall effect with 2.1kVRMS single power and low-voltage current conductor that is simple to use and computes the current utilized.

B. Air and Water Quality

The parameters that are key for measuring water quality are pH value, turbidity and temperature. The main components used-Gravity Analog pH sensor, turbidity sensor, DHT11 sensor. Air quality sensor MQ135 for bulk gases like NOx,NH_3,benzene, smoke, CO_2 has a high responsiveness to Sulfide, Ammonia and Benzene steam and other toxic gases. It is apt for office or factory usage. It is relatively cheaper and especially suitable for air quality monitoring. When the gas gauges in PPM, sensor and module are sufficient.

The flow of the system is depicted in Figure 1. The user interacts with the web application, where they login using their credentials (name, password and house ID). The Web application has its backend linked to database to store the user authentication details. The Spreadsheet ID (For Air and Water Quality) and the Adafruit Dashboard ID (For Electricity usage) is retrieved from the database. The Real-time sensor values are extracted from the spreadsheet and a link to the Adafruit Dashboard for Electricity is supplied to the main Web application dashboard. The water and air quality predictive system presents a forecast for the qualities for the upcoming week (next 7-days) in a tabular and graph form in the dashboard.

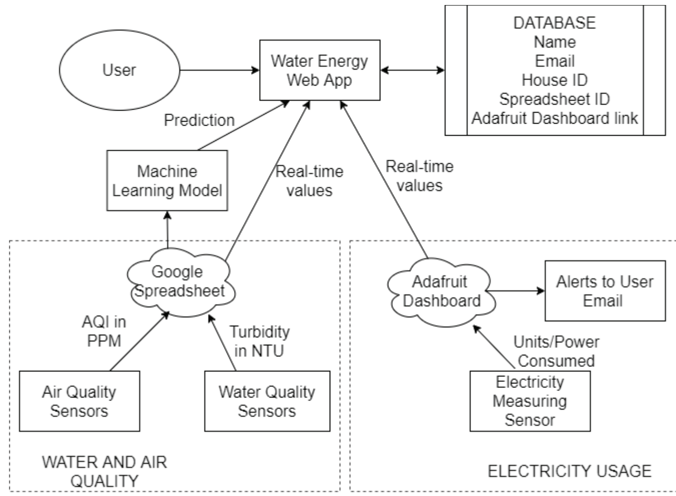

Fig. 1. Proposed System

IV. Methodology

The proposed system employs the use of affordable sensors and provides analysis based on the collected data in real time. In this way, it provides an instantaneous monitoring system that also is collecting data to make predictions on the said water and air qualities and electricity usage. The circuit for the system as shown in Figure 2 consists of: (A) Air Quality Sensor; (B) DHT11 Temperature Humidity Sensor; (C) OLED display; (D) ACS712 Electricity sensor; (E) Arduino Uno; (F) Turbidity Sensor; (G) pH Sensor; (H) Wifi Module (or Node MCU). The water and air quality values read by the sensors

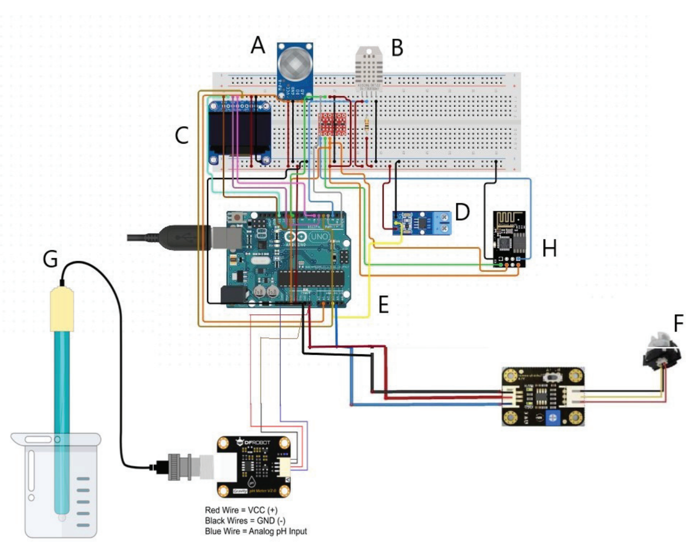

Fig. 2. Circuit Diagram

are used to populate the dataset that is self-replenishing, i.e.

designed to gather the predefined number of most recent values. The electricity is monitored directly by the Adafruit API and does not require a dataset. The sensor values are sent to a Google sheet via Google API as shown in Figure 3. This approach used for data collection is unique since the present predictive systems majorly use previously collected data that may not offer prediction to the most accurate level.

The Google Sheets API uses the sheet ID and a registered service account to recognize the sheet and user. Hence, its access is restricted to the admin ensuring data security. Retrieval, extraction and cleaning of the data is carried out using predefined python libraries. This also involves converting sensor data (in PPM and NTU units for air and water quality, respectively) into appropriate data types, such as float and datetime for the appropriate columns of sensor data and date and time respectively. Also, unwanted NaN values are omitted for further cleaning of the dataset.

Having collected the pre-defined number of datapoints, the model is trained for it, without stopping the data collection for sensor values. Training can then be scheduled at regular intervals according to user needs, following the workflow as shown in Figure 4. To finalise on the model for the system, a comparative study for various machine learning models such as LSTM, ARIMA, RandomForestRegressor was carried out.

Date	Time	Temperature(C)	Humidity(%)	Air Quality(PPM)	Turbidity(NTU)
24/04/2021	18:21:03	25	70	78	866
25/04/2021	14:14:53	25	70	78	866
25/04/2021	14:26:11	28.9	49	75	889
25/04/2021	14:26:14	28.9	49	75	889
25/04/2021	14:26:17	28.9	49	75	889
25/04/2021	14:26:19	28.9	49	75	889
25/04/2021	14:26:22	28.9	49	75	889
25/04/2021	14:26:27	28.9	49	75	889
25/04/2021	14:26:29	28.9	49	75	889
25/04/2021	14:26:31	28.9	48	75	889
25/04/2021	14:26:34	28.9	49	75	889
25/04/2021	14:26:36	28.9	49	75	889
25/04/2021	14:26:39	28.9	49	75	889
25/04/2021	14:26:42	29	50	75	889
25/04/2021	14:26:45	28.9	49	75	889
25/04/2021	14:26:49	29	49	75	889
25/04/2021	14:51:37	29.4	47	75	889
25/04/2021	14:51:40	29.3	47	75	889
25/04/2021	14:51:42	29.3	47	75	889
25/04/2021	14:51:44	29.3	48	75	889
25/04/2021	14:51:47	29.4	47	75	889
25/04/2021	14:52:22	29.4	48	200	889
25/04/2021	14:52:25	29.4	47	198	889
25/04/2021	14:52:27	29.4	46	197	889
25/04/2021	14:52:30	29.4	47	197	889
25/04/2021	14:53:52	29.5	46	171	889

Fig. 3. Dataset in Google Sheets

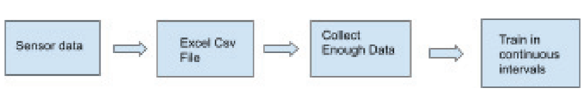

Fig. 4. Training workflow

1) **Long Short-Term Memory (LSTM)** has its core RNN which can rely on learning programs for unexpected learning problems. It is required for complex domains such as machine translation, speech recognition etc. The key feature of LSTM is the shape of its cell, which is similar to a conveyor belt. This runs straight across the cell with a few interactions that are linear. LSTM has the ability to add or delete information to a cell, which is supervised by its ”gates”. Gates are a way of letting information in the cell and consists of a sigmoid neural network and a pointwise multiplication operator.

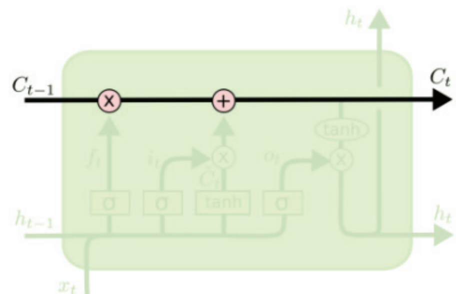

Fig. 5. LSTM Cell [11]

2) **ARIMA (Auto-Regressive Integrated Moving Average)** is a category of mathematical models used to predict and analyze time series data. It clearly provides a list of common structures in time series data. The summary captures key features of the model itself, they are:
AR, which stands for Autoregression. This uses a dependent relationship between the observation made and the number of observations that are delayed.
I, which stands for Integrated. The use of visual separation of raw observations so that the time series data remains stationary.
MA, which stands for Moving Average.This model uses the interval between view and the error that is residual from the central moving model used in the observations that are lagged. [12] The linear regression model is constructed such that it includes the specified number and type of names, it also ensures that the data is adjusted at a variable level to make it stationary, removing styles and frameworks that negatively affect the regression model.
3) **Random Forest Regression** belongs to the class of supervised learning algorithms. A constructive ”forest” is a group of decision-making trees, shown in Figure 6, usually trained using ”bagging” technique. Its fundamental idea is that it combines a bunch of learning models to enhance the total effect. The unplanned forest forms many decision-making trees and combines them together to get accurate and precise predictions. Another great advantage of the random forest is the fact that it can be used for reversal and segregation issues, which are responsible for most of the issues in the current machine learning systems.

These models were investigated and compared for air and water quality predictions which is elaborated under the results. Hence, the sensor values collected by the IoT system comprising of the sensors for air and water quality are able to serve as training and validation data for the model that further generates a forecast of the air and water quality in the upcoming week. In order to show these predicted values to the user, an interface in the form of a dashboard was developed

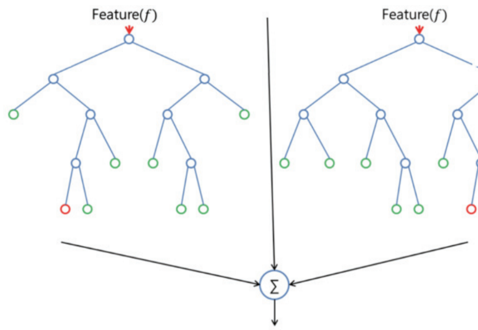

Fig. 6. Random Forest Regressor [13]

that had the functionalities of monitoring the resources quality as well as usage and also displaying the forecasts in the form of a graph as shown in Figure 8. For electricity consumption, another dashboard using Adafruit was developed shown in Figure 7.

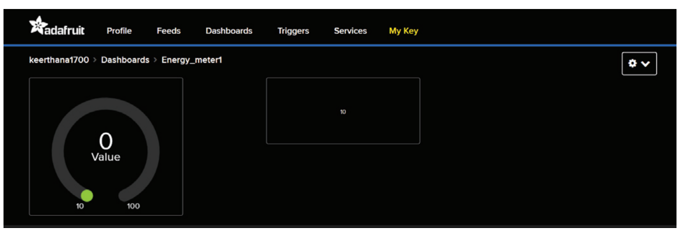

Fig. 7. Electricity Dashboard

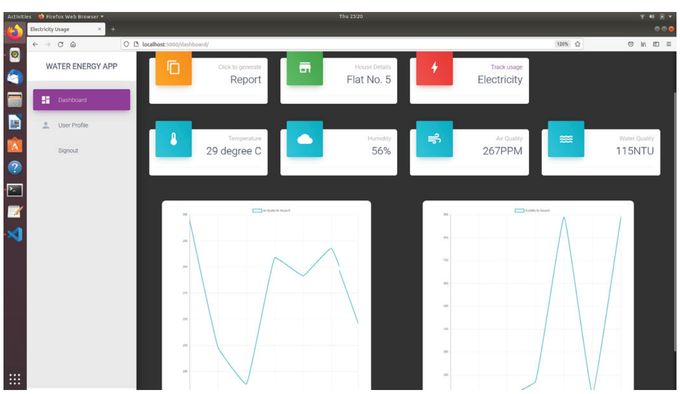

Fig. 8. Dashboard of Web Application

V. Results

The IoT system consisting of the circuit linked to Google Sheet and also to Adafruit dashboard gave real-time value changes within approximately 0.5 seconds. The delay being extremely less can help in receiving instantaneous alerts for air and water qualities and electricity consumption.

In the machine learning system, to obtain the most accurate forecasts, the models were trained using the constructed dataset of around 400 entries. The evaluation metrics used were Root Mean Squared Error (RMSE), Mean Absolute Error(MAE) and Mean Squared Error (MSE) that are calculated based on the comparison of the predicted and expected values. Table I showcases the error values for various implemented models.

TABLE I
ERRORS FOR DIFFERENT MODELS - ROOT MEAN SQUARED ERROR(RMSE), MEAN SQUARED ERROR(MSE), MEAN ABSOLUTE ERROR(MAE)

Error\Model	LSTM	ARIMA	RandomForest
RMSE	48.94	97.49	23.94
MSE	2395.10	9504.55	573.00
MAE	41.46	38.34	15.53

The reason behind high error values for ARIMA and LSTM is that the dataset it is trained on has only 350-400 values (as every month only limited sensor values can be updated on the Google Sheet). However, Random Forest has a significantly lesser error value as the n-estimators parameter, representing the number of decision trees, given to the model for training is 1000, that is, irrespective of the number of data points in the dataset.

The scatter plot and line graph in Figure 9 and 10 are for Random Forest and ARIMA models respectively. It is evident from the scatter plot that the predicted data points (green) are in close proximity to the expected data points (red), hence giving the model a lower error score. The plot in Figure 10 has larger differences in the range of 60-80 data points, hence resulting in larger error scores as shown in Table I.

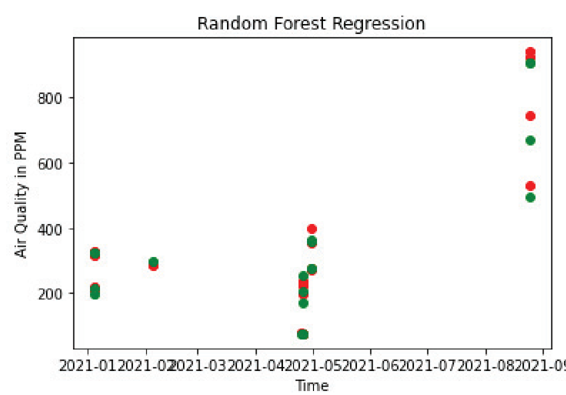

Fig. 9. Random Forest Regression: Scatter Plot for Predicted (green) vs Expected (red) Values

A. *Expected Long Term Results*

- Better monitoring of resources: The underlying intention of this study was to better equip people living in households to track their consumption of vital resources present in their homes.
- Accurate prediction of future consumption: If the prediction of future consumption of resources is accurate,

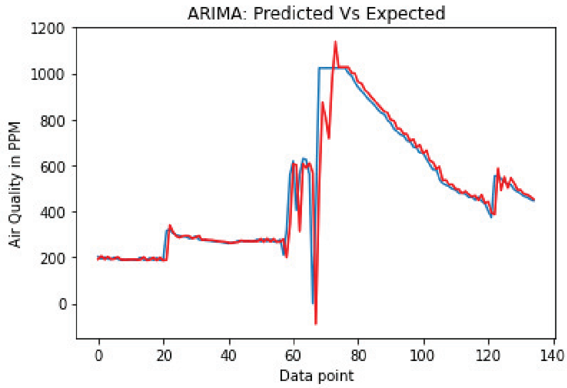

Fig. 10. ARIMA: Predicted (green) vs Expected (blue) Values

it provides a much better picture of their consumption and would make people in that household more conscious about their consumption.

- Dynamic updating of sensor values at regular intervals: In order for the results to be accurate over a long period of time, it is important to ensure that the sensor values taken from the appropriate sensors in a particular household, are updated accurately and at equal intervals into the google sheet and are replenished when necessary.

VI. Conclusion

The system developed is capable of providing updates about changes in air, water quality and electricity usage in real-time. This will further help in not only planning the mitigation of deteriorating resources of air and water quality but also help in keeping these in check and taking measures to avoid the drop in quality. The chosen model of Random forest regression gives optimal results with a Root Mean Squared Error(RMSE) of 23.94, making it suitable for forecasting the values of PPM and NTU for air and water, respectively. Using the comparisons based on error scores and examining the plots, Random Forest when compared to the other models of ARIMA and LSTM, has performed the best on the test data.

Acknowledgment

This research work is funded by IBM Shared University Research grant. The authors gratefully acknowledge the support of IBM and thank them for their able guidance in executing the research work.

References

[1] G. Cotella, S. Crivello, and M. Karatayev, "Chapter 2 - european union energy policy evolutionary patterns," in *Low-carbon Energy Security from a European Perspective*, P. Lombardi and M. Gruenig, Eds. Oxford: Academic Press, 2016, pp. 13–42. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128029701000024

[2] "The importance of energy monitoring." [Online]. Available: https://www.em-solutions.co.uk/insights/the-importance-of-energy-monitoring/

[3] K. Ashna and S. N. George, "Gsm based automatic energy meter reading system with instant billing," *2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)*, 2013.

[4] S. Bourhnane, M. Abid, and R. Lghoul, "Machine learning for energy consumption prediction and scheduling in smart buildings," *SN Appl. Sci*, vol. 2, p. 297.

[5] N. Hossein Motlagh, M. Mohammadrezaei, J. Hunt, and B. Zakeri, "Internet of things (iot) and the energy sector," *Energies*, vol. 13, no. 2, 2020. [Online]. Available: https://www.mdpi.com/1996-1073/13/2/494

[6] E. García-Martín, C. F. Rodrigues, G. Riley, and H. Grahn, "Estimation of energy consumption in machine learning," *Journal of Parallel and Distributed Computing*, vol. 134, pp. 75–88, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0743731518308773

[7] M. Islam, S. Sarker, M. Mazumder, and M. Ranim, "An iot based real-time low cost smart energy meter monitoring system using android application," 01 2020.

[8] M. Kumar, A. F. Minai, A. A. Khan, and S. Kumar, "Iot based energy management system for smart grid," in *2020 International Conference on Advances in Computing, Communication Materials (ICACCM)*, 2020, pp. 121–125.

[9] P. W. Khan, D.-H. Kang, J.-Y. Kang, H.-S. Park, Y. Byun, and S.-J. Lee, "Machine learning-based approach to predict energy consumption of renewable and nonrenewable power sources," *Energies*, vol. 13, p. 4870, 09 2020.

[10] J. M. M. Arce and E. Q. B. Macabebe, "Real-time power consumption monitoring and forecasting using regression techniques and machine learning algorithms," in *2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)*, 2019, pp. 135–140.

[11] C. Olah, "Understanding lstm networks," Aug 2017. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[12] J. Brownlee, "How to create an arima model for time series forecasting in python," Dec 2020. [Online]. Available: https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/

[13] T. S. Srivastava, "Random forest parameter tuning: Tuning random forest," Jun 2020. [Online]. Available: https://www.analyticsvidhya.com/blog/2015/06/tuning-random-forest-model/