A Scheme for Predicting Energy Consumption in Smart Cities Using Machine Learning

Noof Awad Alghamdi
Faculty of Computer and Information Technology, University of Tabuk, Saudi Arabia.
E-mail: 421009774@stu.ut.edu.sa

Israa Mohammed Budayr
Faculty of Computer and Information Technology, University of Tabuk, Saudi Arabia.
E-mail: 421009749@stu.ut.edu.sa

Samar Mohammed Aljehani
Faculty of Computer and Information Technology, University of Tabuk, Saudi Arabia.
E-mail: 421009752@stu.ut.edu.sa

Majed Mohammed Aborokbah
Faculty of Computer and Information Technology, University of Tabuk, Saudi Arabia.
E-mail: m.aborokbah@ut.edu.sa

Received September 18, 2021; Accepted December 16, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V19I1/WEB19230

Abstract

Fluctuating result on weather condition throughout several decades became a global concern due to the direct or indirect effect on energy consumption, and that was well-defined in several sector. Research investigates the use of technology and the speed of obtaining information ، which helps in decision-making. This paper Emphasize the role of data science and their application to monitoring energy consumption, also, explain the importance used and challenges of Internet of Things (IoT). Thus, there is a global concern on data transformation from IoT devices when taking into account deferent weather variations. Cities are a critical part when of energy management, it presents the effect of urbanization and some of the success achievement in several cities around the world. Our Analysis indicate that three dissimilar types of sensors can detect massive amount of information up to four hundred thousand rows, compared to traditional methods for collecting data. The results depict the resilient of IOT performance which provide an aggregate of measures reach around 405,184 rows in a record time, with achieved accuracy up to 99% when implementing the decision tree algorithm, the outcome after applying the algorithm was vary 27.60 per-cent recorded by the first device while the other devices scored 26.14%,46.26% respectively, throughout different circumstances with continuous reading in a short period of times around 8 days.

Keywords

Data Science, Decision Tree, Energy Consumption, Random forests, Smart Cities.

Introduction

Climate change and global warming have a significant impact, the threats by the emissions increased gradually (M. Molina-Solana, 2017). Research has confirmed that global temperatures are rising at an alarming rate from 1.5 to 2 degrees Celsius in the decade, which means that with this pace of global warming, weather change cannot be checked in the near future (D. Pinner, 2020). Nevertheless, its severe impacts can be minimized by the use of technology on a large scale (M. Molina-Solana, 2017).

The recent advancements in Data Science under the aegis of Information and Communications Technology (ICT) has made it possible to extract useful data from (M. Molina-Solana, 2017) diversified energy research and (D. Pinner, 2020) ML algorithms (M. Molina-Solana, 2017).

This has created hopes in finding solutions that would (M. Molina-Solana, 2017) reduce energy consumption, (D. Pinner, 2020) optimize energy management, and (D. Pinner, 2020) offer innovative means to avoid the use of radiations – a leading factor causing accumulation of greenhouse gases in the atmosphere. Among one such solution is to build smart cities whose dynamics would be driven by eco-friendly, green energy.

Building smart cities depends on the stakeholders and decision makers which varies from region to region. It is also energized by the economy of the country's leading to a big question of affordability.

This prompts a question: are there alternatives to technology-driven costlier smart cities? Since troubleshooting is always a preferred choice where one cannot reinvest in new things, modifying and reshaping the existing cities can be a better choice (Nicolas Waern, 2019). For instance, Barcelona city installed beyond than 550 sensors to obtain data from various sources, all that place this city in forefront in by replace the lighting system (S. E. Bibri, 2020).Additionally, implementing the digital technologies decreased the energy consumption and reduce the cost, which made this region mentioned frequently in vast amount of research and articles.

Thus, scientific monitoring of consumptions, analyzing the result through specific periods, then developing and testing the alternatives to reduce energy, time, and resources can be feasible options, such as those provided by IoT telemetry dataset provided by the Amazon Web Services.

Moreover, smart cities will certainly comprise of smart buildings requiring minimum human intervention, but maximum self-optimization capabilities in terms of space usage, energy consumption, waste management and overall ecological improvement. The exponential growth in IoT devices and IoT platforms, long-range transmission protocols, wireless networks, increase in the number of precision sensors and high-tech equipment, has improved the possibility of synchronized services, thereby making the concept of smart cities an undeniable reality. Our research, would be primarily centered on establishing the crucial linkages between the devices efficiency and the speed of obtaining or accessing data to achieve the desired goal which measure the speed of data transmission over different weather condition such as temperature, humidity, LPG, CO, smoke and light motion.

However, instead of building high-tech environment friendly cities from scratch, undoubtedly, modern technology has the potential to open doors for remodeling the existing cities and empower humanity to make data specific decisions accurately and briskly. Replacing all the traditional modes of energy production by setting up powerhouses based on solar energy, wind energy, hydro or wave energy, and hybrid energy can not only solve the environmental issue of climate change but also increase the efficiency and productivity. City-design policies can be made energy-efficient by tracking the data of energy supplies and consumption, leading to further efficient solutions. This will encourage transforming traditional cities into smart cities securely and economically.

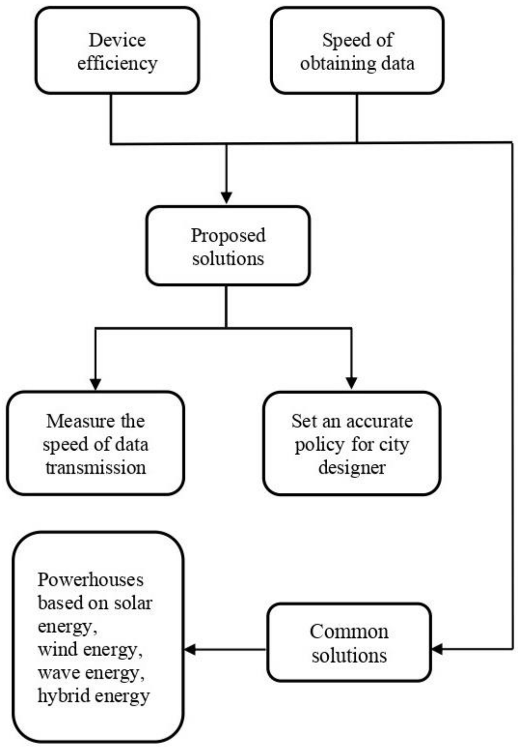

Fig. 1 The efficiency of sensors and their speed to collect the required information lead to two proposed solutions beside the common solutions.

The remainder of this paper is organized as follows. Section II present the methodology implemented to achieve the desired goal and the dataset used in this study. Section III introduces background of the domain including data science, energy consumption, smart cities and IoT. Section IV describes the proposed solution and concerns on applying machine learning algorithms. Section V introduces the data analysis techniques and algorithms used during this study. Section VI describes the results after running the machine learning algorithms. Section VII concludes the research and shows the feature works to be done.

Methodology

This section shows the research method to be utilized in this paper and the source of the dataset to be used in the rest of the article. It also highlights the machine learning algorithms used to predict the speed of transmission of date.

In this research, The data were extracted from Environmental Sensor Telemetry Data that was made by Gary A. Stafford, who is an employee at Amazon Web Services (AWS) were used in this study. Sensor data includes the following features: temperature, humidity, carbon monoxide (CO), liquid petroleum gas (LPG), smoke, light, motion, timestamp, and the unique device ID.

The dataset explored how AWS IoT a analytics could be used to analyze environmental sensor data from a variety of IoT devices in near real-time. Each sensor array was connected to a Raspberry Pi single-board computer (SBC), a popular low-cost Linux computer the size of a credit card.

The IoT devices were effectively placed in locations with different temperature, humidity, and other environmental factors.

The sensor readings were published as a single message, using the ISO standard Message Queuing Telemetry Transport (MQTT) network protocol (Gary A. Stafford, 2020).There were three device IDs that express three places with different environmental conditions as follows:

- Device 1 measures the stable conditions, cooler and more humid.
- Device 2 measures the highly variable temperature and humidity.
- Device 3 measures the stable conditions, warmer and dryer.

To evaluate our proposed concept, different prediction methods can be applied to the dataset, among those methods is the decision tree. A decision tree is a tree structure that looks like a flowchart, with an internal node representing a function or attribute, a branch representing a decision rule, and each leaf node representing the outcome. The root node is the topmost node in a decision tree. It learns to partition based on the value of an attribute. It recursively partitions the tree, which is known as recursive partitioning. This flowchart-like structure helps you in decision making (Z. Yu, 2010).

Background

A. Data Science

Data Science comprises a set of techniques and tools which pursue different goals and some of the most popular techniques are classification, clustering, regression, and association rule mining. These techniques have proven efficient in addressing many problems in the Energy Management industries. Monitoring and optimizing energy consumption is considered as one of the main features of smart cities, therefore a lot of researches and work are applied with the help of data science and machine learning algorithms (Z. Ullah, 2020).

Effective ways for boosting energy conservation and enhancing energy efficiency can be provided through the use of energy big data to analyze the behavior of energy consumption. Consequently, this analysis of big data can be useful to generate understandable sights and new observations that help to make decisions that reduce energy consumption and improve energy efficiency (K. Zhou, 2016).

B. Energy Consumption

The need of reducing energy consumption is increasing as the world energy reserves depletion and environmental pollution are a real global concern. And substantial efforts are being put into the field of energy consumption prediction. According to (K. Amasyali, 2018), the process of developing building energy consumption predictors has two approaches to follow. First, the physical modeling approach depends on engineering methods where physical principles are used to calculate some environmental parameters. And Second, the data-driven approach where the models of building energy consumption prediction are developed using machine learning algorithms based on inserted data.

The global concern about energy sustainability increases the need for smart cities which have substantial requirements and challenges. Where no energy sustainability can be reached without reducing energy consumption, much work has been done to use data science and different technologies to bring intelligent energy management platforms for many real-life subsystems. For instance, the proposed solution in (M. V. Moreno, 2014) presents the analyzed role of buildings in increasing energy consumption and the impact of buildings on energy performance at the city level. The proposal also shows the importance of energy monitoring and providing users with energy consumption information from different subsystems to achieve energy efficiency in buildings.

Intelligent energy management systems can be fed with data that comes from physical places to be analyzed to provide appropriate information that supports users to achieve energy efficiency (M. Castro, 2013).

C. Internet of Things (IOT)

“IoT systems are expected to become much more prevalent in the coming years.” (W. Yaïci, 2020).

The IoT is a network of linked sensors, it contains multiple devices that can communicate with each other over the internet to exchange and transmit information by using unique identification (A.K. Gupta, 2019). Also, it used to automate and simplify various aspects of daily human life, as well as to assist in making informed algorithmic decisions to ensure that various tasks are completed in the most efficient way as possible (W. Yaïci, 2020). According to optimistic forecasts, the IoT world will exceed 26 billion connected devices starting from 2020 (S.K. Gupta, 2020).

IoT is not restricted to a single industry, it has developed applications that enable users to convert non-smart devices into smart devices and access them through the Internet. Moreover, Small-scale applications have made a major contribution to large-scale applications such as waste management, agriculture, traffic management and e- commerce (L. Pawar, 2019) (S. K. Vishwakarma, 2019). It transforms the house into a smart house and gives you a more secure way to manage home appliances (L. Pawar, 2019).

The structure of the IoT is vary and their sizes starting from small number of sensors to reach a million devices. However, each device contains three components which are hardware, software and application layers (W. Yaïci, 2020).

There are enormous number of challenges in implementing the applications including: connection issue due to a massive number of devices, security concerns and heterogeneity in data due to the many amount of protocols (S.K. Vishwakarma, 2019). Despite doubts about data security for both network and end-user, manufacturers have ignored security measures due to the rapid growth of IoT devices and the goal of energy efficiency (S.K. Gupta, 2020). Nowadays, security in IoT has to be upgraded to secure the huge quantity of projects from any damage. The rapid data that generated from IoT have to be used in efficient manner (B. W. Wirtz, 2020).

D. Smart Cities

According to a considerable research and studies about current urbanization and the human daily lifestyle in the current era, individuals are lean to consume and generate more emission by practicing behaviours that are not environmentally friendly (E. Okai, 2018).

Rearrange a traditional cities approach to an innovative concept is urgently needed on this stage due to the remarkable development in technology. Therefore, going to smart cities is one of the critical solutions (B. Zhao, 2020).

Smart cities can be described as places where various industries are integrated and transformed with an emphasis on economic and sustainable growth. Moreover, it is a movement to make cities more resilient, safe, and workable by using smart technology like Artificial Intelligence (AI) among others to transform urbanity through structural reform, with a concentrate on transportation, government, citizens, economy, climate, and living ecosystems (M. Mundada, 2020).

Some of the world's smartest cities are Seattle, Helsinki, Barcelona, Songdo, Milton Keynes (E. Okai, 2018).

The overarching to adopt the idea of smart cities is to enhance the quality of life throughout effective use of smart technology (C. Turhan, 2019).

Smart cities offer solutions to the current and potential problems that rising. By 2050, the United Nations (UN) predicts that there will be about 6 million people living in cities. This puts a huge strain on cities to fulfill their residents' current needs (E. Okai, 2018).

According to Cisco, national network traffic will hit 77 exabytes per month in the upcoming year2022, which mean seven times that of 2017, posing a difficult test for network power systems (B. Zhao, 2020).

Proposed Solution

Many aspects support the idea of smart city by following the standard that provided by ISO 50001, power industry must incorporate a smart energy management system, and the innovative management system should be digitized (O. Laayati, 2020). In devising the smart city infrastructure, it is extremely important to connect the smart devices together and that need a clear framework depends in each region's situations (B. W. Wirtz, 2020). Also, it is important to enhance energy consumption by managing it with an appropriate technique (E. Okai, 2018) such as IoT. Offloading activities to servers not connected to the devices, by implementing emergence of computation offloading (B. Zhao, 2020).

Implementing these recommendations may contribute to reduce greenhouse gas emissions from 10 to 15 per-cent based on one of McKinsey report. Although, pollution tracking system in Beijing, in less than a year, airborne contaminants were decreased by approximately 20% (C. Turhan, 2019). However, enlarged green space in the urban area has been shown positive health effects, including a drop in psychological stress and stress related to illnesses (J. J. Roe, 2013), as well as cognitive and social growth among children (E. Amoly, 2014).

Some of obstacles that smart cities might face include lack of funding and professionals with specific skill sets who are needed to do the tasks now and in the future, which is challenge (E. Okai, 2018). Correspondingly, The protection and privacy aspects of smart services must be discussed by taking into account the laws and regulations of the city in which they are deployed (B. W. Wirtz, 2020). Another concerns that smart cities face are: Energy usage and resource management. These problems would have a huge effect on the quality of services delivered by smart cities (B. Zhao, 2020).

Machine learning is a component of artificial intelligence. A system in a changing environment should be able to learn in order to be clever. Machine learning is the process of programming computers to maximize a performance criterion based on previous experience or examples of data. We have a model specified up to a several parameters, and learning is the process of running a computer program to enhance the model's parameters using training data or previous experience. The model may be predictive or descriptive, or both (E. Alpaydin, 2020). Linear Regression (LR), Support Vector Regression (SVR), Decision Tree (DT) and Random Forest Regression (RFR) are examples of machine learning algorithms (I. El Guabassi, 2021).

There are effective algorithms used in this paper which are decision tree and Random forests. A decision tree is a supervised learning hierarchy model in which the local area is identified through a series of recursive splits in fewer steps. A decision tree is generated from internal decision nodes and final papers.

Most often, decision trees are used for classification rather than regression. They are extremely popular: They are quick to learn and respond, as well as accurate in a variety of areas.

The hierarchical placement of decisions allows for quick localization of the input area. In the best cases, each decision eliminates half of the case if the decision is binary. The interpretability is another advantage of the decision tree (E. Alpaydin, 2020).

Random Forest algorithm, which is derived from decision tree classifier, is an assembled method that grows trees to full size and without pruning using the CART (classification and regression trees) technique. Random forests make decisions by counting component predictors' votes on each class and then selecting the winner class based on the number of votes it receives (B. Yang, 2008). Also, it produces a high level of accuracy by combining several trees in the training data (K. Nugroho, 2019).

Data Analysis

The following approach is used to construct a reliable classification model. The technique is divided into five steps: Collecting relevant features of the topic of the research, preparing data, building the model, evaluating the model using one of the assessment tools, and finally using the model for future prediction (R. Rajalakshmi, 2018).

The process of creating a tree starts with a very simple question of the form, $x \leq d$? that divides the root node into binary nodes. The variables in the data set are represented by x, and d is a real number. Thus, CART (classification and regression trees) implements a computer-intensive algorithm that searches for the best possible split. Binary recursive partitioning is the method that CART employs for tree construction. The following is the tree-building process:

Step 1: CART splits the first variable at all of its potential split points, that is, at all of the values the variable takes in the sample. The sample splits into binary or two child nodes at each possible split point of a variable. Cases that answer "yes" to the question are sent to the left node, while those that answer "no" are sent to the right node.
Step 2: CART then evaluates the reduction in impurity achieved by applying its goodness-of-split criteria to each split point, using the formula:

$$\Delta i(s, t) = i(t) \text{ -}pL\ [i(tL)\text{- } pR\ [I\ (tR)]$$

where s denotes a specific split, pL denotes the proportion of observations at node t that go to the left child node tL, and pR denotes the proportion of observations at node t that go to the right child node tR. The impurities of the left and right nodes, respectively, are i(tL) and i(pR).

Step 3: CART chooses the best split of the variable as the one with the greatest reduction in impurity. At the root node, the three steps above are repeated for each of the remaining variables.
Step 4: CART then ranks all the best splits on each variable by the amount of impurity reduction achieved by each split and chooses the variable and its split point that reduces the root or parent node's impurity the most.
Step 5: CART then assigns these nodes to classes based on a law that lowers the cost of misclassification. CART has a built-in algorithm that considers user-defined variable misclassification costs during the splitting process. The default cost of misclassification is unit or equivalent. The CART procedure is recursive, so steps 1 through 5 are repeated for each non-terminal child node at each point.

When there is only one observation in each of the child nodes, CART avoids splitting.

Following these steps, a CART algorithm-based decision tree will be created (B. Yang, 2008).

Algorithm 1: Decision Tree Induction
Input: A data set (D) with attribute values **Output**: A decision tree 1: Create a root node of the tree 2: **if all** instances belong to the same class, then   Return the node with that class 3: **for** each attribute x in D **do** 4: compute $x \leq d$ question for attribute x 5: Split into internal node, for answer "yes" send to the left node ,for answer "no" send to the right node 6: select the best split for an attribute's value to provide the minimum value 7: **end** for 8: **return** Tree

In the following, we will import libraries and loading data, perform some data analysis, divide the data into training and testing sets, and finally build the algorithms. Environmental Sensor Telemetry Data is the name of the data that we used from Kaggle.

A. Importing Libraries

The script below imports the required libraries:

```
import numpy as np
import pandas as pd
from sklearn import tree
```

```
clf = tree.DecisionTreeClassifier()
clf = clf.fit(X_train,y_train)
```

B. Loading Data

we Loaded the required dataset using pandas' read CSV function. Also, the timestamp feature was dropped.

```
clf = RandomForestClassifier(n_estimators=10)
clf = clf.fit(X_train, y_train)
```

```
df = pd.read_csv("iot_telemetry_data.csv", header = 0)
df = df.drop(['ts'], axis=1)
df.head(10)
```

The output will look like this:

	device	co	humidity	light	lpg	motion	smoke	temp
0	b8:27:eb:bf:9d:51	0.004956	51.000000	False	0.007651	False	0.020411	22.700000
1	00:0f:00:70:91:0a	0.002840	76.000000	False	0.005114	False	0.013275	19.700001
2	b8:27:eb:bf:9d:51	0.004976	50.900000	False	0.007673	False	0.020475	22.600000
3	1c:bf:ce:15:ec:4d	0.004403	76.800003	True	0.007023	False	0.018628	27.000000
4	b8:27:eb:bf:9d:51	0.004967	50.900000	False	0.007664	False	0.020448	22.600000
5	1c:bf:ce:15:ec:4d	0.004391	77.900002	True	0.007009	False	0.018589	27.000000

C. Converting Categorical Data to Numerical Data

The pandas and scikit-learn python tools have a number of methods for transforming categorical data into numeric values. we used the map() function for ('light', 'motion'). We also transformed device names to numbers to simplify the process.

```
d = {True: 1, False: 0}
df['light'] = df['light'].map(d)
df['motion'] = df['motion'].map(d)

d = {'00:0f:00:70:91:0a': 1, '1c:bf:ce:15:ec:4d': 2,\
     'b8:27:eb:bf:9d:51': 3}
df['device'] = df['device'].map(d)
df.head()
```

	device	co	humidity	light	lpg	motion	smoke	temp
0	3	0.004956	51.000000	0	0.007651	0	0.020411	22.700000
1	1	0.002840	76.000000	0	0.005114	0	0.013275	19.700001
2	3	0.004976	50.900000	0	0.007673	0	0.020475	22.600000
3	2	0.004403	76.800003	1	0.007023	0	0.018628	27.000000
4	3	0.004967	50.900000	0	0.007664	0	0.020448	22.600000

D. Dividing the Dataset

We determined the Device ID feature, which represents the environmental conditions, as the target. Then, we used the train_test_split() to split the data as 70% for training, about 283628 data points, and 30% for testing.

```
y = df["device"]
X = df[features]
```

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X,y, test_size = 0.3, random_state = 0)
```

E. Building Decision Tree Model and the Random Forest Model

The final step is to train the decision tree algorithm and the random forest model on our data and make predictions after they have been separated into training and testing sets.

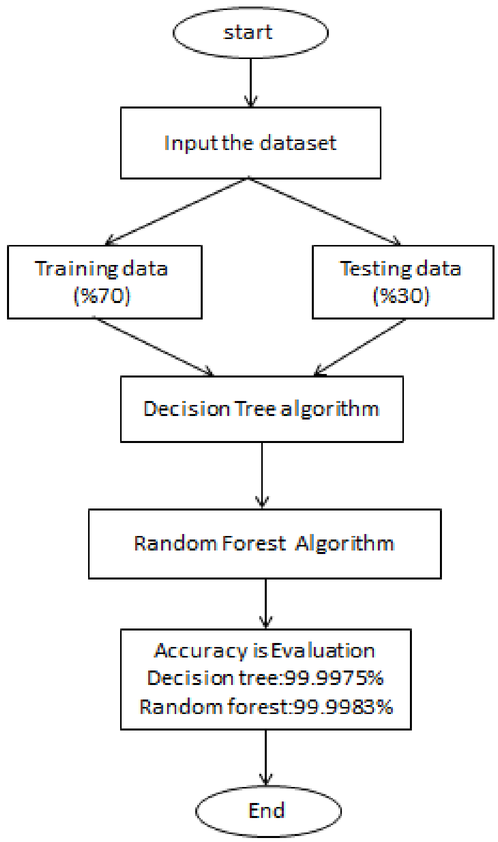

Fig. 2 Random data extraction. After importing the used dataset, it gets split into 70% for training and 30% for testing. The decision tree and random forest algorithms are applied and then the accuracy is evaluated for each classifier algorithm

Results

In this section, we represent the experimental results of the proposed system in terms of sensor readings, the visualization and evaluate the algorithm's performance on our dataset.

The used dataset contains 405,184 data points of sensor readings, unique device ID, and timestamp collected from three devices during 192 consecutive hours. The number of recorded data points from the first device, second device, and the third device is 111815, 105918, and 187451 respectively.

To visualize the sensor readings used in the designed system, the following figures illustrate the relation between some of the measured data against time.

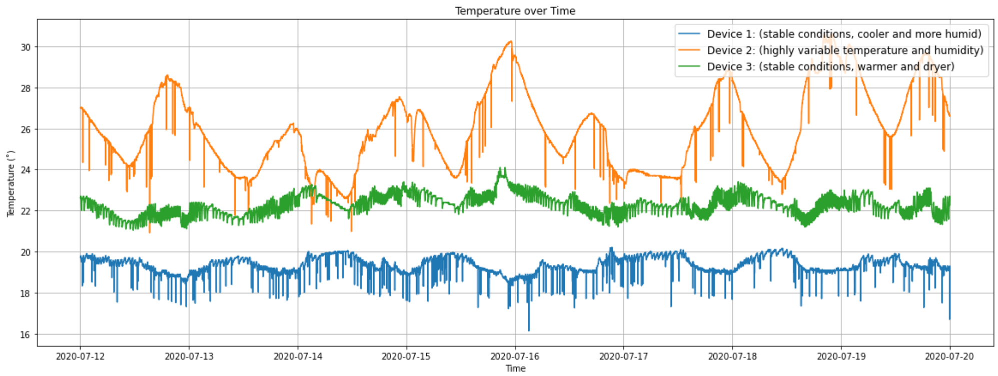

Fig. (a) Display the temperature feature against time

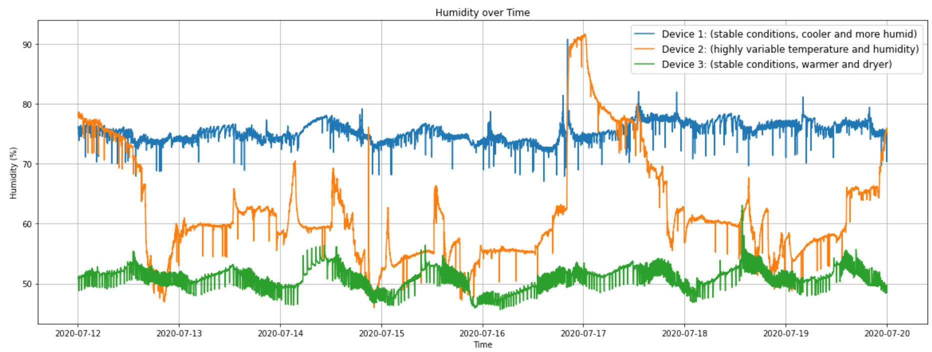

Fig. (b) Display the humidity feature against time

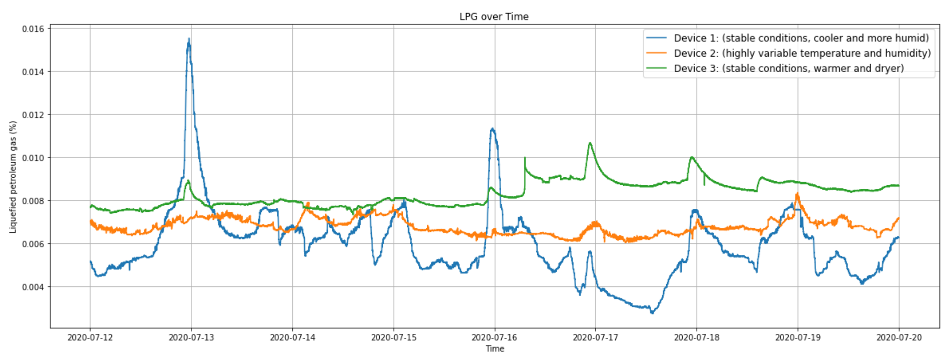

Fig. (c) Display the LPG feature against time

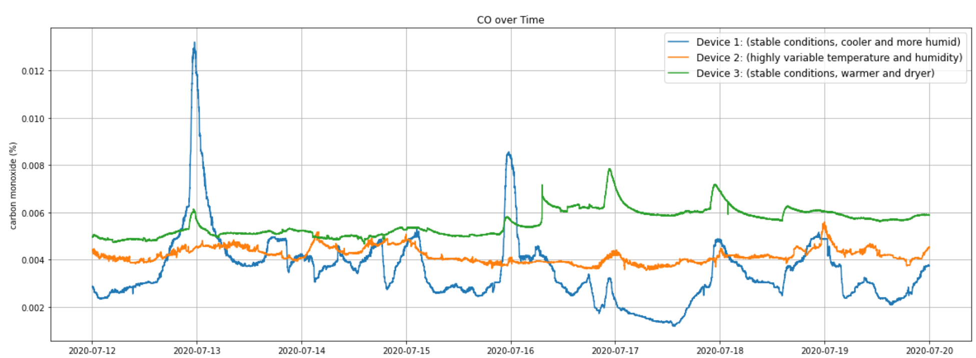

Fig. (d) Display the CO feature against time

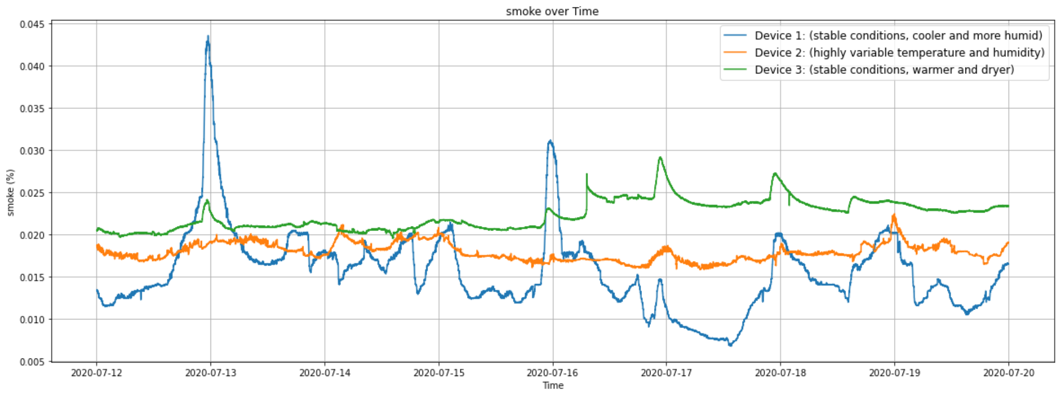

Fig. (e) Display the smoke feature against time

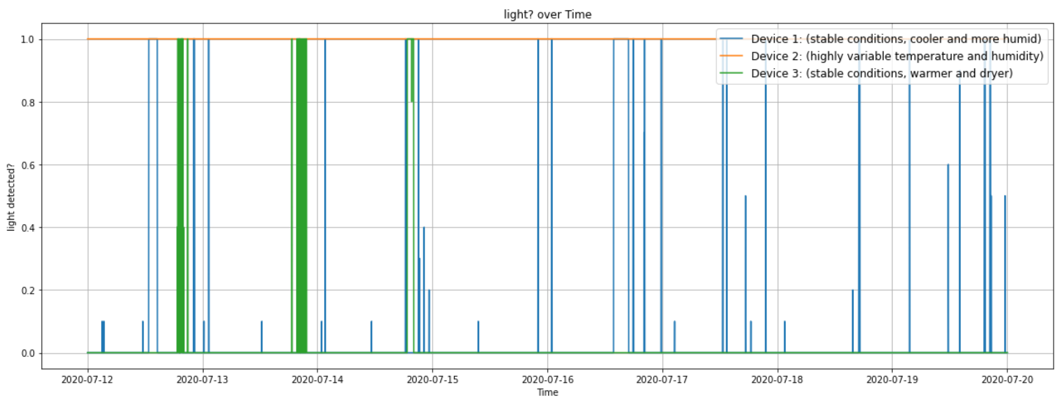

Fig. (f) Display the light feature against time

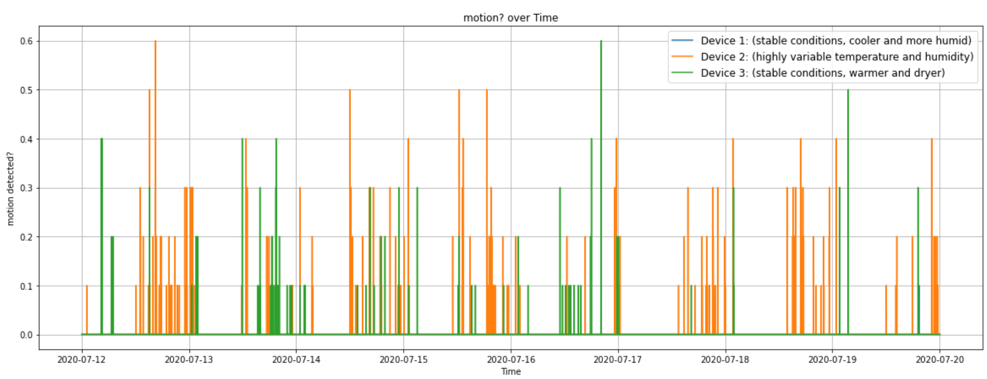

Fig. (g) Display the motion feature against time

Fig. 3 The 7 features of the sensors' readings against time (a) temperature over time (b) humidity over time (c) LPG over time (d) CO over time (e) smoke over time (f) light over time (g) motion over time

The result of the classification shows that the accuracy of the decision tree classifier exceeds 99.9975% and it gets a tiny improvement when the random forest model is used where the model gives an accuracy of 99.9983%.

Conclusion & Future Work

Building smart cities represents a technology-based solution that considered as an impactful solution from many suggested ones that have different requirements. And the need for effective energy management that reduces the cost of building smart cities is increasing every day as the world struggles with environmental problems such as climate changing and rapid waste of resources.

The energy consumption rate which gets increased rapidly can be monitored and reduced by the effective use of data science as this paper shows. This paper presents a way of building a predictive model for environmental conditions based on environmental sensor telemetry Data. The result of our research proves how collected environmental sensor data can be analyzed through a machine learning approach to be used in energy management for efficiency optimization. The results emphasize the significant role of Data Sciences in achieving effective energy management.

Different algorithms can be used built to build classification models that help to optimize the efficiency of energy management. In this paper, the Decision Tree and Forest Random Models where they get excellent results for predicting environmental conditions. While IoT technology has contributed in providing the ability to build real-time technical solutions, the on-board environmental sensors make it possible to the environmental data to be collected accurately in real-time.

The use of data science and information technology in energy management industry represents an effective solution to reduce the energy consumptions and guarantee better outcomes in human life with the absence of effective environmental protection.

As a future work this research suggest to implement different algorithms and use other datasets. In addition, compare the current outcomes with other result within the domain to achieve accuracy.

References

Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., & Martín-Bautista, M. J. (2017). Data science for building energy management: A review. Renewable and Sustainable Energy Reviews, 70, 598-609.

Pinner, D., Rogers, M., & Samandari, H. (2020). Addressing climate change in a post-pandemic world. McKinsey Quarterly April.

A Recipe to Create Smarter Buildings?

https://www.automatedbuildings.com/news/sep19/articles/waern/190828124505waern.html.

Bibri, S.E., & Krogstie, J. (2020). The emerging data–driven Smart City and its innovative applied solutions for sustainability: The cases of London and Barcelona. *Energy Informatics, 3*(1), 1-42.

From concept to applied solutions. Data-driven cities.

Environmental Sensor Telemetry Data. kaggle datasets download -d garystafford/environmental-sensor-data-132k.

(Jul 16,). *Getting Started with IoT Analytics on AWS*. https://towardsdatascience.com/getting-started-with-iot-analytics-on-aws-5f2093bcf704.

Yu, Z., Haghighat, F., Fung, B.C., & Yoshino, H. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings, 42*(10), 1637-1646.

Ullah, Z., Al-Turjman, F., Mostarda, L., & Gagliardi, R. (2020). Applications of artificial intelligence and machine learning in smart cities. *Computer Communications, 154,* 313-323.

Zhou, K., & Yang, S. (2016). Understanding household energy consumption behavior: The contribution of energy big data analytics. *Renewable and Sustainable Energy Reviews, 56,* 810-819.

Amasyali, K., & El-Gohary, N.M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews, 81,* 1192-1205.

Moreno, M.V., Zamora, M.A., & Skarmeta, A.F. (2014). User-centric smart buildings for energy sustainable smart cities. *Transactions on emerging telecommunications technologies, 25*(1), 41-55.

Castro, M., Jara, A.J., & Skarmeta, A.F. (2013). Smart lighting solutions for smart cities. *In 27th International Conference on Advanced Information Networking and Applications Workshops,* 1374-1379.

Yaïci, W., Krishnamurthy, K., Entchev, E., & Longo, M. (2020). Survey of internet of things (IoT) infrastructures for building energy systems. *In Global Internet of Things Summit (GIoTS),* 1-6.

Gupta, A.K., & Johari, R. (2019). IOT based electrical device surveillance and control system. *In 4th international conference on internet of things: Smart innovation and usages (IoT-SIU),* 1-5.

Gupta, S.K., Vanjale, S., Rasal, S., & Vanjale, M. (2020). Securing IoT Devices in Smart City Environments. *In International Conference on Emerging Smart Computing and Informatics (ESCI),* 119-123.

Pawar, L., Bajaj, R., Singh, J., & Yadav, V. (2019). Smart city IoT: Smart architectural solution for networking, congestion and heterogeneity. *In International Conference on Intelligent Computing and Control Systems (ICCS),* 124-129.

Vishwakarma, S.K., Upadhyaya, P., Kumari, B., & Mishra, A.K. (2019). Smart energy efficient home automation system using iot. *In 4th international conference on internet of things: Smart innovation and usages (IoT-SIU),* 1-4.

Wirtz, B.W., Müller, W.M., & Schmidt, F. (2020). Public smart service provision in smart cities: A case-study-based approach. *International Journal of Public Administration*, *43*(6), 499-516.

Okai, E., Feng, X., & Sant, P. (2018). Smart cities survey. *In IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 1726-1730.

Zhao, B., Peng, K., Zhang, H., & Xu, X. (2020). Energy-and Time-Efficient Tasks Offloading and Dynamic Resource Allocation in Smart City. *In International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, 705-712.

Mundada, M., & Mukkamala, R.R. (2020). Smart Cities for Sustainability-An Analytical Perspective. *In Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, 770-775.

Turhan, C., Atalay, A.S., & Akkurt, G.G. (2019). Green Smart Cities: Living Healthily with Every Breath. *In 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*, 114-118.

Laayati, O., Bouzi, M., & Chebak, A. (2020). Smart energy management: Energy consumption metering, monitoring and prediction for mining industry. *In IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, 1-5.

Wirtz, B.W., Müller, W.M., & Schmidt, F. (2020). Public smart service provision in smart cities: A case-study-based approach. *International Journal of Public Administration*, *43*(6), 499-516.

Smart Cities: Digital solutions for a more livable future. https://www.mckinsey.com/business-functions/operations/our-insights/smart-cities-digital-solutions-for-a-more-livable-future#.

Roe, J.J., Thompson, C.W., Aspinall, P.A., Brewer, M.J., Duff, E.I., Miller, D., & Clow, A. (2013). Green space and stress: evidence from cortisol measures in deprived urban communities. *International journal of environmental research and public health*, *10*(9), 4086-4103.

Amoly, E., Dadvand, P., Forns, J., López-Vicente, M., Basagaña, X., Julvez, J., & Sunyer, J. (2014). Green and blue spaces and behavioral development in Barcelona schoolchildren: the BREATHE project. *Environmental health perspectives*, *122*(12), 1351-1358.

Alpaydin, E. (2020). *Introduction to machine learning.* MIT press.

El Guabassi, I., Bousalem, Z., Marah, R., & Qazdar, A. (2021). A Recommender System for Predicting Students' Admission to a Graduate Program using Machine Learning Algorithms.

Yang, B.S., Di, X., & Han, T. (2008). Random forests classifier for machine fault diagnosis. *Journal of mechanical science and technology*, *22*(9), 1716-1725.

Nugroho, K., Noersasongko, E., Fanani, A.Z., & Basuki, R.S. (2019). Improving random forest method to detect hatespeech and offensive word. *In International Conference on Information and Communications Technology (ICOIACT)*, 514-518.

Rajalakshmi, R. (2018). Analysis and Predictions on Blended Learning Readiness among Indian Students at Universities using Decision Tree Classifier in Scikit-learn Environment. *International Journal of Advanced Research in Computer Science, 9*(1).