Second Review Document

Proposing Energy-Efficient Urban Spaces using IoT and Data Visualisation

Anubhav Rakeshwar Singh Guleria
18BCE0186
7758984636
anubhavsingh.guleria2018@vitstudent.ac.in

Ranjan Nitin Pramod
18BCE0272
8529589264
nitin.ranjan2018@vitstudent.ac.in

Srinivasan Sharma
18BCE0931
6385664521
srinivasan.sharma2018@vitstudent.ac.in

Dr. Gladys Gnana Kiruba B
Assistant Professor (Senior)
9597397290
gladys@vit.ac.in

B.Tech.

in

Computer Science and Engineering

School of Computer Science & Engineering

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Introduction

Theoretical Background

Smart Cities are becoming a focus of governments all over the world. They are supposed to be a delicate balance of economics, technology and public planning. However, free spaces to build new cities is a big problem for most sovereign states, smart cities projects are replacing classical infrastructure in existent urban areas.

Additionally, urban regions alone account for 60-80 percent of the total greenhouse gas emissions on the planet and are responsible for the consumption of 80 percent of the global energy. This involves heating, ventilation, air conditioning, lighting, and other major appliances. And as per independent research conducted by the IBM, seven-tenths of the world population shall live in cities by 2050 while 50% of global energy shall be consumed by the buildings in the cities alone.

Motivation

City officials and governments across the globe have struggled so far to effectively analyse, visualise, and translate data from thousands of buildings into policy and program recommendations – partly due to the issue in logistics, partially due to economic and political constraints. Computers, however, provide not only with tools that can be used to benchmark consumption and supply but also to develop models that can compare and demonstrate impacts of energy-efficiency based improvements on the same.

Countries all over the world have also pledged to become carbon neutral or carbon negative before the end of the first half of the 21st century. And governments are increasingly trying to fit both a smart city and green energy resources like the sun and the wind in one single pipeline.

Aim of the Proposed Work

1. Using models of machine learning and statistics to visualise energy consumption of buildings in an urban setup.
2. Developing a computer-aided benchmark and baseline for energy efficiency of buildings in cities.
3. Aiding in the implementation of the following UNDP goals -

Goal 7: Affordable and Clean Energy;
Goal 8: Decent Economic Growth;
Goal 11: Sustainable Cities and Communities;
Goal 12: Responsible Consumption and Production; and Goal
13: Climate Action

Objective

This project aims to study patterns of energy consumption and related costs and to further suggest IoT-based solutions to improvise upon the same by-

1. Visualising energy consumption of buildings in an urban setup.
2. Developing a computer-aided benchmark and baseline for energy efficiency of buildings in cities.

Literature Survey

Literature Reviews and their Analysis

S. No.	Title	Authors	Year	Advantages	Disadvantages	Scope for future work
1	Machine Learning-Based Approach to Predict Energy Consumption of Renewable and Nonrenewable Power Sources	Prince Waqas Khan *et al.*	2020	Compared many load forecasting methods using ML. Another advantage is the models used are compared to hybrid models Moreover simple ML models are used such as Mean absolute error, mean squared error, mean absolute percent error, etc.	The logarithmic error is high. Very simple mathematical models are used making the readings absolute. The idea is based on statistics with no further futuristic use.	The models can be made logarithmic rather than regressive. The data is considered for only a single space, and could be made better for better regression. Could be used hybrid models as well.
2	Energy consumption prediction by using machine learning for smart building: Case study in Malaysia	Mel Keytingan M. Shapi *et al.*	2021	The data collected is for a shopping district that is a very large sample. Skewness and kurtosis values are accrued for the model making it future proof. Accuracy is very high ranging between 90-92%	The models k-NN, SVM, ANN were used in accordance with which are not compatible with each other, and were used as comparators rather than providing solutions. The data testing was normal, which reduces the accuracy of kurtosis. Model development environment is time consuming.	The project takes too much time to emulate SVM hence a faster system could make the data more sublime and ecstatic. Rather than using three different models, using hybrid versions will be better.
3	Machine learning for estimation of building energy consumption and performance: a review	Saleh Seyedzadeh *et al.*	2018	Was based on electric and solar energy. Showed the comparisons using various sensors and iot devices with energy conservation over a long time.	Here they have also used the SVM model hence increasing the computation time. GP modeling making the cost inefficient	Simpler models used due to computation limits , could be improved with better systems. Need an hybrid model for better accuracy as the number of variables increase with IoT devices and sensors.
4	Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction	Zhijian Liu *et al.*	2019	Here ANN and SVM are used and also made use of their hybrid, making improvements in previous papers Since the sample of the data is very large making data more reliable as well.	Since computation of SVM is already hard, making hybridization harder. Though the accuracy for the hybrid is high, the computation time is very slow.	The model can be introduced to fewer constraints,which could make the computation faster. Model structure although simple has many flaws, could be introduced to better structural models.
5	Improving energy consumption of commercial building with IoT	Javed *et al.*	2018	Neural networks were embedded with IoT subsystems. It predicted 68% reduced cost	IoT sub-systems used were expensive and expansive , leading to inflated results. On a minor scale the project	Simpler IoT devices could be used, as they are cost effective, and easier to install and maintain.

	and machine learning			effectiveness over a period of 10 years.	was based on a complex leading to miss aligned perpetuations in the result as well.	The models used for future prediction are simpler using random neural networks. Since smaller sensors could also be used, solar energy could also be incorporated to make profits soar even higher.
6	A Novel Method for Analysing Weather Effect on Smart City Traffic	Aram Nasser and Vilmos Simon	2021	Made use of many real life variables such as wind, sunlight and rain; This gave us the idea to implement a real life visualisation model for the project.	The models used advanced statistical models for rain variables and others.	The model could be simpler. Sensors used could be better.
7	A Systematic Survey on the use of Fuzzy Graph Structures in India's Smart City Development	B. Angel and D. Angel	2021	FGS model was used making the system future proof. Graphs make connectivity easier and vivid.	More inclined towards urban spaces only. Require precise network hence not cost effective.	Can be made more inclined to energy savings rather than quality of life. Needs to be more cost effective.
8	Exploring The Relationship Between Smart City, Sustainable Development And Innovation As A Model For Urban Economic Growth	Procopie Florin Gușul and Alina Ramona Butnariu	2021	Easy explanation of interdependence of economy and innovation. Gave better scope for our own project.	Was completely theoretical, had less mathematical models involved.	Could introduce better results. Data samples were scarce, and needed more data for future integrations.
9	Smart cities and the European Vision	Carmen Florina Fagadar *et al.*	2021	Compared many semi-urban and urban spaces of many scoops of notions including both developing and developed. Good sample space, making comparisons easy to frame and understand.	Solutions in the problems were more inclined towards colder regions; Lacked diversification in results. Was majorly theoretical; Lacked mathematical and statistical models.	Integrating ML will take the study to greater heights. Need more solutions to semi-urban and suburban regions. Lacking a practical approach hence could introduce theoretical variables for better future results.
10	Optimizing Task Allocation for Edge Micro-Clusters in Smart Cities	Yousef Al Hailey et al.	2021	The resource management for clusters is reduced to a mixed integer problem. Greedy algorithm is employed to create the most energy-efficient cluster head selection protocol. The makespan is minimised.	The scope of the project is too large to be evaluated using a quasi-realistic setup on breadboards. As a consequence, the model proposed is not generic.	Scaling up the model to be functional in IoT environments other than just Raspberry Pi would help generalize the model and implement it in more varied and realistic situations.
11	The Network Architecture	Karol Furdik et	2013	A balanced approach towards the human needs of an IoT-	The major drawback is the lack of specialisation of IoT	Creating solutions more compatible with individual

	Designed for an Adaptable IoT-based Smart Office Solution	al.		enabled office space are taken into consideration. The audit of data is possible and the result is a cost-effective solution.	services. Different IoT services are required in different rooms and spaces. This project only explores a limited set of users and rooms.	requirements.
12	Understanding Smart Cities: An Integrative Framework	Hafedh Chourabi, Taewoo Nam,Shawn Walker,J. Ramon Gil-Garcia,Sehl Mellouli,Karine Nahon,Theresa A. Pardo,Hans Jochen Scholl	2012	One of the very few papers which takes into consideration various countries at once Various challenges to the creation of a smart city such as technical,managerial,organisational etc. have been identified Rich literature has been considered Clear and concise	More visual and diagrammatic representations are required so as to address rather complex topics Lack of examples of real-world smart cities,rather countries and their current scenarios are considered	Case study of current smart cities can also be discussed More diagrams and flowcharts can be added so as to make the information more appealing to the readers
13	Conceptualising Smart City with Dimensions of Technology, People, and Institutions	Taewoo Nam & Theresa A. Pardo	2011	More focus on the fundamental building blocks of smart cities Real-world smart cities have been discussed	Different types of smart cities have been mentioned such as ubiquitous cities,hybrid cities,wired cities etc.,but no brief description is given The paper lacks in depth analysis of smart cities,it offers a more broader analysis of the same	Brief discussion about the different types of smart cities In-depth analysis of the various intricacies involved in the creation of a smart city
14	Exploring The Relationship Between Smart City, Sustainable Development And Innovation As A Model For Urban Economic Growth	Procopie Florin Gușul, Alina Ramona Butnariu	2021	Recent work makes the study more likely to chosen as a part of literature Focuses on the environmental impact of smart cities	Lack of information regarding the environmental impact of smart cities The literature hasn't been utilised properly	Proper analysis of scholarly articles can give new insights Brief discussion about the environmental impact of smart cities

15	The Network Architecture Designed for an Adaptable IoT-based Smart Office Solution	Karol Furdik et al.	2013	A balanced approach towards the human needs of an IoT-enabled office space are taken into consideration. The audit of data is possible and the result is a cost-effective solution.	The major drawback is the lack of specialisation of IoT services. Different IoT services are required in different rooms and spaces. This project only explores a limited set of users and rooms.	Creating solutions more compatible with individual requirements.
16	Using Social Network Data To Improve Planning And Design Of Smart Cities	Raquel Pérez-del hoyo , Higinio Mora & José Francisco Paredes	2018	Clear and concise The aspect studied is generally liked by readers	Lack of literature survey; Insufficient data visualisations;	Proper literature survey can be done Addition of more visualisations will give a more detailed study touch to the work
17	From Smart Cities to Human Smart Cities	Álvaro Oliveira,Margarida Campolargo	2015	Eye catching topic has been studied Not much relevant literature exists,as a result,addition of this work will greatly benefit future readers	Although challenges have been identified in this paper,the solutions to handle the same aren't present in sufficient amount; Since the My-Neighbourhood project was implemented in only 4 cities,as a result,it is difficult to believe that the same conclusions are valid for other cities as well	More focus on the methods to tackle the challenges identified Discussion about the results from different projects can be included

Primary Conclusions

1. Most of the existing research leans greatly on the theoretical end of the subject matter.
2. Most of the existing literature is very specific to the dataset chosen. And these datasets often focus on a single city or a single region alone.
3. Very little research focuses on making the products more acceptable and usable to the people en masse and give political and economic incentives to actually bring forth the research to common use.

Proposed System

Related Concepts

1. IoT - Internet of Things is an umbrella term to describe physical objects with sensors, their processing ability and related software and related technologies like communications and similar paradigms.
2. Visualisation - Visualisation is the graphical representation of data. The idea is to let the consumer grasp the information from the data purely by observation.
3. Benchmarking - It is the process of measuring data, services and related products or processes and activities against a set standard or to set such standards and protocols.
4. Time Series Analysis - It involves analysing a sequence of data points on a curve collected against a set interval of time. The most common usage is the prediction of unrecorded data points in the same or extended time frames.

Framework & Architecture

The main idea of the project is an algorithm. The implementation of the same is subject to the hardware availability with the user.

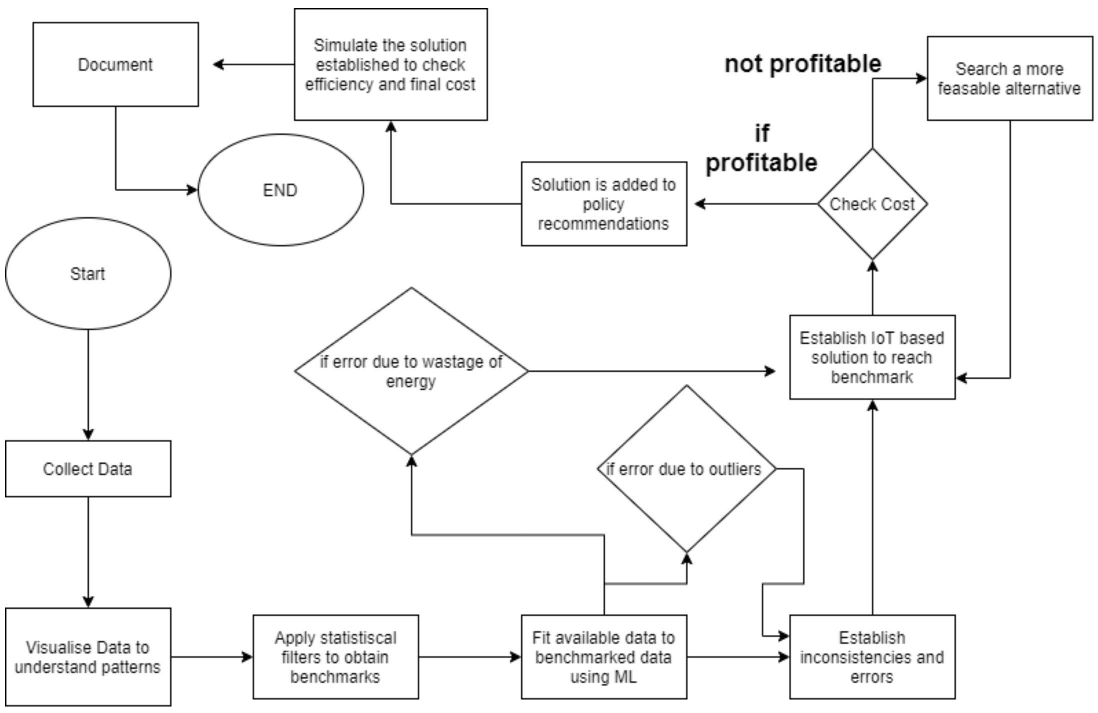

Figure 1The proposed Algorithm

Algorithm

Datasets

This report deals exclusively with the following datasets:

1. London Municipal Council: Daily Household Energy Consumption
2. Government of India: Daily Energy Consumption by state (October 2019 - May 2021)

The visualisations studied for this project report also include the following datasets:

1. IEA: natural gas consumption by region
2. IEA: Percentage Change in energy by source
3. US Gov: Primary Energy Renewable as a source
4. IEA: Share of Low Carbon Energy vs GDP PPP
5. World Coal consumption 1978-2020
6. Government of India Open-Source Data: Power Loss Across States between 2011-19
7. IEA: Petroleum and products demand: 1978-2020
8. IIT Bombay: Energy Usage in Individual Campus buildings Datasets 1-50
9. New York City Council: Hourly energy demand: 2001-2020
10. USA: Hourly Energy Demand over 20 years: PJME Incorporation
11. University of Boston: Hourly Energy Demand on Campus.

Steps involved in detail

1. Collection and visualisation of data –

The tools: Tableau and Python.

Tableau is used to create interactive dashboards in order to consolidate all data. This step is to understand –

a) The most common forms of energy used in cities.

b) The way in which energy consumption has changed over years, thus predicting the patterns of future.

c) The power consumption pattern in buildings in order to prepare benchmarks for the same.

Note: *The datasets we are using are standard government datasets when it is concerned with energy consumption and its sources. For the purpose of benchmarking real urban data, we have used datasets separately generated by IIT Bombay and University of Boston respectively. This allows the study of two different countries with completely different energy usage patterns. Both the datasets have two parts –*

i. The magnitude of phase voltages; and
ii. The power consumption with respect to each phase; over a period of 3 months.

Tableau Link: https://public.tableau.com/views/EnergyTrendsViz/CoverPage?:language=en-US&:display_count=n&:origin=viz_share_link

India Data Exploratory: https://www.kaggle.com/code/nitinr2510/indiadataexploratory

2. Statistical Analysis of Data –

Tools: Python

Step is performed to estimate measures like mean, median, mode, variance and deviation etc. in the data. The most prominent and recurrent derived values shall serve as the benchmarking criterion.

In this project report, we have considered the central tendencies - mean, median, mode followed by ML algorithms that include ARIMA, LSTM and XGBoost.

Note: Since the project aims at creating a model where energy usage habits are studied, in a real-life model of the smart city, the city council/municipal corporation shall receive all the energy usage data in real time through IoT and cloud.

London Data using LSTM and ARIMA: https://www.kaggle.com/code/nitinr2510/energy-consumption-forecast

Applying Regression on Indian Data: https://www.kaggle.com/code/nitinr2510/regression-indian-energy

3. Simulation –

Tools: Cisco Packet Tracer

Since the project is based on an urban landscape, it is important to simulate it to find if it is technologically possible. 4. Cost Estimation – It is important that the IoT tools are not only compatible in terms of technology. But it must also be economically feasible. So, this step involves making them feasible for civilians with respect to the real time prices of IoT devices and machinery in the world. The results shall then be compiled into a report and a research paper be proposed using the same.

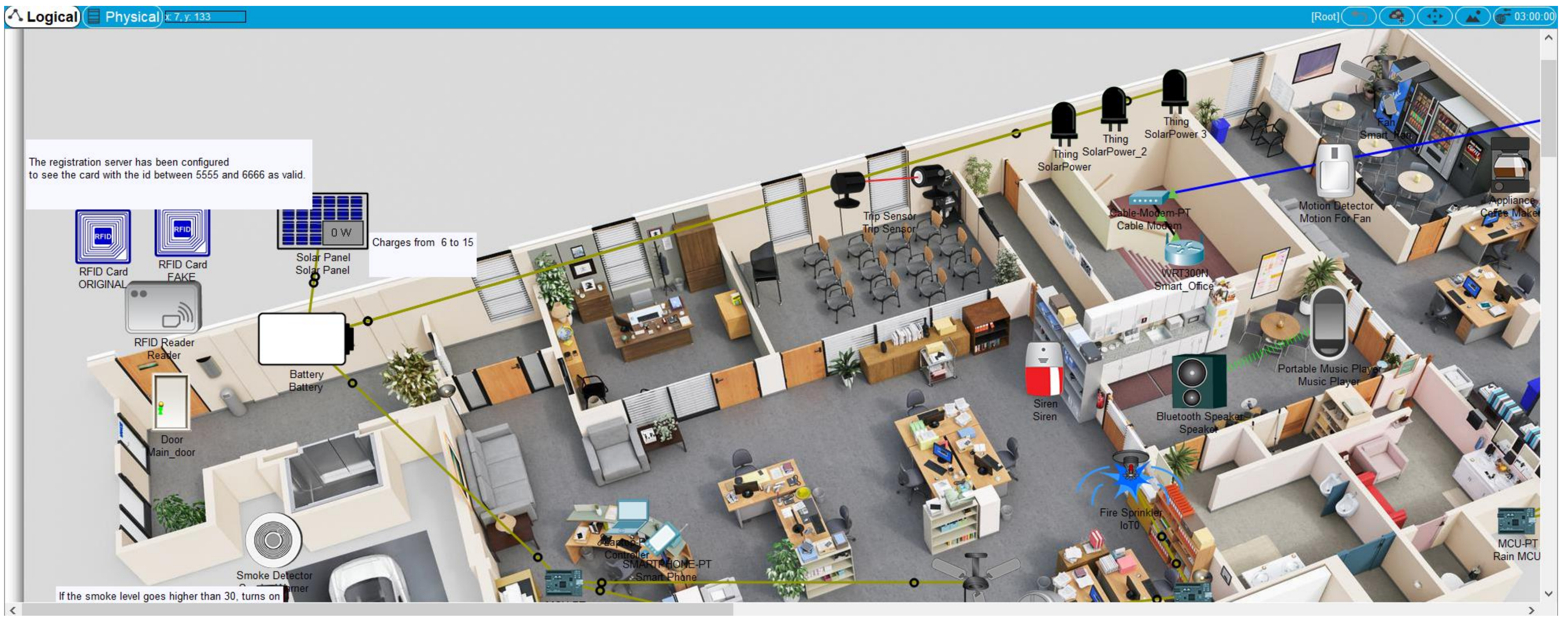

Figure 2The layout of the smart office model proposed to simulate the algorithm

Proposed System Analysis

Requirement Analysis

Functional Requirements

User characteristics

1. The primary user is expected to be a government or public administrator.
2. Common households can too use this algorithm, but the authors do not expect them to be versed with technicalities like benchmarking and visualisation.

Assumption & Dependencies

1. The user has the necessary visualisation and statistical tools.
2. The user and the hub have a communication paradigm in place.

Domain Requirements

1. A cluster head selection algorithm depending on the scale of data and nodes involved.
2. A communication channel to retrieve equipment and electricity usage report and data directly from the buildings to the government servers.

User Requirements

1. Easy to incorporate.
2. Very little breach of privacy.

Non Functional Requirements

Product Requirements

1. **Efficiency (in terms of Time and Space) -** The steps involving benchmarking and calculations are very efficient in both time and space. The datasets only need to be in the form of a simple table. The calculations do not need any special tool or hardware. However, if visualisations are to be stored, the space and time required are greater. Depending on the hardware and software requirements specified, the efficiency should be high and worthy of incorporating into administration.
2. **Reliability -** All the modelling of IoT solutions are very personalised and hence the reliability of the algorithm has only as many problems as human error can allow.
3. **Usability -** The project is specific to urban spaces like residential buildings and offices.
4. **Implementation Requirements (in terms of deployment) -** An agreement between the users and the civic authorities would be good. And a mecha person versed in basic use of computers and knows how to connect to bluetooth or WiFi would be sufficient to install devices.
5. **Engineering Standard Requirements-** A consumer cell to ensure that early grievances are met with fast action so that people adopt the technologies quickly. All devices installed must pass safety regulation.

Operational Requirements

1. **Economic -** The project is economically feasible and actually creates economic incentives for its adoption.
2. **Environment & Sustainability -** The algorithm when employed in urban spaces should provide a sustainable means of energy and cost optimisation for a long run. This in turn is also good for the environment.
3. **Social & Political -** The economic incentives created by the project in terms of long term energy cost optimisations and the long term positive effect on environment should create social and political incentives for its adoption as well.
4. **Legality -** The user’s energy usage and equipment-usage patterns shall be tracked. No other form of data is required to perform the process. So, a simple agreement between the civic authorities and users should suffice.
5. **Inspectability -** The packet tracer demonstration proves that the project components are open to audit and inspection in real time to both civic admins and the household or office owners.

System Requirements

H/W Requirements

1. A computer with RAM greater than 16GB and a harddisk with at least 256GB free storage shall be sufficient for the data of an entire city.
2. IoT devices depending on the scope and scape of installation

S/W Requirements(details about Application Specific Software)

1. Python 3.8 or above
2. R can be used too, but python is far superior in data visualisation tools.

Results and Discussion

With reference to the data for energy consumption patterns in India,

1. The mode is a very unstable factor among all the central tendencies.
2. The mean and median are fairly stable tendencies; however, median has a lower difference to the absolute value of the energy consumed and thus, the seasonal variance is lower.
3. While the nominal error was close to 30%, the actual error at the end of the time interval was only 4%.

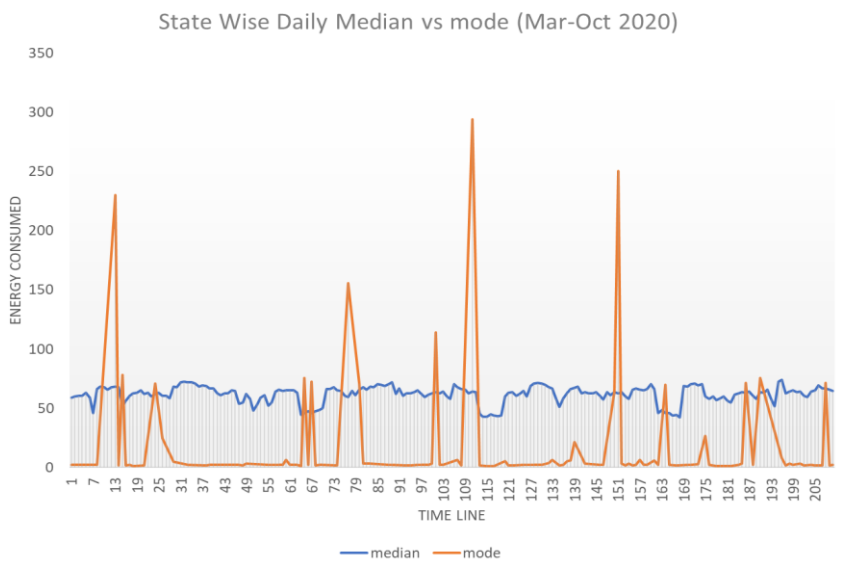

Figure 3Mode vs median for Indian Data

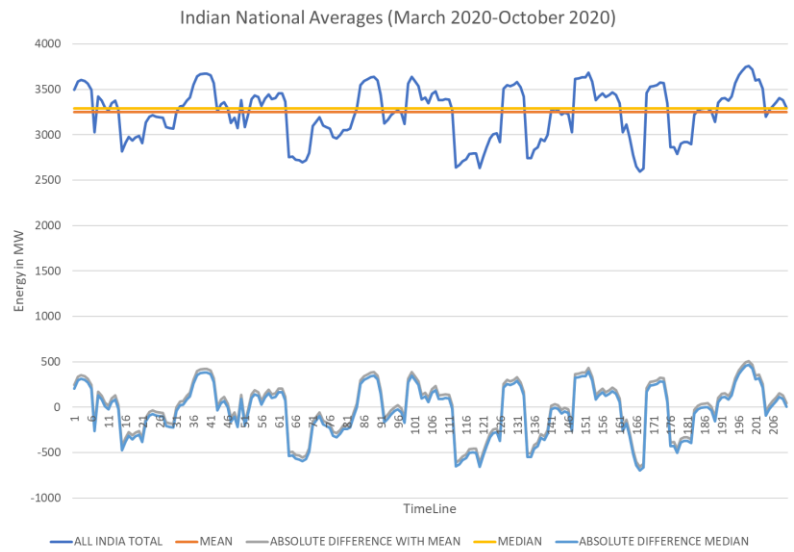

Figure 4Mean, Median and mode for Indian Energy Consumption Data

So, among all central tendencies, median is the choice for the benchmark.

For the London Data,

1. ARIMA enforced with LSTM achieves an accuracy close to 97%.
2. The computational cost of such an algorithm is very high when compared to statistical centralities.

So, at the end, median is considered to be the final benchmarking paradigm for the project report.

Finally, the simulation generated in cisco packet tracer proves that -

1. The project is highly scalable and the IoT devices are easy to program.
2. The project is open to audit and inspection, ensuring data ethics and transparency.

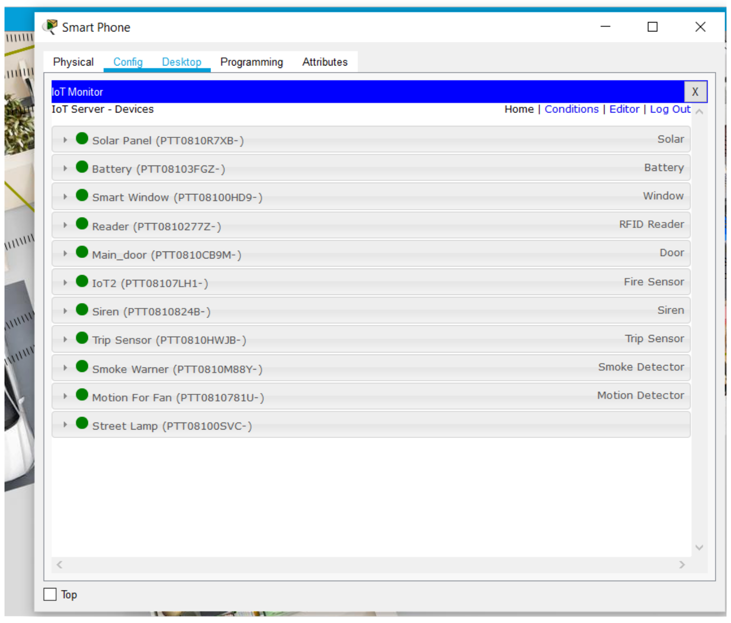

Figure 5The data is easy to audit and monitor

References

[1] Prince Waqas Khan, Yung-Cheol Byun, et al., “Machine Learning-Based Approach to Predict Energy Consumption of Renewable and Nonrenewable Power Sources”, 17 September 2020

[2] Mel Keytingan M. Shapi, Nor Azuana Ramli, Lilik J. Awalin, “Energy consumption prediction by using machine learning for smart building: Case study in Malaysia”, Developments in the Built Environment, Volume 5, 2021, 100037, ISSN 2666-1659

[3] Seyedzadeh, S., Rahimian, F., Glesk, I. et al. “Machine learning for estimation of building energy consumption and performance: a review”. Vis. in Eng. 6, 5 (2018)

[4] Zhijian Liu, Di Wu, Yuamwei Liu, et al., “Accuracy analyses and model comparison of machine learning adopted in building energy consumption prediction”, Sage Journals, Volume 37, Issue 4, January 2019

[5] A. Javed, H. Larijani and A. Wixted, "Improving Energy Consumption of a Commercial Building with IoT and Machine Learning," in IT Professional, vol. 20, no. 5, pp. 30-38, Sep./Oct. 2018, doi: 10.1109/MITP.2018.053891335.

[6] A. Nasser and V. Simon, "A Novel Method for Analyzing Weather Effect on Smart City Traffic," 2021 IEEE 22nd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2021, pp. 335-340, doi: 10.1109/WoWMoM51794.2021.00061.

[7] B. Angel and D. Angel, "A Systematic Survey on the use of Fuzzy Graph Structures in India’s Smart City Development," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1249-1253, doi: 10.1109/ICCMC51019.2021.9418353.

[8] Gusul, Procopie & Butnariu, Alina-Ramona, “EXPLORING THE RELATIONSHIP BETWEEN SMART CITY, SUSTAINABLE DEVELOPMENT AND INNOVATION AS A MODEL FOR URBAN ECONOMIC GROWTH”, The Annals of the University of Oradea. Economic Sciences, 2021

[9] Fagadar, Carmen & Trip, Diana & Darie, Gavrilut & Badulescu, Daniel, “SMART CITIES AND THE EUROPEAN VISION” The Annals of the University of Oradea. Economic Sciences, 2021

[10] Y. Alayza, J. Singer and A. L. Michala, "Optimizing Task Allocation for Edge Micro-Clusters in Smart Cities," 2021 IEEE 22nd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2021, pp. 341-347, doi: 10.1109/WoWMoM51794.2021.00062.

[11] Furdík, Karol & Lukac, Gabriel & Sabol, Tomas & Kostelnik, Peter, “The Network Architecture Designed for an Adaptable IoTbased Smart Office Solution”, International Journal of Computer Networks and Communications Security. 1. 216-224.

[12] H. Chourabi et al., "Understanding Smart Cities: An Integrative Framework," 2012 45th Hawaii International Conference on System Sciences, 2012, pp. 2289-2297, doi: 10.1109/HICSS.2012.615.

[13] Taewoo Nam and Theresa A. Pardo, “Conceptualizing smart city with dimensions of technology, people, and institutions”, Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, Association for Computing Machinery, New York, NY, USA, 282–291.

[14] Gusul, Procopie & Butnariu, Alina-Ramona, “EXPLORING THE RELATIONSHIP BETWEEN SMART CITY, SUSTAINABLE DEVELOPMENT AND INNOVATION AS A MODEL FOR URBAN ECONOMIC GROWTH”, The Annals of the University of Oradea. Economic Sciences, 2021.

[15] Furdík, Karol & Lukac, Gabriel & Sabol, Tomas & Kostelnik, Peter, “The Network Architecture Designed for an Adaptable IoTbased Smart Office Solution”, International Journal of Computer Networks and Communications Security. 1. 216-224, 2013

[16]Pérez-delHoyo, Raquel & Mora, Higinio & PAREDES, JOSÉ, “USING SOCIAL NETWORK DATA TO IMPROVE PLANNING AND DESIGN OF SMART CITIES”, 171-178. 10.2495/UG180161, 2013.

[17] Á. Oliveira and M. Campolargo, "From Smart Cities to Human Smart Cities," 2015 48th Hawaii International Conference on System Sciences, 2015, pp. 2336-2344, doi: 10.1109/HICSS.2015.281.

[18] https://online.hbs.edu/blog/post/data-ethics (accessed on 10th March 2022)

[19] Floridi Luciano and Taddeo Mariarosaria, 2016, “What is data ethics?” *Phil. Trans. R. Soc. A.*3742016036020160360

[19] A vision of smarter cities, IBM Institute for Business Value (https://www.ibm.com/downloads/cas/2JYLM4ZA)