



# Energy Monitoring Platform: A High-Performance Smart Meter Data Analytics Engine

Akshay Venkatesan, Deepanshu Sehgal, Sushant Sharma

**Abstract:** 4th Industrial revolution or Industry 4.0 is a concept of factories in which machines are augmented with wireless connectivity and sensors, connected to a system that can visualize the entire production line and make decisions on its own. To remain competitive, Industries or factories are one of the major factors leading to the consumption of energy. However, these industries are facing two obstacles. The first one is the low availability of energy and the other is increased cost of the current available energy. Different alternatives like Generators and Inverters have their limitations when it comes to cost and supply. The industries need to optimize the usage to stay price in check. It has become a necessity for the industries reduce energy consumption for a comparable level of activity to observe it. The platform is the answer to these problems. It is a solution that puts the client/user in control of their energy consumption by enabling live monitoring of the energy smart grid data via a dedicated online platform.

**Keywords:** Elastic Map Reduce, Flask API, S3 storage, Regression

## I. INTRODUCTION

In the early period of mechanical innovation, the conveyance of vitality as fuel or power was totally relied upon conventional vitality meters or techniques. The utilization of those meters has been continuously diminishing with the progress in development as expeditious changes have been made to face the controversy that occurred by them. The major problem was that the consumers were unable to measure their usage correctly. The monthly bill that the customers were provided was of no use as it couldn't tell how much energy each appliance use. Smart Meters or Smart Grids are the solutions to these problems. Their activity and future vitality management will be progressively information concentrated. They have begun producing an ever progressively huge volume of information. With the utilization of this information, everyday alarms will be given to the customers dependent on hourly usage of vitality. Their essential point is to decrease vitality utilization in the ventures and inhabitants.

Manuscript received on March 10, 2021.

Revised Manuscript received on March 16, 2021.

Manuscript published on April 30, 2021.

\* Correspondence Author

**Akshay Venkatesan\***, Engineering student, North Cap University, Gurugram, Haryana, India.

**Deepanshu Sehgal**, Engineering student, North Cap University, Gurugram, Haryana, India.

**Sushant Sharma**, Engineering student, North Cap University, Gurugram, Haryana, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

While studying the market, the conduct of the consumers can be examined and analyzed. The vitality utilization examples can be encouraged in improving the conduct of clients. The power market can be rebuilt with the establishment of these meters, as it not only conserves the energy yet in addition lessens carbon dioxide outflows. Timely consumption of consumers can be reduced as Smart Meters are connected to online billing. This paper presents a solution to build a more sustainable, highly digitized, and dynamic energy system, but before that, we must begin by fixing how we utilize, devise and calibrate vitality usage our facilities.

**Smart Grids:** Network simulators contain two sub frameworks that are interlinked and incorporated: (a) the data framework, where information flows in the digital piece of intelligent lattices called smart grids. For example, estimation of information and control signals are imparted by set of connected computers; and (b) the framework which conveys flow power in the physical part of keen networks involving smart meters and force gadgets, for example: generators, towers and transformers. The IT segment of the data foundation incorporates demonstrating, analysis, financial viewpoints like business exchanges, data sharing and administration. This work centers around the administration of the meter information and investigation everywhere scales, which are center bits of the IT part of network simulators.

**Data Mining:** Different information mining methods are applied to break down the network simulators information for different application purposes. This would prompt the rev-elation of information and would help in picking up bits of knowledge from big data to help vitality management. However, most work has been shown by using relatively compact data sets, for example: demand or load forecasting, customer segmentation, pattern classification, energy tariff recommendations, energy use of appliances in individual home, and demand side management.

## II. LITERATURE SURVEY

The Literature Survey is performed using research papers published previously and the survey clearly states how the smart meters are used on various appliances to collect data. It also depicts that how latest technologies like Internet of Things, Big Data and Data analytics, Machine Learning and Cloud Computing are used to make effective use of the data for displaying power usage statistics. It helps in discovering the relations between the trend of big data era and that of the



# Energy Monitoring Platform: A High-Performance Smart Meter Data Analytics Engine

new generation green revolution and examining structural and auxiliary and social determinants of vitality utilization.

**Table- I: Literature Survey**

Ref No	Objective	Algorithms used	Dataset	Relevance	Performance Measure
1	A smart home energy management system using Internet of Things and big data analytics approach	Big Data Analytics	Data set is collected using many different Wireless sensors installed in residential units.	Provides energy management system and enhances sustainability.	Handles Big data and meets consumer demands
2	A comprehensive study of big data driven smart energy management	Smart grid and Demand side management (DSM)	Taking smart grid as the research background	Provides information about managing big data related to smart energy management	Specificity, Security and Privacy
3	Discovering the relations between the trend of big data era and that of the new generation green revolution	Internet of Things and Data Mining	Using IOT devices, CISCO reports	Provides information regarding green issues related to systems using big data.	Accuracy and challenges
4	Examining structural and behavioral determinants of residential electricity consumption	Factor analysis and stepwise regression	Data set of sample 1628 households' electricity consumption	Provides information regarding the necessary factors leading to consumption of electricity	Realtime analysis
5	Enables the consumers and the micro-energy producers to take a sparkier role in the electricity market	Cloud Computing, High performance computing and network management	Data set is collected Smart Grids	Provides dynamic energy management strategies.	Efficiency and forecasting
6	To perform clustering analysis of smart meter data	Clustering	Data sets from smart meters	Provides information regarding data analysis strategies for smart meter	Efficiency and challenges
7	Comprehensive study on the compression techniques for smart meter big data	Anomaly Detection and Big Data Analytics	Data sets from smart grids	Helps in mining and analyzing smart meter big data	Accuracy and efficiency

8	Online display of the power usage of solar energy as a renewable energy	IOT and Cloud Computing	This monitoring is done through raspberry pi using flask framework	provides information regarding displaying power usage statistics	Display Accuracy
---	---	-------------------------	--	--	------------------

### III. EXISTING SYSTEM

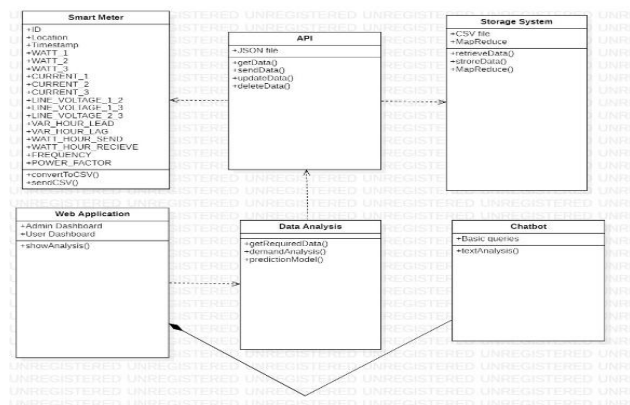
The current system in place for factories in India that use diesel generators rely upon either ordinary meters to compute the metered usage of energy, or smart meter with real-time data but with no avail to any real-time analysis of the live data. There is no seamless interconnection between the real-time data and analysis right now in the current existing system in case of diesel generator factories.

### IV. PROBLEM DEFINITION

Residents as well as factories don't have analysis of their power usage. Users don't have real-time access to power usage statistics. Analysis can often be complicated and general users might not be able to understand the analysis. Current systems don't provide ease of access to the user.

### V. PROPOSED SYSTEM

The proposed system is to collect data from the gateway with the help of an API. The data is present in the form of json packets. These json packets would be converted into Excel and csv files. These files would be analyzed to show different behaviors through different data visualization techniques to show the usage of energy as per timestamp. These visualizations would be represented on an online portal which would be accessed by the company clients. In the later stages predictive modelling of the data would be done to gain insights from the data so that we can show how the customer can save energy and reduce costs. The predictive model would be created on the basis of a Machine Learning algorithm.

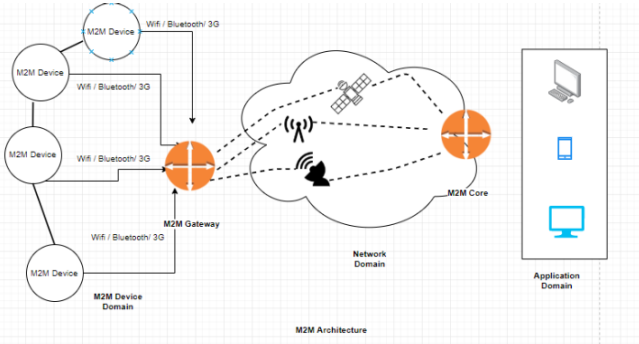


**Fig. 1. Class diagram of the proposed system**



**VI. SYSTEM ARCHITECTURE**

The proposed system comprises of an M2M architecture. A large number of M2M devices together comprise on an M2M area domain. Each of these individual devices are referred to as nodes. These nodes are smart devices that enable us to use a lot of the advanced sensory technologies of the modern day. In our case, these nodes are primarily the smart meter. The smart meters gather data at a set frequency and then with the help of wireless technologies and cloud storage, these data are transmitted to the cloud storage for further analysis. The client thereafter can access the analysis to the real-time data from their application and can easily monitor the data.

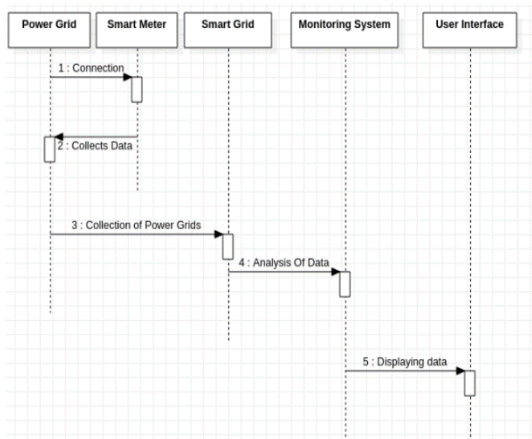


**Fig. 2. Architecture Diagram**

**VII. OBJECTIVE AND SCOPE**

**Objective-** The main objective of this project is to analyze the energy consumption and to suggest whether the current source of power or fuel is suitable and to maximize the output. The objective of project Energy Monitoring Platform is to build an energy efficiency platform along with a chatbot that can deliver personalized information on disaggregated energy usage through an intelligent and interactive virtual assistant.

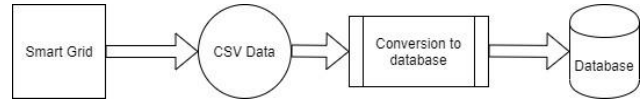
**Scope-** The scope of this project is to analyze the energy consumption and to provide better solutions on the fuel to be used to conserve energy as well as generate higher outputs with minimum use of energy. The analysis would be highly beneficial for the industries in the future and would help in conservation of natural resources.



**Fig. 3. Sequence Diagram of the proposed system**

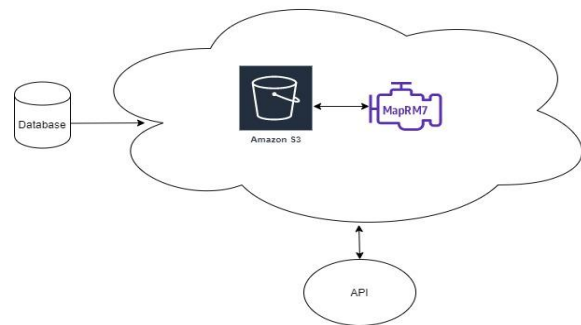
**VIII. MODULES**

**Module 1 (Data pre-processing)** - The smart meters collect live data for each area at a set frequency. This data is collected and stored in the format of a CSV file. This data is then taken up for every smart meter and with the help of UNIX commands stored with the name of the smart meter along with the timestamp marking the date.



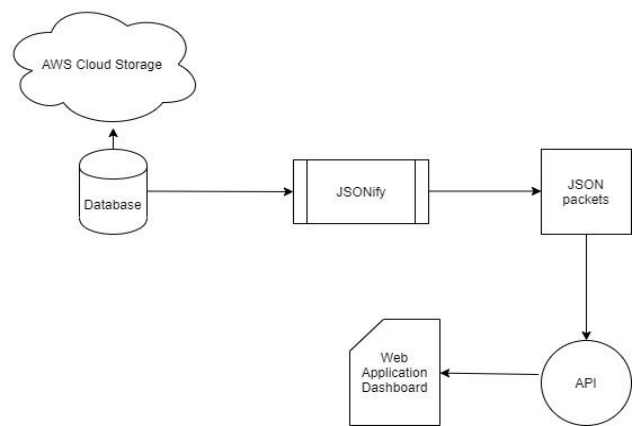
**Fig. 4. Module 1**

**Module 2 (API creation)** - To create an API using FLASK module from Python. Data is obtained from the storage in database format. Data is retrieved using the API and converted into json format. Queries can be run on the data and only required data would then be passed on for analysis.



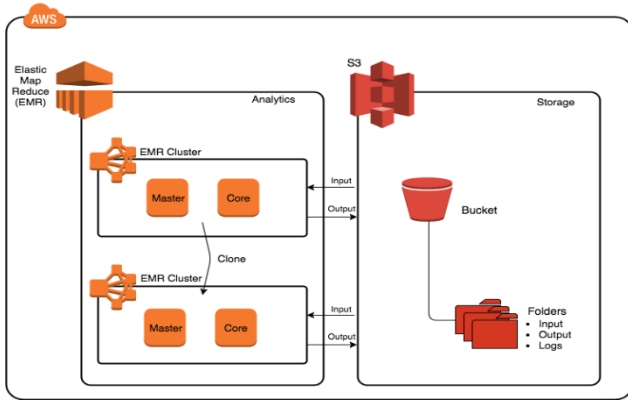
**Fig. 5. Module 2**

**Module 3 (Cloud Storage using EMP)** - The preprocessed data is stored using HDFS architecture. We use the AWS S3 for storage and the AWS Elastic Map Reduce technology to create the necessary architecture to store our large data. The stored CSV files are converted to daily database log files for each smart meter and then uploaded to the cloud. The data is then retrieved using API into database format. The API would then further convert the data into JSON packets.



**Fig. 6. Module 3**

# Energy Monitoring Platform: A High-Performance Smart Meter Data Analytics Engine



**Fig. 7. Elastic Map Reduce**

**Module 4 (Data Analysis)** - Finally, the data analysis would be performed. Many techniques like clustering as well as some ML algorithms would be used. Decision tree would be used to create a predictive model. Queries can be run on the data and only required data would then be passed on for analysis.

**Module 5 (Chatbot)** - A chatbot uses machine learning algorithms in order to respond to user queries. Chatbots are used for purposes such as assistant bots for asking basic questions to the user in order to improve the user experience. In our case, we have used the chatbot for the same purpose. The chatbot in our application is used to process basic user queries such as processing user queries for bill and power consumption and to then either redirect the user to a page or give user an answer for the query. If the bot is unable to process the query, it then displays the contact details for referring the query to a real person. With the help of multiple data sets, the chatbot was then trained with the help of corpus primarily focusing on the conversations between past assistant bots and the users. The following chapter comprises of the essential code for creation of the chatbot.

## IX. DATA VISUALIZATION AND PREDICTIVE MODELLING

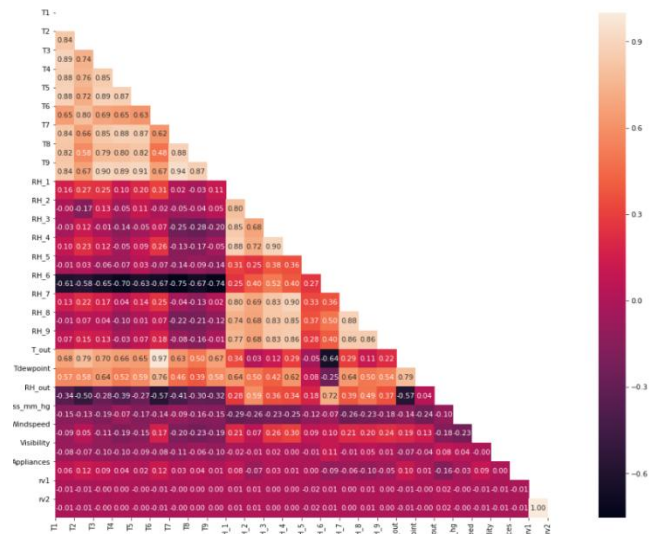
After analyzing the data and performing the data visualization, the data preprocessing was done. For training the data and creating the predictive model, there were 6 algorithms to choose from: -

- (1) Ridge regression
- (2) Lasso regression
- (3) SVM regression
- (4) K-Neighbour regression
- (5) Random Forest Regressor
- (6) Multi-Layer Perceptron Regressor
- (7) Extra Tree Regressor

## X. RESULT AND DISCUSSION

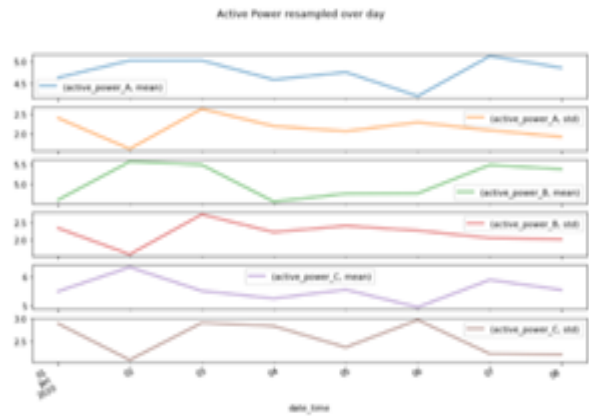
There were two datasets on which the project was implemented- (1) On a dataset comprising of energy consumption in a daily household to check and monitor on the daily appliance consumption of energy and (2) On the dataset of a diesel generator factory. Each of the above chosen models were implemented on the data sets that were pre-processed. The R1, R2 and RMSE scores were measured

for each model. Based on the results of these scores, the most accurate model which was the Extra Tree regressor model for both the cases was chosen. Along with the predicted values, the Platform offers data visualizations such as the correlation graph and feature importance to help understand and analyze the data sets further.

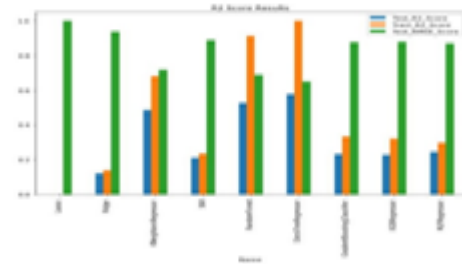


**Fig. 8. Correlation heatmap**

The extra tree regressor model was chosen because it had the highest efficiency and hence that was chosen to create the model. The computed RMSE score of the model was 0.7263



**Fig. 9. Sample data visualization**



**Fig. 10. R2 Scores**

Table- I: Result

S No	Name	Train Time	Train R2 Score	Test R2 Score	RMSE Score
1	Ridge Regression	0.036536	0.137553	0.121391	0.937341
2	Lasso regression	0.055907	0.000000	0.000000	1.000000
3	SVM regression	14.356879	0.235724	0.209934	0.888857
4	K-Neighbour regression	0.025851	0.681464	0.485560	0.717245
5	Random Forest Regressor	3.103549	0.913691	0.525681	0.688708
6	Multi-Layer Perceptron Regressor	3.411786	0.298567	0.243178	0.869955
7	Extra Tree Regressor	0.800048	1.000000	0.577183	0.726266

XI. CONCLUSION

With the help of this portal one can monitor different areas of various departments which are utilizing high energy so as to take preventive measures. The portal will be capable to collect, store and analyze the collected data from different remote monitoring devices to manage electrical cost associated with the electrical systems. The aim is to build a more feasible, eminently digitized, and dynamic energy system, but before that we must begin by fixing how we consume, produce and measure energy within the facilities. Energy squandering will be decreased by elevating the possibility of reduced consumption using analyzed Smart Meter data assigned to perform this research work. The scope of this project is to analyze the energy consumption and to provide better solutions on the fuel to be used to conserve energy as well as generate higher outputs with minimum use of energy. The analysis would be highly beneficial for the industries in the future and would help in conservation of natural resources. With the advancements in technology such as blockchain, this project can further be modified to enable energy diversion in regions of need and the entire process can be automated and the energy would be distributed based on the live demands.

of Science and Technology for his valuable advice and guidance.

REFERENCES

1. Al-Ali, A. R., Zuolkernan, I. A., Rashid, M., Gupta, R., & Alikarar, M. (2017). A smart home energy management system using IoT and big data analytics approach. *IEEE Transactions on Consumer Electronics*, 63(4), 426-434.
2. Zhou, K., Fu, C., & Yang, S. (2016). Big data driven smart energy management: From bigdata to big insights. *Renewable and Sustainable Energy Reviews*, 56, 215-225.
3. Wu, J., Guo, S., Li, J., & Zeng, D. (2016). Big data meet green challenges: Big data toward green applications. *IEEE Systems Journal*, 10(3), 888-900.
4. Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy*, 55, 184-194.
5. Diamantoulakis, P. D., Kapinas, V. M., & Karagiannidis, G. K. (2015). Big data analytics for dynamic energy management in smart grids. *Big Data Research*, 2(3), 94- 101.
6. Flath, C., Nicolay, D., Conte, T., van Dinther, C., & Filipova-Neumann, L. (2012). Cluster analysis of smart metering data. *Business & Information Systems Engineering*, 4(1), 31-39.
7. Patil, S., Vijayalashmi, M., & Tapaskar, R. (2017). Solar energy monitoring system using IOT. *Indian Journal of Scientific Research*, 149-156.
8. Wen, L., Zhou, K., Yang, S., & Li, L. (2018). Compression of smart meter big data: A survey. *Renewable and Sustainable Energy Reviews*, 91, 59-69.

AUTHORS PROFILE



**Akshay Venkatesan**, is an engineering student from SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India. Having done his B.Tech from the university, his main thrust areas are Big Data Analysis, Cloud computing and Text Mining and DBMS. He is currently a Systems Engineer at Tata Consultancy Services.



**Deepanshu Sehgal**, is an engineering student from SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India. Having done his B.Tech from the university, his main thrust areas are Big Data Analysis, Cloud computing, Text Mining and DBMS. He is currently an Application Development Associate at Accenture.



**Sushant Sharma**, is an engineering student from The NorthCap University, Gurugram, Haryana, India. Having done his B.Tech from the university, his main thrust areas are Big Data Analysis, Cloud computing and DBMS. He is currently an Analyst at Ernst and Young- GDS.

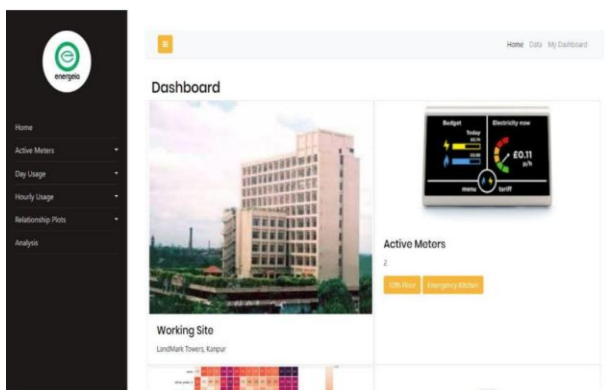


Fig. 11. Dashboard page

ACKNOWLEDGMENT

First and foremost, we would like to thank the SRM Institute of Science and Technology and The Northcap University. Their astonishing support has been the inspiration to the creation of the energy monitoring platform. We would then like to thank the company Energeia for providing us with live data sets. We thank Dr N.Parthiban at SRM Institute