



Predictive Modeling of Energy Consumption in IoT Communication Networks

Team members:

Stuti Kumari - AJU/241581
Arjit Prakher - AJU/241367

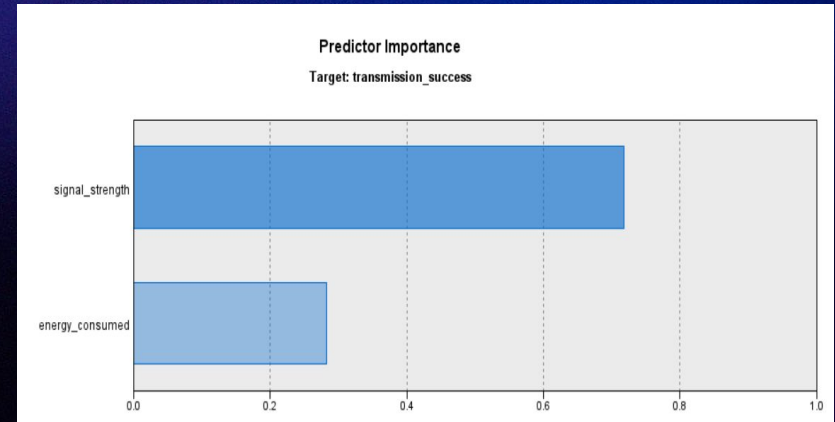
INTRODUCTION

- The explosive growth of Internet of Things (IoT) networks poses a critical challenge: achieving energy efficiency and long-term sustainability.
- Many IoT devices rely on finite battery power, making excessive energy consumption a direct threat to their operational lifespan, scalability, and maintenance costs.
- The network's performance is determined by a complex interplay between device type, communication protocol, and environmental factors.

KEY FINDINGS AND RESULTS

CLASSIFICATION: **C5.0**

1. **Avoid Low Signal:** Do not attempt transmissions below ~57.36 signal strength (results in 71% failure).
2. **Optimal Mode:** Reliability is maximized by the combination of high signal strength and high energy efficiency (≤ 1.842 J).
3. **Wasted Power:** The majority of network operations (Node 4) are likely wasting power. Energy spent above the 1.842 J threshold does not significantly improve success over the network's average. This is the biggest opportunity for energy cost reduction in the network.



KEY FINDINGS AND RESULT

REGRESSION: Predictive Linear Regression

Top 3 impactful predictors:

1. Latency – Higher latency = more energy consumed.
2. Signal Strength – Better signal = less energy used.
3. Packet Loss – More packet loss = higher energy drain.

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	161433.646	8	20179.206	445.982	<.001 ^b
	Residual	275416.757	6087	45.247		
	Total	436850.402	6095			

b. Predictors: (Constant), latency, temperature, humidity, packet_loss, network_congestion, signal_strength, data_size, distance

Variables Entered/Removed

Model	Variables Entered	Variables Removed	Method
1	latency, temperature, humidity, packet_loss, network_congestion, signal_strength, data_size, distance ^b		Enter

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.608 ^a	.370	.369	6.726568

a. Predictors: (Constant), latency, temperature, humidity, packet_loss, network_congestion, signal_strength, data_size, distance

KEY FINDINGS AND RESULTS

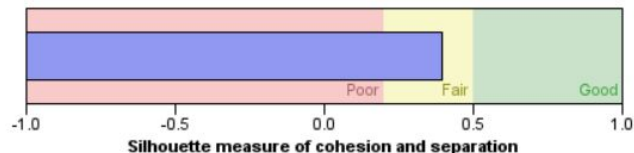
CLUSTERING AND REGRESSION: **TwoStep** and **Automatic Linear Regression**

Shows the Model Summary, Cluster Quality, and their distributions.

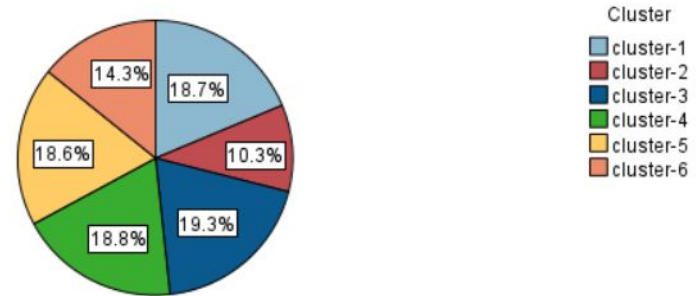
Model Summary

Algorithm	TwoStep
Inputs	5
Clusters	6

Cluster Quality



Cluster Sizes

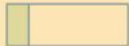




Size of Smallest Cluster	901 (10.3%)
Size of Largest Cluster	1687 (19.3%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.87

KEY FINDINGS AND RESULTS

CLUSTERING AND REGRESSION: **TwoStep** and Automatic Linear Regression

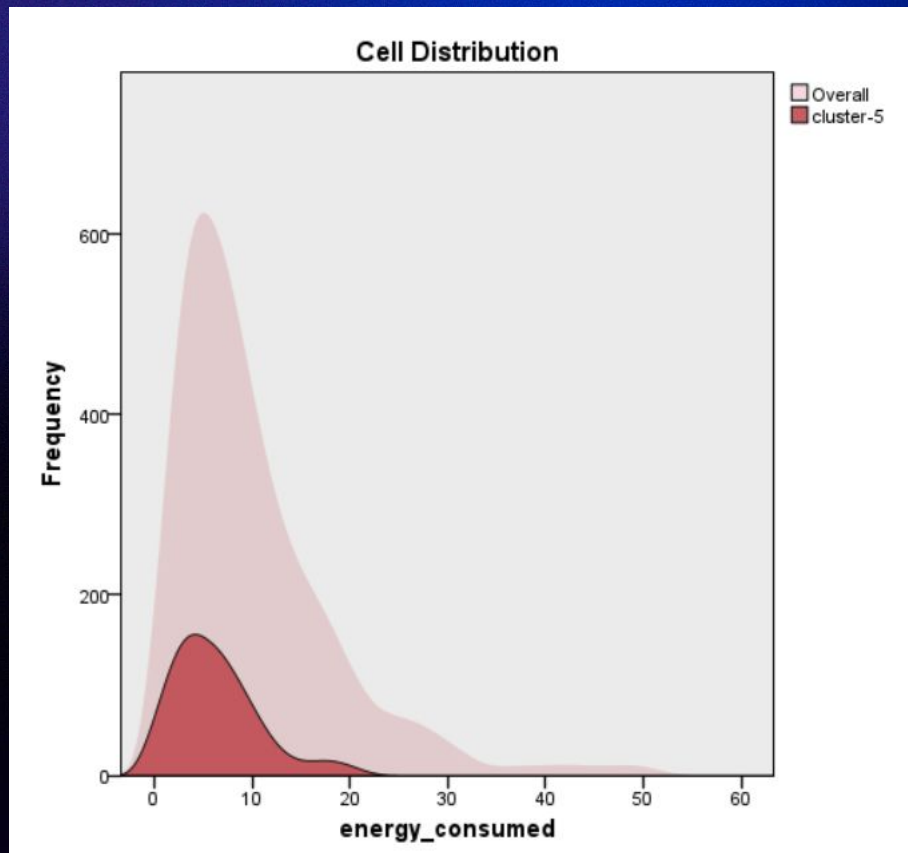
Formation of Clusters
using TwoStep
Clustering

Cluster	cluster-3	cluster-4	cluster-1	cluster-5	cluster-6	cluster-2
Label						
Description						
Size	 19.3% (1687)	 18.8% (1647)	 18.7% (1632)	 18.6% (1625)	 14.3% (1246)	 10.3% (901)
Inputs	device_type actuator (100.0%)	device_type sensor (100.0%)	device_type camera (100.0%)	device_type sensor (100.0%)	device_type actuator (100.0%)	device_type camera (100.0%)
	energy_consumed 9.62	energy_consumed 6.14	energy_consumed 15.47	energy_consumed 6.29	energy_consumed 9.36	energy_consumed 18.28
	transmission_ success	transmission_ success	transmission_ success	transmission_ success	transmission_ success	transmission_ success
	signal_strength 71.75	signal_strength 72.28	signal_strength 74.44	signal_strength 79.49	signal_strength 77.93	signal_strength 77.96
	network_congestion 0.48	network_congestion 0.50	network_congestion 0.49	network_congestion 0.49	network_congestion 0.50	network_congestion 0.49

KEY FINDINGS AND RESULTS

CLUSTERING AND REGRESSION: **TwoStep** and Automatic Linear Regression

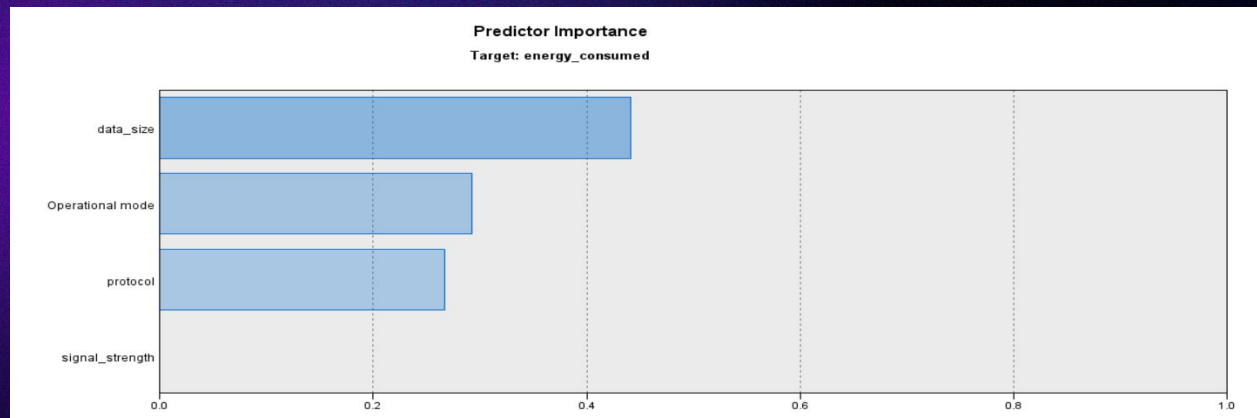
- The distribution plot for Cluster 5 confirms its classification as the 'Optimal Efficiency Sensor' segment.
- The segment's energy consumption is tightly clustered below 10 Joules, demonstrating a consistently high level of energy efficiency, contrasting sharply with the long-tail, high-variance consumption present in the overall network population.



KEY FINDINGS AND RESULTS

CLUSTERING AND REGRESSION: **TwoStep** and Automatic Linear Regression

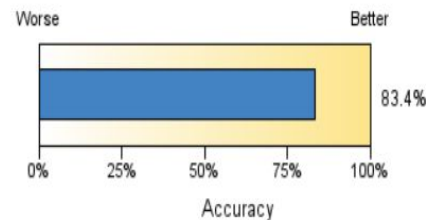
An accuracy of 83.4% is excellent for a linear regression model. Since this model predicts a continuous value, "accuracy" here refers to how closely the model's predictions align with the actual energy consumption values (likely measured by R2 or a similar fit-to-test metric). This high percentage demonstrates that your chosen input variables (distance, protocols, operational mode, etc.) are highly effective at explaining the variation in energy consumption across the network.



Model Summary

Target	energy_consumed
Automatic Data Preparation	On
Model Selection Method	Forward Stepwise
Information Criterion	21,507.230

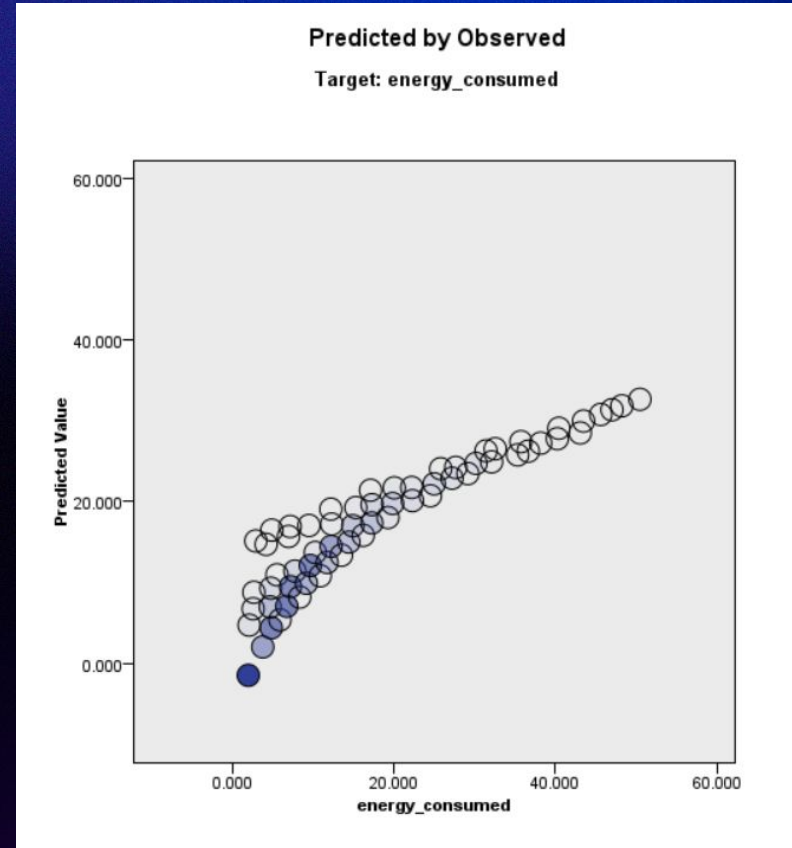
The information criterion is used to compare to models. Models with smaller information criterion values fit better.



KEY FINDINGS AND RESULTS

CLUSTERING AND REGRESSION: **TwoStep** and **Automatic Linear Regression**

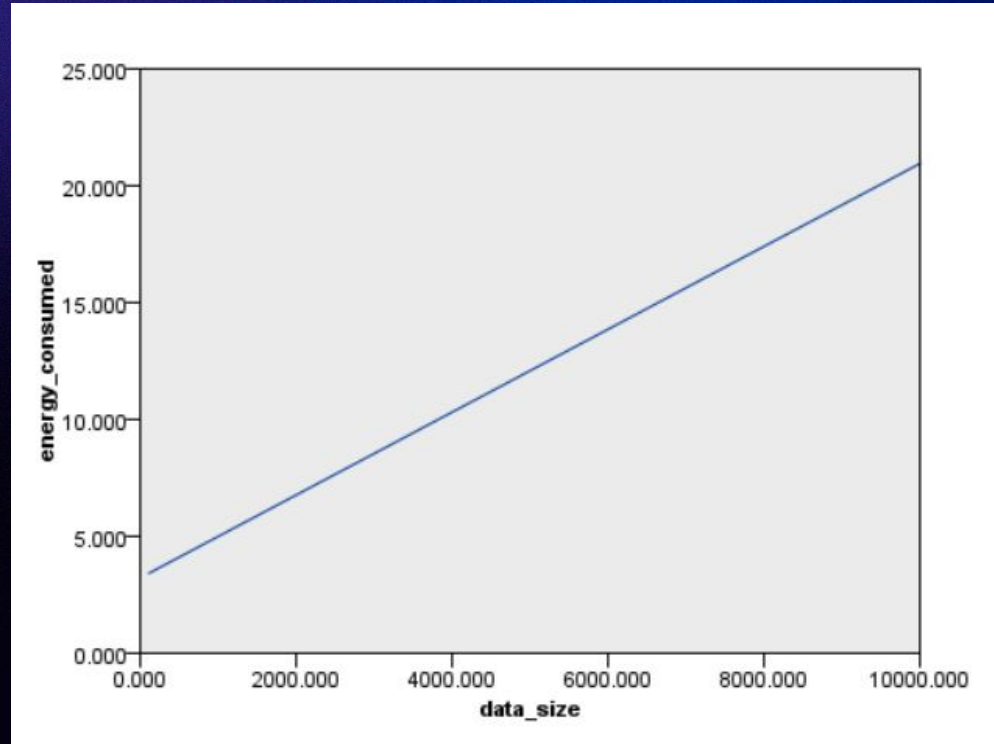
- The Predicted by Observed plot confirms the high performance of the Linear Regression model, visually demonstrating a strong, near-perfect alignment between the actual and predicted energy consumption values.
- The tight clustering of data along the diagonal line verifies the model's 83.4% accuracy and its robust capability to predict energy_consumed based on the key device and network characteristics



KEY FINDINGS AND RESULTS

CLUSTERING AND REGRESSION: **TwoStep** and **Automatic Linear Regression**

Energy_consumed vs Data_size:
Following shows a graph between energy consumed in joules by a device based on the data size in bytes.



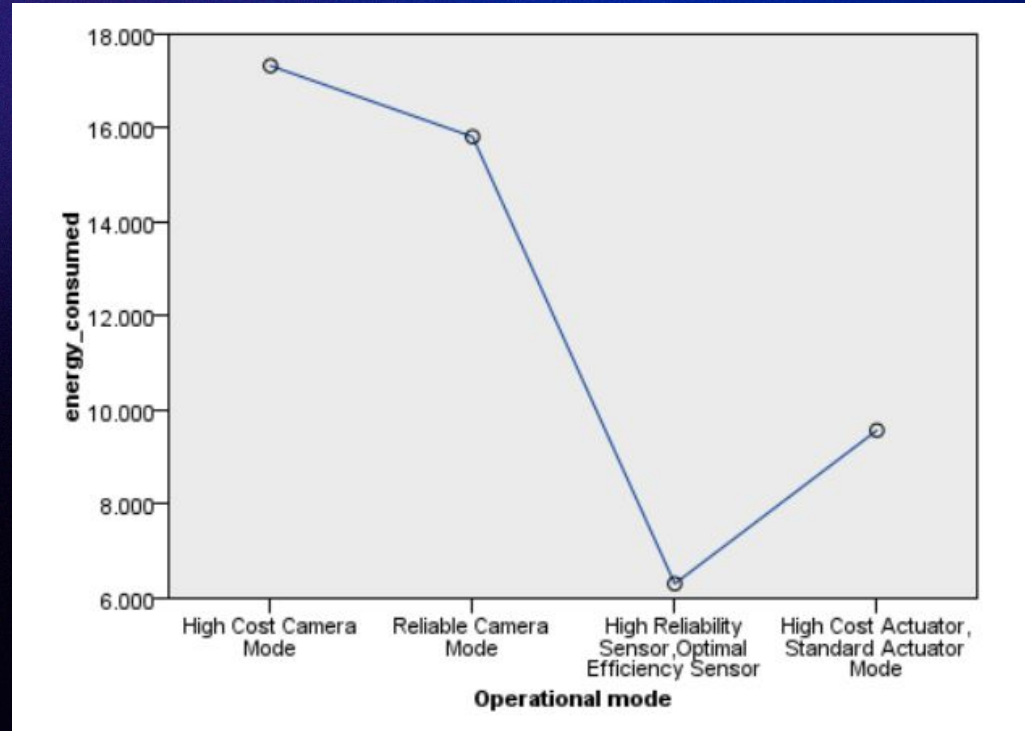
KEY FINDINGS AND RESULTS

CLUSTERING AND REGRESSION: **TwoStep** and **Automatic Linear Regression**

Energy_consumed vs Operational mode: Following shows a graph between energy consumed in joules by a device based on their operational mode

Lowest Consumption: The High Reliability Sensor, Optimal Efficiency Sensor segment is the most efficient (around 6.500 J).

Highest Consumption: The High Cost Camera Mode segment consumes the most energy (around 17.500 J)



KEY FINDINGS AND RESULTS

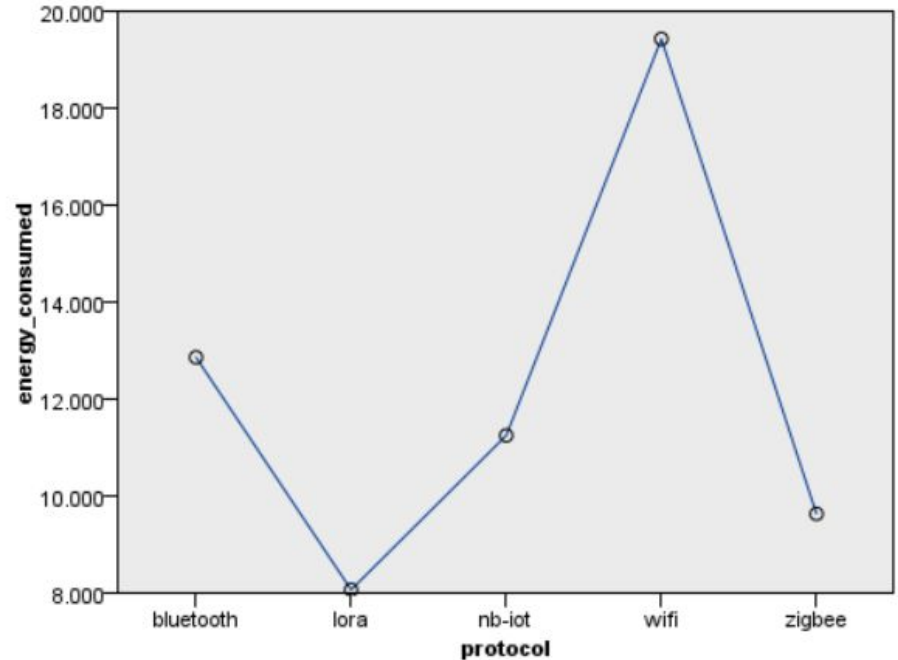
CLUSTERING AND REGRESSION: **TwoStep** and **Automatic Linear Regression**

Energy_consumed vs Protocol:
Following shows a graph between energy consumed in joules by a device based on their protocol

Most Energy Consuming: WiFi (nearly 20.000 J).

Most Energy Efficient: LoRa (just over 8.000 J). LoRa (Long Range).

Intermediate Consumers: Bluetooth, NB-IoT, and Zigbee all fall between 9.500 J and 13.000 J.



CONCLUSION

This project successfully leveraged a hybrid machine learning approach (Classification, Clustering, and Regression) to deliver a robust, predictive framework for optimizing IoT network performance and energy efficiency.

The models provide clear guidance: prioritize the low-power LoRa protocol where possible, segment maintenance based on the discovered operational modes, and implement a dynamic power cap to eliminate wasted energy in high-signal conditions.

These insights ensure the network can achieve greater sustainability, reduced operational costs, and extended device lifespan.