**Materials and methods:**

Para 1:

Study setting: - SIMATS, Saveetha school of engineering.

No of groups :- 2

Group 1 – Xgboost (Extreme Gradient Boosting)

Group 2 – Linear regression

Sample size - 100

Para 2 –

Sample preparation group 1: -

Xgboost (Extreme Gradient Boosting)

* Collection and preparation of data
* Preprocessing of data
* Normalize the data
* Model building
* Model training
* Model testing

Para 3 –

Sample preparation group 2: -

Linear regression algorithm:

* Data collection
* Data transformation
* Remove Collinearity
* Model building
* Model training
* Model testing

Para 4 –

Testing set up: -

* Python idle (3.12)
* Windows operating system
* Intel core i5
* 16 GB RAM
* Google colab

Para 5-

Data collection:

* Kaggle.com
* Numerical dataset with 1000 data

The dataset has been collected from the Kaggle.com. the data set contains more than 1000 numerical data.

Para 6-

* Statistical software used – SPSS
* Independent variables – temperature, general lighting, vents, filtrations, display systems
* Dependent variables – total electricity consumption
* Analysis – T – test, Independent sample test, sample bar means of accuracy by group

Discussion:

Para 1:

Result summary:

In this study, indicates that XGBoost's advanced machine learning methodology performs better at predicting building energy usage than the conventional Linear Regression model.

Para 2: -

Similar findings:

Yucong, Wu, and Wang Bo. "Research on EA-xgboost hybrid model for building energy prediction." (2020).

Accuracy: 89 %

Lu, Hongfang, et al. "Short-term prediction of building energy consumption employing an improved extreme gradient boosting model” (2020)

Accuracy: 70%

Limitations:

* Performance may be affected by the sensitivity of the Extreme Gradient Boosting (XGBoost) algorithm to variations in input data.
* Resource-intensive nature of XGBoost may pose limitations for deployment in resource-constrained environments.

Future scope:

* Exploring ensemble methods to combine the strengths of both XGBoost and Linear Regression for improved predictive accuracy.
* Expanding datasets to include a broader range of building types and locations for more robust models.

Conclusion:

Linear regression model has an accuracy of 76%, whereas extreme gradient boosting shows an impressive 90.47% accuracy. When predicting building energy usage, extreme gradient boosting is more effective.