Project Proposal: Predicting Solar Energy Efficiency of Buildings in Lagos

By Team Nwapa, AI Saturdays Lagos Cohort 9 - Flipped

Problem Statement

Nigeria faces increasing energy challenges as the demand for electricity far exceeds supply, resulting in power shortages and overdependence on non-renewable sources. Despite Lagos' vast rooftop potential for solar energy, many homeowners, developers, and policymakers lack data-driven tools to assess how much renewable energy a building can realistically generate.

Our project aims to build a predictive model to estimate the solar energy efficiency of buildings in Lagos using features such as roof surface area, tilt, building height, capacity factor, and potential installable area. This solution is important because it promotes the use of renewable energy, supports urban sustainability planning, and provides insights that can guide investments in solar adoption across Nigeria.

Existing Solutions

Currently, solar potential assessments in Nigeria rely heavily on manual site surveys or generalized regional estimates. These methods are often time-consuming, expensive, and lack precision at the individual building level.

In contrast, our project applies machine learning to predict solar energy efficiency using real geospatial and building data. Unlike existing static assessments, our model analyzes real-world building features, learns from measurable patterns, and provides a scalable tool that can instantly estimate solar potential for any location in Lagos.

Objectives

- Clean and preprocess the Lagos Rooftop Solar Potential dataset for accurate analysis.

· Perform Exploratory Data Analysis (EDA) to uncover patterns and correlations in rooftop characteristics.

· Engineer additional features such as area utilization ratio, capacity density, energy density, and system efficiency.

· Build regression models to predict annual solar energy potential (kWh/year).

· Develop a classification model to categorize buildings as low, medium, or high solar efficiency.

· Evaluate model performance using metrics such as RMSE, MAE, accuracy, precision, recall, and F1-score.

· Deploy the final model as an interactive web-based tool for public and institutional use.

Proposed Dataset

For this project, we will use the Lagos Rooftop Solar Potential Dataset, which provides detailed rooftop characteristics of over 200,000 buildings in Lagos. [Lagos - Rooftop Solar Potential Mapping - Dataset - ENERGYDATA.INFO]

Why We Chose This Dataset

This dataset contains critical attributes such as surface area, potential installable area, peak installable capacity, building height, tilt, and estimated capacity factor; all directly influencing solar energy generation. It also includes annual energy potential, allowing both regression and classification tasks. The dataset is realistic, large-scale, and region-specific, making it ideal for modeling solar efficiency in Nigeria's most populous city.

Proposed Methodology

Data Sourcing & Cleaning: Inspect the dataset for missing values, outliers, and inconsistencies using Pandas and NumPy. Handle nulls where necessary and retain only relevant columns for energy prediction.

Exploratory Data Analysis (EDA): Use Matplotlib and Seaborn to visualize relationships between rooftop characteristics (area, tilt, height, capacity) and annual energy potential. Explore distributions, correlations, and building type trends.

Feature Engineering: Create new informative ratios and indicators such as;

i. Area Utilization Ratio: ratio of usable to total roof area.

ii. Energy Density: energy potential per square meter of usable area.

iii. Capacity Density: capacity per square meter of usable area.

iv. System Efficiency: ratio of actual energy produced to theoretical maximum (using installed capacity × 8760 hours).

v. Encode building type categories and scale numerical features for model optimization.

Modeling & Evaluation

Train multiple regression models including Linear Regression, Random Forest Regression, and Gradient Boosting to predict annual energy potential.

For classification; train Logistic Regression, Random Forest, and XG Boost to categorize buildings into Low, Medium, or High efficiency.

Evaluate regression models using RMSE and MAE, and classification models using Precision, Recall, F1-score, and ROC-AUC.

Select the best-performing models through cross-validation and hyperparameter tuning.

Modeling Plan

We plan to test and compare multiple models:

Linear Regression: for interpretability and baseline performance.

Random Forest: to capture non-linear feature interactions.

Gradient Boosting (XGBoost/LightGBM): for high predictive accuracy and robustness.

Logistic Regression or Random Forest Classifier: for multi-class solar efficiency classification.

All models will be compared based on consistency, error reduction, and classification fairness across categories.

Deployment Plan

The final model will be deployed as a Streamlit web application, allowing users to input building features (roof area, tilt, height, etc.) and instantly receive predictions for annual solar energy potential and efficiency category.

For wider public access, we plan to host the app on Hugging Face Spaces or GitHub Pages and link it to an open dataset repository.

Expected Outcomes

A robust machine learning model that accurately predicts building-level solar energy potential.

An interactive, web-based platform enabling users to assess rooftop solar feasibility in seconds.

Analytical insights from EDA showing key physical factors that influence solar efficiency in Lagos.

A scalable framework that can later be extended to other Nigerian cities.

Community Impact

This project will empower homeowners, energy companies, and urban planners with actionable data to support renewable energy adoption. By providing an accessible solar assessment tool, it can help reduce dependence on fossil fuels, lower electricity costs, and encourage investments in clean energy infrastructure. Ultimately, this project contributes to environmental sustainability and aligns with Nigeria's transition toward a greener energy future.

Team Members

- Chitom Uzokwe (team lead)
- Precious Ita
- Joy Eromosele
- Toluwani Abiodun
- Precious Odutola

Acknowledgement

This project was inspired by the AI Saturdays Lagos learning initiative and ongoing global research on renewable energy modeling and sustainable urban planning.

References

Lagos Rooftop Solar Potential Dataset (Open Data Portal); Lagos - Rooftop Solar Potential Mapping - Dataset - ENERGYDATA.INFO

Global Solar Atlas – Solar Resource Data

AI Saturdays Lagos – Machine Learning for Social Good Resources

Nigerian Building and Energy Reports

Building Energy Efficiency Guideline for Nigeria (2016) - **By:** Federal Ministry of Power, Works & Housing with GIZ support