NATIONAL INSTITUTE OF TECHNOLOGY MEGHALAYA

Department of Computer Science and Engineering

EE382 – Minor Project

Project Report

On the topic

“Load Management System Using LSTM-Based Solar Power Forecasting”

Report submitted by:

Ayush Raj (B22EE013)
Sidhant Purkayastha (B22CS021)

Under the supervision of

Dr. Piyush Pratap Singh

Abstract

With the increasing reliance on renewable energy sources, especially solar power, efficient energy management has become crucial for ensuring stable and uninterrupted power supply. This project presents an integrated system that combines solar power forecasting with intelligent load management to optimize energy utilization in solar-based microgrids or standalone systems. Using historical weather and solar irradiance data from NIT Meghalaya (Shillong), an LSTM (Long Short-Term Memory) model is trained to accurately predict daily solar power generation. These predictions, along with current battery reserve levels and estimated daily load demand, are used to dynamically manage power distribution across essential and non-essential loads. The proposed system evaluates a calculated viability score to decide whether all loads can be powered, non-essential loads must be curtailed, or backup power sources such as generators or the grid are required. By proactively managing loads based on forecasted energy availability, the system minimizes reliance on backup power while ensuring continuous operation of critical infrastructure. This approach demonstrates how data-driven forecasting and decision logic can significantly enhance the resilience and efficiency of renewable-powered systems.

Abbreviations and Their Full Forms

Abbreviation	Full Form
ANN	Artificial Neural Network
BMS	Battery Management System
BiLSTM	Bi-Directional Long-Short Term Memory
DC	Direct Current
DL	Deep Learning
EMS	Energy Management Systems
FFN	Feedforward Neural Network
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
RES	Renewable Energy Resources
MAE	Mean Absolute Error
MSE	Mean Squared Error
MLP	Multi-Layer Perceptron
PV	PhotoVoltaic
ReLU	Rectified Linear Unit
SVR	Support Vector Regression
SoC	State of Charge

Motivation

The motivation behind this project stems from the growing need for reliable and sustainable energy solutions in areas where solar power is available but underutilized due to inconsistent generation and poor energy management. In many regions, especially remote or semi-urban areas, power outages and overdependence on backup sources like diesel generators lead to increased costs and environmental impact. This project aims to address these challenges by forecasting solar energy availability and intelligently managing energy distribution between essential and non-essential loads. By ensuring critical services always receive power while reducing reliance on backup sources, the system aims to enhance energy efficiency and demonstrates how even small-scale predictive and control systems can make a meaningful difference in energy reliability.

1 Introduction

Accurate forecasting of solar photovoltaic (PV) power generation is essential for managing the inherent variability of renewable energy and ensuring reliable grid or microgrid operation [1]. Deep learning methods—particularly Long Short-Term Memory (LSTM) networks—have demonstrated superior ability to capture nonlinear, time-dependent patterns in meteorological and irradiance data, outperforming traditional statistical and physical models in PV output prediction [1].

In this project, we harness historical weather and solar irradiance measurements from the NIT Meghalaya (Shillong) dataset to train an LSTM model that produces day-ahead solar power forecasts. The model architecture comprises stacked LSTM layers with dropout regularization, optimized using mean absolute error (MAE) and root mean square error (RMSE) metrics to ensure both accuracy and generalization across seasonal variability.

To translate these forecasts into actionable energy-management decisions, we integrate a rule-based Load Management System (LMS) that dynamically allocates power between essential and non-essential loads based on forecasted generation and battery storage levels [2]. According to Rashid et al., simple LMS implementations use priority-driven switching—via relays or microcontroller commands—to shed non-critical loads when renewable supply falls short, thereby preserving battery reserves for vital services [2].

Our unified framework computes a daily "viability score" by summing the predicted solar yield and current battery state of charge (SoC), then subtracting the estimated load demand. This scalar drives a lookup-table engine that selects one of three operating states—full supply, non-essential curtailment, or backup activation—ensuring uninterrupted power for critical loads while minimizing reliance on external sources. By simulating this feedback loop over a 30-day horizon, we demonstrate how combining LSTM forecasting with lightweight LMS logic can substantially improve renewable utilization and system resilience in resource-constrained environments.

2 Dataset Description

2.1 Overview

The dataset used in this analysis consists of locally acquired meteorological and solar energy generation data, collected via ground-based sensors. It captures various atmospheric and environmental parameters relevant to solar energy forecasting and performance evaluation. The dataset is structured in a tabular format, where each row represents a single time-stamped observation, and each column corresponds to a specific measured or derived feature.

2.2 Data Source

The data was collected using on-site sensors installed at the solar energy generation facility. The sensors measured atmospheric and solar conditions at different altitudes and positions, enabling a detailed analysis of the environmental factors affecting solar power output.

Column Name	**Description**
`temperature_2_m_above_gnd`	Air temperature at 2 meters above ground (°C).
`relative_humidity_2_m_above_gnd`	Relative humidity at 2 meters above ground (%).
`mean_sea_level_pressure_MSL`	Atmospheric pressure at mean sea level (hPa).
`total_precipitation_sfc`	Total accumulated precipitation at the surface (mm).
`snowfall_amount_sfc`	Accumulated snowfall at the surface (mm).
`total_cloud_cover_sfc`	Total cloud cover at the surface (%).
`high_cloud_cover_high_cld_lay`	High-level cloud cover (%).
`medium_cloud_cover_mid_cld_lay`	Medium-level cloud cover (%).
`low_cloud_cover_low_cld_lay`	Low-level cloud cover (%).
`shortwave_radiation_backwards_sfc`	Downward shortwave solar radiation at the surface (W/m^2).
`wind_speed_10_m_above_gnd`	Wind speed at 10 meters above ground (m/s).
`wind_direction_10_m_above_gnd`	Wind direction at 10 meters above ground (°).
`wind_speed_80_m_above_gnd`	Wind speed at 80 meters above ground (m/s).
`wind_direction_80_m_above_gnd`	Wind direction at 80 meters above ground (°).
`wind_speed_900_mb`	Wind speed at the 900 millibar pressure level (m/s).
`wind_direction_900_mb`	Wind direction at the 900 millibar level (°).
`wind_gust_10_m_above_gnd`	Wind gusts at 10 meters above ground (m/s).
`angle_of_incidence`	Angle between incoming solar rays and panel surface (°).
`zenith`	Solar zenith angle: angle between the sun and vertical direction (°).
`azimuth`	Solar azimuth angle: sun's direction relative to true north (°).
`generated_power_kw`	Power generated by the solar installation (kW).

2.3 Purpose and Applications

This dataset is suitable for:

- Solar energy forecasting and yield estimation.
- Weather impact analysis on solar power generation.
- Training machine learning models for predictive maintenance and optimization.
- Assessing the influence of meteorological factors on energy output.

3 Evaluation Metrics

To rigorously evaluate the performance of the solar power prediction models, four standard regression metrics were employed: Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2 Score). These metrics collectively assess the accuracy, robustness, and explanatory power of the predictive models.

3.1 Mean Absolute Error (MAE)

MAE calculates the average magnitude of absolute errors between the actual and predicted values. It provides a more interpretable error metric in terms of physical units and is less sensitive to large deviations.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

3.2 Root Mean Squared Error (RMSE)

RMSE is the square root of the MSE, returning the error in the same units as the target variable (solar power). It also gives greater weight to large errors and is useful for comparing model performance in real-world terms.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

3.3 Coefficient of Determination (R^2 Score)

The R^2 score measures the proportion of variance in the actual solar power output that is explained by the model's predictions. A value close to 1 indicates a high level of predictive accuracy.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

4 Comparative Analysis of Prediction Models

This section presents a comprehensive evaluation of six predictive models applied to the task of solar power generation forecasting. The models assessed include: Linear Regression, Random Forest, Gradient Boosting Regressor, Support Vector Regression (SVR), SPFNet Artificial Neural Network (ANN), and Long Short-Term Memory Network (LSTM). The evaluation metrics used are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2), which measures how well predictions approximate actual outcomes.

Linear Regression

Serving as a baseline model, Linear Regression achieved the highest RMSE (0.3011) and MAE (0.4267), with a relatively modest R^2 score of 0.7061. This model is computationally efficient and interpretable but lacks the capacity to capture complex, nonlinear relationships in the data.

Random Forest Regressor

Random Forest significantly improved prediction performance, achieving an RMSE of 0.1945, MAE of 0.2807, and an R^2 score of 0.8102. Its ensemble architecture allows it to handle nonlinearity and feature interactions effectively, making it robust and reliable for solar forecasting tasks.

Gradient Boosting Regressor

This model delivered a solid performance (RMSE = 0.2159, MAE = 0.3162, R^2 = 0.7893). Gradient Boosting's sequential error-correcting mechanism makes it efficient in reducing bias, although it requires more tuning and computational effort compared to Random Forest.

Support Vector Regression (SVR)

SVR demonstrated a balanced performance with an RMSE of 0.2300, MAE of 0.3200, and R^2 of 0.7759. While SVR can capture nonlinear patterns via kernel functions, it is sensitive to parameter tuning and becomes computationally expensive with large datasets.

LSTM (Long Short-Term Memory)

The LSTM model outperformed all others with the best RMSE (0.1300), MAE (0.2000), and the highest R^2 score (0.9546). Its strength lies in capturing temporal dependencies, making it especially effective for time-series forecasting such as solar power prediction. Despite being computationally intensive, LSTM's predictive accuracy justifies its use in high-stakes energy management systems.

Summary

The comparison reveals that deep learning approaches, particularly LSTM, offer the highest prediction accuracy. Ensemble methods like Random Forest also demonstrate strong performance and reliability. Simpler models such as Linear Regression, though useful for quick benchmarking, fall short in accuracy for production-level solar power forecasting.

Table 3: Performance Comparison of Prediction Models

Model	RMSE	MAE	R^2 Score
Linear Regression	0.3011	0.4267	0.7061
Random Forest	0.1945	0.2807	0.8102
Gradient Boosting Regressor	0.2159	0.3162	0.7893
Support Vector Regression	0.2300	0.3200	0.7759
LSTM	**0.1373**	**0.2984**	**0.9546**

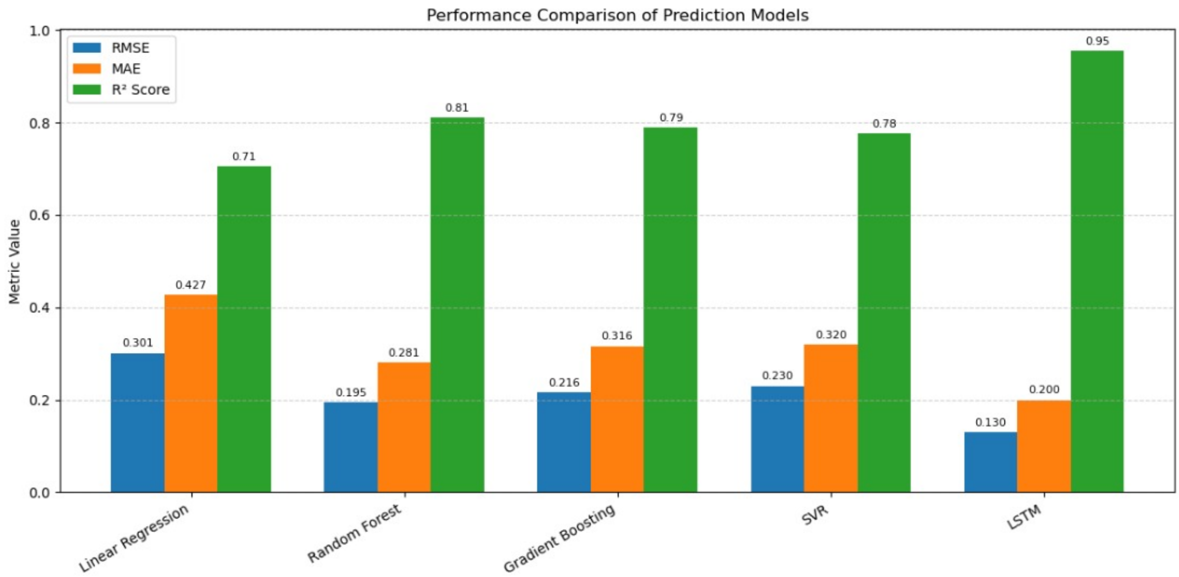

Figure 1: Visual comparison of model performance metrics (RMSE, MAE, R^2).

4.1 LSTM Architectures for Enhanced Solar Power Prediction

After conducting a comparative analysis of various regression models, it was observed that the Long Short-Term Memory (LSTM) model significantly outperformed traditional machine learning models in forecasting solar power output. To understand the reasoning behind this superior performance, and to explore other promising architectures, the following section details the operational principles of LSTM and Bi-LSTM models in the context of sequence prediction tasks such as solar energy forecasting.

4.1.1 Long Short-Term Memory (LSTM):

Long Short-Term Memory (LSTM) networks are a specialized type of Recurrent Neural Network (RNN), designed to learn long-term dependencies in sequential data. Traditional RNNs often suffer from the vanishing gradient problem, which makes it difficult to capture information across long sequences. LSTM addresses this by using memory cells and gated mechanisms to control the flow of information. LSTM networks process input x_t and the previous hidden state h_{t-1} at each time step, using three gates: input, forget, and output gates. These gates regulate how information is stored in or discarded from the memory cell c_t.

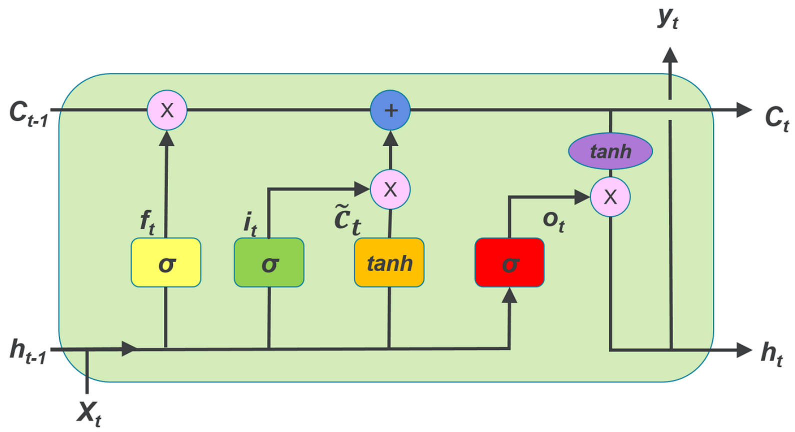

Figure 2: LSTM architecture

Gate equations:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \mathbf{x}_t + \mathbf{U}_f \mathbf{h}_{t-1} + \mathbf{b}_f) \tag{1}$$
$$\mathbf{i}_t = \sigma(\mathbf{W}_i \mathbf{x}_t + \mathbf{U}_i \mathbf{h}_{t-1} + \mathbf{b}_i) \tag{2}$$
$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{U}_c \mathbf{h}_{t-1} + \mathbf{b}_c) \tag{3}$$
$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \tag{4}$$
$$\mathbf{o}_t = \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{U}_o \mathbf{h}_{t-1} + \mathbf{b}_o) \tag{5}$$
$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \tag{6}$$

Symbols:σ is the sigmoid activation function. tanh is the hyperbolic tangent function. $\odot$ denotes element-wise multiplication. W, U are weight matrices and b are bias vectors. c_t is the cell state, h_t is the hidden state.

In solar power prediction, LSTM models are trained on time series data such as irradiance, temperature, and historical power output. The memory cell acts as a contextual buffer, allowing the network to understand patterns influenced by weather dynamics, seasonal cycles, and diurnal variations. This ability to model non-linear and long-range temporal dependencies makes LSTM particularly effective in forecasting photovoltaic (PV) power generation under diverse environmental conditions.

5 Load Management System

5.1 Need for Load Management System

The integration of renewable energy sources, particularly solar power, into residential energy systems introduces variability and unpredictability in power generation due to changing weather conditions and diurnal cycles. This intermittency poses challenges in ensuring a consistent and reliable power supply, especially when attempting to minimize reliance on conventional grid or generator backups. To address these challenges, implementing an intelligent Load Management System (LMS) becomes essential. An LMS can dynamically prioritize and schedule electrical loads based on the availability of renewable energy, storage capacity, and real-time consumption demands. By categorizing loads into essential and non-essential, the system ensures that critical functions remain uninterrupted while optimizing the use of available solar energy. Such strategic load management not only enhances energy efficiency but also contributes to cost savings and reduced environmental impact by decreasing dependence on non-renewable energy sources. Therefore, integrating an LMS is crucial for maximizing the benefits of renewable energy systems and achieving sustainable energy management in residential settings.

5.2 Load Classification Strategy

Efficient energy management, especially in systems integrating renewable sources like solar power, relies heavily on the classification of electrical loads. Load classification strategies aim to prioritize energy distribution based on the criticality and flexibility of connected devices or systems. In this project, the loads are broadly categorized into two main classes: Essential Loads and Non-Essential Loads.

Essential Loads : These are the critical loads that must always be powered to ensure the basic functioning and safety of the system or premises. Examples include:

- Emergency lighting
- Communication equipment
- Medical devices (in hospitals)
- Server infrastructure or control units
- Basic lighting and ventilation

These loads are given the highest priority in energy allocation. The Load Management System ensures that these are never disconnected, even under constrained energy availability, by relying on backup power if necessary.

Non-Essential Loads : Non-essential loads are those that can be deferred, reduced, or completely cut off during periods of low power availability. These include:

- Air conditioning units
- Water heaters
- EV chargers
- Recreational and decorative lighting
- Auxiliary machinery

By shedding or throttling these loads during energy shortfalls, the system can prevent outages of essential loads and extend the usability of battery reserves.

5.3 State Transition Mapping

The State Transition Diagram illustrates the operational flow and decision-making process of the proposed Load Management System (LMS), which is driven by solar power prediction, battery storage status, and estimated load demand. The diagram maps out how the system transitions between different operational states on a daily basis, ensuring optimal utilization of renewable energy while maintaining essential load supply.

5.3.1 System States

The system operates in three main states, based on the Viability Score—a computed value derived from solar power prediction, current battery capacity, and load estimate:

All Loads Supplied (Normal Operation) : In this state, the energy supply (from solar and battery) is sufficient to meet both essential and non-essential load demands. The system functions under optimal conditions with no need for throttling or backup power.

Non-Essential Load Cut (Partial Curtailment) : This state indicates a moderate viability condition where energy supply is insufficient for full load demand but enough for essential services. Non-essential loads are curtailed or deferred to prioritize critical operations.

Switch to Backup Power (Critical Condition) : Under low viability, the system cannot support even the essential loads with available solar and battery power. In this state, the LMS triggers the use of backup sources such as the grid or generators to ensure continuity of essential operations.

5.3.2 Transitions

Each transition in the diagram corresponds to a daily evaluation cycle, as the system processes the inputs and updates its operational state accordingly:

Compute Viability : The system receives input parameters-Load Estimate, Current Battery Capacity, and Predicted Solar Generation (via the Solar Power Prediction Model fed by Weather Data). It computes a Viability Score using a weighted formula.

Update System State : Based on the computed viability, the system transitions to one of the three defined operational states (0, 1, or 2), each associated with specific load-handling logic.

Record Present Day System State : The system logs the current day's state for performance monitoring and analysis purposes, enabling retrospective evaluation and learning.

Update Battery Capacity : After the load is served (or backup is used), the system updates the battery status to reflect discharge/charge activity, preparing for the next day.

Next Day Prediction Feedback Loop : The updated battery data are fed into the Next Day Prediction loop, which utilizes the solar forecasting model and future load estimates to consider residual battery energy for the next n days.

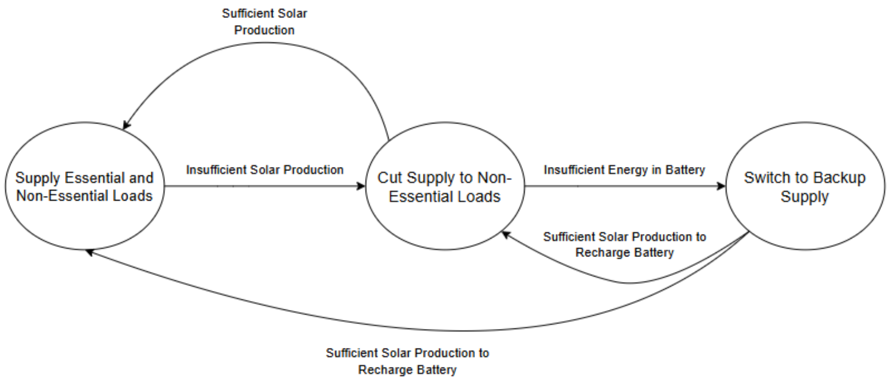

Figure 3: State Transition Diagram

5.4 System Architecture

5.4.1 Overview

The proposed Energy Management System (EMS) integrates day-ahead solar power prediction, battery state monitoring, and load estimation into a unified control framework. At a high level, weather and historical irradiance data are fed into an LSTM module, which produces a 24-hour solar generation forecast. That forecast is combined with instantaneous battery energy and a load estimate to compute a single "viability score." A lightweight lookup table then maps that score into one of three operating modes—full load supply, non-essential load curtailment, or backup power activation. Finally, the chosen state is recorded for analysis and the battery energy is updated to reflect that day's net charge or discharge. This feedback loop repeats each day, adapting to changing weather, consumption patterns, and storage levels.

5.4.2 Data Collection Phase

In the Data Acquisition layer, the system gathers three streams of input every control cycle. First, weather data (solar irradiance, temperature, cloud cover, etc.) is pulled from a historical dataset of weather pattern in Shillong. Second, a battery monitor reports the current State of Charge (SoC) in kilowatt-hours. Third, the daily load estimate is provided by the operator or derived from historical consumption patterns stored in a local database. All raw inputs are stored in memory for the next processing step.

5.4.3 Processing (Forecasting) Phase

The Processing layer transforms raw weather and irradiance timeseries into a reliable solar generation forecast using a Long Short-Term Memory (LSTM) neural network. Prior to training, historical data is, normalized to unit scale, and windowed into sequences of past N days. During runtime, the pre-trained LSTM model—implemented in TensorFlow/Keras—consumes the most recent N-day weather window and outputs a point estimate of next-day solar energy production (in kWh). This predicted value is then fed forward into the decision logic, providing a proactive estimate of renewable supply for the coming 24-hour period.

5.4.4 Decision Logic Phase

The Decision Logic module synthesizes the forecasted solar yield, current battery SoC, and the operator's load estimate into a single viability score:

$$viability = SoC_{kWh} + PredictedSolar_{kWh} - LoadEstimate_{kWh} \tag{7}$$

This scalar score is passed to a table-driven lookup engine—loaded from a simple JSON or Python list—where predefined ranges map "viability" into three states:

- All Loads Supplied (viability >= 0.4)

- Non-Essential Curtailed (0.1 <= viability ¡ 0.4)
- Backup Activated (viability ¡ 0.1)

5.4.5 State Update Phase

Once the decision logic selects an operating mode, the EMS issues control commands—e.g., relay open/close signals or inverter setpoints—to implement the new state. Simultaneously, it logs the day's chosen state along with the input parameters and resulting viability score to a CSV file for later analysis.

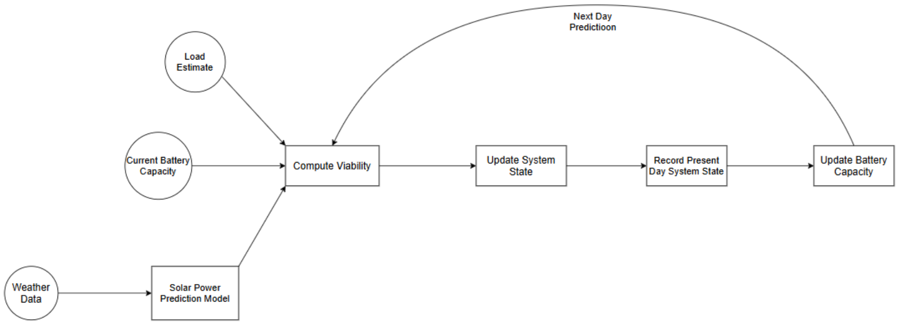

Figure 4: System Control Flow Diagram

5.5 System Operation and Visualization

5.5.1 Feedback Loop Execution

After computing the system state for a given day based on the solar power prediction, current battery SoC, and estimated load, the Load Management System (LMS) updates the battery's SoC based on actual usage. This updated SoC is then passed to the next day's data collection phase, along with new solar power predictions and load estimates, forming a continuous feedback loop.

This cycle is repeated for 30 consecutive days, allowing the system to demonstrate its ability to dynamically manage load prioritization and energy resources over time. The system's adaptability is particularly evident in days with low solar irradiance or high demand, where it automatically curtails non-essential loads or switches to backup power.

5.5.2 Daily State Transition Visualization

The figure below illustrates the daily state transitions of the LMS over a simulation period of 30 days. Each point represents the operating state on a given day—categorized as:

- State-0: Supply both Essential and Non-Essential Loads
- State-1: Cut Supply to Non-Essential Loads
- State-2: Switch to Backup Supply

This visual representation enables system operators to evaluate the effectiveness and responsiveness of the LMS under varying environmental and load conditions.

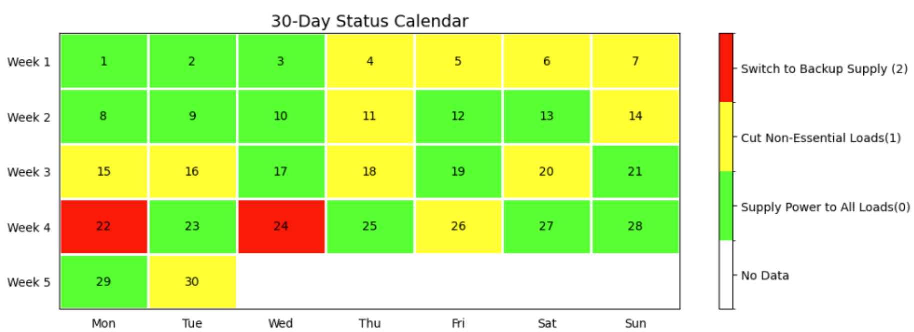

Figure 5: Operational State Visualization Over 30 Days

Conclusion

This project presents an integrated solution for efficient energy utilization in solar-powered systems by combining machine learning-based forecasting with intelligent load management. Using an LSTM model trained on real-world data from NIT Meghalaya, the system predicts daily solar power generation, which, along with current battery reserve and estimated load demand, drives the decision-making process for optimal power distribution. The Load Management System (LMS), implemented in Python, categorizes each day into one of three operational states—supplying all loads, cutting non-essential loads, or switching to backup power—ensuring that essential services remain uninterrupted while minimizing dependency on external sources. A feedback loop updates the battery's state of charge (SoC) daily, enabling realistic multi-day simulation and dynamic adaptation to varying environmental conditions. A 30-day simulation visualizes the system's behavior, offering valuable insights for operators. This approach demonstrates the potential of integrating predictive models and rule-based control to build sustainable, adaptable energy systems.

Future Work

- Adding machine learning models to predict future load demand based on historical usage patterns can allow the control system to plan even more efficiently, avoiding blackouts and battery overdischarge.
- Implement more detailed load profiling that includes time-of-day consumption patterns, user preferences, and critical load thresholds to allow more control.
- Replace the rule-based lookup table with a fuzzy logic controller or trained reinforcement learning agents to make more adaptive, real-time decisions.
- Develop a simple user interface or dashboard to visualize predicted solar generation, current battery level, and the daily system state for easier monitoring by users or operators.

References

1 ”Deep Learning Models for PV Power Forecasting” (https://www.mdpi.com/1996-1073/17/16/3973)

2 ”Load Management System with Integration of Renewable Energy Resources” (https://www.researchgate.net/publication