**Features of Data**

The dataset utilized in this research comprises three key environmental features like temperature, dew point, and humidity. These variables are crucial in understanding and predicting the performance and efficiency of solar panels. Temperature indicates the ambient conditions surrounding the panels and plays a significant role in determining the power output, as higher temperatures can reduce the efficiency of photovoltaic cells. Dew point reflects the moisture level in the air, influencing condensation and the potential for dew formation on panel surfaces, which can obstruct sunlight and reduce power generation. Humidity measures the amount of moisture in the atmosphere, impacting both the clarity of the air (and thus sunlight penetration) and the potential for condensation or corrosion of panel components.

Temperature fluctuations can cause thermal stress on solar panels, leading to material degradation and reduced lifespan. When temperatures rise significantly above optimal operating levels, the electrical efficiency of panels drops due to increased resistance in the photovoltaic cells. Additionally, periods of high dew point can lead to the accumulation of moisture on panels, particularly during early mornings. This moisture may result in dust and dirt particles adhering to the surface, reducing the overall transmittance of sunlight. Humidity levels also affect solar irradiance; higher humidity leads to a denser atmosphere, causing light scattering and reducing the amount of solar energy reaching the panels.

The combined impact of these features on the efficiency of solar panels underscores the importance of monitoring and maintaining optimal environmental conditions. Temperature, dew point, and humidity levels must be carefully analysed to anticipate potential performance issues and schedule cleaning or maintenance activities effectively. By incorporating these features into a machine learning model, it becomes possible to predict the impact of environmental conditions on solar panel efficiency, allowing for proactive and data-driven maintenance strategies that optimize energy production.

**ML Algorithm for Dashboard and Trend Analysis**

For the solar panel care project, we use a Time Series Analysis Algorithm to forecast key environmental features like temperature, dew point, and humidity over time. This approach helps predict future trends and identify any significant deviations that may affect the efficiency of solar panels. The time series model, such as ARIMA (AutoRegressive Integrated Moving Average) or SARIMA (Seasonal ARIMA), is used to analyse historical data and generate accurate forecasts. These algorithms are chosen due to their capability to model temporal dependencies and seasonality, which are common in environmental data. The forecasted values are then visualized on the dashboard, providing a real-time trend analysis and alerting users when predicted conditions may negatively impact panel performance.

To supplement the forecasting model, we employ a Machine Learning Classification Algorithm like Logistic Regression, Random Forest, or Support Vector Machine (SVM) to determine whether the solar panel needs cleaning. This classification model uses the forecasted features (temperature, dew point, and humidity) as input variables. The algorithm classifies the need for cleaning based on threshold values or patterns learned from historical data, such as high humidity levels or consistent dew formation, which may lead to panel surface obstruction. This decision-making model allows for quick and accurate predictions, enabling proactive scheduling of maintenance tasks to optimize energy output.

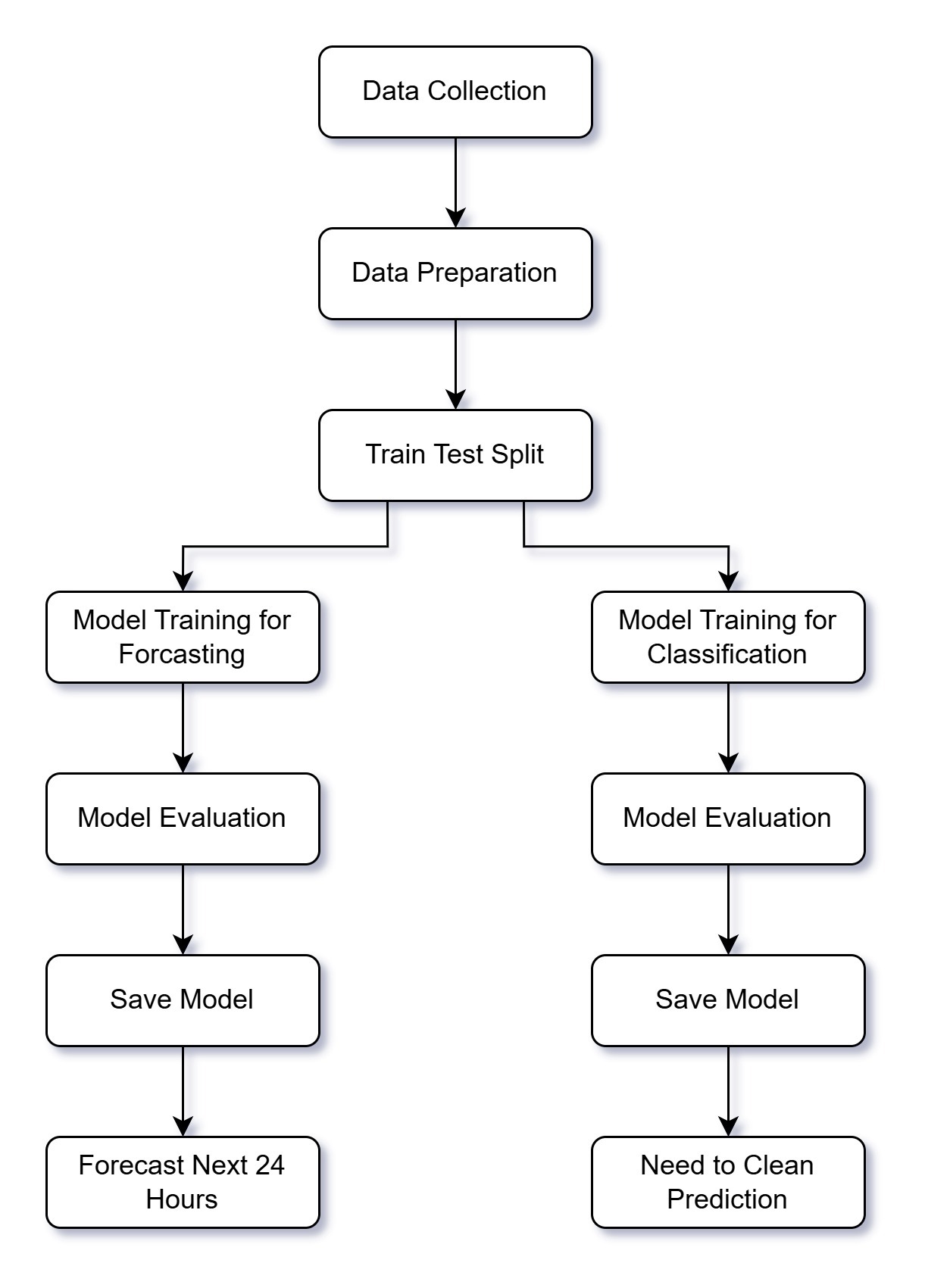
For the dashboard and trend analysis, a combination of visualization tools, such as Plotly or Matplotlib, integrates the forecasted data and classification outcomes. These tools display dynamic charts, showing both historical trends and future forecasts for temperature, dew, and humidity levels. The dashboard not only visualizes these trends but also indicates the probability of panel cleaning being required, based on the classification model’s output. Users can interact with the dashboard to explore specific time frames, view seasonal patterns, and monitor the status of solar panel efficiency in real-time.

The combination of time series forecasting and machine learning classification allows for a comprehensive analysis and maintenance strategy. By leveraging both temporal trends and classification insights, the system ensures that solar panels are cleaned only when necessary, minimizing downtime and maximizing energy production. This integrated approach provides users with a robust tool to manage solar panel efficiency and improve overall system performance, demonstrating the effectiveness of predictive maintenance in renewable energy management.

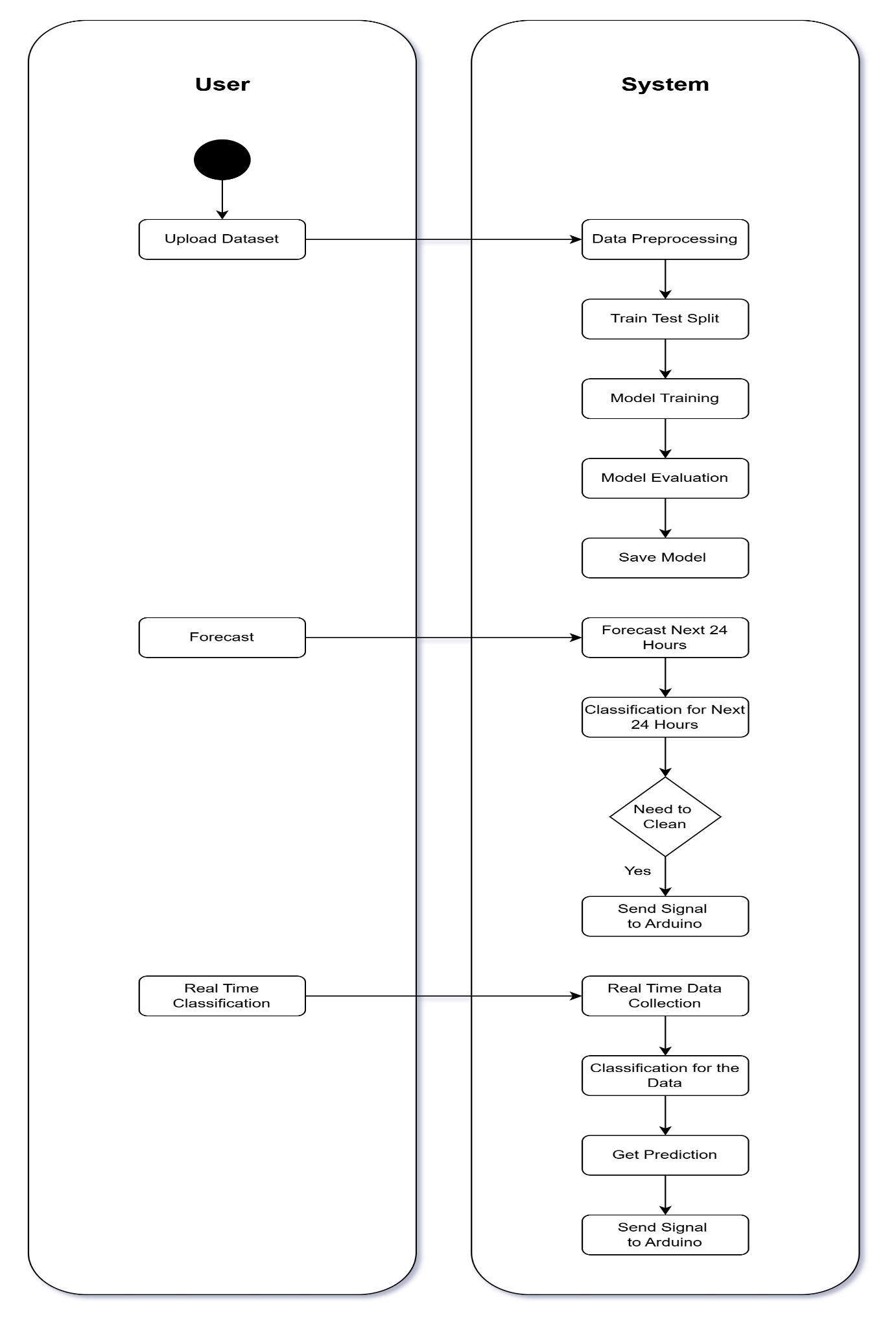
**Recommendations/Predictions by the ML Model**

The time series model in this project forecasts future environmental conditions, such as temperature, dew point, and humidity, which are critical factors affecting solar panel performance. By analysing historical data, the model predicts these conditions for the upcoming hours or days, providing insights into potential environmental changes that may impact solar panel efficiency. For instance, if the forecast indicates high humidity levels or a sharp rise in dew point, the system can predict an increased likelihood of dew or moisture accumulation on the panel surfaces. Such predictions allow users to anticipate these conditions and take preventive measures, like scheduling panel cleaning during optimal times or deploying automated cleaning mechanisms when needed.

The machine learning (ML) classification model uses these forecasted features to determine whether the solar panels need cleaning or maintenance. Based on patterns learned from past data, the model identifies conditions that typically lead to performance degradation, such as consistent high humidity or high temperature accompanied by a high dew point. When such conditions are predicted, the ML model outputs recommendations like “clean panels within the next 12 hours” or “schedule maintenance within the week.” These predictions are crucial for optimizing energy output, as they help minimize the impact of environmental factors on panel efficiency. The combination of time series forecasting and ML classification ensures that cleaning and maintenance actions are both timely and data-driven, enhancing overall system reliability and performance.



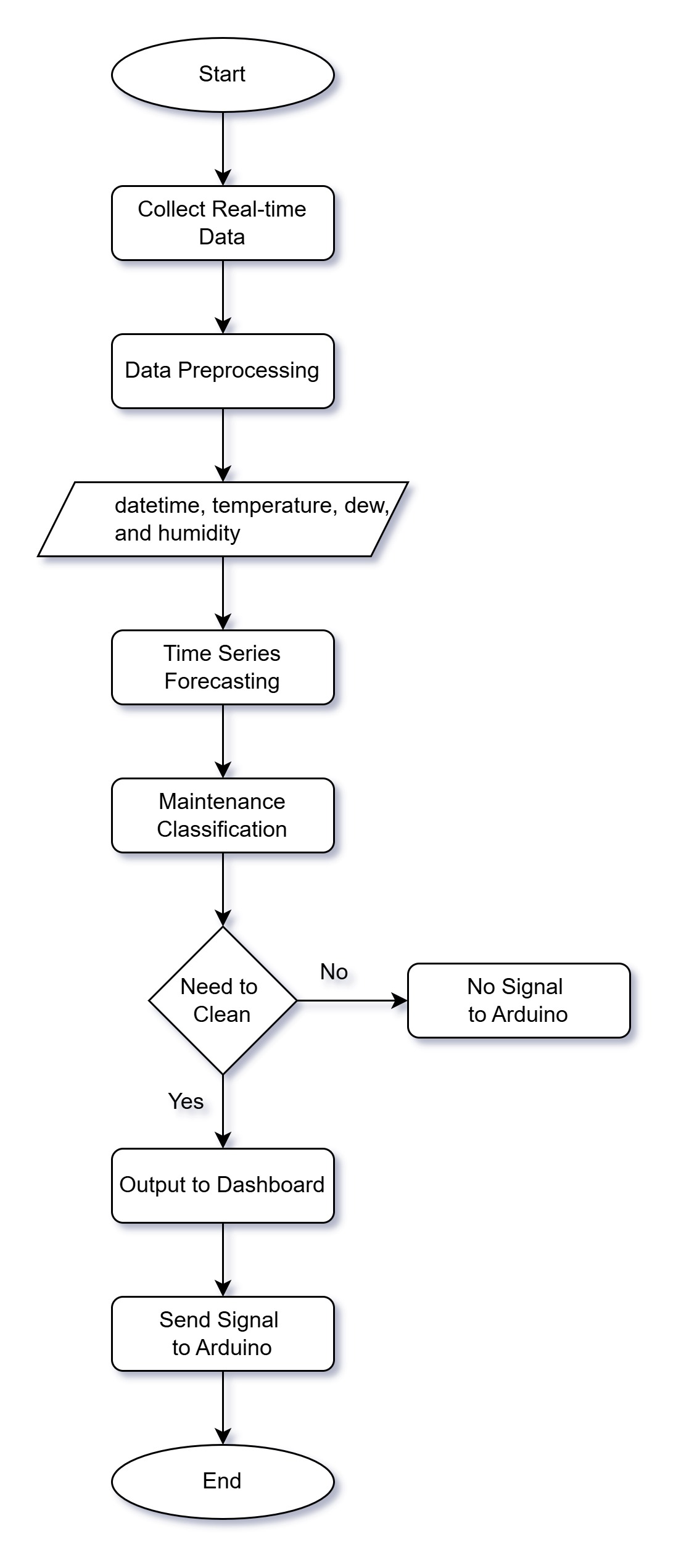
Block Diagram



UML / Activity Diagram

Software Description

The application is designed to predict and monitor the need for solar panel maintenance, focusing on environmental conditions that affect panel efficiency. It uses real-time data on temperature, dew, and humidity and applies time series forecasting and machine learning to anticipate the impact of these conditions on solar panel performance. To achieve this goal Python is chosen for its vast ecosystem of libraries for data analysis, time series forecasting, machine learning, and ease of integrating with web-based applications.



Flowchart

Algorithms

1. Data Preprocessing Algorithm

This algorithm cleans and preprocesses the data for forecasting and classification.

Steps:

- Input: Raw environmental data (temperature, dew, humidity)

- Check for missing values and handle them.

- Generate additional features (lags or moving averages for time series).

- Output: Cleaned and feature-enhanced data.

2. Time Series Forecasting Algorithm (SARIMA)

This algorithm is used to forecast future values of temperature, dew, and humidity based on past trends.

Steps:

- Input: Historical time series data for each parameter.

- Define SARIMA model parameters (p, d, q) for auto-regression, differencing, and moving average.

- Train the SARIMA model on historical data.

- Forecast the next 24-hour values for each parameter.

- Output: Predicted temperature, dew, and humidity values for the next 24 hours.

3. Classification Algorithm (Random Forest)

This algorithm classifies the need for solar panel cleaning based on forecasted environmental data.

Steps:

- Input: Forecasted temperature, dew, and humidity values.

- Train a Random Forest classifier on historical labeled data (indicating if cleaning was required under similar conditions).

- Predict the maintenance label (0 = no cleaning required, 1 = cleaning required).

- Output: Maintenance decision recommendation.