Name: Ravikumar Savani
Student ID: 685167
Assignment: KIT717 – Assignment 2
Title: IoT-Enabled Weather Lamp using Machine Learning
Date: 06/06/2025

Section A: Explanation of Machine Learning Method

For this assignment, I used the Least Squares Regression algorithm from the PHP-ML library to predict the minimum and maximum temperature and humidity for a selected day and site.

Why Least Squares:
Least Squares is a simple and effective regression model suitable for forecasting numerical values, which matches our task of predicting temperature and humidity based on historical time-series data. It does not require large feature engineering and works well with a small dataset, making it appropriate for this context.

Input and Output:

- **Input to the model:** Hour of the day (in decimal format, e.g., 13.5 for 1:30 PM)
- **Target output:** Temperature or Humidity values

The model was trained using filtered historical data from the cleaned dataset. Each row included DateTime, Temperature, Humidity, and SiteID. I filtered only those rows that matched the user's selected day, month, and site, across all years from 2015–2021.

After training, I used the model to predict temperature and humidity at 48 half-hour intervals across a full 24-hour period. These values were then used to calculate the minimum and maximum for the day.

Section B: Discussion of Prediction Accuracy

The predictions provided by the regression model are reasonably accurate for identifying general trends in temperature and humidity over the selected day. However, there are several factors that affect the accuracy:

Factors Affecting Accuracy:

- **Limited Features:** The model only uses time of day as input. It does not consider other factors like season, wind, or weather patterns.
- **Outliers and Incomplete Data:** Some records in the historical dataset may contain noise or unusual values, which can influence the model's output.
- **Daily Variation:** Weather data can be highly variable even for the same day across years, reducing the ability of a simple model to generalize accurately.

Possible Improvements:

- Use a more advanced model like Random Forest or Support Vector Regression.
- Include additional input features like month, average wind speed, or a “season” indicator.
- Perform additional preprocessing to remove outliers or normalize values.

Despite the limitations, the model provides a solid first-level prediction that works well for the scope of this project.