# **Project Report**

# **1.Introduction:**

A greenhouse is a structure with walls and roof made chiefly of transparent material, such as glass, in which plants requiring regulated climatic conditions are grown. These structures range in size from small sheds to industrial-sized buildings. A miniature greenhouse is known as a cold frame. The interior of a greenhouse exposed to sunlight becomes significantly warmer than the external temperature, protecting its contents in cold weather. In this project we aim to develop an IOT based solution to classify the readings using a machine learning algorithm and then raise an alarm if the sample belongs to class 1 and un-raised the alarm if it belongs to the class 0.



Figure . A greenhouse

# **2.Task:**

Our task is to train a machine learning algorithm to classify the data obtained from the sensors and raise the alarm when the data is belongs to class 1 and unraised the alarm when the data belongs to class 0.

# **3.Condition to Turn Red Led On:**

The red led will be turned on when the following condition is met:

1. Point classified by the machine learning algorithm belongs to class 1

# **4.Data Collection Methodology:**

A total of 1300+ data is collected. The data collection methodology is explained in the following sections:

# Light Data:

Data obtained from the sensors after the conversion is send by the client to the border-router from where the data is send to the localhost where the light data is stored in a pickle format file.

# Humidity Data:

Since the value of humidity is independent of temperature and light (rainforests where light and temperature are low, but humidity is high), humidity value is generated by generating a random number between 0 and 100. The following snippet of code is used to generate the humidity data:

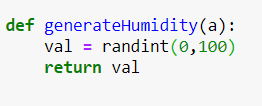


Figure 2.Humidity Generation Code

# Temperature Data:

Temperature is generated by developing a correlation with the light intensity. The following script is used to generate temperature values:

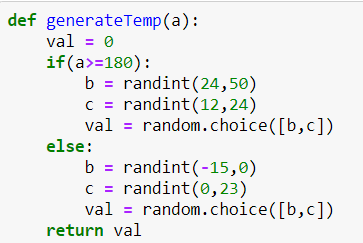


Figure 3.Temperature Generation Code

As seen from the above script if the light intensity is greater than or equal to 180 then two random numbers between the ranges of (24,50) and (12,24) are generated of which only one is then randomly selected. Otherwise, the temperature value is generated by generating two random numbers between ranges of (-15,0) and (0,23) of which then one is randomly selected. This introduces randomization in the data so that overfitting does not happen.

# **5.Data Storage and Labelling:**

Data is stored in a pickle format file “LabelledData”. The following labelling strategy is employed:

* + 1. If the light value is greater than 180 or temperature is greater than 24 then data belong to class 1 otherwise the data belong to class 0. Class 1 is where the alarm needs to be raised and class 0 is where the alarm does no need to be raised. Humidity value is not used because humidity does not have any such effect on the health of plants.

# **6.Machine Learning Algorithm:**

The following machine learning algorithm is used:

# Logistic Regression:

Logistic regression is a non-linear classification algorithm used in this project because real world data is not always linearly separable and non-linear functions also have a linear part so they can work in both cases.

# Train-Test Split:

A train-test split of 80-20 is used for the training of machine learning algorithm.

# Training Accuracy:

A training accuracy of 98% was obtained with L2-regularization strategy.

# Testing Accuracy:

A testing accuracy of 98% is obtained on the 20% split used.

# **7.Latency:**

Classification is done on every incoming sample rather than a batch of samples. The time it takes for classification is of milli-seconds magnitudes, but it varies from 0.3ms to 0.1 ms.