# CHAPTER 4

## PROJECT IMPLEMENTATION

* 1. **IMPLEMENTATION STEP**

In this work, Python is used to perform all of the experiments. Python is the most commonly used machine learning method. The selection of a dataset is important since the entire process is based on the fields and records. Steps for coding:-

Step 1:- Import all the files like pandas, numpy, and seaborn; train the model and visualize the data from pandas and numpy.

Step 2:- Load the dataset using read\_csv function. Step 3:- Display the top five rows.

Step 4:- Perform the explored data analysis.

Step 5:- check the number of data and print in dataset.

Step 6:- Check the null value or not then add all the number of each columns.

Step 7:- Check the information of dataset and describe the dataset, it display the count, mean, min, 25% value of each columns

Step 8:- Fill the null value using the mean value of each columns.

Step 9:- display the probability and check the value counts how much 0 and 1 are there.

Step 10:- visualize the correlation of all the feature of using the heap Map function of seapom.

Step 11: - Now see the out layer using the boxplot function. You can see the soload function contains the out layer but we cannot remove the out layer from the solid features, waste water safer to drink every time. But it contain the out layer for also the not safe to drink.

Step 12: - Divide the dataset except the probability X = contains all the independent features except the probability. Y = contain over target feature “wavelength”.

Step 13:- Split the dataset using the train and testing like some features like independent feature, dependent feature, test\_size is 20% , random\_state is used for ignore the suffle every time.

Step 14:- Train the model using the decision tree classifier and fit the model; fit means it layers tree from the training data.

Step 15:- Predict and check the model how it is performing on the test data and check the accuracy using the predicted dataset.

Step 16:- Predict only single row “how model perform only one dataset is one row”. Step 17:- Apply hyper parameter tuning on decision tree classifier.

Step 18:- Took the three parameter and then create the dictionary on three parameter Grid Service perform the hyper parameter tuning and get the output score.

## PROCEDURE DESCRIPTION

* + - Data pre-processing - The transformations applied to data before it is fed into the algorithm are referred to as data pre-processing. The method of transforming raw data into a clean data set is known as "data pre-processing." In other words, data collected from different sources is received in raw format, making analysis impossible. When using a model in Machine Learning projects, the data must be formatted correctly in order to get better results. Among the implemented Machine Learning algorithms, the best-suited Machine Learning algorithm is found and measured.
    - Classification - When the output has finite and discrete values, classification is the most efficient method of supervised learning since it defines which class data elements belong to. Furthermore, it assigns a class to an input variable. The supervised learning community includes classification. The goals were also shown the input data. Credit recognition, medical diagnosis, and target marketing are only a few of the applications for classification. The targets were also shown the input data.
    - Performance Evaluation - Every data science project requires evaluating the output of a machine learning model. The aim of machine learning model evaluation is to Figure out how well a model generalizes to a given dataset.

## ELECTROCOAGULATION ANALYSIS USING IOT

Electrocoagulation process starts which voltage is given to textile waste water container through anode and cathode plate. Then sludge is forming and then it settle down to bottom. During this process, user need to calculate the voltage, turbidity, pH value through the sensor, using IoT and value would directly send to online google excel. Figure 4.4, show parameter of real time data.

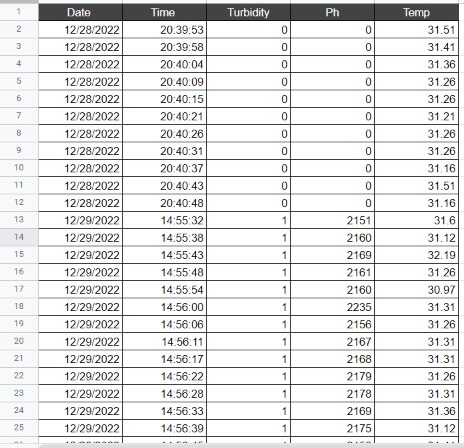


Figure 4.1 Dataset of Real Time through IoT

## EXPERIMENTAL DESCRIPTION

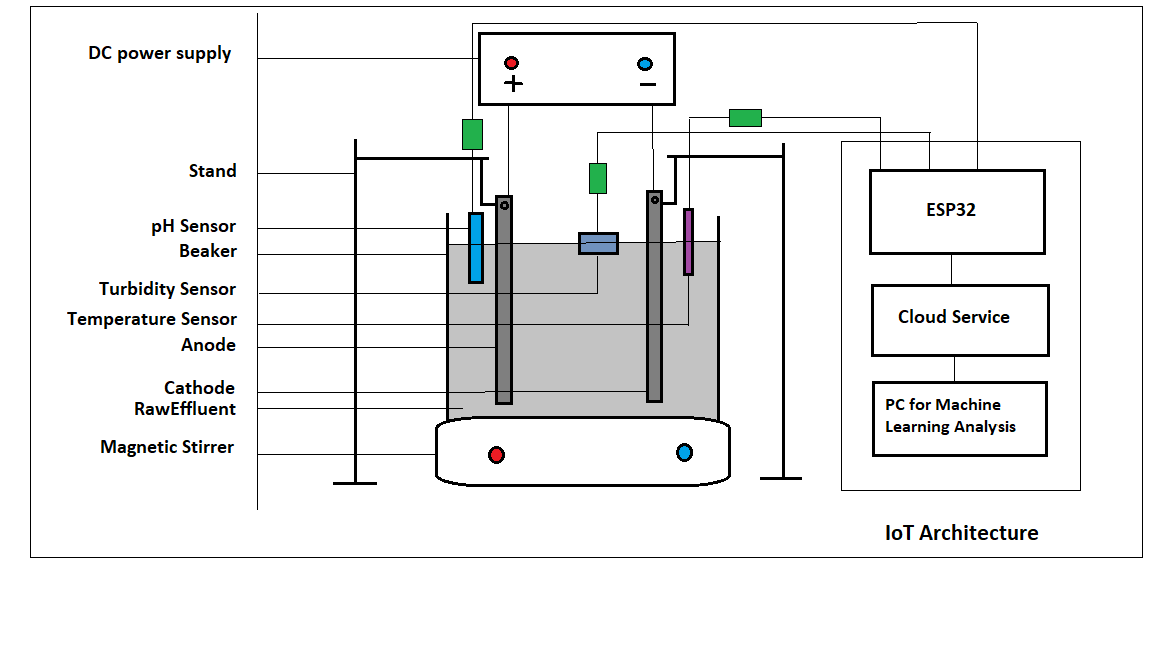


Figure 4.2 Experimental setup for Electrocoagulation Process

During the process of electro-coagulation process, the IoT sensor need to be connected with power supply and stand of beaker as shown in figure 4.2

Components - Dimension/Quality

Iron Electrodes - 21 cm \* 2.5 cm \*0.4cm

[Length\*breadth\*Thickness]

Number of Electrodes - 2

Beaker Capacity - 2L(1.8L of Textile waste water in 2L Beaker)

Distancebetween

the Electrode - 4.5 cm

Sensors - (Turbidity, pH,Temperature)