The document is the outline of how the project s carried out.

The thesis is a classification model in Bangladeshi weather data. The Machine learning algorithms used are run on the dataset that we extracted and modified, to classify between abnormal weather data that is impossible to reach in Bangladesh Context. The following explains the procedure followed.  
  
Sensor data are collected based on Bangladesh Context for 3.5 months.

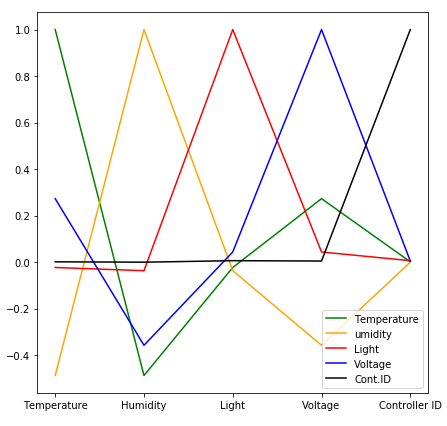
Data is stored in adafruit cloud server

Data is extracted in CSV Format

Data is pre-processed

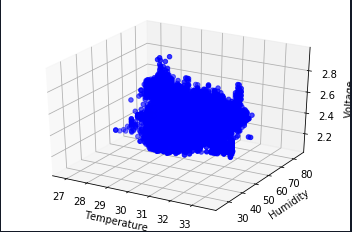
Feature Correlation is found out.

Feature with the least correlation with other features is discarded- Controller ID.



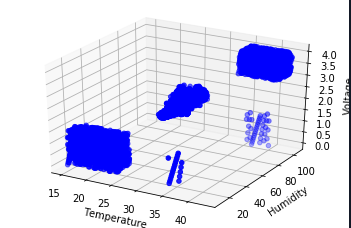
Data is Dropped from duplicates

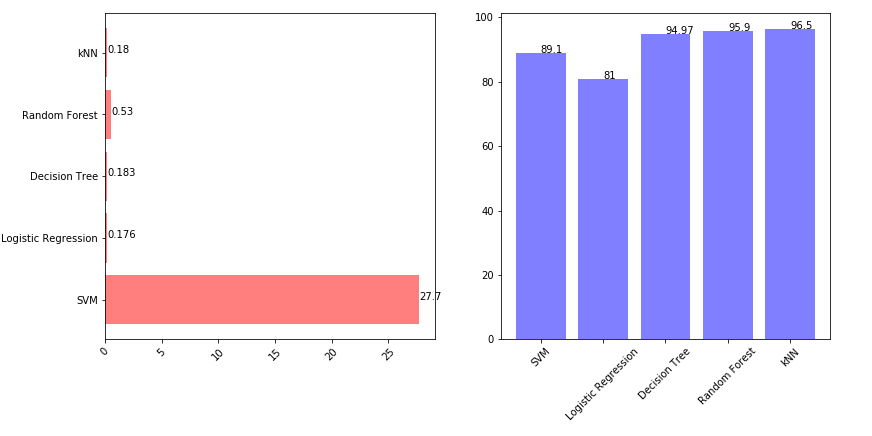
Data is visualized



Since these are sensor data, they are normal data. For performing classification between sensor data and malicious data we need to form malicious data which are not typically weather data and looks aberrant. For example with a temperature of 12 degrees we have a humidity of 90 percent. Or a Temperature of -25 degrees in Bangladesh, - these data standards are false in Bangladeshi Context.

Malicious Data are added artificially

Scatter Diagram is then plotted again

Results Graph on left shows time comparison and graph on right shows accuracy comparison