Assignment 4.2: Final Team Project Progress Report

TEAM 6: Bin Lu, Trevor McGirr, Olympia Saha

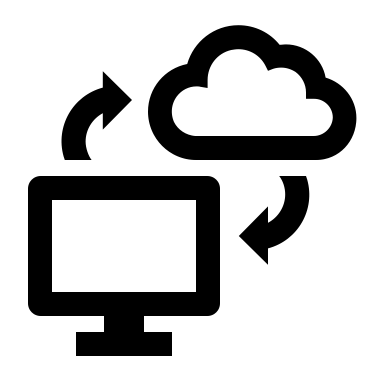
Shiley-Marcos School of Engineering, University of San Diego

AAI-530-03-SP23 – Data Analytics and Internet of Things

Professor An Tran

Feb 6, 2023

**Section 1. Aquaponics Fish Pond IoT System Infrastructure Diagram and Reference Document**



Sump Tank

External Antenna

Wired

Serial Data

Wired

Serial Data

Wired

Serial Data

Wired

Serial Data

Wired

Serial Data

Lead

Lead

Lead

Lead

Lead

Filtration System

Hydroponic Unit

ADC

5. Sensor

MQ-137 Ammonia Sensor

6. Sensor

MQ-135 Nitrate

Sensor

4. Sensor

DF Robot pH V2.2

SEN0161-V2

3. Sensor

DF Robot Dissolved Oxygen SEN0237

2. Sensor

DF Robot Turbidity SEN0189

HTTP

TCP/IP

External RAM

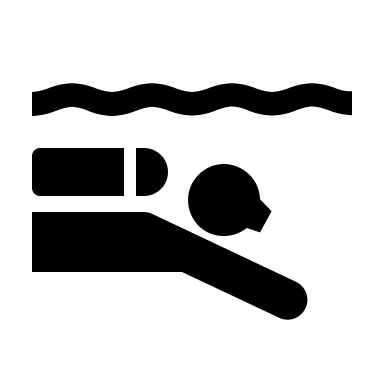
ESP 32 MCU integrated with WIFI and Bluetooth

1. Sensor

Dallas Instrument Temperature DS18D20

Measure temperature in the Air

Lead

Fish Pond



**Reference Documentation**

**Aquaponics Fish Pond IoT System:**

Aquaponics is a type of agriculture that combines raising fish and growing plants in a symbiotic environment. The system will measure the Temperature, Turbidity, Dissolved Oxygen level, pH value, Ammonia, and Nitrate. We got the sensor type from the data set. The biggest issue we ran into during the research was the type of sensor, type of signal, sensor placement, and the research on the use case of ESP 32 MCU.

1. Dallas Instrument Temperature sensor (DS18B20):

a) Placed to measure ambient

b) Aquaculture species ranging between 25.5 °C and 30.5 °C

c) Accuracy +/-0.5 ºC (between the range -10ºC to 85ºC)

2. DF Robot Turbidity (NTU value) sensor (SEN0189):

a) Placed in the fish pond underwater

b) Output Method: Analog; Will connect to ADC

3. DF Robot Dissolved Oxygen sensor (SEN0237):

a) Placed in the fish pond underwater

b) Maintenance Period: Once every month

4. DF Robot pH sensor V2.2 (SEN0161-V2):

a) Placed in the fish pond sump tank

b) Measurement Accuracy: ±0.1@25℃

c) Probe Life: >0.5 year

5: MQ-137 Ammonia sensor:

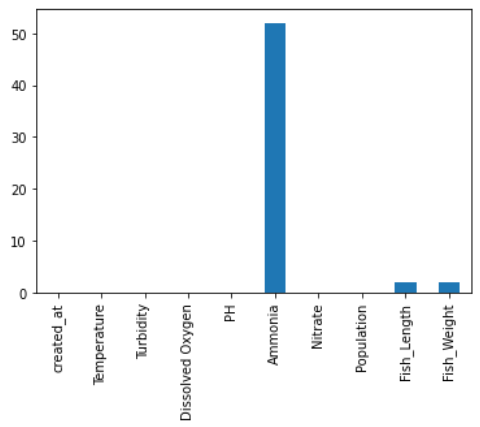
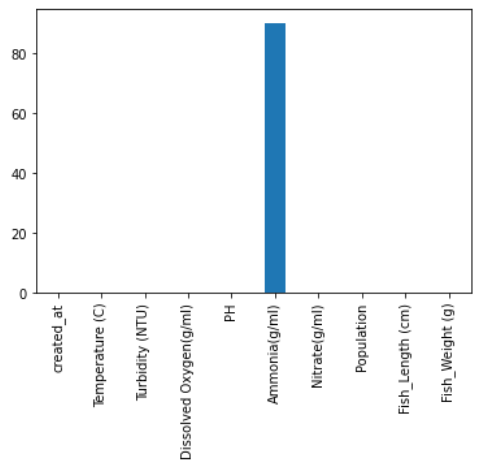
Placed in the filtration system

6. MQ-135 Nitrate sensor:

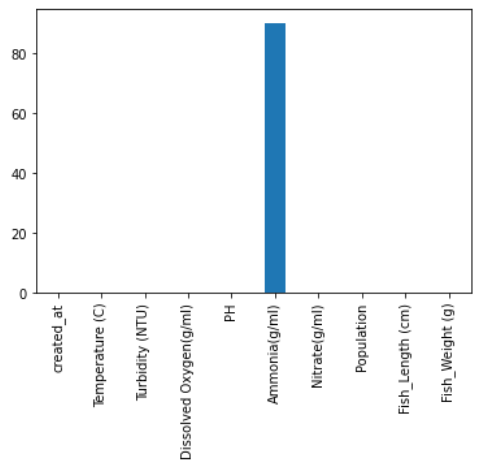
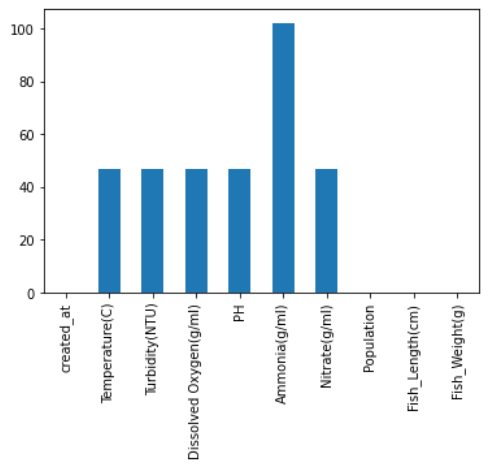
Set in the Hydroponic Unit

**Section 2. Dataset Exploration and Cleaning**

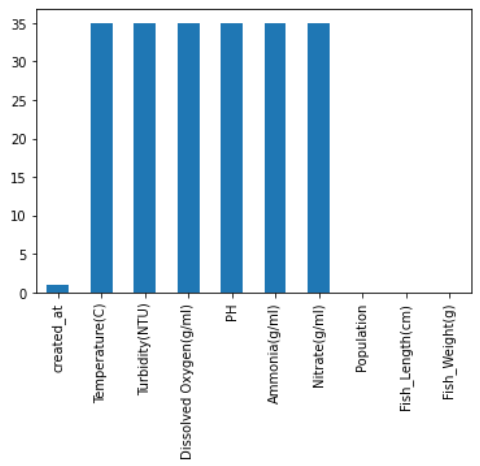
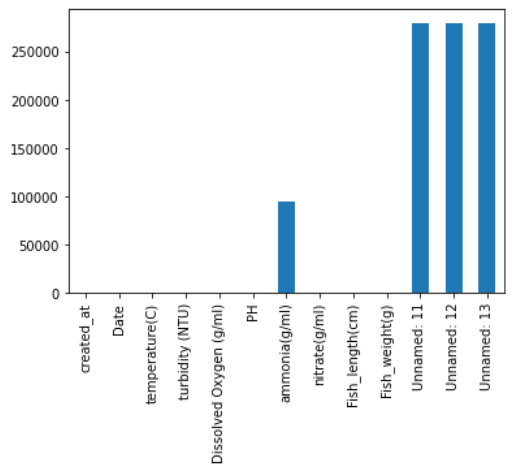
The data files are 12 comma-separated values (.csv) format files, each representing an aquaponics fish pond. However, the IoTPond5 dataset was found to need to be added once downloading the dataset. A graphical view of missing data in each dataset is provided below.

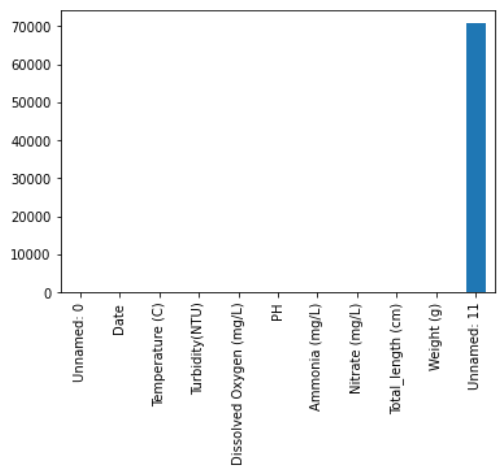
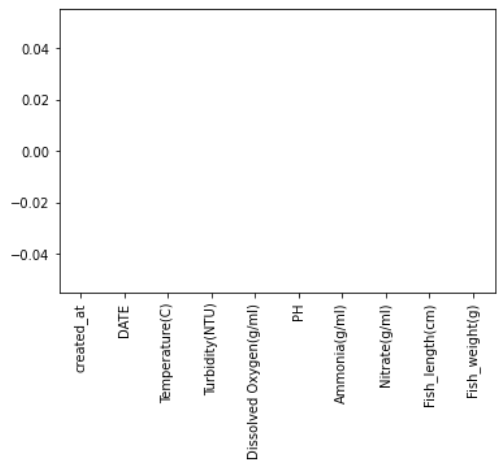
Pond-1 Pond-2

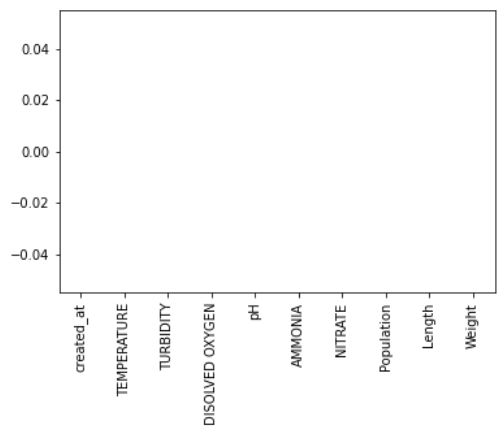
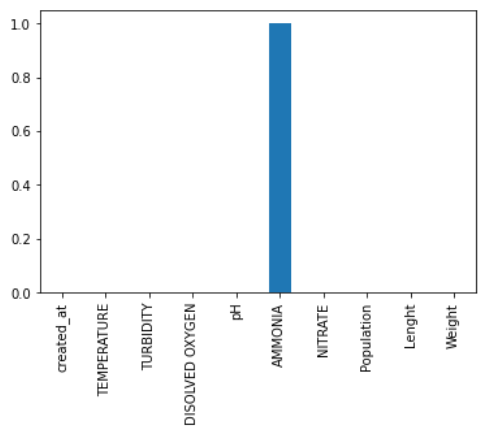
Pond-3 Pond-4

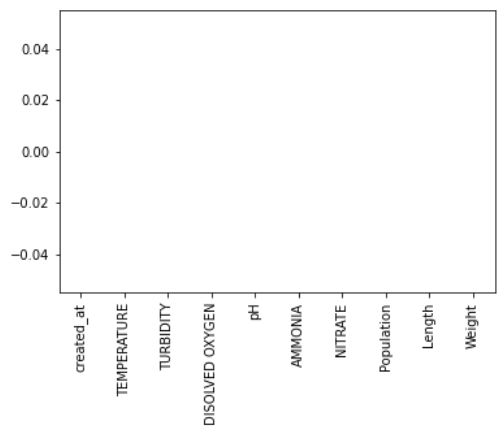
Pond-7 Pond-8

Pond-8 Pond-9

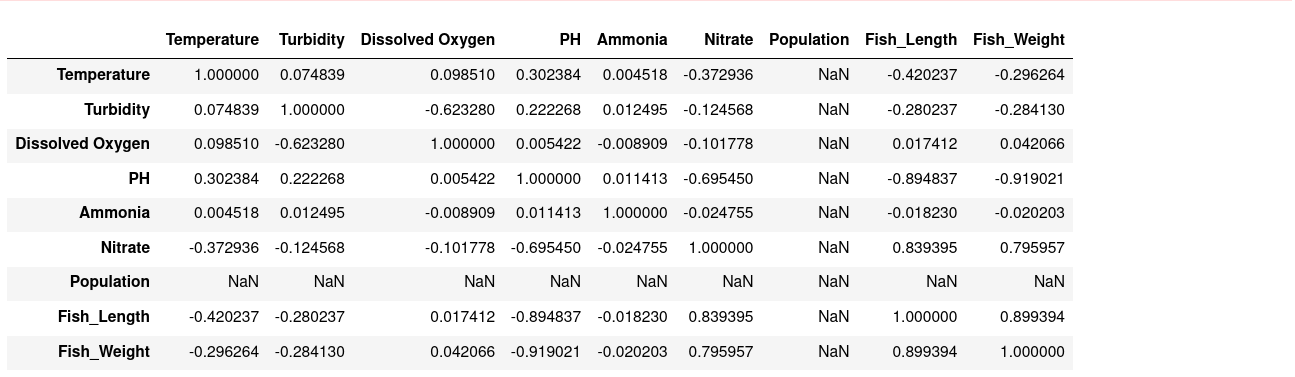
Pond-10 Pond-11

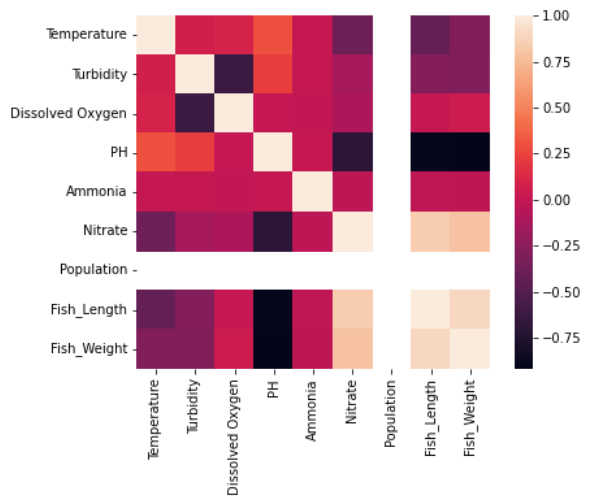


Pond-12

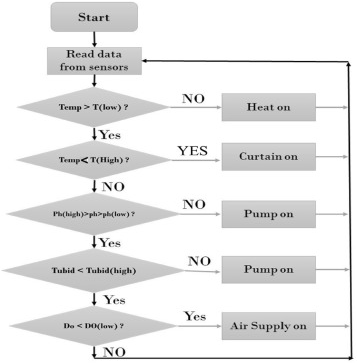
We decided to drop the records with any missing data and unwanted columns.

The correlation heat map shows the interrelationship among various system input parameters.





Upon EDA, we found the following:



pH strongly correlates with fish weight and length increments per week. In addition, nitrate has strong correlations with all the input parameters.

One of the most significant things to monitor is the pond's oxygen level, which strongly correlates with turbidity. Predictive models may provide farmers with the dissolved oxygen concentration value a day in advance, with a high prediction range, so they can adjust feeding or activate aerators as necessary—varying turbidity results in significant variations in the growth of the fish. Because high turbidity reduces the Catfish's feeding abilities, this species' avoidance of the turbid waters could correspond to an increase in its feeding ability. To search for appropriate water turbidity conditions and the dissolved oxygen that can be helpful in fish growth, the correlation between these features with an increment of fish length and fish weight increment will be explored.

**Section 3. Machine Learning Method Choice for Deep Learning Requirement**

We have chosen the Long Short Term Memory (LSTM) deep learning model to satisfy the Deep Learning requirement. The LSTM will be able to capture the characteristics of the water conditions that will display the best growth rate for the fish across the different ponds.

**Section 4. Machine Learning Method Choice for Time Series Requirement**

We have chosen a Regression model to satisfy the time-series prediction requirement. The regression model will be trained on the historical data points in the different ponds over the captured period and allow for an analysis of the optimal water conditions for growth rate.