
Smart Water Management.

Sensors.

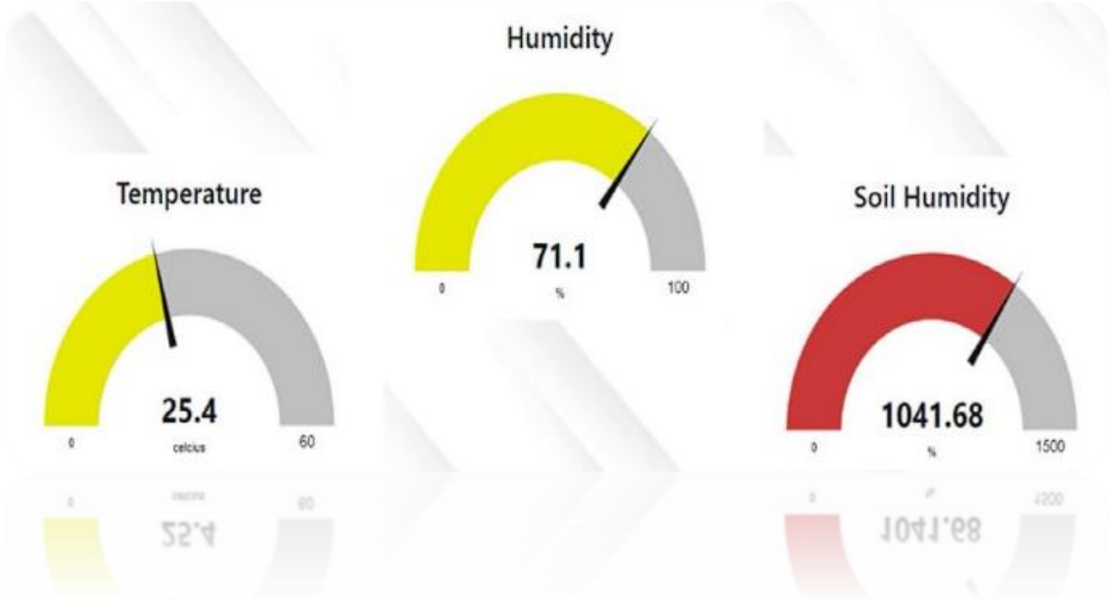
- Soil moister sensors .
- environmental sensor ,
- DIGITEN G1/2 Water Flow Hall Sensor Switch Flow Meter Flowmeter Counter

DataSets.

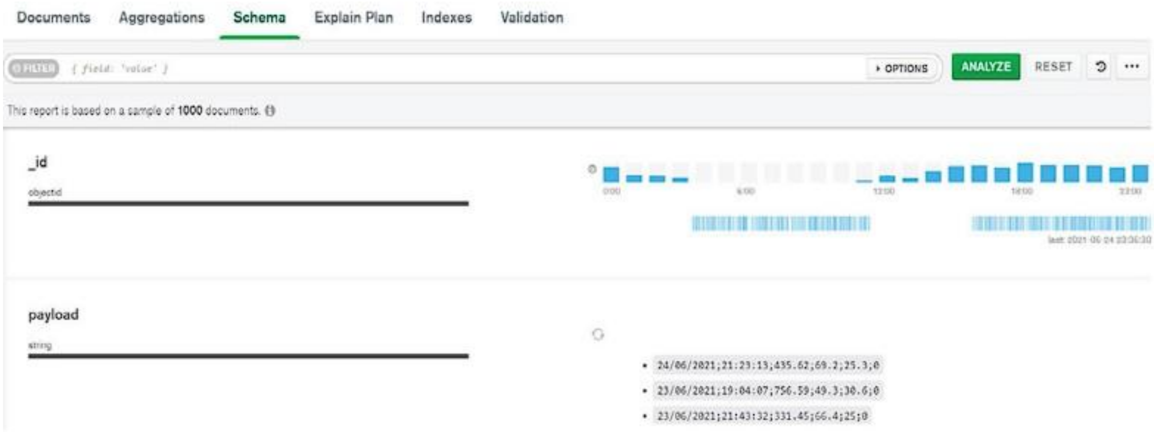
With the help of IoT technologies, made up of a multitude of autonomous devices in the form of sensors

capable of self-organization and working to collect information, we began to implement these devices in various

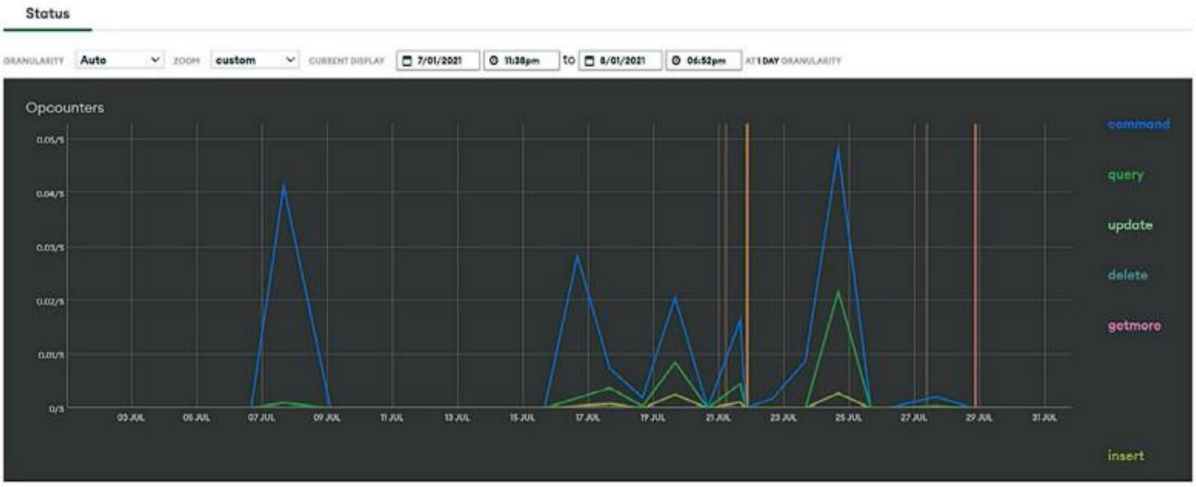
environments containing several domestic plants in the mass collection process for the absolute need of information



Node-red Supervision system



Scheme of data collected in MongoDB Local



Charts presented all used commands in MongoDB Cloud

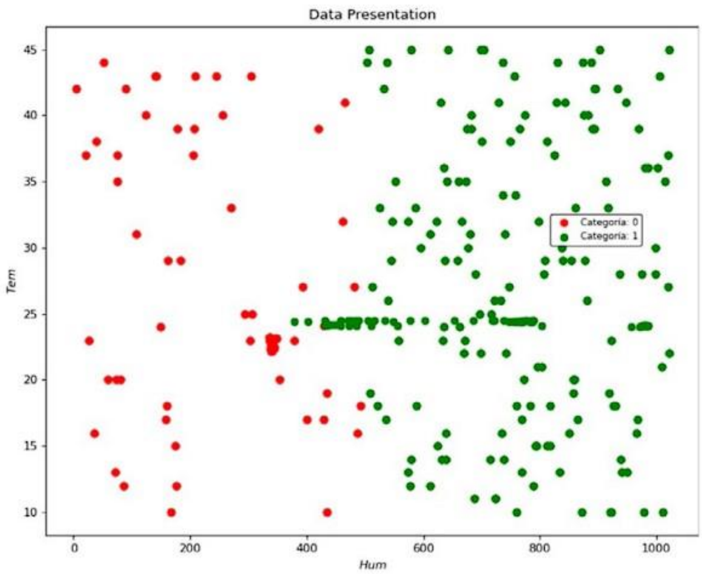
- Soil moisture data: This data is emitted by an analog sensor in a data interval between the value 0 and the value 1023, which is illustrated in the table, we notice that the minimum value is 314.47 and the maximum value is 987.83, so the average value is 384.5.

- Temperature data: These data are becoming more and more important, and they have been collected thanks to a temperature sensor which presents the state of the temperature in Celsius, we see that the average temperature during these months of the collection is 26, 34 ◦C and the minimum value is 18 ◦C without forgetting the maximum value is 39 ◦C, to subsequently overcome its limits proof of expectations.

- Air humidity data: With the same sensor that ensured the collection of temperature values, we managed to collect humidity data, for an analysis passage of these data which is as follows: the average is 66.4%, and the minimum value is 38% and the maximum value is 81.3%, while collects the massive data pass. In the race for computerization, we find the Output data: For this, we have proposed an architecture based on the peer-to-peer principle which resides in categorical data between a value “0” which means that pumping must be stopped, and a value “1” which means that pumping must be activated.To conclude, we have carried out a partial implementation of the final architecture while demonstrating the feasibility to be far from a generalized failure.

Model result.

we see a data manipulation schematic that involves modeling the obtained results to create a representation that can be used for reporting and decision-making. It's important to note that the predictions in this context are largely inductive by nature. Additionally, empirical observations are universally represented through a peer-to-peer identification system that categorizes them into two distinct colors. The color red signifies category "0," indicating a condition where the pumping is judged based on temperature and humidity trends leading to deactivation. On the other hand, the color green represents category "1," which is associated with temperature and humidity patterns that influence active pumping.



Datasets and employing a variety of methods and algorithms for analysis. This approach allows us to extract valuable insights from past events, enhancing our ability to predict future occurrences and make informed decisions for our irrigation system. The likelihood of accurate predictions largely depends on the choice of algorithms, including neural networks, support vector machines (SVM), logistic regression, Naïve Bayes, and K-Nearest Neighbors (K-NN). These algorithms play a crucial role in forecasting and shaping forthcoming trends within our system.

Table 2. Present analysis data.

	Incoming data			Outgoing data
	Soil moisture data	Temperature data	Air humidity data	Pump data
Mean	384.50	26.34 °C	66.4%	–
Min	314.47	18 °C	38%	0
Max	984.83	39 °C	81.3%	1

Table 3. Results data training model.

Models	Parameter	Accuracy	(RMSE)
K-Nearest Neighbors	K = 3	98.3%	0.12
Neural Network	Sequential, Epochs = 50	97,2%	0.16
Naïve Bayes	GaussianNB	97%	0.17
Support Vector Machine	Linear SVC	96,7%	0.17
Logistic Regression	Logistic Regression	96.2%	0.19

Table 3 summarizes the results of various tests conducted to train prediction models, following a thorough exploration of pertinent data. Notably, the K-NN model achieved an impressive accuracy rate of 98.3% in the training set with a low root mean square error (RMSE) of 0.12. In comparison, the Neural Network, Gaussian Naive Bayes, SVM, and logistic regression models yielded consecutive results: (97.2% / 0.16), (97% / 0.17), (96.7% / 0.17), and (96.2% / 0.19) respectively. These findings highlight the potential benefits of accurately determining data (see Table 4).

Before applying these models, data standardization was performed, and the data were divided into test and training sets to enhance model accuracy. A neural network classification approach was employed with specific periods, allowing for a closer examination of data accuracy and loss in each period.

Table 1 presents a summary of the model results, including parameters, accuracy, and RMSE, while Table 2 visualizes the same results in graphical form. For instance, Line 1 in Table 2 represents the KNN model, showcasing a high accuracy rate due to its ability to effectively differentiate between "0" and "1" categories. Line 2 demonstrates a sequential neural network model with a 97.2% accuracy rate, while Line 3 presents the results of the Naive Bayes model, showing an improved performance over the SVM model with a 97% accuracy rate. Line 4 illustrates the SVM model's linear classification, with a slightly lower accuracy rate of 96.7%. Lastly, the logistic regression model achieved an accuracy rate of 96.2% with an RMSE close to 0.2.