An intelligent irrigation system based on internet of things (IoT) to minimize water loss

Samar Amassmir, Said Tkatek, Otman Abdoun, Jaafar Abouchabaka
Computer Sciences Research Laboratory, Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco

ABSTRACT
This paper proposes a comparison of three machine learning algorithms for a better intelligent irrigation system based on internet of things (IoT) for different products. This work's major contribution is to specify the most accurate algorithm among the three machine learning algorithms (k-nearest neighbors (KNN), support vector machine (SVM), artificial neural network (ANN)). This is achieved by collecting irrigation data of a specific products and split it into training data and test data then compare the accuracy of the three algorithms. To evaluate the performance of our algorithm we built a system of IoT devices. The temperature and humidity sensors are installed in the field interact with the Arduino microcontroller. The Arduino is connected to Raspberry Pi3, which holds the machine learning algorithm. It turned out to be ANN algorithm is the most accurate for such system of irrigation. The ANN algorithm is the best choice for an intelligent system to minimize water loss for some products.

Keywords: Arduino, Intelligent, IoT, Irrigation, Raspberry Pi3

1. INTRODUCTION
Morocco's agricultural sector is one of the most vulnerable. It is the primary source of income in some cities. Agriculture is becoming more productive today as a result of advances and new technologies such as sensors, devices, and information technology, which help generate a large amount of data [1]. Network parameter optimization (NPO) algorithm is a new metaheuristic optimization algorithm to simulate the human behavior in their motion when searching for water and food [2]. Perennials include olive, almond, pistachio, and fodder trees. The existing irrigation system is based on uniform water distribution [3] which is inefficient. As a result, agricultural monitoring technologies are being utilised. So, for the protection of farmers, to get the required production, we consider using artificial intelligence (AI), which plays an important role in developing good development and management practices, because agriculture will benefit tremendously from an intelligent approach powered by artificial intelligence. As a result, machine learning [4], a subset of artificial intelligence, plays a critical role in allowing machines to learn without being explicitly programmed. In farm management and production, numerous successful machine algorithms [5]-[7] recognise the input and correspondingly output the connection in agricultural inputs, producing an estimated forecast.

With the advent and growth of machine learning and internet of things (IoT), irrigation techniques are also becoming very modernized. The data retrieved by the sensors and the type of action done by the processor may be watched and modified by the user based on his needs, making the machine process more flexible based on real-time data [8] IoT, AI, and cloud computing [9]. When a precision irrigation system is
An intelligent irrigation system based on internet of things (IoT) to minimize water loss (Samar Amassmir)

used, the traditional system may be enhanced and the cost of farm management can be lowered. As a result, the farmer can decide on irrigation water by estimating the water need [10], [11]. The soil moisture status may be monitored using a soil moisture sensor to identify the present soil water content and irrigation depending on the moisture deficiency [12]. There are various methods for developing automated or smart irrigation systems that use sensors and are supported by IoT to allow for communication [13]-[22].

To perform a job in machine learning, the learning process is categorized into two categories. There are two types of learning: supervised learning and unsupervised learning. The irrigation system makes advantage of both of the learning processes. However, the most often used learning technique is supervised learning, which employs labelled data as an input data for training purposes. It is made up of both input components and the necessary result. The goal of using machine learning is to reduce the waste of water, so that when receiving informations from the sensors, the machine is able to make the right decision to allow irrigation or not. However, the major disadvantage of most of these systems is that they are completely reliant on the machine learning algorithm, which may be correct to some extent but cannot be accurate. Furthermore, irrigation is a land and crop specific process, so complete automation may result in unwanted extra water loss at times and excessive energy consumption from the system. The proposed work aims at comparison of three machine learning algorithms for a better intelligent irrigation system based on IoT for different products which: i) Collecting irrigation data from web for training the machine learning algorithms while waiting for the collection of data of our next experiments which will be realized in the field of Gharb region (Morocco); ii) Specify which of the three algorithms (K-nearest neighbors (KNN), support vector machine (SVM), artificial neural network (ANN)) is more accurate; and iii) Uses IoT in field to evaluate the algorithm.

2. METHOD

In this paper, our objective is to define an accurate algorithm among the three popular machine learning algorithms KNN, SVM, and ANN for an accurate intelligent irrigation system for a specific product. The steps we followed are: i) The very first step is to obtain the data that we need from available data sources and scrubbing it, “clean” and “filter” the data; ii) Split the data into training and test data then we made a model for each of the three algorithms (KNN, SVM, ANN), then calculate the accuracy for each one; and iii) we built a system of IoT devices, the temperature and humidity sensors are installed in the field interact with the Arduino microcontroller, the Arduino is connected to Raspberry Pi3 that holds the machine learning algorithm we found most accurate.

2.1. Collecting and cleaning data

Data collection is a crucial step; without data, we cannot proceed; thus, we must first decide where to collect data, whether it is from existing files or from the internet, using a web scraping tool that easily extracts massive quantities of data from the web. In our case, we will be collecting data from the web and the original owner of the database: Computer Application Department National Institute of Technology, Raipur. After collecting data, the most important step is cleaning data.

In the Table 1 sets of databases, we collected. As it turns out, we focused only on three types Wheat, Maize, and Potato. For each type, the data was checked and cleaned, so we obtained the number of data shown in the Table 1. In Figure 1 after collecting and cleaning the data. It was entered into the work environment where we used Google Colab. Colab is a free Jupyter notebook environment that runs entirely in the cloud. Colab supports many popular machine learning libraries. The Table 2 shows the details of the data collected.

<table>
<thead>
<tr>
<th>Product</th>
<th>The number of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>100</td>
</tr>
<tr>
<td>Maize</td>
<td>125</td>
</tr>
<tr>
<td>Potato</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit of measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CropType</td>
<td>Fixed numerical values</td>
<td></td>
</tr>
<tr>
<td>CropDays</td>
<td>Days after sowing</td>
<td></td>
</tr>
<tr>
<td>SoilMoisture</td>
<td>DHT11</td>
<td>Soil moisture, read by the soil moisture sensor</td>
</tr>
<tr>
<td>temperature</td>
<td>°C</td>
<td>Read by the temperature sensor</td>
</tr>
<tr>
<td>Humidity</td>
<td>%</td>
<td>Read by the humidity sensor</td>
</tr>
<tr>
<td>Pump (class)</td>
<td>1 for irrigated 0 for non-irrigated</td>
<td></td>
</tr>
</tbody>
</table>
2.2. Building model

Generally, we will split our data, for the three algorithms, into two parts: training and testing sets. We train our model with training data, once it is ready to use, test it on test data to calculate the accuracy of each algorithm. We will take a known set of input data and known responses to the data (output) and train a model to generate reasonable predictions for the response to new data. The algorithms used in our study are KNN, SVM, and ANN.

2.2.1. K-nearest neighbors (KNN) algorithm

The KNN method calculates the distance between each training sample and each test sample in the dataset and then provides the k-examples that are the closest. It has a linear time complexity and is guaranteed to discover the precise k-nearest neighbors. However, the linear search method's computing complexity is proportional to the size of the training dataset for each test sample, and it is O(nd), where n is the number of dimensions in the training dataset and d is the number of dimensions in the training dataset [23], [24]. Our first working method is based on theses steps as: i) load the data; ii) separate data into 80% training data and 20% test data; iii) model creation; iv) train the model; v) verification of the accuracy of the model; vi) make the prediction of a recording; and vii) Make the prediction with the test data.

2.2.2. Support vector machine (SVM) algorithm

SVMs are a family of machine learning (ML) algorithms that have good intrinsic generalization ability and are relatively robust to noise in the training data. Given a training data set, SVMs make at each data point an error of at most ‘ε’, where ‘ε’ is a predefined small positive number. SVMs were initially intended for linear models but non-linear ones can be made by transforming input X through a non-linear mapping Φ(X) and then applying conventional linear SVMs over features Φ(X). According to the data, the performance of SVM is the same order, or even better, than a neural network (NN) [25]. Our second working method is based on theses steps as: i) load the data; ii) separate data into 80% training data and 20% test data; iii) model creation; iv) train the model and predict the test data; v) verification of the accuracy of the model; vi) make the prediction of a recording; and vii) make the prediction with the test data.

2.2.3. Artificial neural network (ANN) algorithm

The ANN began with a simpler application in many disciplines and amazing success in pattern recognition (PR), especially in the manufacturing industry. Deep learning is one of the main machine learning technologies. A quite promising technique extends classical ANN by adding more complexity (“depth”) to the model. Deep learning requires a large volume of data and the calculating power of machines to work well [26]. Despite the fact that great progress has been made and reviewed in resolving ANN application to PR concerns. ANN is among the most important algorithms that perform learning tasks. ANN are inspired by biological neural networks. Similarly, the ANN can learn from complicated data and provide insights through predictions or classifications [27]. To use this AI algorithm in data analysis, we will have to follow this working method.

Our third working method is based on theses steps as: i) load data; ii) separate data into 80% training data and 20% test data; iii) model creation (An entry layer with 10 nodes; A hidden layer having 20 nodes, An output layer having a node to predict); iv) compile the model; v) adapt and train the model by optimizing weights and biases; vi) evaluate the model; and vii) make probability predictions with the model.

Figure 1. The data printed in Colab
3. RESULTS AND DISCUSSION

The primary objective of model comparison and selection is definitely better performance of the machine learning software/solution. The objective is to narrow down on the best algorithms that suit both the data and the business requirements. After the implementation of machine learning algorithms, we obtain the results of accuracy. In this work we use three algorithms KNN, SVM, and ANN. However, ANN algorithm gives very high result accuracy and Table 3 shown the comparison of accuracy result of using 3 machine learning algorithms.

Table 3. Accuracy result and comparison of machine learning algorithms

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-Nearest Neighbors KNN</td>
<td>91%</td>
</tr>
<tr>
<td>Support Vector Machine SVM</td>
<td>87%</td>
</tr>
<tr>
<td>Artificial Neural Network ANN</td>
<td>96.81%</td>
</tr>
</tbody>
</table>

In the Table 3 it is clear that ANN machine learning algorithm gives accuracy result better than other applied machine learning algorithms. So according to this result the sensed data is then transmitted serially via serial communication to the Raspberry Pi3 control unit, which is running the machine learning algorithm ANN. The Pi3 control units employ a machine learning algorithm to train a data set pertaining to moisture and temperature for various soil conditions. The projected performance is sent to Arduino as a control signal, which activates the pump to water the field as required. The system is tested in two fields, the potato field and the wheat field, and it produced accurate results as the water pump was only opened when the field needed irrigation. The Figure 2 represents the flow diagram of the overall system, the ANN algorithm running in the Raspberry pi will check for various conditions from the data received from the nodes.

![Figure 2. Intelligent irrigation system based on IoT](image)

It took a lot of new technology to incorporate this application with an internet connection, such as sensors and Arduino, like viewpoints, on the one hand, it would be very important to implement this application that allows this device the ability to be controlled remotely. To improve our results we intend to integrate hybrid methods such as genetic algorithms [28]-[30]. We can use drones in agriculture with their multifunction in areas with high risk of infection can ensure easy and safe distant interventions [31].

In the future, we can derive new ideas to develop our work such as smart urban farming service concept based on an open IoT platform [32], [33]. Using open source IoT platform such as (NodeMcu, Node Red, and message queue telemetry transport (MQTT)) [34]. As an automatic tool for monitoring the availability of water that can help the farmer to monitor the farm-field [35]. There is also a technology a low-cost wireless sensor network (WSN) for detecting soil, environmental, and crop characteristics, which, when appropriately analyzed, are utilised in conjunction with weather forecasts to decide future agricultural operations based on agronomic models incorporated in the software platform [36]. To minimize water loss we can also build an application for automatic schedule-based distribution and monitoring of irrigation by...
applying the waterfall model process [37]. IoT with a fibrous capillary irrigation system based on the climatic demand estimated by the weather condition can also guarantee an accurate irrigation [38]. A modern intelligent gadget is created and used to manage the functioning of irrigation pumps across vast distances utilizing a strong way of communication to send information at the lowest cost and with the longest battery life.

4. CONCLUSION

Factors such as global climate change, food security development and consideration have prompted the company to pursue several innovative approaches to safeguard and improve crop yields. Agriculture may be a big business and thus the base of the economy. As a result, agriculture is changing over time as part of the company’s technical growth. There is tremendous scope for the use of AI in agriculture and irrigation is among these uses, so AI can aid in decision-making on irrigation, applications of fertilizers and pesticides, detection of intrusions, detection of pests, predicting yield, prediction of plant diseases, and detection of fires.

We provided a brief explanation of the importance of artificial intelligence in the field of agriculture in this article, then we chose irrigation as part of this sector. We worked to reduce the water consumption used for irrigation. We are responsible for ensuring automatic watering plants that enabled us to monitor the parameter of soil water quality and content in real time. In the future work, we will be interested in analyzing the data of the real experiences on the agricultural zone which knows a great irrigation activity and more specifically in Gharb region of Morocco.

REFERENCES


An intelligent irrigation system based on internet of things (IoT) to minimize water loss (Samar Amassmir)

**BIographies of Authors**

Samar Amassmir Her master's degree in computer science has been received from Ibn Zohr University. She is currently an Ibn Tofail University, Kenitra, Morocco, PhD student at the Computer Science Research Laboratory (LaRI). She can be contacted at email: samar.amassmir@gmail.com.
Said Tkatek is Professor of Computing Science Research at the University of Kenitra, Faculty of Science, Ibn Tofail. Their study focuses on the optimisation of NP-Hard issues with the metaheuristic approaches in artificial intelligence and big data for decision-making in diverse fields. Includes big data analytics, artificial intelligence, information system, and human resources optimization. He can be contacted at email: said.tkatek@uit.ac.ma.

Otman Abdoun is Professor of Computer Science at Larache Abdelmalek University, Multidisciplinary Faculty. His work focuses on addressing the NP-Complete problem with methods of meta-heuristics to lessen the complexity of the problems investigated. He can be contacted at email: abdoun.otman@gmail.com.

Jaafar Abouchabaka is a Professor of Computer Science at Ibn Tofail University in Kenitra, Faculty of Science. Their research is presently focused on the optimization and implementation of NP-hard Jaafar technology and big data analysis. Abouchabaka is the director of the research laboratory for computer sciences (La.R.I). He can be contacted at email: jaafar.abouchabaka@uit.ac.ma.