International Journal of Computing and Digital Systems
2025, VOL. 17, NO. 1, 1–14

http://dx.doi.org/10.12785/ijcds/1571062172

IoT and Deep Learning-based Approach for An Efficient Land Suitability Prediction in Smart Farms

Toufik Bendahmane[1], Abdelhak Merizig[1], Khaled Rezeg[1], Okba Kazar[1,2] and Saad Harous[3]

[1]*LINFI Laboratory, Computer Sciences Departement, Mohamed Khider University of Biskra, Biskra, ALGERIA*
[2]*College of Computing and Informatics, Department of Computer Science, University of Sharjah, College of Arts, Sciences Information Technology, University of Kalba, Sharjah, United Arab Emirates*
[3]*College of Computing and Informatics, Department of Computer Science University of Sharjah, UAE*

Received 12 August 2024, Revised 14 February 2025, Accepted 15 February 2025

Abstract: The earliest known way for humans to make a living is through farming. Smart farming is a new, improved vision of agriculture that incorporates new technologies. In recent times, farmers have increasingly depended on technology to efficiently carry out their daily responsibilities and enhance the quality of their crops. In agriculture, land suitability is an important aspect, which describes how well the area is conducive for plant growth. Experts in land suitability can determine it or use mathematical tools to make accurate predictions. Artificial techniques have been proven to be efficient prediction tools for this purpose. Empowered by the Internet of Things and Big Data, Artificial Intelligence (AI) is capable of handling these kinds of tasks and easing the burden on farmers and experts. The devices used to improve farming generate data in several formats, which might lead to ambiguous data. This paper proposes an ontology-based solution to deal with the heterogeneity problem. Moreover, this paper uses a deep learning-based solution that uses streamed weather data generated from sensors. Our system uses the long-short-term memory model to predict land suitability. The model exhibited encouraging outcomes that could influence the field of agriculture.
Keywords: Agriculture 4.0, Land Suitability, Internet of Things, Machine learning, Deep Learning, Long-Short Term Memory, Ontology.

1. INTRODUCTION

Agriculture has served as the primary sustenance for both people and animals throughout the past 12,000 years. Due to its significance, authorities and private entities are always trying to improve the return from farming the land, which leads to increasing the incoming resources and meeting the population requirements. The latest technological development, if used properly, can substantially enhance the farming field. Applying new technological developments in traditional agriculture resulted in a new term called smart agriculture [1], [2]. Smart agriculture integrates Internet of Things (IoT) technologies with conventional agricultural practices [3]. This combination will greatly facilitate field monitoring, particularly in the production chain process. In the conventional agriculture paradigm, farmers are responsible for nearly all process aspects, including aggregation and crop collection. In addition, experts face limitations when they want to study different phenomena and/or diseases that affect crops. The emerging concept of agriculture revolves around implementing sensors on the field to collect data about several characteristics of the environment, such as humidity, temperature, wind speed, and so on [4]. Experts and data analysts can use the collected data for decision-making process purposes [5], [6]. In addition, it could be to study different phenomena related to the seeds or the fertility of the land. Agricultural land suitability prediction is a significant concern that producers prioritize greatly [7]. To optimize crop quality, farmers must know precisely the optimal timing and location for seed planting, considering several elements and criteria, including water availability, weather conditions, soil composition, and fertilizer. Experts need a huge amount of historical data to make the prediction as accurate as possible. The new developments in technology can be very helpful in this domain. Using sensors and actuators [8], [9], a huge amount of historical data can be collected, and then, using AI techniques to study this data, accurate information about the land and the crops is deduced. In this work, we aim to design and implement a system that uses AI techniques and IoT technology to collect and analyze data that helps farmers make appropriate predictions about land suitability. Through developing a more sophisticated method for evaluating land suitability in the context of smart farming, this study aims to address these difficulties. This paper specifically intends

E-mail address: toufik.bendahmane@univ-biskra.dz, a.merizig@univ-biskra.dz, k.rezeg@univ-biskra.dz, o.kazar@ukb.ac.ae , harous@sharjah.ac.ae

to construct a dynamic, responsive system that can assess land suitability with improved accuracy and efficiency by integrating IoT and deep learning-based approaches for efficient land suitability prediction in smart farms and advanced data analytics. This approach can transform land evaluation by utilizing AI's predictive skills and real-time data processing, making it more responsive to the demands of modern agriculture. Usually, new farmers do not have a sufficient understanding of the soil properties required for agricultural production [10]. They lack awareness of the necessity of assessing agricultural land before cultivation. Consequently, it is essential to assess the area's suitability before cultivating crops to optimize production [11], [12]. Traditionally, farmers rely on manual data collection and soil testing labs to know the soil's properties. These methods may not achieve the goal because the data may not be accurate [13], [14]. The elements of land can determine whether it is suitable for agriculture and plantations. There is still a crucial gap in how land suitability is determined and incorporated into larger agricultural systems, despite tremendous advancements in smart farming technologies. Conventional land appraisal techniques are frequently labor-intensive, resource-intensive, and prone to human error. Furthermore, these techniques might not sufficiently consider real-time data or adjust to shifting environmental circumstances, reducing their usefulness in contemporary agricultural practices. Nevertheless, terrestrial components are excessively utilized and exploited. Various regions encounter diverse issues, including soil erosion, water logging, groundwater depletion, excessive runoff, and productivity losses. The deterioration of land poses a significant danger to food and energy security, water availability and quality, biodiversity, and human life. Weather conditions can significantly affect agricultural lands. This phenomenon results from a modification in atmospheric conditions that impacts the ability of agricultural operations to flourish and enhance crop production. Several factors that influence the conditions for crops and cattle include temperature, nutrient concentrations, soil moisture, and water accessibility. Research on crop yields has shown that very high temperatures, resulting in elevated vapor pressure deficits, can reduce the yields of rain-fed crops across different crops and areas. Interconnected devices or physical objects via the internet, known as IoT technology, represent the new evolution in artificial intelligence [15]. This approach uses different kinds of communication. Data collected this way may be of different types. This problem is due to the nature of data gathered by devices that are generally heterogeneous [16]. In fact, these data are collected in real-time, which will generate a sort of Big Data [17]. This kind of data is usually handled using a data streaming process [18], [19]. Smart agriculture brings a new vision that injects AI technology into traditional agriculture. This combination increases production and economic profits [20]. The emergence of sensors, which are based on different systems, leads to problems about how this data is exchanged among these systems. Usually, to avoid this issue, which is related to understanding heterogeneous data or exchange signals between sensors, known as the interoperability issue. In fact, land suitability represents an issue for farmers in two manners: the first consists of the quality of food or plants, and the second is related to the nature of the field area. By presenting an AI-driven method for assessing land suitability that outperforms current approaches in terms of precision and adaptability, this research makes a significant advance to the field of smart farming. Although AI has been used in the past to forecast land suitability, these studies frequently used static models that were ineffective at integrating multi-dimensional data inputs or adequately accounting for changes in the environment in real time. This study's method stands out because it uses advanced data analytics and generative AI to produce a dynamic, real-time assessment tool. In contrast to conventional models, which could be constrained by their reliance on past data and set parameters, this study presents a system that can update its predictions in real-time using the most recent information on soil properties, climatic patterns, and environmental variables. This development sets a new benchmark for using AI in agriculture by improving the precision of land suitability predictions and giving farmers access to a more responsive and flexible tool. The primary contributions of this study are:

- Develop a distributed architecture to ensure interoperability among devices.
- Propose an ontology domain to ensure data understanding between different generated data;
- Propose a deep learning model for land suitability prediction;
- Develop a distributed architecture to ensure interoperability among devices;
- Propose a framework to visualize data in real-time;
- Perform several comparisons of our strategy with alternative machine learning classifiers.

The paper is organized as follows: Section 2 describes related works about smart agriculture, including land suitability. In Section 3, we present our proposed framework. Section 4 explains the prediction model for land suitability. The obtained results are discussed in Section 5. Finally, Section 6 concludes this paper and gives perspectives for future works.

2. Literature Review

Since the emergence of sensors and actuators and the idea of the IoT, many developers and researchers have incorporated these technologies in many domains, such as industry or agriculture. In the literature, many works have used the Internet of Things as a solution in different fields. In this section, we present some works that have used the IoT to solve problems in the agriculture field. The agriculture field has gained many advantages from this technology, such as collecting data, using irrigation systems, monitoring the field, and detecting diseases in the crops. In addition, land suitability is a major challenge due to its importance to the countries, especially those with large areas. The work of Hsu et al. in [21] has designed a platform that uses cloud fog computing for agriculture purposes. The main idea is to present a new creative IoT system layer composed of a set of layers instead of the traditional IoT layers. In fact, the authors integrated the fog computing layer into this system. According to the authors, this architecture ensures data gathering and analysis from different types of devices. However, they did not mention how to deal with the heterogeneity of collecting data from these devices. Pathak et al. [22] proposed a field monitoring system based on IoT technology to deal with the problem of crop irrigation. The main objective of the developed system is to optimize the quantity of water used for the corps' irrigation. The main idea is to use the collected climatic parameters to predict the water level in the field and thus decide the amount of water that should be poured to irrigate the crops. Also, Dos Santos et al. in [23] have proposed a model to measure crop productivity and anticipate problems. The proposed model, named AgriPrediction, is based on both the ARIMA model and LoRa IoT technology. The authors combined a set of technologies, such as wireless network range systems, with a prediction method to notify the farmers of some possible recommendations. The system uses different climate factors (soil humidity and temperature) to decide what actions should be taken. However, the authors did not consider all climate factors that could affect crop productivity, which is a shortcoming of this work. Another work focuses on crop recommendation systems in smart agriculture, such as in Shams et al. in [24] have proposed a recommendation system for suitable crops based on various factors such as soil quality and climate conditions. The authors have proposed an approach that stands on explainable AI. With the proposed model, the system can help clear explanations for farmers align AI recommendations with their knowledge and local conditions, allowing for more confident decisions. Where Paudel et al. in [25] the authors have focused on crop yield forecasting, their paper aims to deal with the evaluation of the performance and interpretability of neural network models for crop yield forecasting. They compared LSTM, 1DCNN, and a Gradient-Boosted Decision Trees (GBDT) models to evaluate the system. Chen et al. [26] proposed a platform called agriTalk, which is an IoT-based platform for the precision farming of soil cultivation. The authors ensure connections between the sensors and actuators to preserve farming precision. The objective of their solution is to increase the number of crop cultivation resources through turmeric cultivation. They used network time protocol in their platform. However, interoperability between devices is not mentioned in this work. In [27] Senapaty et al. have applied IoT solution-based machine learning to increase crop productivity. The solution consists of analyzing soil nutrients to enhance precision in agriculture. In addition, the proposal includes a crop recommendation model. The machine learning model is represented as a combination of a support vector machine with a directed acyclic graph and fruit fly optimization technique. Where Shevchenko et al. in [28] have proposed a solution to deal with the impact of climate changes on land suitability. The solution aimed to deal with risks that impact food security. The machine model deals with the problem under different carbon emission scenarios. However, interoperability between devices is not mentioned in these works. Deep learning and machine learning methods have gained significant popularity in the domain of land suitability and the advancement of smart agriculture. Conventional techniques for land cover classification, such as logistic regression, distance measures, and clustering, mainly depend on manually designed characteristics, which restrict their flexibility and adaptability. Deep learning has proven its capacity to automatically collect non-linear and hierarchical information, making it a potent method for diverse fields such as remote sensing and urban planning [29]. Recent advancements in deep learning have resulted in substantial enhancements in tasks such as land cover categorization, data fusion, and reconstructing missing data. This application is valuable for identifying specific locations, analyzing narratives related to landscapes, and providing answers to geographic inquiries. Previous studies have shown that machine learning and IoT benefit from their advantages, such as dealing with large and complex datasets. Yet the heterogeneity in such works does not account for specific locations and dynamics, which can improve the model performance.

As described in this section, the heterogeneity of the data was not considered. This strategy uses modern techniques such as DL parallel processing and real-time analytics to improve the proposed system. In contrast to conventional data processing methods, typical data processing methods can impede processing and fail to capture the range and nature of data acquired from multiple devices. As new technologies, such as ontologies, our solution benefited from them. The system can rapidly grow with increasing data quantities, enabling higher flexibility and efficiency than older systems that struggle with enormous datasets. Optimizing resources and expediting procedures can enable the system to reduce operational expenses by requiring fewer hardware and manpower resources to manage substantial data volumes.

3. Proposed methodology

The agriculture field is evolving with new technologies supported by artificial intelligence. There are many problems related to traditional agriculture that limit farmers' profits. One of these significant problems is land suitability, one of the most known issues in traditional agriculture. The study of the suitability of a given land helps the farmer to decide which kind of crops should be implanted and what should be added to this land to support other crops. Through data analysis and studying historical data related to weather conditions, including soil conditions, this issue should be covered. Similarly to other fields, the agriculture field may benefit from different ideas using new technological developments, such as sensors that collect data about the environment. The main use of new developments in technologies is to reduce human effort in different fields, and this is especially true in the agriculture domain. Farmers nowadays use technology to complete their daily tasks in the field to improve crop quality. To tackle these kinds of problems, solutions based on multi-layer architecture have been proposed. These solutions analyze collected data, visualize these data in an easy-to-understand way for farmers, and predict which crops are more suitable to be implanted in this land. These solutions combine IoT and data analytics to study a new phenomenon. In this section, we propose an architecture that consists of a set of layers. Every layer has a distinct responsibility. The first layer installed in the field collects data from the physical area using sensors and actuators. The second layer consists of data analysis, which covers the preprocessing operations. The last layer is a Cloud IoT approach; this layer represents the model construction process and stores the data on cloud servers. Besides, our proposal ensures data analytics in real-time over historical one through data visualization displayed as data insights for the users.

A. Architecture description

In this subsection, we present the constituents employed in our suggested approach and their respective functions. Our architecture is designed to address the issue of land suitability problems through an interoperability protocol. It is a multi-layer architecture based on IoT to predict land suitability and visualize the data in real-time. In addition, since we are using different farm settings, including the number of sensors and actuators, it might produce various kinds of data. To deal with this issue, our solution is to use an ontology to ensure the interoperability protocol exchanged between inter-systems. Figure 1 shows the proposed architecture composed of the following set of components:

1) **End User:** This component is the interface to our system. The user can access all operations available in the system by clicking on the desired operation. The system supports all devices, such as pads, phones, etc. This component aims to interact with the users with the cloud services, which in our case are data analysis, data visualization, and more importantly, land suitability prediction. To perform a suitability detection for a given land, the user should introduce some parameters related to a blessed land. These parameters are weather conditions and data of the land, such as PH and nature, to check if the given area is suitable for a specific crop.
2) **Sensing layer:** called the physical field layer, its primary role is gathering data from the field and sending it to the sink. This layer is made up of temperature, humidity, soil moisture, UV, and pH sensors. All sensors are controlled by the sink, which is the Raspberry Pi model in our case. The primary role of the sink is to transmit the collected data to the IoT Edge layer. The gathered data can be scheduled weekly, daily, or hourly according to the necessity of the data and the studied area. With this configuration, the farm will be monitored in each lap of time, generating vast amounts of data that will be treated as big data. The resulting data will be passed to the next level for analysis purposes.
3) **IoT Edge layer:** This component is an IoT edge layer. We can call this layer a middleware in our architecture, which connects the processing layer (cloud IoT layer) and the physical layer (sensing layer). This layer also ensures the interoperability between different data collection types from the field. In this layer, we use an ontology domain to unify the data collected from the field. An MQTT broker facilitates the interaction between the physical layer and this layer through the sink. The MQTT broker ensures the exchanged messages between devices are published and subscribed to [30]. It contains the following modules:
   - **Interoperability protocol:** Before we start the data analysis operation, we need to unify the different data types. After receiving the collected data from each farm site, data streaming is transmitted to the MQTT broker. Then we perform the interoperability; we use an ontology domain attributed to this layer, which will be presented in the next section.
   - **Collected data analysis (preprocessing):** After collecting data from different areas, this module occupies the pre-treatment process, which covers data cleaning. This operation is a data streaming-based method that will collect data in real-time. After this operation, the extracted meaningful patterns from the collected data will be transformed into the cloud IoT layer for the model creation.
4) **Cloud IoT layer:** This component is a cloud-based layer. It extends the local server's abilities by using the cloud server's resources. Furthermore, this layer gains from cloud advantages such as resource pooling and service on demand [31], [32]. This layer also offers a platform with multi-functionalities accessible via the Internet.
   - **Cloud storage:** The different sensors collected huge amounts of data (Big Data) [33], which requires a lot of storage space. The cloud provides the user with unlimited data storage space compared to the local server.
   - **Data visualization:** It allows the farmers to visualize data in real-time as it is collected from the field. The data visualization makes data more understandable. At the end, the generated data will be transmitted as XML and JSON formats for later use.
   - **Model construction or machine learning module:** The cloud offers high computing re-

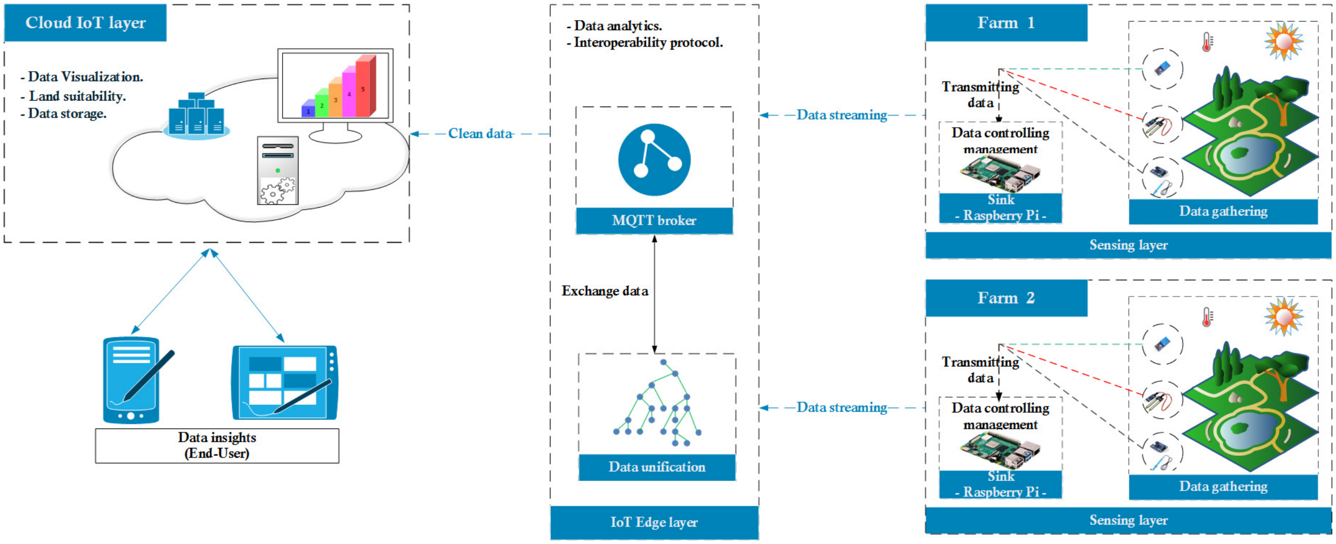


Figure 1. Land suitability monitoring and prediction system.

sources, which can run complex algorithms to predict land suitability. In our work, we use a deep learning method, which is a neural network-based method. In our proposal, we use a machine learning algorithm that handles the land suitability prediction.

B. Dataset description

To construct our model to predict the suitability of the land based on remote sensing data, we utilized meteorological data collected from different regions over twenty years. This dataset includes various features such as soil moisture, precipitation, temperature, relative humidity, and pressure. We focused on using deep learning techniques to predict land suitability across a heterogeneous set of land covers, including natural vegetation, croplands, and human infrastructure.

C. Land Suitability prediction process

To explain the role of each element in the architecture depicted in the previous figure (see figure 1), we use the following flowchart (see figure 2). As described in figure 2, our system consists of settings, especially at the physical layer (sensing layer) installed in each farm. To ensure that a given area can produce healthy plants, our proposal comprises two different subsystems. Our system provides the land suitability prediction in a specific area. We present the previous flowchart's various operations to deal with the interoperability problem between other data generated from sub-systems (different farms). The issue related to interoperability due to the nature of generated data could be presented in other formats (JSON and CSV in our case or maybe include XML format).

1) **Data collection:** This step is important to construct the data, especially data streams. The system receives numerical data from our farm, where the output (raw data) is stored as unstructured data (JSON file). In the end, the local server saves this data using streaming frameworks such as KAFKA [34], [35].

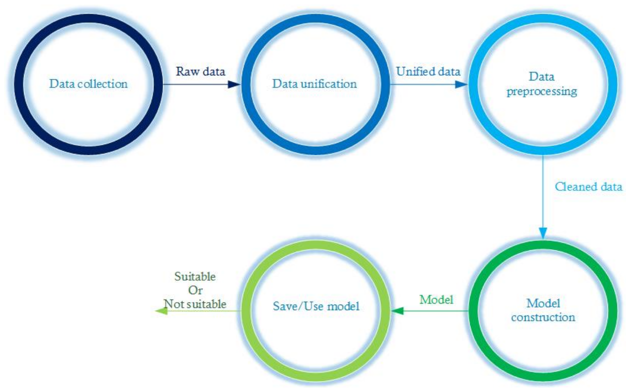


Figure 2. Land suitability process.

2) **Data unification:** Since we have many types of data generated from each farm field, this presents a big challenge related to understanding each type. To ensure an excellent manner to process these data, this step introduces a protocol to unify them. The primary objective of this operation is to prepare data for the pre-processing stage. The MQTT broker handles the unification process, which we also call the interoperability process. The main idea is to use the collected data from the field. In IoT, interoperability can be defined as two systems that can communicate and share information or data services via devices [36]. The devices generate different kinds of data with several data formats, including XML, JSON, or even CSV [15]. As we have seen in our proposed architecture, our system may generate data from

devices as in the previous data format. According to [15], semantic interoperability in IoT can be ensured through different ideas, which are given in three ways: information model, data model, or using ontologies. The existence of different kinds of sensors can generate different data types. To deal with the mentioned problem, semantic interoperability is needed, and gathering data from devices needs the use of sensor ontologies [37]. Since ontology is well known in the semantic field through many works in literature, our interest is to use them [38], [39]. We aim to use domain ontology to unify the collected data to start the data analytics task. In complex datasets like land suitability assessments, which often include climate and soil data, an ontology-based solution is an advanced way to overcome data heterogeneity. By providing a consistent framework for interpreting and combining diverse data, ontologies improve land suitability estimates' accuracy and efficiency. Integrating these sources for a cohesive prediction model is difficult due to their diverse formats, structures, and semantics. To create the proposed ontology, we collect technical information from IoT devices from the literature since an ontology is a formal, explicit specification of a shared conception. It establishes entities, relationships, and rules in an area of knowledge, giving systems a common language. Figure 4 shows how we created the domain idea identification of sensor features and information to build our ontology. Then mapping each data field from heterogeneous datasets to the associated concept in the ontology linked data sources to the ontology. By offering a standardized framework for integrating diverse datasets, the ontology-based solution helped solve data heterogeneity. By establishing semantic coherence and improving inference, land suitability forecasts improved. It also automated data integration and enabled rapid querying, making land suitability assessment more scalable and credible.

To extract information from each file using the ontology and to find the right concept from one data format to another, we execute the equation presented in equation 1 [40]. The primary role of this operation is to calculate the similarity degree from concept to concept and then create a unified file at the end. In this work, we use the XML file type as a unified data file. The idea behind this decision is due to the tools and library that make file generation easier than other types.

$$W(term) = tf{-}idf(term) = tf(term) - log_2\left(\frac{N}{df(term)}\right) \quad (1)$$

W represents the term weight, tf is the term frequency, where N is the number of documents in the collection, and df is the document frequency, which is the number of times the word appears in the other documents. After unified data construction,

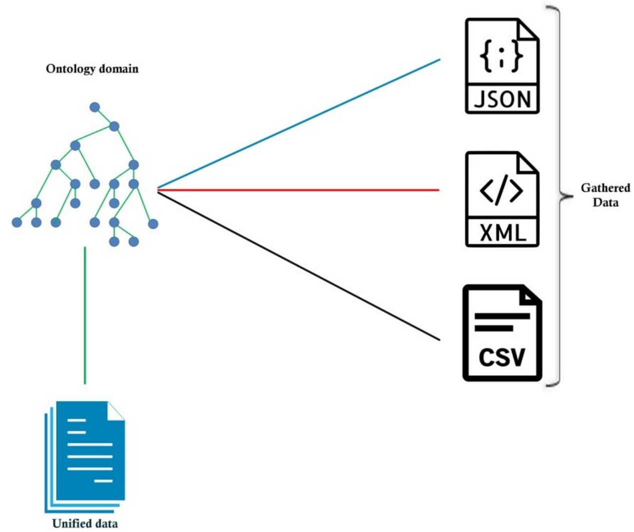


Figure 3. Data unification model.

the MQTT brokers send the result to start the data preprocessing. The following figure represents the ontology used to unify concepts between different data formats. The presented ontology is inspired by [41]. The IoT flow data is fed to the model to create a model. Firstly, as we can see through the proposed solution, we consider the multi-format data generated by sensors. The problem addressed in the paper is handling the heterogeneity of data, especially related to the nature of the generated data since it has different data formats. Our proposal used a new ontology domain that helps to manage streamed data to unify it using Algorithm 1. After the data collection step, the raw weather data from sensors often requires extensive preprocessing (including data normalization and missing data) before being fed into a deep-learning model for time-series prediction purposes. Sometimes, while streaming data, data privacy and security are represented by big challenges in such solutions.

To fetch data from the used ontology given above (see figure 4), we propose the following algorithm to extract the pertinent data. After extracting the classname, we compare them (data format: csv, xml, and JSON files) to unify the data. After that, we send the data for the preprocessing step to create our prediction model. Our algorithm consists of several steps, given as follows:

- Step 1. Given the data format received from the sink, which could be one of the three formats discussed earlier. This data format is presented as a query to our algorithm, and we also give the used ontology.
- Step 2. For each name 'n' from class name 'q' in query Q, we fetch the appropriate concept that is near to 'n'.

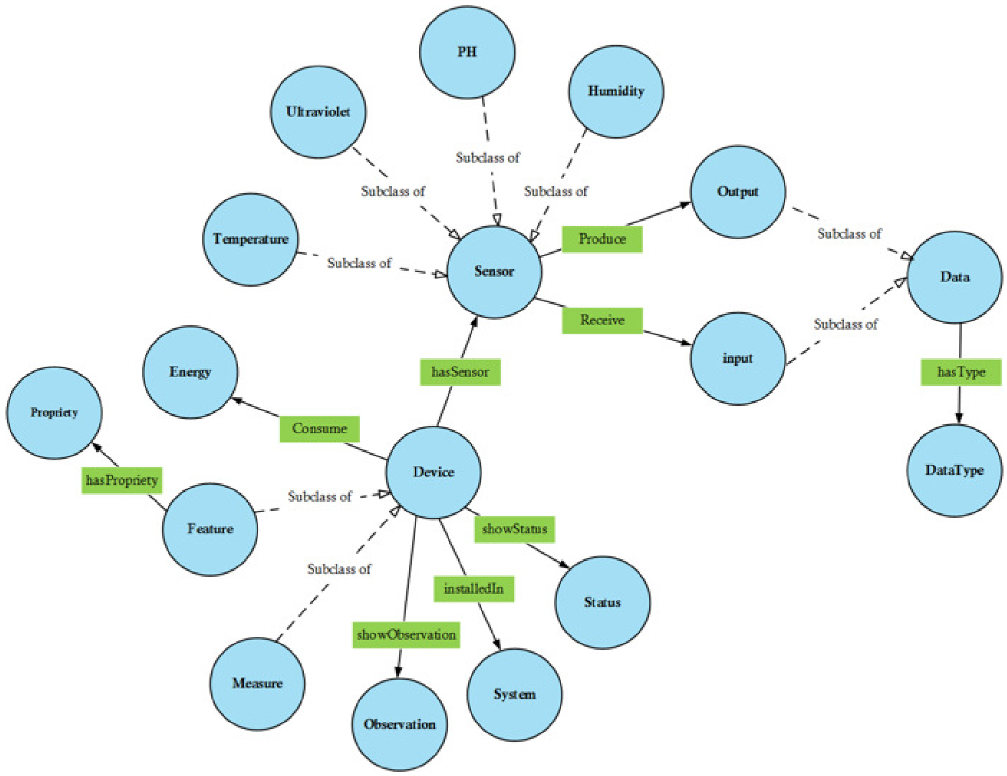


Figure 4. Proposed domain ontology.

Algorithm 1 Fetch in the ontology.

Input: Q: query: data from a source file, O: ontology domain.
Output: C: a list of concepts and their signification.
Begin
Let o *class_node* of data from O;
Let q set of *name_class* of Q;
for each name n in q **do** ▷ extract class name from ontology.
  $i \leftarrow 0$;
  Let S an array of real values
  for each class_node o in O **do** ▷ calculate degree of similarity between concepts. Tf-idf
    $S[i] \leftarrow similarity(o, n)$;
    $i \leftarrow i + 1$;
  end for
  $classes \leftarrow ExtractClassName(S)$; ▷ Extract class name with the highest value.
  $C \leftarrow addtolist(classes, n)$;
end for
Return C;
End.

- Step 3. We calculate the degree of similarity of each name 'n' and compare it with each class node of ontology O and store it in an array S.
- Step 4. Last step, we extract the class node with the highest degree of similarity and save it in a list named 'C' with the current class name 'n'. When we finish this algorithm, we return each concept with its meaning using the ontology, resulting in list 'C'.

3) **Data pre-processing:** Accumulating vast quantities of minuscule numerical values will always result in incomplete datasets. If the raw data contains missing values and is immediately inputted into the model, it will lead to an error. Hence, a preliminary data pre-processing stage is required to ensure data cleanliness. To improve accuracy and efficiency, it is essential to standardize the data to a range of 0 to 1. The result will be a tidy and reliable dataset prepared for usage with the model.
4) **Model construction:** The selection of the optimal model and architecture significantly influences the accuracy of predictions. From the several Machine/Deep learning methods and algorithms available, we have selected an approach that considers time series data. This phase consists of two steps:
   - Designing the model and its architecture, including determining the number of layers, neurons, and activation functions.
   - Performing model training and testing by producing predictions to evaluate the model's performance. The algorithm that was utilized will be examined in the "Prediction algorithm" sec-

tion.

5) **Deploy the model:** Upon creation, the model is deployed to the web platform to be served for users access for prediction purposes. The users can visualize the results in four different decisions: ”Best suitability”, ”Suitable”, ”Moderately suitable”, or ”Unsuitable”.

D. Evaluation metrics

Performance evaluation is an essential stage in validating the efficiency of a machine learning system. Various metrics are employed for this objective, including Recall, Precision, F1- score, Accuracy, and Confusion Matrix [42] [43]. These equations are calculated using the number of correct and incorrect classes, as referred to as True positive/negative and False positive/negative. The metrics used are measured as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \quad (5)$$

4. Prediction model for land suitability

LSTM networks are a specific type of Recurrent Neural Network (RNN). Their purpose was to tackle the problem of prolonged reliance on RNN. LSTMs have an exceptional ability to preserve information over prolonged durations. Due to the potential impact of prior information on model correctness, LSTMs are often used for this purpose [44]. The LSTM architecture consists of a module known as the ”Repeating Module”, which comprises four neural network layers that interact distinctively. The characteristics of the data at hand primarily determined the selection of the LSTM architecture. Given that we are working with time series, which refers to data that is associated with time and has been recorded sequentially, LSTM has demonstrated superior performance in situations when it is capable of retaining and recalling information from past data points. Our circumstances can benefit from this approach, as the prediction of land suitability is not possible using dispersed or unsorted data. The repeating module is equipped with three gate activation functions: σ_1, σ_2, σ_3 and two output activation functions ϕ_1 and ϕ_2 as shown in Figure 5.

As depicted in figure 6 the proposed architecture consists of a set of layers by the definition of LSTM. This study employs a stacked LSTM architecture with numerous LSTM layers, succeeded by dense layers, to translate the sequential outputs into the final prediction. The model receives input characterized by climatic factors, with each time step serving as an input to the model. Dropout layers are implemented subsequent to each LSTM layer to mitigate overfitting. Each LSTM cell comprises three gates: an input gate, a forget gate, and an output gate (measured in equations from 6 to 11). These gates regulate the information flow and allow the network to determine which information to retain or discard over time. The cell state, in conjunction with hidden states, enables the LSTM to learn temporal patterns proficiently. These gates are explained as follows [45].

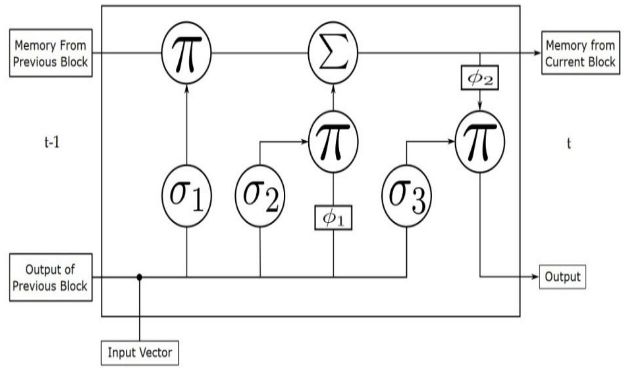


Figure 5. LSTM Repeating Module.

- **Forget Gate:** The forget gate (f_t) determines what information from the previous cell state should be discarded or kept.
- **Input Gate:** The input gate (i_t) updates the cell state with new information from the current input.
- **Output Gate:** The output gate (o_t) controls the output of the cell, which becomes the hidden state for the next time step.

$$f_t = \sigma(W_f[h_{t1}, x_t] + b_f) \quad (6)$$

$$i_t = \sigma(W_i[h_{t1}, x_t] + b_i) \quad (7)$$

$$\tilde{C}_t = tanh(W_C[h_{t1}, x_t] + b_C) \quad (8)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (9)$$

$$o_t = \sigma(W_o[h_{t1}, x_t] + b_o) \quad (10)$$

$$h_t = o_t \times tanh(C_t) \quad (11)$$

Where x_t is the input vector at time step t, h_{t-1} is the hidden state from the previous time step, and W, b are the weight matrices and bias terms, respectively.

The idea behind choosing LSTM model is that this kind of networks are especially appropriate for our goal since they can effectively capture long-term dependencies in time-series data. In contrast to conventional RNNs, LSTMs proficiently address the vanishing gradient issue, which is essential for recognizing patterns throughout prolonged data sequences, such as elements influencing weather conditions. As we mentioned

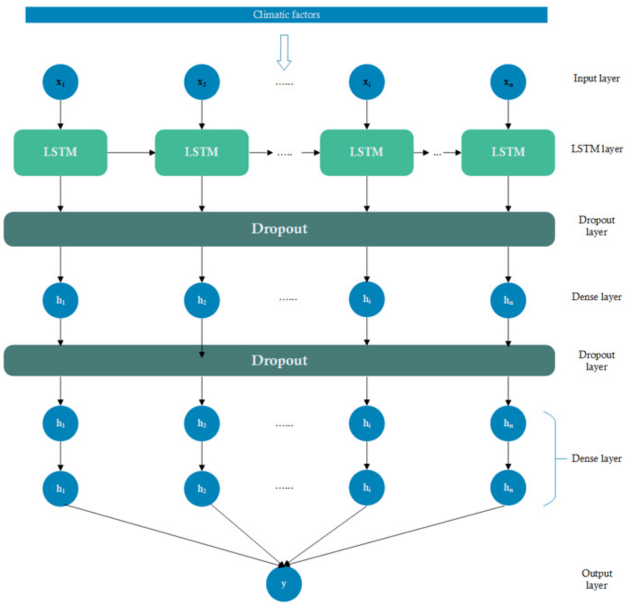


Figure 6. Proposed LSTM Model Architecture.

before, we use the LSTM method to create our prediction model. Our algorithm follows the next steps:

1) *Data collection:* This procedure is carried out, as previously described, using Internet of Things (IoT) devices such as sensors and Raspberry Pi.
2) *Data preparation:* involves several steps before inputting it into the model. These steps include filling in missing values, cleaning the data, normalizing it, and reshaping it into a 3-dimensional array suitable for LSTM input.
3) *Model creation:* Selecting the appropriate model is crucial for developing an effective learning model. The creation process is executed by specifying certain parameters, such as the number of neurons, layers, and the activation functions used.
4) *Training:* The dataset will be partitioned into two segments. The initial segment will be employed for model training, specifically for fine-tuning the weights to align the predictions with the anticipated outcomes.
5) *Evaluation:* The second segment of the dataset will be employed for testing and assessing the model. The testing data is less in size and different from the training data.
6) *Prediction:* Once the training and testing steps are complete, the model can be utilized to make predictions, specifically about land suitability.

5. Results and discussion

This section outlines the configurations that enable us to implement our suggested design, hence our prediction model. Moreover, this section summaries the attained outcomes and provides a discussion related to the subject matter.

Algorithm 2 LSTM prediction model pseudo code.

Input: Datasets of weather conditions.
Output: LSTM_model
Begin
normalize data(0,1); ▹ Normalizing data values between 0 and 1.
$x\_train, y\_train, x\_test, y\_test \leftarrow splitdata(data, 25)$; ▹ Divide the data into training and testing sets (25% testing) reshape data(data) ▹ Reshape data according to LSTM input
$model \leftarrow CreateSequentialModel()$ ▹ Prepare and configure the LSTM model
$model.add\_LSTM\_layer(number\_lstm, sigmoid)$
$model.add\_NN\_layers(nbr\_nn)$ model.compile()
for $epochs\_number$ **do** ▹ Training the model
  for $batch\_size$ **do**
    $model.fit()$;
  end for
end for
$results \leftarrow model.predict(x\_test)$; ▹ Test the model and make predictions
End.

A. Simulation's settings and configurations

Our proposed approach was implemented in Python on a machine with an Intel i7 processor consisting of 16 GB RAM. The architecture of our model consists of an input layer, three hidden layers, and one output layer. The input layer is provided with a three-dimensional array of four features, where a one-dimensional array with four columns represents each feature. The next table (see Table I) explains

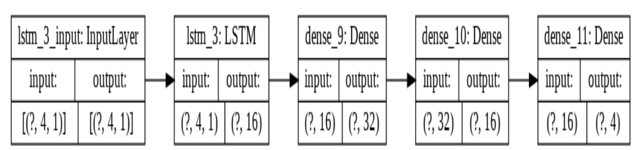


Figure 7. LSTM model architecture and layers.

the used configurations to train the used model LSTM. The explained hyperparameters presented in Table I are selected based on the OptKeras optimization tool. Accordingly, for each hyperparameter, a set of trials is provided to the optimization tool, and for the results, the best combination is chosen.

TABLE I. Hyperparameters settings.

Parameters	Values
LSTM output size	4
Dropout	0.3
Epoch	20
Learning rate	0.001
Batch size	128
Validation split	0.3
Loss function	Cross-entropy
Optimizer	0.80

B. *Obtained results and Discussions*

In this subsection, we present some results about our proposed system. It shows some plots for our model resulting from training and constructing it. Also, it gives data visualization about the degree of suitability for a given area.

1) *Interfaces:* As we mentioned earlier, our system is a web-based application that gives the possibility to the user to use it anywhere. Plus, this application could be used by any means, whether mobile or other device. Our system also gives the possibility to help the experts of farmers to give real-time data streaming from the field with two types of data: weather data and soil data, including PH values and quantity of water from the humidity of the soil.
To check if the given area is suitable or not, the next figures (see figures 8 and 9) help the experts to see the degree of suitability of the chosen area according to its climatic factors. The next figure (see figure 9)

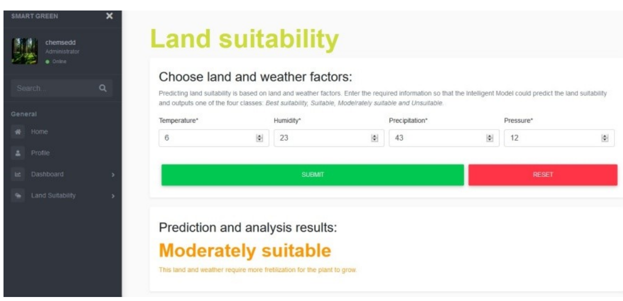


Figure 8. Land suitability from a chosen preference.

helps the experts visualize the degree of suitability of the chosen area as the choropleth map results from the Cloud IoT layer. Figure 9 shows suitability degrees from 1 to 4 (unsuitable, moderately suitable, suitable, and best suitability, respectively).

2) *Model creation:* The next figure shows the results after the training step. As we can see, we have four classes, which correspond to unsuitable, moderately suitable, suitable, and best suitability, respectively. From figure 10 we can see that the testing data presented in green and the prediction results presented in red are almost identical, which shows the effectiveness of our model. Figure 11 shows the difference between the desired output and the predicted output,

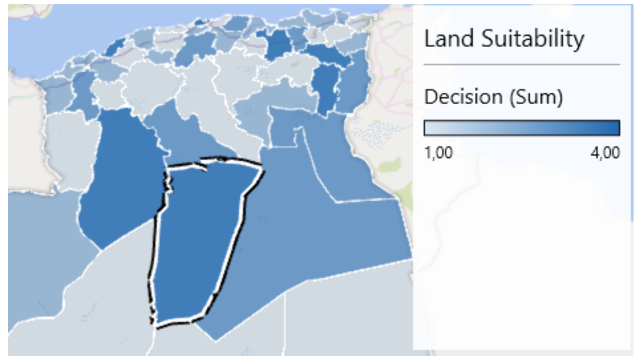


Figure 9. Land suitability using visualization.

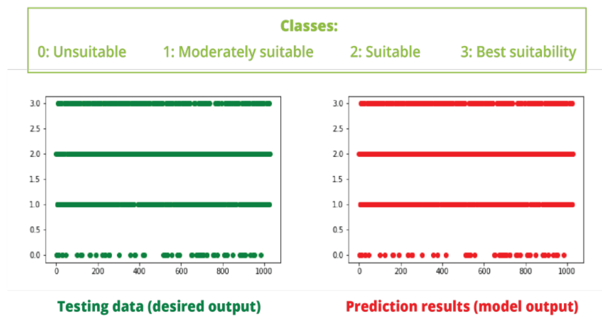


Figure 10. LSTM model -a- Results.

which is given by our model. Google Colab was used for the model's training phase. 200 epochs of training with a five-person batch size were conducted. To accurately identify and predict whether the land will be appropriate or not, our LSTM model was trying to understand the relationship between weather data and the land suitability label. As seen in the image below, accuracy was very poor at the start of the training, with a large loss value.
After 200 epochs, the accuracy at the end of the training has increased to 0.98, while the loss value has decreased to 0.06 compared to the initial values. The LSTM model was able to match the data, as seen by its increasing accuracy over time, which went from 0.44 to 0.98. Predicting land suitability accurately the majority of the time, the validation accuracy, which varies from 0.95 to 0.98, is likewise thought to be very high. The trained model was able to predict accurately for an entirely separate collection of data since it used a different validation set of data that was not part of the training set.
We employed the categorical cross-entropy function to compute the loss. It is the predominant function in the context of categorization difficulties. The categorical cross-entropy metric rises as the anticipated probability deviates from the true label. The model operates by minimizing the loss function, which condenses all components of our algorithm into a single numerical measure that quantifies the

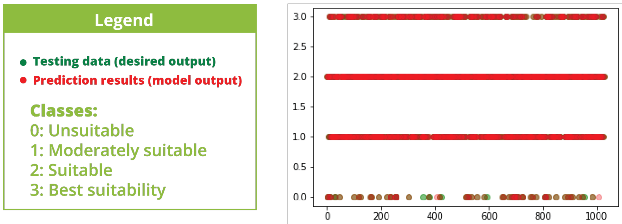


Figure 11. LSTM model -b- Results.

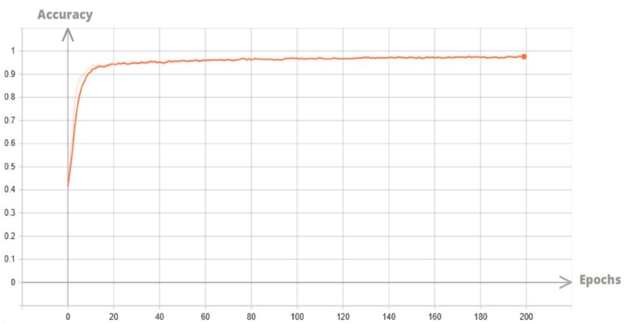


Figure 12. LSTM accuracy.

effectiveness of our model. During the initial training phases, the loss function had a high value of 1.2. The LSTM model strives to learn from the data and reduces the loss by employing backpropagation as the training progresses. After a few epochs, the value drops dramatically to 0.06, indicating that the model successfully extracted information from the dataset and identified the pattern to predict the suitability of the land accurately.

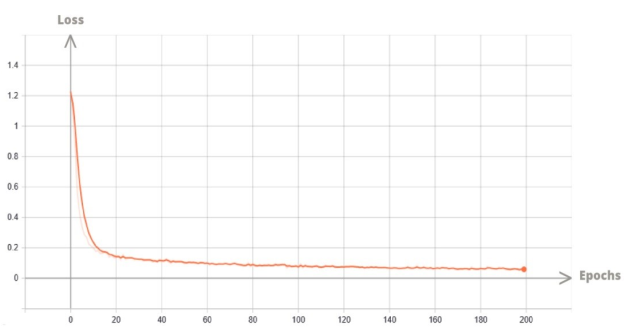


Figure 13. LSTM loss.

As we can see from the obtained results depicted in Table II, which gives a comparative study between LSTM and other ML techniques. Particularly, the LSTM model demonstrated superior performance in capturing the temporal dynamics of the remote sensing data compared to more traditional machine learning approaches. Our deep learning model was able to achieve an overall accuracy of 98% in predicting land suitability, which is a significant improvement over the baseline pixel-based classification

TABLE II. Comparative performance of ML models.

Model	Accuracy	Precision	Recall	F1-score
SVM	0.6533	0.6468	0.6533	0.6338
Linear Regression	0.5012	0.4504	0.5012	0.3949
Logistic Regression	0.4663	0.3683	0.4663	0.4038
LSTM	**0.9805**	**0.9809**	**0.9805**	**0.9805**
Bagging [46]	0.8875	0.8597	0.9108	0.8845
Parallel RF [47]	0.96	-	-	-
MLP [48]	0.945	0.959	0.938	0.946
Crop Suitability Prediction [49]	0.965	0.952	0.9901	0.9709
Soil measures [50]	0.931	0.952	0.943	0.94

accuracy of 85%. Further, the model provided detailed insights into the important drivers of land suitability, with the Sentinel-derived vegetation indices and soil moisture data being the most predictive features. To ensure scalability and generalizability over crops and regions, we tested our model on two datasets. These datasets predict land suitability for a given crop in different regions. These datasets were published on *Kaggle* datasets.

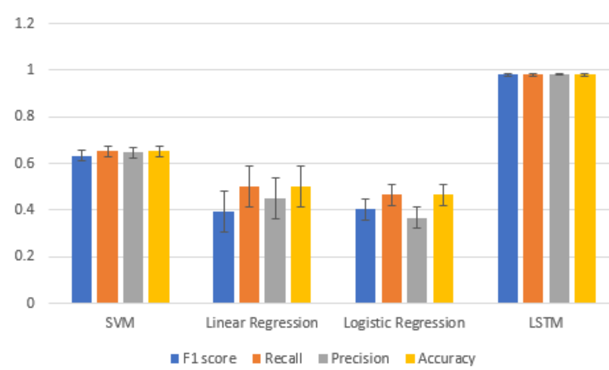
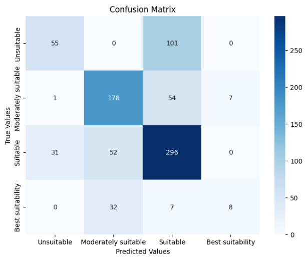
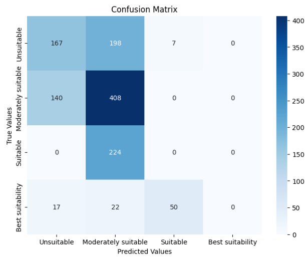


Figure 14. Comparison between different ML models.

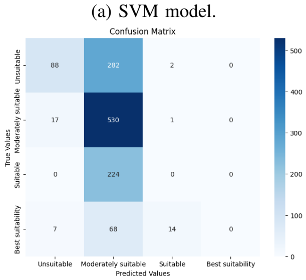
The below figure (see Figure 14) explains the differences between standard deviation changes over the evaluation metrics and the prediction accuracy. Here, the next figure 15 explains the different confusion matrices between ML techniques and LSTM. As we can see, the LSTM has the highest accuracy and lowest errors for different land use classes compared to the other methods. Ensuring device interoperability is crucial to seamlessly integrating new technologies with current systems. Our work has prioritized



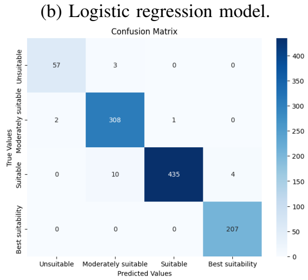
(a) SVM model.



(b) Logistic regression model.



(c) Linear regression model.



(d) LSTM model.

Figure 15. Confusion matrix for different models

interoperability by strictly following standardized data formats derived from the suggested ontology and the fetching mechanism. Our design approach allows for easy integration with other agricultural devices and systems, improving the practical usefulness of our model. Our model's interoperability with existing sensors and data platforms enables seamless real-time data integration, leading to enhanced decision-making and optimized resource management. Preliminary experiments have shown that integrating this approach is beneficial, resulting in improved portability across various platforms and enhanced overall functionality.

C. Discussions

The proposed system aims to deal with land suitability prediction in smart farming to overcome the agriculture field since it plays a crucial role in economic development. As depicted in the obtained results, our solution shows good results compared to other works presented in the literature or different machine learning techniques. The proposed model uses the metrological data during the training step to construct it. The land suitability-based prediction system mainly focuses on smart agriculture for arid areas. The farmers neglected the idea of focusing on such an area due to the costs of land study, the time-consuming nature, and the ability that the area could produce such specific crops. However, the study does not work on specific crops due to the lack of data about each crop, namely pepper, potato, tomato, or even date fruits. These problems are related to the model's generalizability, which can cover all kinds of crops. Unfortunately, to cover such a point, we need an important dataset with various information, such as data on satellites, fertilization, and chemical features for soil in need of crops. To cover the model's scalability, In our subsequent experiments, we have explored how the model can scale by incorporating climatic and soil factors, which significantly influence crop performance. We tested the model on various datasets representing different geographic regions and farming conditions. The results show that, while scaling up the model requires careful consideration of computational resources, it is feasible with the right adaptations, such as model compression techniques and efficient data processing pipelines. We will incorporate a discussion of these experiments in the revised paper to show how the model can be extended to larger datasets and real-world agricultural systems, where the obtained results are presented in Table I. In addition, the solution's scalability should also consider some dynamic changes, especially when dealing with meteorological data. Ultimately, we presume that the challenges of scalability and generalization in agricultural systems are significant, with key issues related to computational capacity, data quality, infrastructure availability, and adaptability to diverse agricultural environments. Overcoming these challenges may require tailored solutions, such as region-specific training of models, improved data collection techniques, and flexible technological infrastructure.

6. Conclusion and Future Work

Land suitability is a crucial factor in agriculture and farming, and our primary objective is to offer accurate predictions. The suggested architecture, which integrates DL and IoT, effectively predicted land suitability by leveraging data on weather and soil conditions. In contrast to the existing literature on smart farming and land appropriateness, our strategy also prioritized the utilization of a data streaming mechanism. Gathering data using the most effective method will undoubtedly enhance outcomes. Our method utilizes the LSTM model to make predictions with time-series data instead of traditional ANN and other ML methods. This is because time series data exhibits patterns over time rather than being composed of isolated data points. As seen in the previous sections, farms generate different types of data, which will create a problem of data heterogeneity. Our solution consists of proposing ontology to deal with the issue. In addition, we have proposed an algorithm to fetch information through data formats. Unfortunately, our solution was hindered by the lack of data. The available historical data span only nine years, which is relatively limited, with only one data point recorded daily.

In the future, we are open to incorporating additional functionalities into our system, like cameras, to keep an eye on the health of the plants. Every day, pictures will be taken and kept on the server databases. Convolutional Neural Networks (CNN) will thereafter receive all of the saved photos to process them and extract pertinent data regarding the health of the plants. Adding more sensors to the field will improve forecast accuracy by enriching the dataset. From an alternative angle, we would like to equip every agricultural truck with a GPS tracker so that the server software can locate each on the field. Giving farmers complete management authority over their farms will increase agricultural output and make the work even easier. To enhance interoperability in the future, we are focused on improving the ability of different systems to work together by partnering with others and creating open interfaces for software applications, which will expand the range of situations where the model may be used and its effectiveness. In general, our emphasis on interoperability tackles important difficulties and improves the practical usefulness of our technology in actual agricultural environments.

References

[1] S. Prathibha, A. Hongal, and M. Jyothi, "Iot based monitoring system in smart agriculture," in *2017 international conference on recent advances in electronics and communication technology (ICRAECT)*. IEEE, 2017, pp. 81–84.

[2] I. Mat, M. R. M. Kassim, A. N. Harun, and I. M. Yusoff, "Smart agriculture using internet of things," in *2018 IEEE conference on open systems (ICOS)*. IEEE, 2018, pp. 54–59.

[3] K. Lakhwani, H. Gianey, N. Agarwal, and S. Gupta, "Development of iot for smart agriculture a review," in *Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018*. Springer, 2019, pp. 425–432.

[4] D. D. K. Rathinam, D. Surendran, A. Shilpa, A. S. Grace, and J. Sherin, "Modern agriculture using wireless sensor network (wsn)," in *2019 5th international conference on advanced computing & communication Systems (ICACCS)*. IEEE, 2019, pp. 515–519.

[5] M. Ghasemaghaei, S. Ebrahimi, and K. Hassanein, "Data analytics competency for improving firm decision making performance," *The Journal of Strategic Information Systems*, vol. 27, no. 1, pp. 101–113, 2018.

[6] J. Ye, "Generalized ordered weighted simplified neutrosophic cosine similarity measure for multiple attribute group decision making," *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, vol. 14, no. 1, pp. 51–62, 2020.

[7] F. Harahap, R. Sitompul, A. Rauf, D. Harahap, H. Walida *et al.*, "Land suitability evaluation for oil palm plantations (elaeisguenensis jacq) on sitellu tali urang julu, pakpak bharat district," in *IOP Conference Series: Earth and Environmental Science*, vol. 260, no. 1. IOP Publishing, 2019, p. 012116.

[8] S. Kim, M. Lee, and C. Shin, "Iot-based strawberry disease prediction system for smart farming," *Sensors*, vol. 18, no. 11, p. 4051, 2018.

[9] R. Khebbache, A. Merizig, K. Rezeg, and J. Lloret, "The recent technological trends of smart irrigation systems in smart farming: a review," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 10 317–10 335, 2023.

[10] A. Somov, D. Shadrin, I. Fastovets, A. Nikitin, S. Matveev, O. Hrinchuk *et al.*, "Pervasive agriculture: Iot-enabled greenhouse for plant growth control," *IEEE Pervasive Computing*, vol. 17, no. 4, pp. 65–75, 2018.

[11] S. Amini, A. Rohani, M. H. Aghkhani, M. H. Abbaspour-Fard, and M. R. Asgharipour, "Assessment of land suitability and agricultural production sustainability using a combined approach (fuzzy-ahp-gis): A case study of mazandaran province, iran," *Information Processing in Agriculture*, vol. 7, no. 3, pp. 384–402, 2020.

[12] A. A. Sulaiman, Y. Sulaeman, N. Mustikasari, D. Nursyamsi, and A. M. Syakir, "Increasing sugar production in indonesia through land suitability analysis and sugar mill restructuring," *Land*, vol. 8, no. 4, p. 61, 2019.

[13] M. M. Subashini, S. Das, S. Heble, U. Raj, and R. Karthik, "Internet of things based wireless plant sensor for smart farming," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 2, pp. 456–468, 2018.

[14] F. Wang, Z. Zhang, C. Liu, Y. Yu, S. Pang, N. Duić, M. Shafie-Khah, and J. P. Catalão, "Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting," *Energy conversion and management*, vol. 181, pp. 443–462, 2019.

[15] M. Noura, M. Atiquzzaman, and M. Gaedke, "Interoperability in internet of things: Taxonomies and open challenges," *Mobile networks and applications*, vol. 24, pp. 796–809, 2019.

[16] M. Al-Osta, A. Bali, and A. Gherbi, "Event driven and semantic based approach for data processing on iot gateway devices," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 12, pp. 4663–4678, 2019.

[17] S. V. Limkar and R. K. Jha, "A novel method for parallel indexing of real time geospatial big data generated by iot devices," *Future generation computer systems*, vol. 97, pp. 433–452, 2019.

[18] M. Munir, S. A. Siddiqui, M. A. Chattha, A. Dengel, and S. Ahmed, "Fusead: Unsupervised anomaly detection in streaming sensors data by fusing statistical and deep learning models," *Sensors*, vol. 19, no. 11, p. 2451, 2019.

[19] Y. Liu, K. Tountas, D. A. Pados, S. N. Batalama, and M. J. Medley, "L1-subspace tracking for streaming data," *Pattern Recognition*, vol. 97, p. 106992, 2020.

[20] Ö. Köksal and B. Tekinerdogan, "Architecture design approach for iot-based farm management information systems," *Precision Agriculture*, vol. 20, pp. 926–958, 2019.

[21] T.-C. Hsu, H. Yang, Y.-C. Chung, and C.-H. Hsu, "A creative iot agriculture platform for cloud fog computing," *Sustainable Computing: Informatics and Systems*, vol. 28, p. 100285, 2020.

[22] A. Pathak, M. AmazUddin, M. J. Abedin, K. Andersson, R. Mustafa, and M. S. Hossain, "Iot based smart system to support agricultural parameters: A case study," *Procedia Computer Science*, vol. 155, pp. 648–653, 2019.

[23] U. J. L. dos Santos, G. Pessin, C. A. da Costa, and R. da Rosa Righi, "Agriprediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops," *Computers and electronics in agriculture*, vol. 161, pp. 202–213, 2019.

[24] M. Y. Shams, S. A. Gamel, and F. M. Talaat, "Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making," *Neural Computing and Applications*, vol. 36, no. 11, pp. 5695–5714, 2024.

[25] D. Paudel, A. De Wit, H. Boogaard, D. Marcos, S. Osinga, and I. N. Athanasiadis, "Interpretability of deep learning models for crop yield forecasting," *Computers and Electronics in Agriculture*, vol. 206, p. 107663, 2023.

[26] W.-L. Chen, Y.-B. Lin, Y.-W. Lin, R. Chen, J.-K. Liao, F.-L. Ng,

Y.-Y. Chan, Y.-C. Liu, C.-C. Wang, C.-H. Chiu *et al.*, "Agritalk: Iot for precision soil farming of turmeric cultivation," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5209–5223, 2019.

[27] M. K. Senapaty, A. Ray, and N. Padhy, "Iot-enabled soil nutrient analysis and crop recommendation model for precision agriculture," *Computers*, vol. 12, no. 3, p. 61, 2023.

[28] V. Shevchenko, A. Lukashevich, D. Taniushkina, A. Bulkin, R. Grinis, K. Kovalev, V. Narozhnaia, N. Sotiriadi, A. Krenke, and Y. Maximov, "Climate change impact on agricultural land suitability: An interpretable machine learning-based eurasia case study," *IEEE Access*, 2024.

[29] S. Nosratabadi, A. Mosavi, R. Keivani, S. Ardabili, and F. Aram, "State of the art survey of deep learning and machine learning models for smart cities and urban sustainability," in *International conference on global research and education*. Springer, 2019, pp. 228–238.

[30] P. Jutadhamakorn, T. Pillavas, V. Visoottiviseth, R. Takano, J. Haga, and D. Kobayashi, "A scalable and low-cost mqtt broker clustering system," in *2017 2nd International Conference on Information Technology (INCIT)*. IEEE, 2017, pp. 1–5.

[31] A. Merizig, O. Kazar, and M. L. Sanchez, "A multi-agent system approach for service deployment in the cloud," *International Journal of Communication Networks and Distributed Systems*, vol. 23, no. 1, pp. 69–92, 2019.

[32] S. A. Butt, M. I. Tariq, T. Jamal, A. Ali, J. L. D. Martinez, and E. De-La-Hoz-Franco, "Predictive variables for agile development merging cloud computing services," *IEEE Access*, vol. 7, pp. 99 273–99 282, 2019.

[33] Y. Drohobytskiy, V. Brevus, and Y. Skorenkyy, "Spark structured streaming: Customizing kafka stream processing," in *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*. IEEE, 2020, pp. 296–299.

[34] M. T. Tun, D. E. Nyaung, and M. P. Phyu, "Performance evaluation of intrusion detection streaming transactions using apache kafka and spark streaming," in *2019 international conference on advanced information technologies (ICAIT)*. IEEE, 2019, pp. 25–30.

[35] B. Blamey, A. Hellander, and S. Toor, "Apache spark streaming, kafka and harmonicio: a performance benchmark and architecture comparison for enterprise and scientific computing," in *International Symposium on Benchmarking, Measuring and Optimization*. Springer, 2019, pp. 335–347.

[36] J. Kiljander, A. D'elia, F. Morandi, P. Hyttinen, J. Takalo-Mattila, A. Ylisaukko-Oja, J.-P. Soininen, and T. S. Cinotti, "Semantic interoperability architecture for pervasive computing and internet of things," *IEEE access*, vol. 2, pp. 856–873, 2014.

[37] S. Aydin and M. N. Aydin, "Semantic and syntactic interoperability for agricultural open-data platforms in the context of iot using crop-specific trait ontologies," *Applied Sciences*, vol. 10, no. 13, p. 4460, 2020.

[38] L. Feng, G. Chen, and J. Peng, "An ontology-based cognitive model for faults diagnosis of hazardous chemical storage devices," *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, vol. 12, no. 4, pp. 101–114, 2018.

[39] M. N. Omri, "Fuzzy ontology-based querying user'requests under uncertain environment," *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, vol. 14, no. 3, pp. 41–59, 2020.

[40] A. Merizig, H. Saouli, M. L. Sanchez, O. Kazar, and A.-N. Benharkat, "An extended data as a service description model for ensuring cloud platform portability," in *Advanced Intelligent Systems for Sustainable Development (AI2SD'2018) Volume 5: Advanced Intelligent Systems for Computing Sciences*. Springer, 2019, pp. 736–744.

[41] D. Hooda and R. Rani, "Ontology driven human activity recognition in heterogeneous sensor measurements," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 12, pp. 5947–5960, 2020.

[42] A. Merizig, H. Belouaar, M. M. Bakhouche, and O. Kazar, "Empowering customer satisfaction chatbot using deep learning and sentiment analysis," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 3, pp. 1752–1761, 2024.

[43] D. Houfani, S. Slatnia, O. Kazar, I. Remadna, H. Saouli, G. Ortiz, and A. Merizig, "An improved model for breast cancer diagnosis by combining pca and logistic regression techniques," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 701–716, 2023.

[44] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Deep recurrent neural network vs. support vector machine for aspect-based sentiment analysis of arabic hotels' reviews," *Journal of computational science*, vol. 27, pp. 386–393, 2018.

[45] B. Jiang, Y. Li, Y. Rezgui, C. Zhang, P. Wang, and T. Zhao, "Multi-source domain generalization deep neural network model for predicting energy consumption in multiple office buildings," *Energy*, vol. 299, p. 131467, 2024.

[46] M. Ganesan, S. Andavar, and R. S. P. Raj, "Prediction of land suitability for crop cultivation using classification techniques," *Brazilian Archives of Biology and Technology*, vol. 64, p. e21200483, 2021.

[47] K. Senagi, N. Jouandeau, and P. Kamoni, "Using parallel random forest classifier in predicting land suitability for crop production," 2017.

[48] D. R. Vincent, N. Deepa, D. Elavarasan, K. Srinivasan, S. H. Chauhdary, and C. Iwendi, "Sensors driven ai-based agriculture recommendation model for assessing land suitability," *Sensors*, vol. 19, no. 17, p. 3667, 2019.

[49] A. WUT. (2023) Crop suitability prediction. [Online]. Available: https://www.kaggle.com/datasets/asadawut/crop-suitability-prediction

[50] M. MOSTAFA. (2024) soil_measures. [Online]. Available: https://www.kaggle.com/datasets/mohamedmostafa259/soil-measures