International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An Integrated Decision-Support System for Increasing Crop Yield Based on Progressive Machine Learning and Sensor Data

[1]Saurabh Bhattacharya, [2]Dr. Manju Pandey

Submitted:26/03/2023 **Revised:**29/05/2023 **Accepted:**12/06/2023

Abstract: The utilization of data has created a massive data tsunami that has affected almost every sector of the economy. The information wave has been amplified by a very significant amount as a direct result of the growing number of man-machine and machine-based digital data handling tools. There has been a significant rise in the number of digital applications in the agricultural industry. These applications are designed to provide improved services to both producers and consumers. The process of crop cultivation is carried out regularly in order to evaluate the various natural factors that affect the development and maintenance of a soil. Although soil has nutrients that are needed for proper cultivation the nutrient levels in the soil are decreasing due to the use of more fertilizer. This issue then led to the reduction of crop production. However, the uncertainty regarding the multiple factors that affect the soil's character can lead to poor decision-making. The ability of agronomists and farmers make informed decision is dependent on the accurate climate and soil data. The concept of precision agriculture is connected to the supervision of the production of crops. DSS is responsible for the collection, organization and evaluation of a wide variety of data types using a variety of mathematical models. various type of crops and environmental data that are collected through sensors in order to improve the decision support system used in agricultural production. To increase the crop yield, an integrated decision support system (DSS) is proposed that takes into account the various components like soil nutrient value, crop details, fertilizer ratio used to predict the ideal crop, fertilizer and rainfall for a region.

Keywords: *DSS, soil nutrient, fertilizer, Machine learning, precision farming.*

1. Introduction

Agriculture is the primary source of gross domestic product and food supply. In India Agricultural sector provides 20.4% of annual GDP in 2022 . It provides food, jobs and fuel for millions of people. According to the UN, "*the global food demand should grow to 70% by 2050 to meet the needs of the world's population. Despite the world's current food production, over 500 million people are still malnourished and over 821 million are going hungry*". This rise in population is seen as a barrier to achieving the Sustainable Development Goals, which have been set out to improve the world's living conditions. It is also possible that the world's water requirements will be difficult to fulfil by the year 2030. Due to the increasing demand for food, farmers are required to increase their production and use more sustainable methods.

Due to the unpredictable weather, improper balanced used of fertilizers, miscalculation of crops, and the increasing demand for food, agricultural industries are looking into using artificial intelligence (AI) to improve the efficiency of their operations. This technology can help them get more out of their harvest. The emergence of the Internet of Things (IoT) and the industrial revolution 4.0 have allowed for the development of new technologies and innovations. These can help improve the efficiency and production of agriculture[1][2].

Through the use of prediction technologies and environmental measurements, smart farming[3] can help farmers improve their operations and reduce their natural resources usage. This technology can also help them achieve greater productivity and improve their bottom line.

The paper presents a framework that uses machine learning and analytics to help farmers identify the ideal seed varieties and allocate the appropriate farmland for their crops. It also proposes that agri-business organizations adopt this framework to create value by providing a decision-support tool. The paper's main contribution is the development of a decision-support tool that combines the capabilities of analytics and machine learning. This tool can help farmers improve their efficiency and increase their profitability. By adopting this framework, agri-business organizations can position themselves as the innovators in their industry.

1Research Scholar, Department of Computer Applications, NIT, Raipur, C.G India
babu.saurabh@gmail.com 0000-0002-9303-3797
2Assistant Professor, Department of Computer Applications, NIT, Raipur, C.G India
mpandey.mca@nitrr.ac.in 0000-0002-5817-4121
****Corresponding Author** – Saurabh Bhattacharya, babu.saurabh@gmail.com*

- The paper's main contribution is the development of a decision-support tool that combines the capabilities of analytics and machine learning. This tool can help farmers improve their efficiency and increase their profitability.
- This paper is regarded as one of the first publications in the field of data science that proposes a data-driven method for crop yield prediction. It includes the segmentation of fertilizer type, crop type, and rainfall prediction.

The goal of the paper is to create a decision-support tool that combines the capabilities of analytics and machine learning. It can help farmers identify the ideal seed varieties and allocate the appropriate farmland for their crops. To start the process of creating this tool, the authors have developed a pilot desktop application that allows a farmer to perform all of the analytics mentioned in the paper.

1.1 smart and precision farming

A smart agriculture concept as depicted in fig.1 is a method that allows farmers to manage the variable aspects of their operations, such as water management, fertilizer production, and production. Modern farming uses various sensors, such as wind, rain, temperature, pH, and moisture. Machine learning is being used in farming to develop new and more effective ways of doing business. Although the technology is still in its early stages, it is widely expected that it will eventually provide a significant advantage to the farming industry. One of the most prominent applications of this type of technology is precision agriculture[4].

In addition to helping farmers improve their farming abilities, artificial intelligence is also being used to guide them to achieve better quality and larger yields. This type of technology is being used in smart agriculture, which aims to reduce the workloads of farmers while increasing their farm output. Some of the modern technologies being used in this field include drones, sensors, and unmanned aircraft systems[5], [6]. The concept of precision agriculture and smart agriculture is being widely used in the farming industry to improve the efficiency and effectiveness of their operations. Through the use of various technologies, such as the internet of things and information technology, farmers can collect data from their surroundings to improve their farm operations.

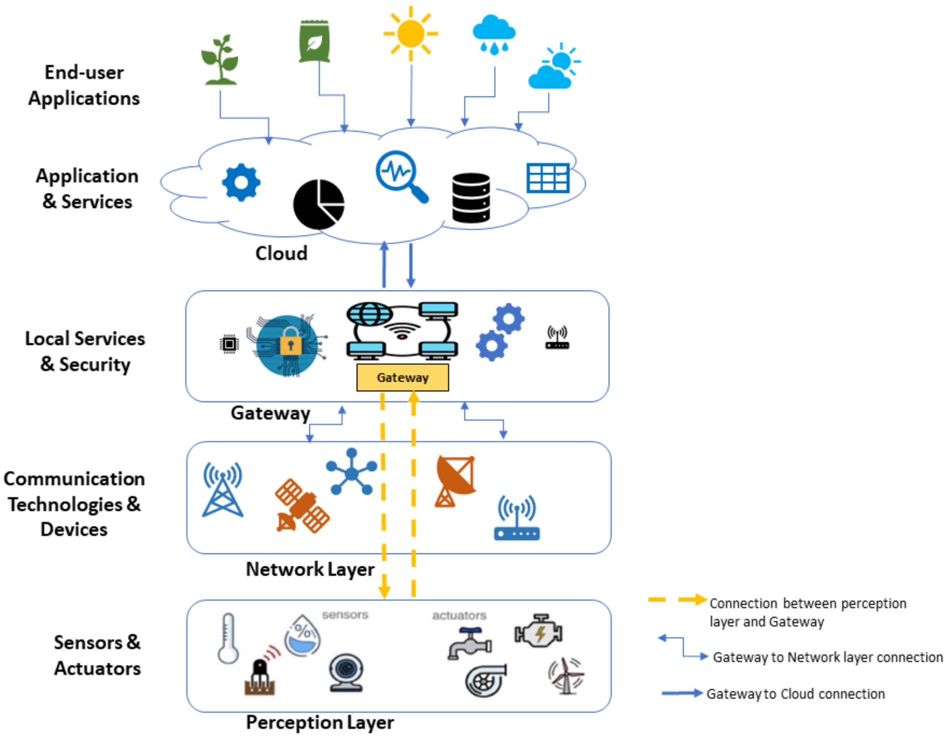

Fig. 1 Precision farming

The use of precision agriculture and the internet of things is being widely used in the farming industry to improve the efficiency and effectiveness of their operations. Through the use of various technologies, such as the internet of things and information technology, farmers can collect data from their surroundings to improve their farm operations[7].

In addition to collecting data about the environment, such as temperature and moisture, sensors also collect other information such as sunlight levels and foliage. This type of technology is commonly referred to as smart agriculture, which is built on the combination of data collection, digital automation, and decision-making. Data analysis and data processing are two of the most common functions of a sensor network. In smart agriculture, the environment is a vital part of the system. Through the use of meteorological data, which is a major factor in the forecasting of climate diseases, farmers can monitor the development of infections in their fields.

The use of data collected for crop protection can help increase the yield and reduce the environmental impact of the operations. Unfortunately, human error can prevent them from properly processing the data. To make the most of the data collected, it is important that the users can use the necessary tools and techniques to improve their decision-making [8]. One of the most critical factors that a farmer must consider when it comes to analyzing and using big data is the ability to identify patterns. The ability to collect data about the environment is a new input that can improve the efficiency of the farming industry. This is because it allows farmers to make informed decisions and improve their farm operations.

1.2 Need for precision farming

Agriculture is regarded as one of India's most important industries and makes a sizeable contribution to the overall gross domestic product of the nation (GDP). It is responsible for 16% of the country's GDP as well as 10% of the country's total exports. Agriculture has some kind of direct or indirect influence on the lives of approximately 75 percent of the country's population. Tomatoes are grown on more than 350,000 hectares of land in India, which is one of the largest producers of this plant and is considered to be one of the largest producers in the world.

India is the world's second-largest producer of potatoes after China, with an annual output of 48.60 million tonnes. China is the world's largest producer of potatoes. In addition to this, it is the sixth largest apple producer in the world, with an annual output of approximately 2311.0 million tonnes of apples produced. Additionally, it is the world's second-largest producer of rice, with an annual output of approximately 116.42 million tonnes of rice, making it the second-largest producer of rice after China.

In India, diseases are known to waste over 15% of the country's crops. This contributes to the country's economic downfall and increases its greenhouse gas emissions. Agriculture accounts for over 10% of Europe's water use and 44% of its greenhouse gas emissions. Pesticides are commonly used to increase fruit crop production, which can have detrimental effects on the environment [9], [10].

Almost 80% of the human diet is composed of plants. This means that there is a growing interest in developing new ways to reduce the water consumption of plants, as well as optimizing the use of pesticides to protect the environment. As a result, agriculture is constantly looking for new ways to increase its crop yields. Unfortunately, due to the various factors that affect its operations, such as weather patterns and population growth, it is difficult to expand its production.

Artificial intelligence is being widely seen as a game-changer in the field of agriculture, and this is why it is important that the technology is used to automate the decision-making process in order to improve the efficiency of the farming industry[3], [11]. This process can be done through the use of various methods such as deep learning and machine learning. In order to identify and monitor crop yield, these technologies can be used to improve the accuracy of the decisions that are made in the field. One of the most beneficial factors that can be considered when it comes to implementing computerized apprehensions is the ability to produce fast and exact results. This technology can also help reduce the time and labor costs associated with the production of crops.

1.3 Adverse effect of chemical

The purpose of the green revolution was to boost the productivity of cereal crops by introducing new, high-yielding varieties of plants and increasing the amount of water that was used for irrigation. In order to ensure that the yield gap in cultivated crops can be closed, proper management of the soil's nutrients is also essential. Utilizing synthetic fertilizers is one of the most efficient approaches to enhancing the nutrient content of soil.. The importance of maintaining a food surplus and enhancing crop productivity has been acknowledged in the past. However, excessive application of pesticides and fertilizers in certain regions has raised concerns about their environmental sustainability.

The continuous application of chemical fertilizers has negative effects on the environment and people's health. This is because reactive nitrogen, which is a volatile element, can disperse through water and air. In addition to this, other problems such as runoffs and nitrate leaching can also cause environmental degradation[12].

Besides harmful effects on the environment, excessive application of nitrogen fertilizers has also been linked to groundwater contamination. Several studies have shown that the application of nitrogen can cause various economic and environmental issues. One of the most critical factors that can affect the sustainability of a country's agricultural systems is the use of proper and balanced input. Unfortunately, the excessive use of chemical fertilizers has raised concerns about the environment. Fig.2 shows the distribution of NPK levels in India.

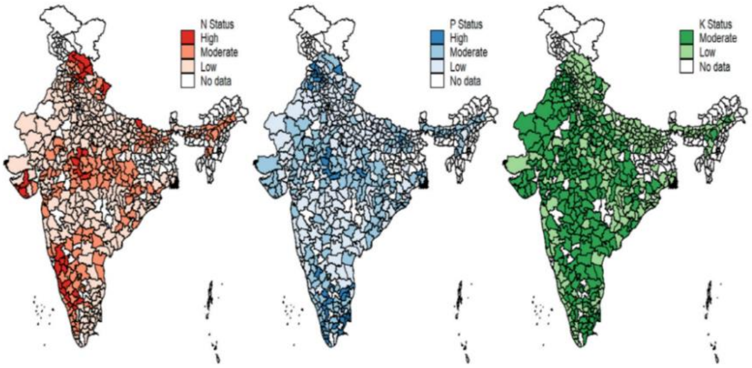

Fig. 2 NPK status in India

Excess fertilizer particles can travel from croplands to bodies of water, which can result in pollution. Continued intensive cultivation can also lead to soil degradation as the nutrients are not being properly utilized.

When soils are depleted of nutrients, the reduction in crop yields can be severe. This is why it is important that the application of chemical fertilizers is carried out in a balanced manner. Injudicious use of chemical fertilizers can also lead to environmental degradation. The optimal use of chemical fertilizers and the application of organic manure can help improve soil health. According to the latest data available, the India's chemical fertilizer consumption during 2017-2018, 2018-19, and 2020-21 was at 54.38, 56.21, 59.88, and 33.85 million tons, respectively and it holds 2nd position globally for using fertilizers as shown in fig.2. Some of the major fertilizers that were used during these years include Urea, DAP, MOP, Complexes, and Single Super Phosphate.

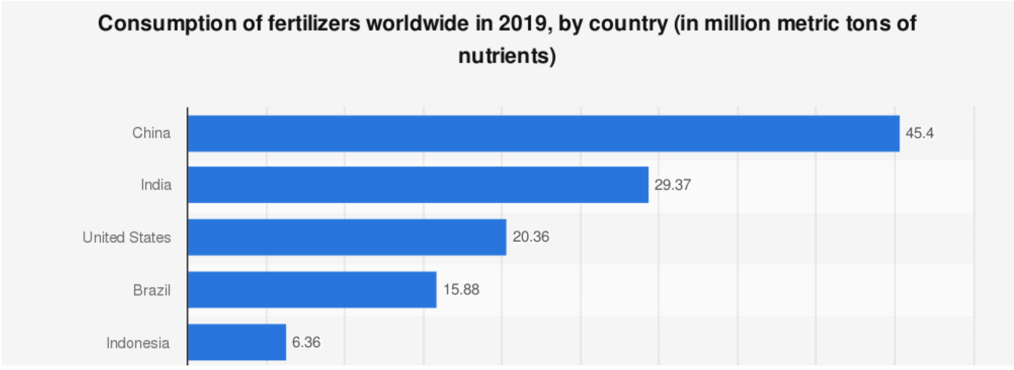

Fig. 3 Global ranking of using fertilizers (src- Statista)

2. Related Works

The use of fertilizers is influenced by various factors such as the type of crop that's being cultivated, the soil type, and the weather conditions. Besides these, other factors such as the farmers' ability to purchase and irrigation are also taken into account to determine the optimal use of these chemicals. The Department of Agriculture and farmers Welfare regularly carry out assessments to determine the chemical fertilizers required for the various crops in the country. These assessments are carried out based on the requirements of the different nutrients, such as nitrogen, potassium, and phosphate.

The failure to identify the appropriate crop identification and nutrient parameters can affect the yield of a plant. This issue can lead to various long-term issues such as global warming and famine. According to studies, implementing more pragmatic farming techniques using scientific methods can help improve the efficiency and reduce environmental damage. In a study conducted by Tawseef Ayoub Shaikh et al.[13] they discussed the various advantages of using ML in farming. They noted that it can help identify the diseases in plants, improve crop yield, and control pesticides. In a study conducted by Kaushik Bora[14], he analyzed the relationship between the application of chemical fertilizers and environmental sustainability. He found that there is a wide range of variations in the fertilizer application across India. The spatial patterns of the different districts revealed that there are both deficit and excessive fertilizer applications. In the Indo-Gangetic plains, many districts have excessive amounts of fertilizer. The results of the assessment revealed that there is a surplus of nitrogen. This issue can be addressed through the implementation of policies aimed at nutrient balance in the soil. A study conducted by Savvas Dimitriadis et al.[8] focused on precision agriculture, which is a method that focuses on plant-driven management. Through the use of a decision support system that can learn about a field's conditions, it can provide treatments such as irrigation, fertilizer, and pesticides in real time. Machine learning techniques can be used to extract new knowledge from the existing data in order to improve the efficiency of the natural resource management process. The method is based on an iterative and inductive process, which involves examining the patterns and associations that have emerged during the discovery stage.

In a study conducted by Loris Francesco Termite et al.[15] they developed a machine learning framework that can describe the temporal behavior of moisture in soil. Two different models were used to train the system. The Soil Water Index is a representation of the moisture content of the soil. It is derived from satellite observations and, in addition to being input, has the last known value. This variable can be used in the scheduling of irrigation. It is a vital part of decision support systems that help manage the use of water resources.

Anat Goldstein et al.[16] used a combination of classification and regression techniques to develop models that can predict the weekly irrigation plan. They were able to achieve this by using eight different variables. After comparing the two models, it was revealed that the Gradient Boosted Regression Trees performed the best. The Boosted Tree Classification model performed well with a 95% accuracy.

E. Esakki Vigneswaran et al.[17] conducted a study on the use of a machine learning algorithm to analyze the soil quality by measuring the amount of nutrients, Urea, potassium, magnesium, and pH. It was able to provide farmers with a better understanding of the potential yield of their crops. The study also analyzed the datasets and collected samples to predict the optimal crop for their agricultural operations. The researchers were able to create a suitable crop rotation sequence by estimating the nutrients available in the soil during the different trimesters.

In a study conducted by Chunying Zeng et al.[18] they discussed the various features of the expert database and how they can be used to design a system. The database can be used to display tree lists and other agricultural data. The researchers additionally found the map path using the loading function and easily uploaded the data. The researchers moved the data collected from the platform to other cities and provinces in order to increase its scalability. They also used a grey decision-making system and a pair of prediction models to predict the future agricultural indicators.

A study conducted by Brahim Jabir et al.[19] utilized a deep learning framework to identify weeds and then apply a herbicide-free method to the environment. Through this method, they were able to reduce the use of large-scale chemicals and preserve the environment. A smart system that was proposed using object detection models was able to identify and classify various objects in a wheat crop area in real time. It was also able to provide decision support by selecting a suitable herbicide based on the detected weed.

In a study conducted by Rubby Aworka et al.[20] they proposed three crop prediction models that are: the Crop Random Forest, the Crop Gradient Boosting Machine, and the Crop Support Vector Machine. They were able to combine the data from various sources to develop a decision system that can predict the crop yield in 14 East African countries. Unfortunately, the lack of reliable agricultural data in Africa has hindered the development of effective decision systems.

In a study conducted by Nesrine Kalboussi et al.[21] they proposed a decision support tool that can help agricultural producers make informed decisions regarding the use of water reclamation. The tool can also help them interpret the Life Cycle Assessment results.The study also highlighted the need for further research regarding the various applications of DL models in agriculture. The researchers noted that the models' performance can be improved by implementing new methods, such as focusing mechanisms and single-stage detection models.

Khadijeh Alibabaei et al.[22] discussed about the novelty of DL models' applications and the various challenges they face in agriculture are some of the factors that need to be studied further. The use of new methods, such as focus mechanisms and single-stage detection models, can help improve the models' performance.

A deep learning system was proposed by Shahbaz Khan et al.[23] to identify crops and weeds in croplands. It was evaluated and implemented using high-resolution UAV images taken over two target fields: strawberry and pea. The system was able to identify the weeds with an accuracy of 95.3%, while the overall accuracy of the system was 94.73%. The developed system exhibited an average kappa coefficient of 0.89. It was able to outperform the existing DL and ML approaches in terms of performance. It can be used as a precision sprayer for implementing the SSWM strategy.

A framework that combines clustering, machine learning, and simulation to help farmers make critical decisions has been developed by a team led by Durai Sundaramoorthi et al.[24]. This approach utilizes a data-driven simulation to estimate the optimal parameters needed to perform portfolio optimization. According to the analysis performed by the researchers, the system could help an average farmer earn up to $177,369 annually by analyzing and implementing it. In addition, they noted that there are various challenges that need to be addressed before deep learning can be widely used in agriculture.

The findings suggest that by utilizing recent developments in the field of machine learning, it will be possible to improve model performance in terms of accuracy while simultaneously reducing the amount of time required for inference. Additionally, this will make the models more applicable to applications in the real world. This motivates to design an integrated DSS which analyze and recommends types of crops based on soil nutrient, fertilizers recommendation as well as predict the rainfall at a glance.

3. System design and deployment

The concept of soil fertility refers to the ability of a soil to provide nutrients to crops. When the nutrients are balanced, the higher the yield, but when the nutrients are not present in sufficient quantities, the crop yield is reduced. Most farmers are not aware of the harmful effects of over-fertilization on the environment when it comes to the application of nutrients such as nitrogen, phosphorus, and potassium. This issue has been identified as one of the main reasons why the country is experiencing severe environmental pollution. The surface runoff from an agroecological system can contaminate water sources and contribute to the pollution of air and soil. In addition, improper use of fertilizers and energy can affect the environment. To minimize the effects of environmental pollution, agriculture should use the necessary resources efficiently and carefully. A decision support system is an information system that helps a business make informed decisions by analyzing and resolving various issues related to its operations. It can also help the agro-based organization and farmers make better decisions for better crop yield. This type of system is either human-powered or automated. It can collect and analyze vast amounts of data to help solve business problems.

3.1. Designing of sensory architecture

The prototype uses an embedded microcontroller known as the ESP32, which is based on the Xtensa platform. Its robust and versatile design enables it to be used in a variety of power scenarios. It also features a single 2.4 GHz Bluetooth chip and a built-in Wi-Fi connection. The sensor, which is called the NPK, is designed to measure the level of nutrients in the soil. This sensor can be used without the need for chemical reagents, which makes it ideal for monitoring various fields such as agriculture. Soil temperature, soil pH value, environment temperature and humidity sensors are integrated and programmed to match the data collected by the device.

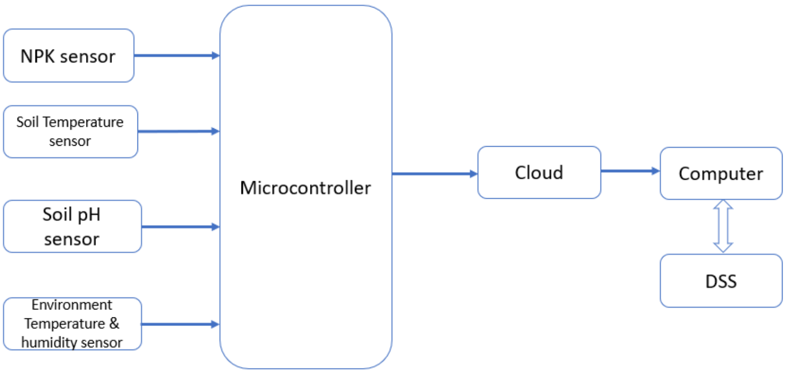

Fig. 4 Prototype - sensor architecture

The prototype can be placed in the soil to improve its accuracy by digging a pit measuring 30 and 15 centimeters deep. Its robust and waterproof design helps keep track of the data collected continuously for several days. It can also help determine the appropriate fertilizer quantities for different crops. Data collected from sensor is transfer to computer via cloud. Further data is used in making proper decision

3.2. Design of DSS

Decision support systems (DSS) have significantly evolved since their early development during the 1970s. They have taken on a broader or narrower definition over the years, with different systems now being created to help decisionmakers with specific kinds of problems. One of the first definitions of these systems was provided by Morton and Keen. They were designed to help improve the quality of decisions by collecting resources and focusing on semi-structured problems[25], [26]. A decision support system is a type of human-computer system that collects, processes, and presents information based on computer models. Decision support system is primarily designed to improve the quality of decisions by collecting and presenting information based on computer models. It can be divided into four main subsystems: the Data Management, the Model Management, the User Interface, and the Knowledge-based. The Data Management component manages the data that will be used to make decisions. The Model component is composed of various models that help decision-makers make effective decisions. The Knowledge-based is the system's hearth, where the solution-making process takes place. The User Interface allows users to interact with the system and obtain information[27].

Due to the increasing number of issues related to climate change and the need to improve the quality of agricultural production, the development of decision support systems has been regarded as an essential tool for the agriculture industry. These systems can help farmers face these challenges and implement effective strategies to increase their production. The increasing number of applications of decision support systems has been attributed to the technological advancements that have occurred over the past decade. These include the emergence of new technologies such as Artificial Intelligence and Machine Learning[28].

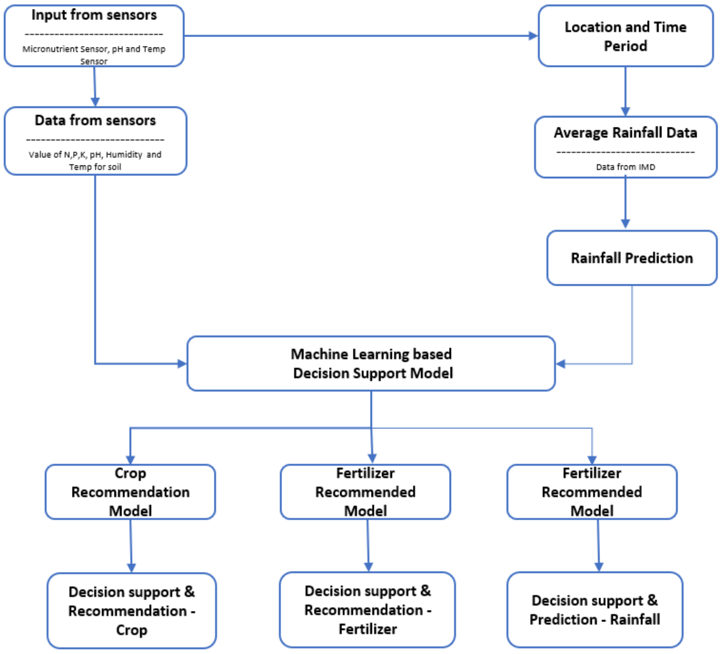

Fig. 5 Prototype - Decision Support System

A decision support system for agriculture is composed of various components that gather, organize, interpret, and analyze information related to a crop. These systems can help farmers make informed decisions and implement effective strategies to increase their production. For instance, they can provide them with the necessary information to schedule effective treatments and improve the efficiency of their operations.

Designing a decision support system can be challenging due to the complexity of the data collected and analyzed. This is because it involves the involvement of various disciplines, such as mathematics, statistics, and crop hardware and software. For instance, it's important to understand how different factors affect crop yield.

Although there is currently no single agricultural decision support system that has been adopted globally, several have been developed over the years for different types of farming practices. These systems were designed to help farmers improve their crop management and irrigation practices.

4. Methodology

The following sections describe the various steps involved in the study, which was carried out using a data mining process. The process included the definition of the data set, extraction-transformation-loading, the application of ML algorithms, the evaluation of the trained models and designing of DSS. Fig.5 shows the steps involved in design

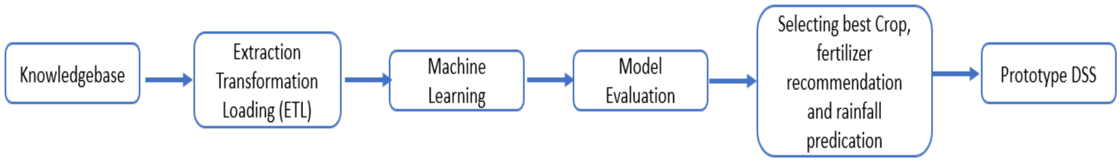

Fig. 6 Steps involved in making DSS with ML

4.1. Data sources

Data are collected from various sensors like NPK, soil temperature and pH values. For training 3 various dataset (www.kaggle.com) are used and for testing sensory data used to analyze and recommend the requirement.

4.2. Machine Leaning Classifiers

Machine learning (ML) is a field of study that focuses on the design and construction of computer programs that can improve their performance after experiencing certain types of conditions. One of the main goals of this discipline is to develop algorithms that can learn from data. These algorithms usually utilize a set of observations known as a training-set. Machine learning algorithms are designed to perform various tasks, such as analyzing the relationships between multiple independent variables and dependent variables. They can also classify an observation into one of a set of classes or clusters, clustering, or association rule learning.

"*Pycharm*" is used for the development of models and data analysis on the Python platform; "*pyQt*" is used to design Decision Support System interface. It is a python-based machine learning framework that can be used for the creation and deployment of applications and services. Its libraries for data manipulation and transformation, as well as its multiple ML toolkits, are designed to help create and evaluate models.

In this study, Decision Tree, Naïve Bayes and Logistic Regression are used to analyze the result and recommend accordingly. The models were presented in the following sections.

4.2.1. **Decision Tree classifier** - The use of a decision tree is a supervised learning technique that is commonly used for analyzing classification and regression problems. It is a structure that consists of a set of nodes that represent the features of a given dataset. Each leaf node represents the outcome of the problem. The two nodes in a decision tree are known as the Decision Node and the Leaf Node. The former is used to make decisions and has multiple branches, while the latter is output of those decisions. A decision tree is a representation of the various possible solutions to a given problem or issue based on the features of the given dataset. It can also perform tests based on the given conditions. A decision tree is typically described as a tree-like structure that starts with the root node and grows to various branches. To build one, we use the classification and regression algorithm known as the CART algorithm.

4.2.2. **Naïve Bayes classifier** – Naive Bayes is a supervised learning method that is based on the Bayes theorem and focuses on the hypothesis that a particular feature class can be missing or present without any other classes. Naive Bayes is an example of the Bayesian inference method. This class is not connected in any way to the other characteristics. The Naive Bayes model takes into account the many different theoretical facets that are involved in the process. For example, the target characteristic Y is the kind of variable that can be forecasted based on the values that come before it in the list: x1, x2,..... When taking into account the K different alternative values, another characteristic of Y can assume the value of the expected feature. For example, the variable Xi is considered to be an independent variable or feature, whereas the variable Y that it is related with is considered to be a dependent variable.

4.2.3. **Logistic Regression classifier -** The output of a variable that is dependent on a certain categorical value can be predicted by using the logistic regression method. This is the approach's purpose. In order to carry out this strategy, one must take into account a number of different independent factors. Either a discrete value or a categorical value must be used for this value. The approach, rather than revealing the precise value, reveals the probability values, which range from 0 to 1. The logistic function is used to make predictions about the chance of a variable being right. Because it can make dependable and accurate forecasts, this strategy has gained a lot of popularity in recent years. It is also capable of categorizing data based on the many forms of data.

5. Results and Discussion

The process of data preparation, also known as data preprocessing, involves performing various steps on raw data to prepare it for other data processing operations. This has been an important step in the process of mining data. Recently, techniques related to data preprocessing have been used in the development of AI models and training machine learning models. Preprocessing is a process that takes data from various sources and formats and makes it easier to process in data mining, machine learning, and other data science tasks. It is typically performed at the earliest stages of the development of AI and machine learning projects. The various steps involved in data preparation include selecting a representative subset of the data, transforming the data, denoising it, and creating a single input. Others include imputation, which produces relevant data for missing values, normalization, and feature extraction. Fig. 7 and fig.8 shows the interface of DSS. DSS provides the flexibility of selecting various dataset of various category which makes it more flexible and robust in nature. After proper pre-processing of data the best classifier helps to recommend the fertilizer and crop to user and also helps in predicting rainfall.

5.1. Screenshot- Prototype DSS

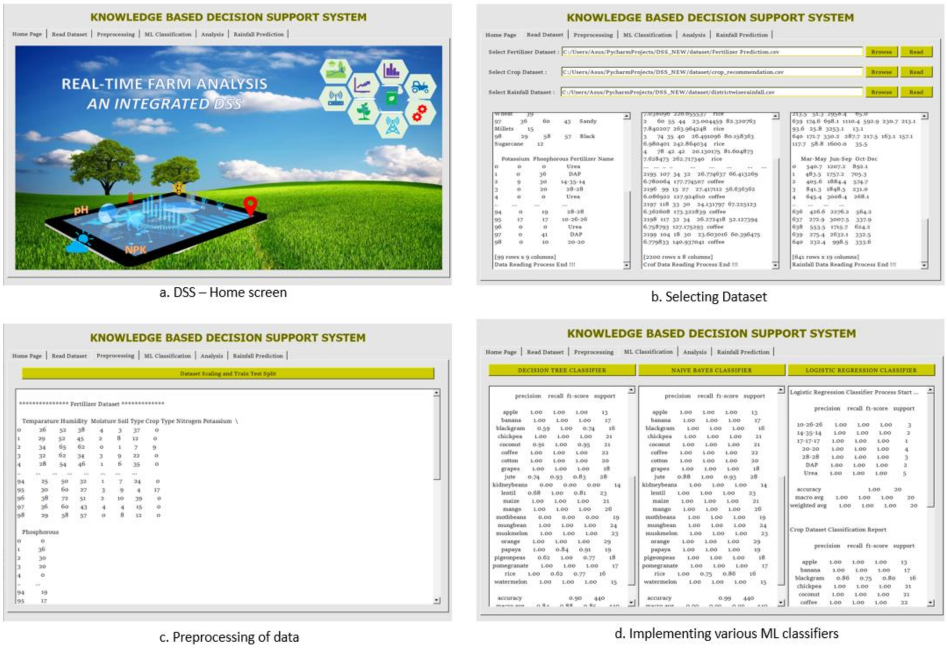

Fig. 7 screenshot - Prototype DSS

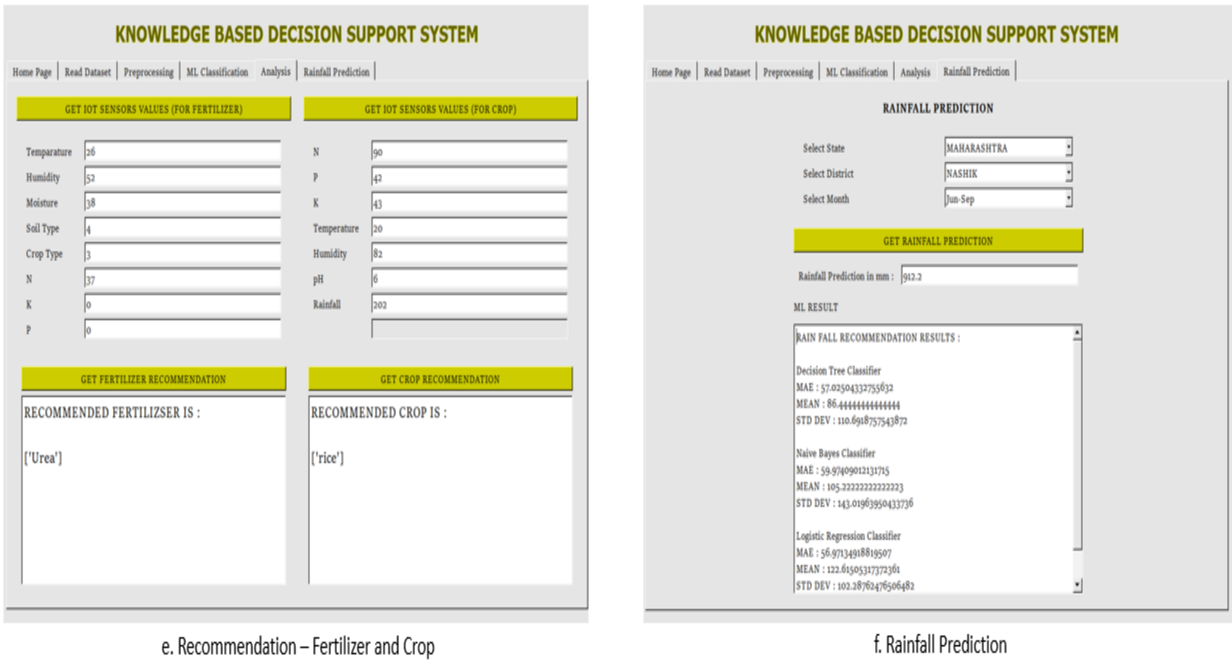

Fig. 8 screenshot - Prototype DSS

The paper discussed the importance of accurate crop and environmental data to be incorporated into the Agricultural DSS. It also highlighted the need for decision support frameworks and inputs that can help facilitate the decision-making process. The paper showed the various types of sensors that can be used for data acquisition. The first step in the process of implementing a decision support system is the acquisition of data. This process involves selecting the appropriate tools and methods for data acquisition. Even though there are numerous sensors and systems that can be used for data acquisition, the accuracy of the system's decision-making is still affected by the selection of the right data acquisition techniques.

5.2. ML classifier Evaluation

a. Accuracy - The accuracy of a prediction is a percentage that is calculated by taking into account all the correct predictions made by the test participants.

$$Accuracy = \frac{Correct\ prediction}{Total\ Prediction}$$

b. Precision- The precision is measured by the fraction of its relevant examples that were predicted to be in a given class.

$$Precision = \frac{True\ positive}{True\ positive + False\ positive}$$

c. Recall- A recall is a fraction of the examples that were predicted to be in a given class. It refers to the number of examples that are truly representative of the class.

$$Recall = \frac{true\ positive}{true\ positive + false\ negative}$$

d. F1-score- The F1 score is calculated as the harmonic mean of both the recall and the precision categories.

$$F1 - score = 2 * \frac{precision * Recall}{precision + Recall}$$

e. MAE- The mean square error is a statistical measure that shows the difference between the predicted and the actual output. It is commonly used to evaluate whether a prediction was too high or too low.

$$\text{MSE}\left(y_{true}, y_{pred}\right) = \frac{1}{n_{samples}} \sum \left(y_{true} - y_{pred}\right)^2$$

f. **Standard Deviation-** A standard deviation is a number that shows how many values are close to the average. Low standard deviations mean that most of the numbers are within the range of the average. On the other hand, a high standard deviation values the values over a wider area.

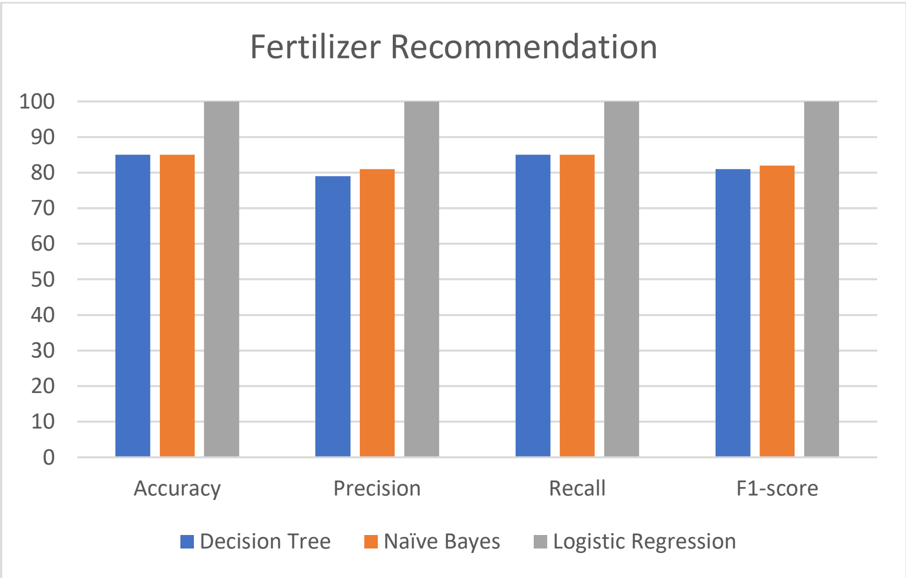

Fig. 9 Fertilizer Recommendation

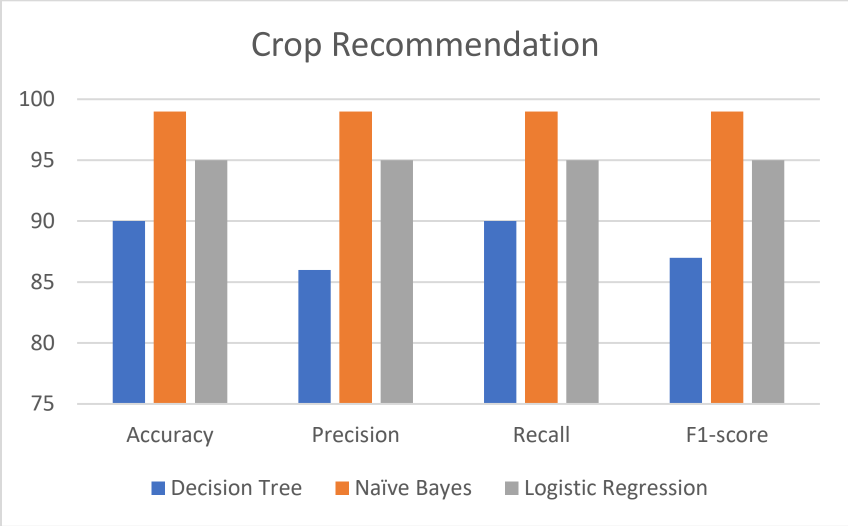

Fig. 10 Crop recommendation

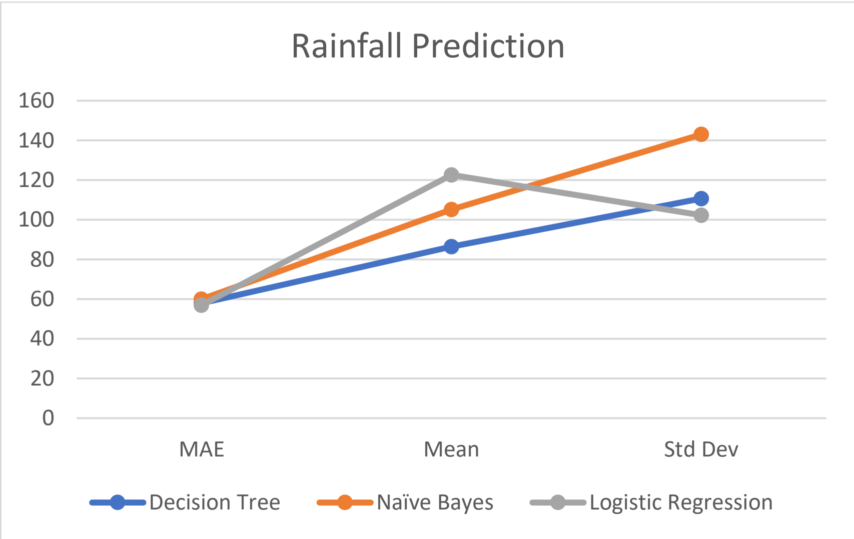

Fig. 11 Rainfall prediction

6. Conclusion

The agriculture industry is a major contributor to the development of developing countries. It is very challenging to extract the necessary knowledge from large amounts of data. When it comes to the cultivation the nutrients needed for fertilization are balanced. Unfortunately, over-fertilization can lead to inaccurate data and environmental issues. Soils become less competent to handle the nutrients needed for the crop yield. It is very important that specialists are able to use tools to calculate the appropriate rates of macronutrients for their crops. This study shows how a system can help them do so. This is why it is important that the techniques used in data mining are designed to improve the efficiency of the agricultural sector. Predicting the amount of harvest that a crop will produce is one of the most difficult challenges that farmers face. They are able to increase the effectiveness of their operations by employing data mining strategies. The purpose of this research is to construct a decision support system that, by making use of historical data, will give farmers the ability to forecast both the amount of crop yield and the amount of rainfall. The purpose of this paper is to provide a comparison of various machine learning and data mining techniques in the hopes of achieving the highest possible level of accuracy..

References

[1] S. Mohapatra, B. Sharp, A. K. Sahoo, and D. Sahoo, “Decomposition of climate-induced productivity growth in Indian agriculture,” *Environ. Challenges*, vol. 7, no. October 2021, p. 100494, 2022, doi: 10.1016/j.envc.2022.100494.

[2] R. Rosati *et al.*, “From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0,” *J. Intell. Manuf.*, 2022, doi: 10.1007/s10845-022-01960-x.

[3] M. K. Saggi and S. Jain, “A Survey Towards Decision Support System on Smart Irrigation Scheduling Using

Machine Learning approaches," *Arch. Comput. Methods Eng.*, vol. 29, no. 6, pp. 4455–4478, 2022, doi: 10.1007/s11831-022-09746-3.

[4] P. Taechatanasat and L. Armstrong, "Decision Support System Data for Farmer Decision Making," *Proc. Asian Fed. Inf. Technol. Agric.*, pp. 472–486, 2014, [Online]. Available: https://ro.ecu.edu.au/ecuworkspost2013/855/.

[5] D. Popescu, F. Stoican, G. Stamatescu, L. Ichim, and C. Dragana, "Monitoring in Precision Agriculture," *Sensors (Switzerland)*, vol. 20, no. 3, p. 817, 2020, [Online]. Available: https://doi.org/10.3390/s20030817.

[6] H. Bagha, A. Yavari, and D. Georgakopoulos, "Hybrid Sensing Platform for IoT-Based Precision Agriculture," *Futur. Internet*, vol. 14, no. 8, 2022, doi: 10.3390/fi14080233.

[7] B. Keswani *et al.*, "Adapting weather conditions based IoT enabled smart irrigation technique in precision agriculture mechanisms," *Neural Comput. Appl.*, vol. 31, no. s1, pp. 277–292, 2019, doi: 10.1007/s00521-018-3737-1.

[8] S. Dimitriadis and C. Goumopoulos, "Applying machine learning to extract new knowledge in precision agriculture applications," *Proc. - 12th Pan-Hellenic Conf. Informatics, PCI 2008*, pp. 100–104, 2008, doi: 10.1109/PCI.2008.30.

[9] S. A. Mir and S. M. K. Quadri, "Climate Change, Intercropping, Pest Control and Beneficial Microorganisms," *Clim. Chang. Intercropping, Pest Control Benef. Microorg.*, 2009, doi: 10.1007/978-90-481-2716-0.

[10] B. J. Van Alphen and J. J. Stoorvogel, "A Functional Approach to Soil Characterization in Support of Precision Agriculture," *Soil Sci. Soc. Am. J.*, vol. 64, no. 5, pp. 1706–1713, 2000, doi: 10.2136/sssaj2000.6451706x.

[11] Ashok Kumar, L. ., Jebarani, M. R. E. ., & Gokula Krishnan, V. . (2023). Optimized Deep Belief Neural Network for Semantic Change Detection in Multi-Temporal Image. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 86–93. https://doi.org/10.17762/ijritcc.v11i2.6132

[12] A. Belhadi, S. S. Kamble, V. Mani, I. Benkhati, and F. E. Touriki, "An ensemble machine learning approach for forecasting credit risk of agricultural SMEs' investments in agriculture 4.0 through supply chain finance," *Ann. Oper. Res.*, 2021, doi: 10.1007/s10479-021-04366-9.

[13] R. Chand and L. M. Pandey, "Fertiliser Use, Nutrient Imbalances and Subsidies: Trends and Implications," *Margin*, vol. 3, no. 4, pp. 409–432, 2009, doi: 10.1177/097380100900300404.

[14] T. A. Shaikh, W. A. Mir, T. Rasool, and S. Sofi, *Machine Learning for Smart Agriculture and Precision Farming: Towards Making the Fields Talk*, no. 0123456789. Springer Netherlands, 2022.

[15] K. Bora, "Spatial patterns of fertilizer use and imbalances: Evidence from rice cultivation in India," *Environ. Challenges*, vol. 7, no. January, p. 100452, 2022, doi: 10.1016/j.envc.2022.100452.

[16] L. F. Termite, A. Garinei, A. Marini, M. Marconi, and L. Biondi, "Combining satellite data and Machine Learning techniques for irrigation Decision Support Systems," *2019 IEEE Int. Work. Metrol. Agric. For. MetroAgriFor 2019 - Proc.*, pp. 291–296, 2019, doi: 10.1109/MetroAgriFor.2019.8909279.

[17] A. Goldstein, L. Fink, A. Meitin, S. Bohadana, O. Lutenberg, and G. Ravid, "Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge," *Precis. Agric.*, vol. 19, no. 3, pp. 421–444, 2018, doi: 10.1007/s11119-017-9527-4.

[18] E. Esakki Vigneswaran and M. Selvaganesh, "Decision Support System for Crop Rotation Using Machine Learning," *Proc. 4th Int. Conf. Inven. Syst. Control. ICISC 2020*, no. Icisc, pp. 925–930, 2020, doi: 10.1109/ICISC47916.2020.9171120.

[19] C. Zeng, F. Zhang, and M. Luo, "A deep neural network-based decision support system for intelligent geospatial data analysis in intelligent agriculture system," *Soft Comput.*, vol. 26, no. 20, pp. 10813–10826, 2022, doi: 10.1007/s00500-022-07018-7.

[20] B. Jabir and N. Falih, "Deep learning-based decision support system for weeds detection in wheat fields," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, pp. 816–825, 2022, doi: 10.11591/ijece.v12i1.pp816-825.

[21] [20] R. Aworka *et al.*, "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries," *Smart Agric. Technol.*, vol. 2, no. March, p. 100048, 2022, doi: 10.1016/j.atech.2022.100048.

[22] N. Kalboussi, Y. Biard, L. Pradeleix, A. Rapaport, C. Sinfort, and N. Ait-mouheb, "Life cycle assessment as decision support tool for water reuse in agriculture irrigation," *Sci. Total Environ.*, vol. 836, no. April, p. 155486, 2022, doi: 10.1016/j.scitotenv.2022.155486.

[23] K. Alibabaei *et al.*, "A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities," *Remote Sens.*, vol. 14, no. 3, pp. 1–43, 2022, doi: 10.3390/rs14030638.

[24] S. Khan, M. Tufail, M. T. Khan, Z. A. Khan, and S. Anwar, "Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer," *Precis. Agric.*, vol. 22, no. 6, pp. 1711–1727, 2021, doi: 10.1007/s11119-021-09808-9.

[25] D. Sundaramoorthi and L. Dong, "Machine learning and optimization based decision-support tool for seed variety selection," *Ann. Oper. Res.*, 2022, doi: 10.1007/s10479-022-04995-8.

[26] K. Czimber and B. Gálos, "A new decision support system to analyse the impacts of climate change on the Hungarian forestry and agricultural sectors," *Scand. J. For. Res.*, vol. 31, no. 7, pp. 664–673, 2016, doi:

10.1080/02827581.2016.1212088.

[27] M. Debeljak *et al.*, “A Field-Scale Decision Support System for Assessment and Management of Soil Functions,” *Front. Environ. Sci.*, vol. 7, no. August 2019, pp. 1–14, 2019, doi: 10.3389/fenvs.2019.00115.

[28] Gabriel Santos, Natural Language Processing for Text Classification in Legal Documents , Machine Learning Applications Conference Proceedings, Vol 2 2022.

[29] R. Rupnik, M. Kukar, P. Vračar, D. Košir, D. Pevec, and Z. Bosnić, “AgroDSS: A decision support system for agriculture and farming,” *Comput. Electron. Agric.*, vol. 161, no. November 2017, pp. 260–271, 2019, doi: 10.1016/j.compag.2018.04.001.

[30] R. Katarya, A. Raturi, A. Mehndiratta, and A. Thapper, “Impact of Machine Learning Techniques in Precision Agriculture,” *Proc. 3rd Int. Conf. Emerg. Technol. Comput. Eng. Mach. Learn. Internet Things, ICETCE 2020*, no. February, pp. 18–23, 2020, doi: 10.1109/ICETCE48199.2020.9091741.